Collaborative tagging, metadata creation and learning: A study within a higher-education course

Rachel Or-Bach
orbach@yvc.ac.il
Emek Yezreel College

Collaborative tagging and user created metadata are usually discussed either from a social perspective (social classification) or as an alternative to ontologies. In this paper we present a study regarding metadata creation by students and discuss the implication of such learning activities. The students' tasks involve creation of metadata for computer programming examples that they use and reuse in different ways throughout a course.

The iterative activities of refining the metadata schemes are done both individually and collaboratively in teams. Our findings show the potential of metadata creation for students' learning. Iterative processes of metadata creation and modification encouraged and supported reflection activities and brought students to revisit, rethink and reinterpret the various code examples and the various relations between them.

Introduction

Collaborative tagging refers to the process by which users add metadata in the form of keywords to shared content. This process gained popularity lately on web sites that allow users to tag bookmarks, photographs and other content. The use of keywords or tags is a common way for organizing content in a way that enables future search, filtering and navigation. Organization by categorizing or indexing is usually performed by an authorized person (a librarian for example) or is derived from the material provided by the authors of the content. In contrast, collaborative tagging is the practice of allowing anyone – especially consumers – to freely attach keywords or tags to content. This kind of collaborative tagging offers an interesting alternative to current efforts at semantic web ontologies (Shirky, 2005). Sites that feature collaborative tagging allow users to publicly tag shared content so that they can not only categorize information for themselves, but they can also browse the information as categorized by others. Golder and Huberman (2006) conclude from their study of user activity within collaborative tagging systems that the prevalence of tagging with a very large number of tags and according to information intrinsic to the tagger demonstrates that a significant amount of tagging is done for personal use rather than public benefit. Tagging or metadata creation involves cognitive processes of categorization, labeling and sense making, which makes metadata creation a preferred learning activity. Collaborative tagging adds the collaboration facet which provides opportunities for negotiation of meaning, reflection etc.

Metadata created by a student for any learning material captures her current understanding of the material. Metadata creation is expected to be a dynamic process that captures the growing understanding and fosters the dynamic nature of knowledge.

Comparisons between learners' metadata schemes can provide an anchor for meaningful discussions. The potential resulted negotiation of meaning can support the development of a community of practice. The notion of a "Community of Practice" emphasizes the view on knowledge within a social constructivist framework. Allert (2004) proposed the concept of community-oriented metadata, based on Wenger's statement of learning: "Meaning within communities" (Wenger, 1998). The idea that learning involves a deepening process of participation in a community of practice has gained significant ground in recent years and the collaborative refinement of a metadata scheme fosters this idea. Our study explores learning processes during individual and collaborative creation of metadata schemes for given computer programming examples that were used and reused by students during a course.

The Study - Task, Context and Measuring Tools

The course "Introduction to Expert Systems" is an elective course for students in our Computer Science and Information Systems department. During this course students are introduced to issues of knowledge acquisition, knowledge representation and expert systems design along with actual implementation. Implementation involves programming with CLIPS, which is an expert system shell (http://www.ghgcorp.com/clips/CLIPS.html). This rule-based system with a declarative programming paradigm is very different from programming paradigms that the students are familiar with. The course includes a hands-on laboratory, where we employ a lot of programming examples. Some of the examples are solutions to exercises that involve various modifications of a given program to achieve different goals. Program examples play an important role in teaching programming and the use of a library of annotated examples or templates is highly advocated by computer science educators (Linn and Clancy, 1992). All the learning materials for our course are available to the students through the web. We have a mechanism that enables the teacher to hide solution files from the students while they are working on the laboratory tasks and expose them when it is suitable. Students find themselves with a multitude of files of potentially useful code examples, which they have difficulties to retrieve when needed. This is an example of a case where students feel the need for a knowledge management tool and the creation of metadata seems not as an artificial requirement. Very soon students find that a simple tag or indication, such as in which lesson the specific code example was given, is not informative enough. A code example, and hence a useful metadata tag, might relate to syntax issues, to the use of various constructs in CLIPS, to different ways to solve a given problem, to ways to extend a solution etc. Examples can be linked by a CLIPS related issue, by a problem statement, by being a part of a bigger problem, by being one of several possible solutions, by relating to a specific pre-requisite etc. The student's challenge is to express explicitly the various roles and links among the example programs, actually to define categories. The required activities provide opportunities for exercising cognitive skills and activities that are important for developing programming skills: generalization and abstraction, use of "best practice" examples, appreciation of reuse; and linking syntax to semantics and to function. The measuring tools we employed for investigating the evolving learning processes were: homework assignments that involved metadata definition utilizing the CLIPS formalism (definition of a general deftemplate with respective slots); observations that were conducted during laboratory/class assignments; and written reflection reports that students were required to submit along with the homework assignments. The written reflection reports consisted of explanations about the choices, the alternatives and the respective considerations that were employed while defining the metadata scheme (the slots' names, the list of possible values or the range of values, default values for various slots etc.).

