New Tagging Paradigms for Enhancing Collaboration in Web 2.0 Communities

Andreas Nauerz
Matthias Brück
Martin Welsch
IBM Research and Development
71032 Böblingen, Germany
{andreas.nauerz|mbrueck|martin.welsch}
@de.ibm.com mail

Fedor Bakalov
Birgitta König-Ries
University Jena
07743 Jena, Germany
{fedor.bakalov|koenig}
@informatik.uni-jena.de

Abstract

In this paper we present new sophisticated tagging paradigms and their influence on users collaboration behavior and the construction of user– and context–models.

We present paradigms like alien tagging which allows one user to apply tags for another user, reputation-based tagging which allows users' expertise to influence tags' weights, quantitative tagging which allows users to manually manipulate tags' weights, anti tagging which allows users to specify "negative tags", tag voting to solve the tag space littering problem by e.g. allowing users to vote against tags, tag expiry which allows tags to have a lifetime, contextual tagging which allows tags to be associated to certain context profiles, and so forth and describe how these can be used to refine our models and to perform even more valuable adaptations or to issue more valuable. We also allow for mechanisms to follow users' tagging "trails" in order to learn from what they are tagging.

All these techniques aim to provide the user with more advanced ways, to add, filter, group and view tags.

The concepts presented are currently been prototypically implemented within IBMs WebSphere Portal and can be presented in a live demo at the workshop.

1 Introduction

In recent years Enterprise Information Portals have gained importance in many companies. As a single point of access they integrate various applications and processes into one homogeneous user interface. Today, typical Portals are comprised of a huge amount of content. They are no longer exclusively maintained by an IT department, instead, Web 2.0 techniques are used increasingly, allowing user generated content to be added. These systems grow quickly and in a more uncoordinated way as different users possess different knowledge and expertise and obey to different mental models. The continuous growth makes access to really relevant information difficult. Users need to find task- and role-specific information quickly, but face information overload and often feel *lost in hyperspace*. Thus, users often miss out on resources that are potentially relevant to their tasks, simply because they never come across them. On the one hand, users obtain too much information that is not relevant to their current task, on the other hand, it becomes cumbersome to find the right information and they do not obtain all the information that would be relevant.

The recent popularity of collaboration techniques on the Internet, particularly tagging and rating, provides new means for both semantically describing Portal content as well as for reasoning about users' interests, preferences and contexts. It can add valuable meta information and even lightweight semantics to web resources.

In our previous work [Nauerz *et al.*, 2008] we proposed a framework which allowed arbitrary annotators, e.g. human users or analysis components (for automated tagging), to annotate any of these resources. Analysis of the tagging behavior allowed us to model interests and preferences of users as well as semantic relations between resources, and thus to perform reasonable recommendations and adaptations.

In this paper we present paradigms like alien tagging which allows one user to apply tags for another user, reputation-based tagging allows users expertise to influence tags' weights, quantitative tagging which allows users to manually manipulate tag's weights, anti tagging which allows users to specify "negative tags", tag voting to solve the tag space littering problem by e.g. allowing users to vote against a tag, tag expiry which allows tags to have a lifeime, contextual tagging which allows tags to be associated to certain context profiles, and so forth and describe how these can be used to refine our models and to perform even more valuable adaptations or to issue more valuable. We also allow for mechanisms to follow users tagging "trails" in order to learn from what he is tagging.

2 Related Work

A lot of work is currently underway to experiment with different techniques to improve working with tag engines.

Most researchers try to improve the quality of tags being presented. Tagging systems must often select a subset of tags to be displayed to the user due to limited screen space. Thus they must determine the most valuable ones. [Sen *et al.*, 2009a] present a tag selection algorithm based on users' implicit and explicit (tag rating) behaviour, to select the right tags to be displayed. [Liu *et al.*, 2009] present a tag ranking algorithm used on Flickr[1]. Other approaches for tag selection algorithms are presented in [Sen *et al.*, 2009b] and [Zhang *et al.*, 2009]. Other researchers aim to improve tag quality by recommending and suggesting

[1] http://www.flickr.com

tags. Such approaches are e.g. described in [Suchanek *et al.*, 2008]. [Garg and Weber, 2008a] present a new algorithm called *Hybrid* to determine reasonable tags to be recommended to users on Flickr. [Symeonidis *et al.*, 2008] present a tag recommendation algorithm based on tensor dimensionality reduction. Other tag recommendation work is described in [Garg and Weber, 2008b], [Sigurbjörnsson and van Zwol, 2008], [Song *et al.*, 2008], and [Vig *et al.*, 2009].

