SOCIAL LEARNING: DEFINING LEARNING OBJECTS FROM SOCIAL TOOL

André Luís Andrade Menolli[1], Sheila Reinehr[2] and Andreia Malucelli[2]
[1]*Computer Science Department - State University of North Paraná – UENP - Polytechnic School*
Postgraduate Program in Computer Science - PPGIa Pontifical Catholic University of Paraná – PUCPR, Brazil
[2]*Polytechnic School, Postgraduate Program in Computer Science - PPGIa Pontifical Catholic University of Paraná – PUCPR, Curitiba, Brazil*

ABSTRACT

Social learning is focused for that learning occurs within a social context, in a place where people can work and learn collaboratively. New technologies, such as social network, wiki, blogs, among others social tools, enable collaborative work and are important facilitators of social learning process. These tools provide an easy mechanism for people to communicate and collaborate, which help in the creation of knowledge. However, collaboration is one of the several necessary components for learning. It is important that all acquired knowledge be organized to be reused faster, easily and efficiently. This paper aims to propose an approach to generate learning objects from social tool, in order to organize the information to be easily reused, improving social learning.

KEYWORDS

Social Learning, Learning Objects, Social Tool, Ontology, Web Semantic

1. INTRODUCTION

The advent of new technologies, known as Web 2.0 tools, like social network, wiki and blogs, facilitating the communication among people. O'Reilly (2011) said that the term "Web 2.0" means putting the user in the center, designing software that critically depends on its users, since the content is contributed by thousands or millions of users. This is the reason for the "Web 2.0" is also called the "participative Web". Wherefore, O'Reilly (2011) defined Web 2.0 as "the design of systems that get better the more people use them." Therefore, often the term participative "Web" or "Web 2.0" is used interchangeably with "social software" (Vassileva, 2009). In Web 2.0, the software recedes into the background; it provides the framework or the infrastructure, like the electricity or plumbing (Vassileva, 2009).

Through these new technologies and social environments, virtually anyone can create knowledge and make it available to be accessible and possibly useful to others. Therefore, the learning happens socially, with people creating and sharing knowledge dynamically. However, for social learning occurs some characteristics are required, such as trust in the social relationships and a way for discussions and ideas exchanging leading to collective knowledge construction. Hence, instead of designing technologies that "teach" the learner, the new social learning technologies will perform three main roles: 1) support the learner in finding the right content; 2) support the learner to connect the right people; and 3) motivate/incentivize people to learn (Vassileva, 2009).

Nevertheless, although most of social tools support collaborative work, this kind of tool does not provide means to achieve the required characteristics for social learning occurs satisfactorily. These tools often provide an efficient way to collaborate and create knowledge, such as wiki, that according to Kimmerle, et al. (2009) may help both the process of internalization and externalization of knowledge, using the constructivist approach (Piaget, 1977). Externalization occurs through the writing of texts, which leads to the realignment or improvement of cognitive schemes. Internalization occurs through bits of information from wiki, which are decoded and incorporated in internal structures of existing knowledge. This creates new knowledge entities in the person's cognitive system, new associations among knowledge entities and new schemes.

However, besides social tools, it is necessary other mechanisms for social learning takes place within

organizations. Thus, some authors present certain approaches to achieve this. Among works found in the literature, some have tried to improve communication among software development teams, while others have contributed to knowledge management in organizations and they are related to e-learning environments using semantic resources.

In this way, we can find some researches that focus in social learning. Vassileva (2009) illustrates how social learning technologies can be designed using some existing and emerging technologies: ontologies versus social tagging, exploratory search, collaborative versus self-managed social recommendations, trust and reputation mechanisms, mechanism design, and social visualization. Carreras, et al. (2009) cite that resources such as wiki, blogs and social networks are being used as substitutes for intranets within companies, creating an environment in which communication and collaboration among workers take place more effectively, offering a collaborative environment in which organizational learning is possible. Finally, Capuano, et al. (2009) introduce an e-learning solution called Intelligent Web Teacher (IWT), which is capable of modeling knowledge about educational domain, users preferences and competences through the Web Semantics approach, to develop customized and contextualized learning activities.

However, none of these researches is specifically focused on generating learning objects from social tool, as proposed in this paper. Polsani (2004) states that a learning object is a content independent and autonomous unit, which may be reused in several teaching contexts. We present an approach to generate learning objects from social tool, using semantic technologies, in order to allow users communicate, cooperate and dynamically create new content. Thereby, the knowledge may be organized didactically, instructionally and contextualized in specific domains, so that knowledge can be reused and easily available whenever necessary.

