Exploiting the Power of Social Tagging Systems: A Semantic Flickr Approach for Tutoring and Knowledge Management

Martin Atzmueller and Alexander Hörnlein

University of Würzburg,
Department of Computer Science,
Am Hubland, 97074 Würzburg, Germany
{atzmueller, hoernlein}@informatik.uni-wuerzburg.de

Abstract. In general, social tagging systems are a versatile tool for the collaborative collection, annotation, organization, and distribution of resources and information. This paper proposes the design of a social tagging system for tutoring and general knowledge management. The system is specifically targeted at supporting the management of image resources. We describe a prototypical implementation of the system, and discuss its practical application in two real-world scenarios.

1 Introduction

With the advent of the so-called Web 3.0, i.e., the integration of Web 2.0 techniques with methods of the semantic web [1], social tagging systems [2] are prominent systems for the collection, annotation, organization and distribution of resources. Then, both of the Web 2.0 and the semantic web contribute to various degress to the specific applications.

In this paper, we propose the KNIZR (pronounced [naɪzər]) system (i.e., *KN*owledge organ*IZR*) – a social tagging system for tutoring and general knowledge management applications: We specifically target the management of image data that can be specified as resources, in addition to standard document types such as text, links (URIs) or PDF documents. Thus, in analogy to *Flickr* [3], we focus on the management of meta-data assigned to images that can be uploaded by the users and can be organized in various ways. Additionally, we also provide several powerful semantic browsing options for accessing the stored resources and content.

The proposed KNIZR system is based in the *gnizr* [4] framework and extends it by more diverse resource types, most importantly image data, and by including sophisticated knowledge-based techniques. For embedding semantics into the set of tags used for taggging the resources we utilize SKOS [5] (Simple Knowledge Organization System) for organizing the set of applicable tags, and for a better tag recommendation. We demonstrate the applicability and benefit of the presented approach in two real-world application scenarios.

The rest of the paper is organized as follows: Section 2 describes the general approach: First, we discuss the general context of social tagging systems, before we describe the KNIZR framework in more detail. After that, we describe two application scenarios in Section 3. Finally, we conclude the paper in Section 4 with a summary and point out interesting issues for future research.

2 A Semantic Flickr Approach

For organizing information and knowledge, the notion of *tagging* has recently received much attention: It has become a useful way for the collection, annotation, organization, and distribution of resources by users. The tags assigned to specific resources are used for navigation, locating the resources and for serendipitous browsing. Due to its ease of use, tagging system provide an immediate benefit for users: The allow a transparent access to various resources. Additionally, tagging can help for communicating interesting nuggets of information [6].

In this paper we focus on resources given by images as well as common resources, for example, provided by textual or PDF documents. Similar to the *Flickr* [3] system users can upload images as well as link other resources to the system. In the following, we shortly introduce the main features of social tagging systems, before we describe the KNIZR system in a nutshell.

2.1 Social Tagging Systems

Social tagging systems are based on resources, e.g., bookmarks or images, users and tags. Thus, an entry e in such a system can be regarded as a 3-tuple composed of a resource $r \in R$, a user $u \in U$ and a set of tags $t_1, \ldots t_n \in T$, where R specifies the set of resources, U specifies the set of users and T specifies the set of valid tags.

In general, the set of tags T is unbounded and can be extended by the users as needed. This is the case for folksonomies [7], for which arbitrary tags can be assigned. However, there are also other possibilities, especially for *closed communities*: In this case, the set of users U is fixed, and new users cannot join the community on their own. In such circumstances, often a restricted vocabulary is more appropriate. Then, the set T is pre-specified by a *super-user* and cannot be freely extended. Usually, only a selected group of users, e.g., domain specialists, are enabled to modify this set of tags. In a knowledge-acquisition step, it is usually generated according to the specific domain and the targeted closed-community of users.

2.2 KNIZR in a Nutshell

As outlined above, KNIZR relies on a set of tags that are assigned to resources by the users. To enable a knowledge-based approach, KNIZR utilizes SKOS [5], i.e., the *Simple Knowledge Organization System*, a W3C standard for knowledge organization for the web and especially the semantic web. SKOS is based on (simple) web-standards such as RDF, and provides a formal language for constructing, for example, thesauri or concept taxonomies. At its core it is quite simple, and thus usually easily extensible and maintainable.