Other work goes a little bit more into the direction of what we do and provides users with new means to directly influence tag quality. [Lee and Han, 2007] introduce *QTag* a qualitative tagging system that allows users to tag in order to rate content and express opinions.

Other work similar to ours also experiments with new visualization techniques. [Gwizdka and Bakelaar, 2009] presents a technique for preserving and presenting context and history while navigating web resources described by keywords using tags and tag clouds as application area.

Many researchers also try to improve tag quality by enriching tags with more semantics. [Echarte *et al.*, 2009] introduced methods to group tag variations with matching techniques. Another collaborative Web 3.0 approach was presented by Kreiser et al., which allows users to augment tags with semantics and collaboratively model relations between tags.

In the end quality tags are often used not only to recommend other tags but finally to recommend content. This kind of personalized recommendation of content is based on the content's relatedness to certain tag terms. E.g., [Wu *et al.*, 2006] proposes a modified version of the HITS algorithm to determine experts and high-quality documents related to a given tag.

3 Concepts

3.1 Alien tagging

As said before Web 2.0 communities can be rather heterogeneous. The expertise of users contributing (and consuming) content can vary a lot. What might be obvious for one user might be completely unknown to others. *Alien tagging* allows more experienced users to tag content for less experienced ones. In our prototypical implementation tag widgets allow power users to apply tags to resources on behalf of other users (or even user groups). Next time one of the users for which alien tags have been applied logs-in, he or she is notified about the availability of these and can inspect the underlying resources. The same way we used "normal" tags in our previous work [Nauerz *et al.*, 2008] to refine user models that describe users interests and preferences we can use these alien tags, too. In real environments *alien tagging* could be used e.g. by managers pretagging content for their new hires, by team- or technical leads to point their team members to relevant content which they otherwise might have missed. Thus *alien tagging* opens another opportunity to prevent users from missing out content by issuing recommendations provided by "alien" users. To identify the different kinds of tags in the tag cloud, each kind is encoded in a different color (figure 1). Green tags are tags added and only visible to the user who applied them, the private tags. Blue tags are tags added and visible to the whole community, the public tags. Orange tags are added from one user for another user, the alien tags. An alien tag has two additional icons in the upper right corner, a *plus* sign to transfer the alien tag to the private tag store

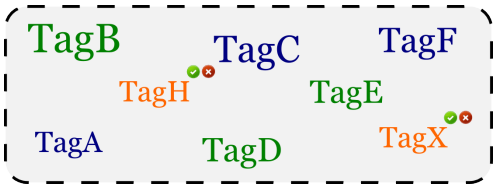

Figure 1: View alien tags in the tag cloud

of a user and the cross sign to discard the alien tag (afterwards, the tag is deleted from the tag cloud). An alien tag is applied by selecting the target user from a drop-down box and specifying the tag name (figure 2).

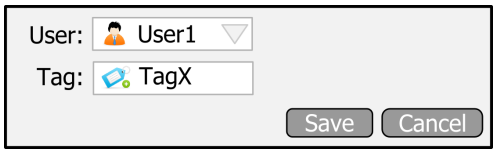

Figure 2: Apply an alien tag

3.2 Tag following

As mentioned previously expertise and interests of users in a community can vary a lot. Therefore it can be interesting for a user to see how other taggers work with the tagging system and in what kind of resources they are interested. If a user thinks that he could benefit from following other taggers, because of overlapping interests with respect to some topics or just to see in what kind of topics a more experienced user is interested in, we provide him with means to follow the tagging "trail" of this user. The user just needs to select the user to follow from a drop-down box (figure 3). Afterwards a notification informing about any newly added public tag of the user being followed pops up, every time the user following logs in (figure 4). The following user can visit the resources and even transfer the tags of the user being followed to his own tag store. It is also possible to follow only certain kinds of tags, e.g. a specific tag bag (see section 3.6), to narrow the focus of the tagging stream.

