Fatih Gedikli, Mouzhi Ge and Dietmar Jannach

Explaining Online Recommendations Using Personalized Tag Clouds

Erklärung von Empfehlungen mithilfe von Schlagwortwolken

Empfehlungssysteme_Erklärungen_Visualisierung_Personalisierung

Zusammenfassung. Empfehlungssystem sind verkaufsunterstützende Anwendungen, welche gewöhnlich in Webshops integriert sind und den Konsumenten auf Produkte und Dienstleistungen hinweisen, an denen er oder sie interessiert sein könnte, aber noch nicht gekauft hat. In den letzten Jahren wurden zahlreiche Algorithmen zur Erhöhung der Vorhersagegenauigkeit solcher Systeme entwickelt. Es gibt jedoch auch andere Faktoren, die die durch den Benutzer empfundene Qualität des Systems beeinflussen. Es hat sich insbesondere gezeigt, dass systemseitig erzeugte Erklärungen, warum ein bestimmtes Produkt vorgeschlagen wurde, besonders dazu geeignet sind, die Zufriedenheit des Benutzers mit dem System sowie die Effizienz des Empfehlungsprozesses zu steigern. In diesem Aufsatz berichten wir von den Ergebnissen einer ersten Studie, in welcher untersucht wurde, inwiefern Schlagwortwolken ein geeignetes Mittel sind, um Empfehlungen visuell zu erklären. Die Studie zeigt, dass mithilfe von Schlagwortwolken sowohl die Zufriedenheit des Benutzers mit dem System als auch die Empfehlungseffizienz im Vergleich zu anderen existierenden schlagwort-basierten Verfahren messbar gesteigert werden kann.

Summary. Recommender systems are sales-supporting applications that are usually integrated into online shops and are designed to point the visitor to products or services she or he might be interested in but has not bought yet. In the last decade, many techniques have been developed to improve the predictive accuracy of such systems. However, there are also factors other than accuracy that influence the user-perceived quality of such a system. In particular, system-generated explanations as to why a certain item has been recommended have shown to be a valuable tool to improve both the user's satisfaction and the system's efficiency. This paper reports the results of a first user study which was conducted to evaluate whether personalized tag clouds are an appropriate means to visually explain recommendations. The evaluation reveals that using tag clouds as explanation mechanism leads to higher user satisfaction and recommendation efficiency than previous keyword-style explanations.

1. Introduction

On most online shopping sites and e-commerce platforms a sub-area of the shop's user interface is used to point the visitor to potentially interesting items in the shop that he or she might have not bought yet. Instead of simply relying on static lists that often contain top-selling or currently price-reduced products, more and more platform providers now aim to exploit the potential of providing personalized shopping recommendations. Amazon.com probably was one of the first large online retailers who successfully relied on recommender systems (RS) technology on a large scale to boost sales ("Customers who bought this item also bought"). Since then, RS have been applied in a variety of domains and different studies demonstrated the business value of personalized sales recommendations, see for example (Jannach and Hegelich, 2009).

Since the mid-1990s, the research community has for a long time focused on improving the predictive accuracy of recommender systems only, that is, the degree to which the system is capable of predicting the degree of how much a user will like an item. However, it soon became evident that precise predictions are not enough. Recommending Rocky II to someone who liked Rocky I might be highly precise, but probably not valuable to the customer, because the recommendation is obvious. Therefore, other factors that influence the user-perceived quality of an RS such as list diversity or serendipity moved into the research focus, see for example (Ziegler et al., 2005)

DOI 10.1524/icom.2011.0002

or (McNee et al., 2006). Beside these aspects, in particular the system's capability of providing explanations as to why a certain item has been recommended has been identified as a valuable instrument to increase the quality of an RS in different dimensions (Bilgic and Mooney, 2005; Sinha and Swearhingen, 2002; Symeonidis et al., 2009; Tintarev and Masthoff, 2007a/2007b). System-side explanations can for example help to increase the user's trust in the system when the user can view a justification of the system's recommendations. Beyond that, explanations can also help the user to make decisions more quickly and thus increase the efficiency of the overall sales process.

