A Knowledge Management Framework
to Support Online Communities Creation

Catarina Lucena[1,2], João Sarraipa[1,2], and Ricardo Jardim-Goncalves[1,2]

[1] Departamento de Engenharia Electrotécnica, Faculdade de Ciências e Tecnologia, FCT,
Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
[2] Centre of Technology and Systems, CTS, UNINOVA, 2829-516 Caparica, Portugal
`{cml,jfss,rg}@uninova.pt`

Abstract. Nowadays, there are constant increases of people that actively produce and consume online content. They use various tools to create and share information with others, namely: blogs; forum; photo & video sharing and social networking sites. As a consequence the development of services or mechanisms for the acquisition of knowledge from Internet users is currently observed as crucial. However, most of such knowledge is often unstructured and is only available in a tacit form. Thus, this paper presents a framework to facilitate knowledge management to support communities' establishment. Its main objective is to build a community knowledge base from cyclical conversion of individual to organizational knowledge representations creation, and consequently, to allow specific reasoning and decision-making.

Keywords: Knowledge Management, Collective Awareness, Ontology.

1   Introduction

Social Web is a set of relationships that link together people over the Web [1]. Social Web encompasses how social software is designed and developed in order to support and foster social interactions [2]. This social software tools typically handle the capturing, storing and presentation of communication and focus on establishing and maintaining a connection among users [3]. Social Web focuses on human communication, and as consequence the vast majority of the information is in human readable format only. Thus software programs cannot understand and process this information, and much of the potential of the Web remains unusable. This leads to the necessity of having an approach or solution to concretize the semantic web concept.

The original *Scientific American* article on Semantic Web concept appeared on 2001 [4]. It described the evolution of a Web that consisted in a large amount of documents for humans to read and that included data and information for computers to manipulate [5]. Semantic Web can be seen as a set of technologies that help knowledge sharing across the web between different applications. Ontologies based technologies play a prominent role in the Semantic Web [6]. They make possible the widespread publication of machine understandable data, opening opportunities to automatic knowledge reasoning able to potentiate the generation of context awareness systems.

In this paper an initial assessment related to the necessity of gathering knowledge from social web applications is made. This gathering and storing of knowledge in a formalized structure will make possible to software applications to enrich the knowledge retrieved from individuals, adjusting it to specific objectives enabling its systems to have the notion of a context. This will contribute, for instance, to help users to find others to interact with about a specific purpose, which will facilitate to online communities creation and also to enable them to have contextual awareness able to make decisions supported by the knowledge handled. Thus, based on this, it is then proposed an ontology-based framework to contribute to such online communities creation. Afterwards, it is presented an architecture based in an application scenario of the proposed framework, followed by some conclusions and future work statements.

2   Relationship to Collective Awareness Systems

In today's internet, many applications arise in the context of "Web 2.0", a concept popularized by O'Reilly Media [7]. Web 2.0 focuses in collaboration of users and sharing information between them. On the one hand, Web 2.0 focuses on enhance creativity, information sharing and collaboration among users. On the other hand, Semantic Web refers to the intelligent interaction among systems and applications by deploying ontologies, semantic annotation of Web content, and reasoning. Its ultimate goal is to make data understandable to computers [8]. Together they contribute for software applications with knowledge from social software. This leads to the Web 3.0 appearance, where Semantic Web technologies are integrated into Web 2.0 technologies, or powering, large-scale Web applications [9].

Folksonomies are inventive Web 2.0 tools, which aggregates both web 2.0 and semantic web characteristics being great for categorizing documents and resources in a collaborative way [10]. They arise when a large number of people are interested in particular information and are encouraged to describe it [5]. The term was originally defined by Thomas Vander Wal [11] as the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The person consuming the information does the act of tagging.

By other hand, ontologies specify a conceptualization of a domain in terms of concepts, attributes, and relations [12]. Consequently, semantic web and its applications rely heavily on formal ontologies to structure data for comprehensive and transportable machine understanding. In addition, the Semantic Web's success is then dependent on the quality of its underlying ontologies [13], [14].