The remaining parts of the paper are organized as follows: Section 2 presents main technologies that can support social learning. Section 3 describes the approach to generate learning objects from social tool. Section 4 shows an experiment and section 5 concludes the paper.

2. NEW TECHNOLOGIES FOR SOCIAL LEARNING

In any organization, it is extremely important that the knowledge generated by the developers can be assimilated across the corporation, thus creating organizational learning. One way to foster this is to leverage the existing social relationships within the company to promote social learning. Social learning theory concentrates on the learning that occurs within a social context. It considers that people learn from each other, including such concepts as observational learning, imitation, and modeling (Ormrod, 1999). New technologies, which support collaborative work, may help create and facilitate social learning.

Each company must evaluate the available technologies in order to choose the most adequate to promote and facilitate learning in the organization. Besides the traditional technologies already being used by companies, such as, intranet, online communication tools, shared data banks and other technologies that give support to knowledge communication and storing, there are new technologies that may be applied in the search for social learning.

As the Web 2.0 technologies emerged, the process of building powerful social learning environments was simplified, because Web 2.0 is a new platform for developing Internet applications (Vassileva, 2009). The user is no longer a viewer, a recipient, or a consumer; the user is an actor, self-centered and rational (in the economical sense), but, surprisingly often, a collaborative and altruistic contributor (Vassileva, 2009).

Among the new Technologies, the Web 2.0 stands out since, according to Rech and Ras (2008), Web 2.0 technologies promote distributed collaboration, motivating the free reuse of information, experiences, or products and give support to knowledge workers by dealing with the information overload, integrating and reusing information spread out by several sources of content. Of all the Web 2.0 technologies, the most representative to support the social learning are (Rech and Ras, 2008): Social Network, Wikis, Web Blogs, Discourse Systems, Folksonomies and Mashups.

Besides the Web 2.0 technologies, other technologies may help in the construction of a social learning environment, such as ontologies. Another technique that may be used is text mining, an emerging research area interested in knowledge extraction process or interesting rather than trivial standards of the text documents (Tan, 1999).

3. PROPOSED APPROACH FOR GENERATION OF LEARNING OBJECTS FROM SOCIAL TOOL

In this work, we propose an approach to generate learning objects from social tool in a semi-automatic way, aiming to better organize the content put in this tools. This organization can facilitate the reuse of knowledge, and enabling the creation of units of learning (Polsani, 2004), maximizing the potential of these contents in educational factor.

For social learning to occur among distributed people through a web platform, some characteristics must be implemented by the environment that supports social learning. Therefore, a social learning environment is not just a set of tools and technologies, but rather an organized and configured environment to achieve the necessary characteristics for that learning occurs in a specific context (Menolli et al, 2011). Therefore, we projected set learning objects from content of social tools. To achieve this, we proposed an ontology called Ontology for Organizational Learning Objects (OOLO) (Menolli et al, 2012). It is based on the IEEE LOM standard (IEEE, 2002), and it has important properties to define a learning object, so that content may be reused (Menolli et al, 2012). For the establishment of OOLO, we analyzed LOM IEEE (IEEE, 2002), Dublin Core (NISO, 2007) and SCORM (ADL, 2004) standards. A mapping between the existing properties was established in order to find the correspondences.

The proposed ontology was need, because this work aims to generate learning objects in enterprise environments. The metadata standards studied were focused on e-learning environments and there are many differences between learning in educational and corporate environments (Menolli, 2011). Therefore, many of the properties described in the LOM were not used, as this is not a requirement for their use (IEEE, 2002). In addition, some attributes were added, like as: Source, Context, Use, Scope. The concepts and properties, proposed by OOLO, are presented in Table 1.

Table 1. OOLO basic structure. (Menolli et al, 2012)

Concepts	General	Life Cycle	Technical	Educational	Rights
Properties	Identifier Title Language Description Keyword	Version Status Role Date	Format Format Source Type Artifact Type	Interactivity Type Learning Resource Type Context	Copyright Use Scope

The learning objects also must include instructional role properties, which allow the classification and identification of the type of knowledge they provide to learners. So, we propose an instructional model in this approach to provide the type of instructional resource, which is fundamental to classify the knowledge, so helping in a learning process

This model provides a way to organize the instructional content, and it is defined through an instructional schema. Instructional elements are additional information that complements the concepts and other items of information already defined (Borges and Barbosa, 2009). Therefore, instructional model is created in a way that may guide learning, so, the knowledge inserted in the social tools will be classified and organized not only in domain area, but also instructionally, and such instructional element is represented in order to better understand and assimilate the domain knowledge, which may help in constructing knowledge. As example of instructional elements, Barbosa and Maldonado (2006) cite examples, suggestions, studies, exercises, tools, simulations, among others. Thus, this instructional scheme helps those who insert materials to classify them correctly.

