Providing knowledge recommendations: an approach for informal electronic mentoring

Ricardo Colomo-Palacios
(Universidad Carlos III de Madrid, Leganés (Madrid), Spain
ricardo.colomo@uc3m.es)

Cristina Casado-Lumbreras
(Universidad Complutense de Madrid, Madrid, Spain
cristinacasadolumbreras@pdi.ucm.es)

Pedro Soto-Acosta
(Universidad de Murcia, Murcia, Spain
psoto@um.es)

Sanjay Misra
(Atilim University, Ankara, Turkey
smisra@atilim.edu.tr)

Abstract: The use of Web 2.0 technologies for knowledge management is invading the corporate sphere. The Web 2.0 is the most adopted knowledge transfer tool within knowledge intensive firms and is starting to be used for mentoring. This paper presents IM-TAG, a Web 2.0 tool, based on semantic technologies, for informal mentoring. The tool offers recommendations of mentoring contents built upon personal competencies of the mentee, combined with content and opinion tagging. To validate the tool, a case study comparing recommendations from the IM-TAG and a group of experts was conducted. Results show that the accuracy of IM-TAG's recommendations is notable and satisfactory. The main conclusions of this research may be valuable to organizations immersed in mentoring programs.

1   Introduction

The second phase in the Web evolution, the Web 2.0, is attracting the attention of (information technology) IT professionals, businesses and Web users (Murugesan, 2007). Today, millions of people interact through blogs, collaborate through wikis, play multiplayer games, publish podcasts and video, build relationships through social network sites and evaluate all the above forms of communication through feedback and ranking mechanisms (Warschauer & Grimes, 2007). The Web 2.0 provides the software to both inspire and support these new ways of interaction (Kuswara & Richards, 2011). As a result, the social Web brings an ever-growing number of social networks which host and share all types of contents, knowledge and expertise in a number of areas (Poblet et al., 2011). The advantages of "the Web 2.0" have raised the interest of companies which seek to obtain the benefits derived from this technology (Ferreira, 2010). In an scenario in which the Web 2.0 technologies are invading the corporate sphere, MacAfee coined the term Enterprise 2.0 to summarize

the interest in the use of these tools for generating, sharing and refining knowledge in a global setting (McAfee, 2006). Several works (e.g. Richards, 2009) analyzed the impact of the social web on knowledge management, while others (e.g. Levy, 2009; Paroutis & Al Saleh, 2009) coined the term Knowledge 2.0 to summarize the upcoming trend in knowledge management.

The offering of rewards, incentives, mentoring and induction, within the organizational context, may be seen as formal mechanisms to facilitate knowledge sharing between individuals (Bosua & Scheepers, 2007). With regard to mentoring, the Web 2.0 is the most adopted knowledge transfer tool for knowledge intensive firms (Fong and Choi, 2009), enabling the transfer of tacit knowledge from experienced employees to new recruits (Srikantaiah & Koenig, 2000). Despite its advantages and the decreasing costs of technology, the implementation costs of mentoring programmes are one of the main disadvantages (Grybek, 1997). Informal mentoring, a type of mentoring created spontaneously, is also expensive because of personnel costs which in many cases are high. Social web tools are being seen as enablers of both informal and formal mentoring (e.g. Kirkwood, 2010; Stanton-Salazar, 2011; Wheeler, 2009).

Another issue which affects the Web 2.0 is the information overload phenomenon (Bawden & Robinson, 2009), which also influences knowledge management based on web 2.0 tools (Kirchner and Sudzina, 2009). Information overload represents a state of affairs where an individual's efficiency in using information is hampered by the amount of relevant, and potentially useful, information available (Bawden & Robinson, 2009). Recommendation systems will leverage the current information overload by providing a reliable knowledge management oriented set of strategies. Such systems can reduce search efforts (Liang et al., 2006)), so solving the problem of information overload (Kuo, Chen & Liang, 2009). The aim of this paper is to present the IM-TAG tool, a web 2.0 system, based on semantic technologies, which provides knowledge recommendations for informal mentoring.

The paper consists of four sections and is structured as follows. Section 2 reviews the relevant literature about mentoring and the use of semantic technologies in content tagging. Section 3 describes the tool paying attention to its architecture and main features. Section 4 describes the evaluation process carried out. Finally, the paper ends with a discussion of research findings, limitations and concluding remarks.

2   Literature Review

2.1   Semantic Technologies and Content Tagging

The "Semantic Web", described by Berners-Lee Hendler, and Lassila (2001), has resulted in a considerable amount of research and development initiatives to extend the current Web technology by using machine-understandable metadata. Semantic, from the Greek "sēmantikos", involves giving significance or meaning to words or symbols, enabling distinctions between the meanings of different words or symbols. The Semantic Web proposes the idea that web contents are defined and linked not only for visualization but for being used and processed by applications (Castellanos-Nieves et al., 2011).

Semantic technologies are based on ontologies (Fensel, 2002) and provide a common framework that enables data integration, sharing and reuse from multiple sources. Ontology can be defined as "a formal and explicit specification of a shared conceptualisation" (Studer, Benjamins & Fensel, 1998). The main objective of ontologies is to establish ontological agreements, which serve as the basis for communication between either human or software agents, hence, reducing language ambiguity and knowledge differences between agents, which may lead to errors, misunderstandings and inefficiencies (Blanco et al., 2011).

The research on semantic technologies, apart from ontologies, relies on a number of key methodologies such as knowledge representation languages or reasoning algorithms (Hitzler & Janowicz, 2011). The importance of these technologies has produced semantics-based solutions in a wide scope of environments including customer relationship management (García-Crespo et al., 2010), multimedia (Paniagua-Martín et al., 2011; García-Barriocanal et al., 2011), trading (Rodríguez-González et al., 2011), research and development activities (Colomo-Palacios et al., 2010), digital libraries (García-Crespo et al., 2011a) and education (Yang et al., 2011; Yessad et al., 2011).

Web content is readable by humans, but, unless it is semantically annotated, it is not machine readable, in the sense that it cannot be automatically interpreted in any reasonable manner (Dotsika, 2010). Semantic annotations go beyond textual annotations of the documents. They identify concepts and relations between concepts within documents intended primarily to be used by machines (Uren et al., 2006).). In this context, the current focus of the semantic web research has turned to investigating methods to redesign the Web in order to add semantics to data, manually or automatically. These methods enable easier machine processing of information (Benjamins et al., 2008).

Although automatic annotation of content is producing good results, assuring near-to-perfect quality is still beyond the limits of state of the art systems (Frank et al., 2012). Taking this into account, according to Siorpaes & Simperl (2010), there is a wide range of approaches that allow semi-automatic annotation of texts, with most of them using natural language processing and information extraction techniques. Thus, IM-Tag, following Noh et al.'s (2010) work, adopts a semi-automatic annotation method to provide knowledge recommendations.

