What Do My Colleagues Know?
Dealing with Cognitive Complexity in
Organizations through Visualizations

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

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

Abstract. In order to cope with the growth of information complexity,
organizations have started to implement various forms of knowledge man-
agement applications. Approaches range from file-, data-, information-
centric software to information retrieval, search engines, and decision
support systems. Thereby, the data presentation plays often a crucial
part in making knowledge available in organizational settings. We ex-
amine two visualizations and investigate their capabilities to support or-
ganizational knowledge and their usability. One is a document-keyword
centric graph-based visualization, while the other is person-institute cen-
tric. Both were evaluated positively in supporting improvement of orga-
nizational knowledge.

Keywords: Social Portals, Knowledge Discovery, Recommender Sys-
tems, Visualization, User-Study, Trust

1 Introduction

Knowledge grows. The growth of knowledge has continued to accelerate over the
recent years and is expected to do so in the future [1]. As a result, it becomes
more and more difficult to extract meaningful data from the available knowl-
edge. For this extraction, it is necessary to turn knowledge into information and
the information into data.[2]. The field of information science and information
management has developed a copious amount of research on how to store data
effectively as information (e.g. meta-data). Nevertheless, the amount of informa-
tion and data is growing rapidly, therefore hindering the user from acquiring his
desired knowledge from data and information.

Coping with this ever increasing amount of information and data is the cen-
tral challenge of big data and knowledge discovery. The challenge shifts from or-
ganizing information to finding information that is both relevant [3] and helpful

to the user in order to create knowledge. This search process becomes increasingly important in the organizational context.

Organizations are becoming ever more complex and employee fluctuation makes knowledge management increasingly harder to do [4]. To prevent knowledge loss, knowledge management systems are increasingly used. However, these address mainly explicit knowledge and fail to capture tacit knowledge. Social software solutions have tried to address this topic by capturing communication processes as they happen and storing this data for later evaluation.

Nonetheless retrieval of this knowledge from the stored data is still a great challenge. Whenever information is connected and complex, it becomes necessary to not only transfer the information from system to user, but also to shape the mental model [5] the user has of the information. It is furthermore necessary to take into account what the model of the technology behind knowledge management software is, due its strong effect on acceptance [6]. The composition of the software must be clearly communicated. Here various forms of systems come into play that address different aspects of knowledge management and knowledge transfer. Still, if the mental model of both system and content are clear, success of such a system can not be guaranteed. Often the creator of information and the benefactor are different persons [7] thus a sense of community or even locality [8] are important criteria for their success.

2   Related Work

In this paper we look at various forms of knowledge management and discovery systems and address their applicability in an organizational setting. Typical forms of knowledge management systems that deal with big data are the following [9]:

1. *Intranet and Groupware* software [10] are designed to support organizational collaboration and integrate a network of clients. Typically summarized under the term CSCW they are designed with work tasks in mind and often based on well known protocols like HTTP, SMTP and FTP. They are most of the time file-centric.
2. *Data Warehousing & OLAP* [11] are data centered solutions. Transaction and Process data [12] is integrated and stored in data cubes, which can later be analyzed for reporting purposes. Data warehouses implement a single-source of truth policy and keep track of data history. Tools are required for extraction, cleaning and loading data into the Online Analytical Processing (OLAP) system. Afterwards, the OLAP system analyzes the given data cubes to find patterns or possible trend candidates. Since queries are often multidimensional, meta-data management and query management is important and often tool-assisted.
3. *Content Management Systems* are content-centric and focus on publishing processes. Content can be created, updated, published and deleted. In addition, editing workflows are implemented to address publishing responsibilities. Generally versioning and authorship meta data is maintained. Recently

CMS have been used in enterprise content management, as internal documents often follow similar procedures as publishing.

4. *(Collaborative) Search engines* [13, 14] or enterprise search engines merge the joint efforts of users in locating and tagging information. This allows the retrieval of more relevant information by learning from user input and the relations of users interests.

