

TagCare: A Personal Portable Tag Repository

Evgeni Golov

Heinrich-Heine-University Düsseldorf
Dept. of Information Science
Universitätsstr. 1
40225 Düsseldorf, Germany
evgeni.golov@uni-duesseldorf.de

Katrin Weller

Heinrich-Heine-University Düsseldorf
Dept. of Information Science
Universitätsstr. 1
40225 Düsseldorf, Germany
katrin.weller@uni-duesseldorf.de

Isabella Peters

Heinrich-Heine-University Düsseldorf
Dept. of Information Science
Universitätsstr. 1
40225 Düsseldorf, Germany
isabella.peters@uni-duesseldorf.de

ABSTRACT

This paper presents TagCare¹ as a tool which allows users to maintain their personal tagging vocabulary and to carry it along different platforms. TagCare collects the tags which a user has applied within several social software tools. These personal tags may be edited and structured, e.g. interrelated with hierarchical and other semantic relations.

Categories and Subject Descriptors

H.1.2 User/Machine Systems: *Human information processing*; H.3.1 Content Analysis and Indexing

General Terms

Management, Documentation, Human Factors.

Keywords

Folksonomy, persononomy, social tagging, tag gardening, lightweight ontology, Web 2.0, Social Semantic Web.

1. INTRODUCTION

Ways to organize tags in folksonomy-based systems have recently been arousing interest in the respective research community. Activities to edit and structure tags have been described as “tag gardening” [1], [2]. The image of tag gardening is used to describe processes of manipulating and re-engineering folksonomy tags in order to make them more productive and effective.

The aspect of gradually creating semantic representations out of folksonomies is related to the approaches in “ontology maturing” [3] and “semantic enrichments” [4] and can thus be a basic foundation of (collaborative) ontology engineering. Furthermore, gradual refinement of folksonomy tags and a stepwise application of additional vocabulary control and semantic structure to tagging systems [5] is a promising approach for handling the current problems [6] of folksonomies like inconsistent vocabularies and varieties of synonyms, spelling variants, misspellings, language variants and a general lack of semantics.

These problems do not only occur in systems with large user communities but may even be a problem on a personal level. It

happens, that even one single user applies different spelling variants of the same word (e.g. “SemanticWeb”, “semanticweb” and “semantic_web”) or different synonyms (e.g. “Web2.0” and “Social Web”) to different documents in his collection – either within one single platform or when using different folksonomy-based platforms in parallel. This causes additional expenses when searching personal document collections. Searching for all documents on the topics *Web 2.0* and *Semantic Web* would now require the user to enter all his spelling variants and synonyms as search terms. Thus, particularly someone using different folksonomy-based tools in parallel will profit from using his own terminology very consistently.

Motivated by this, TagCare has been designed to help users to apply the same tags uniformly in different platforms. This will create a consistent cross-platform personal tagging vocabulary or “persononomy” [5]. The user can assemble all tags which he has used within different systems and may furthermore create his own vocabulary hierarchy, synonym collections and cross-references to related terms to establish some lightweight form of controlled vocabulary or ontology.

2. TagCare: The Personal Tag Repository

TagCare currently supports Flickr [7], Bibsonomy [8] and del.icio.us [9]; i.e. it allows a user to import his personal tags from these platforms into TagCare and to maintain them all in one place. TagCare has been developed in PHP; with JavaScript and Ajax for the user interface. It is realized as a web application (a browser plug-in is under development) and is using a MySQL database to store the information. It collects tagging data from the mentioned applications via their APIs².

The database includes three types of data: users, tags, and tag interrelations (with the latter two always being assigned to exactly one user). The following sections describe the functionalities for editing and structuring tags which have been included so far, extensions are planned.

2.1 Tag Maintenance: Collecting and Editing Tags

When a user signs up for TagCare he provides his login data for the different social tagging services he uses and can then import

¹ This paper presents work in progress. TagCare is still under development. The latest demo for testing some of the functionalities is available at <http://www.demo.tagcare.org>.

² Using the API implementations phpFlickr and php-delicious, to integrate Bibsonomy an own implementation (phpBibsonomy) had to be written. So far, only the import function is used.

his tags from Flickr, Bibsonomy or del.icio.us. Basic statistics are provided on how often the user has applied single tags in total for all services and separately for Bibsonomy and del.icio.us³. This information can be displayed as tag cloud or as a ranking of the most frequently or least frequently used tags. This can help to detect tags which have been used too often (and thus become too general) or very seldom (which may indicate that they should be bundled with others).

Basic editing functionalities for tags comprise renaming and deleting of tags as well as directly creating new terms in TagCare. In future, users should be enabled to predefine their preferred spelling variants (e.g. preferring singular over plural, preferring British English over American English spelling, or separating compound words by underscore or camelCase). Coupled to an underlying dictionary or even some software for speech pattern recognition, the user could be warned if he is deviating from the favored settings (e.g. using a plural form although singular is preferential).