2.2 Mentoring

The mentoring topic has received considerable attention in both the academic and popular press as well as in the highly public venue of the Internet (Haggard et al., 2011). As a result, the number of articles published on mentoring in the social sciences and education literatures has increased "exponentially" in the last 20 years (Kirchmeyer, 2005).

The origin of mentoring dates back to the earliest stages of human civilization (Kammeyer-Mueller & Judge, 2008). More specifically, it dates back to Homer's Odyssey when Odysseus, before leaving to fight in the Trojan War (traditionally dated 1193 BC-1183 BC), entrusted his older friend Mentor to teach and educate his son, Telemachus (Gentry, Wever, Sadri, 2008). The description of Mentor includes the following characteristics: half-human, half-God; half-male, half-female; believable yet unreachable; wisdom personified; and a paradoxical union of both goal

and path. So, the image of a mentor as a wise person, a guide, and a stand-in parent who assists in the protégé's growth and development has its roots way back in literature (Bierema & Merriam, 2002).

Mentoring can be defined as the matching of a novice with a more experienced person in the same role (Reiss, 2007). The People-Capability Maturity Model (P-CMM) states that the purpose of mentoring is to transfer the lessons learned from experienced personnel to other individuals or workgroups (Curtis, Hefley & Miller, 2009). Furthermore, mentors use their experience to provide not only skills but also personal support and guidance. Mentoring activities are organized around knowledge, skills and processes to deploy competency-based competences (Curtis, Hefley & Miller, 2009).

There are many papers in the literature which highlight the benefits that mentoring relationships bring for mentors and mentees. With regard to mentees, research confirmed that individuals who are mentored have greater opportunities to advance in their professional career, get higher salaries and achieve better satisfaction (Knouse, 2001). Also, there is research (e.g. Allen, 2007) which reports a set of benefits from mentoring to mentors as well, including higher performance, satisfaction, rejuvenation and higher promotion opportunities. However, as suggested by Singh, Ragins and Tharenou (2009), mentoring, while mattering for career success, represents only a part of a constellation of career resources embedded within the relationships.

In contrast to spontaneously-derived informal mentoring relationships, formal mentoring programs (imposed by the organization) are usually in the form of the assignment of a mentee to a mentor (Blake-Beard, 2001). Ragins and Cotton (1991) discussed the differences between formal and informal mentoring following three dimensions: initiation of the relationship; structure of the relationship; and processes in the relationship. The initiation of formal mentoring relationships is externally controlled while informal mentoring relationships are initiated when two people are attracted to one another based on the foundation of perceived similarity. With respect to the structure and the processes of the relationship, the main differences are as follows: informal mentoring relationships last from three to six years; meetings and activities in informal mentoring occur when desired as opposed to a set schedule; and the goals of relationships evolve over time in informal mentoring (Blake-Beard, 2001).

The literature suggests that informal mentoring relationships may be more valuable than the formal ones (e.g. Raabe & Beehr, 2003; Casado-Lumbreras et al., 2011), since informal mentoring relationship tends to be more natural and spontaneous, hence, happening more on an ad hoc basis (Ragins, 1999). Informal mentoring has been found to be positively and significantly associated with knowledge sharing (Karkoulian, Halawi & McCarthy, 2008). In addition, previous research confirmed that informal mentoring relationships take more time, although they outperform formal relationships in terms of professional development (Chao, Waltz & Gardner, 1992).

The increasing importance of the internet has led to new forms of mentoring (Soto-Acosta, Casado-Lumbreras & Cabezas-Isla, 2010). This statement can be extended to the informal mentoring panorama. Now, many organizational efforts are devoted to integrating the Web 2.0 within organizational learning scenarios (Wang,

2011) as a part of their knowledge worker support (Schneckenberg, 2009). On the other hand, Universities are also following this path (García-Peñalvo et al., 2011) . This paper goes beyond these initiatives and presents IM-TAG, a tool in which social content published by informal mentors in a corporate intranet is semantically tagged. The main output of this approach is the recommendation of informal mentoring contents performed by a system, which is adapted to the needs of users.

3   IM-Tag

It is widely acknowledged that Mentoring programs are expensive (Grybek, 1997). This also occurs for the case of informal mentoring because, although it is not planned, personnel costs are usually high. The Social web may provide a solution to reduce costs. More specifically, blog posts can act as informal mentoring instruments, since a single post may be consulted by thousands of informal mentees all over the world. At the same time, blog tagging is today a very active line of research in the literature (e.g. Tsai, 2011; Tsai & Chan, 2011). Thus, making use of both perspectives, the IM-TAG tool proposes the collaborative tagging of internal blog-posts to produce content recommendations in an intranet environment. This methodology approaches the concept of semantic blogging, which aims to describe semantic information about individual content items within blog posts (internal semantic) using RDF (Bojārs et al., 2008). However, beyond adopting or redefining the concept, the purpose of the system is to give exact recommendations based on consistent profiles and reliable comments in a scenario in which corporate culture is uniform.

Next, to introduce the system, the architecture of the tool is described and, then, a case study is presented.

3.1   Architecture

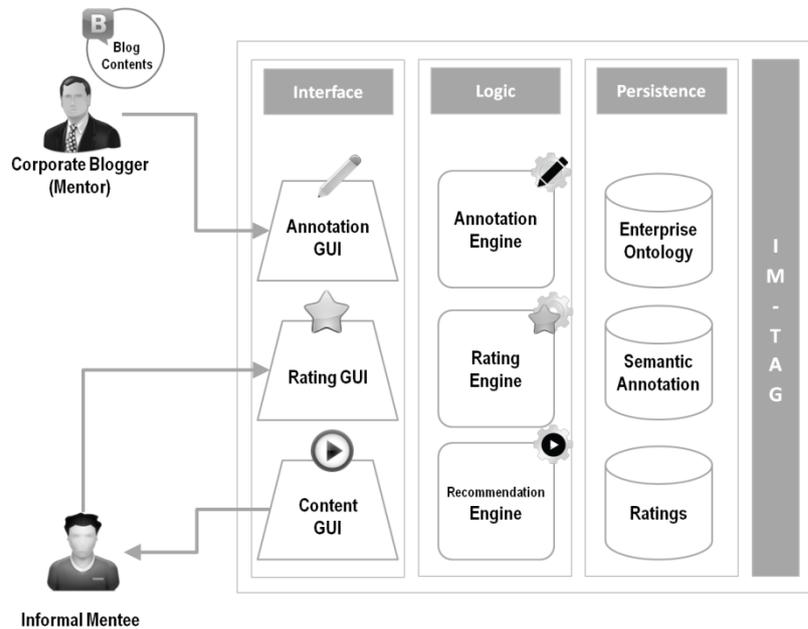The architecture presents three main elements, which are described below. Figure 1 shows how the IMG-TAG tool works.

