Learner-Generated Annotation of Learning Resources – Lessons from Experiments on Tagging[1]

Steffen Lohmann
(University of Duisburg-Essen, Germany
steffen.lohmann@uni-duisburg-essen.de)

Stefan Thalmann
(University of Innsbruck, Austria
stefan.thalmann@uibk.ac.at)

Andreas Harrer
(Catholic University of Eichstätt-Ingolstadt, Germany
andreas.harrer@ku-eichstaett.de)

Ronald Maier
(University of Innsbruck, Austria
ronald.maier@uibk.ac.at)

Abstract: The importance of comprehensive annotations for learning resources is widely recognized. However, it is often unclear how these annotations should be created. A promising solution might be that learners annotate learning resources during execution of learning processes. This paper examines tagging as a learner-driven approach that can be used to get annotations on a collaborative level. The characteristics of learner-generated annotations in learning scenarios and the applicability of these annotations to describe learning resources are investigated. As empirical grounding, the results of three independent experiments are presented in order to inform decisions concerning the establishment of institutionalized settings deploying tagging approaches.

Keywords: Annotation, Learning Resources, Tagging, Metadata, Collaboration
Categories: L6.2, L6.1, L3.6, L3.5, L1.2

1 Introduction

Adding metadata to learning resources is a prerequisite for sophisticated administration, exchange, and reusability of these resources. Even though benefits and requirements of metadata are well-described and formulated in standards such as IEEE LOM sufficient annotations are still not provided in many cases [Brooks and McCalla 2006]. Often, no organizational routines or responsibilities are defined. Authors of learning resources are not willing to spend additional effort to add metadata and professional metadata authors are too expensive. With the advent of Web technologies allowing large numbers of users to participate in content

[1] Lohmann, S.; Thalmann, S.; Harrer, A.; Maier, R.: Learner-Generated Annotation of Learning Resources – Lessons from Experiments on Tagging. In: Tochtermann, K.; Maurer, H. (Hrsg.): Proceedings of I-KNOW ´08. 8th International Conference on Knowledge Management, Graz, Austria, September 3-5, Journal of Universal Computer Science, 2007, 304-312.

production, collaborative ways to annotate Web resources, such as tagging, become popular. The application of learner-generated annotation could be a promising solution to get suitable descriptions of learning resources.

In this paper we examine the applicability of tagging as a form of learner-generated annotation of learning resources. We first give a brief overview on different ways of annotating learning resources in Section 2. In Section 3, we report on studies we conducted to investigate this phenomenon. Subsequently, we draw conclusions in Section 4 based on our insights gained in the studies. The paper ends with an outlook on future research and development tasks with respect to this topic.

2 Annotation of Learning Resources

Annotation of resources has a long history in librarianship and was taken over to digital resources. Traditionally, a small group of experts categorizes or indexes resources on the basis of an agreed, structured catalogue of keywords, a taxonomy, in order to make resources accessible [McGregor and McCulloch 2006]. With the rapidly increasing amount of digital resources, time and cost required for professional metadata creation are unsustainable for many organizations. As a consequence, resources are frequently not or only partially annotated and descriptions often do not fit the users' needs.

Automatic processes can resolve this problem in part by reducing the number of metadata elements that have to be manually edited [Duval and Hodgins 2004], thus relieving the metadata author from part of the tedious work of filling in metadata forms [Cardinaels et al. 2005]. The basic idea of automatic annotation of resources is to use (1) different sources available to the authoring system, (2) the resource itself or (3) the context in which the resource is used [Motelet et al. 2006]. Automatic metadata generation offers good results for technical descriptions such as size or format but is no feasible approach for more qualitative descriptions, like summarizing keywords or level of difficulty [Bauer et. al. 2008]. Particularly, if annotations for multimedia content and different usage contexts are required, the exclusive use of automatic approaches is not sufficient, in general, but can be used to support the manual annotation process.

