Christoph Rensing, Philipp Scholl, Doreen Böhnstedt, Ralf Steinmetz: *Recommending and Finding Multimedia Resources in Knowledge Acquisition Based on Web Resources*. In: Proceedings of 19th International Conference on Computer Communications and Networks, p. 1-6, IEEE eXpress Conference Publishing, August 2010. ISBN 9781424471140.

# Recommending and finding multimedia resources in knowledge acquisition based on Web resources

Christoph Rensing, Philipp Scholl, Doreen Böhnstedt, Ralf Steinmetz
Multimedia Communications Lab
Technische Universität Darmstadt
Darmstadt, Germany
{First_Name.Last_Name}@kom.tu-darmstadt.de

***Abstract*—Personal knowledge acquisition in a world where people have to acquire knowledge for task completion on the job constantly is done increasingly based on Web resources. Web resources ranging from mainly text based forms like Wikis up to multimedia resource like videos document knowledge and can be used for knowledge acquisition. Due to the participation of communities in creating content in Web 2.0 applications the mass of available resources increases steadily. Searching for high quality and reliable resources using a search engine becomes impossible. Thus, recommending relevant resources becomes more and more important. For this learning scenario, we present and compare different approaches for recommendation of multimedia resources we integrated in an application for collaborative management of Web resources for learning purposes, called CROKODIL.**

***Recommender Systems, Knowlegde Acquisition, Web resources, Social Networks***

## I. INTRODUCTION

People have to acquire knowledge for task completion on the job constantly due to the evolving and changing knowledge and the fast development of new technologies [1]. Knowledge can't be acquired once for lifetime. The Web offers a huge quantity of resources that can be used for knowledge acquisition and learning. On the one hand, Web users document knowledge in mostly text-based formats like in FAQs, Wikis, Blogs or documents (like word- or pdf-files). On the other hand, we have multimedia resources like videos, especially lecture recordings, or interactive flash-based movies (e.g. for slide presentations) also provided by Web users.

One major challenge, besides some others [2], for the learner is to identify reliable and relevant resources which match his needs. Searching for resources using a common search engine often does not bring sufficient results. We have developed a concept and a first prototype of a platform which supports the user in the management of Web resources for learning and knowledge acquisition purposes. Finding relevant resources in a collaborative scenario is supported in the platform by integrating different methods for recommendation of resources.

This paper is organized as follows. Section 2 summarizes different state of the art approaches for recommender systems. In section 3 we present an overview of the functionality and main concepts of our CROKODIL platform. How different methods for recommending and finding resources are integrated in the platform is described in section 4. The paper ends with a summary and outlook.

## II. RECOMMENDER SYSTEMS

Recommender systems serve to recommend so far unknown resources or information to a user which could be interesting and relevant for her. Generally they filter and rank resources or information based on implicit and explicit preferences of the user. Systems and mechanisms, used in different scenarios like recommendation of product items in electronic stores, can be distinguished as Collaborative Filtering systems and contents-based systems [3]. Content-based methods suggest new resources which are similar concerning the contents of resources which a user has used or preferred so far [4]. They usually apply technologies from Information Retrieval to determine similarity of the content of resources. However, Collaborative Filtering methods do not analyze contents. They are based on the similarity between preferences of users [5]. They first detect users with similar interests and suggest resources which have been judged positively from these like-minded users as recommendations. Information about the user can be explicitly specified by the user in a profile or implicitly collected by tracking the user's behavior.

Web users are increasingly members of so-called social networks in the Web. The number of users of social network sites grows continuously. They build groups based on fields of interest quite often. Thus, information about the fields of interest in social networks can be used to improve Collaborative Filtering methods for recommendation substantially [6][7].

Increasingly, in Web platforms like social bookmarking sites (e.g. Delicious) or video sites (e.g. YouTube) or picture sites (e.g. Flickr), resources are tagged by the users. Tagging is a simple means for users to organize their items in without the constraints imposed by strict hierarchical organization (e.g. having to choose a unique category for items). By aggregating the tags of all users, folksonomic structures emerge. These structures can be used for recommendation of similar resources to the users. Usually, the rank of the recommended resources is determined by the frequency of utilization. Additionally some

This work is supported by the German Federal Ministry of Education and Research. The responsibility for the contents of this publication lies with the authors.

platforms (e.g. Delicious) allow users to recommend resources to another user by adding a special "for:username" tag to a resource.