5. *Recommender Systems* [15] are used to actively suggest interesting content to the user. Often used in Internet sales to suggest other products that are of interest. Furthermore, they are used to recommend documents, books[16], scientific literature[17] or even teams [18]. Recommender System often learn from users previous choices [19] but may also rely on multi-criteria filtering [20]. How suggestions are generated, should be clearly explained. [21] This is especially important in the case of hybrid systems [22] that integrate collaborative filtering and machine learning approaches.

6. *Decision Support Systems (DSS)* [23] derive significant information and possible emerging patterns from raw data. By considering the extracted information, the system assists the decision making process. DSS often incorporate visualizations and can be fully automatic, fully human dependent or combine both efforts.

Still all of these systems rely on various forms of information presentation to allow the user to acquire knowledge from the information or data presented. In all cases the type of visualization is critical to improve the transfer of deep structure from machine to user [24]. The concept "overview, zoom, detail on demand" [25] summarizes a core paradigm of visualizations that are based on full information display. The field of *HCI-KDD* [26] addresses the need for research of the interaction of Human-Computer Interaction and Knowledge Discovery in Databases.

For the case of organizational knowledge it is also important to understand the complexity of knowledge available in an organization that is shared between employees. Finding an employee or a document with critical knowledge or information is a challenge when organizational structures are not well understood. Users must learn the intricacies of overview, structure, and detail along the hierarchy of an organization.

2.1   Visual Recommender Systems

In order to ease the understanding of information, visual approaches can be used. In our case we focus on visual recommender systems. The recommender component, serves the purpose of increasing the transparency of the underlying system. The only similar solution that we could find to our prototypes is proposed by O'Donovan et al. [27]. They propose a graph-based visual collaborative filtering tool called PeerChooser that uses multiple criteria to allow users to find movie suggestions. Montaner et al. [28] propose a taxonomy of recommender systems spanning seven criteria. These should be used in order help in designing a recommender system. The instances are task dependent but are for most cases interchangeable.

- *Representation* describes how data is represented (e.g. historical, feature vector, etc.)
- *Initial Information* describes how data is preloaded into the algorithm before the user interacts with it (e.g. none, manual, training set)
- *Learning* refers to how the algorithm improves on usage (e.g. TF-IDF, ID3, etc.)
- *Feedback* describes how the user can give feedback to the algorithm (e.g. rating systems, choice).
- *Adaptation* refers to how the algorithm adapts to the feedback (e.g. add new, natural selection, GFF).
- *Filtering* indicates how data is filtered before used in the algorithm (e.g. collaborative, hybrid, demographic)
- *Matching* describes how items are matched with the users requests or feedback (e.g. nearest neighbour, cosine similarity, etc.)

3   Visualization Prototypes

By considering Montaner et al. criteries, we investigate two different types of knowledge discovery systems in a research setting addressing these topics. We present a user evaluation of these systems and their particular visualizations. The first system is a visual recommender system that recommends documents to read that are relevant to the research interest of the user. The second system presents a visual collaboration support system that allows finding collaborators in a research organization that can contribute to the user's topic. Both systems are integrated in a social portal that is used within the research organization.

Graph-based Document Recommender System. The graph-based publication recommender system TIGRS uses a mixed-node graph [29] connecting publications with their keywords, with respect to their relative relevance (see Fig. 1 and [30]). Users can now filter for keywords and their relative relevance in order to find relevant documents. The system uses the users previous keywords to suggest only documents that are relevant to the user. It allows to browse through content, while at the same time seeing connections between documents sharing mutual keywords.

