

Expert Recommendation for the Academic Field

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie (Dr. phil.)
durch die Philosophische Fakultät der
Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Tamara Heck
aus
Simmern/Hunsrück

Betreuer: Prof. Dr. Wolfgang G. Stock

Düsseldorf, Juni 2015

Contents

Introduction	1
State of Research	2
Scope and Research Questions of this Work	5
Research Methodology	7
Outline of this Work	7
References	11
1 Scientific Collaboration and Communities.....	15
1.1 Collaboration between Scientists	15
1.2 Communities of Practice for Academic Purposes	18
1.2.1 Principles of Communities of Practice	20
1.2.2 Academic Knowledge Creation.....	27
1.2.3 Survey: Usage of Academic Services and Communities of Practice.....	34
References	39
2 Detection of Scientific Communities.....	48
2.1 Implicit Relations via Bibliographic Coupling and Co-Citation	50
2.2 Expanding Perspectives for Enhancing Academic Networks.....	59
2.3 Implicit Relations via Collaborative Filtering	64
References	74

3	Introduction to Recommender Systems	84
3.1	The Task of Recommender Systems	88
3.2	Types of Recommender Systems	91
3.2.1	Content-Based Approaches	91
3.2.1.1	Examples of Content-Based Text Retrieval Methods.....	93
3.2.1.2	Limitations of Content-Based Approaches.....	95
3.2.2	Knowledge-Based Approaches	96
3.2.2.1	Limitations of Knowledge-Based Systems.....	100
3.2.3	Collaborative Filtering Approaches	102
3.2.3.1	User-Based and Item-Based Approaches	103
3.2.3.2	Limitations of Collaborative Filtering Systems	108
3.3	Comparison of the Approaches	109
3.4	Explicit and Implicit User Ratings	116
3.4.1	Tag-Based Approaches	121
3.5	Expert Recommendations.....	129
3.5.1	General Approaches in Expert Recommendation	130
3.5.2	Expert Recommendation in Social Tagging Systems.....	133
3.6	Evaluation of Recommender Systems.....	136
	References	143
4	Case studies on Expert Recommendation.....	159
4.1	Structure of Bookmarking Data and Similarity Metrics.....	160
4.1.1	Data Set	160
4.1.2	Data Structure.....	162
4.1.3	Differences Between Similarity Models and Coefficients	165
4.2	Qualitative Evaluation of Recommendations for Researchers	177
4.2.1	Data Set and Recommendation Network Model	179
4.2.2	Evaluation and Discussion	180

4.3	Recommending Authors Using Multiple Kinds of Social Information.....	183
4.3.1	Data Set for Implicit Relations Based on Bibliometric Data	185
4.3.2	Data Set for Implicit Relations Based on User-Generated Data	192
4.3.3	Recommendation Models and Evaluation Method	197
4.3.4	Results and Discussion I: Relevance of Recommended Authors.....	208
4.3.5	Results and Discussion II: Relevance of Visualized Graphs	217
4.3.6	Combination of Applied Models.....	226
	References	229
Conclusion	238
	References	243

List of Figures and Tables

Figures

Introduction

Figure 0.1. The need for expert recommendation in the academic sphere.6

Chapter 1

Figure 1.1. Characteristics of a community of practice. Figure adapted from Wenger (2008, p. 73).21

Figure 1.2. Stages of community development. Figure from Wenger, McDermott, and Snyder (2002, p. 69).....22

Figure 1.3. The duality of participation and reification. Figure adapted from Wenger (2008, p. 63).23

Figure 1.4. Interrelations between human characteristics. Figure from Stock and Stock (2013, p. 52).....30

Figure 1.5. Relations in the academic environment and support through community awareness.38

Chapter 2

Figure 2.1. CORDIS research partner search, retrieved 06/21/2016 from <https://cordis.europa.eu/partners/web/guest>.49

Figure 2.2. ArnetMiner/AMiner. The service offers a search for scientists and shows similar authors. Users have to sign up. Retrieved 06/21/2016 from <https://aminer.org/profile/tamara-heck/53f31a07dabfae9a8442f298>50

Figure 2.3. Single view of a surrogate in Web of Science with links to citations and related records. Retrieved 06/21/2016 from <http://apps.webofknowledge.com/>.....56

Figure 2.4. Scopus search results page with links to citations and references. Indirect relations are searchable by selecting items in the results list.	57
Figure 2.5. Scopus reference view: Same reference, formatted differently in the reference lists of two articles (doi:10.1515/libri-2013-0024 // doi:). Retrieved 06/21/2016 from https://www.scopus.com/	58
Figure 2.6. Source for determining implicit social information about a target researcher.	62
Figure 2.7. Relations between two authors considered for expert recommendation.	63
Figure 2.8. CiteULike user profile with bookmarks and tags (top). The bottom screenshot shows related articles (plus tags and other users) found via a user’s tag.....	69
Figure 2.9. Distinction between diverse tag types adapted from Sen et al. (2006).	72

Chapter 3

Figure 3.1. The principle of content-based filtering, where a target user gets recommendations of items that have similar content features.....	92
Figure 3.2. Car recommender system “myproductadvisor”, in which users are able to determine preferred characteristics as a basis for recommendations. Retrieved 06/21/2016 from http://www.myproductadvisor.com	98
Figure 3.3. The principle of user-based collaborative filtering, where a target user gets recommendations for unknown items from nearest neighbors who have purchased the same products.....	105
Figure 3.4. The principle of item-based collaborative filtering, where a target user gets recommendations for unknown items that are similar to those he or she already knows.	106

Figure 3.5. A personal recommendation page on CiteULike. A user needs 20 bookmarks as a basis for any recommendations. He or she is able to select item- or user-based recommendations. http://www.citeulike.org/profile/tamarah/recommendations	109
Figure 3.6. Recommendation variants in a tagging system. Figure adapted from Balby Marinho et al. (2012, p. 9).	122
Figure 3.7. Types of recommender system predictions in evaluation processes. Figure adapted from Jannach et al. (2011, p. 171).	139

Chapter 4

Figure 4.1. Power law distribution of bookmarks (n=13,608) (y-axis) per user (unique users: 2,437).	163
Figure 4.2. Distribution of common bookmarks (y-axis) for the 6,491 user-user pairs.	166
Figure 4.3. Distribution of similarity values for all 6,491 user-user pairs, ordered by Dice.	169
Figure 4.4. Similar users for “dchen” based on Dice (above) and on cosine coefficient (below). Graph generated with Gephi, algorithm: Force Atlas (see section 4.3).	172
Figure 4.5. Distribution of common tags for the first 1,000 user-user pairs.	173
Figure 4.6. Distribution of common bookmarks and common tags for all the top 141 user-user pairs who have the most bookmarks in common.	175
Figure 4.7. Single-link cluster for target researcher a, with a cosine coefficient threshold of <math><0.1</math>.	177
Figure 4.8. Complete-link cluster of target researcher a, with a cosine coefficient threshold of <math><0.1</math>.	178
Figure 4.9. Concept of author recommendation based on different sources of social information.	184

Figure 4.10. Scopus export of citing resources, including references that cite at least one of the target author’s publications. www.scopus.com.....	187
Figure 4.11. Author network based on bibliographic coupling (BICO) in Web of Science for scientist “R. Zorn” (a10), cosine coefficient threshold 0.2.....	199
Figure 4.12. Author network based on author co-citation (ACC) network of author “Zorn”, cosine coefficient threshold 0.05.	200
Figure 4.13. Author network based on CiteULike common users (CULU) for researcher “Zorn”, cosine coefficient threshold 0.49-0.99.....	201
Figure 4.14. Author network based on common tags in CiteULike (CULT) for researcher “R. Zorn” (a10), cosine coefficient threshold 0.49-0.99.....	202
Figure 4.15. Number of recommended authors and their valuation of each target researcher, where the number shown along the x-axis legend is the total number of recommended authors per target researcher.....	209
Figure 4.16. Distribution of relevant authors for all scientists for the three services Web of Science (WoS), Scopus and CiteULike, respectively.	213
Figure 4.17. Coverage of relevant authors per target scientist (a1-a10), where the number shown on the x-axis is the total number of relevant authors named by the target researchers.....	214
Figure 4.18. Coverage (similar to precision) of relevant authors per target scientist (a1-a10) for the top 20 recommended authors according to the cosine coefficient.....	215
Figure 4.19. Relevance feedback (rating of 1-10) of visualized author networks.	218
Figure 4.20. Author network of author a3 based on common tags in CiteULike (CULT) (extract of main cluster), cosine coefficient threshold 0.45-0.99, network rating = 9.	221
Figure 4.21. Author network of author a7 on the basis of common tags in CiteULike (CULT), network rating = 1.	222

Figure 4.22. Author network for author a6 on the basis of author co-citations (ACC) (extract of main cluster), with keywords assigned by a6 during the evaluation.	224
Figure 4.23. Tag cloud containing all tags assigned to the target researchers' bookmarked publications in CiteULike.	225

Tables

Chapter 2

Table 2.1. Search fields in CiteULike. Retrieved 06/21/2016 from www.citeulike.org/search_help	68
---	----

Chapter 3

Table 3.1. Tasks of recommender systems, adapted from Herlocker et al. (2004) and Ricci et al. (2011).	91
---	----

Chapter 4

Table 4.1. Chosen journals (abbreviations) for the data set.	161
Table 4.2. Number of bookmarks and users for all three bookmarking services.	162
Table 4.3. Similarity between target user “dchen” and 30 similar users, based on shared bookmarks, ordered by Dice coefficient.	171
Table 4.4. Similarity between target user “dchen” and the 30 most similar users, based on common tags, ordered by Dice.	174
Table 4.5. Basis for generated data sets: numbers of publications from 2006-2011/2012 per target author, numbers of their surrogates found in Scopus and Web of Science, numbers of citing resources in Scopus, and numbers of citations derived via Scopus csv files.	186

Table 4.6. Basis for generated data set in CiteULike: number of publications found, number of bookmarked publications, number of tags (all tags and unique tags) assigned to target authors' publications, number of users (all users and unique users) who bookmarked.....	195
Table 4.7. Questions concerning the top recommended authors in evaluation part b) (original questions in German).	205
Table 4.8. Questions concerning the visualized author graphs in evaluation part c) (original questions in German).....	206
Table 4.9. Evaluation values: novelty based on recommended author lists. Recall and precision values (based on named relevant authors who are appropriate collaborators) are measured according to the top 20 list ranked by cosine coefficient.	216
Table 4.10. Comparison: modified cosine and cosine values and ranks for the 16 important authors for a7.	227

List of Image Sources

Person (colors modified) in figures 0.1, 2.6, 2.7, 3.1, 3.3, 3.4, 3.6, 3.7, 4.9. By GNOME icon artists (HTTP / FTP) [CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>) or LGPL (<http://www.gnu.org/copyleft/lgpl.html>)], via Wikimedia Commons. Retrieved 06/10/2016 from https://upload.wikimedia.org/wikipedia/commons/0/0a/Gnome-stock_person.svg

Image in figure 2.9. Own image

Shoes in figures 3.1, 3.3, 3.4. Retrieved 06/10/2016 from <https://openclipart.org/detail/215041/shoes-lineart>

Sneaker shoe in figures 3.1, 3.3, 3.4. Retrieved 06/10/2016 from <https://openclipart.org/detail/1735/gym-shoe>

Camera in figures 3.1, 3.3, 3.4. Retrieved 06/10/2016 from <https://openclipart.org/detail/1586/digital-camera>

Jacket in figures 3.1, 3.3, 3.4. Retrieved 06/10/2016 from
<https://openclipart.org/detail/11220/various-clothing>

Printer in figures 3.3, 3.4. Retrieved 06/10/2016 from
<https://openclipart.org/detail/24819/printer>

Computer in figure 3.7. Retrieved 06/10/2016 from
<https://openclipart.org/detail/169006/computer>

Introduction

Collaboration and communication are the basis of a researcher's scientific craft. During a scientific career a scientist's peer network grows – he or she gets to know other researchers, with whom they exchange interests and work on common projects. Often, new contacts are established through already known peers, like one's advisor or local colleagues. A researcher's network is influenced by his or her existing environment. Without exchange between researchers, research would not be possible. Interaction processes between scientists foster knowledge creation and sharing (see Suorsa & Huotari, 2014). But closed environments – in most cases they are unintended – can hinder research development. The challenge is to be open to enable new connections to foster one's own and other's research. However, openness can best be practiced if a researcher notices his or her opportunities within research networks.

New developments and techniques seem to make collaboration and communication easier. The Web 2.0 (O'Reilly, 2005) offers new opportunities to share and exchange resources. But these do not come as a matter of course. A big challenge is the problem of information overload (e.g. Borchers, Herlocker, Konstan, & Reidl, 1998). Here, support is needed to facilitate collaboration and communication, which drive scientific craft and innovation. Information overload hides collaboration and communication options. Therefore, it is important to show researchers relevant resources and potential relations. New approaches like recommender systems offer solutions for this support.

Seeking for recommendations is natural among human beings. If one needs advice, a good friend and trustful person may give recommendation and helps making the right decision (e.g. Resnick & Varian, 1997; Gärtner, 2012). Recommendation systems pursue the same goal: To help finding the right resources and to help deciding between diverse resources – were challenges have minor, but not less relevant differences depending on a concrete task and a system's environment (see Ricci, Rokach, & Shapira, 2011). These systems need to have the right information about the user to be able to support him or

her. Using the right information, filtering it and applying it depending on a researcher's need is within the scope of this work, which concentrates on the task to find relevant peers in the academic field – not just any who happen to be available but the “right” ones. It investigates the scientific environment and relations between researchers to support scientists in finding their peers and foster communication and collaboration.

State of Research

Recommender system research began in the 1980s (Klahold, 2009). One aim of these systems is to order and filter the huge amount of user-generated data that arose alongside new developments in Web 2.0 (O'Reilly, 2005; Peters, 2009). Research on recommender systems adapts methods applied in information retrieval and filtering systems. However, the purpose of recommender systems lies in automated personalized recommendation (Ricci et al., 2011). Another focus is on the recommendation of new and unknown items. A retrieval system aims to give good results based on users' information needs. Thus, such a system “has to translate their information need into a query” (Baeza-Yates & Ribeiro-Neto, 2011). However, here is a major point of concern, as all searches begin with a user. A formulation of an information need – especially if it is problem-oriented and not clearly definable – can be difficult for a user: “To formulate clearly and precisely, one would have to know what one does not know” (Stock & Stock, 2013, p. 107). If a user knows what to search for, he or she still searches on the basis of their pre-knowledge and skills. Other perspectives are not considered. A recommendation system tries to help users in this way, as it takes historic user data and from this, tries to derive potential user interests and needs. The idea is to make users aware of things that they would not have found on their own. The second aspects is that such services aim to overcome information overload. “[Recommender systems] are primarily directed towards individuals who lack sufficient personal experience or competence to evaluate the potentially overwhelming number of alternative item[...].” (Ricci et al., 2011, pp. 1–2). Such a system for the academic sphere could help researchers find relevant resources and experts. Approaches are

based on relations between objects, which are implicit for the users, and make them explicit and visible. The approaches applied and techniques used are manifold and concentrate on diverse user needs. Recommendations for the academic field concentrate on literature recommendation, where there are various approaches. Approaches towards finding experts, for example, combine text retrieval models with citation analysis and therefore use online databases like DBLP¹ (Deng, King, & Lyu, 2008) (see also Renugadevi, Geetha, Gayathiri, Prathyusha, and Kaviya (2014). These approaches are derived from the field of expertise retrieval (Balog, 2012), which has a slightly different focus (see chapter 3). Furthermore, there exists research considering user-generated data from collaborative information services such as bookmarking services (Au Yeung, Noll, Gibbins, Meinel, & Shadbolt, 2009) and Twitter² (Saito & Yukawa, 2011). Bogers (2009), who discusses recommender systems based on social bookmarking data, aims at recommending relevant scientific literature to users (see also Bogers and Van den Bosch (2008)). McNee (2006) proposes literature recommendation based on citation data by researchers. His focus lies on users' acceptance of recommendations, which are evaluated via a survey, and on different algorithmic comparisons. The first aspect is also a focus of this work, while the comparisons of diverse algorithms for similarity measurement are not discussed. This work consults professional information services for measurements on citation data. In addition, it draws on another scientometric approach that is adapted for a recommender system.

An important aspect in providing researchers with valuable recommendations is reputation. Recommendations of potential collaborators rely on accurate information about them. In their study concerning collaboration in a scientific environment, Hara, Solomon, Kim, and Sonnenwald (2003, p. 957) find that "respondents typically considered acceptance as a scientist and intrinsic recognition that one's knowledge is valued as prerequisites for collaboration."

¹ <http://dblp.uni-trier.de>

² <https://twitter.com/>

One valuable source of information about researchers are their publication and citation data. In another study, Farooq, Ganoë, Carroll, and Giles (2007) suggest support strategies to enhance collaboration in online communities. They conducted a survey asking users with whom they would prefer to collaborate. Co-citation and references were important aspects for the users. Another good indicator that someone may be an appropriate collaborator is if they search for and read similar papers. Farooq et al. (2007, p. 3) conclude: “potential collaborators can share common ideas that focus on the papers they look for or cite.” Both aspects are picked up on in this work. On the one hand, citation data is gathered, while the aspect of reader similarity is considered via users’ bookmarking behavior. Farooq et al. (2007) propose detecting relations based on social network analysis and weak ties. This work regards diverse perspectives for detecting new and hidden researcher relations to foster collaboration.

The field of scientometrics concentrates on scientific information analyses such as productivity (published documents), paper topics (word and co-word analyses), and reader reception and formal communication (via references and citations) (Stock, p. 447). Diverse metrics such as co-authorship, citation, co-citation and bibliographic coupling analysis are applied to map-scientific fields (Cronin, Shaw, & La Barre, 2003; Leydesdorff, 1998; Schneider & Borlund, 2007a, 2007b). Analysis should provide a picture of actual patterns of researcher collaboration and communication (Ardanuy, 2012; Gazni, Sugimoto, & Didegah, 2012; Velden, Haque, & Lagoze, 2010) and of how a research field develops over time, based on researchers’ activities in terms of publication and citation. Scientometrics also discusses similarities between papers and authors (Leydesdorff, 2005, 2008), but there are few examples (Blazek, 2007; Guns & Rousseau, 2013, 2014) where such data is used for recommendations. The aim of mapping scientific fields is to give a realistic picture – more specifically, *the* true picture – of scientific communication and field development (Price, 1986; Small & Sweeney, 1985; Small, Sweeney, & Greenlee, 1985). Expert recommendation aims at personalized partner suggestions for a target researcher as these have the greatest value. Similarity measurements that apply data from a

large scientific field cannot show relations of personal relevance for a researcher. Recent discussions consider different perspectives on standard metrics. So-called “altmetrics”, or alternative metrics, are applied to measure the impact of researchers and journals (Priem & Hemminger, 2010), for instance. One approach is to base a journal’s impact not only on a single factor (such as the impact factor), but to consider user-generated data from bookmarking services (Haustein, 2012; Haustein & Siebenlist, 2011). This work goes in a similar direction, but uses scientometric approaches to expand the scope of vision of a target researcher with regard to his scientific collaboration network.

The need to recommend potential new collaborators does not derive from the fact that researchers do not collaborate. In fact, collaboration has grown immensely over the last decades (Cronin et al., 2003; Luukkonen, Persson, & Sivertsen, 1992; Persson, Glänzel, & Danell, 2004). The need to recommend collaborators derives from the fact that a researcher’s scientific craft depends on interaction with colleagues and can best unfold in realms where communities of practice are able to develop (Wenger, 2008). When researchers become aware of interaction potentials, they are able to participate and engage in their community and develop their own scientific work as well as the work of their colleagues. The purpose of a recommender system is to suggest new and as-yet undetected aspects that they would not have noticed themselves. “The task is to identify such groups and help them come together as communities of practice” (Wenger & Snyder, 2000, p. 144). Thus it is proposed to use such a system to make researchers aware of their colleagues in order to foster scientific collaboration. This work considers three different perspectives of expanding a researcher’s own view on his potentials.

Scope and Research Questions of this Work

Finding research partners for collaboration is a common task for any scientist. When a researcher has found the right scientific community, he or she interacts with fellow scientists in developing his or her research. It can be said that researchers act in communities of practice (Wenger, 2008) that offer a realm for

their scientific craft. However, collaboration is not as easy as it may seem (Hara et al., 2003). First and foremost, researchers need to be aware that their communities exist. It would seem that more collaboration would be possible if there were services for detecting such communities. Nowak and Wurst (2004, p. 244) speak of “reflective awareness”, which builds on a basis of previously existing information to facilitate the detection of new information to be used by all members.