Figure 1. IMG-TAG Architecture

INTERFACE

There are two different user profiles, each with a specific role: mentor and mentee. IM-TAG provides specific elements in order to ease the work of each type of user. The Interface contains the following three modules:

- Annotation GUI. The interface allows easy annotation of blog contents. These annotations will be made by means of semi-automatic annotation methods. Based on the previous work SOLAR (García-Crespo et al., 2010b), this feature is implemented as an Annotation Plug-in that can be inserted in environments such as Drupal or WordPress. This setup facilitates the annotation of blog contents in the same environment in which these contents are produced. Annotation functionalities were easily hardcoded and integrated in the plug-ins taking advantage of the flexible and dynamic properties of the Wordpress framework. In addition, the authors developed an Asynchronous JavaScript And XML (AJAX, for short) interface, both in the Search and Navigation software components, to provide a set of loosely-coupled features which ensure efficiency and ease of use.

- Rating GUI. Once the content is annotated by the informal mentors, it needs to be rated by one or more mentees. Using the GUI rating, informal mentors are able to perform that task. These ratings constitute the rank in which blog contents are ordered and classified.

- Content GUI. The content GUI is the means by which blog contents are sent to mentees. These contents are delivered to users in the news

section located within the private intranet home page, thus allowing them to read and rate contents.

LOGIC

Under the interface, the logic layer hosts the engines which allow the operation of IM-TAG as a whole. This layer presents three modules:

- Annotation Engine. This module provides the means to translate user interactions (posts and comments) into semantic annotations stored in the persistence layer in OWL-DL defined ontologies. The objective of the annotation is to provide advanced searches and facilitate the retrieval of information. Given that this component must suggest an annotation to a specific text, it includes a Text Processor and a Natural Language Processor developed by the authors in previous works (García-Crespo et al., 2009) and based on GATE (General Architecture for Text Engineering).

- Rating Engine. This module allows the rating of contents provided by mentees. This rating is described as a weighted mean of the ratings provided by all mentees. The rating of an individual user is calculated considering aspects such as his/her professional role and rating history joint to build a reputation mechanism.

- Recommendation Engine. This module consisted in first versions of the system of an OWL Description Logics based Reasoner, the Renamed ABox and Concept Expression Reasoner (RACER). Several Description Logic (DL) axioms were developed in order to automatically classify individuals of the ontology depending on some parameters such as the rating levels. However, current version have been changed to automatically perform these tasks using rule-based engines instead of DL axioms due to performance issues such as the ones described by Rodríguez-González et al. (2012). The new inference schema makes use nowadays of Jena Rules engine to perform the recommendation tasks based on the rules defined.

PERSISTENCE

The organizational ontology has been defined using the OWL (Bechhofer et al., 2004). The OWL language presents three variants: OWL-Lite, OWL-DL and OWL-Full. OWL-Lite provides a small set of features, while OWL-DL is more expressive than OWL-Lite providing decidability based on description logics. OWL-Full allows full expressivity but decidability is not guaranteed. For this reason, IM-TAG employs OWL-DL for the ontology definition. The OWL DL flavour builds on the formal foundations of Description Logics (DL). OWL DL supports the use of subsumption as the selected reasoning. The storage and ontology reasoning were developed based on the Jena framework.

With respect to enterprise ontology, this can be seen as a collection of terms and definitions relevant to an enterprise to ensure that all parties involved have a shared understanding of the relevant aspects of that enterprise (Uschold et al., 1998). An *Enterprise Ontology* is fundamental because it defines a common vocabulary to guide the description of any organization. There are several enterprise or business

ontologies, but the most cited one is the Enterprise Ontology (Uschold, King, Moralee, & Zorgios, 1998). This ontology has been used in various knowledge management scenarios (e.g. Chen, 2008; Han & Park, 2009). Moreover, this is the root of many software engineering ontologies, such as the Software Enterprise Ontology (Villela et al., 2005) and constitutes the base of the enterprise ontology included in IM-TAG.
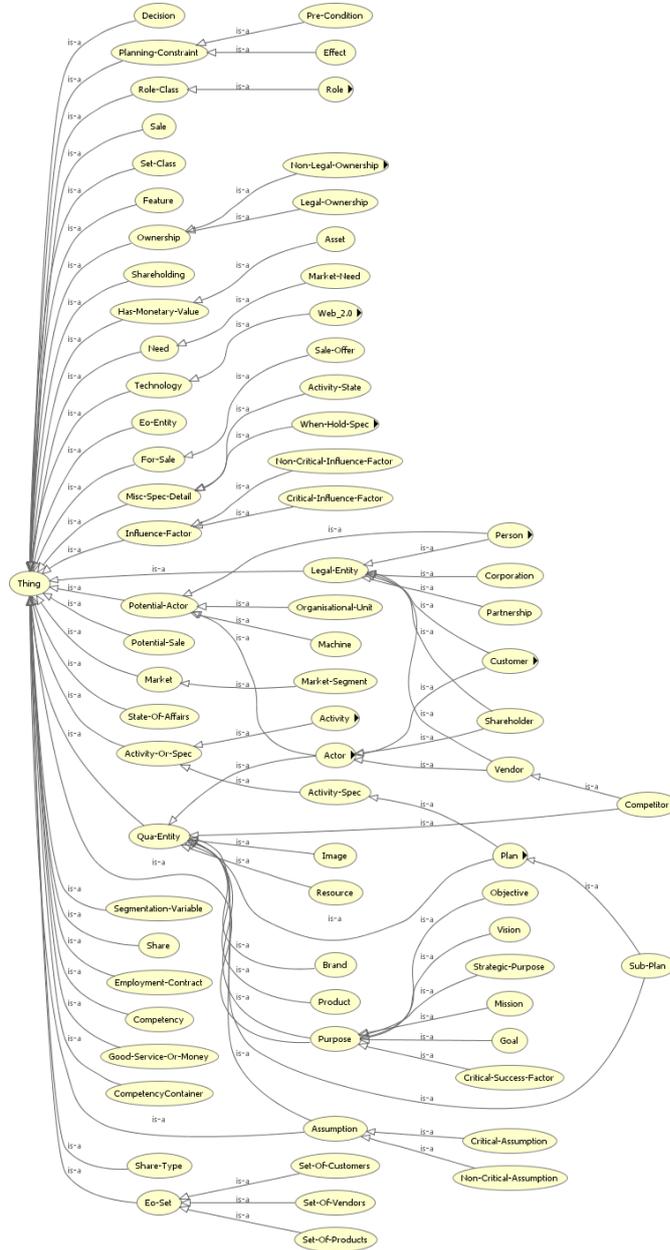
Figure 2. Excerpt of the ontology schema

The main modifications which have been made over enterprise ontology consist in the definition of several classes which are necessary in order to store the knowledge related with the domain and provide recommendations. First of all, it is necessary to mention that the most up-to-date version of the enterprise ontology is still developed in Ontolingua (Gruber, 1992). The main contribution for hence was the development of the structure of enterprise ontology using OWL as representation language. Some of the relations which were originally defined in enterprise ontology have been also included in our OWL version, paying special attention to those which are related with the recommendation process described in the paper.