Collaborator Suggestion System. The collaborator suggestion system proposed by Yazdi et al. [31] is used to suggest fruitful collaboration in a research cluster by analyzing previous collaboration and mutual keywords. By using social network analysis possible coauthors are visually suggested when hovering over a bubble-based graph. Using a bubble-bag layout (see Fig. 2) it additionally conveys information about where a suggested collaborator works. This further conveys organizational structure information, allowing users to understand who is who in their organization and what they work on.
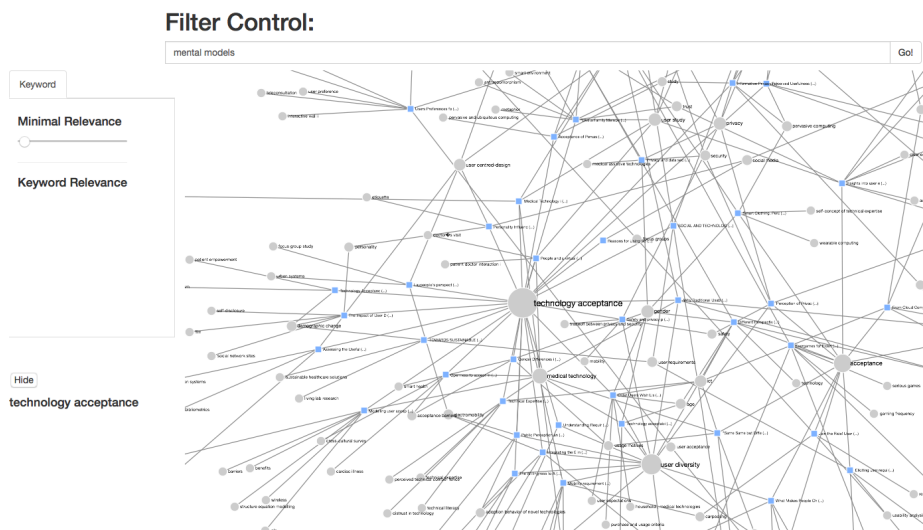
Fig. 1. Publication and keyword centric visualization of collaboration [30]

4   Method

We tested both systems in a large research facility (i.e. over 180 researchers) with a sample of 16 and 20 members from different fields for each system. Both prototypes were used in a user-studies ($N = 16$, $N = 20$) determining both the overall usability (SUS [32]) and the likelihood of being recommended to a fried (NPS [33]). We investigated user factors (e.g. age, discipline, research expertise, track record) and evaluated how both prototypes complement each other in a scientific setting.

With regard to recommender systems, finding appropriate metrics for their evaluation is critical [34] in order gain an understanding of what needs to be optimized. Here we only look at general usability and qualitative insights. For further detailed analysis please refer to the original works of the prototypes (cf. [30, 31]).

For both prototypes users were invited to take part in a user study in our laboratory. They were given time to get accustomed to the prototypes and their handling and were then given tasks to complete (i.e. find suitable publication/collaborator). The whole process was recorded and then analyzed. Additionally further quantitative analyses were performed derived from a questionnaire completed by the participants.

In this paper we want to focus on reporting qualitative findings from both prototypes and their implementation in a social portal. Both prototypes are to be integrated in the so-called "Scientific Cooperation Portal" (SCP). [35] a tool devised to tackle the high staff volatility in a large research cluster. The SCP is a social portal which centralizes communication, file-exchange, member profiles,
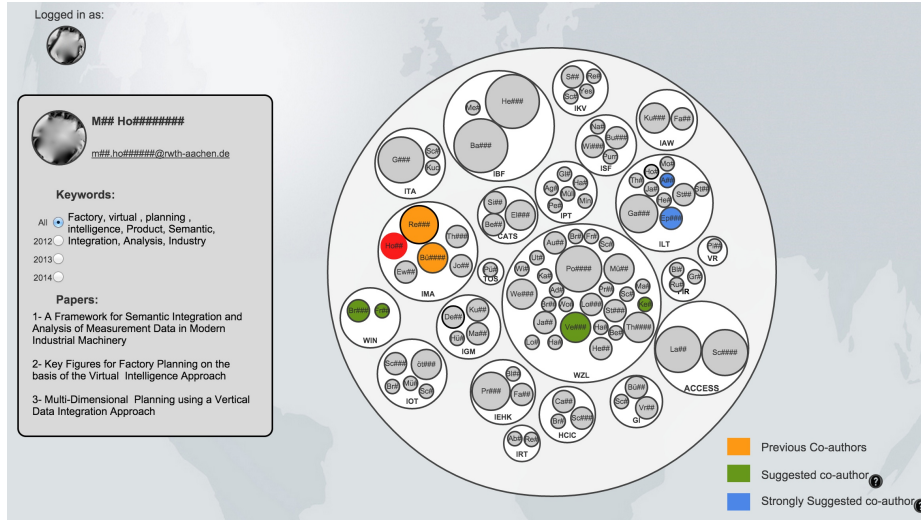
Fig. 2. Author and institute centric visualization of collaboration [31]