The aim of this work is to analyze and evaluate a proposal for an expert recommender system for researchers, while taking into consideration scientometric and alternative approaches (figure 1.1). The focus here lies on the usage of appropriate data and the application of new expert recommender models. The evaluation analyzes the usefulness of the models based on target user feedback. The work will address neither aspects concerning the implementation of an operative recommender system, nor technical issues with regard to the automation of the proposed model. It concentrates on the evaluation of the results derived from the experimental model and suggests aspects to be considered for future implementation.

Regarding the need to develop further approaches to recommend research experts, this work addresses the following research questions:

1. Can researchers be supported in finding relevant experts for collaboration with the help of expert recommendation on the basis of

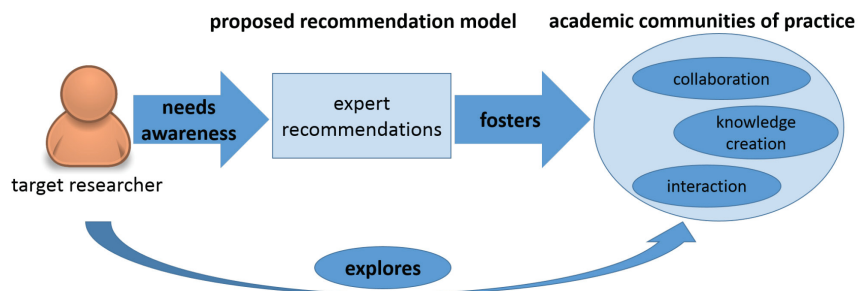


Figure 0.1. The need for expert recommendation in the academic sphere.

- scientometric and alternative approaches?
2. Is professionally indexed data from information services, based on scientometric approaches, appropriate for use in expert recommendation?
 3. Is user-generated data from social bookmarking services, based on collaborative filtering models, appropriate for use in expert recommendation?
 4. Are there any differences in outcome relative to the approaches and the datasets?

Research Methodology

To answer these research questions, an evaluation is conducted based on qualitative interviews. The focus lies on a target researcher who is looking to receive valuable results. Personal opinions and relevance feedback can only be derived from direct user statements (Berendsen, De Rijke, Balog, Bogers, & Van den Bosch, 2013; McNee, 2006). The focus on a target researcher implies that correct data is available for him or her. Thus, different datasets are manually derived from three sources in order to base the resulting approaches on the most accurate data possible. The evaluation outcomes incorporated as direct researcher feedback in the results on expert recommendations is gained via the evaluation interviews.

The work emphasizes a qualitative user study in which expert recommendations and data construction are based on data for target researchers. To this end, the evaluation was conducted in collaboration with scientists from *Forschungszentrum Jülich*. Ten physicists participated in the main study and evaluated their personal recommendations in qualitative semi-structured interviews. The focus of the evaluation lies on the usefulness gleaned from the recommendations and their representation.

Outline of this Work

The work is organized in four main chapters:

Chapter 1

This chapter discusses the general perception of collaboration and communication among researchers and introduces a broader view on this aspect. The change of collaboration among researchers is discussed, as well as the main reasons for collaborating and the direct advantages for scientists. Many activities of a researcher depend on communication and exchange with others. Collaboration and communication lead to rewarding benefits for science, and thus for researchers. The synergy of these will be described with the principle of communities of practice, which offer a realm where collaboration and communication gets possible. Next, the concept of academic knowledge creation as a researcher's scientific craft will be introduced to understand the need to cultivate communities of practice as a creative realm for scientific collaboration.

Chapter 2

Considering the cultivation of communities of practice, this chapter discusses the expansion of a researcher's perspectives regarding his or her scientific network. Expanding perspectives is essential to foster broader collaboration as it offers new opportunities for researchers to find appropriate peers. The chapter introduces the main social information services, which contain citation data revealing researcher relations, as well as bookmarking systems, which offer new data based on user-generated content. Citation and user-generated data are discussed to be able to detect implicit and explicit relations between scientists and thus help expanding a researcher's perspective. The introduced perspectives based on researcher relations are discussed regarding possible recommendation approaches that serve as foundation of the conducted case studies in chapter 4.

Chapter 3

This chapter introduces the recommender system task and gives an overview of recommender approaches with a focus on collaborative filtering concepts and expert recommendation in tagging systems. Recommender systems established to help people find the right resources based on personalized interests and perceptions. Depending on the task, information about users and resources is

needed to satisfy a target user and offer him the right recommendations. Recommender system research distinguishes three main recommender approaches, which will be introduced and compared, focusing on the principles of collaborative filtering as these are adapted for recommendation in the academic field. Explicit and implicit user ratings regarding their meaning and relevance as well as tag-based approaches are also introduced. After this overview, the focus lies on approaches that concentrate on expert recommendation and expert retrieval. The chapter concludes with principles of recommender system evaluation discussing user- and system-based approaches.

Chapter 4

This chapter focuses on the case studies. The main study, in section 4.3, is preceded by two pre-studies analyzing social bookmarking data structure as well as similarity metrics, and conducting first evaluation cases. The first pre-study discusses the structure of bookmarking data and the difference of similarity metrics regarding user recommendation. The second pre-study includes a first user evaluation of researcher recommendations based in first approaches on social information about researchers. Finally, the findings in these studies are then brought together in the main case study, which carries out the model for expert recommendation proposed in this work. The evaluation, done by physicists who received personalized recommendations, focuses on the relevance of recommended expert lists as well as visualized networks that show relations between scientists. The last part discusses possible combinations of the proposed approaches.

Acknowledgements

This work could not have been done without the support and collaboration of my colleagues and friends. Many thanks for your advice, support and collaboration. I would particularly like to thank Oliver Hanraths for helping me gathering the research data, and Paul Becker for proofreading my work. Many thanks to my advisor Isabella Peters for her support and fruitful discussions. Many thanks to my doctoral father Wolfgang G. Stock for supporting me and

convincing me to carry on. Thanks to Jens Terliesner for supporting me in any technical issues. I would especially like to thank Tobias Siebenlist, who not only helped me with any difficult problem I had, but also supported me as a friend. Likewise thanks to Lisa Orszulok for her tiny elephantastic attentions, which made me smiling and keep going on. In particular I would like to thank my officemate Sarah Hartmann for her patience, for helping me with my software fights, for teaching my seminars in the last phase of writing, and most of all for enduring my moods and quirks.

At last, thanks to my parents Ulrike and Manfred for letting me go my own way, and thanks to my brother Alexander, my best friend Nicole, my godchild Laura Marie and my other family members and friends for cheering me up and not forgetting me far-off the Hunsrück Mountains.

Formal Remarks

In the following work, the term “scientometric approach” refers to approaches conducted for similarity measurements, which are common standards in the field of scientometrics. Alternative approaches refer to methods that use alternative metrics with regard to researcher similarity. Thus, new collaborative filtering approaches are seen as alternative approaches to standard scientometric models. Collaborative filtering is understood as one of many models that are applied in recommender system research (Resnick & Varian, 1997). Additionally, this work has a focus on expert recommendation for researchers in the academic field, where academic researchers are defined as researchers who publish their scientific work. They may work in scientific institutions or universities, or in knowledge-intensive companies. Researchers who do not publish their work cannot be considered in the proposed model as recommendations exclusively rely on a scientist’s publications. The term “expert” and “expertise” is used throughout the work in reference to the diverse approaches described and their definition of the terms. Similarly, the proposed model defines “expertise” with reference to a researcher.

All parts of this work have their own list of references. If not indicated

otherwise, sources cited in the references sections and in other parts of this work have last been accessed on June 10, 2016.

References

- Ardanuy, J. (2012). Scientific collaboration in library and information science viewed through the Web of Knowledge: the Spanish case. *Scientometrics*, *90*(3), 877–890. doi:10.1007/s11192-011-0552-1
- Au Yeung, C. M., Noll, M. G., Gibbins, N., Meinel, C., & Shadbolt, N. R. (2009). On measuring expertise in collaborative tagging systems. In *Proceedings of the Web Science Conference: Society On-Line*. Retrieved from <http://www.websci09.org/proceedings/>
- Baeza-Yates, R., & Ribeiro-Neto, B. (2011). *Modern Information Retrieval: The Concepts and Technology Behind Search*. New York, NY, USA: Addison Wesley.
- Balog, K. (2012). Expertise Retrieval. *Foundation and Trends in Information Retrieval*, *6*(2-3), 127–256. doi:10.1561/15000000024
- Berendsen, R., De Rijke, M., Balog, K., Bogers, A. M., & Van den Bosch, A. (2013). On the assessment of expertise profiles. *Journal of the American Society for Information Science and Technology*, *64*(10), 2024–2044. doi:10.1002/asi.22908
- Blazek, R. (2007). *Author-statement citation Analysis Applied as a Recommender System to Support Non-domain-expert Academic Research: Doctoral dissertation*. Nova Southeastern University. Retrieved from http://nsuworks.nova.edu/gscis_etd/416
- Bogers, A. M. (2009). *Recommender Systems for Social Bookmarking: Doctoral dissertation*. Retrieved from <http://ilk.uvt.nl/~toine/phd-thesis/>
- Bogers, A. M., & Van den Bosch, A. (2008). Recommending scientific articles using CiteULike. In *Proceedings of the 2nd Conference on Recommender Systems (RecSys)* (pp. 287–290). New York, NY, USA: ACM. doi:10.1145/1454008.1454053
- Borchers, A., Herlocker, J., Konstan, J., & Reidl, J. (1998). Ganging up on information overload. *Computer*, *31*(4), 106–108. doi:10.1109/2.666847
- Cronin, B., Shaw, D., & La Barre, K. (2003). A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology*, *54*(9), 855–871. doi:10.1002/asi.10278
- Deng, H., King, I., & Lyu, M. R. (2008). Formal models for expert finding on DBLP bibliography data. In *Proceedings of the 8th International Conference on Data Mining (ICDM)* (pp. 163–172). New York, NY, USA: IEEE.

- doi:10.1109/ICDM.2008.29
- Farooq, U., Ganoe, C., Carroll, J., & Giles, C. (2007). Supporting distributed scientific collaboration: Implications for designing the CiteSeer collaboratory. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences, HICSS'07* (p. 26). IEEE.
doi:10.1109/HICSS.2007.506
- Gärtner, K. (2012). *Analyse von Recommendersystemen in Deutschland. B.I.T. online - Innovativ: Vol. 38*. Wiesbaden, Germany: Dinges & Frick.
- Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. *Journal of the American Society for Information Science and Technology*, 63(2), 323–335.
doi:10.1002/asi.21688
- Guns, R., & Rousseau, R. (2013). Predicting and recommending potential research collaborations. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.): *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference (ISSI)* (pp. 1409–1418). Wien, Austria: Facultas Verlags- und Buchhandels AG. Retrieved from <http://www.issi2013.org/proceedings.html>
- Guns, R., & Rousseau, R. (2014). Recommending research collaborations using link prediction and random forest classifiers. *Scientometrics*, 101(2), 1461–1473. doi:10.1007/s11192-013-1228-9
- Hara, N., Solomon, P., Kim, S.-L., & Sonnenwald, D. H. (2003). An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, 54(10), 952–965.
doi:10.1002/asi.10291
- Haustein, S. (2012). *Multidimensional Journal Evaluation: Analyzing Scientific Periodicals beyond the Impact Factor* (1. Aufl.). Knowledge and Information. Berlin/Boston: De Gruyter Saur.
- Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, (5), 446–457.
doi:10.1016/j.joi.2011.04.002
- Klahold, A. (2009). *Empfehlungssysteme: Recommender Systems - Grundlagen, Konzepte und Lösungen*. Wiesbaden, Germany: Vieweg + Teubner in GWV Fachverlage.
- Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, 43(1), 5–25.
doi:10.1007/BF02458391
- Leydesdorff, L. (2005). Similarity measures, author cocitation analysis, and information theory. *Journal of the American Society for Information Science and Technology*, 56(7), 769–772. doi:10.1002/asi.v56:7
- Leydesdorff, L. (2008). On the normalization and visualization of author co-

- citation data. Salton's cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology*, 59(1), 77–85. doi:10.1002/asi.v59:1
- Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology & Human Values*, 17(1), 101–126. doi:10.1177/016224399201700106
- McNee, S. M. (2006). *Meeting User Information Needs in Recommender Systems: Doctoral dissertation*. Minnesota, MN, USA: University of Minnesota. Retrieved from <http://www-users.cs.umn.edu/~mcnee/mcnee-thesis-preprint.pdf>
- Nowak, J., & Wurst, M. (2004). Supporting knowledge creation and sharing in communities based on mapping implicit knowledge. *Journal of Universal Computer Science*, 10(3), 235–251. doi:10.3217/jucs-010-03-0235
- O'Reilly, T. (2005). *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. Retrieved from <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>
- Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421–432. doi:10.1023/B:SCIE.0000034384.35498.7d
- Peters, I. (2009). *Folksonomies. Indexing and Retrieval in Web 2.0*. Berlin, Germany: De Gruyter Saur.
- Price, D. J. d. S. (1986). *Little Science, Big Science ... and Beyond*. New York, NY, USA: Columbia Univ. Press.
- Priem, J., & Hemminger, B. H. (2010). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*, 15(7). doi:10.5210/fm.v15i7.2874
- Renugadevi, S., Geetha, T. V., Gayathiri, R. L., Prathyusha, S., & Kaviya, T. (2014). Collaborative search using an implicitly formed academic network. *Aslib Journal of Information Management*, 66(5), 537–552. doi:10.1108/AJIM-12-2013-0142
- Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56–58. doi:10.1145/245108.245121
- Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 1–35). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_1
- Saito, J., & Yukawa, T. (2011). Extracting user interest for user recommendation based on folksonomy. *IEICE Transactions on Information and Systems*, 94(6), 1329–1332. doi:10.1587/transinf.E94.D.1329
- Schneider, J. W., & Borlund, P. (2007a). Matrix comparison, Part 1: Motivation

- and important issues for measuring the resemblance between proximity measures or ordination results. *Journal of the American Society for Information Science and Technology*, 58(11), 1586–1596.
doi:10.1002/asi.20643
- Schneider, J. W., & Borlund, P. (2007b). Matrix comparison, Part 2: Measuring the resemblance between proximity measures or ordination results by use of the mantel and procrustes statistics. *Journal of the American Society for Information Science and Technology*, 58(11), 1596–1609.
doi:10.1002/asi.v58:11
- Small, H., & Sweeney, E. (1985). Clustering the science citation index using co-citations: I. A comparison of methods. *Scientometrics*, 7(3), 391–409.
doi:10.1007/BF02017157
- Small, H., Sweeney, E., & Greenlee, E. (1985). Clustering the science citation index using co-citations: II. Mapping science. *Scientometrics*, 8(5), 321–340.
doi:10.1007/BF02018057
- Stock, W. G., & Stock, M. (2013). *Handbook of Information Science*. Berlin/Boston: De Gruyter.
- Suorsa, A., & Huotari, M.-L. (2014). Knowledge creation and the concept of a human being: A phenomenological approach. *Journal of the Association for Information Science and Technology*, 65(5), 1042–1057.
doi:10.1002/asi.23035
- Velden, T., Haque, A.-U., & Lagoze, C. (2010). A new approach to analyzing patterns of collaboration in co-authorship networks: mesoscopic analysis and interpretation. *Scientometrics*, 85(1), 219–242. doi:10.1007/s11192-010-0224-6
- Wenger, E. (2008). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, MA, USA: Cambridge Univ. Press.
- Wenger, E., & Snyder, W. (2000). Communities of practice? The organizational frontier. *Harvard Business Review*, 78(1), 139–145. Retrieved from <https://hbr.org/2000/01/communities-of-practice-the-organizational-frontier>

1 Scientific Collaboration and Communities

1.1 Collaboration between Scientists

An important task for researchers in academic settings and in knowledge-intensive companies is to find the “right” people who can work together to successfully solve a scientific or technological problem. During their professional life, researchers establish and expand their social professional network to include colleagues with similar and complementary research interests. Due to the fact that scientists require other people’s research labor to substantiate and defend their own work, they have to deal with the works of other colleagues. But this is not the only reason for collaboration. Further contact to other scientists is essential for one’s personal career and for science per se (Price, 1986), as Price and Beaver (1966, p. 1014) concluded in the 1960s: “The most prolific man is also by far the most collaborating [...]”

Research collaborations offer a scientific worker the option to improve his own scientific achievements and reputation, which is explicitly apparent in a researcher’s number of publications and citations. Collaboration is, in some cases, a predictor for a researcher’s productivity (Lee & Bozeman, 2005). Studies show that collaborations have a positive effect on a researcher’s or a research institution’s reputation. Guerrero Bote, Olmeda-Gómez, and De Moya-Anegón (2013) analyzed the reputation of scientific publications by examining the number of citations of scientific articles and the internationality of their authors based on author country information. The impact of a country is calculated via the number of citations of its publications, applying normalized citation weights to consider different citation behaviors in diverse scientific disciplines. The researchers conclude “[...] that international collaboration increases the impact of a country’s scientific production [...]” (Guerrero Bote et al., 2013, p. 402), independently of any national bias (see also Lancho Barrantes, Guerrero Bote, Rodríguez, & De Moya-Anegón, 2012). Further studies confirm that influential research centers with a high impact have more collaborations with researchers from other

institutions (Gazni, Sugimoto, & Didegah, 2012).

Besides this explicit factor, researchers gain more benefits from collaborations. Lee and Bozeman (2005, p. 693) talk of “spillovers beyond the publication of papers.” In fact, collaboration tends to be a factor that promotes and transmits “scientific and technical human capital” (Bozeman & Corley, 2004, p. 599). Collaboration bundles scientific competencies and resources as well as relevant knowledge (Stevens & Campion, 1994). This leads to enormous benefits for researchers, as they can exchange competencies with others and complement their own ones. It is not possible for a single scientist to hold all relevant resources and knowledge – he or she needs support from others. Explicit relevant resources to be exchanged among researchers are, for example, literature (including gray literature), data gleaned via experiments, laboratory samples, and technical equipment. The researcher him- or herself will profit from this collaboration, and finally their scientific work and science as a whole (Finholt, 1999; Kling & McKim, 2000) – of course one essential pre-condition being that these collaborations are fruitful. Hence, it is obvious for most researchers that collaborations with colleagues are part of their scientific working life. Additionally, and especially for young scientists like MA and PhD students, collaborations are an important step into the scientific community. These collaborations are also a confirmation for newcomers that they are accepted in this community (Hara, Solomon, Kim, & Sonnenwald, 2003).

Collaboration can be defined in different ways (see for example Kagan, 1991; Schrage, 1995). Mattessich and Monsey (1992, p. 11) define collaboration as a “mutually beneficial and well-defined relationship entered into by two or more organizations to achieve common goals.” The authors also distinguish between collaboration, cooperation, and coordination, which are often used as synonyms. Cooperation is a kind of informal relationship that aims at sharing information. However, it does not focus on a common mutual mission. Coordination is similar, but includes the latter aspect of common mission. In contrast, collaboration aims at a more “durable and pervasive relationship” (Mattessich & Monsey, 1992, p. 42) with a strong focus on personal commitment as well as shared and new

goals both parties are willing to reach together. Thus, collaboration requires a stronger commitment, which is more difficult to establish, but leads to results that are more valuable. Two crucial elements found in definitions of collaboration are the “working together for a common goal” and the “sharing of knowledge” (Hara et al., 2003, p. 953).

The number of collaborations between researchers has grown in recent years (Cronin, Shaw, & La Barre, 2003; Luukkonen, Persson, & Sivertsen, 1992; Persson, Glänzel, & Danell, 2004; Price & Beaver, 1966). This trend is determined by the numbers of co-authored published articles. However, Cronin et al. (2003) emphasize that there are major differences between disciplines. Their study analyzing co-authorships in psychology and philosophy showed that the latter discipline’s trend towards collaboration is far less distinct, even unobservable (Cronin et al., 2003). With these few exceptions, researchers in academic settings and in knowledge-intensive companies search for the “right” people with whom they can work together to successfully solve any scientific problem. The reasons for searching these partners are manifold. Scientists search relevant colleagues to

- advance the solution of a research problem,
- advance investigations in an upcoming research field,
- establish a (formal) working group in a large university department or company,
- bring together researchers to prepare a project proposal for a research grant (within and outside of the department and company),
- form a community of practice, independently of any affiliation or institution, following only shared interests (Wenger, 2008),
- accredit colleagues in preparation of a congress, a panel or a workshop,
- ask colleagues for contributions to a textbook or specialized journal issue,
- find appropriate co-authors for any scientific publication.

Researchers get to know each other in different ways. The interviewed scientists who participated in this work’s experimental studies stated that they met future collaborators at scientific conferences, through colleagues working in the same

institute as well as by contacting people who were already in the community network of their boss or PhD supervisor (Heck, 2012a). However, it is difficult – for young researchers as well as for experienced ones – to find collaborators for a specific purpose. This is the case with temporary tasks, like the preparation of a panel for a conference with a fixed submission deadline. Especially researchers with no experience do not find it easy to find an appropriate collaborator. Blazek (2007) calls them “domain-novice researchers”, who can be young or older, and are trying to dig into a specific research field for the first time. The former are simply too young to know their future scientific network. The latter are the established researchers, who are already integrated in a community. However, they might choose to change their core field of interest and thus have to negotiate their way in a new community. Besides this fact, established researchers look for new collaborations, especially interdisciplinary ones, to expand their scientific network.

