What Should I Read Next?
A Personalized Visual Publication Recommender System

Simon Bruns[1], André Calero Valdez[1], Christoph Greven[2], Martina Ziefle[1], and Ulrik Schroeder[2]

[1] Human-Computer Interaction Center, Campus Boulevard 57, RWTH Aachen University, Germany
`{bruns,calero-valdez,ziefle}@comm.rwth-aachen.de`
[2] Learning Technologies Research Group, Ahornstr. 55, RWTH Aachen University, Germany
`{greven,schroeder}@cs.rwth-aachen.de`

Abstract. Discovering relevant publications for researchers is a non-trivial task. Recommender systems can reduce the effort required to find relevant publications. We suggest using a visualization- and user-centered interaction model to achieve both a more trusted recommender system and a system to understand a whole research field. In a graph-based visualization papers are aligned with their keywords according to the relevance of the keywords. Relevance is determined using text-mining approaches. By letting the user control relevance thresholds for individual keywords we have designed a recommender system that scores high in accuracy ($\bar{x} = 5.03/6$), trust ($\bar{x} = 4.31/6$) and usability (SUS $\bar{x} = 4.89/6$) in a user study, while at the same time providing additional information about the field as a whole. As a result, the inherent trust issues conventional recommendation systems have seem to be less significant when using our solution.

Keywords: Recommender Systems, Visualization, User-Study, Trust, Usability

1 Introduction

The rapid growth of information on the world wide web enables users to get information from multiple sources. However, it can still be difficult to evaluate which sources are trustworthy [1]. This phenomenon is not only limited to standard users, but also influences specialists such as researchers. Researchers are required to be up to date on any recent technology or research that is relevant to their own work. The main source to find information on current research projects are scientific publications. However, this is getting more and more difficult, due to the sheer amount of publications released every year. The underlying problem of this development is addressed by Bradford's law [2]. Bradford states that the effort to find relevant publications for oneself increases exponentially over time

and the number of available publications. As a result, the chance to miss useful research information is very high.

The miss rate is even further increased, due to the fact that not enough publications are taken into account. Most researchers mainly focus on publications that have been published in their own research field. By also considering publications with different research backgrounds, new aspects or questions for already well known problems can arise [3].

In a large research cluster Social Portals are used as an approach to assist interdisciplinary collaboration in order to increase the awareness of research generated within an organization [4]. Publications and their relationships can be visualized [5, 6] in order to improve access to research results from within an organization. Nonetheless, this requires researchers to put their own effort into searching for relevant publications.

In order to provide researchers with the necessary means to increase their finding rate, a recommender system can be utilized. However, to be beneficial to the researcher not only is it necessary to generate good recommendations, but also to convince the users that the system is trustworthy and beneficial for them. The success of both aspects greatly depends on the recommendation algorithm, visualization of its results and the systems look and feel [7]. A recommender system that provides valuable suggestions most of the time may still be perceived poorly if its results are difficult to access or understand [8]. Web-based recommender system also require to appeal to the hedonic needs of the user to be successful [9], thus overall visual appeal is highly important.

2   Related Work

Several approaches have been used in recommender systems to improve their outcome on different aspects. The initial aim for our system was to create a highly transparent recommender system, where the user himself explores the data graph to find appropriate content. In order to ensure that the system meets user requirements, we identified two critical aspects for our project: visualization and recommender logic.

Telling Stories with Visualizations Visualizations for data exploration or recommender systems have recently started to employ techniques that tell stories with data. For example, Wu et al. [10] have used a tree branch visualization to show the development of career paths of researchers. By visualizing which topics were published in which year, the development and shift of interest of researchers can be seen. A similar approach has been tried by Liu et al.[11], who using co-word analysis have used a visualization to track the change of research topics over time.

Segel & Heer [12] propose the use of so-called narrative visualization for recommender systems. The main focus is put on the data itself and its arrangement. Depending on the query and the user's preferences, the system generates a result screen, which consists of multiple items connected to the initial query. The items

are aligned so the user does not perceive them as simple facts, but more as a story told to him.