With the advent of new Web concepts under the "Web 2.0" umbrella [O'Reilly 2005], tagging seems to be a promising approach that can be used for annotating learning resources. Thereby, a tag is a freely chosen text-based annotation to a resource that is not restricted to a controlled vocabulary. Collaboratively added tags can be represented in different kinds of structures that have come to be called "folksonomies" [Rollett et al. 2007]. In contrast to the top-down approach of expert annotation, this bottom-up approach reflects the learner perspective on the resources and can thus lead to a broad spectrum of descriptions for learning resources.

A major concern that has been voiced against this solution is that untrained people cannot achieve the metadata quality professionally trained staff reaches. However, Surowiecki argues for the superiority of the "wisdom of the crowds" which is conceptualized as "collective intelligence", i.e. the ideas of many people bring in a wider variety compared to one single person [Surowiecki 2004]. Similarly, the theory of the "expert blind spot" [Nathan et al. 2001] poses that the expert's judgement can be oriented towards a focused perspective that is not well connected to the

perspectives of the learners, particularly to unconventional perceptions. Specifically for metadata a broad spectrum of descriptions is valuable because the community of users of learning resources represents the potential seekers as well as several application contexts exist in which learning resources are used. Figure 1 shows the different approaches for generation of annotations schematically.

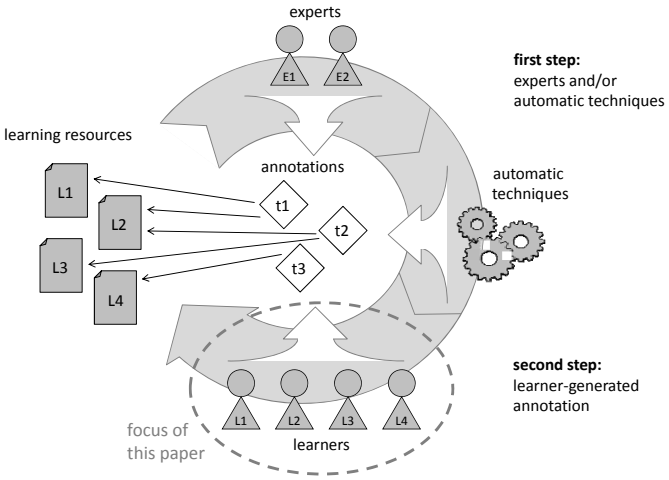

Figure 1: Annotation life-cycle

3 Experiments on Tagging of Learning Resources

We conducted studies to investigate the applicability of tagging as an approach to annotate learning resources. In the following, these studies are briefly described.

3.1 Tagging of Lecture Slides in a Pen-and-Paper Study

We investigated the reliability and usefulness of students' tags compared to expert tags for lecture materials in a small-scale experiment. To get a controlled scenario – without technical influences on the question we investigate in – we chose a pen-and-paper design in a classroom setting, where the tagging procedure and the intended navigational feature of shared tags was emulated by circulating learning materials randomly among the students.

The study was conducted as a voluntary experiment in an advanced course on “computer-supported learning and work in groups” with 8 students and 14 exemplary slides taken from selected lectures of the course. Each student was assigned to a specific colored pen so that we can take a look at individual tagging behavior without violating data security. The slides were directly annotated resulting in 286 student tags of which 123 were different. In our analysis, we compared the collection of student tags with the tags given independently by an expert in the field, the lecturer of the course (and author of the study). The expert tag set was comprised of 74 tags of

which 70 were different (the other 4 were more abstract concepts characterizing multiple slides). Based on this material, we elaborated on three research questions:

1. *"Expert's Recall": Which slides can be found by a simulated query with the expert tags on the complete set of learner tags and which relevance (number of matching tags) does it have?* With respect to this question, we found that in the 14 simulated queries (the expert tag set for each slide against the whole tag set of learners) all relevant slides could be found; the ratio of matches was in average 66% of the expert tags also assigned to a slide by the students.
2. *"Learner's Recall": Which slides can be found by a simulated learner query on the set of expert tags and which relevance does it have?* The result was that 95 of 100 simulated queries with the student tags got the relevant slide as a match. Interestingly, the only missed queries had all the same slide as focus, where the expert had given just 2 tags, which resulted in only 2 out of 7 of the students that tagged this slide finding a match.
3. *Abstraction / keyword ratio: How many keywords assigned as tags are already present as text in the slides?* This question addresses if the tags given to the learning resources stay close to the original resource or if abstraction and aggregation processes are performed by the tagger. The ratio between tags present as keywords in the slides (excluding the problematic slide mentioned above) and all tags given is 86% for the expert (62/72) and 94% (234/250) for the students. While the results do not allow reliable claims, a hypothesis for further research could be to check if experts tend to tag more abstractly and conceptual, thus having a smaller keyword ratio.