As shown above, researchers do not only search for partners with the aim of co-writing scientific papers. Collaboration leads to benefits that are more fruitful and essential for science. The synergy of these can best be described via the principle of communities of practice and the concept of knowledge creation (Hara et al., 2003), to be introduced in the next section.

1.2 Communities of Practice for Academic Purposes

Communities of practice are groups of people who share the same interests, exchange information and knowledge, and collaborate with one another (Wenger, 2000). Three aspects are of importance: Firstly, members of communities of practice have a joint enterprise of what their community is about. Secondly, the community is built on mutual engagement, and thirdly, the members produce a shared repertoire of resources.

Research on communities of practice is manifold (Wenger, 2010). Bolisani and Scarso (2014), who review research on communities of practice in the field of knowledge management, state that this field is mostly concerned with Wenger’s concept, including aspects of organizational learning and knowledge creation in

companies (Bertels, Kleinschmidt, & Koen, 2011; Brown & Duguid, 1991; Corso, Giacobbe, & Martini, 2009; Davenport & Hall, 2002; Huysman, 2002; Huysman & Wit, 2002; Lesser & Storck, 2001; Pattinson & Preece, 2014; Ramchand & Pan, 2012; Swan, Scarbrough, & Robertson, 2002). Additionally, case studies are concerned with community building and fostering, while analyzing specific fields of employment, such as nursing (Valaitis, Akhtar-Danesh, Brooks, Binks, & Semogas, 2011), caregiving (Fenton et al., 2007), diverse other public health services (Mabery, Gibbs-Scharf, & Bara, 2013), digital humanities (Green, 2014) and librarianship (Henrich & Attebury, 2010). Hara (2009) intensively studied knowledge sharing and communication between public defense attorneys. His main concern was the exchange of knowledge and junior professionals' occupational learning. The learning and integration of apprentices was also one of Lave's and Wenger's (Lave & Wenger, 1991) starting points for the development of their concept of peripheral legitimate learning and, later, for the concept of communities of practice. Related to this is the notion of tacit and explicit knowledge (Nonaka & Takeuchi, 1995; Polanyi, 1967), which will be discussed later in this work. Hara (2009) points to the importance of the interaction between newcomers and oldtimers within a community that entails the gathering and sharing of knowledge.

Current approaches analyze the value of technical support for communities of practice, such as specific applications to facilitate communication and knowledge exchange (Fenton et al., 2007; Pan & Leidner, 2003). Some studies show that a high usage of IT alone does not strengthen a community, which is mainly due to a lack of identity formation of its members (Hara, 2007; Hara & Kling, 2002). However, changes in working and everyday life as well as new technologies like mobile solutions might offer new ways of community development (Kietzmann et al., 2013).

The interpretation of a community of practice might have changed over the years, as community structures are interpreted differently, which also depends on the respective field of research and a community's environment (see Cox (2005) and Murillo (2011) for detailed discussions). However, research in most cases

analyzes the main claims of Wenger's original concept. In the following, this concept is discussed with regard to academic communities of practice (mainly groups in schools and universities) and the most important aspects are summarized.

1.2.1 Principles of Communities of Practice

The principle of communities of practice was introduced by Etienne Wenger and Jean Lave 1991 while discussing the concept of "legitimate peripheral participation" (Lave & Wenger, 1991). The aim was to analyze and characterize the dynamic in an apprenticeship and the relations between masters and apprentices, or students. The researchers claim that learning is an integral and inseparable part of social practice. Questions involved the learning process and transmission of knowledge between masters and students. Legitimate peripheral participation indicates the process of entering a community of practice. 'Peripheral' here "[...] suggests an opening, a way of gaining access to sources for understanding through growing involvement" (Lave & Wenger, 1991, p. 37). It guarantees full participation. 'Legitimacy' suggests that new members are accepted in a community and that they get the chance to participate actively. Wenger (2008, p. 101) gives an example of legitimacy referring to academics:

"Today, doctoral students have professors who give them entry into academic communities. Granting the newcomers legitimacy is important because they are likely to come short of what the community regards as competent engagement."

A community of practice should therefore create situations in which legitimate peripheral learning is possible. Referring to academic communities, legitimate peripheral learning as a process means the entry and future acceptance of researchers in their scientific community. Practice is here seen as a learning process that changes constantly and has continuities and discontinuities (Wenger, 2008, p. 49). Its history includes three main characteristics (figure 1.1) (Wenger, 2000; Wenger, 2008; Wenger, McDermott, & Snyder, 2002).

The first characteristic is mutual engagement. It should not be understood as a harmonic coexistence – rather, the opposite is the case. Wenger (2008, p. 75)

emphasizes that diversity often generates engagement in practice and even facilitates it in the first place. The social complexity of this characteristic is unlimited. Nevertheless, it strives for community maintenance. Concerning academic communities, this may be the most crucial characteristic. Scientific communities would not exist without diverse forms of engagement. Researchers develop new theories, prove and refute existing perceptions, and engage in the creation of new knowledge. Via mutual shared engagement, researchers push forward their scientific field.

The second characteristic of practice is joint enterprise, which is a result of mutual engagement. Joint enterprise is not to be seen as a number of fixed goals, but rather as an agreement to reach mutual accountability. It can be defined on the basis of mutual engagement in practice. The development of a joint enterprise lies in the hands of its members. They develop and determine conditions and constraints, which are either explicit or implicit. The idea of a community in a specific research field also includes a joint enterprise. The characteristics are different for each community. Common objectives may include their agreement to push forward their scientific field, or their agreement upon research topics included in this field.

The third characteristic is a shared repertoire, which is the source of a community

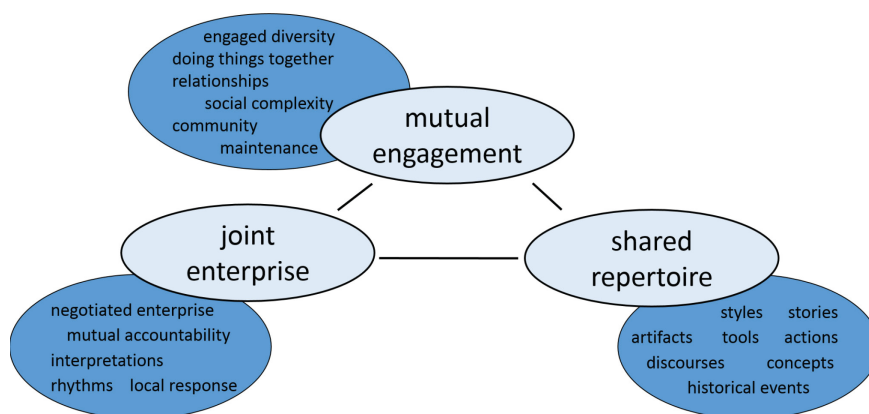


Figure 1.1. Characteristics of a community of practice. Figure adapted from Wenger (2008, p. 73).

and is used in its practice. ‘Repertoire’ refers to all actions performed in practice, including language, symbols, concepts, and tools (Wenger, 2008, p. 83). For an academic community, examples include terminology, evaluation techniques, associations, and conferences. The repertoire can be regarded as the most visible characteristic of a community for non-members.

Members in communities of practice engage on various levels. A community has a coordinator at the head of a central core group, a small active group, many members on the periphery, and outsiders who are not direct members (Wenger et al., 2002). The level corresponds to a member’s activities within the community. The participation level of each member changes and there are no fixed borders between those levels. For example, if the focus of a community changes, core members shift to the periphery and other members take their place.

Furthermore, communities of practice have no fixed structures. As members and their activities change constantly (Brown & Duguid, 1991), the community itself goes through various stages (figure 1.2). If there is potential, the community starts to develop, and generally, loose networks of people evolve into more connected ones. Members begin to share and develop common knowledge. The group’s energy level grows, as does its visibility, while more members participate and

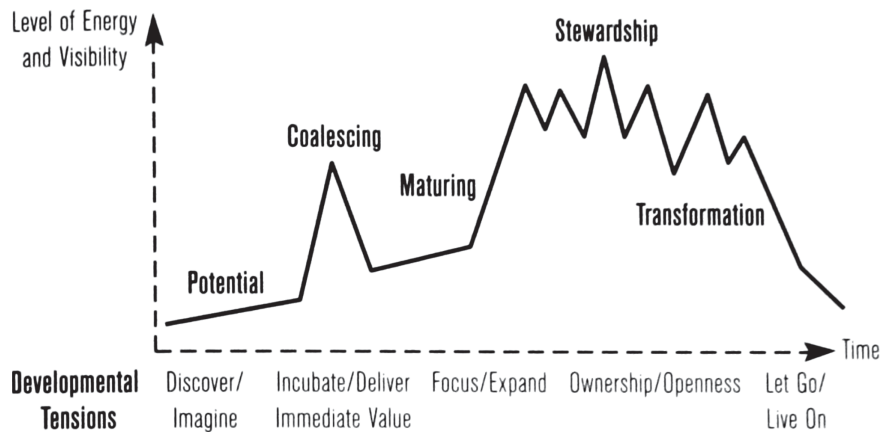


Figure 1.2. Stages of community development. Figure from Wenger, McDermott, and Snyder (2002, p. 69).

more knowledge is created (coalescing and maturing stage). However, a community does not rest its development at this stage. Even well-established communities go through critical changes and tensions. In the end, the community might either die or radically transform itself. According to Wenger et al. (2002), this is a natural process, but does not come without an emotional component. However, the fading of a community can be a chance to relive the same process and activate its engagement.

Communities of practice exist in many places. According to Wenger, each individual is a part of one or more communities. People form their identity via engagement in a community of practice, and even non-participation shows a form of identity (Wenger, 2008, p. 164). Participation is related to the concept of reification, which Wenger defines as a complement of participation. The border between both elements is fluid and often not explicitly distinguishable by the members, with both elements influencing the negotiation of meaning. More concretely, participation contextualizes meaning (Wenger, 2008, p. 63), which then flows into a reification process. A community's shared repertoire is part of its reification, which includes documents, instruments, and forms (figure 1.3). Both elements eliminate the other's misleading negotiations. Reification tries to

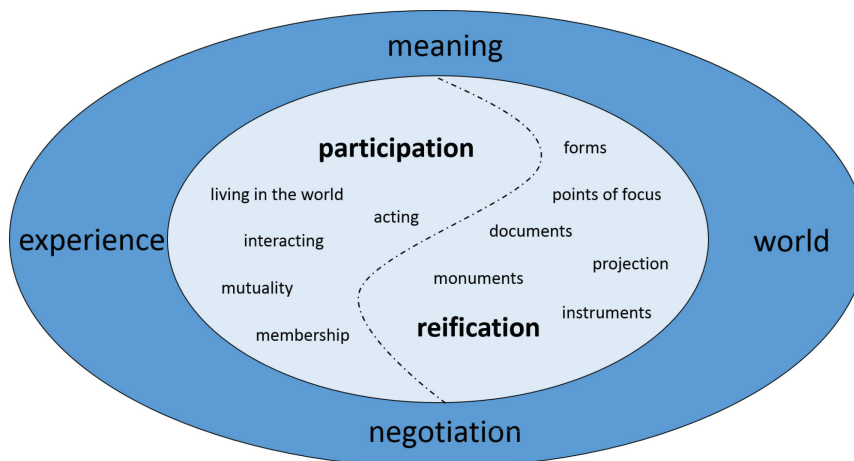


Figure 1.3. The duality of participation and reification. Figure adapted from Wenger (2008, p. 63).

make meaning tangible, but may have flaws. For example, written law (reification) is interpretable (done via participation), similarly to scientific work. That is why scientists not only publish articles, but also participate in conferences to make their written work clearer. Participation is needed to repair reification failures, which leads, for instance, to the adaptation of scientific models or even the adoption of new models. Conversely, reification is needed to make participation possible because it establishes a shared repertoire and makes processes tangible. Thus, participation and reification enable community members to interact with each other.

With these preconditions, communities of practice act as social constructs that foster learning and knowledge exchange and allow people to engage with one another (Wenger, 2008). Engagement occurs to different degrees and means collaboration in its widest sense. Community members develop a form of identity and a sense of belonging to their group. Identity is formed not only through individual preferences, but also through community participation and interaction in practice. While doing research, a scientist automatically interacts with his community, for example by citing other researchers or discussing at conferences – and all these activities form his own identity process and, finally, his work. His active participation contributes to and forms the community of practice and thus its knowledge, which becomes shared practice for all other members. More concretely put, participation enables the creation of knowledge (Lave & Wenger, 1991; Tobbell, O'Donnell, & Zammit, 2010). Knowledge of a community of practice becomes tangible via objects, more concretely, the processes of participation and reification in the practice within communities “leave a historical trace of artifacts – physical, linguistic, and symbolic – and of social structures, which constitute and reconstitute the practice over time” (Lave & Wenger, 1991, p. 58). Nistor and Fischer (2012) talk of domain knowledge within a community, including members' knowledge about academic research, publications, fundraising, teaching, collaboration and young researcher support. The scope of domain knowledge depends on the community and its objectives. The aim of a community here is to make its knowledge accessible to all members, through what Lave and Wenger call “legitimate peripheral participation”. New members, or

novice researchers, need to achieve this status of full participation, in order to become community experts. This suggests that participation and expertise may be interconnected and depend on one another (Nistor & Fischer, 2012; Tobbell et al., 2010). However, even newcomers can have expert knowledge and enrich a community (Fuller, Unwin, Felstead, Jewson, & Kakavelakis, 2007, 2012), although early assumptions by Lave and Wenger (1991) suggest that expertise is only gained in a community of practice, with new members learning from established ones as implicit knowledge is turned into explicit knowledge. Other researchers claim that participation in an academic community of practice is only possible with a minimum of domain knowledge (Brown, 2001; Tobbell et al., 2010).

Nevertheless, in both cases participation and expertise seem to be connected. Nistor and Fischer (2012) used a quantitative model to measure the correlation between expertise and participation. In their surveys, individuals working at two university institutions (scientists as well as technical and administrative staff) were asked about aspects of domain knowledge, participation, and contribution to “artifact development” (defined as assigning information to researcher profiles on the university’s web pages, such as publications). Results show that knowledge depends on participation and the individual’s expert status (as evaluated by other community members) is influenced by his expertise and participation. However, a high expert status does not mean high expertise. The expert status suggests that a person is well integrated into a community, while the expertise level gives insight into a person’s domain expertise. If there are discrepancies regarding both aspects, Nistor and Fischer (2012) suggest fostering knowledge sharing and interaction to better integrate persons with high expertise and a low expert status. Furthermore, results reveal that novice persons have a much lower expert status and participate less than experts, although their level of experience (measured in years) is similar to that of the experts. Thus, domain knowledge, expertise and participation depend on one another and influence the “artefact development” of an institution. The crucial point is to allow each member of a community to reach full participation in order to improve their expertise, and attain expert status. Wenger et al. (2002) speak of cultivating communities of practice relative to their

specific stages (figure 1.2). They name seven main principles of cultivation. The most important point for academic communities is to “open a dialogue between inside and outside perspectives” (Wenger et al., 2002, p. 54). This aspect emphasizes the notion of communities of practice as open social structures that are distinct from time-limited cooperation teams as seen in co-authorships, for instance. The full value of communities of practice evolves through long-term participation and interaction among their members. With their cultivating principle, Wenger et al. (2002) stress the importance of open groups, which are provided with new insights by outsiders and start a dialogue with them to foster the advancement of a community and the creation of new knowledge.

Communities of practice consist of intra- and extra-departmental employees of a company or institute, individuals who work at the same location or in different places or even people who do not belong to the same company. The important factor is that communities of practice establish and organize themselves. Members meet willingly and are not compelled by authority, as enforcement could lead to a refusal of collaboration (Blair, 2002). That is the main difference between such communities and teams established by company managers (Wenger & Snyder, 2000). In research environments, social academic communities start developing at research-intensive “hubs” like universities or research institutes (Tian, Nakamori, & Wierzbicki, 2009). In the best case, these small communities expand and open up to include people from different communities. Such communities and their interaction within the group and between diverse groups are vital for science, as these environments facilitate the exchange of existing knowledge and the creation of new knowledge. Thus, small social academic communities need to connect, or rather interact, with other scientific groups, both within their own research field and beyond, for the purpose of interdisciplinary exchange. The small communities evolve into bigger international communities of practice, which share a mutual engagement. Wenger et al. (2002) emphasize that those distributed communities have to overcome various boundaries. There are hard boundaries, such as geographical distances and different time zones, but also social boundaries, like different languages and cultural differences, which also occur within a single institution. Distributed communities tend to be less

present for their members (Wenger et al., 2002, p. 116), and are slightly invisible. Boundaries such as language differences increase this effect. Distributed and virtual online communities (Cassidy, 2011; Cheung, Lee, & Lee, 2013; Davenport, 2001; Hara, Shachaf, & Stoerger, 2009; Rosenbaum & Shachaf, 2010) differ from “face-to-face” communities as knowledge sharing might be more difficult and the exchange of implicit knowledge is not possible in the former. “Virtual meetings” might not be enough to replace face-to-face meetings between community members (Gust von Loh, 2009). As research on collaboration shows that academic communities gain value from collaborating internationally, problems arising from distributed communities hinder the exchange of knowledge, and thus academic knowledge creation (Bos et al., 2007), which is a researcher’s scientific craft.

1.2.2 Academic Knowledge Creation

Academic knowledge creation is a part of a scientist’s work, but more generally, knowledge creation is a part of what every human being does. The discussion of all facets of knowledge creation goes beyond this work, but it will briefly introduce the basic elements for academic knowledge creation.

Suorsa and Huotari (2014) give an overview of topics related to knowledge creation, including technology-based support, economic value and the huge field of knowledge management within companies. Important perspectives concerning a researcher’s activities and the meaning of collaboration for knowledge creation are derived from research, which discusses knowledge creation in relation to interactive processes. Knowledge creation come about via a process of interaction within communities (Jakubik, 2008, 2011; Nonaka & Takeuchi, 1995; Tsoukas, 2009). Research aims at a better understanding of concrete processes of knowledge creation with a consideration of the human factor and the structure of communities. Suorsa and Huotari’s hermeneutic discussion demonstrates a theoretical concept to further examine interactive processes as well as knowledge creation and sharing, which can serve as the basis for a deeper understanding of communities of practice (Suorsa & Huotari, 2014).

Interactions as a basis for knowledge creation are supposed to be enhanced in a community of practice. Lave and Wenger (1991, p. 34) state: “The generality of any form of knowledge always lies in the power of renegotiate the meaning of the past and future in constructing the meaning of present circumstances.” Lave and Wenger talk of “situated learning”, in which a human being creates knowledge in his specific context. An important aspect here is interaction between experience and competence (Wenger, 2008), which opens new ways for further development:

“On the one hand, a community of practice is a living context that can give newcomers access to competence and also invite a personal experience of engagement by which to incorporate that competence into an identity of participation. [...] On the other hand, a well-functioning community of practice is a good context to explore radically new insights without becoming fools or stuck in some dead end.” (Wenger, 2008, p. 214).

Conditions for establishing these realms are the tension between experience and competence as well as a respect for experience and a strong anchor of mutual engagement. This is the realm in which the acquisition and creation of knowledge can occur. Thus communities of practice, with their potential to develop knowledge (Choi, 2006), “provide the social context for individual interactions” (Jakubik, 2008, p. 6).

Academic knowledge creation, described in a simplified way, includes three basic elements: Hermeneutics, debate and experimentation (Wierzbicki & Nakamori, 2006a, 2006b). Wierzbicki and Nakamori discuss and combine diverse models of academic knowledge creation processes. Another famous model that shows knowledge creation and distribution is that of Nonaka and Takeuchi (1995) (also Nonaka, 1994), who propose the SECI model (socialization, externalization, combination and internalization) to describe the flow of tacit and explicit knowledge between individuals and groups. Wierzbicki and Nakamori (2006b) (see also (Wierzbicki & Nakamori, 2005) introduce what they call the “creative space model”, a kind of meta-model, which expands the SECI model. Hermeneutics, debate and experimentation as the three basic elements are summarized in the “triple helix” (Wierzbicki & Nakamori, 2006a) and form the

basis of knowledge creation in normal science (as defined by Kuhn, 1996). Hermeneutics includes all processes that are crucial to the development of new ideas, like searching, analyzing, comparing, and reflecting on results, while referring to scientific literature. Gadamer (1975) defines hermeneutics (not only relating to science) and stresses the essence of being aware that “the virtue of H(ermeneutics) always fundamentally consists in transposing a complex meaning from another “world” into one’s own” (Gadamer, 1974, p. 1061; translation from Stock & Stock, 2013, p. 50). The theory explains the complex interrelations between a human being and the world and the resulting conditions concerning knowledge creation and knowledge exchange. One explicit phenomenon is the individual use of language and the development of personal concepts, which is crucial for information science research (Stock & Stock, 2013). Stock and Stock (2013) show the complex interrelation between a human being and his world with reference to a holistic perspective on information science (figure 1.4). They also include aspects of commitment, which is likewise discussed in the concept of communities of practice (Wenger (2008) emphasizes ‘identification’ as a more fundamental process) and their members’ engagement (see for example McLure Wasko and Faraj (2000) and Wasko and Faraj (2005)).

