L Zaina, J Rodrigues-Jr, A Amaral (2015) **Social tagging for e-learning: an approach based on the triplet of learners, learning objects and tags** In: *Learning Technology for Education in Cloud*, 104-115 Springer isbn:978-3-319-22629-3.

Original version: http://link.springer.com/chapter/10.1007%2F978-3-319-22629-3_9

Social tagging for e-learning: an approach based on the triplet of learners, learning objects and tags

Luciana A M Zaina[1] – José F. Rodrigues Júnior[2] – Anderson R. do Amaral[1,3]

[1]Federal University of São Carlos – Rdv João L. Santos, Km110 – Sorocaba – SP – Brazil
[2]University of São Paulo –Av. Trabalhador São Carlense, 400 – São Carlos – SP – Brazil
[3]ETEC FernandoPrestes - R. Natal, 340 – Sorocaba – SP – Brazil

lzaina@ufscar.br – junio@icmc.usp.br – anderson.amaral@etec.sp.gov.br

Abstract. The emerging of the Web 2.0 has allowed users more interactivity with Web applications. Social tagging has been recognized as an important solution to the description of resources available on the Web. In the context of e-learning it may be used as an auxiliary mechanism to the composition of learning object metadata. This paper presents an approach based on the triplet of learners, learning objects and tags for providing the social tagging for e-learning. We performed an experiment with 336 technician students that marked 218 electronic learning objects for about 4,985 times. Although our results have shown that social tagging is a promising practice for e-learning some challenges on how to implement it has to be overcome.

Keywords: folksonomy, social tagging, learning objects, Web 2.0

1 Introduction

The term Web 2.0 has been created to refer to a new generation of Web applications mainly characterized by providing support for collaboration and sharing of user-generated content [13]. Usually, companies developing applications for Web 2.0 use the Web as a platform to create collaborative and community-based websites, such as social networks, blogs, wikis, and others. The idea is to make the online environment more dynamic, where users can play a more active role and work together for producing and organizing the content, unlike the traditional Web (Web 1.0) where users were mostly readers of information.

The use of Social Web applications for different purposes increase since the emerging of Web 2.0. Specially, applications that promote the interrelationship of people and of knowledge through the Web have gained popularity. The social networks have been seen as an e-learning environment [22]. Accordingly, academic and commercial e-learning systems have adapted to the characteristics of these applications, becoming more appealing to students [5].

In the social Web, a remarkable feature is the possibility of tagging online content, what allows the users to create vocabularies that categorize the resources - or learning objects [9] - they interact with. From the freedom of marking the objects arise the concept of folksonomy

L. Uden et al. (Eds.): LTEC 2015, CCIS 533, pp. 104-115, 2015
DOI: 10.1007/97

(folk, as for people; sonomy, as for taxonomy), which can be defined as a classification system outlined by people, without rules for terms' creation [3]. In the folksonomy the users freely choose keywords (also called tags) to identify, describe and classify the resources [19]. The folksonomy manifests in the form of social tagging systems [21]. In the practice of social tagging, users collaboratively use tags to annotate and make sense of content, a valuable source of information that has the potential of bringing order (indexing and classification, or cataloguing) to vast volumes of information [1].

Social tagging systems assume that users will express their impressions by means of tags that they use to classify the content they use [11]. Social tagging fulfills the impracticable classification that would be performed by specialists [7], its main features include: flexibility, as the users use their everyday dynamic vocabulary; pattern identification, as the users spontaneously choose the words that best describe the content; and collaboration, as predicted by Social Web applications.

Within the context of electronic learning (e-learning), the tagging systems may provide a process for indexing the resources based on the tags attributed from the user (teacher or student). The educational resorces usually are called learning objects (LOs). Formally, learning objects refer to entities used in the teaching-learning process; videos, images, simulator software, and text, among other possibilities. In the electronic-learning domain, it is desirable that learning objects be reusable for different learning objectives, or be combined to build up more complex objects [10]. To this end, the objects must carry metadata that contextualize and describe their use in a standard manner [20]. The main standard for learning objects is the IEEE LOM (Learning Object Metadata) [20]. Each category has a specific purpose, such as describing general attributes of objects, and educational objectives.