Combining the results on questions 1 and 2, we consider the correspondence of the student-generated tags and the expert tags as a good fit, in contrast to the results given in a similar study in [Bateman et al. 2007]. Yet, there are several differences in the studies, such as a synchronous vs. asynchronous setting, voluntary vs. obligation to participate. Question 3 goes along a cognitive perspective to explore the conceptual understanding and ability for abstraction while tagging, i.e. how strong the tendency is to cling to the exact words contained in the resource when choosing tag names.

3.2 Tagging of Learning Resources via an Online Questionnaire

Next to the pen-and-paper study, a multi-round lab experiment was realized within a course on introduction to information systems. Altogether, 174 Bachelor students of business and economics participated. The experiment was realized in two experimental series with seven course groups of between 20 and 28 students each. Every student was asked to tag 10 digital learning resources (2 videos, 3 presentation slides, 3 screenshots and 2 pictures). Overall, 4246 tags (404 different tags) were assigned to the 10 resources. As common in many available tagging tools, the most frequently used tags were presented and could be chosen easily. Also, additional tags could be entered. Technically, the study was realized by a questionnaire tool. Typical linguistic errors such as wrongly written tags and confusion by plural/singular forms were cleaned by the researchers. A more detailed description of the study can be found in [Bauer et. al. 2008].

The aim of the experiment was to investigate the applicability of tagging to gain suitable descriptions of learning resources. In order to achieve this, three premises were formulated and tested in the experiment:

4. *Commitment: Which tags are generally accepted and which tags are specialized and individual?* To apply tagging in organizations and to use learner-generated descriptions a set of meaningful tags must be extracted. Therefore, a reliable separation of generally accepted tags from specialized tags has to be achieved. The results of the experiment show that this separation seems to be realizable. In the experiment a 5% criteria could be applied to extract the meaningful tags suitable for describing resources on an institutional level. The statistical distribution shows a power law curve.
5. *Convergence: Is there a stable set of frequently used tags and which set of tags is it?* In order to apply criteria for commitment a stable set of tags and thus convergence is required. Convergence of tags could not be shown finally but several findings indicate a tendency towards convergence. The number of changes in the top 5 ranking of tags was monotonically decreasing. Additionally, the ratio of distinct tags to number of all tags decreased but the total number of distinct tags increased over all experimental turns. To show a total convergence a very large number of taggers is required. In organizational settings, this large number of taggers is normally not available and thus indicators for sufficient convergence should be identified in further research.
6. *Coordination: Is there an effect of initial tags on the semantic orientation that user-generated tags evolve?* The use of tagging in organizational settings should consider respective conventions and management directives. Therefore, coordination of the 'semantic direction' of descriptions by varying the starting set would be very useful. In order to investigate this aspect we used two different directions of semantic description (context- and content-related descriptions) for the two experimental settings. In the series with start tags describing the content of learning resources, further tags describing the content were entered by the learners. Likewise, in the series with start tags describing the context, further tags describing the context were entered. Overall, the textual input differs between 71% and 93% with 81% on average for the same resources over both experimental series. This means that on average 81% of the tags were different between the series. The 19% tags that were identical on average were mainly words that occur in the resource itself, such as headlines. These observations support the proposition on coordination, i.e. the formation of the starting set of tags can indeed influence the semantic direction of the metadata descriptions.

Altogether, in this lab experiment the extraction of meaningful descriptions by applying a 5% criteria and the coordination of tagging behavior by changing the initial set of tags could be shown. Indicators suggest a trend towards convergence in larger user groups.

3.3 Using Tagging Systems in University Courses

In addition to the two controlled experimental settings, we applied tagging systems in field studies over a longer period to support blended learning in university courses.

The general aim of this research was to discover commonalities and differences between application of tagging in open Web and focused learning contexts.