In summary, different methods for recommendation of resources in the Web exist. Most applications using recommender systems can be characterized by a recommendation of homogeneous resources which have a uniform format. Algorithms for content detection which are used in content based recommendation methods differ for different types of resources. As our scenario includes recommending heterogeneous resources, these algorithms would need to be used in combination, making it very complex. Therefore, we need methods that do not analyze the content itself. Collaborative Filtering methods perform quite well if there is enough data in form of ratings available. Otherwise the so-called cold-start problem exists. Our scenario takes place in a more specialized application domain compared to common social communities, thus this critical mass of data is not necessarily given. This can not be assumed in our more specialized application scenario than in common social communities. In addition, Collaborative Filtering methods have the disadvantage that results are biased by the equalizing force of the masses. In our scenario, where users follow individual tasks, this could be a limitation.

Subsumed, all commonly used methods for recommendation of Web resources have a limited usability for our scenario which is characterized by heterogeneous resources and high individuality of users. How we address this challenge is presented in chapter 4.

## III. CROKODIL – A PLATFORM FOR COLLABORATIVE KNOWLEDGE ACQUISITION BASED ON WEB RESOURCES

### A. *Overview*

To support the user in acquiring knowledge based on Web resources we have developed a concept and prototype of a platform for collaborative knowledge acquisition, called CROKODIL. CROKODIL combines functionalities for the management of resources and functionalities which are part of social networks.

Building blocks of the CROKODIL platform are a (1) plugin for the Web browser Firefox which is used by a user during online search, (2) a semantic network as the back-end storing the data about resources and users and (3) a Web application, called Knowledge Portal, which provides the functionality to build a social network and to navigate through the net of resources and the community graph. Recommendations are integrated both in the plugin as well as in the Knowledge Portal.

The plugin, shown in figure 1, is used to add Web resources or snippets (textual Web resource fragments which contain the relevant and useful information only) to the net of resources. In general, all kind of resources can be used within the CROKODIL platform. Added resources are tagged by the user. Specifically, we apply the concept of semantic tagging [8], where a tag is associated with its semantic meaning. The semantic meaning is assigned to a tag by selection of a type. We support a predefined set of types: *Topic*, *Person*, *Location*, *Event*, *Learning Goal*, resource *Type* and also a non typed tag.

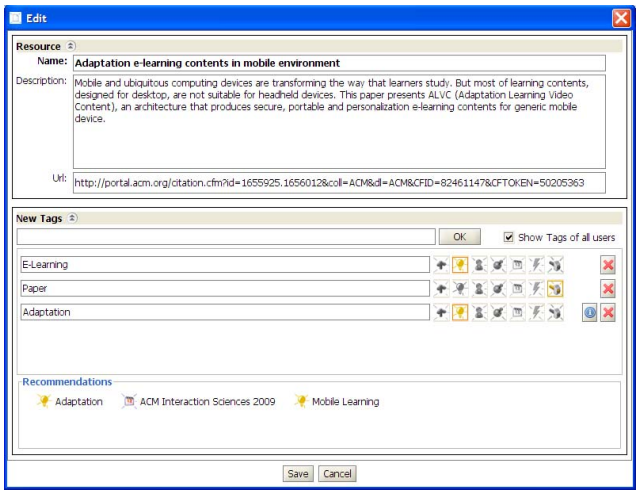


Figure 1. CROKODIL plugin for tagging of Web resources

Due to extensibility of the predefined semantic types of tags users can implement their own mental concepts for tagging resources. Tagging in this fashion allows the user to adapt the structure to her own needs, while being simple enough to be handled by non-domain experts for structuring information generically.