Apart of the creation of a reduced OWL version of enterprise ontology, several classes which are very related with the domain have been introduced to support the representation of specialized knowledge and for hence being able to provide recommendations. A list of the most relevant classes which have been generated is provided. Figure 2 shows an excerpt of the new ontology schema.

However, the inclusion of some classes related with the domain doesn't allow storing the knowledge of the domain with enough accuracy. A few relations and data properties have been also included to improve the quality of the ontology. Figure 3 shows an excerpt of these relations and the domain and range which affects.

Figure 3. New relations provided in the ontology
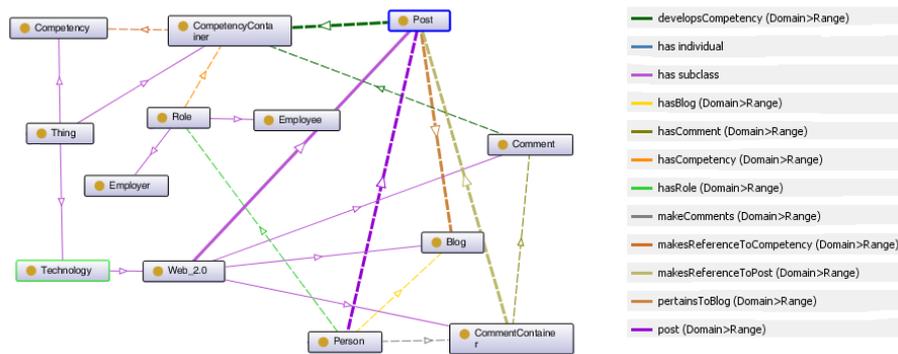
Defined Relations:

<Role> hasCompetency <CompetencyContainer> : NON FUNCT[1]
<CompetencyContainer> makesReferenceToCompetency <Competency> : FUNCT

<Person> hasRole <Role> : NON FUNCT
<Person> hasBlog <Blog> : NON FUNCT
<Person> post <Post> : NON FUNCT
<Post> pertainsToBlog <BLOG> : FUNCT[2]

---

[1] Non functional relation
[2] Functional relation

<Person> makeComments <CommentContainer> : NON FUNCT
<CommentContainer> hasComment <Comment> : FUNCT
<CommentContainer> makesReferenceToPost <Post> : FUNCT
<Post> developsCompetency <CompetencyContainer> : NON FUNCT
<Comment> developsCompetency <CompetencyContainer> : NON FUNCT

Defined Properties:

<CompetencyContainer> realLevel: FUNCT : int [0-10]
<CompetencyContainer> virtualLevel: FUNCT : int [0-10]

In figure 4 it is possible to see a representation of the entire ontology with the new relations and classes which have been described.

Figure 4. Ontology Representation

3.2 Use Case

MINT-INC (fiction name) is a software development company. During the last years MINT-INC has developed projects for different clients including other consulting companies as well as public and private end-clients. The company has its headquarters in Madrid (Spain) but has branches in several cities around the country. Apart from this, the company has three delegations in Latin-America: Argentina, Mexico and Colombia. The total workforce of the company is 2,500 employees.

MINT-INC has designed and implemented formal mentoring programmes in order to facilitate the transfer of tacit knowledge among employees. Apart from formal and informal mentoring relationships, there are two mentoring sources that employees use to understand the company to a wider extent: CEO and CIO blogs. These blogs present posts related to the company and its environment and are a good source of information about the culture and values of MINT-INC. The company realized the potential of the social media for knowledge management and promoted the use of blogs among its personnel. Thus, the company decided to make use of social web contents in order to provide good information to newcomers and veteran employees. To achieve that objective, the IM-TAG tool was installed on Wordpress to provide bloggers with a tool to tag content and facilitate its sharing among workers. Before the tool was installed, the enterprise ontology was adapted to cover all issues relative to MINT-INC knowledge and culture. This ontology presents concepts that cover company competences and culture issues.

The first phase of the process begins when content (a post) is written using IM-TAG and the blogger tags the post using the ontology. The second phase of the process comes when a user, recognized as an employee of MINT-INC, reads the content and wants to express his or her opinion about it. This can be done using the classic comment tool that almost every blog presents, but it can be enriched by using IM-TAG too. In this case, the user expresses his or her impressions about the content and usefulness for other employees. This is done through rich content rating.

Finally, IM-TAG gathers all contents and matches employees' profiles with blog content tagging and rating. The result of this process is a recommendation that IM-TAG provides to users in form of contents published in the intranet. This socio-semantic approach goes further the closed learning platforms. Recommendation flows mean real informal learning outcomes for the novice workers and are useful for mentees and, as an extension, to organization.

4 Evaluation

With the aim of getting feedback concerning the tool, an evaluation was carried out. The methodology and results are described below.

4.1 Design

Once the system had been developed and tested from the point of view of the development process, the second step was to test the validity of the IM-TAG tool. That is, to investigate the accuracy of its recommendations.

To evaluate the accuracy of the system, precision, recall and F1 measures were used. Recall and precision measures reflect the different aspects of annotation performance. These measures were first used to measure an information retrieval system by Cleverdon et al. (1966). The F1 measure was later introduced by van Rijsbergen (1979) in order to combine precision and recall measures, with equal importance, into a single parameter for optimization. These measures must be complemented in order to get a full view of the system from an evaluation perspective. Thus, the second dimension is the coverage test. The coverage of a system is a measure of the domain of items in the system over which the system can form predictions or make recommendations (i.e. the percentage of items for which a recommender system can provide recommendations) (Herlocker et al., 2004). A high coverage value means that the system provides assistance in selecting among most of the items, while a low coverage means that the system may be less valuable to users, since its decision-making capacity is limited (Park & Chang, 2009).

To carry out both procedures, a set of software development organizations were contacted. The focus on software organizations is rooted on their inner complexities (Colomo-Palacios et al., 2011) and their dependence on human capital. Up to four organizations answered positively to a personal invitation sent by the authors selected from their personal and academic contacts. From these four organizations, two of them were selected: on the one hand, their CEOs regularly publish blog posts and on the other hand, they present a set of junior professionals enrolled in a formal mentoring program. Taking into account the aims of the evaluation, two different tasks were performed. On the one hand, a set of potential informal mentees were identified and classified in the system. On the other, several posts were annotated in order to feed the system with enough information (mentoring contents). The semantic annotation was performed by researchers assisted by post authors on both blogs using semi-automatic annotations. This joint annotation was adopted in order to avoid tool adoption problems. Overall, a set of eighty posts were annotated - forty per blog. Posts selection criteria was based on finding the first forty posts that can be classified under this categories: how-to/tutorial posts, standard list posts, case studies, problems-and-solutions posts, stories, controversial posts, inspiring posts or research posts.