The project Bohemian Bookshelf [13] shows the potential of creating explorative interfaces. It emulates a digital book shelf. Instead of using the hard covers as visualization means, the user can choose between multiple styles of visualization. One of these styles is a clustered bubble graph. When a user inspects a bubble, he is not only shown the referenced book, but also all books aligning to this bubble. In a case study, it was shown that users were highly motivated and genuinely excited to use the system, since they felt more integrated.

Since data arrangement in graphs are not arbitrary, but contextual, graphs themselves provide information. For this purpose Miller et al. [14] developed a cluster graph. The graph was accumulated of numerous papers, which were analyzed in respect of meaningful words. The whole corpus of papers was then rearranged into word clusters, which consisted of their respective papers. The goal was to give users an overview of possible current trends, but also to motivate them to work interdisciplinary with other facilities.

Integrating the User in Recommender Logic Another trend in recommender systems is to integrate the user in the recommender logic. The user gets control over various aspects of the system or his behavior is analyzed to optimize recommendations. Loepp et al. [15] base the recommendations on user choices that have been done previously, thus applying a mix of collaborative filtering and user analysis. But analyzing the users choices his preferences are elicited by factor analysis.

The other approach is to implement the user as the recommender logic. Mühlbacher et al. [16] found that the main challenge is to identify significant steps in the system and to visualize them in a understandable form. While such systems provide high level of interaction, it is difficult to select the right steps for the user to influence according to Yi et al. [17].

In other projects, we have also encountered the idea of parametrization of the recommender process. However, the parametrization is rather limited and the results are only shown as a list. Examples of this approach are existing in team recommender systems. T-Recs [18] is a system, which suggests developing teams for upcoming projects. Thereby, the user can influence the importance of specific requirements. Another team recommender HR Database for team recommendation [19] also generates suggestions, however it requires the user to input their requirements at the start of each query.

For our approach we decided to use a combination of both ideas. On one hand enabling the user to explore the data, on the other hand giving him the ability to directly influence the query parameters and the result visualization.

3   The Recommender System - TIGRS

In this paper we propose a user-centered recommender system for researchers. The recommender system supports the user in identifying publications suited for

his research interests. In addition, it enables him to explore the set of publications on his own (see Fig. 1).
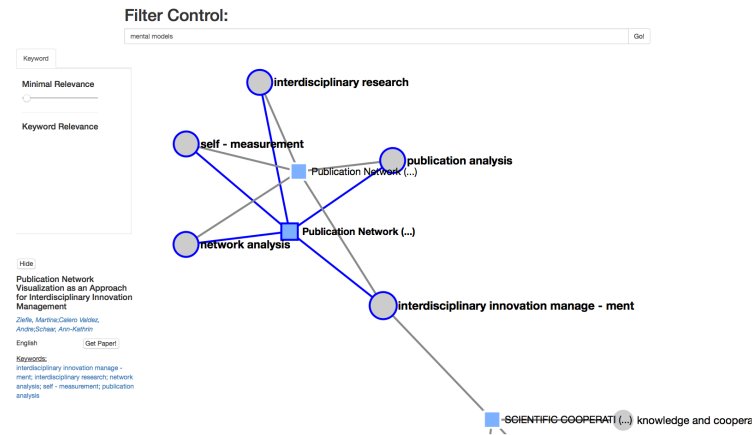

Fig. 1. Exemplary visualization of a user selecting a publication.

In contrast to conventional recommender systems, the proposed one has a user-centered interaction model. Thereby, the system provides the user with an interface that allows to directly influence the behavior of the recommender algorithm and thus immediately observe the impact of adjustments. Therefore, the system is separated into two parts, the first one is responsible for the adjustment of the algorithm behavior and the second part deals with the results and their visualization.

The behavior adjustments range from influencing the topic weighting in the filtering process to focusing on content-based recommendations. In addition the system also enables the user to filter for only keywords but also for specific authors and their respective research fields.