### B. *Personal and community net of resources*

The resources, the tags, the relationships between tags and the relationships between a resource and its tags are stored in the back-end in form of a semantic network. Semantic Networks are graphical notations of knowledge representations consisting of nodes and relations between those nodes [9]. In our case, the semantic network consists of different types of nodes (the semantic tags, e.g. *Person* or *Topic,* and the resources) that are connected by relations, as shown in Figure 2. A semantic meaning (e.g. *is author of*, *has topic, is subtopic*) can be assigned to the relations also.

By tagging different resources, the user constructs his personal net of resources, called Personal Resource Net (PRN). This PRN represents the user's personal mental organization of the knowledge domain. There are no experts upfront who model the semantic networks in advance. The representation, the semantic network, can be used for a semantic search of resources at a later time or for browsing through the network. Semantic searches enhance the possibilities of full-text searches because it can search along relations, e.g. a resource can be found with an event tag, which the user has associated to the resource, like the "*VIS*" tag in figure 2 – even if the resource doesn't contain the event. Not only resources but also all other semantic types of nodes (esp. topics) can be searched in the network. This is very relevant in our scenario, where users have no prior knowledge about the domain and the topics. For searching and browsing the Knowledge Portal, shown in figure 2 and figure 3, is used.

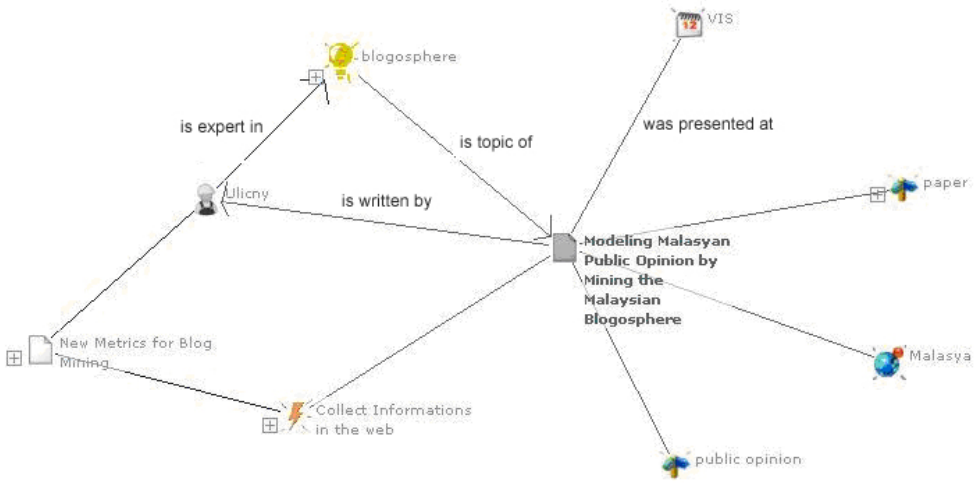


Figure 2. Example for representation of resources, tags, and their relations

In CROKODIL different users of a social community work together using the CROKODIL platform. By linking nodes of different PRN's a combined net of resources is originated. We call this linked net the Community Resource Net (CRN). The union of the PRN's is realized so far by linking resources which are identical (e.g. having the same URL) or by linking same or similar tags. More complex algorithms have to be integrated in future. This CRN offers a more complex view of a knowledge domain and contains more resources and more information about the resources than in an individual PRN.


Figure 3. Representation of resources in the Knowledge Portal

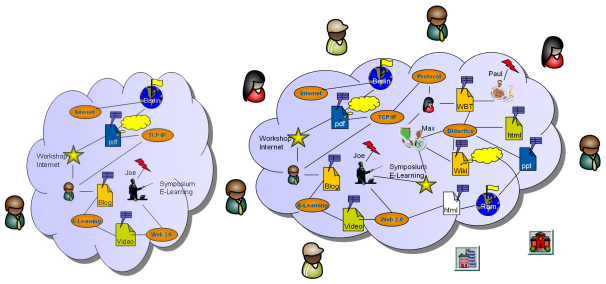

Figure 4. Increase of domain knowledge in a CRN

Figure 4 shows this increase of information about a knowledge domain in a schematic view.