One of the main problems of explaining recommendations is that the reasons why a certain item is included in a recommendation list can be rather complex and for example be the result of some machine learning process. In the past, several methods have been proposed for generating user-understandable explanations based on different visualization approaches, see (Herlocker et al., 2000) for collaborative filtering RS or (McSherry, 2005; Pu and Chen, 2006) for case-based reasoning RS. Recently, Vig et al. explored how user-contributed tags can be exploited in the explanation process (Vig et al., 2009). Their online study revealed that so-called „tagsplanations" can for example help to improve an RS's effectiveness.

In Vig et al.'s work, the items' tags and their relevance are displayed in tabular form in the explanation process. In our work, however, we hypothesized that tag clouds are a more effective way of visualizing explanations and conducted a first corresponding user study in which we contrast our approach with previous keyword-style explanation techniques.

The paper is organized as follows. In the next section, we review the different goals, trade-offs and existing works in explanation in RS. After that, we describe our approach, the experimental setup and the results of our user study. The paper ends with a summary and a short outlook on future work.

1.1 Background

The capability of intelligent systems to explain their reasoning and problem solving strategy to their users has been considered as one of the important and valuable features already in early knowledge-based systems (Berry and Broadbent, 1987). With the help of explanations, knowledge-based systems (KBS) have shown to be able to increase order accuracy, produce more credible predictions or improve user satisfaction (Wanninger, 1998; Doyle et al., 2003). Furthermore, an explanation facility enabled KBS to provide understandable accountable decision support in a various domains such as financial and medical industry (Rowe and Wright, 1993; Ong et al., 1997). In recent years, the concept of explanations has also been investigated and adopted in the area of RS. In that context, the explanation facility of an RS is for example used to expose the reasoning behind a recommendation (Herlocker et al., 2000) or enable more advanced communication patterns between a selling agent and a buying agent (Jannach et al., 2010).

In order to better understand the different aspects of the concept of explanation, different classification approaches have been proposed. Chandrasekaran et al. (1989) for example identified three top-level functions for the explanation generation problem: basic content, responsiveness and human-computer interface. Generating the basic content amounts to selecting the appropriate information and produce a justification that is independent of user's decision process. Based on the basic content, responsiveness means organizing and shaping the explanation content to match the user's knowledge. The final problem is to effectively present and display the information to the user in an appropriate way. Gregor and Benbasat (1999) later on adapt this taxonomy and structure the explanation problem into content, presentation format and provision mechanism. Regarding the content, they further distinguish between reasoning, justification, control and terminology. The presentation format is detailed as text-based or multimedia; the provision mechanism can be user-invoked, automatic and intelligent. A user-invoked mechanism means that the explanations are provided only upon request of the user. As an example, consider the travel recommender system based on case-based reasoning technology from (McSherry, 2005). In this system the explanations contain reasoning and justification aspects, are user-invoked and presented in a text-based form. Automatic mechanisms, in contrast, are out of the user control and provide the explanations also without being explicitly requested. Herlocker et al. (2000), for example, propose 19 types of explanations for collaborative filtering RS. Their explanations can be classified as being related to reasoning, justification and terminology; the explanations are automatically presented using both text-based and visual representations. Beyond that, even more intelligent mechanism can be used to provide personalized explanations. For example, Vig et al. (2009) proposes a tag-based explanation approach that intelligently presents the explanations in a text-based form. In this paper, we aim to further evolve the tag-based explanation scheme and present explanations in the form of tag clouds, which combine the text-based representation with a visual effect.