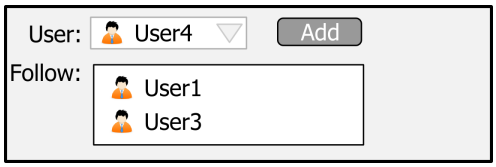

Figure 3: Follow a users tagging stream

3.3 Reputation-based tagging

In our previous solutions we always assumed that the weight (i.e. the importance) of tags only depends on the frequency of their occurrence. I.e. a tag applied more often with respect to a certain scope was regarded of higher importance than a tag applied less often. In our new prototype we additionally assume that the weight of a tag can depend on the reputation (or expertise) of a user. I.e. that tags applied by more experienced users have higher weights, and thus higher influence on what content the community is presented (or recommended) with, than tags from less

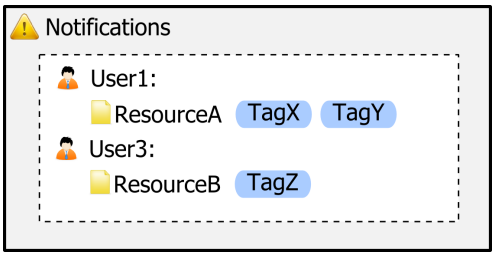

Figure 4: Notification about newly added tags

experienced users. This way we can point users to more relevant content as we assume experts to know better what the community should focus on. E.g., in development team we assume the tagging behavior of the team- or technical lead of higher importance. With *reputation-based tagging* we also ensure that "incorrect or less suited" tags perceive lower weights (influence). E.g., a newbie might apply a more "incorrect/less suited" tag as he just misunderstands (due to his insufficient knowledge) what he is looking at. The way we determine users' expertise has already been described in [Nauerz *et al.*, 2008]. In figure 5 the weight of the tags displayed in the tag cloud only reflect the count of the tag. The magenta colored tags are tags applied by user "UserA" and the cyan colored tags are applied by user "UserB". The tag cloud in figure 7 also considers the reputation level of the user, which applied the tag, in order to calculate tag weights. Therefore we allow users to apply ratings to tags and to users of the community. The reputation level of a user can be determined by, e.g. calculating the median over all ratings applied to the user or over all ratings applied to tags this particular user has applied. Figure 6 shows, that user "UserB" got a better average rating then user "UserA". We see the impact of this difference in both users reputation in the tag cloud in figure 7. Even though, the tag "TagB" is applied as often as the tag "TagF", it is displayed with a lesser weight than tag "TagF", just because of the better reputation level of "UserB" compared to "UserA".

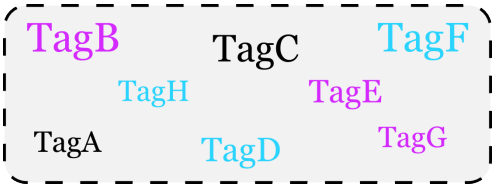

Figure 5: Tag cloud without reputation based weighting

Figure 6: Average rating of user "UserA" and "UserB"

3.4 Quantitative tagging

Previously we also assumed that tags can only have "positive character". I.e. that we assumed that a resource can be tagged with a term to describe that the resource has something to do with this term, but also assumed that a resource

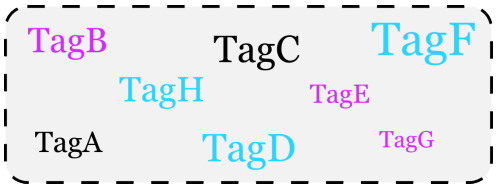

Figure 7: Tag cloud with reputation based weighting

cannot be tagged with a term to describe that the resource has nothing to do with it. In addition to that aspect we did not provide means allowing single users to express that a certain tag is of less relevancy for them. *Quantitative tagging* provides a solution to both problems: in our prototypical implementation a *plus*- and a *minus* sign is presented besides each tag being displayed. In addition, when applying a tag, a *not* sign is available (figure 9). Clicking the *not* sign when applying a tag allows users to express that a resource has nothing to do with the term applied, a helpful feature for more fine-granular categorization of resources: e.g., users could tag some resources with the term *Web 2.0* and a few of them with "not" *scientific*. This helps users to quickly find all Web 2.0 related resources and to quickly distinguish between the scientific and non scientific ones among them. Clicking the *plus*- and *minus* signs when working with tags allows single users to express that they are less interested in a tag (or a certain tag associated to a certain resource) or can additionally express that a tag is of less relevancy for the entire community (figure 8). In the tag cloud an anti tag is displayed red colored and with a not sign in the upper right corner . Thus, these mechanisms allow for further refinements of our user models.