The possibility of tagging system is converging as a new model for cataloging learning objects based on the tags provided either by students or by professors. In the case of students, the tagging process concerns a reflection experience in which students tag the objects based on their own experience [2,15]. Then, the repositories of tagged learning objects can be searched by the very students or by other people in the course of learning [4]. For achieving a cataloging that effectively describes the learning objects is mandatory to follow a tagging process that fulfill learning purposes. In e-learning area the vocabulary of tags must be suitably heterogeneous in order to extensively describe the objects.

The goal of this work is to present an approach based on the triplet of learners, learning objects and tags aiming to provide the social tagging for e-learning. To empirically examine the proposal, we performed an experiment with 336 technician students who have tagged 218 learning objects marked 4,985 times. In our experiment, we did not use a particular pedagogical learning model. But we believe that our proposal can be extended to different pedagogical theories. The focus is the cataloging of learning objects which facilitates the access to such objects.

The remainder of this paper is organized as follows: Section 2 presents the related work; the proposed model is presented in Section 3; Section 4 outlines the experiment results; finally, Section 5 remarks the main conclusions.

2 Related Works

Social tagging systems grew in popularity in the last years due to their simplicity to categorize and retrieve content based on tags. The increase in the number of users that provide information to such mechanisms caused the emergence of systems that assume the users express their preferences by means of the tags they create and use [11].

The main features of social tagging are: communication and immediate feedback, fast adaptation to vocabulary alterations, single or collective organization of objects, potential of cataloging, and assistance in the recommendation of content, among others. The tags entered in the system allow users to freely explore objects and other users' profiles without having to follow a rigid predefined hierarchy of concepts [14].

A collaborative model for construction of learning object repository is proposed by Monge, Ovelar and Azpeitia [12] which pointed out that the educational materials should be available in an open and multidisciplinary environment. They consider that technique of social tagging adds a rich-semantic for the description of a material that can improve the social dynamics of learning repositories, building - teachers, students and research institutes - a large network of knowledge. The approach proposes by Sierra and Valmayor [16] aim at the creation and extension of metadata of learning objects using the social tagging to pinpoint the attributes of the metadata. We consider the idea of cataloging the learning objects using the tags an opportunity to create large repositories that reflect the perception of the users.

Bateman, Brooks and Mccalla [2] analyze social tagging applied to e-learning by using tags collected from the interaction of students and professors with learning objects. In their study, they observe that the professors use a more specialized terminology than the students, and that an initial set of tags (a seed) must be provided during the earlier stages of the system. We follow their advice providing our students with such a set during the experiments; differently, we further discuss the behavior of the students, and the resulting vocabularies of tags, tracing some relevant hypotheses.

In a recent work, Zervas and Sampson [23] evaluate how the motivation affects the enlargement of tagged learning repositories. Although they discuss some interesting issues about the influence of the profiles of the students, they do not put conclusive considerations about this interesting topic. In our research, we follow a similar investigation to settle, as much as possible, revealing remarks about how social tagging occurs in the educational domain.

Sinclair and Cardew-Hall [17] claimed that a tag cloud reflects the folksonomy's vocabulary in a social navigation tool, reducing the cognitive effort on the understanding of a tag and promoting the relationship of tag to learning objects. In a review of social tagging, Gupta et al [7] suggest that in the representation of the tag cloud can be used font sizes and colors to draw a distinguish on the frequency of the use of tags. In our work we adopted the visualization through tag clouds where the font size increases according the use of the tag.

3 e-Learning Social Tagging Approach

Based on the presented concepts and related works, we have defined a model to conceive the tagging system as a triplet made of students, tags, and learning objects that, together, interact to form a descriptive vocabulary (Figure 1). The descriptive vocabulary represents the stu-

dents' understanding about the topics of the learning object adding to the repository the feature of social dynamics as proposed by Sierra and Valmayor [16].

The proposed model introduces an evolutive cycle that allows the learners the creation and the reuse of tags. The cycle starts from the search by t and retrieval of learning objects (1); proceeds to the creation and reuse of tags (2); and evolves through the refinement achieved by the association of tags to the objects (3). These three steps cycle for some indefinite time, during which a repository of tags is built. The repository is organized as two sets: "my tags" - the tags created by a specific user who is logged in the system; "global tags" - the universe of all the distinct tags created by all users for a specific learning object. Along time, some tags are recurrently used for describing the learning objects. The convergence of the tags to a stable descriptive set defines a vocabulary (4).