As mentioned before, social software normally shows limitations related to use or in the formalization of user's knowledge. For that reason a new framework to gather and formalize knowledge from social software, allowing reasoning and support to decision making is proposed (see Figure 1). This framework uses ontologies and folksonomy concept principles to enable knowledge management features, able to facilitate the achievement of collective awareness. Thus, this proposed framework represents the author's research, which contributes itself to collective awareness systems establishment. Collective awareness refers to a common and shared vision of the whole team's or community context that allows members to coordinate implicitly

their activities and behaviors through communication [15]. Making decisions with awareness will reduce the effort to coordinate tasks and resources by providing a context in which to interpret utterances and to anticipate actions [16]. This is also an objective that the framework intends to contribute to, where collective awareness is supported by the knowledge acquisition features from social networks through specific reasoning over that knowledge. This facet will facilitate communities' coordination by helping them in decision-making. Decision-making can be regarded as the cognitive process resulting in the selection of a specific resource or action among several alternative scenarios. Every decision making process produces a final choice [17].
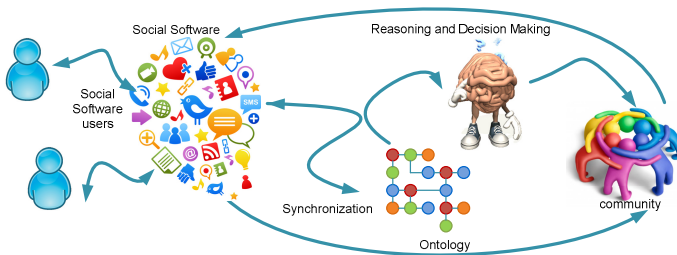

Fig. 1. Framework for Knowledge Management to support online communities creation

As can be seen in Figure 1, the proposed framework is composed by four modules: 1) Social Software; 2) Synchronization; 3) Ontology; and 4) Reasoning and Decision Making. The input of the framework is the Social Software users and the output is the community in general, which in this case raises some awareness using the knowledge represented in the system. The first module is social software, which is characterized by being collaboratively handled by social web users. This module can be a blog, wiki, forum, etc. However, as mentioned before, the content of social software is human readable only. This means that it is not formalized to facilitate its computerized use. For that reason, the synchronization module has incorporated some changes to that social software in order to transform its content to machine interpretable information. This synchronization is made between the knowledge represented in the used social software and the one in the ontology.

The role of the ontology in the proposed framework is to hold the gathered knowledge about a domain in a structured way, with a specific purpose to facilitate reasoning over that knowledge. Such reasoning is made by the module Reasoning and Decision Making, which is able to supply personalized or specific knowledge, representing to communities an access to structured contextual information. Contextual information is used by humans when communicate with each other. By improving computers access to such context, we increase the richness of communication human-computer and make it possible to produce more useful computational services [18]. Thus, these communities can, between other options, use the retrieved knowledge, and consequently to facilitate the interaction with other users with same interests, experience, etc.

3   ALTER-NATIVA's Knowledge Management Architecture

The proposed approach can be applied to various domains. In this case it is described the application of the framework to develop a Knowledge Management architecture for the ALTER-NATIVA community (Figure 2).

ALTER-NATIVA is one ALFA III project. ALFA is a program of cooperation between Higher Education Institutions of the European Union and Latin America. ALFA III retains the original objective of the previous phases of the ALFA Program, that is, to promote higher Education in Latin America as a means to contribute to the economic and social development of the region [19]. The main goals of the project were to: 1) Give education for everyone; 2) Give an environment of formation to professors when leading with persons with some disabilities; 3) Balance the inequalities of opportunities when accessing information.
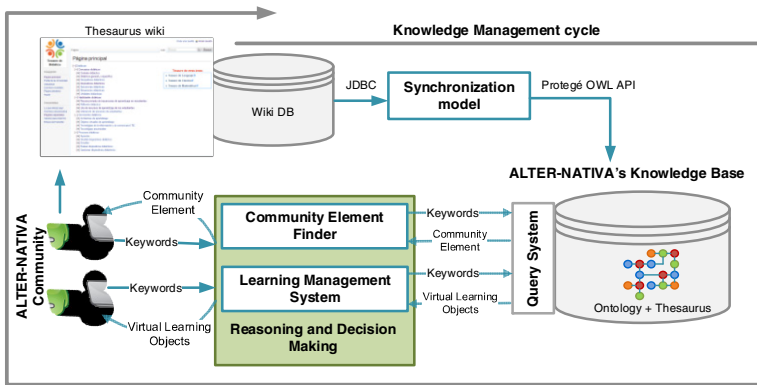