3.1   Text-Mining for Keyword Relevance

For our recommendations we use a keyword based approach. All full-text PDF files are required for the text mining approach. Furthermore we have access to our institutional library that allows API based access to meta data (when available) to ensure correctness of data. Similar APIs are provided by arXiv[3] and Mendeley[4]. When no meta-data is available TIGRS scans the PDF for keywords in the document header.

---

[3] `http://arxiv.org/help/api/index`
[4] `http://dev.mendeley.com/`

We then use Apache OpenNLP[5] to mine the full text of the PDF. Using language and noun phrase detection we reduce the amount of words considered for further data processing. To ensure, that words have no duplicates, the system calculates the Levenshtein-Distance for the potential duplicate pairs. If the resulting value exceeds a predefined tolerance level, the respective words are merged together. For the remaining words we perform term frequency-inverse document frequency (TF*IDF) to establish the words relevance for the document in contrast to the corpus.

For every keyword that we find, we gather the distinctive words from all documents that refer to that keyword in a global category. Then we use an iteration of TF*IDF, term frequency-inverse category frequency (TF*ICF), where we calculate the relevance of a word in a category in contrast to the whole category-corpus. The resulting word sets do not only describe, how distinctive a word is for a paper, but also the word's relevance as a representative of its category. Finally, the keyword relevance for each paper is calculated by adding up all TF*ICF values. Using this approach we can identify the relative importance of each keyword for each document.

3.2   Visualizing Results

The adjustments and recommendations are accessible from the visualization UI. The visualization UI consists of a responsive graph. Each node within the graph represents either a keyword or a publication that match the users' research profiles or their interests. The graph reacts to every interaction of the user, thereby immediately displaying the consequences of the user's actions. Additionally, the graph acts as a substitute for the conventional ranking visualization of results. Because of that, the user is able to better distinguish between the recommended items in respect of their value to the user and also their discerning factors between one another [20].

Besides the graph, the system allows the user to explore the whole database on his own by traversing the links of the graph. In doing so, the connection between publications and topics are further clarified. Furthermore, the publications are put into context to one another.

Our graph based visualization has two type of nodes (see Fig. 2). The first type of nodes are publication nodes. For them shortened titles are displayed. Publication nodes are connected to keyword nodes, when the keyword is listed on the publication. Node size of keywords depends on node degree. This makes keywords that are used in multiple documents larger than less frequently used keywords (see Fig. 3). For each edge a relative importance is stored as a double value indicating the relative relevance of the keyword for the document.

The UI has a filter that allows auto-complete assisted selection of keywords. Furthermore from any given node all its meta descriptors such as keywords or authors can be added by a single click. Adding a keyword to the filter adds a relevance selector to the left part of the screen. By moving the selector the

---

[5] http://opennlp.apache.org/
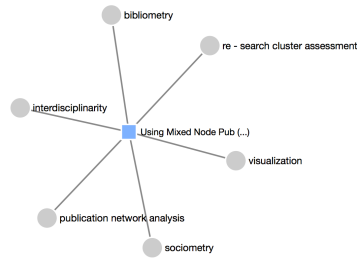
Fig. 2. Publications are displayed as blue squares and keywords as gray boxes, using [6] as an example.

user can select a minimum threshold of relevance of a keyword. The author filter retrieves the research profile of the selected author and adds it, similar to the keyword filter, as a unique filter with a relevance selector. This limits the amount of publications displayed and allows to dynamically adjust weighing of filter keywords by the user.
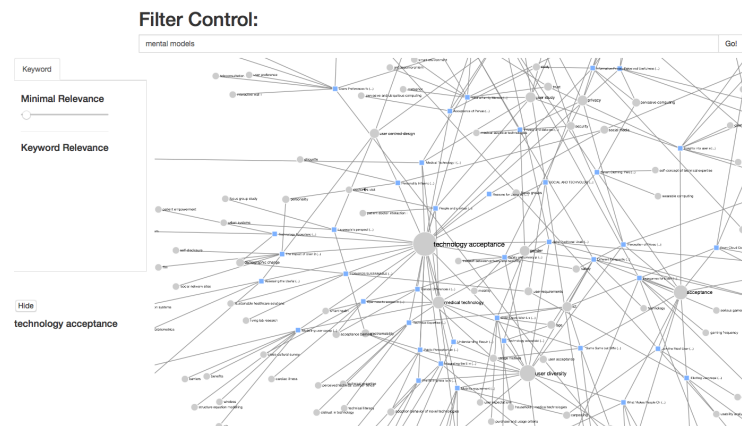
Fig. 3. By visualizing all research of a group prominent topics become more apparent