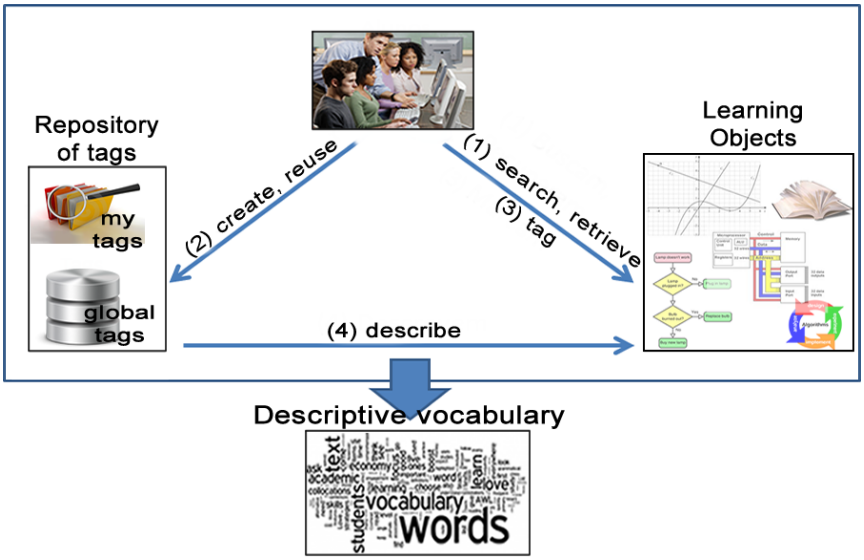

Figure 1. e-Learning Social Tagging Model.

In the model proposed the student becomes an active agent in the process of learning objects description, contributing to the maintenance and organization of objects. The repository of tags can be seen as a "living" dictionary, constantly updated by the users. The use of tags as descriptors of objects favors the indexing, navigation and recommendation of different learning objects.

After we drew the cycle, we seek for tools that matched to our model and could support our experiment. We noticed that previous works did not offer an adequate (open source and accessible) environment for experimentation according to our model. For this reason, our research group designed and developed the TagLink tool [18], which is able to (1) retrieve learning objects, (2) create/reuse tags, (3) support the tagging of objects, (4) manage a descriptive vocabulary of tags, and (5) display the tag in a tag cloud.

TagLink was designed following the scheme presented in Figure 1. It allows students to search and retrieve learning objects from the Web using Google's Custom Search API [6]. To do so, the students provide search terms and TagLink returns the objects and their corre-

sponding links. Each link is processed as a learning object to which students can associate tags.

TagLink tool is configured with a set of relevant repositories of learning objects. The repositories are registered in TagLink together with a priority indicator that specifies in which order they are to be searched. The step for tagging the learning objects is showed in Figure 2. First the student search the keyword (A) and the results returned by TagLink correspond, each, to one learning object; the student select the object that it (B) can be tagged with a new (E) or with an existing tag (C) and (D). The field "tags of the object" (C) shows all the tags that were attributed to the object, regardless of who did; and the field "my tags" (D) reports the tags that the student has been used in any object. These attributes encourages the vocabulary construction collaboratively.

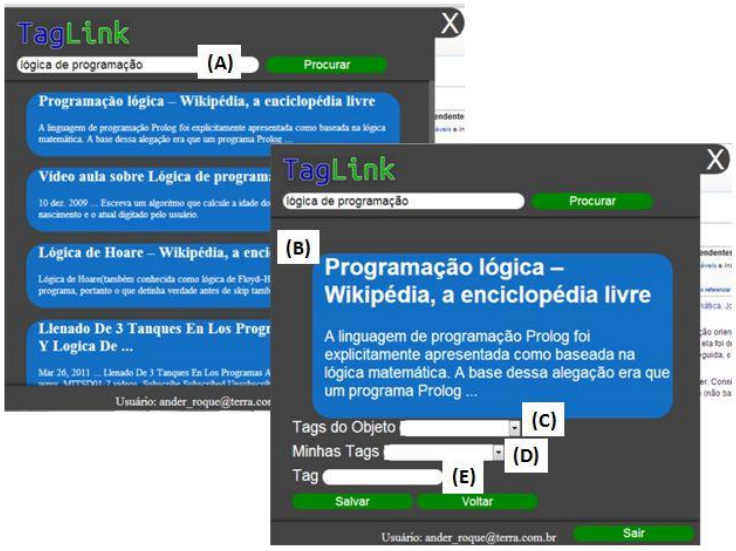

Figure 2. TagLink tool – search and tagging view.