Regarding the purpose of explanation in RS, an explanation can be considered as a piece of information presented in a communication process which can serve different goals (Jannach et al., 2010). Tintarev and Masthoff (2007) conduct a systematic review on the goals of providing explanations in RS. They identify seven factors which are transparency (explaining why a particular recommendation is made), scrutability (allowing interaction between user and system), trust (increasing the user's confidence in the system), effectiveness (helping the users make better decisions), persuasiveness (changing the user's buying behavior), efficiency (reducing the time used to complete a task) and satisfaction. In this paper, we focus on two goals: satisfaction and efficiency. Satisfaction refers (a) to the extent of how the presented explanation helps the users to assess the quality of a recommended item and (b) the extent of how users find the explanations to be helpful and the experience enjoyable. One direct measurement approach is simply to ask if users are satisfied with the explanations. Additionally, Bilgic and Mooney (2005) argue that satisfaction is more important than persuasiveness in the long run as greater satisfaction can help to establish trust in the RS and attract further users. Efficiency, on the other hand refers to the ability of an explanation to help decreasing the user's decision-making effort. One possi-

	Taxonomy of explanation	Classification of RS	Main goals to analyze
Herlocker et al. (2000)	Reasoning, terminology, text-based, automatic	Collaborative filtering	Persuasiveness
Bilgic and Mooney (2005)	Reasoning, terminology, text-based, automatic	Hybrid	Effectiveness Persuasiveness
McSherry (2005)	Reasoning, justification, text-based, user-involved	Knowledge-based	Effectiveness
Pu and Chen (2006)	Reasoning, text-based, automatic	Collaborative filtering	Trust Efficiency
Cramer et al. (2007)	Reasoning, text-based, automatic	Content-based	Trust Persuasiveness
Vig et al. (2009)	Reasoning, terminology, tag-based, intelligent	Hybrid	Effectiveness
Our paper	Reasoning, terminology, tag clouds, intelligent	Hybrid	Efficiency Satisfaction

Tab. 1. A review of related works studying explanation in RS.

ble evaluation approach is to measure the time used to complete the same task with and without an explanation facility or with different types of explanations and compare the time difference between the two or several scenarios. In their user study, Pu and Chen (2006) for example provided two explanation interfaces to users and compared the time needed to locate a desired item using the different interfaces.

Note that in Tintarev and Masthoff's literature review, effectiveness and efficiency are the evaluation factors that most literature (50%) focuses on. One possible reason why the two factors are frequently used is that they are crucial in evaluating explanations in RS and that they are also easy to manipulate. Additionally, one advantage of using effectiveness and efficiency as dependent variables is that there is limited correlation between the two factors. That means that it is possible to find a type of explanation that is both effective and efficient at the same time.

Table 1 summarizes our short review and categorizes our work and previous works in explanation in RS along different dimensions: the taxonomy of explanations, the type of the RS and the goal of the explanation. As for the taxonomy of explanations, we mainly rely on the work of Gregor and Benbasat (1999). Regarding the type of the RS, we use the typical categorization in literature: collaborative filtering, content-based and knowledge-based approaches as well as hybrid systems. With respect to the goals of explanation, we use the classification of (Tintarev and Masthoff, 2007). Regarding the taxonomy, our work is most similar to the one reported in (Vig et al., 2009). The main differences are the presentation format (tag-based vs. tag clouds) and the goals that have been analyzed in the study.

Overall, research on explanation in RS requires us to build on and further develop existing work from different areas such as intelligent systems, human-computer interaction and information systems. Considering the different views from the related communities, we consider tag-based explanation as a promising way to improve the performance of RS. In this paper, therefore we extend the work of Vig et al. (2009) and aim to provide an innovative and personalized user interface to achieve higher user satisfaction and efficiency.

2. Explanation Interfaces

In our work we aim to evaluate whether (personalized) tag clouds are an appropriate means for explaining recommendations in RS. We therefore conducted a study, in which we compared three different explanation interfaces: keyword style explanations (KSE), tag clouds (TC), and personalized tag clouds (PTC). We use keyword-style explanations (KSE) as a baseline because this visualization approach performed the best in the study by Bilgic and Mooney (2005). The new methods TC and PTC use user-contributed tagging data for explaining the recommendations; the KSE approach relies on keywords which are automatically extracted from item descriptions. In the following, we will discuss of the three explanation interfaces in more detail.