Experimentation, or experimental testing in the “triple helix”, includes all processes concerned with the interpretation of results and their selection to develop new ideas (related in figure 1.4 to the cluster of “understanding documents”). Finally, debate includes all processes concerned with discussions on ideas in researcher groups, the advancement of results and options for developing new ideas on the basis of these group discussions (Wierzbicki & Nakamori, 2006b). The purpose of these models is to understand the processes a researcher goes through and his needs during these processes in order to be able to support his creative development. Wierzbicki and Nakamori (2006b) refer to computerized decision support systems. Recommendation systems also help users in making decisions and supporting them, for example in finding literature or experts, where both processes belong to either the hermeneutical or the debate

part of a researcher's knowledge creation process¹.

Studies show that researchers, especially novice researchers, have difficulties in some tasks concerning knowledge creation. Tian et al. (2009) conducted two surveys with novice researchers (master and doctoral students, post-docs and research assistants) asking them about their academic knowledge management and some of their most difficult tasks. The researchers aim to improve the creative environment in order to help people with their knowledge management tasks. Besides difficulties in seeking information, finding new ideas, and applying supportive IT tools, the participating students from a Japanese institute stated that they concentrate on self-study instead of collaborations with other researchers. The second survey, asking about any improvement the participants wished for in relation to their work, revealed that language differences impedes communication with international colleagues and tacit knowledge is hardly shared amongst them at all. The most crucial point for the students is the guidance

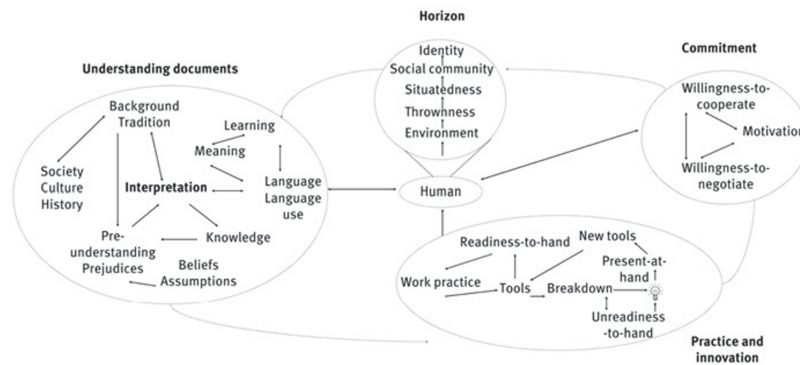


Figure 1.4. Interrelations between human characteristics. Figure from Stock and Stock (2013, p. 52).

¹ Note that research generally distinguishes between recommender and classic decision support systems, although both fields overlap (see for example Liang (2008); Malinowski, Weitzel, and Keim (2008)).

from a supervisor and frequent communication within a research group, which they rarely experienced during their work. To summarize this result:

“[...] a lot of responders work alone and also get much less timely encouragement and help from others at the same time; there is not good enough critical feedback, questions and suggestions in group discussions, and so on. [...] [But] for individual researchers, communication is not enough; they have to check their ideas through debate” (Tian et al., 2009, p. 87).

An important aspect here is knowledge sharing, which is supposed to be improved (Tian et al., 2009). However, not all knowledge is easy to share. Nonaka and Takeuchi (1995) speak of implicit (or tacit, see also Polanyi, 1967) and explicit knowledge while discussing their model of knowledge management. In their approach of explaining the sharing of knowledge within a company community, they introduce the four steps of socialization, externalization, combination and internalization. In these successive steps, knowledge changes and evolves from implicit to explicit and back again to implicit knowledge while it is shared amongst employees. Polanyi (1967, p. 8) coined the statement “one can know more than one can tell”, and refers to the fact that humans cannot express certain knowledge explicitly, but only have some notion of it that derives from one’s awareness of related factors. Thus tacit knowledge is not tangible, but crucially, it is the starting point for a human being creating knowledge. Polanyi (1967, p. 21) gives an example: “Therefore: a mathematical theory can be constructed only by relying on *prior* tacit knowing and can function as a theory only *within* an act of tacit knowing, which consists in our attending *from* it to the previously established experience on which it bears.” This also leads to the fact that a kind of objective knowledge cannot exist, but knowledge is always interrelated with personal commitment (Polanyi, 1982). For science itself, Polanyi (1967, p. 25) concludes: “To accept the pursuit of science as a reasonable and successful enterprise is to share the kind of commitments on which scientists enter by undertaking this enterprise”. The notion of Polanyi’s aspects, as well as Wenger’s and Lave’s theory of situated learning (Lave & Wenger, 1991), provide an understanding of the importance of knowledge sharing as the core basis for further knowledge

creation and the development of the validity of scientific knowledge. Considering the notion of tacit knowledge, knowledge sharing might not be an easy task and could be more complex than just sharing written scientific papers. That is why Nonaka and Takeuchi (1995) stress the importance of making implicit knowledge explicit in order to be shared amongst individuals. Only then is new knowledge created, as it is also grounded in the assumptions of academic knowledge creation (Wierzbicki & Nakamori, 2006b). Therefore, the crucial task is to make the process of knowledge exchange between people flow and support the sharing of knowledge, while establishing practice environments to foster this flow (Brown & Duguid, 2002).

Besides the concept of communities of practice, which aim to develop such environments, research discusses the meaning of “*ba*” (Nonaka & Konno, 1998; Nonaka, Toyama, & Konno, 2000). “*Ba*” can be understood as a “place” of “shared context” for interaction and knowledge creation similar to the concept of situated learning. A core aspect here is that the creation of knowledge is connected to a specific context, as it is in situated learning. One research focus is on analyzing scientific behavior and a scientist’s view of his work and community in order to foster “*ba*” (Hautala, 2011). Hautala (2011, p. 13) sums up “*ba*” by saying that “although it is not possible to plan discourse and interaction, certain aspects of *ba* can be planned.” Thus both models aim to offer a realm for knowledge creation. However, Wenger stresses that communities of practice cannot be created or explicitly planned, but only cultivated (Wenger, 2008; Wenger et al., 2002). The understanding of knowledge, its sharing, and its transmission between two persons is central for any collaboration. Tian et al. (2009) recognize that their students miss collaborations between their peers and propose diverse strategies to foster a creative environment supporting collaboration and knowledge sharing. Here they refer to “hard environments”, which also include a proper working place for researchers to meet in for the purposes of sharing and discussion. Communities of practice, according to Wenger (Wenger, 2008), focus on a more social, structured environment. Nevertheless, the intentions of both approaches are equal. A study by Hara et al. (2003) aimed at analyzing researchers’ views on collaboration via interviews. The

results show that participants value the importance of collaboration, although their opinions differ. Undergraduate students are not seen as collaborators, with senior researchers seeing their relations to the former on a rather educational level. However, Hara et al. (2003) stress that the exchange and collaboration amongst junior and senior researchers is crucial, and that the “learners” also contribute to the development of the research community. They refer to the process of learning in a community of practice and propose a model showing diverse forms of interaction levels between scientists. This concept is quite similar to Wenger et al. (2002), who describe a community of practice’s engagement levels regarding a core group and a peripheral group, as well as external persons. Hara et al. (2003) further distinguish between different levels of awareness of these engagement types on the part of community members who see a collaboration either on a more complementary level (project divided into separate units) or on a more integrative level (shared project). They conclude that integrative collaboration needs time to emerge and collaboration among members must be enhanced, especially regarding the political trend toward supporting collaboration (for example, National Research Council, 1993; ClusterCollaboration²; Cordis³).

Not only novice researchers have difficulties in finding the right community and sharing knowledge, as Tian et al. (2009) showed. Bos et al. (2007) define three barriers that scientists face in their work (for recurring barriers see also Allen-Meares & Pugach, 1982; Cooley, 1994), namely the aggregation of knowledge, a scientist’s preference for working independently, and institutional restrictions. The first refers to the difficulty of disseminating tacit knowledge, as stated above. The second barrier relates to the fact that in most cases, a researcher understands his work as independent, which makes the establishment of mutual communities and common goals more difficult than in other environments. The third barrier refers to external restrictions a researcher has to cope with, as, for example,

² www.clustercollaboration.eu

³ www.cordis.europa.eu

funding regulations that hinder collaborations (see also Hara et al., 2003, for factors influencing collaboration on various levels). In their study, Bos et al. (2007) try to define types of scientific collaboration environments so as to give advice on how these environments work and should be established with a view to these barriers. The examples name best practices, such as establishing a virtual community of practice and showing a need to develop environments to foster community building and knowledge sharing.

1.2.3 Survey: Usage of Academic Services and Communities of Practice

Before approaching an academic recommender system, a survey was conducted to explore researchers' collaboration behaviors and their searching for new collaborators (Heck & Peters, 2010). Additionally, the participants were asked about their usage of collaborative management systems such as social bookmarking systems, which are considered for the academic recommendation model. Thus, the survey gives a first impression of the usage and understanding of social bookmarking systems and academic communities of practice.

The link to the online questionnaire was sent to researchers working at the *Forschungszentrum Jülich*, Germany, which agreed to collaborate with the study⁴. The survey was sent to 363 employees in the field of solid-state physics, of which 43 (11.85 %, 25-62 years old) participated in the survey. The results of the questions concerning social bookmarking systems were appalling, but also revealing: Only one participant actually uses one single bookmarking system (Del.icio.us⁵). Being asked why they do not use collaborative bookmarking systems, one exemplary answer was: "For me, search engines, normal bookmarks and literature databases like JabRef work fine." It seems that several systems for academic information management and retrieval have become established, and new web services have difficulties convincing users of their advantage. Seven

⁴ The same participants also evaluated the experimental approaches in Chapter 4.

⁵ www.delicious.com

respondents thought that bookmarking systems are “less important”, while fourteen stated that they are “not important”.

Two out of 27 participants said that they work in an academic community of practice (Heck & Peters, 2010). One community was established via “contact through a conference or a colleague”, the other “through connections by people formerly working at the same institute.” This shows that communities of practice, although not generally common between researchers, have found their way into the scientific workplace. 8 % (n=25) said that communities of practice and working groups make their work a lot easier, 16 % said that they make them easier. Asked about bookmarking systems and their help in socializing with other colleagues and researchers to establish working groups, four out of twelve participants gave a positive feedback. One comment was that the “communication is made easier.” 35 % (n=28) thought that a recommendation system that points out scientists with shared interests for possible collaboration could be helpful. Comments on these questions show that most participants consider personal contact to be the most important factor for collaboration. However, a “seriously structured” recommendation system could support researchers’ work, “especially for younger scientists” (Heck & Peters, 2010).

In addition to the survey from Heck and Peters (2010), researchers from Jülich, who evaluated the expert recommender system (n=10), participated in a general survey as part of the approach’s evaluation (Heck, 2012b, 2013). This survey consisted of a semi-structured interview regarding questions about the scientist’s research behavior and his purchases of relevant literature, as well as working behavior – for instance, do they work in teams, and if so, with whom do they collaborate?

The results show that most of the physicists work in research teams (in groups generally no larger than five people). Regular face-to-face meetings are

important, although difficult if international partners are involved. Novice researchers, such as PhDs or newcomers in a field, often come into contact with new potential collaborators at meetings such as scientific conferences and workshops, or get introduced to them via senior colleagues. However, it is more difficult for novice researchers to find new relevant collaborators as they have not yet established their social scientific network. The researchers' choice of possible collaborators highly depends on their research interests. On the one hand, there must be a high thematic overlap. On the other hand, an overlap that is too high could be disadvantageous. It is more valuable if scientific interests complement each other. This aspect is important when it comes to author similarity detection, as discussed further on. Moreover, a scientist's professionalism is important. The interviewees stated that a "professional collaboration" is crucial and a colleague's "style of work" and motivation must fit their own expectations and preferences. Furthermore, collaboration is difficult if people who do not know each other are "stuck together" in a team. In contrast, personal relations and mutual interests facilitate collaboration. If a collaboration team exists, most researchers claim that face-to-face meetings are more fruitful than video conferences or purely digital exchange, which confirms studies on distributed and virtual communities.

The interviewees regard collaborations with international institutions as desirable. However, institutions may also hinder further collaborations, either consciously or unconsciously. Some collaborations between institutions have a long tradition or have developed historically. They are durable, but may prevent other valuable collaborations between institutions. These statements from the interviews confirm past studies (Bos et al., 2007; Stokols et al., 2003; Stokols, Harvey, Gress, Fuqua, & Phillips, 2005). Concerning academic knowledge management and search behavior, the physicists showed similar preferences. No-one used a social tagging system. However, one participant uses Endnote⁶, which enables them to share stored literature with a team or other Endnote users. When searching for literature, researchers apply diverse strategies. A classic method involves using information

⁶ www.endnote.com

services like Web of Science⁷, PubMed⁸ or Google Scholar⁹. When citation and reference information is available, scientists do not only rank publications by their number of citations to find relevant literature, but they also look at publications that cite themselves to gain new insights. Some researchers have a fixed number of journals (generally about four or five) important to their field, which they read regularly to find new relevant literature. Additionally, they set literature alerts to get new information. To assess the relevance of a publication, one interviewee stressed the importance of the author's name. Only if he or she does not know the author, the scientist will read the abstract and further text passages of a publication to get an impression of the relevance of author and paper.

The interviews showed that collaboration and engagement in communities of practice is not the standard among researchers. Although some scientists stated that finding collaborators is not difficult, and that relevant partners are met at conferences or via other colleagues, other participants claimed that collaboration could be improved. Especially during their first year as novice PhDs, they did not know their community very well and did not have many collaborators. Collaboration opportunities developed through interaction with their local colleagues and their further integration into the research community. This scenario corresponds to Lave and Wenger's (1991) concept of situated learning and Wenger's (2008) idea of communities of practice. Nevertheless, integration into a network and full participation must be fostered. Furthermore, academic knowledge creation requires elements that thrive through more interaction with other researchers and an openness toward existing knowledge in the academic field. Thus, collaboration beyond local communities is needed.

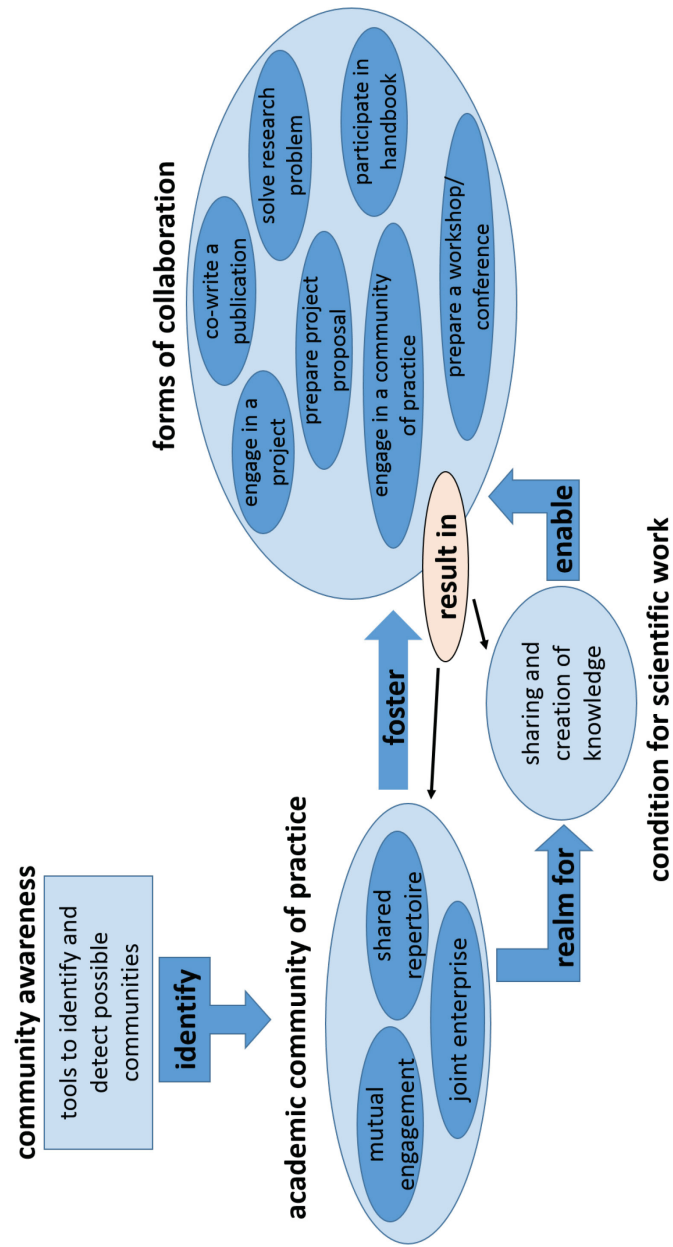
To summarize, communities of practice offer a realm – held together by mutual engagement, joint enterprise and a shared repertoire – where members share, exchange and create knowledge. This condition is important for researchers to

⁷ <http://apps.webofknowledge.com>

⁸ <http://www.ncbi.nlm.nih.gov/pubmed>

⁹ <https://scholar.google.de>

Figure 1.5. Relations in the academic environment and support through community awareness.



unfold their knowledge creation in the best possible way. Figure 1.5 shows the relation between the concept of a community of practice and knowledge creation. For a researcher, communities of practice are spaces in which to develop and advance scientific work through engagement, participation and interaction. This valuable environment leads to research collaborations, which are crucial for science per se and for the development of a researcher's knowledge creation with regard to the underlying theories of knowledge creation and knowledge exchange. However, as Hara et al. (2003, p. 953) state: "Collaboration is neither easily achieved nor guaranteed to succeed even though the nature of scientific work requires working together for a common goal and sharing of knowledge." Thus, the question now is which support can be provided in order to enhance those environments that foster communities of practice and interaction among researchers? How can the process of establishing communities of practice for academics be supported and initiated? How can services introduce researchers to each other to support scientific collaboration? A first step is to make scientists aware of their environments and of the potential for future communities. Only then can researchers fully participate in and support fledgling communities of practice in order to contribute to their own knowledge creation and the development of science. Wenger claims: "The task is to identify such groups and help them come together as communities of practice" (Wenger & Snyder, 2000, p. 144). Taking into account the sense of cultivating communities of practice, new technologies and web data might foster this awareness of future communities and help to detect them.