The students can search the learning objects that are stored in the repository through the tag cloud (Figure 3). The words represented by a larger font, are the tags that were attributed to more than one object. How much larger the font of the tag is the meaning is that more objects were assigned to it. When a tag is selected (A) in the cloud TagLink displays the list of the objects that were marked by this tag (B). The student can then click on it to access it.

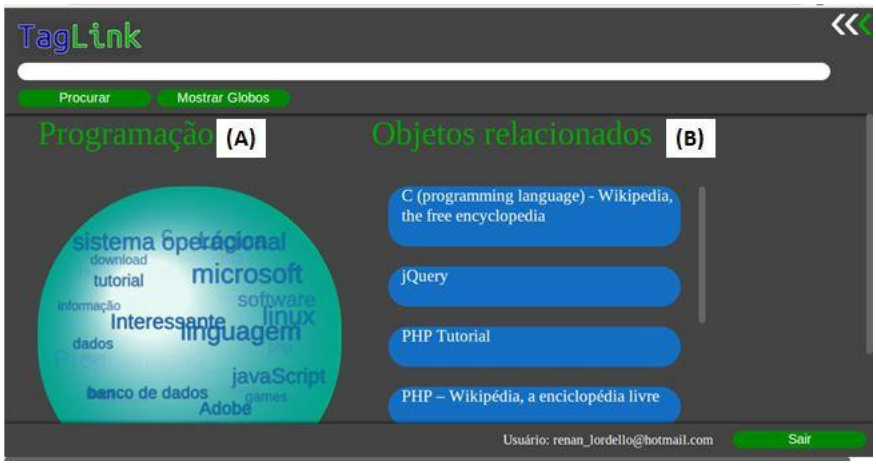

Figure 3. TagLink tool – tag cloud view.

In TagLink, it is possible to register users and to organize them in groups - classes, or workgroups, for instance - so that it becomes possible to observe the behavior of specific sets of students. It also supports the retrieval of data about the tags: creation date, how many times it was used, who created it, corresponding objects, and so on; and the retrieval of data about the learning objects: which tags, times of use, origin, and so on.

4 e-Learning Social Tagging Approach in Action

We carried out an experiment with 336 students from the vocational education level (information technology technicians) at a country side school in São Paulo, Brazil. The students were instructed to search and retrieve learning objects related to their current courses and to tag these objects through TagLink tool. The activity of the students was recorded for analysis.

A preliminary poll revealed that most of the students were regular users of social networks; and that they were familiar with tagging, but they had never used such functionality for educational purposes. Based on this, we decided to provide the students with an initial training before using the system.

4.1 Planning the Experiment

We chose to split the student in two groups: Group A (GA) consisting of the older students, and Group B (GB) consisting of the younger students. We also defined two sets of search terms - both with 5 terms each - according to the students course. These sets answer for the initial seed necessary at earlier stages of tagging systems as suggests Bateman, Brooks and Mccalla [2] and we reported in Section 2:

- Ta: contained generic terms for information technology, like logic, C#, Databases, Windows, and Linux; and

- Tb: contained web programing terms, like JQuery, PHP, XNA, Android and Flash.

The experiment should have 2 phases. In the Phase1, the students would be told to search the object from the TagLink tool using all the terms of the set that was assigned to him/her. They should also select at least one learning object from each search result, and to create three or more tags for this object. In Phase 2, the students would be allowed to use existing tags, either of their own (my tags) or of other students (tags of the object), or to create more tags. The goal will be to have the students build a vocabulary of tags. Prior to the experiment we had an introductory period when the students learned about how to use TagLink and about the importance of the experiment.

4.2 Experiment in Action

We conducted the two phases of the experiment during 2 months, switching the sets of search terms to the groups: first, GA has used set Ta and GB has used Tb; and after GA has used set Tb and GB has used Ta. Our intention were to observe if one group would use the tag of other group, showing us the evolution and the estabilization of the vocabulary. At the end we had 2,019 distinct tags for 218 learning objects selected by 336 students.