and offers interdisciplinary collaboration support. Additionally, it tracks research output of individual researchers by tracking their publications. This latter feature is also used to enable steering the cluster from a management point of view [29]. In the prototypes we use this data to construct visualizations that help facilitate collaboration and understanding the organization.

5   Results

Our analyses (univariate analysis of variance) show that both systems address different aspects of understanding how an organization works. The first allows understanding how different departments work on various topics, while overlapping in their content and methodology. The second system allows understanding the scientific content that researchers work on in depth.

5.1   Sample Description

For the first prototype we asked $N = 16$ researchers from an interdisciplinary research facility whose average age was $\bar{x} = 33.6$ years ($\sigma = 6.14$, range$= 23-52$) and 56% of whom were female to take part in our study. Ten had finished their undergraduate training (Masters) while five already had graduate training (PhD or Professor). Most researchers came from the fields of linguistics or communication science (see Tab. 1). Most researchers had published about 5-6 papers in their careers with some outliers of over 150 (i.e. a professor) and some none (new colleagues). The facility has a total of 25 researchers.

Field	Count
Linguistics and Communication Science	6
Psychology	5
Computer Science	4
Sociology	3
Architecture	1

Table 1. Research fields in sample for the first prototype (multiple selections allowed)

The second prototype was testes with $N = 20$ researchers from an interdisciplinary research cluster (out of 180 employees). Forty participants were approached at seven different institutes mostly from engineering sciences, but also including communication science and computer science.

5.2 Quantitative Results

The first prototype received very high ratings in usability. Overall SUS was high ($\bar{x} = 81.5$, $\sigma = 2.17$) indicating a good usability of the system. Nonetheless the NPS was relatively low (-7). We got 4 detractors, 8 passives, and 4 promoters. This means further development of the system needs to be performed to align with user requirements.

Quantitatively the SUS showed a mean of $\bar{x} = 82.5$ ($\sigma = 24.4$) indicating a high acceptance of the prototype. The NPS analysis yields 4 Promoters, 6 Passives and 0 Detractors. The overall NPS is 40 indicating good usability [32, 33].

5.3 Qualitative Results

In addition to the quantitative evaluation we screened the user study recordings for mentions of various categories. We also analyzed the behavior how users were using the prototypes and in particular how surprised they were. Furthermore we asked users what they learnt about their organization and how much the visualization improved their knowledge of the organization.

Looking for Relevant Publications. An interesting observation during the usage of the first prototype was the different styles of how it was used. We identified three different approaches how users used the filtering mechanism.

The first style was a *drill down* approach. Users that applied this approach first looked at the full graph including all publications from their institute, resetting all filters before looking for a recommendation. They then used generic terms that were of interest to them and played with the relevance sliders to further drill down on interesting suggestions. Then more experienced users were trying to look for items they did not know yet while keeping their focus on the center of the graph were items are that are connected to all relevant filter terms.

The second style was an *incremental bag* approach. They started with a very specific term that was of current interest to them often dissatisfied with the few results they gradually increased the bag of filters with specific terms. Interestingly these users reported to find very relevant suggestions albeit often previously known suggestions.

The third style was a *traverse related work* approach. Users applying this style looked for single items that they found interesting and sequentially added keywords that were relevant to that single item, repeating this process several times. This can be seen as a traversal along keywords approach often leading to utterances like "we have someone writing about this, I was looking for something like that", indicating very serendipitous finds.