### C. *Community Graph*

As we said before, we combine the functionalities to manage resources as described with functionalities of social networks in the CROKODIL platform. The main common features of social networks in the Web are [10]:

- representation of a member in form of a profile,
- modeling of social relationships by linking members as friends,
- organization of groups as a subset of all members,
- provisioning of means for communication between members, like instant messaging or e-mail.

We integrated these functionalities in CROKODIL and combined it with the CRN. The social relationships and the affiliation to a group are modeled in the semantic network also. They build the so called Community Graph. Therefore, we added to types of relations "*is friend of*" and "*belongs to group*". Figure 5 shows the combination of the CRN and Community Graph in a schematic view.

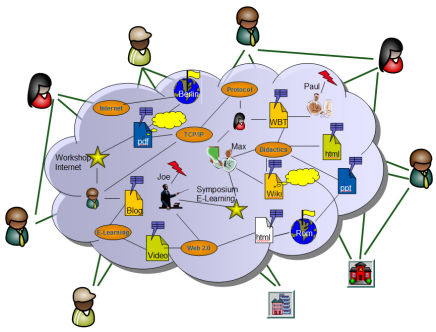

Figure 5. Combination of Community Graph and CRN

## IV. RECOMMENDING AND FINDING RESOURCES IN CROKODIL

The CROKODIL platform supports the management of heterogeneous resources ranging from text-based resources up to multimedia resources. A commonly used context-based recommendation method is restricted to a dedicated type of resources. Therefore the integration of such a method in CROKODIL can only suggest resources of the selected type to the user. The user maybe interested in resources of different types. Therefore we have to add not only one but different methods.

In the following, we distinguish two types of methods for finding and recommending resources. First we present the methods which are activated by the user to find relevant resources and by a member of the community to suggest resources to a dedicated user. After that we summarize the methods which recommend resources related to the currently activated node (esp. resource) inside the semantic network. How the recommendations are integrated in the plugin as well as in the knowledge portal is presented in the third part of this chapter.

### A. *Methods for finding resources and active recommendation*

The Community Graph represents social relationships and stores information about the learners using the CROKODIL platform. It helps the learner to identify people who share the same interests. If they are member of the same interest- or learning group, similar interests can be assumed. The user can browse through a part of a CRN, generated by filtering only this part which shows resources of his friends or of the members of a group he belongs to. Figure 6 shows this exemplarily in a detailed view of the Knowledge Portal. In this way the user finds personally unknown and for his own task potentially relevant resources. The navigation through the Community Graph is not limited to the user's direct friends or to the members of his group, called "strong ties". He can also browse to the resources of friends of friends and so on. These "weak ties" are very relevant to find new resources, which are not known by the strong ties so far [6].

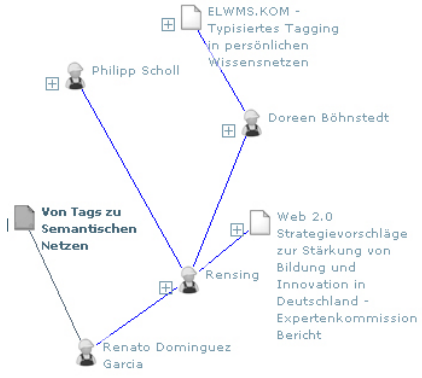


Figure 6. Finding resources by browsing the Community Graph

Active recommendations by friends can be realized in the CROKODIL community by linking resources to persons, using a new type of relation "*maybe interesting for*". The relation can be added by tagging the resource like it is realized in Delicious "*for:username*". The user for whom the new resource is suggested has to be able to mark the relation or to remove it.

### B. *Methods for recommendations based on information about the current node*

Using the CROKODIL platform the user browses through resources and concepts as shown before. If he has activated a resource which is part of the CRN, by clicking on it, the platform presents her resources which are related to the active resource.

First of all, we offer recommendations based on the structural properties of the semantic network. Therefore we use the semantic information which has been modelled in the semantic network by the users of the platform. In the example, shown in figure 7, we have two different resources in two different PRNs which are tagged with "*Communication Networks*". Based on this information, the "*Wiki*" resource can be recommended to the user shown on the left hand side. The relations between tags are used in addition to the relations between resources and tags. Based on the semantic knowledge that "*Protocols*" "*is a sub topic of*" "*Communication networks*", the "*script*" can also be recommended. The advantage of this method compared to content based methods is that it can be used for heterogeneous resources because it is independent of the type a resource has.