4.3 Analysing the Outcomes

Aiming to verify the potencial of use of social tagging to catalog learning objects, we analyze two aspects: the number of times and of students that created new tags; and the number of times and of students that reused existing tags. The two aspects can show us the involvement of the students with the process of tagging learning objects. We verified that 1.279 out of 2.019 tags (63.34%) were used just once, when they were created; the other 740 tags (36.65%) were reused. Table 1 summarizes how the students used the system in relation to the creation and to the reuse of tags. It is important to highlight that the summarization does not consider only tags used just once, reflecting the students activities in the system.

Group	Phase1		Phase2		Total	
	Created	Reused	Created	Reused	Created	Reused
A	1,126	648	144	303	1,270	951
B	1,329	933	139	363	1,468	1,296
Total	2,455	1,581	283	666	2,738	2,247

Table 1. The relevant terms according to the responses of students.

In Figure 4, we can verify one first concentration in the data defining a nearly Gaussian-peaked distribution around 2 tags per student; and one main second concentration defining a nearly Gaussian-smooth distribution center around 12 tags per student. The first concentration is expected; since the participation in the experiment was optional, a significant fraction of the students created no more than 3 tags. The second main concentration revealed that 4 times

as much students took part of the experiment creating from 4 up to 18 tags each. The participants of the second main concentration created 2,420 tags - not necessarily distinct - or nearly 11 tag creations per learning object.

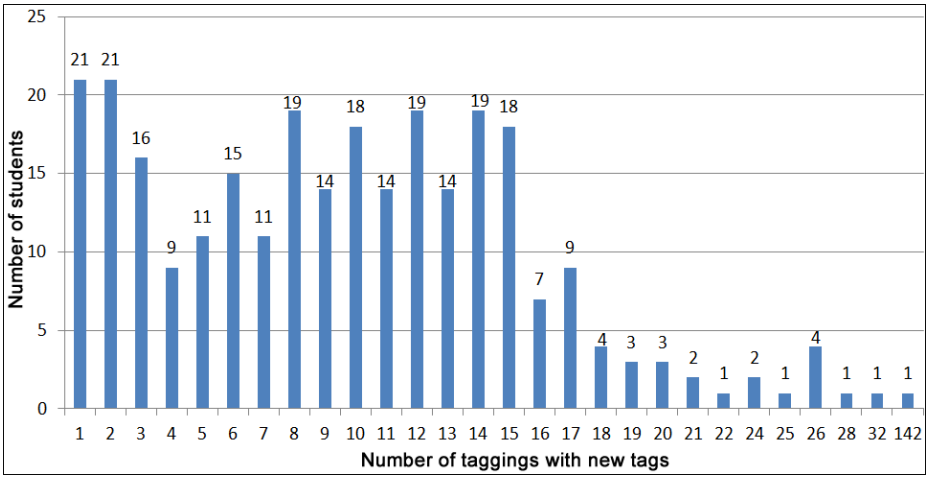

Figure 4. Distribution of the number of students per number of tags created.

In Figure 5, we can verify two peaks, one around 5 reused tags per student, and another one around 15 reused tags per student. The distribution now is shifted if compared to the distribution of new tags per student - Figure 5. In the place where there was a peak, now there is a valley; the events indicate that there was an increase in the participation of the students who did not create tags before - more to the left in the distribution. Meanwhile, a smaller set of students engaged even more in the experiment and increased the expected value - more to the right in the distribution. This is a curious observation, it shows three behaviors for the students, as derived from Table 1: a set with students that only created new tags (11.3%), a set with students that only reused tags (8.9%), and a set of students that did both things (79.7%).