Finding Fruitful Collaborators. When using the second prototype most users were astonished with how much the visualization revealed about the organizational structure. Users were surprised to see that others in the organization (out of 180 researchers) were working on similar topics that they were. In particular seeing the who has worked with whom was interesting as this information was somewhat opaque to find from publication lists.

Discovering New Knowledge About the Organization. Both prototypes were able to reveal new knowledge by visualizing publicly available information in a new fashion. Both prototypes caused users to have serendipitous finds either as publications or possible collaborators from a pool that was theoretically available to them. Nevertheless, the effort to manually look for this information had been a barrier to do so (both organizations existed for more than 5 years).

In both cases the head of the organization was asked to use the prototypes and talk about the benefits of the visualization for managing purposes. Both mentioned the benefits of getting overview knowledge about their organization. They were also surprised to see how much publications had been written since the funding of the organization and realized the scope of their organization. Interestingly, a need to communicate publications to teams arose during these experiments.

6   Conclusion

In this paper, we have examined two knowledge systems. The first system recommended relevant documents to the user in the form of a visual recommender system. Thereby, the system provides the user with the means to directly influence the recommender algorithm and the recommender visualization. Whereas the second system, supports the user in finding suitable future collaborators, who can contribute to the user's topic.

In conjunction both systems provide insights into what the colleagues do and how their work can be useful in respect to the user's own work. This knowledge can be used either for collaboration or as a basis for one's own work. The systems

help in creating knowledge from data and information through their specifically adapted visualizations.

During the use various types of new applications arose. The need to extend the visualization to other types of documents was seen. In particular seeing not only publications but also grant proposals for a whole university and collaboration suggestion within the university was mentioned as a possible application.

6.1 Limitations

Both tools require PDFs and Full-Texts with meta information to work properly. The conducted studies were done with relatively small user groups because of the intensive analysis required after the test. Therefore, the experiment is not able not reveal any effects between user-factors and usage behavior. Differences are too small to be statistically significant, thus they must be assumed non-existent. Inspite of that the experimenter felt the need to report, that less experienced researchers were using the tool differently than experienced researchers. Further analysis of the videos might reveal these differences at a later point in time.

6.2 Outlook

The user studies indicate that both prototypes can be used not only to assist scientific collaboration but also in organizational knowledge management. Here they can be used to interlink documents from an enterprise content management system in order to find relevant documents when working on another. The aim of our prototypes is to bring the various levels of collaboration support together. Documents can be served from the intranet or CMS and be connected with collaborative search and tagging from a social portal. The publication recommender system TIGRS brings these together by providing relevant documents to the user integrated into the social portal. The collaboration suggestion system even goes a step further as it actively recommends suitable collaborators for the users of the social portal. Nonetheless, only in conjunction can they help in understanding the organizational structure, the employees and the topics that are being worked on collaboratively.

References

1. Metcalfe, J.S.: Knowledge of growth and the growth of knowledge. Journal of Evolutionary Economics **12**(1-2) (2002) 3–15