Keyword Style Explanations (KSE)

An example of the KSE interface is shown in Figure 1. The interface consists of an ordered list of 20 keywords extracted from the movie descriptions, which are assumed to be the most important one for the user ("BUSCEMI", "POLICE", etc.). The importance or strength of a keyword k is determined by the following formula: strength(k) = t * userStrength(k), where t corresponds to the number of times the keyword appears in the movie's content description and userStrength(k) measures the target user's affinity towards the given keyword, which is basically computed by measuring the odd ratios P(k | positive classification) / P(k | negative classification) for a given user, i.e., how much more likely a keyword will appear in a positively rated example than in a negatively rated one. The probabilities are estimated using a naïve Bayesian text classifier. Internally, a movie's content description is based on five different slots. Each slot consists of a "bag of words" containing an unordered

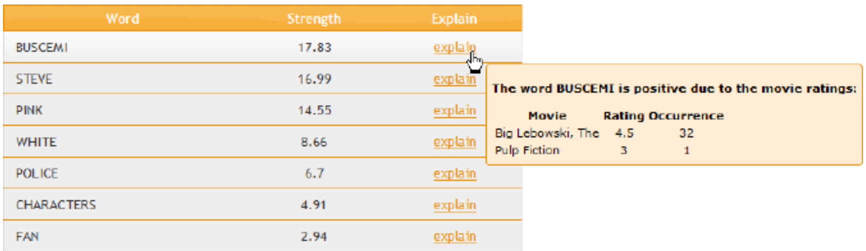

Fig. 1. Keyword style explanation (KSE).

set of words together with their frequencies. In our study in the movie domain, we considered the slots director, actors, genre, description and related-titles. The data about directors, actors, genres and related titles was taken from the IMDb website (1) and the MovieLens data set (2). For the movie description slot we considered all available movie reviews by crawling Amazon.com (3) as well as synopsis information collected from Amazon, Wikipedia (4) and moviepilot (5).

Beside the list of important keywords, the KSE explanation interface features a link („Explain") for each keyword that opens a pop-up window containing more information. The popup window shows all the movies that the user has rated that contain the respective keyword. The user is presented both with his rating for the movie and the number of times the keyword appears in the content description.

Note that in (Bilgic and Mooney, 2005), the KSE approach performed best in the book domain with respect to effectiveness (enabling users to make good decisions). However, the evaluation of efficiency (enabling users to make fast decisions) and satisfaction (the extent to which users enjoy explanations) was not part of their work but will be analyzed in our study.

Tag Clouds (TC)

Tag clouds as shown in Figure 2 have become a frequently used visualization and interaction technique on the Web. They can be often found on Social Web platforms such as Delicious (6) and Flickr (7) and are used to visually present a set of words or user-generated tags. In such tag clouds, the font size, weight and the color of tags is varied according to the relevancy or frequency of a keyword or tag. Additionally, the position of tags can be automatically adjusted based on some heuristics, but usually the tags are sorted alphabetically from the upper left corner to the lower right corner.

In our basic approach of using tag clouds as a not-yet-explored means to explain recommendations, we simply used the number of times a tag was attached to a movie as a metric of its importance

Fig. 2. Tag cloud (TC).

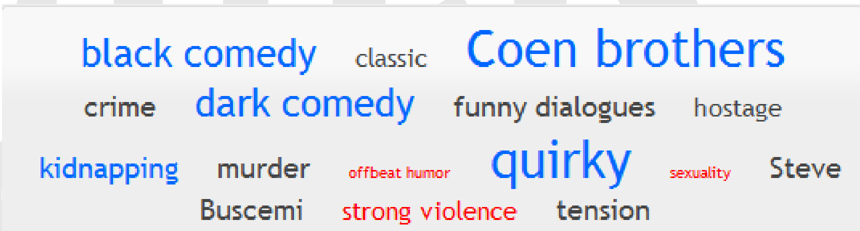

Fig. 3. Personalized tag cloud (PTC).

assuming that a keyword that is often used by the community is suited to characterize its main aspects. When a user clicks on a recommended item, we display the tag cloud of the movie as shown in Figure 2. Tags such as „black comedy" or „quirky" have been used by many people and are thus displayed in a larger font size. Tag positions and font colors are not varied in this visualization approach, although these attributes possibly have an additional effect on the user's perception on the explanation interface, which could be considered in future studies.