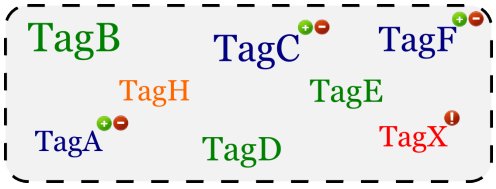

Figure 8: Rate a tag

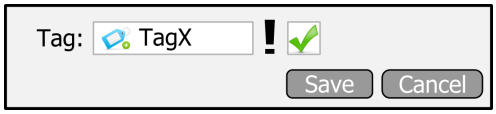

Figure 9: Create a "Not-Tag"

Anti tagging

Anti tagging describes an enhancement to *quantitative tagging* (cp. 3.4). Here we automatically increase or decrease tags' relevancy for the entire community by analyzing tags semantics (cp. [Nauerz *et al.*, 2008]). One option we have evaluated is to take into consideration antonyms. E.g., when a resource is tagged with "good" and "bad" we regard it as not tagged at all with either of these two terms as they annihilate each other. Antonyms can e.g. be found using the antonym thesaurus [2]. As *anti tagging* is not trivial to be realized as most examples are much more complicated

[2] http://www.synonym.com/synonyms/

and less obvious than the one just provided we have not yet incorporated it in our prototype.

Tag voting

Tag voting is a further enhancement to *quantitative tagging* (cp. 3.4). A user can vote in favor of a tag in order to let it become a favorite tag or against a tag. While the user hovers over a tag, a click on the heart icon (figure 10), expresses that he regards the tag as important and should thus be higher weighted than the other tags. The tag is automatically stored in a tag bag called *My Favorites* providing quick access to all favorite tags. A click on the trash icon (figure 10), indicates that the user regards the tag as inappropriate to describe the resource it has been assigned to. We provide another seperate view which is free of the tags the user has voted against. Depending on the system configuration if enough user vote against a tag the tag can be entirely removed from the resource. Thus, voting against tags gives the community the power to correct errors and solve the tag space littering problem autonomously.

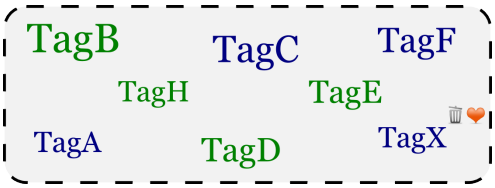

Figure 10: Vote against a tag or favor a tag

3.5 Tag expiry

In our previous work we also assumed that tags can be applied once and stay alive until they are manually deleted again. This let to tag-space littering as most users never deleted tags anymore even if they became obsolete. The fact that tags do not remain valid forever occurs in Portals that provide dynamic content very often. This resulted in having a lot of tags assigned to resources that did not describe the resource adequately nor express the resources relevancy to the community appropriately anymore. In our prototype *tag expiry* allows users to specify a chronological validity for tags when assigning them to a resource. Taggers can give tags a start date, an end date or a time frame in between they live. We also allow tags that are assigned a ”lifetime” to become more (or less) important as time passes by. E.g. if there is a page in the Portal system providing information about the Olympic Games 2012, this page might become more and more interesting to users as we get nearer to the year 2012 and less interesting after 2012. Thus users can specify that the tag should not be available before 2011, vanish after 2013 and become more important from 2011 till 2012 and less important from 2012 till 2013. Thus, *tag expiry* is yet another mechanism to help the community to focus on what is currently really relevant. Moreover, *tag expiry* allows us to neglect ”invalid” tags from being considered when doing content adaptation or recommendation. A clock icon in the upper right corner of a tag in the tag cloud indicates that a lifetime has been applied to the tag (figure 11). If only one user has applied a lifetime to the tag, a tooltip appears during hovering over the tag displaying the dates of the lifetime. Otherwise, the clock icon implies that multiple lifetimes from different users have been applied to the tag. The tag cloud can be filtered by specifying a date or dragging the date slider into

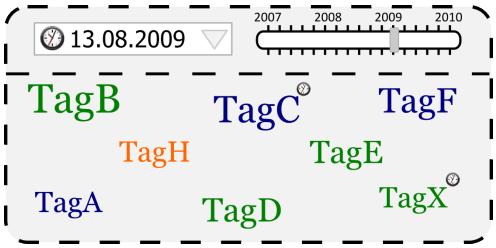

Figure 11: View tags with lifetime in the tag cloud

the past or future. Latter feature allows simulating past or future representations of the tag cloud; i.e. it provides ”filtered” views of the cloud with respect to a certain point in time. To apply a tag with a lifetime, the user selects, using a date picker the start date, the end date or both (figure 12).