3.1 Proposal Approach

Figure 1 shows our approach to generate learning object from social tool. First, users who are related through social networks, introduce content in social tools. The social network is proposed because it is a technology that may facilitate and improve the sharing information and maximize socialization. Social networks, especially trust networks, provide a new paradigm for knowledge management in which users “outsource” knowledge and beliefs via their social networks (Ding et al, 2003).

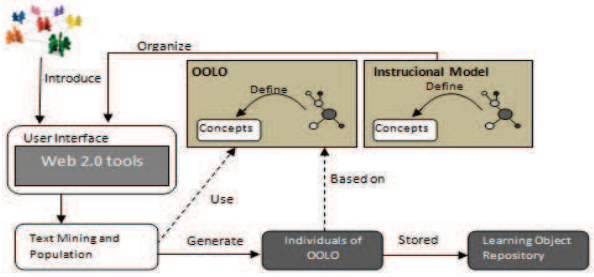

Figure 1. Schema to Generate Learning Objects from Social tools

From the content inserted into tools, a text-mining component is used in order to extract and classify attributes according to the OOLO and population component generated individuals in Organization Learning Object (OLO) format, after that these individuals are stored in a repository.

Some attributes, like educational attributes are not extracted automatically, for this reason, the tools utilized by users to insert contents, are prepared and organized according to the instructional model. This allows, when the content is being inserted, the user can choose the right educational and context attributes to classify it. Thus, the knowledge is organized and classified properly, which will facilitate the creation of learning objects.

Therefore, after all attributes have been found, some automatically and others manually, it is generated individuals that defines the content as an object OLO.

The challenge of this proposal is that for each type of social tool it is necessary a different implementation to organize the content as learning objects. Among all Web 2.0 tools, we choose Wiki to evaluate our approach.

4. EXPERIMENTS

Although there are several Web 2.0 tools that provide social learning, facilitating the creation of content, in this work we choose wikis to implement and evaluate our proposal. This choice was mainly due to the existence of Wikipedia.

Wikipedia was selected in the work as content source because it is one of the world's largest collaborative knowledge bases. Most web users know Wikipedia due to its high visibility from major search engines. Although there are only a few contributors (less than 10% of all users) to the content of Wikipedia (Priedhorsky et al, 2007), it has a huge pool of readers due to its high accessibility.

The use of Wikipedia to generate learning objects can be justified based on its features in terms of research value. Among these features, we may highlight quality, diversity, associations and dynamics (Chang and Quiroga, 2010).

A comparison regarding content quality between Wikipedia and Britannica has been made and reported. The study found a similar level of quality for both encyclopedias, which reveals the potential of Wikipedia (Giles, 2005).

4.1 Analyzing the Proposal Approach

Aiming at analyzing and evaluating the proposed approach, an implementation was done using Wikipedia. Wikipedia pages were selected in the Software Engineering area, in order to define OLO for each page. The analyzed pages were in the Portuguese language. In order to let the proposal approach implementation clear, we use the following definitions:

Definition 1 (Web document) A Web document or Web page *d* is a semi-structured document, following the Wikipedia format.

Definition 2 (Tags Content) A tag content is a content inserted into a HTML tag. The content *p* in a tag *t* is the content from a root tag element *t* until the end of the tag *t*, *<t> p</t>*.

In this Wiki approach, the OOLO attributes were extracted from three ways: automatically from the page properties; automatically from the page content; and manually by users.

4.1.1 Attributes Extracted Automatically from the Page Properties

As we are working with Wikipedia and just Portuguese pages, some properties are always the same, as *Language*, *Source*, *Interactivity Type* and *Copyright*. Others were extracted from URL, like *Location* and *Format*. In addition, some properties were created automatically by the text mining and population component, like *Version* and *Identifier*.

Finally, some properties were not captured because the proposed approach is forecast to be used in an enterprise wiki. Therefore, we could have access to the user's records and some properties could be assigned automatically, for instance, *Status* and *Role*.