Personalized Tag Clouds (PTC)

The PTC explanation interface is an extension to the basic tag cloud interface presented above. It provides more information by using additional "tag rating data" which was reported in Gedikli and Jannach (2010) as an additional knowledge source for recommender systems. In Gedikli and Jannach (2010) the authors present a recommendation approach, in which users rate items by rating their attached tags. While the general idea of "tag preferences" was also reported in Vig et al. (2009) the novel idea consists in allowing users to rate tags in the context of an item. The intuition behind this idea is that the same tag may have a positive connotation for the user in one context and a negative in another. For example, a user might like action movies featuring the actor Bruce Willis, but at the same time this user might dislike the performance of Bruce Willis in romantic movies. In (Gedikli and Jannach, 2010) the authors show that the predictive accuracy of recommender algorithms can be improved when incorporating such user- and item-specific tag rating data. In the PTC explanation interface, we pick up on this idea but aim to use the tag rating data to improve the quality of explanations for recommendations. An example of the PTC interface for a comedy movie is shown in Figure 3. In contrast to the TC interface,

1 http://www.imdb.com
2 http://www.grouplens.org/node/73
3 http://www.amazon.com
4 http://www.wikipedia.org
5 http://www.moviepilot.de
6 http://www.del.icio.us
7 http://www.flickr.com

in the PTC approach, we vary the color of the tags according to the user's affect attached to the tag. For example, in our study, blue colored tags are used to highlight aspects of the movie toward which the user has a positive feeling. Tags with a negative connotation are shown in red; tags, for which no particular preference is known, are shown in black. Similar to the TC approach, the font size is used to visualize the importance or quality of a tag. In order to determine the positive or negative feeling attached to a tag, we analyze the tag rating distribution of the target user's nearest neighbors in order to decide whether the target user will like, dislike or feel neutral about the item features represented by these tags.

2.1 Experimental Setup / Procedure

We have conducted a within-subjects user study, in which each subject was confronted with all explanation interfaces presented above. A total of 19 subjects have participated in our experiment. During the experiment we observed the subjects while performing their tasks. Our evaluation procedure extends the procedure proposed in (Bilgic and Mooney, 2005) and had two parts. In the first part, preference information about users and tags was gathered and a user-profile was built. In the second part, which was executed a few weeks after the first session, the subjects used an RS which presented them with item proposals based on the data collected in the first part. In addition, the different explanation interfaces are shown to the user.

Experiment – Part 1

In the period between 11/22/2010 and 12/10/2010 we collected sample movie ratings and tag ratings from the participants, who were asked to rate at least 15 out of 100 movies. We have limited the number of movies to 100 in order to be able to find nearest neighbors in the PTC approach. When the user rates a movie, a screen appears (Figure 4) in which the tags of the movie are shown. On this screen the user can rate up to 15 tags of the movie. The tags were taken from the "MovieLens 10M Ratings, 100k Tags" data set ([8]). Users could rate an arbitrary number of tags (we have not asked users to rate a certain number of tags); skip tags, in case they thought that they are not suitable for a given movie; or explicitly mark tags as inappropriate for rating. Note that in the experiment the users were not allowed to apply their own tags. We made this decision in order to ensure that we have a reasonable overlap in the used tags given the relatively small number of participants.

Experiment – Part 2

The collected rating data served as a basis for recommendations and explanations in the second part of our experiment, which was conducted between 12/11/2010 and 01/20/2011. In the second part, we used a classical user-based collaborative filtering algorithm to generate a set R of movie recommendations for each participant. Then, the following procedure was followed, see also (Bilgic and Mooney, 2005).