Once the annotations had been stored in the system, the recommendation engine provided a set of recommendations to informal mentees. Results were compared with the recommendations that a set of experts provided to the group of mentees using the same blog posts that were annotated.

4.2   Sample

The sample consisted of two groups of subjects. Twelve junior professionals were selected (eight men and four women) as mentees. The average age of mentees was 26.2 years. Subjects were selected from those who answered positively to a personal invitation sent by the authors to contacts working in the two companies. The sample was distributed equally among companies.

The second group of subjects (mentors) consisted of three experts (two women and one man) working in the software sector as human resource management professionals for more than 8 years, and their average age was 39.6 years. They provided the human recommendations, which were then compared with IM-TAG's recommendations.

4.3 Results

Table 1 shows results of recommendations process performed by both the IM-TAG tool and the experts. All recommendations were divided into three concepts: competencies, skills and experiences.

Table 1. Recommendations: IM-TAG vs. experts

	IM-TAG			Experts		
	A Company	B Company	TOTAL	A Company	B Company	TOTAL
Competencies	34	26	**60**	39	31	**70**
Skills	48	40	**88**	46	38	**84**
Experiences	13	19	**32**	20	26	**46**
TOTAL	**95**	**85**	**180**	**105**	**95**	**200**

IM-TAG provided a set of 180 recommendations - 95 for company A and 85 for company B - while the experts provided 200 recommendations - 105 for Company A and 95 for Company B. With respect to the distribution of recommendations by concepts, skills lead the ranking with 88 recommendations from IM-TAG and 84 from experts, followed by the concepts of competencies and experiences, respectively. In average, every user received 15 recommendations from IM-TAG and 16.7 from expert's side.

To evaluate the quality of recommendations from the two sources, precision, recall and F1 were used as measures. These metrics have been widely used to evaluate the quality of recommendations (Liu & Shih, 2005). A recommendation method may recommend interesting or uninteresting topics. The recall-metric indicated the effectiveness of a method for locating interesting topics, in our case, posts. The precision-metric represented the extent to which the items recommended by a method really are interesting to mentees compared to the ones detected by experts. F1 is calculated as the sum of the weights of precision and recall. Table 2 depicts the values of these three metrics with respect to recommendations from the IM-TAG tool and the experts.

Table 2. Precision, recall and F1 measures

	A Company	B Company	TOTAL
Precision	0,68421	0,51765	**0,60556**
Recall	0,61905	0,46316	**0,54500**
F1	0,65000	0,48889	**0,57368**

Results show differences between the measures by company. In fact, Company A present higher values for all measures. Besides this tendency, it is also worth noting the different values between precision and recall. This may be explained by the differences regarding total recommendations given by the experts and the IM-TAG tool (200 vs. 180).

It is also interesting to measure the usefulness of recommendations. This measure goes beyond accuracy and includes suitability of the recommendations to users. That

is, the coverage or percentage of a dataset that the recommender system is able to provide predictions for (Herlocker et al., 2004). The IM-TAG tool offered a result (coverage of the system) of 86.25%.

4.4 Discussion

The implementation of the IM-TAG tool offers reliable recommendations. Results show that the precision, recall and F1 values obtained outperform those from previous studies in the literature. In this sense, there is no study devoted to informal mentoring using these metrics, although extensive research in related areas already exists. Table 3 shows the comparison of IM-TAG's results with recent efforts in the literature regarding recommendations.

Table 3. Comparison of IM-TAG's accuracy with that of similar works

Reference	Domain	Precision	Recall	F1
IM-TAG	eMentoring Recommendations	0.606	0.545	0.574
García Crespo et al. (2011b)	Tourism	0.480	0.480	0.480
García-Crespo et al. (2010)	Customer Relationship Management	0.637	0.794	0.707
García-Crespo et al. (2012)	Portfolio Recommendations	0.320	0.320	0.320
García-Peñalvo et al., (2011b)	Software Engineering	0.824	0.741	0.780
Miao, Li & Dai (2009)	Sentiment Mining	0.876	0.874	0.874
Morales-del-Castillo et al. (2009)	Digital Libraries	0.500	0.707	0.582
Porcel & Herrera-Viedma (2010)	Digital Libraries	0.675	0.613	0.635
Zanker & Jessenitschnig (2009)	Conversion Rates	0.097	0.291	0.139

Results show that IM-TAG recommendation accuracy is remarkable. The comparison focuses on the F1 measure given that the complexity of the domain is high (it is important to note that the sources of information are not focused and must be classified using natural processing techniques). Results obtained from IM-TAG are comparable with some efforts (e.g. Morales-del-Castillo et al., 2009; Porcel & Herrera-Viedma, 2010), are better than others (e.g. García Crespo et al., 2011b; García-Crespo et al., 2012), and somewhat less accurate than others (e.g. García-Crespo et al., 2010; García-Peñalvo et al., 2011b; Miao, Li & Dai, 2009). The reason for the latter may be that these are more focused than the IM-TAG tool, so the domain in which the Natural Language Processing (NLP) Tools perform their job can be handled more precisely using automatic tools. In the case of IM-TAG, this is not possible, since it is based on the analysis of social web contents and these contents are naturally heterogeneous.

In spite of the acceptable results, there is still room for improvement. A possible origin of this misbalance could be rooted on the differences that can be found between Company A and Company B. A more in-depth analysis regarding the lack of accuracy in recommendations must be performed. Differences are important between companies, reaching 17% in F1 scores and are presented in all accuracy metrics. A deeper analysis of such differences should be performed in order to clarify if the

ontology definition was not fully applicable to this company or, for instance, if Company B contents are more difficult to process with the NLP engine.

5   Conclusions and Future Work

The Web 2.0 phenomenon has made the social Web possible, initiating an explosion in the number of Web users, but also empowering employees with a huge autonomy in adding content to web pages, labeling it, creating folksonomies of tags, and enabling millions of users to build their own web page (Breslin & Decker, 2007). The Web has started to play a fundamental role not only as a means for providing and searching for information, but also for creating and sharing knowledge (Eiter et al., 2008). In this environment, both users and organizations can benefit from the application of semantic technologies, since they enable the access and storage of information in a unified way. Semantic technologies have started to solve many challenging and cost-intensive problems within firms and are being positioned as one of the most important research topics within the current generation of Web applications (Nixon et al., 2008). In this sense, Focusing on knowledge management and human development issues, the IM-TAG offers a new solution to a well-known objective for the personnel development domain: the offering of wise advice to less experienced employees. This tool, based on semantic annotations of social web contents, provides content recommendations to users based on their profiles and tags, supporting informal mentoring, and as a consequence of this, informal learning. The validation conducted through a case study shows that the accuracy of IM-TAG's recommendations is notable and satisfactory. Thus, and agreeing with Clough et al (2007), it can be said that blogs and their contents can be considered as enablers for informal learning in organizational settings. However, and in order to fight against information overload or as defined in words of Mason and Rennie (2007), the overload of attention-grabbing opportunities., tools like IM_TAG can attract user's attention in a more direct and precise way.