References

- Allen-Meares, P., & Pugach, M. (1982). Facilitating interdisciplinary collaboration on behalf of handicapped children and youth. *Teacher Education and Special Education*, 5(1), 30–36.
doi:10.1177/088840648200500106
- Bertels, H. M. J., Kleinschmidt, E. J., & Koen, P. A. (2011). Communities of practice versus organizational climate: Which one matters more to dispersed collaboration in the front end of innovation? *Journal of Product Innovation Management*, (28), 757–772. doi:10.1111/j.1540-5885.2011.00836.x
- Blair, D. C. (2002). Knowledge management: Hype, hope, or help? *Journal of*

- the American Society for Information Science and Technology*, 53(12), 1019–1028. doi:10.1002/asi.10113
- Blazek, R. (2007). *Author-statement citation Analysis Applied as a Recommender System to Support Non-domain-expert Academic Research: Doctoral dissertation*. Nova Southeastern University. Retrieved from http://nsuworks.nova.edu/gscis_etd/416
- Bolisani, E., & Scarso, E. (2014). The place of communities of practice in knowledge management studies: a critical review. *Journal of Knowledge Management*, 18(2), 366–381. doi:10.1108/JKM-07-2013-0277
- Bos, N., Zimmerman, A., Olson, J. S., Yew, J., Yerkie, J., Dahl, E., & Olson, G. M. (2007). From shared databases to communities of practice: A taxonomy of collaboratories. *Journal of Computer-Mediated Communication*, 12(2), 652–672. doi:10.1111/j.1083-6101.2007.00343.x
- Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy*, 33, 599–616. doi:10.1016/j.respol.2004.01.008
- Brown, J. S., & Duguid, P. (1991). Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. *Organization Science*, 2(1), 40–57. doi:10.1287/orsc.2.1.40
- Brown, J. S., & Duguid, P. (2002). *The Social Life of Information*. Boston, MA: Harvard Business School Press.
- Brown, R. E. (2001). The process of community-building in distance learning classes. *Journal of Asynchronous Learning Networks*, 5(2), 18–35. Retrieved from <http://onlinelearningconsortium.org/read/journal-issues/>
- Cassidy, L. (2011). Online communities of practice to support collaborative mental health practice in rural areas. *Issues in Mental Health Nursing*, 32(2), 98–107. doi:10.3109/01612840.2010.535648
- Cheung, C. M. K., Lee, M. K. O., & Lee, Z. W. Y. (2013). Understanding the continuance intention of knowledge sharing in online communities of practice through the post-knowledge-sharing evaluation processes. *Journal of the American Society for Information Science and Technology*, 64(7), 1357–1374. doi:10.1002/asi.22854
- Choi, M. (2006). Communities of practice: an alternative learning model for knowledge creation. *British Journal of Educational Technology*, 37(1), 143–146. doi:10.1111/j.1467-8535.2005.00486.x
- Cooley, E. (1994). Training an interdisciplinary team in communication and decision-making skills. *Small Group Research*, 25(1), 5–25. doi:10.1177/1046496494251002
- Corso, M., Giacobbe, A., & Martini, A. (2009). Designing and managing business communities of practice. *Journal of Knowledge Management*, 13(3), 73–89. doi:10.1108/13673270910962888

- Cox, A. M. (2005). What are communities of practice? A comparative review of four seminal works. *Journal of Information Science*, 31(6), 527–540. doi:10.1177/0165551505057016
- Cronin, B., Shaw, D., & La Barre, K. (2003). A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology*, 54(9), 855–871. doi:10.1002/asi.10278
- Davenport, E. (2001). Knowledge management issues for online organisations: 'Communities of practice' as an exploratory framework. *Journal of Documentation*, 57(1), 61–75. doi:10.1108/EUM0000000007077
- Davenport, E., & Hall, H. (2002). Organizational knowledge and communities of practice. *Annual Review of Information Science and Technology*, 36(1), 171–227. doi:10.1002/aris.1440360105
- Fenton, S. L., Covvey, H. D., Mulholland, D. W., Cowan, D. D., Shamian, J., & Schroeder, B. (2007). A web-based communities of practice support system for caregivers. In K. A. Kuhn, J. R. Warren, & T.-Y. Leong (Eds.): *Studies in Health Technology and Informatics. Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems (Medinfo)* (129th ed., pp. 993–996). Amsterdam, Netherlands: IOS Press. Retrieved from <http://search.informit.com.au/documentSummary;dn=781782981827801;res=IELHEA>
- Finholt, T. A. (1999). Collaboratory life: Challenges of internet-mediated science for chemists. In *National Research Council (US) Chemical Sciences Roundtable. Impact of Advances in Computing and Communications Technologies on Chemical Science and Technology: Report of a Workshop*. (pp. 97–108). Washington, DC, USA: National Academy Press. Retrieved from <http://www.ncbi.nlm.nih.gov/books/NBK44967/>
- Fuller, A., Unwin, L., Felstead, A., Jewson, N., & Kakavelakis, K. (2007). Creating and using knowledge: an analysis of the differentiated nature of workplace learning environments. *British Educational Research Journal*, 33(5), 743–759. doi:10.1080/01411920701582397
- Fuller, A., Unwin, L., Felstead, A., Jewson, N., & Kakavelakis, K. (2012). Creating and using knowledge. In D. W. Livingstone & D. Guile (Eds.): *The Knowledge Economy and Lifelong Learning* (pp. 191–206). Rotterdam, NL: SensePublishers. doi:10.1007/978-94-6091-916-2_10
- Gadamer, H. G. (1974). Hermeneutik. In J. Ritter (Ed.): *Historisches Wörterbuch der Philosophie. Band 3* (pp. 1061–1074). Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.
- Gadamer, H. G. (1975). *Truth and Method*. London, UK: Sheed & Ward.

- Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. *Journal of the American Society for Information Science and Technology*, 63(2), 323–335. doi:10.1002/asi.21688
- Green, H. E. (2014). Facilitating communities of practice in digital humanities: Librarian collaborations for research and training in text encoding. *Library Quarterly*, 84(2), 219–234. doi:10.1086/675332
- Guerrero Bote, V. P., Olmeda-Gómez, C., & De Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of the American Society for Information Science and Technology*, 64(2), 392–404. doi:10.1002/asi.22754
- Gust von Loh, S. (2009). *Evidenzbasiertes Wissensmanagement*. Wiesbaden, Germany: Gabler.
- Hara, N. (2007). Information technology support for communities of practice: How public defenders learn about winning and losing in court. *Journal of the American Society for Information Science and Technology*, 58(1), 76–87. doi:10.1002/asi.20445
- Hara, N. (2009). *Communities of Practice: Fostering Peer-to-Peer Learning and Informal Knowledge Sharing in the Work Place*. Berlin/Heidelberg, Germany: Springer.
- Hara, N., & Kling, R. (2002). Communities of practice with and without information technology. In *Proceedings of the 65th ASIS&T Annual Meeting* (pp. 338–349). ASIS&T. doi:10.1002/meet.1450390137
- Hara, N., Shachaf, P., & Stoerger, S. (2009). Online communities of practice typology revisited. *Journal of Information Science*, 35(6), 740–757. doi:10.1177/0165551509342361
- Hara, N., Solomon, P., Kim, S.-L., & Sonnenwald, D. H. (2003). An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, 54(10), 952–965. doi:10.1002/asi.10291
- Hautala, J. (2011). International academic knowledge creation and ba. A case study from Finland. *Knowledge Management Research & Practice*, 9(1), 4–16. doi:10.1057/kmrp.2010.23
- Heck, T. (2012a). Analyse von sozialen Informationen für Autorenempfehlungen. *Information - Wissenschaft & Praxis*, 63(4), 261–272. doi:10.1515/iwp-2012-0048
- Heck, T. (2012b). Recommendation for social networking in academia. In G. Widén & K. Holmberg (Eds.): *Library and Information Science Volume 5: Social Information Research* (pp. 239–265). Bingley, UK: Emerald Group Pub. doi:10.1108/S1876-0562(2012)0000005013

- Heck, T. (2013). Combining social information for academic networking. In A. Bruckman, S. Counts, C. Lampe, & L. G. Terveen (Eds.): *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW)* (pp. 1387–1398). New York, NY, USA: ACM. doi:10.1145/2441776.2441932
- Heck, T., & Peters, I. (2010). Expert recommender systems: Establishing communities of practice based on social bookmarking. In *Proceedings of 10th International Conference on Knowledge Management and Knowledge Technologies (I-Know)* (pp. 458–464). Graz, Austria: J.UCS. Retrieved from <http://www.isi.hhu.de/heck>
- Henrich, K. J., & Attebury, R. (2010). Communities of practice at an academic library: A new approach to mentoring at the University of Idaho. *Journal of Academic Librarianship*, 36(2), 158–165. doi:10.1016/j.acalib.2010.01.007
- Huysman, M. (2002). Organizational learning and communities of practice: A social constructivist perspective. In *Proceedings of 3rd European Conference on Organizational Knowledge, Learning and Capabilities*. Retrieved from http://apollon1.alba.edu.gr/OKLC2002/Proceedings/pdf_files/ID149.pdf
- Huysman, M., & Wit, D. de. (2002). *Knowledge Sharing in Practice. Information science and knowledge management*. Dordrecht, Netherlands: Kluwer Acad. Publ.
- Jakubik, M. (2008). Experiencing collaborative knowledge creation processes. *The Learning Organization*, 15(1), 5–25. doi:10.1108/09696470810842475
- Jakubik, M. (2011). Becoming to know. Shifting the knowledge creation paradigm. *Journal of Knowledge Management*, 15(3), 374–402. doi:10.1108/13673271111137394
- Kagan, S. L. (1991). *United We Stand: Collaboration for Child Care and Early Education Services. Early childhood education series*. New York, NY, USA: Teachers College Press.
- Kietzmann, J., Plangger, K., Eaton, B., Heilgenberg, K., Pitt, L., & Berthon, P. (2013). Mobility at work. A typology of mobile communities of practice and contextual ambidexterity. *Journal of Strategic Information Systems*, 22(4), 282–297. doi:10.1016/j.jsis.2013.03.003
- Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science and Technology*, 51(14), 1306–1320. doi:10.1002/1097-4571(2000)9999:9999<:AID-ASI1047>3.0.CO;2-T
- Kuhn, T. S. (1996). *The Structure of Scientific Revolutions*. Chicago, IL, USA: University of Chicago Press.
- Lancho Barrantes, B. S., Guerrero Bote, V. P., Rodríguez, Z. C., & De Moya-Aneón, F. (2012). Citation flows in the zones of influence of scientific

- collaborations. *Journal of the American Society for Information Science and Technology*, 63(3), 481–489. doi:10.1002/asi.21682
- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation. Learning in doing*. Cambridge, MA, USA: Cambridge Univ. Press.
- Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673–702. doi:10.1177/0306312705052359
- Lesser, E. L., & Storck, J. (2001). Communities of practice and organizational performance. *IBM Systems Journal*, 40(4), 831–841. doi:10.1147/sj.404.0831
- Liang, T.-P. (2008). Recommendation systems for decision support: An editorial introduction. *Decision Support Systems*, 45(3), 385–386. doi:10.1016/j.dss.2007.05.003
- Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology & Human Values*, 17(1), 101–126. doi:10.1177/016224399201700106
- Mabery, M. J., Gibbs-Scharf, L., & Bara, D. (2013). Communities of practice foster collaboration across public health. *Journal of Knowledge Management*, 17(2), 226–236. doi:10.1108/13673271311315187
- Malinowski, J., Weitzel, T., & Keim, T. (2008). Decision support for team staffing: An automated relational recommendation approach. *Decision Support Systems*, 45(3), 429–447. doi:10.1016/j.dss.2007.05.005
- Mattessich, P. W., & Monsey, B. R. (1992). *Collaboration: What Makes it Work. A Review of Research Literature on Factors Influencing Successful Collaboration*. St. Paul, MN: Amherst H. Wilder Foundation. Retrieved from <http://eric.ed.gov/?id=ED390758>
- McLure Wasko, M., & Faraj, S. (2000). "It is what one does" why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems*, 9(2-3), 155–173. doi:10.1016/S0963-8687(00)00045-7
- Murillo, E. (2011). Communities of practice in the business and organization studies literature. *Information Research*, 16(1), 464. Retrieved from <http://InformationR.net/ir/16-1/paper464.html>
- National Research Council. (1993). *National Collaboratories: Applying Information Technology for Scientific Research*. Washington, D.C., USA: National Academy Press. Retrieved from <http://www.worldcat.org/title/national-collaboratories-applying-information-technology-for-scientific-research/oclc/43475647>
- Nistor, N., & Fischer, F. (2012). Communities of practice in academia: Testing a quantitative model. *Learning, Culture and Social Interaction*, 1(2), 114–

126. doi:10.1016/j.lcsi.2012.05.005
- Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1), 14–37. Retrieved from <http://www.jstor.org/stable/2635068>
- Nonaka, I., & Konno, N. (1998). The concept of "Ba": Building a foundation for knowledge creation. *California Management Review*, 40(3), 40–54. doi:10.2307/41165942
- Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. New York, NY, USA: Oxford Univ. Press.
- Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and leadership: A unified model of dynamic knowledge creation. *Long Range Planning*, 33(1), 5–34. doi:10.1016/S0024-6301(99)00115-6
- Pan, S. L., & Leidner, D. E. (2003). Bridging communities of practice with information technology in pursuit of global knowledge sharing. *Journal of Strategic Information Systems*, 12(1), 71–88. doi:10.1016/S0963-8687(02)00023-9
- Pattinson, S., & Preece, D. (2014). Communities of practice, knowledge acquisition and innovation: a case study of science-based SMEs. *Journal of Knowledge Management*, 18(1), 107–120. doi:10.1108/JKM-05-2013-0168
- Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421–432. doi:10.1023/B:SCIE.0000034384.35498.7d
- Polanyi, M. (1967). *The Tacit Dimension*. Garden City, NY, USA: Doubleday (Anchor Books).
- Polanyi, M. (1982). *Personal Knowledge: Towards a Post-Critical Philosophy* (8. ed.). Chicago, IL, USA: University of Chicago Press.
- Price, D. J. d. S. (1986). *Little Science, Big Science ... and Beyond*. New York, NY, USA: Columbia Univ. Press.
- Price, D. J. d. S., & Beaver, D. D. (1966). Collaboration in an invisible college. *American Psychology*, 21(11), 1011–1018. Retrieved from <http://psycnet.apa.org/journals/amp/21/11/>
- Ramchand, A., & Pan, S. L. (2012). The co-evolution of communities of practice and knowledge management in organizations. *Data Base for Advances in Information Systems*, 43(1), 8–23. doi:10.1145/2139011.2139013
- Rosenbaum, H., & Shachaf, P. (2010). A structuration approach to online communities of practice: The case of Q&A communities. *Journal of the American Society for Information Science and Technology*, 61(9), 1933–1944. doi:10.1002/asi.21340

- Schrage, M. (1995). *No More Teams!: Mastering the Dynamics of Creative Collaboration*. New York, NY, USA: Currency Doubleday.
- Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. *Journal of Management*, 20(2), 503–530. doi:10.1177/014920639402000210
- Stock, W. G., & Stock, M. (2013). *Handbook of Information Science*. Berlin/Boston: De Gruyter.
- Stokols, D., Fuqua, J., Gress, J., Harvey, R., Phillips, K., Baezconde-Garbanati, L., . . . Trochim, W. (2003). Evaluating transdisciplinary science. *Nicotine & Tobacco Research*, 5(1), 21–39. doi:10.1080/14622200310001625555
- Stokols, D., Harvey, R., Gress, J., Fuqua, J., & Phillips, K. (2005). In vivo studies of transdisciplinary scientific collaboration Lessons learned and implications for active living research. *American Journal of Preventive Medicine*, 28(2), 202–213. doi:10.1016/j.amepre.2004.10.016
- Suorsa, A., & Huotari, M.-L. (2014). Knowledge creation and the concept of a human being: A phenomenological approach. *Journal of the Association for Information Science and Technology*, 65(5), 1042–1057. doi:10.1002/asi.23035
- Swan, J., Scarbrough, H., & Robertson, M. (2002). The construction of 'communities of practice' in the management of innovation. *Management Learning*, 33(4), 477–496. doi:10.1177/1350507602334005
- Tian, J., Nakamori, Y., & Wierzbicki, A. P. (2009). Knowledge management and knowledge creation in academia: a study based on surveys in a Japanese research university. *Journal of Knowledge Management*, 13(2), 76–92. doi:10.1108/13673270910942718
- Tobbell, J., O'Donnell, V., & Zammit, M. (2010). Exploring transition to postgraduate study: shifting identities in interaction with communities, practice and participation. *British Educational Research Journal*, 36(2), 261–278. doi:10.1080/01411920902836360
- Tsoukas, H. (2009). A dialogical approach to the creation of new knowledge in organizations. *Organization Science*, 20(6), 941–957. doi:10.1287/orsc.1090.0435
- Valaitis, R. K., Akhtar-Danesh, N., Brooks, F., Binks, S., & Semogas, D. (2011). Online communities of practice as a communication resource for community health nurses working with homeless persons. *Journal of Advanced Nursing*, 67(6), 1273–1284. doi:10.1111/j.1365-2648.2010.05582.x
- Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *Management Information Systems Quarterly*, 29(1), 35–58. Retrieved from <http://www.jstor.org/stable/25148667>

- Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, 7(2), 225–246. doi:10.1177/135050840072002
- Wenger, E. (2008). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, MA, USA: Cambridge Univ. Press.
- Wenger, E. (2010). Communities of practice and social learning systems: The career of a concept. In C. Blackmore (Ed.): *Social Learning Systems and Communities of Practice* (pp. 179–198). London: Springer London. doi:10.1007/978-1-84996-133-2_11
- Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating Communities of Practice: A Guide to Managing Knowledge*. Boston, MA, USA: Harvard Business School Press.
- Wenger, E., & Snyder, W. (2000). Communities of practice? The organizational frontier. *Harvard Business Review*, 78(1), 139–145. Retrieved from <https://hbr.org/2000/01/communities-of-practice-the-organizational-frontier>
- Wierzbicki, A. P., & Nakamori, Y. (2005). Knowledge creation and integration: Creative space and creative environments. In *Proceedings of the 38th Annual Hawaii International Conference (HICSS)* (p. 91). New York, NY, USA: IEEE. doi:10.1109/HICSS.2005.370
- Wierzbicki, A. P., & Nakamori, Y. (2006a). *Creative Space: Models of Creative Processes for the Knowledge Civilization Age*. Berlin/Heidelberg, Germany: Springer.
- Wierzbicki, A. P., & Nakamori, Y. (2006b). Towards creative environments: Conclusions from creative space. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS)* (p. 46). New York, NY, USA: IEEE. doi:10.1109/HICSS.2006.495

2 Detection of Scientific Communities

Support in finding relevant collaboration partners is offered by diverse services. Cordis¹, for example, is a research partner database by the European Union that enables to create a profile and add information about European-funded projects. Via a search mask, it is possible to search for potential partners, or inversely, a researcher can search for projects and suggest him- or herself as a potential partner. Currently (June 2016), Cordis includes 11,513 partner profiles (figure 2.1) (numbers from May 2015 showed 4,685 partner profiles). 364 German partnership requests and partners are included in the database. The number is growing steadily, as in May 2014 only 89 potential German collaborators were found, and in May 2015 there were 352. However, only individuals who have created a profile in Cordis can be found, of course. Other potential partners will remain invisible, and relevant future collaboration unfulfilled.

Other expert search services, like Microsoft Academic Search² or ArnetMiner³ (Tang et al., 2008), are more elaborate. They try to gather all information about researchers, from their institutional and private webpages to data gleaned from information services, including publication and citation data. A scientist does not need to be active to be listed in those services, which is an advantage over offers such as Cordis. On expert search sites, a user is given the opportunity to search for topics or conferences via terms contained in the description, or for researchers via their names. If a user can formulate his or her information need and search query, these services can represent a viable option for finding experts. However, many users are unable to formulate their need. More importantly, however, results in those services depend on the user's perspective. The following section will discuss the need to expand this perspective. Furthermore, these services do not recommend personalized results based on a

¹ <http://cordis.europa.eu/>

² <http://academic.research.microsoft.com>

³ <http://aminer.org>



The screenshot shows the CORDIS website interface. At the top, there are navigation links: "About CORDIS | Contact | Advanced Search | Legal Notice | English (en)". Below this is the CORDIS logo and the text "Community Research and Development Information Service". A blue navigation bar contains "European Commission > CORDIS > Partners Service > Guest > Home" and a "Sign in" button.

The main content area is titled "Research Partners". Under "You can:", there are several options:

- Search for partners (with a "Search" button)
- Query more with an advanced search
- Browse these active profiles and collaboration requests to build your network:
 - 11513 Partner profiles
 - 49 Open Calls for Proposals
 - 5210 Partnership requests
 - 1261 Proposing project
 - 3949 Offering collaboration
 - 387 Groups
- Contact National Contact Point networks to get further support to find partners in your specific theme
- View the help pages and video tutorials

On the right side, there is a "Create or update your profile" form with fields for "Username:" and "Password:", a "Log in" button, and links for "Forgot your username or password?" and "Not yet registered?".

Figure 2.1. CORDIS research partner search, retrieved 06/21/2016 from <https://cordis.europa.eu/partners/web/guest>.

target person (figure 2.2). In the best case, they show similar researchers when a target person searches for their own profile, but those results are based on explicit relations such as co-authorship or topical overlap via terms including author keywords. In ArnetMiner, novice researchers do not get recommendations based on “similar authors”, likely because the system does not have enough researcher information and there is a set threshold.

In addition to retrieval services or person databases, expert recommender systems that focus on personalized recommendations include tasks such as “designing for a network” (Reichling, Veith, & Wulf, 2007, p. 431) or “mining social networks” (Alguliev, Aliguliyev, & Ganjaliyev, 2011, p. 229) for collaboration. The starting point of an expert recommendation system in this work is social information concerning researchers. The main question here is: which information must be determined regarding its relevance in order to find the “right” collaborator?