1. R = Set of recommendations for the user.
2. E = Set of explanation interfaces {KSE, TC, PTC}.
3. For each randomly chosen (r, e) in R x E do:
4. Present explanation using interface e for recommendation r to the user.
5. Ask the user to rate r and measure the time taken by the user.
6. For each recommendation r in R do:
7. Show detailed information about r and ask the user to rate r again.
8. Ask the user to rate the explanation interfaces.

Instead of displaying the movie itself, the system randomly picked one of the recommendations and one of the possible explanation styles and presented the user with the explanation for the movie. We randomized the selection process for the recommendations and interfaces in order to minimize the effect of seeing recommendations or interfaces in a special order. Next, the user was asked to rate the recommended movie by solely relying on the presented explanation for the recommendation, i.e., the title of the movie was hidden. If the users thought that they have recognized one of the recommended movies, they could inform the system about this fact and the rating for this movie/interface combination was consequently not taken into account. We additionally measured the time it took the user to submit a rating as to measure the efficiency of the user interface. Figure 5 shows an example of the TC interface with the movie title hidden.

After these steps had been completed for all recommended items, we presented the recommendations again to the user, this time showing the complete movie title and links to the corresponding movie information pages at Wikipedia, Amazon and IMDb, see Figure 6. Users were instructed to read the detailed information about the recommended movies and then asked to rate the movies again. At the end of the experiment, the users could give feedback on the different explanation interfaces (as to measure satisfaction with the system) by rating the system as a whole on a 0.5 (lowest) to 5 (highest) rating scale. Again, we randomized the order to account for biasing effects.

2.2 Hypotheses, Results and Discussion

According to Bilgic and Mooney (2005), an explanation that minimizes the difference between the ratings based on the explanation only and the rating based on more knowledge is desirable as it increases the perceived effectiveness of the explanation interface. In case the rating based on the explanation interface is higher than the "informed" rating, the explanation presented causes the user to overestimate his or her own informed rating of an item, which is equivalent to a persuasive explanation. In the following section we will report and discuss the results regarding efficiency and satisfaction of the different explanation interfaces TC, PTC and KSE.

We tested two hypotheses. First, we hypothesized that users make decisions faster when using the tag cloud interfaces TC and PTC (H1: Efficiency). We believe this as we think that the visual nature of a tag cloud allows the user to grasp the content information inside a cloud more quickly. In the KSE approach, in contrast, the explanatory information is organized in a tabular view with same-size table entries and a strength-field, which has to be

8 http://www.grouplens.org/node/73

Fig. 4. Rating tags of a given movie on a Likert scale of 0.5 to 5.

Fig. 5. Rating movies by solely relying on the explanation on a Likert scale of 0.5 to 5 (step 5).

Fig. 6. Rating movies after acquiring detailed information about the movies recommended before on a Likert scale of 0.5 to 5 (step 7).

interpreted by the user first. On the other hand, the KSE approach provides more detailed information about the movies that influenced the strength of a keyword, i.e., the user's affinity towards the given keyword. Due to higher complexity of the KSE approach and the way the information is presented there, we however conjectured that tag clouds can help users to decide faster. We further assumed that users enjoy explanations in the form of a tag cloud or personalized tag cloud more than in the KSE style as we assumed that tag cloud explanations are easier to interpret for the end user.

Efficiency

To test our hypothesis of improved efficiency of tag clouds, we analyzed the time measurement data which was automatically collected during the second part of the experiments. Table 2 shows the mean times (in seconds) for submitting a rating after seeing the corresponding explanation interface. We have run the Friedman test in conjunction with a post-hoc Nemenyi test in order to decide whether the reported differences are significant or occurred by chance.

We can see in Table 2 that the time period for the tag cloud approaches is significantly shorter than for KSE. Thus, we can conclude that the data supports hypothesis H1 at a significance level of $\alpha = 0.05$. The data in Table 2 also indicates that the PTC method helps users to make decisions slightly faster than the TC approach, but the difference was not statistically significant.