Several authors have examined the use of micro-blogs to facilitate process-oriented learning and informal learning in higher education (e.g. Ebner et al., 2010) and its use within personal learning environments (e.g. Dabbagh & Kitsantas, 2012). Also, wikis are oriented to the transmission and formalization of knowledge. Thus, two separate lines of future research are proposed. First, the expansion of the IM-TAG tool to include other Web 2.0 contents: micro-blog posts and wiki articles and, second, the increase in the scope of IM-TAG to a multi-corporation environment. This expansion must be performed carefully, since culture spread is one of the main outputs of mentoring, and culture is different among organizations.

Acknowledgements

References

Allen, T. D. (2007). Mentoring relationships from the perspective of the mentor. In B. R. Ragins & K. E. Kram (Eds.). The handbook of mentoring at work: Theory, research and practice (pp. 123-147). Thousand Oaks, CA: Sage.

Bawden, D., & Robinson, L. (2009). The dark side of information: overload, anxiety and other paradoxes and pathologies. Journal of information science, 35(2), 180-191.

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D.L., Patel-Schneider, P. F., & Stein, L.A. (2004). OWLWeb Ontology Language Reference. http://www.w3.org/TR/owl-ref/

Benjamins, V.R., Davies, J., Baeza-Yates, R., Mika, P., Zaragoza, H., Greaves, Gómez-Pérez, J.M., Contreras, J., Domingue, J., & Fensel, D. (2008). Near-Term Prospects for Semantic Technologies. IEEE Intelligent Systems, 23(1), 76-88.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. Scientific American, 284(5), 34-43.

Bierema, L.L., & Merriam, S.B. (2002). E-mentoring: Using Computer Mediated Communication to Enhance the Mentoring Process. Innovative Higher Education, 26(3), 211-227.

Blake-Beard, S.D. (2001). Taking a hard look at formal mentoring programs. A consideration of potential challenges facing women. Journal of Management Development, 20(4), 331-345.

Blanco, C., Lasheras, J., Fernández-Medina, E., Valencia-García, R., & Toval, A. (2011). Basis for an integrated security ontology according to a systematic review of existing proposals. Computer Standards & Interfaces, 33(4), 372-388.

Bojārs , U., Breslin, J.G., Finn, A., & Decker, S. (2008). Using the Semantic Web for linking and reusing data across Web 2.0 communities. Web Semantics: Science, Services and Agents on the World Wide Web, 6(1), 21-28.

Bosua, R., & Scheepers, R. (2007). Towards a model to explain knowledge sharing in complex organizational environments. Knowledge Management Research & Practice, 5(2), 93–109.

Breslin, J.G., & Decker, S. (2007).The Future of Social Networks on the Internet: The Need for Semantics. IEEE Internet Computing, 11(6), 86-90.

Casado-Lumbreras, C., Colomo-Palacios, R., Soto-Acosta, P., & Misra, S. (2011). Culture dimensions in software development industry: The effects of mentoring. Scientific Research and Essays, 6(11), 2403-2412.

Castellanos-Nieves, D., Fernández-Breis, J.T., Valencia-García, R., Martínez-Béjar, R., & Iniesta-Moreno, M. (2011). Semantic Web Technologies for supporting learning assessment. Information Sciences, 181(9), 1517-1537.

Chao, G. T., Walz, P. M., & Gardner, P. D. (1992). Formal and informal mentorships: A comparison on mentoring functions and contrast with nonmentored counterparts. Personnel Psychology, 45 (3), 619-636.

Chen, T.Y. (2008). Knowledge sharing in virtual enterprises via an ontology-based access control approach. Computers in Industry, 59(5), 502-519.

Cleverdon, C.W., Mills, J., & Keen, E.M. (1966). Factors Determining the Performance of Indexing Systems, Cranfield, UK: College of Aeronautics.

Clough, G., Jones, A.C., McAndrew, P., & Scanlon, E. (2007). Informal learning with PDAs and smartphones. Journal of Computer Assisted Learning, 24 (5), 359-371.

Colomo-Palacios, R., Fernandes, E., Soto-Acosta, P., & Sabbagh, M. (2011). Software product evolution for Intellectual Capital Management: The case of Meta4 PeopleNet. International Journal of Information Management, 31(4), 395-399.

Colomo-Palacios, R., García-Crespo, Á., Soto-Acosta, P., Ruano-Mayoral, M., & Jiménez-López, D. (2010). A case analysis of semantic technologies for R&D intermediation information management. International Journal of Information Management, 30(5), 465-469.

Curtis B, Hefley WE, Miller SA (2009). People Capability Maturity Model (P-CMM®) Version 2.0, Second Edition. CMU/SEI-2009-TR-003.

Dabbagh, N., & Kitsantas, A. (2012). Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. The Internet and Higher Education, In press.

Dotsika, F. (2010). Semantic APIs: Scaling up towards the Semantic Web. International Journal of Information Management, 30(4), 335-342.

Ebner, M., Lienhardt, C., Rohs, M., & Meyer, I. (2010). Microblogs in higher education — A change to facilitate informal and process-oriented learning?. Computers in Education, 55(1), 92–100.

Eiter, T., Ianni, G., Lukasiewicz, T., Schindlauer, R., & Tompitsa, H. (2008). Combining answer set programming with description logics for the Semantic Web. Artificial Intelligence, 172 (12-13), 1495–1539.

Fensel, D. (2002). Ontologies: A silver bullet for knowledge management and electronic commerce. Berlin, Germany: Springer-Verlag.

Ferreira, N. (2010). Social Networks and Young People: A Case Study. International Journal of Human Capital and Information Technology Professionals, 1(4), 31-54.

Fong, P.S.W., & Choi, S.K.Y. (2009). The processes of knowledge management in professional services firms in the construction industry: a critical assessment of both theory and practice. Journal of Knowledge Management, 13(2), 110-126.

Frank, A., Bögel, T., Hellwig, O., & Reiter, N. (2012). Semantic Annotation for the Digital Humanities-Using Markov Logic for Annotation Consistency Control. Linguistic Issues in Language Technology, 7(8).

García-Bariocanal, E., Sicilia, M.A., Sánchez-Alonso, S., & Lytras, M. (2011). Semantic annotation of video fragments as learning objects: a case study with YouTube videos and the Gene Ontology. Interactive Learning Environments, 19(1), 25-44.

García-Crespo, A., Colomo-Palacios, R., Gómez-Berbís, J.M., & García-Sánchez, F. (2010b). SOLAR: Social Link Advanced Recommendation System. Future Generation Computer Systems, 26 (3), 374-380.

García-Crespo, A., Colomo-Palacios, R., Gómez-Berbís, J.M., & Mencke, M. (2009). BMR: Benchmarking Metrics Recommender for Personnel issues in Software Development Projects. International Journal of Computational Intelligence Systems, 2(3), 257-267.