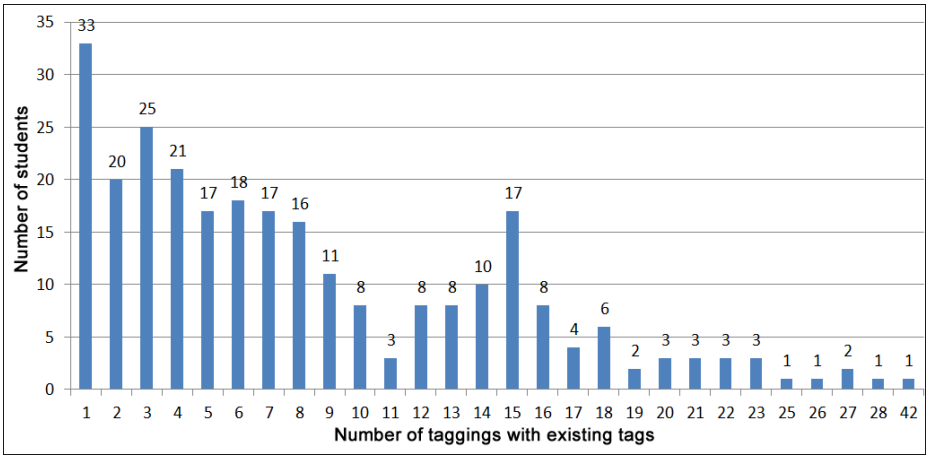

Figure 5. Distribution of the number of students per number of tags reused.

Observing the aspect of convergence to a well-defined descriptive vocabulary of tags, we had to analyze the number of new tags created in the system along the time of the experiment (two months). In Figure 6, we can see that the number of new tags behaves according to a Normal distribution with a peak close to the middle of the period - in the 28th day. The Normal distribution suggests that after the peak, the students create just a few more new tags, a number tending to zero as we get far from the expected value.

The Normal distribution, while valid, is not as strong as a descending power-law distribution would be; nevertheless, this fact is also interesting. Why did not the process behave like a power-law in which most of the new tags were created at the beginning of the period? The answer is quite straight when we consider that the experiment goes over a human-computer interface. In such systems, the user goes through a learning curve with three phases: slow beginning, steep acceleration, and plateau. We speculate, hence, that the left side of the Normal distribution - until nearly the 20th day - was affected by the learning period and that, only after, the users were able to fully work on TagLink and demonstrate their tagging profiles.

While Figure 7 shows that the tags are to stabilize after the 28th day, Figure 6 shows that a small subset of tags dominates the usage in the system. More precisely, there were 2,019 different tags; among them, 1,972 of them were used less that 17 times - not shown in the figure; and only 47 were used 17 times or more - shown in the figure. These characteristics describe a long heavy-tailed distribution indicating a strong imbalance as the students concentrate on a very small subset. Table 2 complements our view about the tags used by the groups (GA and GB) in each phase.

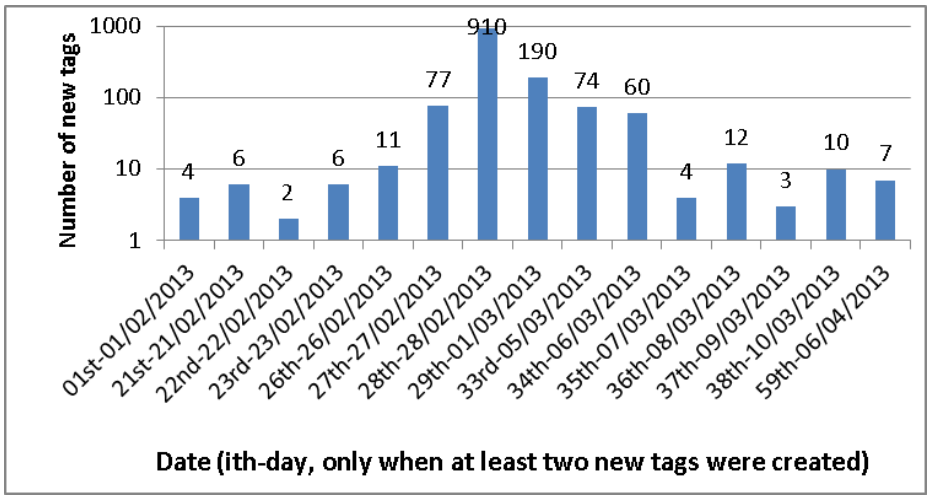

Figure 6. Distribution of the number of new tags created along time.