The system then relies on the following items: Users, resources, tags, links, and content (i.e., the textual content of the resources and optional annotations). For the resources, we provide a convenient interface for upload and annotation. An example is shown in Figure 1. KNIZR facilitates the knowledge-based browsing, resource management (various resources such as: images, documents, URIs, ...), community management, tag management, and a powerful search interface.

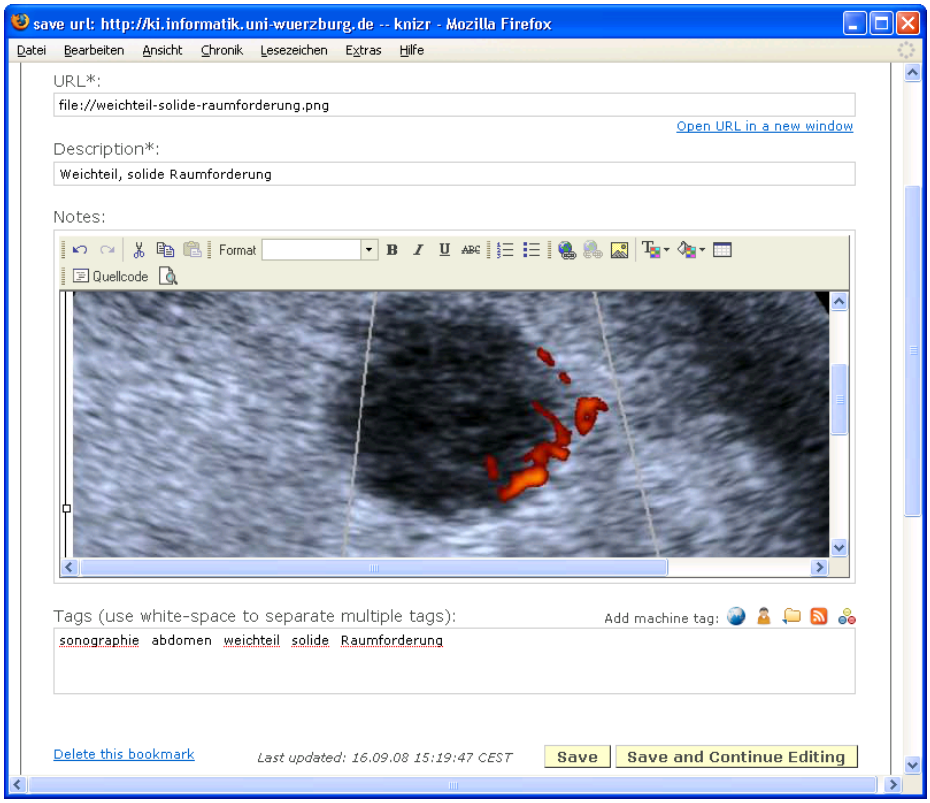

Fig. 1. KNIZR: Creating a resource/image with annotation and tags (medical example, in german).

Essentially, KNIZR extends common social tagging system on the resource side. While resources are represented by URI, the resources can also be textually be described and commented. Furthermore, KNIZR also allows for files, i.e., an URI representing a file can be utilized as a resource. Then, the file is automatically uploaded to the server of the system (if a maximal allowed size is not exceeded). As an optional extension, the system can be configured for closed-communities such that controlled vocabularies for the tags, i.e., the tag ontology of the system can be defined. As described above, a closed/predefined system of tags is then used. Usually, this set of tags is edited by subset of users ("experts"). This setting is especially applicable for closed communities, i.e., for specific applications projects with a fixed specific set of contributing users.

For general resource management, KNIZR offers the following functionality:

- Resources can be created, textually described and a set of tags can be assigned to them.
- Resources are defined for certain groups of users.
- Resources can be rated (ranked) by the users.
- Resources can be commented by the users.

Connections between resources/tags are thus either established using (explicit) links between the resources, implicit relations/links given by the a matching set of tags of different resources, or by considering the relations between the tags defined in a domain ontology. Since the KNIZR system utilizes the SKOS system for the representation of the ontological relations, the tags are defined using a standard semantic, and the set of tags can be transparently distributed and organized.

For the tag assignments, KNIZR utilizes a tag recommendation system that takes a short sentence as input, and then proposes a set of suitable tags. This feature is especially important for the targeted medical domain, since the users often do not have a lot of time when entering a new resource, a situation which is quite typical, for example, for medical doctors.