2.2 Vocabulary Structuring with Semantic Relations

The *advanced editing* options in TagCare concern the organization of tags. Knowledge relations between concepts are the structures that add semantics to a tag collection. Different standard relations are used for classical knowledge organization systems and may be adopted to enhance folksonomies [10]. The fundamental types are hierarchical relations. Hierarchies can easily be established in tag care by drag and drop principle. Currently, TagCare does not distinguish *is_a* und *part_of* hierarchies.

Furthermore, a relation of equivalence is of importance as it interlinks synonyms and quasi-synonyms, i.e. words that have exactly or almost the same meaning or can be regarded as being the same within a certain context. In TagCare, synonyms can be interlinked via a pre-defined relation of synonymy. Finally, one may label two tags as being generally related terms (figure 1). This unspecified relation should in future be complemented by some more specific semantic relations, such as “*is_opposite_of*” – and further tag interrelations which can be freely named by the user.

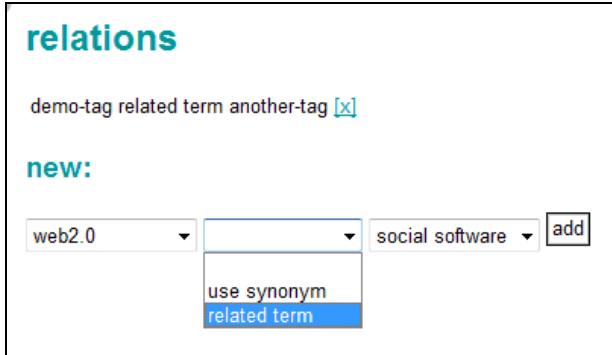


Figure 1. Connecting two tags as related terms in TagCare.

³ These statistics are based on usage data (how often a user has applied a certain tag) as submitted by the services’ API. Currently, Flickr does not provide this type of information.

3. CONCLUSION & FUTURE WORK

So far the basis is provided for collecting, editing and structuring tags platform independently on a personomy level. While this currently means importing tags from different social software services, the next steps will be the other way round: to enable directly searching social software collections via tags from TagCare as well as directly tagging documents out of TagCare.

Tag suggestions may be derived from different folksonomy statistics, automatic clustering mechanisms may be implemented to suggest tag interrelations. In the long run, the exchange of tagging vocabularies between different TagCare users is envisioned. Additionally, the evaluation of functionalities and usability tests are planned as future work. Data from other tools such as YouTube and Connotea should be embedded.

4. REFERENCES

- [1] Governor, J.: On the Emergence of Professional Tag Gardeners. Blog Post, 10.01.2006, retrieved 22. July, 2008, from: <http://www.redmonk.com/jgovernor/2006/01/10/on-the-emergence-of-professional-tag-gardeners/> (2006).
- [2] Weller, K., Peters, I.: Seeding, Weeding, Fertilizing. Different Tag Gardening Activities for Folksonomy Maintenance and Enrichment. In: Triple-I Conference, Proceedings of I-Semantics ‘08, Graz, Austria, pp. 110–117 (2008).
- [3] Braun, S., Schmidt, A., Walter, A., Zacharias, V.: The Ontology Maturing Approach for Collaborative and Work Integrated Ontology Development. Evaluation, Results and Future Directions. In: Proceedings of the International Workshop on Emergent Semantics and Ontology Evolution at ISWC/ASWC 2007, Busan, South Korea, pp. 5–18 (2007).
- [4] Angeletou, S., Sabou, M., Specia, L., Motta, E.: Bridging the Gap between Folksonomies and the Semantic Web. An Experience Report. In: Bridging the Gap between Semantic Web and Web 2.0, SemNet 2007, pp. 30–43 (2007).
- [5] Hotho, A., Jäschke, R., Schmitz, C., Stumme, G.: Bibsonomy. A Social Bookmark and Publication Sharing System. In: Proceedings of the Conceptual Structure Tool Interoperability Workshop at ICCS, Aalborg, Denmark 2006, pp. 87–102 (2006).
- [6] Peters, I., Stock, W.G.: Folksonomy and Information Retrieval. In: Proceedings of the 70th ASIS&T Annual Meeting. Vol. 44, pp. 1510–1542 (2007).
- [7] Flickr: <http://www.flickr.com>.
- [8] Bibsonomy: <http://www.bibsonomy.org>.
- [9] Del.icio.us: <http://del.icio.us>.
- [10] Peters, I., Weller, K.: Paradigmatic and Syntagmatic Relations in Knowledge Organization Systems. Information – Wissenschaft und Praxis 59 (2), pp. 100–107 (2008).