Findings and discussion

The submitted homework provided evidence about individual learning products and processes, while observations during team discussions provided evidence about both individual and collaborative learning processes. Findings show a shift from more objective and general metadata towards more subjective and specific metadata: from issues like date and lesson number towards issues like the program topic and the CLIPS constructs that are involved; and then towards more subjective characteristics such as expected contexts of use, comparisons with other programs, associations etc. Findings show also a shift towards more consideration of retrieval during tags' definition instead of just considering the storage. The activities of metadata creation brought students to rethink and reinterpret the various code examples and the various relations between them. With regard to the vocabulary we found terms that related to the following categories which are not mutually disjunctive: CLIPS syntax, CLIPS knowledge constructs, the respective “problem story”, the learning sequence, the origin (e.g. class example), expected reuse, “part of” relation, input/output mechanisms and value assignment procedures. We observed processes of abstraction, when students modified their scheme by replacing a slot (tag) or more that has "exists"/"not exists" values by a new slot with new respective values (Or-Bach, 2005). Observations showed that negotiation of meaning had a major impact on learning. It also showed the evolution of a community of practice; giving students the opportunity to experience processes similar to what they may encounter in future work situations.

Studies regarding collaborative tagging usually deal with social aspects and refer to social classification that is sometimes called folksonomy (Vander Wal, 2005). Our approach is different; it emphasizes the cognitive aspects of tagging or metadata creation and their potential role in individual and collaborative learning activities. The collaboration in our case has also a different emphasis; we are less interested in the widely-used tag but in the process of defining and redefining tags as a manifestation of a learning process. An important aspect of a folksonomy is that it is comprised of terms in a flat namespace: that is, there is no hierarchy, and no directly specified parent-child or sibling relationships between these terms. In an educational context, and especially in any computing related education, the idea of abstraction is fundamental and the construction of hierarchies is of great importance in a learning process. The dynamic construction and reconstruction of relevant hierarchies is an external representation of sense making processes. In our approach we used the formalism of CLIPS to enable these abstraction processes; to show the dynamic nature of understanding and the dynamic nature of knowledge; and also to reinforce the course topics (CLIPS programming). We have investigated metadata creation mainly as a mindtool (Jonassen, 1999) to support leaning, but the on-going process of metadata creation can also serve as input for relevant student modeling processes focusing on individual progress and not necessarily with reference to some pre-defined scheme or ontology. The resulted student model can in turn be used for adapted tutoring interventions.

References

Allert, H. (2004). Coherent social systems for learning: An approach for contextualized and community-centered metadata. *Journal of Interactive Media in Education*, 2.

Golder, S. and Huberman, B.A. (2006). Usage Patterns of Collaborative Tagging Systems. *Journal of Information Science*, 32(2). 198-208.

Jonassen, D. H. (1999) *Computers as Mindtools for Schools: Engaging Critical Thinking*, 2nd Edition.

Linn, M. C. and Clancy, M. J. (1992). The Case for Case Studies of Programming Problems. *Communications of the ACM*, 35(3), 121-132.

Or-Bach, R. (2005). Educational Benefits of Metadata Creation by Students. *Inroads, ACM SigCSE bulletin*, 37(4), 93-97.

Shirky, C. (2005). Ontology is Overrated: Categories, Links and Tags. http://www.shirky.com/writings/ontology_overrated.html (retrieved November 2006).

Vander Wal, T. (2005). Tagging for Fun and Finding. http://www.ok-cancel.com/archives/article/2005/07/tagging-for-fun-and-finding.html (retrieved November 2006).

Wenger, E. (1998). *Communities of Practice – Learning, Meaning and Identity*. Cambridge: Cambridge university press.