After a set of resources has been defined, a key point are efficient and effective search and browsing capabilities for these resources. According to Shneiderman [8] an effective user interface should implement features for getting an overview first, then zoom on some details, finally getting the details on demand. In this way, the so-called *visual information seeking mantra* is implemented. KNIZR implements these features by both a powerful search facility and also by using a tag cloud visualization. Additionally, for each resource a set of related tags/resources is shown such that the navigation between these is easily implemented. Then, the user can apply a dynamic query refinement for obtaining a refined set of resources within a more detailed cluster of information.

So far, we have partially implemented a prototype of the system that is currently being extended. First results of the application are quite promising and we are continously refining the system for a better usability and performance.

3 Application Scenarios

In the following, we describe two real-world application scenarios for the KNIZR system. We first describe an application in a tutoring context, where KNIZR is coupled to an eLearning system. After that, we describe an application for training and general knowledge management in the medical domain of sonography.

3.1 CaseTrain

The first application scenario is concerned with supporting a large case-based learning project at the University of Würzburg. The CaseTrain project [9] is part of an eLearning initiative of the University of Würzburg which aims at establishing a central eLearning platform for all faculties and supporting teaching staff. The goal is to create eLearning content and to make it available for all students, focusing on interactive online content for case-based or problem based learning (PBL [10]).

PBL is a learner-centered instructional strategy where learners are challenged with diagnostic or therapeutic problems with partial information only. Thus - apart from finding the correct solutions - the learners have to decide which information is needed to help to establish or exclude possible solutions. Requested information may not be given directly, but instead the learners have to find the actual findings themselves by

analyzing intermediary results (e.g. X-ray images in the medical domain or process sheets in economics). Since such images are essential for CaseTrain, the KNIZR system is an ideal complement for the core CaseTrain application. One example of a created case is shown in Figure 2 for the medical domain.

Fig. 2. CaseTrain-Player showing example case (from the medical domain, in german). Left side: Intermediary examination results (video); right side, top: Question about visible findings; right side, bottom: List of alternatives with result. Rightmost: Avatar.

KNIZR can easily complement CaseTrain and vice versa, since the KNIZR system can utilize resources contained within the CaseTrain system. For instance, resources and/or images can be linked to the system, and they can be tagged and annotated as usual. Furthermore, CaseTrain already contains an application ontology that can be mapped to the KNIZR system. Then, the annotation of CaseTrain cases can be utilized for locating suitable images within the system.

Additionally, CaseTrain can also utilize the KNIZR capabilities: For instance, specific (CaseTrain) cases can link to KNIZR images, videos or general resources for exemplifying certain findings of the case. Additionally, the user can search for suitable tags in order to locate similar images for a certain image contained in the case. In this respect, we can also utilize the CaseTrain ontology for locating the specific resources contained in the tagging system.

3.2 SonoConsult

The second scenario concerns the intelligent documentation and consultation system SONOCONSULT system [11] – a medical system for sonography. The system is in routine use in several hospitals, for example, in the DRK-hospital in Berlin/Köpenick and

in the University Hospital of Würzburg; the collected cases contain detailed descriptions of findings of the examination(s), together with the inferred diagnoses.

For a sonography (ultrasound) examination, the examiners need to closely inspect the ultrasound images, for example, see Figure 3 in order to document the correct findings. Since this process is highly subjective and also significantly dependent on the experience of the examiner, there are often discrepancies between beginners and more senior examiners concerning the correct findings, c.f., [12].