2.1 Implicit Relations via Bibliographic Coupling and Co-Citation

A critical aspect for choosing the right collaborator is reliability. If a person needs advice or is searching for something, he or she will ask reliable people, like his best friends. Depending on the individual need, the reliability of the people in one's environment changes. For example, a person searching for medical advice will trust a physician over his or her best friend, but will ask a friend for advice concerning financial issues if that friend works at a bank. Thus, reliability depends on specific needs in a concrete situation. A person (or any other source) must be credible to the person receiving information from them. In respect to computer services, such as recommender systems, credibility can be seen as a "judgement made by a message receiver concerning the believability of a communicator" (Yoo & Gretzel, 2011, p. 457), see also Fogg, Lee, and Marshall (2002)). Among other factors, credibility includes trustworthiness and expertise (Fogg, 2002). The idea is to make services more credible for users and thus increase their usage (Yoo & Gretzel, 2011) (see also aspects of trust in chapter 3). People generally trust other people who have expertise or

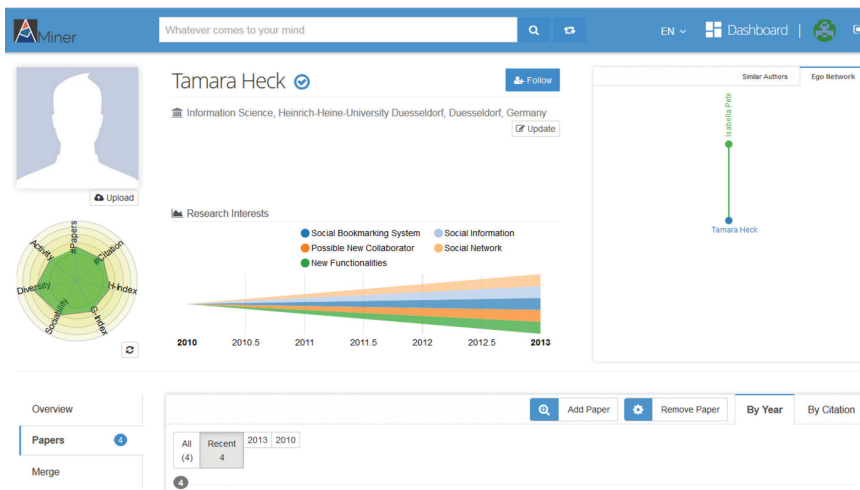


Figure 2.2. ArnetMiner/AMiner. The service offers a search for scientists and shows similar authors. Users have to sign up. Retrieved 06/21/2016 from <https://aminer.org/profile/tamara-heck/53f31a07dabfae9a8442f298>

proficiency in a specific field. A person can gain expert status by claiming that status for him- or herself, but others will not trust them on that basis alone. To the contrary, a person gains expert status by being labeled as such by other people. Thus, it is important to achieve a good reputation within a community.

Jøsang, Ismail, and Boyd (2007) discuss the differences between trust and reputation with regard to the task of reputation systems:

“Trust is the subjective probability by which an individual, A, expects that another individual, B, performs a given action on which its welfare depends” (Jøsang et al., 2007, p. 619) (compare Gambetta, 1988).

“Reputation is what is generally said or believed about a person’s or thing’s character or standing. [...] [It] is a quantity derived from the underlying social network which is globally visible to all members of the network” (Jøsang et al., 2007, p. 620).

The authors provide an example to make the differences between both elements clear: a person can trust another person *because* of their good reputation. On the other hand, a person can trust another person *despite* their reputation. This means that the concept of trust involves a strong subjective component. A person decides whether or not to trust another person on the basis of their experience and relationship with that person (Jøsang et al., 2007). Other influences derive from a human being’s interaction with his or her environment (see figure 2.4). However, if a person’s interpretation repertoire (knowledge, prejudices, assumptions, beliefs) is not sufficient to make a clear decision and if no experience can be consulted, a decision about trust or non-trust is then based on external sources such as reputation, which reflects the opinions of others. Reputation becomes crucial wherever personal relationships do not exist, as in virtual communities found in web forums, online shops, blogging portals etc. Reputation-aware systems try to generate trust among the users of those communities, for instance by applying techniques to show a user another user’s reputation level. A simple model would be an evaluation system where users

can state if they trust other users, as on Amazon⁴, for example. Here, if a user reviews a product other users can state if their comment was helpful or not. Users who write helpful reviews get a better reputation. Amazon lists their statements under “the most helpful reviews” at the top of a review list. In a recommender system, recommendation for a target user should be based on trustworthy users. Cruz, Motta, Claudia, L. R., Santoro, and Elia (2009) suggest implementing reputation mechanisms on a platform to support virtual communities of practice for students, practitioners and researchers working at a university. They refer to the idea, discussed by Wenger (2008), that a newcomer first has to gain a reputation within a community to be accepted. Thus a user’s profile, which includes publications and information about participation in projects, is analyzed and a ranking implemented that identifies a user as a “beginner”, “intermediate” or “expert”. Furthermore, the system analyzes a user’s posts and their reviews by other users based on three dimensions (agreement to a post, usefulness and post relevance for the community). Giving users information about other users’ reputation raises credibility and should lead to more interaction among users. Recommendations of publications within the system are based on a target user’s personal trust network, which he or she can actively change in the case study by Cruz et al. (2009). This user network should guarantee trustworthy recommendations for a target user and lead to greater system acceptance. Ziegler and Lausen (2004) analyzed the correlation between trust and user similarity for decentralized recommender systems, which means systems that refer to diverse types of sources to overcome data sparsity. Although the authors had to conduct two studies due to data biases derived from the service they used, the final results showed that similarity between users who trust each other, is higher than between users who do not (explicit trust statements were available). For recommendation systems, this means that trust can be derived from user similarity (Montaner, López, & de la Rosa, 2002). Thus, similarity between two people must be defined appropriately in order to

⁴ <http://www.amazon.com>

lead to more valuable recommendations.

Trust-aware recommender systems should therefore base their decisions on trustworthy sources. This aspect becomes crucial for expert recommendations that aim to suggest the “right” collaboration partners. In an academic environment, these partners should be reliable researchers and models to find them should therefore rely on trustworthy sources. A researcher’s trustworthiness and reputation depends on their scientific craft, which is shown by their publications. These publications have references and citations, and the relations between them display diverse perspectives that allow one to make statements about their reputation and appropriateness as a collaboration partner. Considering this social information about a researcher leads to the first models that can be used for expert recommendation.

Citation and publication information contain statements about a researcher’s reputation. Generally, it is stated that a researcher’s reputation grows with the number of their publications and citations. Cronin (1984) describes the development of the meaning and position of citations within science. He says that citations have become established as quality indicators: “The most common means of bestowing credit and recognition in science is via citations” (Cronin, 1984, p. 2). However, Cronin warns, the research community must be aware of pitfalls as it is difficult to accurately interpret citations and quantitative numbers do not explicitly refer to qualitative aspects. Researcher citation behavior depends on external restrictions and constraints. One example is formatting and editor rules that may lead scientists to leave out any references. Furthermore, researchers are not always aware of their own citation behavior and the meaning of the citation process for science: “To put it another way: authors may not be clear in their own minds why it is that they cite the way they do” (Cronin, 1984, p. 5). Apart from these epistemological aspects, which should be kept in mind while discussing the meaning of citations, the citing of other scientific work is one of a researcher’s obligations and forms part of their knowledge creation process (see chapter 1). A researcher makes statements about his or her work and research field via their publications and the references therein. This

information directly pertains to their scientific craft and interests as a researcher. The assumption is that such social information about a researcher points to relations to other scientists with similar interests. Additionally, scientists cite other works and build up more relations among their colleagues. To point out such relations would support community building among those similar researchers.

Between any two authors there are four relations with regard to their publications, references and citations, respectively: co-authorship, direct citation, bibliographic coupling of authors, and author co-citation. From a target researcher's point of view, co-authorship and citations are direct connections (Stock & Stock, 2013, p. 751). That means the target person is aware of the relations because he or she knows an author with whom they have published a document, and knows the authors he or she has cited in a publication or who have cited him or her (see interviewee statement in chapter 1). Hence, recommendations based on these relations might lead to relevant results, but these results are too explicit and a researcher will be able to find them without the support of a system. For example, a recommender system based on co-authorship would suggest a co-author of a researcher's co-author as a potential relevant collaborator. However, it is highly probable that a researcher already knows their co-authors' networks and other authors with whom they have collaborated. Furthermore, co-authors themselves act as recommenders for these suggestions. In contrast to co-authorship and citations, relations via bibliographic coupling and author co-citation are undirected connections (Stock & Stock, 2013, p. 751) and more implicit for a target researcher.

Bibliographic coupling (Kessler, 1963) means that two scientific papers are linked if they include the same references. If they have many references in common, the probability that they refer to similar scientific topics increases. Hence, common references point to similarity, allowing a user to find related papers that are important for his or her research (Stock & Stock, 2013, p. 753). If this assumption is aggregated to the author level, it means that two authors are similar if they have many shared citations in their respective papers. Here we

must consider one crucial aspect: either the total number of shared references is important or the number of shared references per publication is important. For example, let author *a* have six references in common with authors *b* and *c*. These six shared references are found in two unique documents by author *a* and author *b*, respectively, but for author *c* they are distributed across six individual documents. Counting the total number of shared references results in author *a* having the same degree of similarity with authors *b* and *c*. If the number of shared references per document is counted, authors *a* and *b* are more similar to one another than authors *a* and *c*, as the reference lists of *a*'s and *b*'s documents are more similar. In the first case, similarity would be based on false assumptions. The number of total shared references might show similarity between two authors who, for instance, refer to a scientific method they both used in diverse cases. Instead, a high level of shared references per paper is more significant to show author similarity and common interests.

In co-citation analysis (Leydesdorff, 2005; Marshakova, 1973; Schneider & Borlund, 2007a, 2007b; Small, 1973; White & Griffith, 1981, 1982), two scientific papers are regarded as linked when they are cited by the same document. This assumes that both papers contain similar topics because another document related to a specific topic refers to them. These co-cited works “form classes of similar works” (Stock & Stock, 2013, p. 753). Aggregating this assumption to the author level, if a third author cites two authors in the same paper, these authors might be similar in their scientific work because the third author names them in one and the same publication. Thus, co-citations point to similarity. This assumption is also valid for all co-authors, which means that all co-authors are similar to another author if all are cited in the same paper (Persson, 2001; Zhao & Strotmann, 2007).

Citation Information Services

For citation indexing and the measurement of similarities between researchers,

publication and citation data is required. The two largest scientific information services are hosted by Elsevier⁵ and Thomson Reuters⁶, namely Scopus⁷ and Web of Science. Web of Science (figure 2.3, formerly named Web of Knowledge) includes diverse databases that are a primary source for scientometric analyses (Cronin & Atkins, 2000; Stock & Stock, 2013). The main database is the Web of Science Core Collection (hereafter referred to as Web of Science), which includes the most important periodicals for a scientific field based on diverse indices (Garfield, 1955; Thomson Reuters, 2016). These indices, which measure a journal's impact and decide whether a periodical is recorded in Web of Science, are necessary as it is practically impossible to analyze citation data for all existing academic journals (Garfield & Stock, 2002; Linde & Stock, 2011). According to Thomson Reuters, Web of Science currently includes 12,000 analyzed journals (Thomson Reuters, 2016). Scopus (figure 2.4) is a database offered by Elsevier and currently includes, among other publications, articles from 21,500 peer-reviewed journals (Elsevier, 2016). Both services offer searches for citations, co-citations and common references.

The screenshot displays the Web of Science interface for a specific article. The article title is "The anatomy of a large-scale hypertextual Web search engine" by Brin, S. (Brin, S); Page, L. (Page, L). The article is from the journal "COMPUTER NETWORKS AND ISDN SYSTEMS", Volume 30, Issue 1-7, Pages 107-117, published in APR 1998. The DOI is 10.1016/S0169-7552(98)00110-X. The abstract states: "In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at http://google.stanford.edu/". Below the abstract, it notes: "To engineer a search engine is a challenging task. Search engines index tens to hundreds of millions of Web pages involving a comparable number of". On the right side, the "Citation Network" section shows "2,332 Times Cited", "12 Cited References", and "View Related Records". There are also links for "View Citation Map" and "Create Citation Alert". At the bottom right, the "All Times Cited Counts" section shows "2,471 in All Databases" and "2,332 in Web of Science Core Collection".

Figure 2.3. Single view of a surrogate in Web of Science with links to citations and related records. Retrieved 06/21/2016 from <http://apps.webofknowledge.com/>

⁵ <http://www.elsevier.com>

⁶ <http://thomsonreuters.com>

⁷ <http://www.scopus.com>

In Scopus, relations of bibliographic coupling, or co-citations, can be retrieved by selecting the results and viewing their references, and then viewing the citations of those references (figure 2.4). In Web of Science, the direct link called “related records” (figure 2.3) shows similar publications based on common references (Cawkell, 2000; Stock, 1999). Thus the numbers of those relations available for single publications are directly available to users. Co-citations cannot be found via direct link in Web of Science.

Scopus and Web of Science include author IDs to identify unique authors and to reduce author ambiguity, which is a great concern in research studies that compile data related to scientific authors. However, as data analysis in the case study in chapter 4 shows, author IDs are not always correct. Web of Science cooperates with ResearcherID⁸, a service that provides it with unique author information. However, a researcher first has to sign in to ResearcherID to get a unique number. If an author has not created a profile, no ID will be available in Web of Science.

Another shortcoming is that data in both services is not complete. With regard to reference and citation data, this issue leads to incomplete relations between authors and thus to inconsistent recommendation. The case study in section 5.3 shows an exemplary analysis of missing author information. Besides author information, there are two other aspects in favor of applying either Scopus or



Figure 2.4. Scopus search results page with links to citations and references. Indirect relations are searchable by selecting items in the results list.

⁸ <http://www.researcherid.com>

Web of Science for citation and reference analysis. The aim of author co-citation is to consider all authors of a publication. In Web of Science, only the first author of any cited document was listed in the reference section of its bibliographic entries (Zhao & Strotmann, 2011) at the time the case study in 5.3 was conducted (2011-2012). This leads to the loss of all secondary and tertiary authors of a publication, who will then not be considered for recommendation. It also means that if a target author is only the secondary author, their publication will not be assigned to them. Thus Web of Science was not an appropriate source of data for co-citation analysis at that time. However, Scopus named all authors in a publication's reference list and was thus considered for gathering co-citation data.

However, bibliographic coupling in Scopus appeared difficult for several reasons. The reference strings in the Excel files, which were downloaded for reference analysis, were not identical, for one, as the same reference can have diverse strings (see figure 2.5). Thus it is difficult to detect unique references. It seems that Scopus directly adopts reference strings from publications. Depending on the format of an article's reference list, unique references are cited differently. As there is no identical format, or any consistency of form to detect unique reference strings, Scopus was not used for bibliographic coupling analysis.

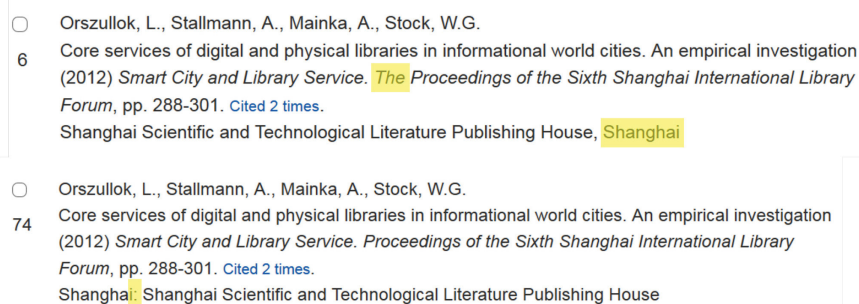
- 
- Orszulok, L., Stallmann, A., Mainka, A., Stock, W.G.
6 Core services of digital and physical libraries in informational world cities. An empirical investigation (2012) *Smart City and Library Service*. **The** *Proceedings of the Sixth Shanghai International Library Forum*, pp. 288-301. Cited 2 times.
Shanghai Scientific and Technological Literature Publishing House, **Shanghai**
- Orszulok, L., Stallmann, A., Mainka, A., Stock, W.G.
74 Core services of digital and physical libraries in informational world cities. An empirical investigation (2012) *Smart City and Library Service*. *Proceedings of the Sixth Shanghai International Library Forum*, pp. 288-301. Cited 2 times.
Shanghai Shanghai Scientific and Technological Literature Publishing House

Figure 2.5. Scopus reference view: Same reference, formatted differently in the reference lists of two articles (doi:10.1515/libri-2013-0024 // doi:). Retrieved 06/21/2016 from <https://www.scopus.com/>

Although data in Web of Science and Scopus includes some mistakes, both services seem to offer the most complete citation sets (compare, for example, Chadegani et al. (2013) and Li, Burnham, Lemley, and Britton (2010)). In contrast to these two, other services such as CiteSeer⁹ (Bollacker, Giles, & Lawrence, 1998) and ResearchIndex¹⁰ apply automated citation indexing. The latter was the basis of McNee's (2006) work on paper recommendation, in which he also applied co-citation analysis. This work uses professional information databases and asks whether datasets derived from those services are appropriate for expert recommendation.

2.2 Expanding Perspectives for Enhancing Academic Networks

Both bibliometric methods introduced above refer to implicit relations between researchers, which a recommender system could apply to make researcher relations explicit and present a target researcher with as-yet undetected collaborators in order to foster community building and interaction. Publication and citation data relates to a researcher's reputation, which is important for the aspects of credibility and trust in recommender systems. With respect to the results of research literature, the best performance in terms of representing research activities is achieved by combining bibliographic coupling and co-citation analysis (Boyack & Klavans, 2010; Gmür, 2003). However, bibliographic coupling and author co-citation limit the perspective on scientific work and its relations.

Relations of bibliographic coupling represent an author's perspective on social relations. He or she decides which papers to cite, and with this choice, influences his or her social relations to other researchers. On the other hand, author co-citation data takes into account a third party. A third author brings two researchers together when citing them. With the choice of papers to be cited, he or she influences social relations among researchers, and thus potential

⁹ <http://citeseerx.ist.psu.edu>

¹⁰ <http://www.researchindex.org>

recommendations. To summarize, using co-citation and reference data to build author networks takes into account the researchers' perspective: whom do they cite, by whom are they cited and who is co-cited?

However, to look only at common references or co-citations might be inappropriate, as data may also be sparse for some researchers. This is especially the case for new researchers who have just started their academic career and have little scientific reputation to be measured. Blazek (2007) calls them "domain-novice researchers", or academics who enter a new domain (see chapter 1). Those newcomers suffer from the cold-start problem: citation analysis can hardly be applied to novice researchers as long as there are no or only few references and citations. Furthermore, there is a time delay when measuring citations and author co-citations because an author's article will be cited several months after the publication at the earliest, with differences in time span varying between scientific disciplines. This means that a researcher who has published a recent article and would be a good collaborator may not be considered by a recommender system based on author co-citation analysis.

To overcome the limitations of data scarcity, further social information from the web can be used to make better decisions about the right collaboration partners. In collaborative services, users contribute to the system's data (Peters, 2009). They get involved in the data collection of a system and are able to add content. On social networking services, users add personal information. The amount of information in general grows with the amount of users in the system (Kipp, 2006b, 2011a). Using this information has advantages compared to data found in information services such as Web of Science and Scopus:

1. There is a greater variety of data available
2. The users' perspective is taken into consideration

Social information about researchers gleaned from web services takes into account the users' perspective because it is based on the content they have contributed. Many users' perspectives are being considered: for instance, if a lot of users refer to works by two authors, this might be a hint that these authors

are similar and have similar research interests, and thus are potential collaboration partners. This social aspect is similar to author co-citation (McNee, 2006), where a third author co-cites two other authors, who might therefore be similar to each other. Therefore, the method of author co-citation analysis is aggregated to collaborative services. The difference is that the users of such services are not necessarily scientific authors, but users of the social web. Hence, considering web users not only includes the opinions of more people but leads to new relations between researchers and expands the view of a researcher's known social network. Here the focus lies not only on novice researchers, but also on senior scientists who have established their community network, as they also benefit from new ways of generating social information relations to point out new collaborators.

The following issue is related to this discussion: recommender systems not only help users filter information, but also aim to point out new relevant resources that the users would not have found by themselves (Ricci, Rokach, & Shapira, 2011, p. 5). This means that a user is given resources outside of their own individualized realm, for instance when a recommender system takes into account the web users' perspective. This counteracts the phenomenon of the "filter bubble", a term discussed by Pariser. Pariser refers to the personalization of retrieval systems, such as the search engine Google, and the danger of such systems influencing a user who is not aware of this influence. When this happens, a user becomes imprisoned in their own filter bubble, which means that they start to live in their own world and develop a limited view of the outside (Pariser, 2011). "Your filter bubble is your own personal unique universe of information that you live in online", says Pariser (Bosker, 25.05.2011). A recommender system has options to overcome or at least work against this filter bubble, for instance by considering other users' perspectives in collaborative filtering systems.

Figure 2.6 summarizes a proposal for expert recommendation based on diverse perspectives. Author co-citation only takes into account the perspective of a third author citing two authors. Bibliographic coupling only considers the

perspective of a target researcher, which is marked in their choice of references. Collaborative filtering based on the content of web users considers the users' perspective, which stands outside the "citation game" (compare Cronin (1984)).

Thus, regarding the users' perspective in an expert recommender system would be one option for expanding a target researcher's view of their work and network. For example, a user in an academic bookmarking system (introduced further on) – who might be a researcher or only a general reader (Haustein, 2012) – generates different relations between researchers by using other tags, and bookmarking other publications, than a target researcher would. These new relations lead to new recommendations for the target, who is made aware of new possible research partners and networks, out of which communities of practice can develop. Researchers who become aware of new connections have the chance to then participate and interact with their colleagues.