Recommendation based on structural information is not restricted to recommendation of resources in contrast to content based methods. It can be supported for nodes of every type in our semantic network. Thus, e.g. persons which are related to topics or resources which are related to an evant can also be recommended. Nodes of which type are recommended depends form the node which is activated currently (see figure 10 and 11).

In addition to structural recommendations we have realized content-based recommendations. The methods we use are limited to text-based resources so far. Semantic similarity between two textual resources can be derived from comparing resources using standard Information Retrieval approaches. However, a problem that often arises when taking into account documents prepared for different audiences is the challenge of vocabulary gap, meaning that different vocabulary is used in these resources, thus diminishing reliability of similarity measures.

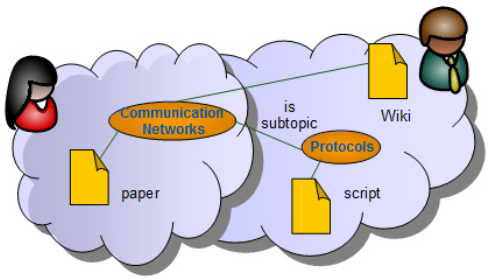


Figure 7. Resource recommendation based on structural properties

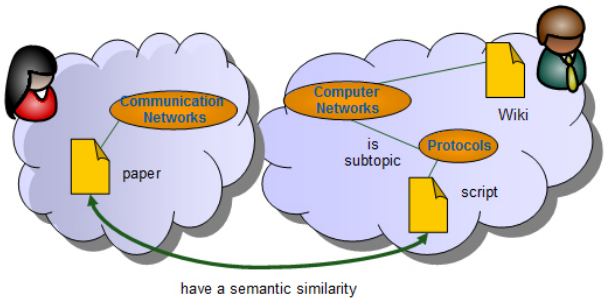


Figure 8. Resource Recommendation based on semantic similarity

Thus, we employ a semantic similarity measure that abstracts from the exact terminology used in the resources [11].

Figure 8 shows an example of the usage of content based recommendation. The two users use different tags "*Communication networks*" and "*Computer networks*" for the same topic. Based on structural properties the two PRNs are not connected by these concepts. Based on a semantic comparison of the documents in the CRN a similarity between the "*paper*" and the "*script*" can be detected. Therefore the resources "*script*" and the resource "*Wiki*" based on the structural information can be recommended nevertheless.

### C. *Presentation of recommendations in CROKODIL*

Recommendations are presented to the user in CROKDOIL in the Firefox plugin and in the Knowledge Portal as well. Integration in the Firefox plugin, like shown in figure 9, helps the user to detect relevant resources during his online research using a search engine. He can see directly, whether the resource he has selected is part of his Personal Resource Network or the joint Community Resource Network. In addition recommended resources are shown in a box captioned "More similar resources".

We assume that the user uses the web portal not during the actual research task but in a later time when she browses through the joined Community Graph and Community Resource Net. In the main part of the Knowledge Portal, all information about the currently active node of the semantic network is shown.

Recommendations are presented in boxes in addition to this information. As can be seen in figure 10 not only similar resources are recommended but also topics which could be relevant to the resource.

The type of the recommendations which are shown in the boxes depends on the type of the node which is currently shown. In contrast to figure 10, where a resource is shown, in figure 11 a topic is shown which results in a recommendation of a person which is a potential expert in this topic.