Fig. 2. ALTER-NATIVA's Knowledge Management architecture

In this architecture a complete Knowledge Management cycle was implemented. The beginning of the cycle or the input of the architecture is the knowledge of the ALTER-NATIVA community. This knowledge is processed and consumed by the community, which then uses it to re-feed this cycle with more knowledge. The gather of knowledge from the community is focused on the creation of a Thesaurus to represent the information related to the concepts/keywords (lexicon) of four distinct areas, namely Literature, Mathematics, Science and Teaching.

The Social Software used in the developed architecture was the Wikipedia. Wikipedia is an open encyclopedia that is collaboratively edited by its users. This software provides so called categories that are used to classify articles and other pages. The process of assigning categories to Wikipedia articles is a kind of collaborative tagging like in a folksonomy, but in this case, it is related to the categorization of concepts based on its definitions. One particularity is that the resulting category system is a thesaurus. A Thesaurus can be represented by a set of classes organized as a taxonomy representing domain reference concepts with associated meanings about domain in a semantic related structure. However, despite this strong usability, Wikipedia is not fully prepared to be integrated with other

information systems. For that reason, in ALTER-NATIVA were implemented and used two extensions to allow a structured insertion of knowledge. The MediaWiki community developed the CategoryTree extension. It facilitates an organized introduction of concepts in a hierarchical form, representing the wanted concept tree. The other extension was developed to avoid the insertion of data with HTML codification (e.g. in the middle of tags), which resulted in an easy and coherent knowledge insertion form for the community users. This extension is composed by four textboxes (Figure 3 a)): in the first textbox it is inserted the father of the concept (category of which this concept is member); in the second it is inserted the definition of the concept; in the third the creation date; and in the forth the author/authors identification.

The synchronization module of this architecture uses Wikipedia Data Base to detect any changes in the log of information that could have occurred in the Wiki Thesaurus since its last run. JDBC (Java Database Connectivity) is used to querying Wiki Thesaurus database. If any change occurred, it then updates the Knowledge Base's Thesaurus accordingly. The OWL API provides classes and methods to load and save OWL files, to query and manipulate OWL data models. This module runs periodically thanks to a "cron job" implemented in the Wikipedia's server. Thus, changes made by the users are constantly updated establishing in this way, a dynamic knowledge handling feature since its gathering from web users till its availability to the community.

The establishment of a Knowledge Base arises from the necessity to formalize the knowledge related to the ALTER-NATIVA's domain and allowing its communities to actively participate in its maintenance. The ALTER-NATIVA's Knowledge Base is split in two distinct parts: the "Ontology" and the "Thesaurus", which are connected by the relation "has keywords". This relation helps to characterize the resources of this community e.g. Virtual Learning Objects (VLO) making use of the knowledge gathered by the architecture. The Ontology is dedicated to represent the Knowledge of this community's resource. It main objective is to facilitate resources representation (categorization), searching and recommendation. Resources categorization uses the lexicon represented in the thesaurus. Each resource can be related to a set of concepts of the thesaurus.

The resources searching functionality use specific characteristics of the resource e.g. level of expertise of a professor, and the concepts that characterize them, which then can be used to find a professor with a pre-determined profile. The resources' recommendation functionality is an advanced feature that automatically records specific patterns usage information, as an example, a professor of Mathematics normally has interest in differential calculus. Thus, next time a professor of this kind of profile enters the system, the system automatically recommends to him a set of resources. It could be, for instance, a specific type of VLO related to differential calculus.