4.1.2 Attributes Extracted Automatically from the Page Content

After the contents have been inserted into the Wiki tool, this content must be analyzed and manipulated so that some OOLO attributes are extracted. Thus, the pages texts were analyzed, and for each distinct attribute, a different approach was performed.

To extract the *Title* and *Date* attributes it was used Regular Expressions, since with the *Title*, there are HTML tags that identify it, and *Date* is always preceded by a standard text. Regular expressions themselves, with a general pattern notation almost like a mini programming language, allow you to describe and parse text. With additional support provided by the particular tool being used, regular expressions can add, remove, isolate, and generally fold, spindle, and mutilate all kinds of text and data (Friedl, 1997).

To extract *Keyword* attribute it was used text mining. The text mining refers generally to the process of extracting interesting and non-trivial patterns or knowledge from unstructured text documents. It can be viewed as an extension of data mining or knowledge discovery from (structured) databases (Tan, 1999).

Hence, for the extraction of keywords using the **Definition 2,** we first extract tag content from categorical labels content - tags <*h2*> and <*h3*>*;* links content - tag <*a href*>; terms in bold italics - tags <*b*> and <*i*>. After that, we extract terms with a high Term Frequency.

We used the *tf–idf* weight (term frequency–inverse document frequency) (Aizawa, 2003), to extract the terms with a high Term Frequency. In our experiment, *tf-id* is defined as:

Definition 3 (Tf-id) are the terms t in a document d, which reflects how important the term t is to a document d. The tf-id weight of t increases with the number of times the term t appears in the document d.

Then, was created a algorithm to extract a collection of keywords $A = \{k_1,\ldots k_n\}$. Therefore, the algorithm 1 reflects the way A is obtained from d.

Algorithm 1 Extract keywords from a wiki page

Input: Web document *d*
Output: Collection $A(k, v)$ of set of keywords and their values
1: extract from *d* the set $Z=Z(D)$ of *t*, where *t* is a tag content , according to Definition 2
2: extract from *d* the set $Y= Y(D)$ of (f, w)-pairs, $Y=\{(f, w)_1,...(f, w)_{10}\}$, where *f* is a frequent term, according to Definition 3, and *w* is the weight of *f* in *d*
3: $A=\emptyset$
4: **for all** $t \in Z$ **do**
5: value =0
6: **for all** $f \in Y$ **do**
7: **if** $f \subset t$ **do**
8: $value = value + w$
9: **end if**
10: **end for**
11: insert (*t, value)* to *A*
12: **end for**
13: remove similarTerms(*A*)
return *A*

The last step of algorithm 1 was remove similar terms. This method was implemented using Jaccard similarity. If two terms have a similarity not less than 80%, the term that has the lowest value is discarded.

To extract the *Description* attribute it was necessary a different approach. The textual description of the contents of a learning object should be in a language and terms appropriate for those that decide whether the learning object being described is appropriate and relevant for the users. Therefore, it was used automatic text summarization to extract the description attribute.

Summary is defined as a "text that is produced from one or more texts, that conveys important information in the original text(s), and that is no longer than half of the original text(s) and usually significantly less than that" (Radev et al, 2002).

As we were working with pages in Portuguese language, it was necessary to find specific tools to this language. In Rino, et al (2004) it is presented a comparison of automatic summarizers of texts in Brazilian Portuguese language. After downloading and performing tests with some tools, we choose the GistSumm System.

4.1.3 Attributes Extracted Manually

Some attributes need to be set manually in page creation, such as *Context*, *Type*, *Use*, *Scope*, and *Learning Resource T*ype. Figure 2 shows how this could be done adapting a semantic wiki.

Figure 2 presents a screen of an adapted wiki. When the users insert content, it is necessary to choose the *Context* and *Learning Resource Type* properties. These properties are defined in the instructional model. After this, the knowledge is already classified according to its instructional role. We have done an experiment, using Mediawiki and Semanticwiki, in order to verify the viability of adapting a wiki in relation to the instructional model. However, we did not use this implementation to define the OLO objects in our initial experiments, since we use existing pages from Wikipedia.

4.2 Results

As main result, learning objects were defined using OOLO format. First, we created the Ontology for Organizational Learning Objects using the ontology editor *Protegé*. After, by the implementation described previously, we extracted attributes from 15 wiki pages, and created individuals in an OOLO for each wiki page as showed in Figure 3.

First, as we use Wikipedia pages, some attributes could be extracted automatically without any problems. It is possible because all Wikipedia pages are structured similarly and contains some categorical and temporal metadata that help in the extraction of some attributes. In addition, the HTML tags enable to find some important properties inside the content.