4   Evaluation

The recommender system was tested in an interdisciplinary research facility with a sample of 16 members from different fields. In a user study we evaluated usability [21](SUS) in respect to user factors (e.g. age, gender, track record). We particularly evaluated the effectiveness of recommendations by measuring

trust and accuracy for recommendations [22]. Additionally, we evaluated supplemental factors of relevance of the visualization (i.e. structure and overview, topic discovery, information about colleagues) and compared the visualization to a list-based recommendation. At last we evaluated the visualization using the NPS [23] (NPS).

4.1   Method

First participants were handed a questionnaire to elicit user factors. They were given access to the visualization and given a short introduction into the general mechanics of the visualization (What are node types? What do mouse gestures do? etc.). Then they were given two tasks. First they were asked to play around with the visualization until they felt comfortable using the visualization. Then they were asked to look for a publication in the recommender system that was relevant to them and previously unknown. The whole process was recorded by video and later analyzed. After the interaction users were given another questionnaire to evaluate the prototype.

The assessed metrics for the prototype are partially taken from ResQue, [22] accuracy (A.1.1, $\alpha = .745$), relative accuracy (A.1.2, $\alpha = .362$) and generated from own items (see Tab. 1). All were measured on six-point Likert scales. For all used scales we assessed the Cronbach's α when more than one item is used. The SUS had a reliability of $\alpha = .731$.
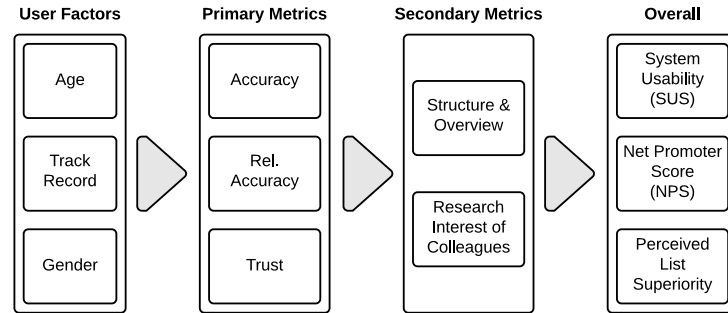
Fig. 4. Research model overview

Additionally, we assessed whether our visualization was seen as superior to a list based presentation in regard to four aspects. Does the visualization help when one is looking for *new content*? Does it help in *understanding the research group*? Does it provide more *overview* and provide *more information* in general than a list based presentation? Those were assessed on a six-point Likert scale (1=disagree completely, 6=agree completely). The investigated relationships can be seen in Fig. 4.

| Scales and Items |
| --- |
| **Research interest of colleagues** Cronbach's $\alpha = .858$
- The suggested publications help me understanding my colleagues research interests.
- The research interests of my colleagues can be derived from the visualization.
- The visualization helps to understand my colleagues research interests.
Trust Cronbach's $\alpha = .808$
- I believe that the system can give sensible recommendations.
- I trust that the system gives me sensible recommendations.
- I would rather trust my colleagues to give me recommendations than the system.*
Structure and overview Cronbach's $\alpha = .594$
- The suggested publications provide an overview of my teams work.
- The visualization structures the content for me to help me maintain an overview.
- The visualization helps aligning my work content with my team.
- The visualization supports me in staying consistent with my colleagues work. |

Table 1. Scales and their item texts. *=inverted items.