	KSE	TC	PTC
Mean time [sec]	30.72	**13.53**	**10.66**
Standard deviation	19.72	8.52	5.44

Tab. 2. Mean time for submitting a rating. Bold figures indicate explanations with a mean time that is significantly different from the base cases (N = 60, α = 0.05, Friedman test with a post-hoc Nemenyi test).

	KSE	TC	PTC
Mean rating	1.87	**3.74**	**3.87**
Standard deviation	0.90	0.65	0.62

Tab. 3. Mean response of 19 users to each explanation interface based on a Likert scale of 0.5 to 5. Bold figures indicate explanations with a mean rating significantly different from the base cases (N = 19, α = 0.05, Friedman test with a post-hoc Nemenyi test).

Satisfaction

Table 3 shows that users prefer the PTC approach over the TC presentation style and the TC style over the KSE method, which supports hypothesis H2. Again, the differences between the keyword-style explanations and the tag cloud interfaces TC and PTC are significant ($\alpha = 0.05$) but no significant difference among the tag cloud interfaces could be found although the data indicates that users favor PTC-style explanations. One possible reason is that tag clouds are in general capable of visualizing the context in a concise manner and can thus help users to reduce the time needed to understand the context which in turn increases user satisfaction. As no significant difference is found between using TC and PTC, one possible explanation can be that instead of paying much attention to the color of the tags, some users may have directly made their judgment only based on the content and size of the tags.

3. Summary and Outlook

In this paper, we introduced tag clouds as an explanation interface to recommender systems and have shown based on a first user study that visualizing explanations of recommendations based on this well-known Web 2.0 concept can help to increase both the users' satisfaction with the system as well as the systems efficiency measured in the time needed by users to make a decision. In practice, we see this as a further step to build more efficient and effective recommender systems in the future.

In detail, our results show that users prefer tag cloud interfaces over keyword-style explanations. We found this fact somewhat surprising as users preferred even the non-personalized explanation interface TC over the personalized KSE interface. We assume that there are factors other than personalization such as the graphical representation, which play a crucial role for effective explanation interfaces. Our experiment also revealed that users need less time to come to a conclusion when they are confronted with a tag cloud explanation interface.

Our future work includes an analysis of further quality dimensions of explanations such as effectiveness and persuasiveness. We also aim to analyze in more detail, whether varying tag cloud attributes such as tag position or font color influences the effectiveness of explanations. Finally, we plan to conduct a larger user study in order to find out whether there are significant differences between the TC and PTC approaches.

References

Adomavicius, G.; Kwon, Y.: New Recommendation Techniques for Multicriteria Rating Systems, IEEE Intelligent Systems 22(3) (2007) 48–55.

Adomavicius, G; Tuzhilin, A.: Extending Recommender Systems: A Multidimensional Approach, Proceedings of the Workshop on Intelligent Techniques for Web Personalization, Acapulco, Mexico, 4(6) (2001).

Berry, D. C.; Broadbent, D. E.: Explanation and Verbalization in a Computer-assisted Search Task, Quarterly Journal of Experimental Psychology 39(4) (1987) 585–609.

Bilgic, M.; Mooney, R. J.: Explaining Recommendations: Satisfaction vs. Promotion, Proceedings Beyond Personalization 2005: A Workshop on the Next Stage of Recommender Systems Research at the 2005 International Conference on Intelligent User Interfaces, San Diego, CA (2005) 13–18

Chandrasekaran, B. ; Tanner, M. C. ; Josephson, J. R.: Explaining Control Strategies in Problem Solving, IEEE Expert: Spring 1989.

Cramer, H.; Evers, V. ; Ramlal, S.; Someren, M, Rutledge; L., Stash, N.; Aroyo, L.; Wielinga, B.: The Effects of Transparency on Trust in and Acceptance of a Content-based Art Recommender, User Modeling and User-Adapted Interaction 18 (5) (2008) 455–496.

Doyle, D.; Tsymbal, A.; Cunningham, P.: A Review of Explanation and Explanation in Case-based Reasoning, Computer Science Technical Report, Trinity College Dublin, 2003.

Gregor, S.; Benbasat I.: Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice. MIS Quarterly 23(4) (1999) 497–530.