Therefore, particularly to wiki pages, the attributes *Title*, *Data* and *Location* were extracted without any problems. However, to other kind of Web 2.0 tool, probably, this approach will not properly work, needing adaption, or even a new approach.

To the attributes that were extracted automatically, there are two main concerns. First, the quality of content and second, the method used to extract attributes.

If the page content is not consistent with the real purpose, or it is very poor, it is not possible to extract a description that provides subsidies to users decide whether the learning object described is appropriate and relevant. Therefore, the description attribute is dependent of summarization tool and of the content quality.

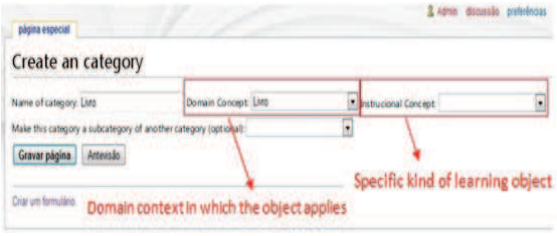

Figure 2. Addition of Context and Learning Resource Type in a Wiki

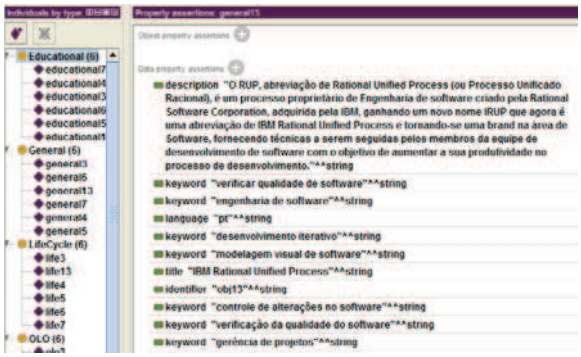

Figure 3. Individuals of OOLO in Protegé Editor

The *Keyword* attribute also depends on the content quality, since terms with a high Term Frequency is one of the methods used to extract this attribute.

For easy reuse of learning objects, besides the *Title*, *Description*, *Keyword*, some other attributes need to be well defined, like as context and learning resource type. Nevertheless, to extract these attributes, it was not assigned automatic extraction method, but proposed to modify the social tool, so that at the moment of inclusion content, the user must define these attributes. We consider that these attributes need to be classified according to a model, so that they will be easily organized, helping in a future reuse.

Hence, we consider that with the appropriated modifications in the social tool, enabling user inserts correctly the instructional and context attributes, and with quality content, it is possible to define appropriate learning objects. Thus, the content reuse can be facilitated, or even materials could be created from these objects, as online course.

5. FINAL CONSIDERATIONS

The work presented here focus on the creation of learning object from content inserted into social tool, aiming to organize the knowledge for facilitating its reuse. This is necessary because more and more these tools enhance the way of people communicate, but often do not care how the knowledge generated by them is used.

The use of collaborative tools is a trend in software development companies. The use of such tools, which assists in the creation of knowledge, mainly through social learning, is having a large growth and provides important resources that can enhance organizational learning. However, the knowledge organization can be improved. To do this, we propose to organize the content in learning objects. So, it can improve the search of content and help to create a logical sequence of contents, which can facilitate the learning from contents generated within a company, increasing the organizational learning.

For that learning objects creation to be possible, it is necessary to use a different approach for each Web 2.0 tool, which is a limitation of our proposal. Besides, it is necessary to modify the Web tool, in order that it will be possible to classify the context domain and the instructional role of the content in the moment that it is created.

As the scope of our work is to create learning object for organizational environment, mainly software engineering companies, it was necessary to define an ontology based on LOM standard and to add some specifics properties of this kind of environment.

To validate our proposal there was as preliminary implementation using Wiki. Wikipedia pages were selected in the Software Engineering domain, in order to define OLO objects for each page. To extract each OLO attribute, different techniques were necessary, as text mining, regular expressions and text summarization, and in some cases, it was not possible to extract the attributes automatically.

Therefore, we may conclude that it is possible to extract data that represent attributes defined by OOLO and created OLO individual from content inserted in the social tool as learning objects. Thus, it is possible to use wiki pages as learning object in any Learning Management System that accept LOM format. Furthermore, this approach can be used in any Wiki based on Mediawiki, and can be extensible to other kinds of Web 2.0 tools.