García-Crespo, A., Colomo-Palacios, R., Gómez-Berbís, J.M., & Ruiz-Mezcua, B. (2010). SEMO: a framework for customer social networks analysis based on semantics. Journal of Information Technology, 25 (2), 178-188.

García-Crespo, Á., Gómez-Berbís, J.M., Colomo-Palacios, R., & García-Sánchez, F. (2011a). Digital Libraries and Web 3.0. The CallimachusDL Approach. Computers in Human Behavior, 27(4), 1424-1430.

García-Crespo, Á., López-Cuadrado, J.L., Colomo-Palacios, R., González-Carrasco, I. & Ruiz-Mezcua, B. (2011b) Sem-Fit: A semantic based expert system to provide recommendations in the tourism domain. Expert Systems with Applications, 38(10), 13310-13319.

García-Crespo, A., López-Cuadrado, J.L., González-Carrasco, I., Colomo-Palacios, R., & Ruiz-Mezcua, B. (2012). SINVLIO: Using Semantics and Fuzzy Logic to provide individual investment portfolio recommendations. Knowledge-Based Systems, In press.

García-Peñalvo, F.J., Conde, M.A., Alier, M., & Casany, M. (2011a). Opening Learning Management Systems to Personal Learning Environments. Journal of Universal Computer Science, 17(9), 1222—1240.

García-Peñalvo, F.J., Colomo-Palacios, R., Soto-Acosta, P., Martínez-Conesa, I., & Serradell-López, E. (2011b). SemSEDoc: Utilización de tecnologías semánticas en el aprovechamiento de los repositorios documentales de los proyectos de desarrollo de software. Information Research, 16(4) paper 504.

Gentry, W.A., Wever, T.J., & Sadri, G. (2008). Examining career-related mentoring and managerial performance across cultures: A multilevel analysis. Journal of Vocational Behavior, 72(2), 241-253.

Gruber, T. R. (1992). Ontolingua: A mechanism to Support Portable Ontologies. Stanford University, Knowledge Systems Laboratory, Technical Report KSL-91-66, March 1992.

Grybek, D. D. (1997). Mentoring the gifted and talented. Preventing School Failure, 41 (3), 115-119.

Haggard, D.L., Dougherty, T.W., Turban, D.B., & Wilbanks, J.E. (2011). Who Is a Mentor? A Review of Evolving Definitions and Implications for Research. Journal of Management, 37(1), 280-304.

Han, K.H., & Park, J.W. (2009). Process-centered knowledge model and enterprise ontology for the development of knowledge management system. Expert Systems with Applications, 36(4), 7441-7447.

Hitzler, P., & Janowicz, K. (2011). Semantic Web tools and system. Semantic Web, 2(1), 1-2.

Kammeyer-Mueller, J.D. & Judge, T.A. (2008). A quantitative review of mentoring research: Test of a model. Journal of Vocational Behavior, 72(3), 269-283.

Karkoulian, S., Halawi, L.A., & McCarthy, R.V. (2008). Knowledge management formal and informal mentoring. An empirical investigation in Lebanese banks. The Learning Organization, 15(5), 409-420.

Kirchmeyer, C. (2005). The Effects of Mentoring on Academic Careers over Time: Testing Performance and Political Perspectives. Human Relations, 58(5), 637–660.

Kirkwood, K. (2010). The SNAP Platform: social networking for academic purposes. Campus-Wide Information Systems, 27(3), 118-126.

Knouse, S.B. (2001). Virtual mentors: Mentoring on the Internet. Journal of Employment Counselling, 38(4), 162-169.

Kuo, M.H., Chen, L.C., & Liang, C.W. (2009). Building and evaluating a location-based service recommendation system with a preference adjustment mechanism. Expert Systems with Applications, 36 (2) Part 2, 3543–3554.

Kuswara, A.U., & Richards, D. (2011). Realising the potential of Web 2.0 for collaborative learning using affordances. Journal of Universal Computer Science, 17(2), 311-331.

Levy, M. (2009). Web 2.0 Implications on Knowledge Management. Journal of Knowledge Management, 13(1), 120-134.

Liang, T.P., Lai, H.J., & Ku, Y.C. (2006) Personalized content recommendation and user satisfaction: theoretical synthesis and empirical findings. Journal of Management Information Systems, 23(3), 45-70.

Liu, D.R., & Shih, Y.Y. (2005). Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences. Journal of Systems and Software, 77(2), 181–191.

Mason, R., & Rennie, F. (2007). Using Web 2.0 for learning in the community. The Internet and Higher Education, 10 (3), 196-203.

McAfee, A. (2006). Enterprise 2.0: The Dawn of Emergent Collaboration, MIT Sloan Management Review, 47(3), 21-28.

Miao, Q., Li, Q. & Dai, R. (2009). AMAZING: A sentiment mining and retrieval system. Expert Systems with Applications, 36(3), 7192-7198

Morales-del-Castillo, J.M., Peis, E., Moreno, J.M. & Herrera-Viedma, E. (2009). D-Fussion: a semantic selective disssemination of information service for the research community in digital libraries. Information Research, 14(2) paper 398.

Murugesan, S. (2007). Understanding Web 2.0. IT professional, 9(4), 34-41.

Nixon, L.J.B., Simperl, E., Krummenacher, R., & Martin-Recuerda, F. (2008). Tuplespace-based computing for the Semantic Web: a survey of the state-of-the-art.

Knowledge Engineering Review, 23(2), 181-212.

Noh, T.G., Park, S.B., Park, S.Y., & Lee, S.J. (2010). Learning the emergent knowledge from annotated blog postings. Web Semantics: Science, Services and Agents on the World Wide Web, 8(4), 329-339.

Paniagua-Martín, F., García-Crespo, Á., Colomo-Palacios, R., & Ruiz-Mezcua, B. (2011). SSAAMAR: Semantic Annotation Architecture for Accessible Multimedia Resources. IEEE Multimedia, 18(2), 16-25.

Park, Y.J., & Chang, K.N. (2009). Individual and group behavior-based customer profile model for personalized product recommendation. Expert Systems with Applications 38 (2, Part 1), 1932–1939.

Paroutis, S., & Al Saleh, A. (2009). Determinants of Knowledge Sharing Using Web 2.0 Technologies. Journal of Knowledge Management, 13(4), 52-63.

Poblet, M., Casanovas, P., López-Cobo, J.M., & Casellas, N. (2011). ODR, Ontologies, and Web 2.0. Journal of Universal Computer Science, 17(4), 618-634.

Porcel, C., & Herrera-Viedma, E. (2010). Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries. Knowledge-Based Systems, 23 (1), 32-39.

Raabe, B., & Beehr, T.A. (2003). Formal mentoring versus supervisor and coworker relationships: differences in perceptions and impact. Journal of Organizational Behavior, 24 (3), 271–293.