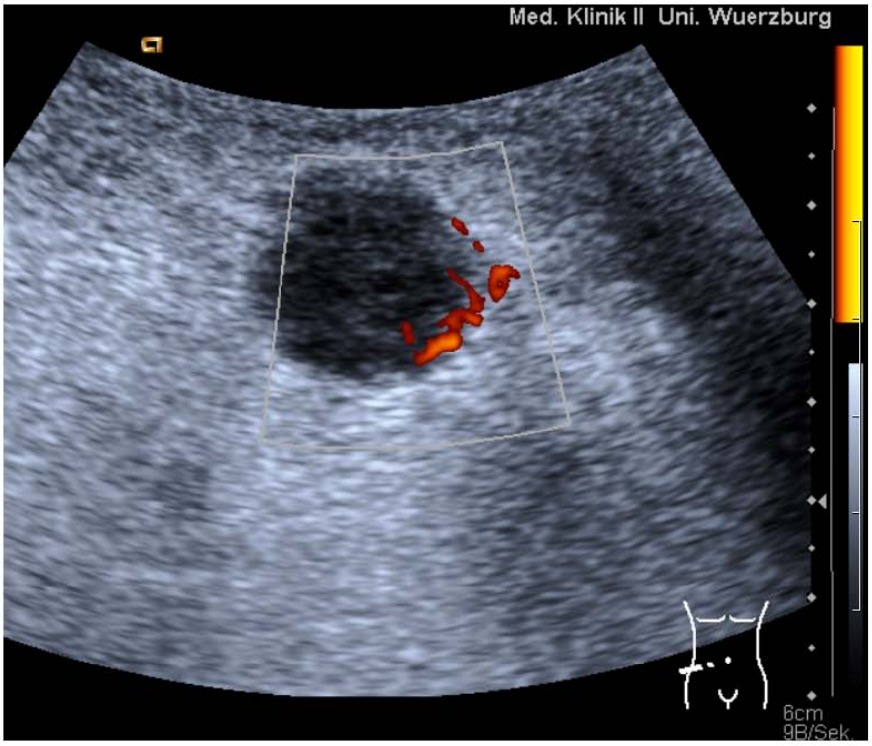

Fig. 3. Exemplary image from an ultrasound examination.

In order to improve this situation, KNIZR provides an ideal framework for implementing an image pool for tutoring, targeted training and general knowledge management. Special cases of ultrasound phenomena can be uploaded to the system. They can be easily described, tagged and annotated, and junior examiners can simply search for specific (difficult) situations. Additionally, hints can be easily communicated using the annotation and/or comment function of the system. Furthermore, using the SKOS features of the system the terminology of the different organ systems and diseases can be directly utilized for an effective retrieval of resources.

4 Conclusion

Social tagging systems can provide powerful and intuitive solutions to various knowledge management problems. In this paper, we have introduced the KNIZR system for tutoring and general knowledge management. We discussed the different components and features of the system, and we have described its application in two real-world application scenarios. So far, we have implemented a partial prototype of the system: First results are quite promising and we are continously refining and improving the system.

For future work, we aim to perform a comprehensive evaluation of the system in the sketched application scenarios. Additionally, we want to utilize as much background knowledge as possible in order to exploit the semantic features at their full level. We also plan to enhance the tag recommendation feature by discovering certain user subgroups and applying their features for a better recommendation user experience.

Acknowledgements

This work has been partially supported by the German Research Council (DFG) under grant Pu 129/8-2.

References

1. Antoniou, G., Harmelen, F.v.: A Semantic Web Primer. MIT Press, Cambridge, Mass. (2004)
2. Golder, S., Huberman, B.A.: Usage Patterns of Collaborative Tagging Systems. Journal of Information Science **32**(2) (April 2006) 198–208
3. Mislove, A., Koppula, H.S., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Growth of the flickr social network. In: Proceedings of the 1st ACM SIGCOMM Workshop on Social Networks (WOSN'08). (August 2008)
4. gnizr: http://code.google.com/p/gnizr/
5. Brickley, D., Miles, A.: SKOS core vocabulary specification 2005-11-02. W3C working draft, W3C (November 2005) http://www.w3.org/TR/swbp-skos-core-spec.
6. Hammond, T., Hannay, T., Lund, B., Scott, J.: Social Bookmarking Tools (I) - A General Review. D-Lib Magazine **11**(4) (April 2005)
7. Jäschke, R., Marinho, L.B., Hotho, A., Schmidt-Thieme, L., Stumme, G.: Tag Recommendations in Folksonomies. In: Knowledge Discovery in Databases: PKDD 2007. Volume 4702 of Lecture Notes in Computer Science., Berlin, Heidelberg, Springer (2007) 506–514
8. Shneiderman, B.: Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces. In: Proc. Second Intl. Conference on Intelligent User Interfaces. (1997) 33–39
9. Puppe, F., Hörnlein, A.: CaseTrain. Online, available at: http://www.casetrain.de.
10. Schmidt, H.G.: Foundations of Problem-Based Learning: Some Explanatory Notes. Medical Education **27** (1993) 422–432
11. Huettig, M., Buscher, G., Menzel, T., Scheppach, W., Puppe, F., Buscher, H.P.: A Diagnostic Expert System for Structured Reports, Quality Assessment, and Training of Residents in Sonography. Medizinische Klinik **99**(3) (2004) 117–122
12. Atzmueller, M., Puppe, F., Buscher, H.P.: Profiling Examiners using Intelligent Subgroup Mining. In: Proc. 10th Intl. Workshop on Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP-2005), Aberdeen, Scotland (2005) 46–51