The ALTER-NATIVA Knowledge Base is able to support on the suggestion of users to interact with, based on common interests and levels of expertise potentiating communities' creation. This is done by the advanced services of the Community Element Finder of the architecture (Figure 2). Other feature is the recommendation of VLO to which is done by the advanced services of the module Learning Management System. Whether the contextual information is formal, seed by the keywords and

associated meaning or by the ontology, the expertise of the platform users is available for the implemented architecture be able to recommend with some context awareness. This could be reached by associating for instance the levels of expertise of users to the collected by associating the number of times that the user consults a specific topic. Then, by specific inference it is possible for the system to be aware of some context. One professor with a specific profile normally has interests or needs about a particular thing (e.g. VLO).
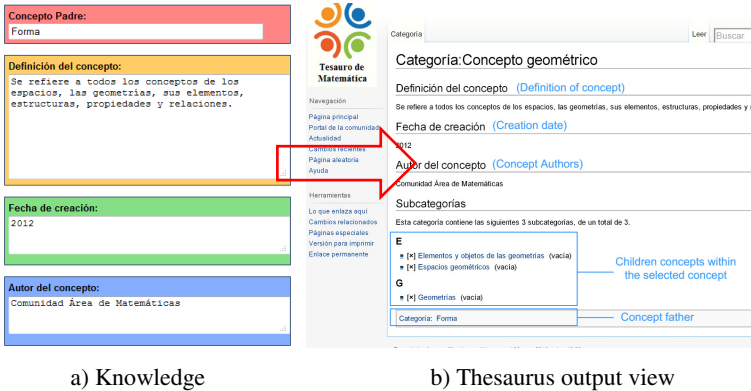

a) Knowledge   b) Thesaurus output view

Fig. 3. Knowledge gathering in the wiki insertion form

4 Demonstration Scenario

The ALTER-NATIVA's project main goal is to define curricular guidelines with technological support for higher education in the areas of language, mathematics and science with specific focus in supporting people in context of diversity. Thus, one of the main scenarios of this project pursue the idea of supporting teachers from different disciplines, many of them from pedagogical careers, with accessible tools that would facilitate them in the creation of working groups for discussion and for the generation of specific VLOs supported by a recommendation system [20]. As a consequence, the importance of developing services or mechanisms to gather knowledge of the domain experts was observed. Since in this scenario, professors are the main actors and better know how to characterize this project's resources: the Thesaurus for each of the ALTER-NATIVA project areas (literature, mathematics and science areas) gathered the knowledge directly from the professors.

An example of a contribution to the thesaurus creation is shown in Figure 3. An expert inserts the information corresponding to de concept "Concepto geométrico" in the form of Figure 3 a). When the user submits it, the output is the web page of Figure 3 b). After having the thesaurus with substantial contents, it is possible trough Community Element Finder services, search for specific profile of people. As an example, it is possible to search for professors that are experiment in "Geometric concepts" ("Concepto geométrico"). Another similar contribution in the possibility of search for a specific VLO characterized for keywords using the Learning Management System, as an example a VLO, which has the keyword "Geometric

concepts". The system is also able of advanced recommendations based on the historical information of users. Taking as an example, if a user consults several times a VLO, is most likely that he/she wants to consult other ones with similar characterization in terms of keywords. The same kind of recommendation is also made when searching for a user to interact with. If a user wants to develop a VLO characterized by a set of keywords, a possible suggestion of user to interact are other users that use to develop Virtual Learning Objects with the same keywords characterization.

5   Conclusions

The proposed framework establishes a set of components with specific characteristics to guide on the creation of software platforms able to support online communities' creation. Its main characteristic is related to its capability of gathering and formalization of knowledge from its users to afterwards use it for specific reasoning services. Due to the use of such knowledge, these services aggregate a collective view of a domain knowledge. Depending on its design, such services are able to provide its related systems with specific reasoning and Decision-making ability.