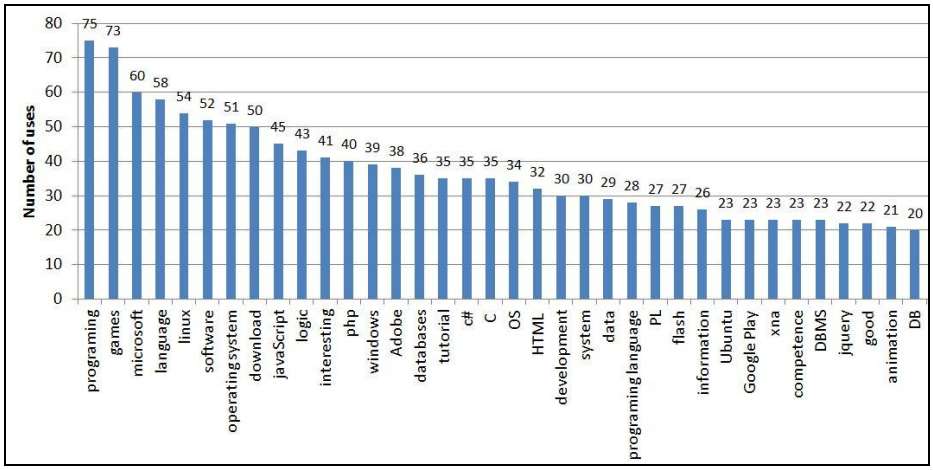

Figure 7. Distribution of the number of times each tag was used.

By comparing this dominant set of terms with the set of seed terms provided in the beginning of the experiment (Ta and Tb) it is possible to observe a great intersection. This fact indicates that the seed of terms strongly influenced the vocabulary and that, possibly, this seed accelerated the process as suggested in other works. Therefore, based on the evidences of Figures 6 and 7, we argue that social tagging is supposed to converge to a well-defined set of tags. We also argue that an appropriate set of seed terms may provide some control over this process, influencing the definition of the set of most frequent tags and, consequently, influencing how descriptive they will be.

Tag x Group x Phase					
	Phase 1		Phase 2		
tag	Grup A	Grup B	Grup A	Grup B	Total
programming	30	7	6	11	54
Microsoft	20	17	6	8	51
Linux	23	13	2	2	40
logic	27	0	1	5	33
software	13	10	7	10	40
language	16	19	3	9	47
operating system	14	10	7	10	41
JavaScript	0	28	9	0	37
SO	10	3	2	9	24
interesting	1	19	3	0	23
Total	154	126	46	64	390

Table 2. Tags used by groups in each phase.

Aiming to enrich our conclusions, after the experiment we asked to other students' group (200 students of 3 different course) to elaborate a study evolving the terms described on Ta and Tb (see Section 4.1). However, instead of inform exactly the terms we suggested some topics as "visual effects in web pages", "formatting of web pages", and so on. In this way, the

participants defined and chose the keywords they would use in searches.The students used the Taglink's tag globe. During the navigation the students reported they found or not the learning object according to the terms they searched. We had 98% of positive feedback.

Therefore we suggest, with significant evidence, that social tagging can successfully be used in e-learning. By considering the experiment, we can affirm that, for our specific setting, the students satisfactorily participated in the tagging process by defining enough tags for the description and latter retrieval of objects.

5 Conclusions and Further Works

The folksonomy through the social tagging allows the creation of a vocabulary of learning objects collaboratively. This possibility brings the student to a more active role in the process of organization and maintenance of learning objects.

Considering this opportunity, this paper presented an approach based on the triplet of learners, learning objects and tags for providing social tagging is the e-learning. The proposal could be achieved by the construction of the TagLink tool which address the proposed approach and supported us in our evalutation phase.

In a real setting with 336 students we found that: (1) the use of social tagging is viable in the sense that students are inclined to build extensive catalogs over the learning objects; (2) despite their colloquial experience with tagging content in social networks, students will tag learning objects using descriptive (formal and general) terms that aid the posterior use of objects catalogs; and (3) the vocabulary of terms converges to a "rich" subset of terms that answers for over 95% of the tags crated and/or reused by the students, indicating that guidance (an initial set of terms) can lead to a faster convergence and to an improved control over the process.

In any case, our experience brings insights that could be used as first assumptions in motivating the construction of social tagging learning systems. Looking at the lessons learned in the experiment, we observed that the reuse of tags can also be seen as a suggestion process of tags. This encourages us to study the possibility to include a tag recommendation cycle on our approach, improving the tagging process.

Acknowledgments

We thank the the students of Fernando Prestes school for participanting of our experiment.