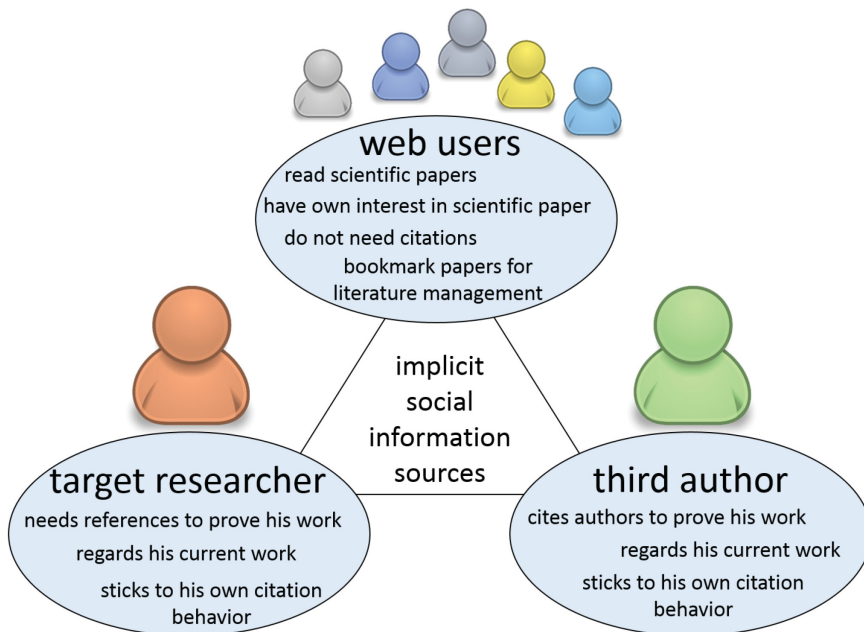


Figure 2.6. Source for determining implicit social information about a target researcher.

Figure 2.7 shows the various relations between two researchers based on social information. On the right-hand side, there is data from information services that deal with scientific publications, references, and citations. On the left-hand side, there is information from an academic social bookmarking system. Since authors are directly related to their publications, their bookmarked works establish indirect connections between their respective authors and the tags assigned to their publications, as well as the users who have bookmarked them. Third authors and references directly related to authors' publications establish

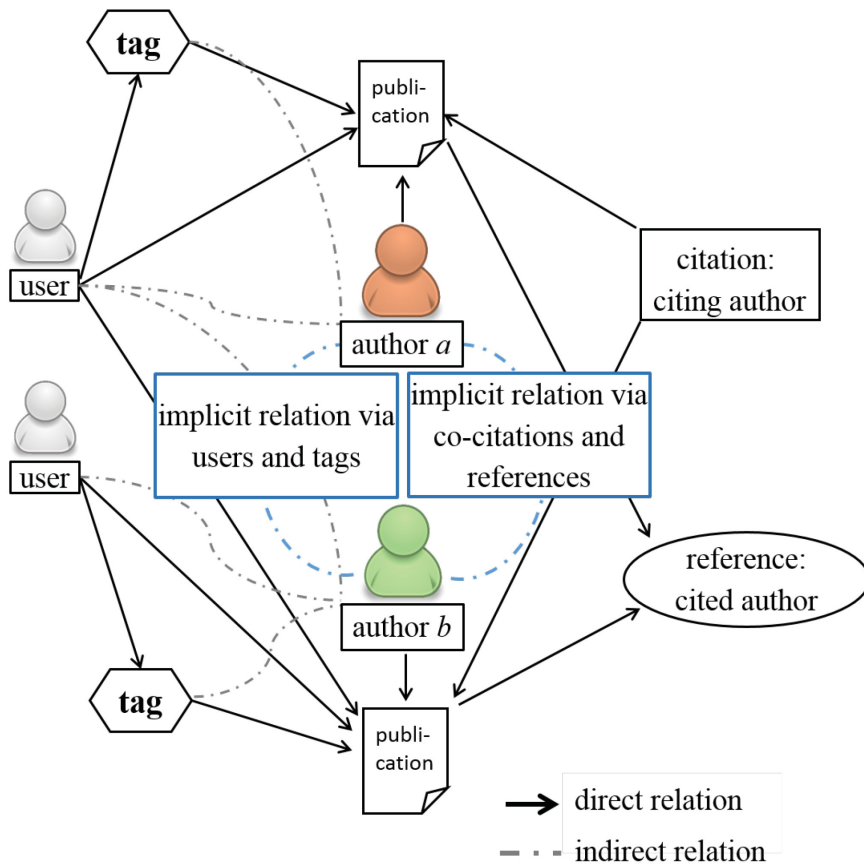


Figure 2.7. Relations between two authors considered for expert recommendation.

an indirect connection between two authors. The direct relations between authors, references and citations on the one hand, and bookmarks, users and tags on the other hand, lead to information about indirect, implicit relations between two authors that both are not yet aware of. Similarities can thus be based on common references, co-citations, users and tags.

Other studies consider diverse social information data, such as researchers' attendance of past conferences (Hornick & Tamayo, 2012). Ben Jabeur, Tamine, and Boughanem (2010) designate co-authorship and friendship as two additional social relations among researchers (see also Cabanac (2011)). However, both of these are direct relations that are explicit to a researcher. Recommendations based on these relations would probably suffer from over-specialization (Lops, Gemmis, & Semeraro, 2011; see also chapter 3). Furthermore, these relations do not build on collaborative filtering. The relations in figure 2.7 rely on the expanded perspectives introduced above, which are based on aspects of trust concerning researchers' reputations. Therefore, reference and citation data as well as the corresponding social information contributed by web users in bookmarking systems is applied. The latter set is derived from implicit relations on the internet that build on the concept of collaborative filtering.

2.3 Implicit Relations via Collaborative Filtering

Using information generated by web users draws on the principle of the "wisdom of the crowds" or "collective intelligence" (Surowiecki, 2005; Weiss, 2005). O'Reilly (2005) summarizes: "The central principle behind the success of the giants born in the Web 1.0 era who have survived to lead the Web 2.0 era appears to be this, that they have embraced the power of the web to harness collective intelligence." The assumption is that if many users, or a large user community, share one single opinion, this opinion is likely to be right. A crucial characteristic of this user community is that it does not have to share common goals, or any knowledge, but is smarter than any one individual (McFedries, 2006). Galton, who speaks of "vox populi", conducted the first experiment on this concept (Galton, 1907). Peters (2009) indicates that general aspects of

earlier assumptions about collective intelligence are discussed in research and adapted to new Web 2.0 environments (see, for example, Weiss (2005) for further discussions). Thus, one can also speak of “collaborative intelligence” ((Peters, 2009, p. 167), compare also Vander Wal (2008)).

Kozinets, Hemetsberger, and Schau (2008) speak of collective and individual creativity, whereas collective user creativity is distinct as it derives from “social interaction”, which offers new interpretations and discoveries. “We can say that collective creativity has occurred when social interactions between individuals trigger new interpretations and new discoveries of distant analogies that the individuals involved, thinking alone, could not have generated” (Hargadon & Bechky, 2006, p. 489). Web 2.0 environments allow users to participate: They can post comments, bookmark and tag resources, and rate resources (and users). Via these social interactions, users generate value for a Web 2.0 service as this service can then apply user data, for example to make recommendations. Furthermore, users collaboratively filter relevant information under the aspect of collective intelligence via their social interactions.

“Collaborative filtering simply means that people collaborate to help one another perform filtering by recording their reactions to documents they read” (Goldberg, Nichols, Oki, & Terry, 1992). The term is related to information filtering, with both concepts aiming at filtering on the basis of relevance. Information filtering in knowledge representation is concerned with determining terms (information filters) for indexing via a knowledge organization system (Stock & Stock, 2013, p. 527). In information retrieval, users set filters to receive information in a database, for example via selective dissemination of information (SDI) or really simple syndication (RSS). Thus, a user only receives the information on which the filters have been set. Belkin and Croft (1992) emphasize that this process can lead to drawbacks, as users tend to become passive in their information seeking behavior. Furthermore, the process of information filtering depends on a user’s settings. If these are inappropriate, the user will not find the information needed (Belkin, 1980). Here, collaborative filtering sets a counterexample because users are helped by

other people who collaboratively engage in the filtering process, for example by assigning tags.

The notion of collaborative filtering is based on the idea of a referral chain, in which a user requires his or her relations and contacts to obtain relevant information or find an expert in order to solve a problem. These relations can be found via similarity values. Kautz, Selman, and Shah (1997) mention a great advantage of collaborative filtering systems: “A user is only aware of a portion of the social network to which he or she belongs. By instantiating the larger community, the user can discover connections to people and information that would otherwise lay hidden over the horizon.” The same argument is true for researcher networks. Svensson, Laaksolahti, Höök, and Waern (2000) state: “By making other users’ action visible we can take advantage of the work they have done to find their way around and to solve problems.” Collaborative information systems show an inherent networking structure (Peters, 2009), which allows the discovery of such connections. They can be used to detect new relations and make people aware of them to expand their view – and, in the academic field to expand their network and be aware of new potential collaborators. The general claim for such approaches is: More like me – find trustworthy users who are similar to me so that I may get relevant information from them (Heck & Peters, 2010a, 2010b; Peters, 2009; Smith, Barash, Getoor, & Lauw, 2008). For a researcher recommender system, this principle is adapted to: More like me – find trustworthy potential collaborators based on the opinions of web user communities.

Social Bookmarking Systems and Folksonomies

Web 2.0 offers a realm where “services especially focalize communication and exchange of resources between users” (Peters, 2009, p. 14). This focus enables users to contribute to the development of services on the web while generating user data. Users increase the value of information in Web 2.0 services, either directly or indirectly (Tredinnick, 2006). Examples of the latter are analyzed to detect new implicit relations (see figure 2.7). Collaborative information systems

are particularly noteworthy among Web 2.0 services. Peters (2009) states that these services focus on resource management and allow personal as well as collaborative information creation. She further distinguishes sharing and social bookmarking offers, where the latter include any e-commerce and recommender services. A broader definition of collaborative information services considers the idea of collaborative filtering and includes all services from which collaborative filtering models can be derived. All recommendations systems containing user-generated data (with or without tags) belong to these services.

Social bookmarking systems offer platforms on which users can archive their references in order to have access to and manage them from any web-accessible device. Systems focus on specific purposes and user groups. Services like del.icio.us¹¹ offer a platform on which users can share links to interesting and relevant web pages and online resources. Besides these services, bookmarking systems for academic purposes established, as for example BibSonomy¹² (Hotho, Jäschke, Schmitz, & Stumme, 2006a, 2006b; Jäschke et al., 2007; Regulski, 2007; Schmitz, Hotho, Jäschke, & Stumme, 2006), CiteULike¹³ (Capocci & Caldarelli, 2008; Kipp, 2011a, 2011), and Mendeley¹⁴ (Lo Russo, Spolveri, Ciancio, & Mori, 2013; Reiswig, 2010). Connotea¹⁵ is another service analyzed in the case study in section 5.1. Unfortunately, this service was discontinued in 2013. Academic bookmarking services focus on the management of scientific literature. All systems rely on the aspects of collaborative filtering. In a bookmarking system, users bookmark references and are able to assign keywords (tags) to those references. Information stored in a bookmark varies according to the resource in question. For scientific literature, including articles, a bookmark contains the bibliographic data of a contribution (title, author names, journal, volume, issue etc.). Thus a user is provided with a

¹¹ <https://delicious.com>

¹² <http://www.bibsonomy.org>

¹³ <http://www.citeulike.org>

¹⁴ <https://www.mendeley.com>

¹⁵ <http://www.connotea.org/>

bookmark list for resources he or she has bookmarked. There is also a tag list, or tag cloud, containing tags previously used. The collaborative aspect means that all users' bookmarks and tags are available to all others. Users can search for bookmarks by others in order to find relevant literature, and they can add bookmarks by others to their own reference lists (also referred to as posts) and assign tags to them. Thus a bookmarking system becomes social, and is not only a reference management system for an individual but a collaborative system in which community users act via combined resources. They help to organize their own literature database. Hence, bookmarking systems exploit the main features of social networking (John & Seligmann, 2006).

title	issn
journal	isbn
abstract	publisher
username	year
tag	notes
surname (prefer author)	ownpub
author	bibkey
journal	doi possible
booktitle	PubMed ID possible

Table 2.1. Search fields in CiteULike. Retrieved 06/21/2016 from www.citeulike.org/search_help

Figure 2.8 shows the CiteULike bookmarking system, which was founded by Richard Cameron in 2004¹⁶. The profile of a user contains their bookmarks and tags. The user also sees which other users bookmarked the same literature, and which tags they used. The service allows a general search for bookmarks, but search fields can also be used (see table 2.1). CiteULike is constantly updating its search fields to not only offer a literature management system, but also

¹⁶ <http://www.citeulike.org/faq/faq.adp>

search functions comparable to a classic information service (Kipp, 2011a). However, detailed search field descriptions are not available yet. CiteULike also allows search via Boolean operators and wildcards. Currently, the service claims to have stored eight million articles (June 2015). The service Mendeley has recently become quite popular. It also offers a desktop application for diverse operation systems. In bookmarking services, users are able to establish groups in order to be able to find relevant literature via input from similarly interested peers. For their users, bookmarking systems replace the traditional offline reference management system that does not allow any collaborative action. Today, Mendeley and CiteULike cooperate with huge publishers, namely Elsevier (Mendeley) and Springer (CiteULike), respectively. Bookmarking services offer other diverse functions to their users. Peters (2009) describes various services in more detail, introducing BibSonomy as an example of a bookmarking system for storage of scientific literature.

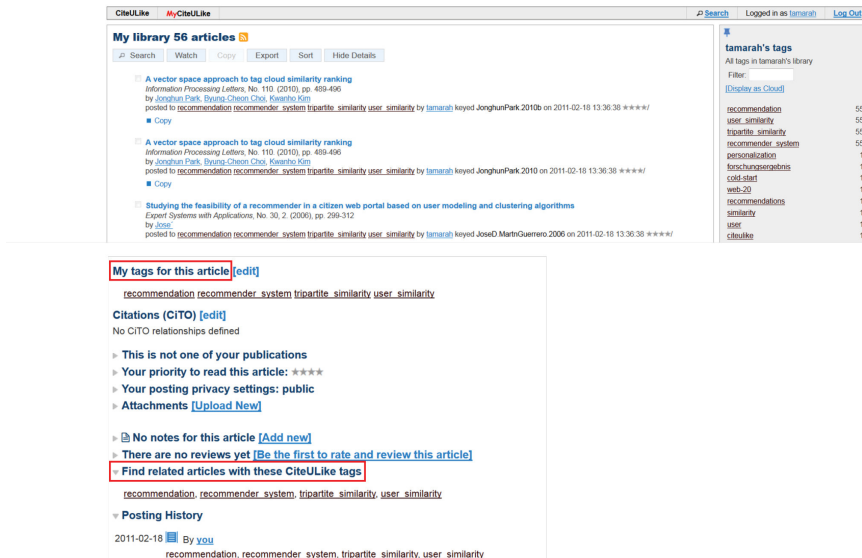


Figure 2.8. CiteULike user profile with bookmarks and tags (top). The bottom screenshot shows related articles (plus tags and other users) found via a user’s tag.

Collaborative database organization and analysis of relations, direct or indirect, are made possible by bookmarking systems' inclusion of a folksonomy structure. A folksonomy structure develops via user activities, although users do not explicitly have to interact with each other. Spiteri (2007) emphasizes "that folksonomies are created in an environment where, although people may not actively collaborate in their creation and assignment of tags, they may certainly access and use tags assigned by others." This aspect stresses the importance of the user perspective (figure 2.6), as users do not have any specific intentions concerning collaboration, as may be the case when using references and citations. In fact, users bookmark resources for personal knowledge management first and foremost (Marlow, Naaman, Boyd, & Davis, 2006a, 2006b). Sinha (2006) states that "[...] tagging captures our individual conceptual associations, but does not force us to categorize. It enables loose coordination, but does not enforce the same interpretation of a concept." Although this fact leads to some difficulties concerning the detection of tagging similarity (see chapter 3), tagging represents the opinions of users.

Peters (2009, p. 155) defines a folksonomy as consisting of "freely selectable keywords, or tags, which can be liberally attached to any information source." Similarly, Golder and Hubermann (2006, p. 198) state that "collaborative tagging describes the process by which many users add metadata in the form of keywords to shared content." Thus a folksonomy includes user-generated data. Collaborative filtering models for recommender systems use the relations between this data, including users, resources (also referred to as items) and tags, to measure similarity between them. For example, the relational structure of users, items and tags can be represented in a graph (Balby Marinho et al., 2012), where the nodes are users, items, or tags and the edges represent relations. 'User-item relation' means that a user has bookmarked or purchased an item, 'tag-item relation' means that a tag was assigned to an item, and so on. Thus a folksonomy – or, more specifically, a dataset based on a folksonomy structure – can be defined as a tuple $F = (U, T, R, Y)$, where U , T and R are finite sets with the elements of 'user name', 'tag' and 'resource', and Y is a ternary relation between them: $Y \subseteq U \times T \times R$, with the elements being called 'tag actions' or

‘assignments’ (Balby Marinho et al., 2011; Jäschke, Marinho, Hotho, Schmidt-Thieme, & Stumme, 2007). In the case study in chapter 4.3, the folksonomy tuple is expanded to use the structure for expert recommendation.

Relations between all elements are possible and can be analyzed (see chapter 3). Here the tags of a folksonomy play an important role because they deliver more content to resources. Tags can be categorized according to users’ needs and tagging behavior. Categorization happens retroactively and is not chosen by users during the tagging process (Peters, 2009). Peters (2009, pp. 196–203), giving a detailed description of diverse categorization proposals, suggests using the most frequently recognized categories from tag analysis studies on diverse platforms, such as people, things, and events. However, the main issue in using tag relations between users and items is finding out what a tag means. A twofold distinction of tags is possible, where one class refers to a resource’s aboutness – similarly to the document aboutness discussed in knowledge representation (Stock & Stock, 2013, p. 519) – and the other class exists independently from this issue. “Aboutness tags” (Peters, 2009) describe a resource and its content. Pluzhenskaia (2006) (see also Kipp, 2006a) emphasizes that these tags are independent from a single user’s context and thus understandable by all other users. Individual tags like “interesting”, “funny” or “toread” do not have any shared meaning and are strongly user-centric (Brooks & Montanez, 2006; Kipp, 2006a). Kipp (2006a) also speaks of ‘affective’ or ‘emotional’ tags by individual users. In contrast, tags describing a resource’s aboutness can act as additional indexing terms for resources. These tags are able to show topical relations between two resources.

Figure 2.9 shows three main diverse tag types according to the tag categories proposed by Sen et al. (2006) (see also Al-Khalifa & Davis, 2007). The first type is unsuitable for detecting topical relations. The second type shows a user’s opinion. This type is quite similar to a user’s rating as regarded in recommender systems. Such tags could therefore be used as a sort of rating in a collaborative filtering approach (see chapter 3). The latter is the most important type concerning the detection of resource topics and similar user interests.

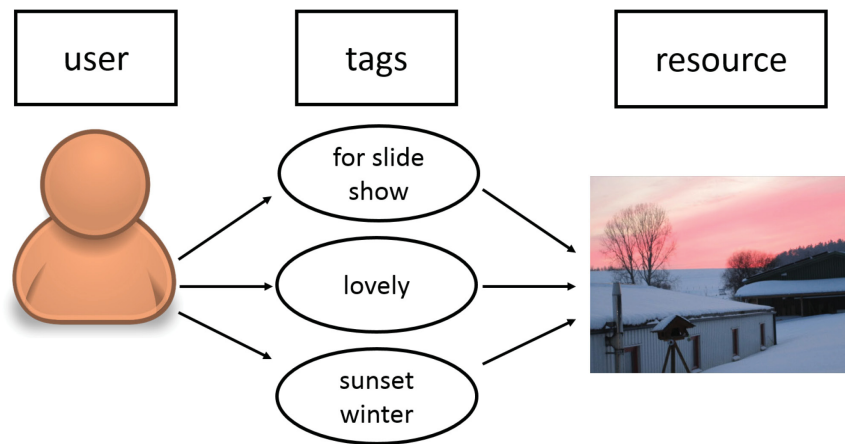


Figure 2.9. Distinction between diverse tag types adapted from Sen et al. (2006).

This means that if two resources share the same tags, they are probably quite similar with regard to their content. Such topical relations can then be used for recommendation (Capocci & Caldarelli, 2008; Kammergruber, Viermetz, & Ziegler, 2009). For knowledge representation, Peters (2009) discusses the use of such tags in order to complement classic indexing with concepts. Several studies explored approaches to creating knowledge organization systems from folksonomies or expanding other systems to make them more acceptable to common users.