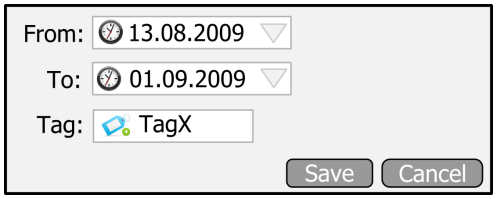

Figure 12: Apply a tag with specific lifetime

3.6 Tagging tags and meta-tagging

Previously we have also worked on solutions to solve major problems of tagging systems: most of these problems discussed dealt with synonyms (multiple tags having the same meaning) and polysemies (a single tag having different meanings). Current tag engines often try to overcome these issues by applying stemming and normalization algorithms which most often only solve problems resulting from morphological variations. Semantical variations can most often not be detected to be a synonym e.g. In our latest prototype we allow the community to resolve the resulting tag-space littering. In our tag-clouds we allow users to drag and drop tags on each other to consolidate them (figure 13). If a tag is dragged onto another tag, the user specify which one of the tags is the representative tag. In addition to that we allow users to create meta-tags (or meta-tag bags as we call them) under which other tags can be organized (figure 14) .Users can create private meta-tag bags only they can see or community meta-tag bags all users part of the community can see. That way users can e.g. create a meta-tag bag ”sports” and drag all sports related tags into that bag; users can also create a meta-tag bag ”favorite-stuff” and just drag what he/she likes most into it.

3.7 Contextual tagging

We also allow for contextual tagging where we can associate tags a certain context (for our context modeling approaches refer to [Nauerz *et al.*, 2008]) to prevent irrelevant tags (irrelevant in a certain context) to appear. This helps focusing on currently relevant content again.

Summarized, a context is described by a context profile which is defined by a set of context attributes permanently observed (e.g. location, day of week, end user device being used to access the system, and so forth). Context management portlets can be used to define new context profiles

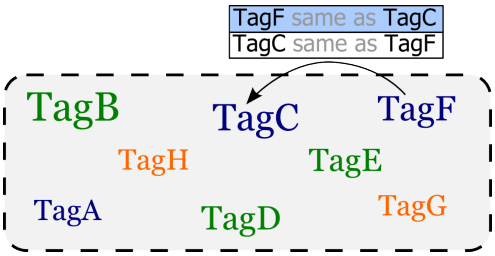

Figure 13: Consolidate two tags

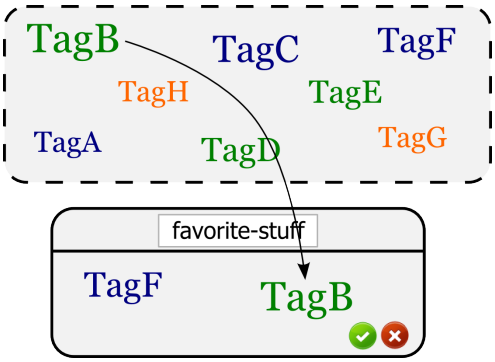

Figure 14: Create a tag bag

or to manually switch between contexts. Alternatively, the system can automatically switch between contexts.

The rationale behind the idea of associating tags to contexts is that some tags are often used or applied in certain contexts only. E.g. some tags might only be applied when doing daily business and working in the office, whereas others might be applied when travling (e.g. tags like "weather_information" or "traffic_information"). Of course tags needed depends on the context, too. So, the tags that have been applied when having been traveling might be irrelevant when being in office.

Thus always displaying all tags part of a tag space is often not reasonable. Hence, we allow for context-senstitive tag widgets, especially tag clouds that only display tags relevant in the current context.

3.8 Other concepts

We are also allowing for tag sharing among subcommunities. Most current tagging systems allow to either create public or private tags but do not allow for a granularity in between. Our prototype allows to share tags with a dedicated set of other users.

The tagging paradigms presented can be combined with one or more of the other paradigms presented. For example, an alien tag can be created which is only valid for a certain period of time.

4 Conclusion and Future Work

In this paper we have presented tagging paradigms which we are using to refine our user- and context modeling approaches presented in our previous work [Nauerz *et al.*, 2008] in order to perform content adaptation and recommendation. The concepts described have already been prototypically implemented and can be presented at the workshop. We have not yet performed in-depth evaluation on these early ideas described in this short paper but are looking forward to discuss them and receive initial feedback. Of course, especially the usefulness of each single concept has still to be evaluated.