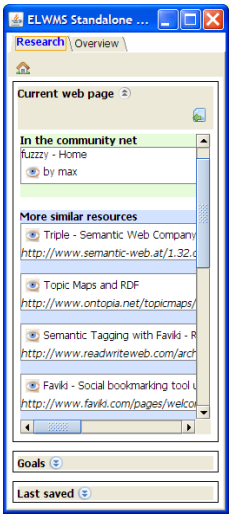


Figure 9. Recommendation of resources integrated in the plugin


Figure 10. Recommendations of nodes related to a "*resource*" node in the Knowledge Pportal

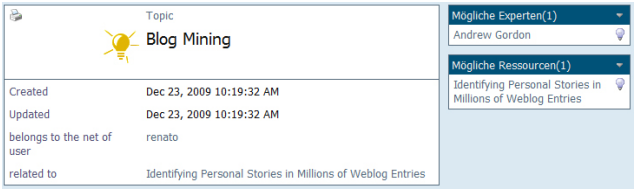


Figure 11. Recommendations of nodes related to a "*topic*" node in the Knowledge Portal

## V. SUMMARY AND OUTLOOK

In this paper we presented a new platform for collaborative management of heterogeneous web resources which is designed for support of the knowledge acquisition based on web resources. Due to the mass of available resources using a search engine for searching high quality and reliable resources for knowledge acquisition becomes impossible. Users of the CROKODIL platform can perform a semantic search for new resources based on all resources which are part of the Community Resource Net. Recommendation of resources, as well as of other nodes of the CROKODIL net, is based on different methods. The semantic knowledge about resources and tags, describing the resources, allows the recommendation based on the structure of the overall semantic net. This property allows recommendation of heterogeneous resources.

So far we implemented the prototype. It will be used in two different scenarios at university and industry during the next months. This deployment will be the base for the long-term evaluation of the usability of the CROKODIL platform and of the acceptance and quality of the different recommendation methods.

In future we will investigate how the information stored in the Community Graph can be used to improve the methods for recommendation. Another aspect we are working on is tagging itself. We want to support the user in tagging by recommending tags and by automatic identification of tag types.

## REFERENCES


[1] M. Eraut: Informal learning in the workplace. In: Studies in Continuing Education, Vol. 26, No. 2, p. 147-273, 2004.

[2] P. Scholl, D. Mann, C. Rensing, R. Steinmetz: Support of Acquisition and Organization of Knowledge Artifacts in Informal Learning Contexts. In: European Distance and E-Learning Network: EDEN - Book of Abstracts, p. 16, June 2007.

[3] G. Adomavicius, A. Tuzhilin: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 6, p. 734-749, 2005.

[4] R.L. Mooney, L. Roy: Content-based book recommending using learning for text categorization. In: Proceedings of the fifth ACM conference on Digital libraries, p. 195-204. 2000

[5] J. B. Schafer, D. Frankowski, J. Herlocker, S. Sen: Collaborative Filtering Recommender Systems. In: The Adaptive Web, Springer LNCS, p. 291-324, 2007.

[6] R. Schenkel, T. Crecelius, M. Kacimi, T. Neumann, J. X.r Parreira, M. Spaniol, G. Weikum: Social Wisdom for Search and Recommendation. In: *IEEE Data Engineering Bulletin*, Vol. 31, No. 2, p. 40-49, 2008.

[7] Y. Xu, L. Zhang, W. Liu: Cubic Analysis of Social Bookmarking for Personalized Recommendation. In: Frontiers of WWW Research and Development – Proceedings of the APWeb 2006, Springer LNCS, p. 733-738, 2006.

[8] D. Böhnstedt, P. Scholl, C. Rensing, R. Steinmetz: Collaborative Semantic Tagging of Web Resources on the Basis of Individual Knowledge Networks. In: Proceedings of First and Seventeenth International Conference on User Modeling, Adaptation, and Personalization UMAP 2009, Springer LNCS, p. 379-384, 2009.

[9] J. F. Sowa: Semantic networks. In: SC Shapiro (Edi.), Encyclopedia of Artificial Intelligence 2, John Wiley, New York, p. 1493-1511, 1992.

[10] I. O'Murchu, J. G. Breslin, S. Decker: Online Social and Business Networking Communities. DERI Technical Report 2004-08-11, online available http://www.deri.ie/fileadmin/documents/DERI-TR-2004-08-11.pdf., 2004.

[11] J. Grimm: Computing Semantic Relatedness of Small Text Fragments using Wikipedia. Master Thesis, Technische Universität Darmstadt, 2009.