In the example of ALTER-NATIVA project the services were designed to recommend specific resources in relation to a determined profile of user. Thus they give some support to what is defined as "context awareness". Depending on the profile (context) of the user, the system is able (aware) to suggest the most appropriate resources to such kind of people, which will help them in their (online) community creation.

As future work, the authors want to integrate ontology learning in the framework. The concept Ontology learning represents methods and techniques used to build, semi-automatically or automatically, ontology from scratch, enriching, or adapting an existing ontology using several sources [21]. Thus, he authors' intend to create a prototype to allow an autonomous gather of knowledge from social software, which capable afterwards used to automatically build its related ontology. One example of a step in execution/implementation is to consider Wikipedia categories as ontological classes and the corresponding pages as instantiations of the knowledge base. This would allow a higher dynamic conversion from the individuals to community knowledge.

References

[1] Appelquist, D., Brickley, D., Carvahlo, M., Iannella, R., Passant, A., Perey, C., Story, H.: A Standards-based, Open and Privacy-aware Social Web. W3C Incubator Group Report, W3C (December 2010)
[2] Porter, J.: Designing for the social web. New Riders Press (2008)
[3] Allen, C.: Tracing the Evolution of Social Software (2004)
[4] Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Sci. Am. 284(5), 34–43 (2001)
[5] Shadbolt, N., Berners-Lee, T., Hall, W.: The Semantic Web Revisited. IEEE Intell. Syst. 21(3), 96–101 (2006)

[6] Taye, M.M.: Understanding Semantic Web and Ontologies: Theory and Applications. CoRR, vol. abs/1006.4 (2010)

[7] O'Reilly, T.: What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software (2005)

[8] Devedzic, V., Gasevic, D. (eds.): Web 2.0 & Semantic Web, vol. 6. Springer (2009)

[9] Hendler, J.: Web 3.0 Emerging. IEEE Comput. 42(1), 111–113 (2009)

[10] Kaminski, J.: EDITORIAL Folksonomies boost Web 2.0 Functionality 13(2), 1–8 (2009)

[11] Vander Wal, T.: Folksonomies definition and wikipedia (2005),
`http://www.vanderwal.net/random/entrysel.php?blog=1750`
(accessed: October 16, 2013)

[12] Fensel, D.: Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce, 2nd edn. Springer-Verlag New York, Inc., New York (2003)

[13] Maedche, A., Staab, S.: Ontology Learning for the Semantic Web. IEEE Intell. Syst. 16(2), 72–79 (2001)

[14] Hazman, M., El-Beltagy, S.R., Rafea, A.: Article: A Survey of Ontology Learning Approaches. Int. J. Comput. Appl. 22(8), 36–43 (2011)

[15] Daassi, M., Favier, M.: Groupware and team aware: Bridging the gap between technologies and human behaviour. Encycl. Virtual Communities Technol. (2005)

[16] Gutwin, C., Greenberg, S., Roseman, M.: Workspace Awareness in Real-Time Distributed Groupware: Framework, Widgets, and Evaluation. In: Proceedings of HCI on People and Computers XI, pp. 281–298 (1996)

[17] Reason, J.: Human Error, 302 p. Cambridge University Press, Cambridge (1990)

[18] Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a Better Understanding of Context and Context-Awareness. In: Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing, pp. 304–307 (1999)

[19] European Commision. Development and Cooperation - EUROPEAID,
`http://ec.europa.eu/europeaid/where/latin-america/`
`regional-cooperation/alfa/publications_en.html`
(accessed: October 17, 2013)

[20] Sarraipa, J., Baldiris, S., Fabregat, R., Jardim-Gonçalves, R.: Knowledge Representation in Support of Adaptable eLearning Services for All. In: Procedia CS, vol. 14, pp. 391–402 (2012)

[21] Gómez-Pérez, A., Manzano-Macho, D.: A survey of ontology learning methods and techniques (2003)