In one scenario we used the social bookmarking service *del.icio.us* in a student project with 22 participants. The students were asked to register for this service and use it to bookmark Web pages that they consider relevant to the project context. In each case, the students should use a predefined 'project tag' in addition to their freely chosen tags. That way, it was possible to view the entire collection of annotated Web pages simply by inputting this 'project tag'.

In another scenario, a Web-based application prototype has been implemented that enables the students to tag lecture slides during the teaching process or afterwards. The students attending the lecture were enabled to organize the slides from their personal perspective or to use the group view to explore the slides via the aggregated tags of the learning community. A more detailed description of the developed application prototype can be found in [Fienhold et al. 2007].

In both cases we provide an initial set of tags as starting point and illustration. With respect to the findings of research question 6 (see Section 3.2), we thereby guide the annotation in a certain 'semantic direction', in this case, towards content-related descriptions. Two research questions were of particular interest in these field studies:

7. *Participation: How many learners use the tagging system actively and how many passively?* Since participation was not mandatory or extrinsically motivated in both settings, similar to open Web contexts only a small number of the students actively used the tagging feature whereas the larger part can be regarded as what is called "lurkers" [Nielsen 2006] – they did not actively contribute but read the contributions of others. In the first case, 5 of 22 participants bookmarked 19 Web pages in addition to the Web pages that have already been bookmarked in advance. In the second case, 20 students used the tagging system to access the lecture slides, but only 4 students actively participated in the tagging process.
8. *Acceptance: Do learners appreciate the possibility to tag resources and to access learning contents via tags?* All students that used the tagging systems give positive feedback on this additional way to annotate or access learning resources. Most students find the feature to explore learning contents by navigating tags particularly useful. In the case of *del.icio.us*, some students even use tags to discover further web pages that have not been bookmarked by the project members but by other *del.icio.us* users. The initially given tag set was regarded as a helpful starting point and illustration. However, in the second scenario, most students prefer to annotate the lecture slides with pen and paper as usual.

Overall, in these focused field studies the basic user behavior and observed effects were similar to what is known from open web contexts. The additional form of accessing learning resources via learner-generated annotations was generally appreciated, but active participation in tag generation was limited though.

4 Conclusions

Generalizing from our studies with different research foci in a multi-case approach [Stake 2005] we gained several indicators that learner-generated annotation of

learning resources can be a valuable contribution for all parties involved in the learning process, i.e. the learner her-/himself, peer learners, lecturers, and institutions. Though our results are first indicators from university settings, similar effects can be expected in other organizational contexts. In the following, the results are systematized as theses for the design of institutionalized settings deploying tagging in businesses and organizations:

Use stable set of tags for agreed description of resources: Based on our results on establishment of generalized tags (question 4), tag convergence (question 5), and the matching of learner and expert tags (questions 1 & 2) we observe a tendency towards a stable set of tags that helps to index and easily find the relevant learning resources. Furthermore, this stable set could be used to gain valuable metadata.

Guide the tagging process: Besides the emergence of tags from a learning community, the addition of expert tags can contribute to the activation of the tagging process. By giving an initial description of learning resources, guidance in a certain direction can be realized (questions 6 & 8), which could be useful for maintaining organizational guidelines and principles.

Use text extracted from resources for starting set of tags: Furthermore, learners tend to use prominent words occurring in the learning resource, like headlines or bold text, as tags in many cases (questions 3 & 6). As a consequence, automatic approaches could be used to extract these words and use them as a starting point for the tagging process.

Use small set of selectable tags to support convergence: In addition to these initial sets, selectable tags could facilitate the convergence and thus the commitment within a group. This orientation can be specifically given by an expert, as the results on the coordination aspect (question 6) show, which could be used by a teacher productively to steer the tagging process and choice of tag words along the line of orientation that is best in tune with the lecture.

Motivating learners to participate, the need for electronically provided learning material, and integration into the existing ICT infrastructure are challenges in this topic area. A further challenge of learner-generated annotation is the high probability of noise in the tags [McGregor and McCulloch 2006]. For using tagging over a longer period of time, institutional life-cycles and procedures for cleaning and consolidation of tag sets are required.