The next step is to generate learning units from defined objects, so creating materials in a correct order to a specific domain. Thus, the contents inserted by users may be organized instructionally, didactically and in an appropriate sequence, helping in the organizational learning and improving social learning.

REFERENCES

ADL, Advanced Distributed Learning, 2004. *Scorm 2004 4th Edition*, http://www.adlnet.gov/Technologies/scorm/default.aspx.

Aizawa, A. 2003. An information-theoretic perspective of tf–idf measures, I*n Information Processing and Management*, Vol. 39, pp. 45–65.

Barbosa, E. F.and Maldonado, J. C., 2006. An integrated content modeling approach for educational modules, *Proceedings of World Computer Congress – International Conference on Education for the 21st Century*, Santiago, Chile, pp. 17–26.

Borges, V. A. and Barbosa, E. F., 2009. A Contribution to Modeling and Automatic Generation of Educational Content, *Proceedings of Workshop sobre Educação em Computação*, Bento Gonçaves, Brasil.

Capuano, N., et al, 2009. IWT: A Semantic Web-based Educational System. *Proceedings of Workshop of the Working Group on*, pp. 11-16.

Carreras M., et al, 2011. Towards a Semantic-Aware Collaborative Working Environment. *In Computing and Informatics*, Vol. 30, No 1, pp. 7-30.

Chang, P. C. and Quiroga, L. M., 2010. Using Wikipedia's Content for Cross-Website Page Recommendations that Consider Serendipity, *Proceedings of Conference on Technologies and Applications of Artificial Intelligence*.

Ding, L. et al, 2003. Trust Based Knowledge Outsourcing for Semantic Web Agents, *Proceedings of Int. Conf. Web Intelligence*, IEEE CS Press, pp. 379–387.

Friedl, J. F., 1997 *Mastering Regular Expressions: Powerful Techniques for Perl and Other Tools*, O'REILLY, USA.

Giles, J. 2005. Internet encyclopedias go head to head, *In Nature*, Vol. 438, No 70, pp. 900-901.

IEEE, 2002. *Draft Standard for Learning Object Metadata*, http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

Kimmerle, J. et al, 2009. Learning and Knowledge Building with Social Software, *Proceedings of International Conference On Computer Supported Collaborative Learning*, pp. 459-468.

Menolli, A. L. et al, 2011. Towards a Semantic Social Collaborative Environment for Organizational Learning, *Proceedings of International Conference on Information Technology and Applications*, IEEE Press, pp. 65-70, Australia.

Menolli, A. L. et al, 2012. Ontology for Organizational Learning Objects based on LOM Standard, *Accepted at Latin American Conference en Informática, CLEI*, Colombia.

NISO, 2007. *Niso Standards. ANSI/NISO Z39.85 - The Dublin Core Metadata Element Set*, http://www.niso.org/kst/reports/standards?step=2&gid=&project_key=9b7bffcd2daeca6198b4ee5a848f9beec2f600e..

O'Reilly, T., 2001. *On the Future of Social Media. Talk of the Nation Science*, http://www.npr.org/templates/story/story.php?storyId=98499899.

Ormrod, J. E., 1999. *Human learning*. Prentice-Hall, Upper Saddle River, USA.

Piaget, J., 1977. Problems of equilibration, *Proceedings of Topics In Cognitive Development*," Vol. 1, New York, pp. 3-14.

Polsani, P. R., 2004. Use and abuse of reusable learning objects, *In Journal of Digital information*, Vol. 3-4.

Priedhorsky, R. et al, 2007. Creating, destroying, and restoring value in Wikipedia, *Proceedings of International ACM conference on Supporting Group Work*.

Radev, D. R. et al, 2002. Introduction to the special issue on summarization, *In Computational Linguistics*, Vol. 28, No 4, pp. 399-408.

Rech, J. and Ras, E., 2008. The Future of Learning Software Organizations: Semantics – Collaboration – Aggregation, *Proceedings of Workshop on Learning Software Organizations*, Rome, Italy.

Rino, L. H. et al, 2004. A Comparison of Automatic Summarizers of Texts in Brazilian Portuguese, *Proceedings of Brazilian Symposium on Artificial Intelligence*, Brazil.

Tan, A., 1999. Text Mining: The state of the art and the challenges, *Proceedings of* W*orkshop on Knowledge Disocovery from Advanced Databases*, pp. 65-70.

Vassileva, J., 2009. Toward social learning environments. *In IEEE Transactions on Learning Technologies*, Vol. 1, No4, pp.199–214.