4.2 Sample Description

As a sample of $N = 16$ researchers from an interdisciplinary research facility were selected at random. The average age of the researchers was $\bar{x} = 33.6$ years ($\sigma = 6.14$, range$= 23 - 52$) and 56% of the participants were female. 10 had finished their Masters (or similar) while 5 already had a Ph.D. In total we had six communication scientists, five psychologists, four computer scientists, three sociologists and one architect in our sample (multiple selections allowed). When looking at the track record distribution of experience was mixed ($\bar{x} = 4.25$, $\sigma =$, 0=no publications, 7=more than 30 publications). Although most researchers had a focus on conference proceedings ($\bar{x} = 4.0$, $\sigma = 2.0$). Journal articles ($\bar{x} = 2.57$, $\sigma = 1.55$) and book chapter contributions were less frequent ($\bar{x} = 1.93$, $\sigma = 1.54$, see also Fig. 5).
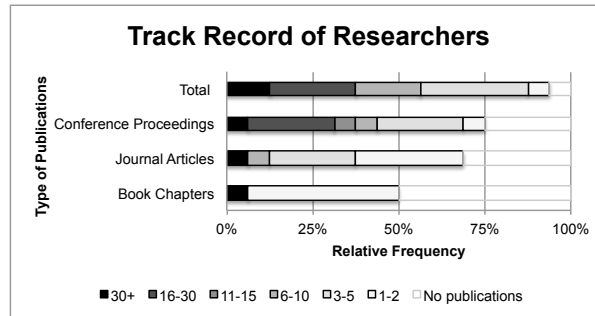

Fig. 5. Overview of the track records of the individual researchers

4.3   Descriptive Results

When looking at the results descriptively, we can say that the *accuracy* of the system is very high ($\bar{x} = 5.03$, $\sigma = 0.15$), while the *relative accuracy* is relatively low ($\bar{x} = 3.13.$, $\sigma = 0.19$). This means that the system does give good recommendations, but colleagues recommendations are still seen as superior to the visualization. Interestingly the relative accuracy showed a very low reliability, indicating that the phrasing of the items leads to differing answers between the items.

The *trust* in the given recommendations is relatively high ($\bar{x} = 4.31$, $\sigma = 0.20$). Users were able to get the impression that the given recommendations were actually sensible. When looking at the secondary metrics *structure and overview* showed a high agreement ($\bar{x} = 4.72$, $\sigma = 0.22$) and *colleagues research interest* as well ($\bar{x} = 4.69$, $\sigma = 0.22$). This means besides giving adequate recommendations the system was able to inform the user about the structure of the research group and the research interests of their colleagues. Overall *SUS* was high ($\bar{x} = 4.89/6$, $\sigma = 0.13$) indicating a good usability of the system. Nonetheless the *NPS* was relatively low (-7). This means further development of the system needs to be performed to align with user requirements. In regard to a *comparison over lists* our visualization was considered superior in all four aspects (see Fig. 6).
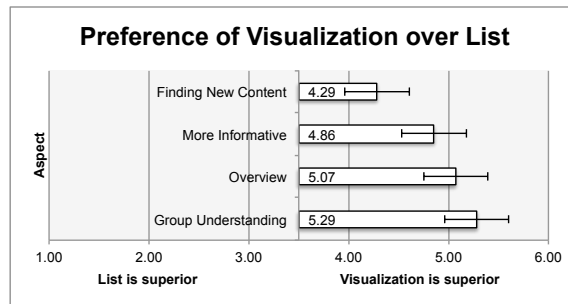
Fig. 6. Perceived preference of our visualization over lists in four aspects. Error bars denote standard errors. A value of 3.5 would indicate a neutral judgement.

4.4   Interaction Effects

When looking at age, gender and track record no interaction with any measured scale could be found ($p > .05$). This means that all our users were able to use the system and evaluated it independently from our user factors.