Gedikli, F.; Jannach, D.: Rating Items by Rating Tags, Proceedings of 2nd Workshop on Recommender Systems and the Social Web at ACM RecSys'2010, Barcelona, Spain (2010) 25–32.

Herlocker, J. L; Konstan, J. A., Riedl, J.: Explaining Collaborative Filtering Recommendations. Proceedings ACM Conference on Computer Supported Cooperative Work. Philadelphia, PA, USA (2000) 241–250.

Jannach, D.; Hegelich K.: A Case Study on the Effectiveness of Recommendations in the Mobile Internet, Proceedings of the ACM Conference on Recommender Systems, New York (2009) 205–208.

Jannach, D.; Zanker, M; Felfernig, A.; Friedrich, G.: Recommender Systems – An Introduction: Cambridge University Press, 2010.

McNee, S.M.; Riedl, J.; Konstan, J.A.: Being Accurate is Not Enough: How Accuracy Metrics Have Hurt Recommender Systems. In CHI '06 extended abstracts on Human factors in Computing Systems: Montreal, Canada, 2006.

Mcsherry, D.: Explanation in Recommender Systems. Artificial Intelligence Review 24(2) (2005) 179–197.

Ong, L.S.; Sheperd, B.; Tong, L.C.; Seow-Choen, F. ; Ho Y.H.; Tong L.C. ; Ho Y.S; Tan K.: The Colorectal Cancer Recurrence Support System, Artificial Intelligence in Medicine 11(3) (1997) 175–188.

Pu, P.; Chen, L.: Trust Building with Explanation Interfaces. Proceedings of the 11th International Conference on Intelligent User Interfaces, Sydney, Australia, (2006) 93–100.

Rowe, G.; Wright, G.: Expert Systems in Insurance: A Review and Analysis, Intelligent Systems in Accounting, Finance and Management 2(2) (1993) 129–145.

Sinha, R.; Swearingen, K.: The Role of Transparency in Recommender Systems. In CHI ,02 extended abstracts on Human Factors in Computing Systems, (2002) 830–831.

Symeonidis, P.; Nanopoulos, A.; Manolopoulos, Y.: MoviExplain: A Recommender System with Explanations. Proceedings of 3rd ACM Conference on Recommender systems, New York, NY, USA, (2009) 317–320.

Tintarev, N.; Masthoff, J.: Effective Explanations of Recommendations: User-centered Design. Proceedings of the 2007 ACM Conference on Recommender systems, Lausanne, Switzerland, (2007a) 153–156.

Tintarev, N.; Judith Masthoff, J.: A Survey of Explanations in Recommender Systems. Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop, Istanbul, Turkey, (2007b) 801–810.

Vig, J.; Sen, S.; Riedl, J.: Tagsplanations: Explaining Recommendations Using Tags. Proceedings of the 13th Onternational Conference on Intelligent User Interfaces. Sanibel Island, USA (2009) 47–56.

Wanninger, L.: Profitable Electronic Commerce: Framework, Examples, Trends, Proceedings of the Eleventh International Bled Electronic Commerce Conference „Electronic Commerce in the Information Society", Slovenia (1998) 3–27.

Ziegler, C. N.; McNee, S.M; Konstan, J. A.; Lausen, G.: Improving Recommendation Lists through Topic Diversification. Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, (2005) 22–32.


1


2


3

1 Dipl.-Inf. Fatih Gedikli, Researcher, Computer Science Department, TU Dortmund, Germany. Main research area: Recommender Systems.
e-mail: fatih.gedikli@tu-dortmund.de

2 Dr. Mouzhi Ge, Researcher, Computer Science Department, TU Dortmund, Germany. Main research area: Information System Quality.
e-mail: mouzhi.ge@tu-dortmund.de

3 Prof. Dr. Dietmar Jannach, Professor, Computer Science Department, TU Dortmund, Germany. Main research areas: Recommender Systems, Intelligent Systems in Business.
e-mail: dietmar.jannach@tu-dortmund.de