References

1. Ahn, L., Dabbish, L. (2004) Labeling images with a computer game, In: CHI '04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 319–326.
2. Bateman, S., Brooks, C., Mccalla, G. (2007) Applying collaborative tagging to e-learning, In: Proceedings of ACM WWW.
3. Dahl, D., Vossen, G. (2008) Evolution of learning folksonomies: social tagging in e-learning repositories,". International Journal of Technology Enhanced Learning, V. 1, p. 35–46.
4. Garcia-Penalvo, F. J., Morales, E., Barron, A. (2007) Learning objects for e-activities in social web, WSEAS Transactions on Systems, V. 6, N. 3, p. 507–513.
5. Glavinic, V., Granic, A. (2008) HCI research for e-learning: Adaptability and adaptivity to support better user interaction, In: HCI and Usability for Education and Work. LNCS 5298, p. 359–376.

6. Google's Custom Search API Available at: https://developers.google.com/custom-search/?hl=en. Accessed in: October, 2014.
7. Gupta, M., Li, R., Yin, Z., Han, J. (2010) Survey on social tagging techniques, In: Newsletter ACM SIGKDD Explorations, V. 12, N. 1, p. 58–72.
8. IEEE LOM, Draft standard for learning object metadata, Available at: http://ltsc.ieee.org/wg12, Accessed in: December, 2014.
9. Jaschke, R., Hotho, A., Mitzlaff, F., Stumme, G. (2012) Challenges in Tag Recommendations for Collaborative Tagging Systems, Recommender systems for the social web, pp. 65–87.
10. McGreal, R. (2004) Learning objects: A practical definition, International Journal of Instructional Technology and Distance Learning, V. 1, N. 9.
11. Milicevic, A. K., Nanopoulos, A., Ivanovic, M. (2010) Social tagging in recommender systems: A survey of the state-of-the-art and possible extensions, Artificial Intelligence Review, V. 33, N. 3, p. 187–209.
12. Monge, S., Ovelar, R., Azpeitia, I. (2008) Repository 2.0: Social dynamics to support community building in learning object repositories, In: Interdisciplinary Journal of E-Learning and Learning Objects, V. 4, I. S. Institute, Ed. Informing Science Institute, p. 191.
13. O'Reilly, T. (2005). What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, Available at: http://oreilly.com/web2/archive/what-is-web-20.html.
14. Shepitsen, A., Gemmell, J., Mobasher, B., Burke, R. (2008) Personalized recommendation in social tagging systems using hierarchical clustering, In: Proceedings of the 2008 ACM conference on Recommender Systems.
15. Shih, W., Tseng, S. (2012) Folksonomy-based indexing for retrieving tutoring resources, In: 2012 Seventh IEEE International Conference on Wireless, Mobile and Ubiquitous Technology in Education, p. 97–101.
16. Sierra, J.L., Valmayor, A. F. (2008) Tagging learning objects with evolving metadata schemas, Eighth IEEE International Conference on Advanced Learning Technologies, ICALT 2008, Santander, Spain.
17. Sinclair, J., Cardew-Hall, M. (2008) The folksonomy tag cloud: When is it useful?, Journal of Information Science February, V. 34, N. 1, p. 15-29.
18. TagLink tool. Available at: http://200.133.238.112/tag/.
19. Trant, J. (2009) Studying social tagging and folksonomy: A review and framework, Journal of Digital Information, V.10, N.1.
20. Vazquez, A. R., Ostrovskaya, Y. A. (2006) Analysis of open technological standards for learning objects", In: Proceedings of 4th LA-Web, Puebla Cholula, Mexico, p. 105–108.
21. Wal, T. V. (2005) Explaining and showing broad and narrow folksonomies,". Off the Top [blog], Available in: http://www.vanderwal.net/random/entrysel.php?blog=1635, Accessed: 13 December 2013.
22. Zaina, L. A. M., Almeida, T. A., Torres, G. M. (2014) Can the online social networks be used as a learning tool? A case study in Twitter. Communications in Computer and Information Science (Print), V. 446, p. 114-123.
23. Zervas, P, Sampson, D. G. (2014) The effect of users' tagging motivation on the enlargement of digital educational resources metadata, Computers in Human Behavior, V. 32, p. 292–300.