

Discovering Keyword Relations

Peter KAJAN*

Slovak University of Technology
Faculty of Informatics and Information Technologies
Ilkovičova 3, 842 16 Bratislava, Slovakia
peto.kajan@gmail.com

One of the popular approaches for discovering new knowledge about the users or the content on the Web is using the wisdom of the masses. We already know folksonomies, resulting from collaborative assignment of tags or keywords to resources. Analyses of this information can reveal relations between users, resources or tags. Similar results can be achieved by analyzing how users navigate on the Web and access web resources. For example, if significant number of users visit page with keyword A and right after they visit page with keyword B, we can conclude that this keywords may describe similar concept. In our work we investigate the potential of folksonomies and similar data structures for acquiring various relations between keywords.

We know different kinds of possible relations. If only used type of relation is a *parent-child relation* forming a hierarchy of keywords, we get a taxonomy. By adding other relations like *synonyms*, the relation between keywords with similar meaning, a taxonomy turns into an ontology.

Relations between the keywords could be also found using web resources like the lexical database Wordnet [1] or users created encyclopaedia, Wikipedia [2]. Our idea is to take advantage of such resources on the Web, which provide at least basic semantic grounding. For instance, Linked Data initiative associates data represented by URIs over the Web and there are plenty of tools and services, built on the top of Linked Data. Useful for dealing with homonyms might be Open Calais which is able to disambiguate a keyword with multiple meanings by using its context.

Another idea is to explore the dynamics of the Web usage. Users' clickstreams may provide sequences of keywords with similar meaning, describing the same concept or keywords forming hierarchy.

There are plenty of works presenting algorithms for taxonomies creation which are based on statistics, clustering or set theory [1,3]. Many authors also published their approaches of turning folksonomies into ontologies by either automatic [4] or semiautomatic processes [2].

* Supervisor: Michal Barla, Institute of Informatics and Software Engineering

Our goal is to propose a method of turning folksonomies into ontologies, which combines several approaches in order to build a knowledge base characteristic for a group of users, who contributed to the creation of underlying folksonomy. The method consists of two steps: ontology creation and ontology enrichment. Our approach takes a folksonomy-like structure based on the dynamic records of the Web usage as an input and employs existing solution combined with usage dynamics analysis in the first step. Next, semantic web services are used for finding new relation and enrich the emergent data structure. This process is outlined in the Figure 1.

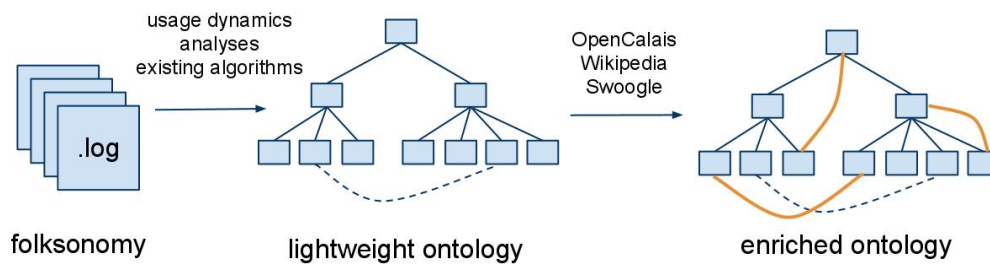


Figure 1. Proposed process of turning folksonomies into ontologies

Another idea is to analyse the relations, their properties like commutativity and transitivity or relations between relations. Evaluation will be done in the domain of the personalized web search. Personalized search results will be presented by analysing the keywords in search queries and extending them using created knowledge base.

Acknowledgement. This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. KEGA 345-032STU-4/2010

References

- [1] Barla, M., Bieliková, M.: On Deriving Tagsonomies: Keyword Relations Coming from Crowd. In *Proc. of the 1st Int. Conf. on Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems*, Springer-Verlag, Berlin, Heidelberg, pp. 309-320, 2009.
- [2] Van Damme, C., Siorpaes, K.: FolksOntology: An Integrated Approach for Turning Folksonomies into Ontologies, In *Bridging the Gap between Semantic Web and Web 2.0*, SemNet 2007, pp. 57-70, 2007.
- [3] Lu, C., Chen, X., Park, E.K.: Exploit the tripartite network of social tagging for web clustering, In *Proc. of the 18th ACM Conf. on Information and Knowledge Management*, New York, NY, USA, ACM, pp. 1545–1548, 2009.
- [4] Tang, J., Leung, H., Luo, Q., Chen, D., Gong, J.: Towards ontology learning from folksonomies, In *Proc. of the 21st Int. Jont Conf. on Artificial Intelligence*, San Francisco, CA, USA, Morgan Kaufmann Publ. Inc., pp. 2089–2094, 2009.