Ragins, B.R, & Cotton, J.L. (1991). Easier said than done: gender differences in perceived barriers to gaining a mentor. Academy of Management Journal, 34(4), 939-951.

Ragins, B.R. (1999). Mentor functions and outcomes: a comparison of men and women in formal and informal mentoring relationship. Journal of Applied Psychological, 84(4): 529-550.

Razmerita, L., Kirchner, K., & Sudzina, F. (2009). Personal knowledge management. The role of Web 2.0 tools for managing knowledge at individual and organisational levels. Online Information Review, 33(6), 1021-103.

Reiss K (2007). Leadership and coaching for educators, Corwin Press, Thousand Oaks, CA.

Richards, D. (2009). Social Software/Web 2.0 Approach to Collaborative Knowledge Engineering. Information Sciences, 179 (15), 2515-2523.

Rodríguez-González, A., García-Crespo, Á., Colomo-Palacios, R., Guildrís-Iglesias, F. & Gómez-Berbís, J.M. (2011). CAST: Using neural networks to improve trading systems based on technical analysis by means of the RSI financial indicator. Expert Systems with Applications 38(9), 11489–11500.

Schneckenberg, D. (2009). Web 2.0 and the empowerment of the knowledge worker. Journal of Knowledge Management, 13(6), 509-520.

Singh, R., Ragins, B.R., & Tharenou, P. (2009). What matters most? The relative role of mentoring and career capital in career success. Journal of Vocational

Behavior, 75(1), 56-67.

Siorpaes, K., & Simperl, E. (2010). Human Intelligence in the Process of Semantic Content Creation. World Wide Web, 13)1), 33-59.

Soto-Acosta, P., Casado-Lumbreras, C., & Cabezas-Isla, F. (2010). Shaping human capital in software development teams: the case of mentoring enabled by semantics. IET Software, 4(6), 445-452.

Srikantaiah, T.K., & Koenig, M.E.D. (2000). Knowledge Management for the Information Professional. Information Today Inc., Medford, NJ,

Stanton-Salazar, R.D. (2011). A Social Capital Framework for the Study of Institutional Agents and Their Role in the Empowerment of Low-Status Students and Youth. Youth & Society, 43(3), 1066-1109.

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: Principles and methods. Data & Knowledge Engineering, 25(1-2), 161–197.

Tsai, F.S. (2011). A tag-topic model for blog mining. Expert Systems with Applications, 38(5), 5330-5335.

Tsai, F.S., & Chan, K.L. (2010). Redundancy and novelty mining in the business blogosphere. The Learning Organization, 17(6), 490 – 499.

Uren, V.S., Cimiano, P., Iria, J., Handschuh, S.,Vargas-Vera, M., Motta, E., & Ciravegna, F. (2006). Semantic annotation for knowledge management: Requirements and a survey of the state of the art. Web Semantics: Science, Services and Agents on the World Wide Web, 4(1), 14-28.

Uschold, M., King, M., Moralee S., & Zorgios, Y. (1998). The enterprise ontology. The Knowledge Engineering Review, 13(1), 31–89.

Van Rijsbergen, C.J. (1979). Information Retrieval, Newton, MA: Butterworth-Heinemann.

Villela, K., Santos, G., Schnaider, L., Rocha, A.R., Horta Travassos, G. (2005). The Use of an Enterprise Ontology to Support Knowledge Management in Software Development Environments. Journal of the Brazilian Computer Society, 11(2), 45-59.

Wang, M. (2011). Integrating organizational, social, and individual perspectives in Web 2.0-based workplace e-learning. Information Systems Frontiers, 13(2), 191-205.

Warschauer, M., & Grimes, D. (2007). Audience, authorship, and artifact: the emergent semiotics of web 2.0. Annual Review of Applied Linguistics, 27(1), 1-23

Wheeler S. (2009). Learning Space Mashups: Combining Web 2.0 Tools to Create Collaborative and Reflective Learning Spaces. Future Internet, 1(1), 3-13.

Yang, S.J.H., Zhang, J., Su, A.Y.S., & Tsai, J.J.P. (2011). A collaborative multimedia annotation tool for enhancing knowledge sharing in CSCL. Interactive Learning Environments, 19(1), 45-62.

Yessad, A., Faron-Zucker, C., Dieng-Kuntz, R., & Laskri, M.T. (2011). Ontology-based semantic relatedness for detecting the relevance of learning

resources. Interactive Learning Environments, 19(1), 63-80.

Zanker, M. & Jessenitschnig, M. (2009). Case-studies on exploiting explicit customer requirements in recommender systems. User Modeling and User-Adapted Interaction, 19, 133-166.

Rodríguez-González, A., Labra-Gayo, J.E., Colomo-Palacios, R., Mayer, M.A., Gómez-Berbís, J.M., García-Crespo, A. (2012). SeDeLo: Using Semantics and Description Logics to Support Aided Clinical Diagnosis. Journal of Medical Systems, in press

Biographies

Ricardo Colomo-Palacios is an Associate Professor at the Computer Science Department of the Universidad Carlos III de Madrid. His research interests include applied research in Information Systems, Software Project Management, People in Software Projects and Social and Semantic Web. He received his PhD in Computer Science from the Universidad Politécnica de Madrid (2005). He also holds a MBA from the Instituto de Empresa (2002). He has been working as software engineer, project manager and software engineering consultant in several companies including Spanish IT leader INDRA. He is also an Editorial Board Member and Associate Editor for several international journals and conferences and Editor in Chief of International Journal of Human Capital and Information Technology Professionals.

Cristina Casado-Lumbreras is Adjunct Professor at Faculty of Education, Universidad Complutense of Madrid. She received her PhD in Psychology from the Universidad Autónoma of Madrid (2002). She has been working in several research institutions and Universities since nineties including Universidad Autónoma of Madrid, Universidad Complutense of Madrid and Escuela de Administración de Empresas (EAE). Her research interests include human capital development, staffing along with emotions in organizational contexts.

Pedro Soto-Acosta is a Professor of Management at the University of Murcia (Spain). He holds a PhD in Management Information Systems (MISs) and a Master's degree in Technology Management from the University of Murcia. He received his BA in Accounting and Finance from the Manchester Metropolitan University (UK) and his BA in Business Administration from the University of Murcia. He attended Postgraduate Courses at Harvard University (USA). His work has been published in journals such as the European Journal of Information Systems, the International Journal of Information Management, the Information Systems Management, and the Journal of Enterprise Information Management, among others.

Sanjay Misra obtained M.Tech. Degree in Software Engineering from Motilal Nehru National Natioanal Institue of Technology, Allahabad India and D.Phil. from University of Allahabad, India. His current researches cover the areas of: Software

quality, software measurement, software metrics, software process improvement, software project management, object oriented technologies, XML, SOA, Web Services, and cognitive informatics. He published more than 70 papers in these areas and got several awards for outstanding publications. He is founder of several international workshops annd is serving as editor in chief, associate editor, and editorial board member of several journals of international repute.