2. Bellinger, G., Castro, D., Mills, A.: Data, information, knowledge, and wisdom. URL: http://www.systems-thinking.org/dikw/dikw.htm (2004)
3. Wang, J., De Vries, A.P., Reinders, M.J.: Unified relevance models for rating prediction in collaborative filtering. ACM Transactions on Information Systems (TOIS) **26**(3) (2008) 16
4. Holsapple, C.W., Whinston, A.B.: Knowledge-based organizations. The Information Society **5**(2) (1987) 77–90
5. Calero Valdez, A., Ziefle, M., Alagöz, F., Holzinger, A.: Mental models of menu structures in diabetes assistants. In Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I., eds.: ICCHP (2). Volume 6180 of Lecture Notes in Computer Science., Berlin, Heidelberg, New York, Springer (2010) 584–591
6. Orlikowski, W.J., Gash, D.C.: Technological frames: making sense of information technology in organizations. ACM Transactions on Information Systems (TOIS) **12**(2) (1994) 174–207
7. Grudin, J.: Groupware and social dynamics: Eight challenges for developers. Communications of the ACM **37**(1) (1994) 92–105
8. Olson, G.M., Olson, J.S.: Distance matters. Human-computer interaction **15**(2) (2000) 139–178
9. Sharma, S.K., Gupta, J.N., Wickramasinghe, N.: A framework for building a learning organisation in the 21st century. International Journal of Innovation and Learning **2**(3) (2005) 261–273
10. Ellis, C.A., Gibbs, S.J., Rein, G.: Groupware: some issues and experiences. Communications of the ACM **34**(1) (1991) 39–58
11. Chaudhuri, S., Dayal, U.: An overview of data warehousing and olap technology. ACM Sigmod record **26**(1) (1997) 65–74
12. Dumas, M., Van der Aalst, W.M., Ter Hofstede, A.H.: Process-aware information systems: bridging people and software through process technology. John Wiley & Sons (2005)
13. Sugiyama, K., Hatano, K., Yoshikawa, M.: Adaptive web search based on user profile constructed without any effort from users. In: Proceedings of the 13th international conference on World Wide Web, ACM (2004) 675–684
14. Morris, M.R., Horvitz, E.: Searchtogether: an interface for collaborative web search. In: Proceedings of the 20th annual ACM symposium on User interface software and technology, ACM (2007) 3–12
15. Burke, R.: Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction **12**(4) (2002) 331–370
16. 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
17. 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
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. 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

20. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions on **17**(6) (2005) 734–749
21. Herlocker, J.L., Konstan, J.A., Riedl, J.: Explaining collaborative filtering recommendations. In: Proceedings of the 2000 ACM conference on Computer supported cooperative work, ACM (2000) 241–250
22. Gunawardana, A., Meek, C.: A unified approach to building hybrid recommender systems. In: Proceedings of the third ACM conference on Recommender systems, ACM (2009) 117–124
23. Shim, J.P., Warkentin, M., Courtney, J.F., Power, D.J., Sharda, R., Carlsson, C.: Past, present, and future of decision support technology. Decision support systems **33**(2) (2002) 111–126
24. Gretarsson, B., O'Donovan, J., Bostandjiev, S., Hall, C., Höllerer, T.: Smallworlds: Visualizing social recommendations. In: Computer Graphics Forum. Volume 29., Wiley Online Library (2010) 833–842
25. Shneiderman, B.: The eyes have it: A task by data type taxonomy for information visualizations. In: Visual Languages, 1996. Proceedings., IEEE Symposium on, IEEE (1996) 336–343
26. Holzinger, A.: Human-computer interaction and knowledge discovery (hci-kdd): What is the benefit of bringing those two fields to work together? In Cuzzocrea, A., Kittl, C., Simos, D.E., Weippl, E., Xu, L., eds.: CD-ARES. Volume 8127 of Lecture Notes in Computer Science., Springer (2013) 319–328
27. O'Donovan, J., Smyth, B., Gretarsson, B., Bostandjiev, S., Höllerer, T.: Peerchooser: visual interactive recommendation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2008) 1085–1088
28. Montaner, M., López, B., De La Rosa, J.L.: A taxonomy of recommender agents on the internet. Artificial intelligence review **19**(4) (2003) 285–330
29. 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. In: AMT. (2012) 606–617
30. Bruns, S., Calero Valdez, A., Greven, C., Ziefle, M., Schroeder, U.: What Should I Read Next? A Personalized Visual Publication Recommender System. In: Proceedings of the HCI International 2015. (2015)
31. Yazdi, M.A., Calero Valdez, A., Lichtschlag, L., Ziefle, M., Borchers, J.: Visualizing Opportunities of Collaboration in Large Organizations. In: Manuscript submitted for publication. (submitted)
32. Brooke, J.: SUS - a quick and dirty usability scale. Usability evaluation in industry **189** (1996) 194
33. Reichheld, F.F.: The one number you need to grow. Harvard business review **81**(12) (2003) 46–55
34. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems (TOIS) **22**(1) (2004) 5–53
35. 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 Lecture Notes in Computer Science. Springer International Publishing (2014) 298–309