5 Further Research

Further research includes deeper investigation and better understanding of the interplay and mutual enhancement of automatic, expert, and learner annotation. As already illustrated in Figure 1, an evolutionary process that combines these three types of annotation appears to be promising. First steps in this direction can be found in [Dahl and Vossen 2008].

Furthermore, organizational routines and processes for tagging should be investigated in order to enable an efficient management of these descriptions, including life-cycles that ensure quality. Since seamless integration of annotation into existing ICT infrastructures is crucial for the success of learner-generated annotation, further research and experiences regarding design and success factors are desirable.

Another promising research topic is the processing and utilization of the learner-generated annotations. On the one hand, sophisticated visualizations (apart from tag clouds) can provide multidimensional access and exploration possibilities on the learning resources by means of graphical navigation through a space representing tags and resources. On the other hand, recommendation features could allow for learner-centered selection and presentation of learning resources. In both cases, all parts of the tripartite model encompassing the learning resources, the learners, and the annotations can be used as a source for relationship discovery.

Acknowledgements

We are grateful to all our students participating in the studies.

References

[Bauer et al. 2008] Bauer, M., Maier, R., Thalmann, S.: "Metadata Generation for Learning Objects – An Experimental Comparison of Automatic and Collaborative Solutions"; MKWI 2008, München, Germany, Springer (2008).

[Bateman et al. 2007] Bateman, S., Brooks, C.: McCalla, G., Brusilovsky, P.: "Applying Collaborative Tagging to E-Learning"; Proceedings of WWW 2007 Workshop on Tagging and Metadata for Social Information Organization (2007).

[Brooks and McCalla 2006] Brooks, C.; McCalla, G.: "Towards Flexible Learning Object Metadata"; Int. J. of Continuing Engineering and Lifelong Learning, 16, 1/2 (2006), 50-63.

[Cardinaels et al. 2005] Cardinaels, K., Meire, M., Duval, E.: "Automatic Metadata Generation: The Simple Indexing Interface"; Proceedings of the International World Wide Web Conference (WWW05), Chiba, Japan, ACM (2005).

[Dahl and Vossen 2008] Dahl, D., Vossen, G.: "Learning Object Metadata Generation in the Web 2.0 Era"; Int. J. Information Communication and Technology Education, 4, 3 (2008).

[Duval and Hodgins 2004] Duval, E., Hodgins, W.: "Metadata Matters"; International Conference on Dublin Core and Metadata Applications, Shanghai (2004).

[Fienhold et al. 2007] Fienhold, J.; Riechert, T.; Lohmann, S.: "Der Lecture Tagger: Social Tagging in E-Learning-Umgebungen"; Proceedings of the Workshop on e-Learning 2007, HTWK, Leipzig (2007), 251-261.

[McGregor and McCulloch 2006] McGregor, G., McCulloch, E.: "Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool"; Library Review, 55, 5 (2006), 291-300.

[Motelet et al. 2006] Motelet, O., Baloian, N., Pino, J. A.: "Learning Object Metadata and Automatic Processes: Issues and Perspectives", Informing Science Press (2006).

[Nathan et al. 2001] Nathan, M., Koedinger, K.R., Alibali, M.: "Expert Blind Spot: When Content Knowledge Eclipses Pedagogical Content Knowledge"; Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Seattle (2001).

[Nielsen 2006] Nielsen, J.: "Participation Inequality: Encouraging More Users to Contribute"; http://www.useit.com/alertbox/participation_inequality.html (last access April 14th, 2008).

[O‘Reilly 2005] O‘Reilly, T.: “What is Web 2.0. Design Patterns and Business Models for the Next Generation of Software”; http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html (last access April 14th, 2008).

[Rollett 2007] Rollett, H., Lux, M., Strohmaier, M., Dösinger, G., Tochtermann, K.: “The Web 2.0 Way of Learning with Technologies”; Int. J. of Learning Technology 3, 1 (2007), 87-107

[Stake 2005] Stake, R.E.: “Multi Case Study Analysis”; New York: The Guildford Press (2005)

[Surowiecki 2004] Surowiecki, J.: “The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations”; B&T Press (2004)