For the future we plan to merge our Web 2.0 collaborative tagging approaches with Semantic Web ideas heading towards the Web 3.0.

References

[Echarte *et al.*, 2009] Francisco Echarte, Jos Javier Astrain, Alberto Crdoba, and Jess E. Villadangos. Improving folksonomies quality by syntactic tag variations grouping. In Sung Y. Shin and Sascha Ossowski, editors, *SAC*, pages 1226–1230. ACM, 2009.

[Garg and Weber, 2008a] Nikhil Garg and Ingmar Weber. Personalized, interactive tag recommendation for flickr. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 67–74, New York, NY, USA, 2008. ACM.

[Garg and Weber, 2008b] Nikhil Garg and Ingmar Weber. Personalized tag suggestion for flickr. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 1063–1064, New York, NY, USA, 2008. ACM.

[Gwizdka and Bakelaar, 2009] Jacek Gwizdka and Philip Bakelaar. Tag trails: navigation with context and history. In *CHI EA '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 4579–4584, New York, NY, USA, 2009. ACM.

[Lee and Han, 2007] Sung Eob Lee and Steve SanKi Han. Qtag: introducing the qualitative tagging system. In *HT '07: Proceedings of the eighteenth conference on Hypertext and hypermedia*, pages 35–36, New York, NY, USA, 2007. ACM.

[Liu *et al.*, 2009] Dong Liu, Xian-Sheng Hua, Linjun Yang, Meng Wang, and Hong-Jiang Zhang. Tag ranking. In *WWW '09: Proceedings of the 18th international conference on World wide web*, pages 351–360, New York, NY, USA, 2009. ACM.

[Nauerz *et al.*, 2008] Andreas Nauerz, Stefan Pietschmann, and Rene Pietzsch. Social recommendation and adaptation in web portals. In *Proceedings of the International Workshop on Adaptation for the Social Web*, Hannover, Germany, 2008.

[Sen *et al.*, 2009a] Shilad Sen, Jesse Vig, and John Riedl. Learning to recognize valuable tags. In Cristina Conati, Mathias Bauer, Nuria Oliver, and Daniel S. Weld, editors, *IUI*, pages 87–96. ACM, 2009.

[Sen *et al.*, 2009b] Shilad Sen, Jesse Vig, and John Riedl. Learning to recognize valuable tags. In *IUI '09: Proceedings of the 13th international conference on Intelligent user interfaces*, pages 87–96, New York, NY, USA, 2009. ACM.

[Sigurbjörnsson and van Zwol, 2008] Börkur Sigurbjörnsson and Roelof van Zwol. Flickr tag recommendation based on collective knowledge. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 327–336, New York, NY, USA, 2008. ACM.

[Song *et al.*, 2008] Yang Song, Ziming Zhuang, Huajing Li, Qiankun Zhao, Jia Li, Wang-Chien Lee, and C. Lee Giles. Real-time automatic tag recommendation. In *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 515–522, New York, NY, USA, 2008. ACM.

[Suchanek *et al.*, 2008] Fabian M. Suchanek, Milan Vojnovic, and Dinan Gunawardena. Social tags: meaning and suggestions. In *CIKM '08: Proceeding of the 17th ACM conference*

on Information and knowledge management, pages 223–232, New York, NY, USA, 2008. ACM.

[Symeonidis *et al.*, 2008] Panagiotis Symeonidis, Alexandros Nanopoulos, and Yannis Manolopoulos. Tag recommendations based on tensor dimensionality reduction. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 43–50, New York, NY, USA, 2008. ACM.

[Vig *et al.*, 2009] Jesse Vig, Shilad Sen, and John Riedl. Tagsplanations: explaining recommendations using tags. In *IUI '09: Proceedings of the 13th international conference on Intelligent user interfaces*, pages 47–56, New York, NY, USA, 2009. ACM.

[Wu *et al.*, 2006] Harris Wu, Mohammad Zubair, and Kurt Maly. Harvesting social knowledge from folksonomies. In *Proc. of the 17th Conf. on Hypertext and hypermedia*, pages 111–114, New York, NY, USA, 2006. ACM Press.

[Zhang *et al.*, 2009] Shaoke Zhang, Umer Farooq, and John M. Carroll. Enhancing information scent: identifying and recommending quality tags. In *GROUP '09: Proceedings of the ACM 2009 international conference on Supporting group work*, pages 1–10, New York, NY, USA, 2009. ACM.