Trust was a factory correlating with most other evaluated metrics. Trust and accuracy showed a high correlation ($r = .761$, $p < .01$), similar as trust and overview ($r = .825$, $p < .01$). The SUS only correlated with accuracy and *research interest of colleagues*, while NPS correlated with trust directly (see also Fig. 7).
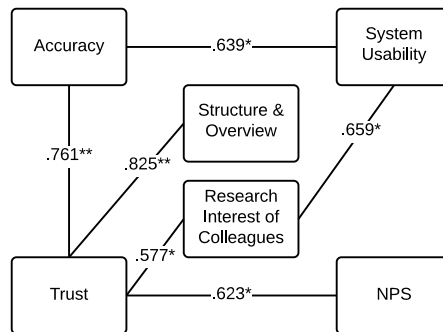
Fig. 7. Correlations between scales. Numbers denote Pearson's r

4.5   Summary

Overall we can say that our visualization and recommender approach was evaluated quite positively. The usability was rated as good and the visualization was judged superior over list based presentations. Interestingly the NPS was relatively low indicating the need for further improvements.

Our visualization is particularly good at assisting in understanding the research group while giving the user information on the research structure and an overview of the institute. Trust seems to be a major factor in influencing adoption by the user because it correlates with secondary metrics and the perceived accuracy of the recommendations.

5   Limitations and Future Work

The prototype of the recommender system was tested in a first iteration. Naturally, there are some limitations which have to be considered when further developing the tool. A technical limitation regards the fact that the system is limited to visualizing publications to which non-encrypted PDF files are available for the text-mining to work. This might work in a research setting in which all relevant publications are available, as e.g. within research groups that might work together for an extended period of time. Further limitations of the current version of the prototype are directed to usability. During the user studies some improvements were mentioned, mostly with respect to the interaction possibilities. Users should be able to change the graph density and the amount of recommendations directly (in order to prevent visual overload and cognitive complexity). Furthermore, we want to improve on the transparency of the relevance thresholds. Changes on the slider should directly highlight changes in the graph to improve the understanding of how the relevance slider works. Finally, functions scope was quite narrow. It could be helpful to add filtering based on article source (outlet) and publication date, in order to support the search process and to match user expectations. This also reflects the nature of the approach here:

It was a computer science approach (to automatize the search for publications) that was then tested and evaluated with users. Another way to improve the tool would be the vice versa approach: observing natural users during academic work, what they are looking for, and why and how the keywords are connected semantically. The findings then could be matched with the recommendation systems finding.

Future work might be directed to the understanding of different search scenarios. One could be to examine different levels of domain knowledge and to study the perceived usefulness of the system. Here we could expect that the system should be extremely helpful for getting a fast overview for novices, while it could be even detrimental for experts with an elaborated mental model. Also it seems worth to study different search approaches across different target scenarios: looking up information in a quick and dirty style (searching for a specific information), or, getting an overall picture of a sample of papers (learning about major research topics of a group), or just looking for interesting papers within a given field.

References

1. Beaudoin, C.E.: Explaining the relationship between internet use and interpersonal trust: Taking into account motivation and information overload. Journal of Computer-Mediated Communication **13**(3) (2008) 550–568
2. Hjorland, B., Nicolaisen, J.: Bradford's law of scattering: Ambiguities in the concept of "subject". In Crestani, F., Ruthven, I., eds.: Context: Nature, Impact, and Role. Volume 3507 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2005) 96–106
3. Maglaughlin, K.L., Sonnenwald, D.H.: Factors that impact interdisciplinary natural science research collaboration in academia. In: Proc. of the ISSI. (2005) 24–25
4. Calero Valdez, A., Schaar, A.K., Ziefle, M., Holzinger, A.: Enhancing interdisciplinary cooperation by social platforms. In Yamamoto, S., ed.: Human Interface and the Management of Information. Information and Knowledge Design and Evaluation. Volume 8521 of LNCS. Springer International (2014) 298–309
5. Schaar, A.K., Calero Valdez, A., Ziefle, M.: Publication network visualization as an approach for interdisciplinary innovation management. In: Professional Communication Conference (IPCC), 2013 IEEE International, IEEE (2013) 1–8
6. Calero Valdez, A., Schaar, A.K., Ziefle, M., Holzinger, A., Jeschke, S., Brecher, C.: Using mixed node publication network graphs for analyzing success in interdisciplinary teams. Active Media Technology (2012) 606–617
7. Knijnenburg, B., Willemsen, M., Gantner, Z., Soncu, H., Newell, C.: Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction **22**(4-5) (2012) 441–504

8. Ortega, F., Bobadilla, J., Hernando, A., Rodríguez, F.: Using hierarchical graph maps to explain collaborative filtering recommendations. Int. J. Intell. Syst. **29**(5) (2014) 462–477
9. Alagöz, F., Calero Valdez, A., Wilkowska, W., Ziefle, M., Dorner, S., Holzinger, A.: From cloud computing to mobile internet, from user focus to culture and hedonism - The crucible of mobile health care and wellness applications. In: ICPCA10 - 5th International Conference on Pervasive Computing and Applications. (2010) 38–45
10. Wu, M.Q.Y., Faris, R., Ma, K.L.: Visual exploration of academic career paths. In: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ACM (2013) 779–786
11. Liu, Y., Goncalves, J., Ferreira, D., Xiao, B., Hosio, S., Kostakos, V.: Chi 1994-2013: mapping two decades of intellectual progress through co-word analysis. In: Proceedings of the 32nd annual ACM conference on Human factors in computing systems, ACM (2014)
12. Segel, E., Heer, J.: Narrative visualization: Telling stories with data. Visualization and Computer Graphics, IEEE Transactions on **16**(6) (Nov 2010) 1139–1148
13. Thudt, A., Hinrichs, U., Carpendale, S.: The bohemian bookshelf: Supporting serendipitous book discoveries through information visualization. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12, New York, NY, USA, ACM (2012) 1461–1470
14. Miller, L.J., Gazan, R., Still, S.: Unsupervised classification and visualization of unstructured text for the support of interdisciplinary collaboration. In: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing. CSCW '14, New York, NY, USA, ACM (2014) 1033–1042
15. Loepp, B., Hussein, T., Ziegler, J.: Choice-based preference elicitation for collaborative filtering recommender systems. In: Proceedings of the 32nd annual ACM conference on Human factors in computing systems, ACM (2014) 3085–3094
16. Mühlbacher, T., Piringer, H., Gratzl, S., Sedlmair, M., Streit, M.: Opening the black box: Strategies for increased user involvement in existing algorithm implementations. to appear in IEEE Transactions on Visualization and Computer Graphics ? Proceedings IEEE VAST 2014 (2014)
17. Yi, J.S., ah Kang, Y., Stasko, J., Jacko, J.: Toward a deeper understanding of the role of interaction in information visualization. Visualization and Computer Graphics, IEEE Transactions on **13**(6) (Nov 2007) 1224–1231
18. Datta, A., Tan Teck Yong, J., Ventresque, A.: T-recs: Team recommendation system through expertise and cohesiveness. In: Proc. of the 20th Intern. Conf. Companion on WWW. WWW '11, New York, NY, USA, ACM (2011) 201–204
19. Brocco, M., Hauptmann, C., Andergassen-Soelva, E.: Recommender system augmentation of HR databases for team recommendation. In: Database and Expert Systems Applications (DEXA), 2011 22nd Intern. Workshop on. (2011) 554–558
20. Cramer, H., Evers, V., Ramlal, S., van 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) (November 2008) 455–496
21. Brooke, J.: SUS - a quick and dirty usability scale. Usability evaluation in industry **189** (1996) 194
22. Pu, P., Chen, L., Hu, R.: A user-centric evaluation framework for recommender systems. In: Proceedings of the fifth ACM conference on Recommender systems, ACM (2011) 157–164
23. Reichheld, F.F.: The one number you need to grow. Harvard business review **81**(12) (2003) 46–55