However, such topical relations based on common tags are only possible in broad folksonomy structures (Vander Wal, 2005a, 2005b). Broad folksonomies allow users to assign a common tag to one single resource several times over. Hence, the importance of a tag for a resource can be measured with regard to its frequency of assignment by diverse users (Marlow et al., 2006a), and tag distributions are possible (Peters, 2012). In narrow folksonomies, on the other hand, a tag can be assigned only once by one single user. Other users can see that tag, but are not able to assign it to the same resource. A resource folksonomy then includes diverse individual tags, but does not reflect the users'

choice and tagging behavior. Hence, Smith (2008, p. 62) states that only broad folksonomy structures refer to “collaborative tagging”, while narrow folksonomies allow “simple tagging”. Quintarelli summarizes the potential of broad folksonomies for the detection of topical relations: “The power of folksonomies is connected to the act of aggregating, not simply to the creation of tags. [...] The term-significance relationship emerges by means of an implicit contract between the users”. The academic bookmarking systems named allow multiple tagging and have broad folksonomies. However, services such as CiteULike also allow users to assign private tags to resources that are not searchable for others.

Tag categories and user behavior depend on the collaborative system’s environment (Golbeck, Koepfler, & Emmerling, 2011). Studies show differences within the scientific field, for example. Heckner, Mühlbacher, and Wolff (2008) analyzed tags from Connotea and state that 92% of the tags refer to a resource’s aboutness or describe the resource type (such as article, or book). Thus, the different tagging behaviors for scientific resources mean that tags in academic social bookmarking systems are appropriate for the detection of topical relations. In her study, Kipp (2011a) compared author keywords and descriptor terms derived from PubMed with tags assigned to the same article in CiteULike. She concludes that many author keywords and terms are equivalent, but that tags also include terms that were not found in the author keyword collection. The latter tags include additional information beyond the authors’ perspectives. These results hold true for the consideration of the web users’ perspective (figure 2.6).

To summarize the findings, the generation of scientific networks to detect collaborators should take into account a researcher’s reputation with regard to the aspect of trust. Furthermore, it should focus on the detection of implicit relations with regard to expanding a researcher’s network and foster the establishment of new communities of practice. Expansion requires diverse perspectives to be considered – a target researcher, their colleagues, and web users. Sources for gathering this data are professional information services,

which offer valuable and valid citation data, as well as social bookmarking services, which correspond to classic citation structure, but take into account the concept of the “wisdom of the crowd”.

Based on the concept of collective intelligence and the notions of similarity relations between users, items and tags, diverse collaborative filtering models established in the context of recommender systems. In chapter 3, collaborative filtering is discussed as one basic approach to designing a recommendation system¹⁷. Recommender systems aim at suggesting personalized recommendations based on historic user data. Social information about researchers is historic data, either generated by researchers (references and citation) or web users (bookmarks and tags). Thus the concept of recommender systems is suitable and adaptable for expert recommendation.

References

- Alguliev, R., Aliguliyev, R., & Ganjaliyev, F. (2011). Investigation of the role of similarity measure and ranking algorithm in mining social network. *Journal of Information Science*, 37(3), 229–234. doi:10.1177/0165551511400946
- Al-Khalifa, H. S., & Davis, H. C. (2007). Towards better understanding of folksonomic patterns. In S. Harper, H. Ashman, M. Bernstein, A. Cristea, H. C. Davis, P. de Bra, . . . D. Millard (Eds.): *Proceedings of the 18th Conference on Hypertext and Hypermedia (HT)* (pp. 163–166). New York, NY, USA: ACM. doi:10.1145/1286240.1286288
- Balby Marinho, L., Hotho, A., Jäschke, R., Nanopoulos, A., Rendle, S., Schmidt-Thieme, L., . . . Symeonidis, P. (2012). *Recommender Systems for Social Tagging Systems*. Boston, MA, USA: Springer.
- Balby Marinho, L., Nanopoulos, A., Schmidt-Thieme, L., Jäschke, R., Hotho, A., Stumme, G., & Symeonidis, P. (2011). Social tagging recommender systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 615–644). New York, NY, USA:

¹⁷ Note: The terms “recommender system” and “collaborative filtering” are sometimes used as synonyms (Höhfeld and Kwiatkowski (2007)). However, not all recommender systems rely on collaborative user opinions (see chapter 3 and statement in Resnick and Varian (1997)).

- Springer. doi:10.1007/978-0-387-85820-3_19
- Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, 5, 133–143.
- Belkin, N. J., & Croft, B. W. (1992). Information filtering and information retrieval: two sides of the same coin? *Communications of the ACM*, 35(12), 29–38. doi:10.1145/138859.138861
- Ben Jabeur, L., Tamine, L., & Boughanem, M. (2010). A social model for literature access: Towards a weighted social network of authors. In *Proceedings of 9th International Conference on Adaptivity, Personalization and Fusion of Heterogeneous Information (RIA0)* (pp. 32–39). Retrieved from <http://dl.acm.org/citation.cfm?id=1937063>
- Blazek, R. (2007). *Author-statement citation Analysis Applied as a Recommender System to Support Non-domain-expert Academic Research: Doctoral dissertation*. Nova Southeastern University. Retrieved from http://nsuworks.nova.edu/gscis_etd/416
- Bollacker, K. D., Giles, C. L., & Lawrence, S. (1998). CiteSeer: An automatic citation indexing system. In *Proceedings of the 3rd Conference on Digital Libraries (DL)* (pp. 89–98). New York, NY, USA: ACM. doi:10.1145/276675.276685
- Bosker, B. (2011, May 25). Facebook, Google giving us information junk food, Eli Pariser warns. *Huffington Post Online*. Retrieved from http://www.huffingtonpost.com/2011/03/07/eli-pariser-facebook-google-ted_n_832198.html
- Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, 61(12), 2389–2404. doi:10.1002/asi.21419
- Brooks, C. H., & Montanez, N. (2006). An analysis of the effectiveness of tagging in blogs. In *Computational Approaches to Analyzing Weblogs. Papers from the AAAI Spring Symposium* (pp. 9–14). AAAI. Retrieved from <http://www.aaai.org/Papers/Symposia/Spring/2006/SS-06-03/SS06-03-002.pdf>
- Cabanac, G. (2011). Accuracy of inter-researcher similarity measures based on topical and social clues. *Scientometrics*, 87(3), 597–620. doi:10.1007/s11192-011-0358-1
- Capocci, A., & Caldarelli, G. (2008). Folksonomies and clustering in the collaborative system CiteULike. *Journal of Physics A: Mathematical and Theoretical*, 41(22), 224016. doi:10.1088/1751-8113/41/22/224016
- Cawkell, T. (2000). Methods of information retrieval using Web of Science: pulmonary hypertension as a subject example. *Journal of Information*

- Science*, 26(1), 66–70. doi:10.1177/016555150002600107
- Chadegani, A. A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). A Comparison between two main academic literature collections: Web of Science and Scopus databases. *Asian Social Science*, 9(5), 18–26. Retrieved from <http://ssrn.com/abstract=2257540>
- Cronin, B. (1984). *The Citation Process. The Role and Significance of Citations in Scientific Communication*. London, UK: Taylor Graham.
- Cronin, B., & Atkins, H. (Eds.). (2000). *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*. Medford, NJ: Information Today.
- Cruz, C. C. P., Motta, Claudia, L. R., Santoro, F. M., & Elia, M. (2009). Applying reputation mechanisms in communities of practice: A case study. *Journal of Universal Computer Science*, 15(9), 1886–1906. doi:10.3217/jucs-015-09-1886
- Elsevier. (2016). *Scopus Content Coverage Guide*. Retrieved from https://www.elsevier.com/__data/assets/pdf_file/0007/69451/scopus_content_coverage_guide.pdf
- Fogg, B. J. (2002). Persuasive technology: Using computers to change what we think and do. *Ubiquity*, (Dec), 5. doi:10.1145/764008.763957
- Fogg, B. J., Lee, E., & Marshall, J. (2002). Interactive technology and persuasion. In *The Persuasion Handbook: Developments in Theory and Practice* (pp. 765–788). Thousand Oaks, CA: SAGE Publications. doi:10.4135/9781412976046.n34
- Galton, F. (1907). Vox Populi. *Nature*, 75(No 1949), 450–451. Retrieved from <http://wisdomofcrowds.blogspot.de/2009/12/vox-populi-sir-francis-galton.html>
- Gambetta, D. (Ed.). (1988). *Trust: Making and Breaking Cooperative Relations*. Oxford, UK: Blackwell.
- Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, 122(3159), 108–111. doi:10.1126/science.122.3159.108
- Garfield, E., & Stock, W. G. (2002). Citation consciousness. *Password*, (No 6), 22–25.
- Gmür, M. (2003). Co-citation analysis and the search for invisible colleges: A methodological evaluation. *Scientometrics*, 57(1), 27–57. doi:10.1023/A:1023619503005
- Golbeck, J., Koepfler, J., & Emmerling, B. (2011). An experimental study of social tagging behavior and image content. *Journal of the American Society for Information Science and Technology*, 62(9), 1750–1760. doi:10.1002/asi.21522
- Goldberg, D., Nichols, D. M., Oki, B. M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of*

- the ACM*, 35(12), 61–70. doi:10.1145/138859.138867
- Golder, S. A., & Hubermann, B. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2), 198–208. doi:10.1177/0165551506062337
- Hargadon, A. B., & Bechky, B. A. (2006). When collections of creatives become creative collectives: A field study of problem solving at work. *Organization Science*, 17(4), 484–500. doi:10.1287/orsc.1060.0200
- Haustein, S. (2012). *Multidimensional Journal Evaluation: Analyzing Scientific Periodicals beyond the Impact Factor* (1. Aufl.). *Knowledge and Information*. Berlin/Boston: De Gruyter Saur.
- Heck, T., & Peters, I. (2010a). Experten-Empfehlungen mit Social Bookmarking-Services. *i-com -Zeitschrift für interaktive und kooperative Medien*, 9(3), 7–11. doi:10.1524/icom.2010.0027
- Heck, T., & Peters, I. (2010b). Implizite Digitale Soziale Netze als Basis für Expertenempfehlungssysteme. In K. P. Fähnrich & B. Franczyk (Eds.): *Informatik 2010, Service Science – Neue Perspektiven für die Informatik (1). Lecture Notes in Informatics* (pp. 613–618). Bonn, Germany: Köllen Druck+Verlag. Retrieved from <https://www.gi.de/service/publikationen/lni.html>
- Heckner, M., Mühlbacher, S., & Wolff, C. (2008). Tagging tagging. Analysing user keywords in scientific bibliography management systems. *Journal of Digital Information*, 9(2). Retrieved from <https://journals.tdl.org/jodi/index.php/jodi/article/view/246>
- Höfeld, S., & Kwiatkowski, M. (2007). Empfehlungssysteme aus informationswissenschaftlicher Sicht – State of the Art. *Information - Wissenschaft und Praxis*, 58(5), 265–276. Retrieved from <http://docplayer.org/5309215-Empfehlungssysteme-aus-informationswissenschaftlicher.html>
- Hornick, M., & Tamayo, P. (2012). Extending recommender systems for disjoint user/item sets: The conference recommendation problem. *IEEE Transactions on Knowledge and Data Engineering*, 24(8), 1478–1490. doi:10.1109/TKDE.2011.90
- Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006a). BibSonomy: A social bookmark and publication sharing system. In *Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures (ICCS)* (pp. 87–102). Aalborg, Denmark: Aalborg University Press. Retrieved from <http://conceptualstructures.org/cstools/cstiwl/proceedings.htm>
- Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006b). Das Entstehen von Semantik in BibSonomy. In *Social Software in der Wertschöpfung*. Baden-Baden, Germany: Nomos. Retrieved from <http://www.kde.cs.uni->

- kassel.de/stumme/papers/2006/hotho2006entstehen.pdf
- Jäschke, R., Grahl, M., Hotho, A., Krause, B., Schmitz, C., & Stumme, G. (2007). Organizing publications and bookmarks in BibSonomy. In *Workshop on Social and Collaborative Construction of Structured Knowledge (CKC 2007) at WWW'07*. Retrieved from http://www2007.org/workshops/paper_25.pdf
- Jäschke, R., Marinho, L. B., Hotho, A., Schmidt-Thieme, L., & Stumme, G. (2007). Tag recommendations in folksonomies. In J. N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenic, & A. Skowron (Eds.): *Vol. 4702. Lecture Notes in Computer Science. Knowledge Discovery in Databases: PKDD* (pp. 506–514). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-74976-9_52
- John, A., & Seligmann, D. (2006). Collaborative tagging and expertise in the enterprise. In *Proceedings of the Collaborative Web Tagging Workshop at the 15th International Conference on World Wide Web (WWW)*. New York, NY, USA: ACM. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.134.296>
- Josang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2), 618–644. doi:10.1016/j.dss.2005.05.019
- Kammergruber, W. C., Viermetz, M., & Ziegler, C.-N. (2009). Discovering communities of interest in a tagged on-line environment. In *Proceedings of the International Conference on Computational Aspects of Social Networks (CASoN)* (pp. 143–148). Washington, D.C., USA: IEEE. doi:10.1109/CASoN.2009.22
- Kautz, H., Selman, B., & Shah, M. (1997). Referral Web: combining social networks and collaborative filtering. *Communications of the ACM*, 40(3), 63–65. doi:10.1145/245108.245123
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10–25. doi:10.1002/asi.5090140103
- Kipp, M. E. (2006a). @toread and Cool: Tagging for time, task and emotion. In *Proceedings of the 17th Annual ASIS&T SIG/CR Classification Research Workshop*. Austin, TX, USA. Retrieved from <http://eprints.rclis.org/13909/>
- Kipp, M. E. (2006b). Complementary or discrete contexts in online indexing: A comparison of user, creator, and intermediary keywords. *Canadian Journal of Information and Library Science*, 29(4), 419–436. Retrieved from <http://hdl.handle.net/10760/8771>
- Kipp, M. E. (2011a). Tagging of biomedical articles on CiteULike: A comparison of user, author and professional indexing. *Knowledge Organization*, 38(3), 245–261. Retrieved from <http://hdl.handle.net/10760/15227>

- Kipp, M. E. (2011). User, author and professional indexing in context: An exploration of tagging practices on CiteULike. *Canadian Journal of Information and Library Science*, 35(1), 17–48. Retrieved from <http://hdl.handle.net/10760/14867>
- Kozinets, R. V., Hemetsberger, A., & Schau, H. J. (2008). The wisdom of consumer crowds: Collective innovation in the age of networked marketing. *Journal of Macromarketing*, 28(4), 339–354. doi:10.1177/0276146708325382
- Leydesdorff, L. (2005). Similarity measures, author cocitation analysis, and information theory. *Journal of the American Society for Information Science and Technology*, 56(7), 769–772. doi:10.1002/asi.v56:7
- Li, J., Burnham, J. F., Lemley, T., & Britton, R. M. (2010). Citation analysis: Comparison of Web of Science, Scopus SciFinder, and Google Scholar. *Journal of Electronic Resources in Medical Libraries*, 7(3), 196–217. doi:10.1080/15424065.2010.505518
- Linde, F., & Stock, W. G. (2011). *Information Markets: A Strategic Guideline for the I-Commerce* (Knowledge & Information. Studies in Information Science). Berlin, New York, NY: De Gruyter Saur.
- Lo Russo, G., Spolveri, F., Ciancio, F., & Mori, A. (2013). Mendeley: an easy way to manage, share, and synchronize papers and citations. *Plastic and reconstructive surgery*, 131(6), 946. doi:10.1097/PRS.0b013e31828bd400
- Lops, P., Gemmis, M. de, & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 73–105). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_3
- Marlow, C., Naaman, M., Boyd, D., & Davis, M. (2006a). HT06, tagging paper, taxonomy, Flickr, academic article, to read. In U. K. Wiil, P. J. Nürnberg, & J. Rubart (Eds.): *Proceedings of the 17th Conference on Hypertext and Hypermedia (HYPERTEXT)* (pp. 31–40). New York, NY, USA: ACM. doi:10.1145/1149941.1149949
- Marlow, C., Naaman, M., Boyd, D., & Davis, M. (2006b). Position paper, tagging, taxonomy, flickr, article, to read. In *Proceedings of the Collaborative Web Tagging Workshop at the 15th International Conference on World Wide Web (WWW)*. New York, NY, USA: ACM. Retrieved from <http://www.danah.org/papers/WWW2006.pdf>
- Marshakova, I. (1973). Система связей между документами, построенная на основе ссылок: по данным Science Citation Index [A System of document connections based on references (data of SCI)]. *Nauchno-Tekhnicheskaya Informatsiya Seriya 2/Scientific and Technical Information Processing*, 2(6), 3–8. Retrieved from <http://garfield.library.upenn.edu/marshakova/marshakovanauchtechn1973.pd>

- f
- McFedries, P. (2006). Folk Wisdom. *IEEE Spectrum*, 43(2), 80.
doi:10.1109/MSPEC.2006.1584367
- McNee, S. M. (2006). *Meeting User Information Needs in Recommender Systems: Doctoral dissertation*. Minnesota, MN, USA: University of Minnesota. Retrieved from <http://www-users.cs.umn.edu/~mcnee/thesis-preprint.pdf>
- Montaner, M., López, B., & de la Rosa, J. L. (2002). Opinion-based filtering through trust. In G. Goos, J. Hartmanis, J. Van Leeuwen, M. Klusch, S. Ossowski, & O. Shehory (Eds.): *Lecture Notes in Computer Science: Vol. 2446. Cooperative Information Agents VI. 6th International Workshop, CIA'02* (pp. 164–178). Berlin/Heidelberg, Germany: Springer.
doi:10.1007/3-540-45741-0_15
- O'Reilly, T. (2005). *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. Retrieved from <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>
- Pariser, E. (2011). *The filter bubble: What the Internet Is Hiding from You*. New York, NY, USA: Penguin Press.
- Persson, O. (2001). All author citations versus first author citations. *Scientometrics*, 50(2), 339–344. doi:10.1023/A:1010534009428
- Peters, I. (2009). *Folksonomies. Indexing and Retrieval in Web 2.0*. Berlin, Germany: De Gruyter Saur.
- Peters, I. (2012). Power tags as tools for social knowledge organization systems. In W. A. Gaul, A. Geyer-Schulz, L. Schmidt-Thieme, & J. Kunze (Eds.): *Studies in Classification, Data Analysis, and Knowledge Organization. Challenges at the Interface of Data Analysis, Computer Science, and Optimization* (pp. 281–289). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-642-24466-7_29
- Pluzhenskaia, M. (2006). Folksonomies or fauxsonomies: How social is social bookmarking? In J. Tennis & J. Furner (Eds.): *Proceedings of the 17th ASIS SIG/CR Classification Research Workshop* (pp. 23–24).
doi:10.7152/acro.v17i1.12497
- Quintarelli, E. *Folksonomies: power to the people*. Meeting of ISKO Italia – UniMIB, Milan, Italia. Retrieved from <http://www.iskoi.org/doc/folksonomies.htm>
- Regulski, K. (2007). Aufwand und Nutzen beim Einsatz von Social-Bookmarking-Services als Nachweisinstrument für wissenschaftliche Forschungsartikel am Beispiel von BibSonomy. *Bibliothek Forschung und Praxis*, 31(2), 177–184. doi:10.1515/BFUP.2007.177
- Reichling, T., Veith, M., & Wulf, V. (2007). Expert recommender: Designing for a network organization. *Computer Supported Cooperative Work*

- (CSCW), 16(4-5), 431–465. doi:10.1007/s10606-007-9055-2
- Reiswig, J. (2010). Mendeley. *Journal of the Medical Library Association : JMLA*, 98(2), 193–194. doi:10.3163/1536-5050.98.2.021
- Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56–58. doi:10.1145/245108.245121
- Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 1–35). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_1
- Schmitz, C., Hotho, A., Jäschke, R., & Stumme, G. (2006). Kollaboratives Wissensmanagement. In T. Pellegrini & A. Blumauer (Eds.): *Semantic Web. Wege zur vernetzten Wissensgesellschaft* (pp. 273–289). Berlin/Heidelberg, Germany: Springer. doi:10.1007/3-540-29325-6_18
- Schneider, J. W., & Borlund, P. (2007a). Matrix comparison, Part 1: Motivation and important issues for measuring the resemblance between proximity measures or ordination results. *Journal of the American Society for Information Science and Technology*, 58(11), 1586–1596. doi:10.1002/asi.20643
- Schneider, J. W., & Borlund, P. (2007b). Matrix comparison, Part 2: Measuring the resemblance between proximity measures or ordination results by use of the mantel and procrustes statistics. *Journal of the American Society for Information Science and Technology*, 58(11), 1596–1609. doi:10.1002/asi.v58:11
- Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., . . . Riedl, J. T. (2006). Tagging, communities, vocabulary, evolution. In *Proceedings of the 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW)* (pp. 181–190). New York, NY, USA: ACM. doi:10.1145/1180875.1180904
- Sinha, R. R. (2006). *A social analysis of tagging*. Retrieved from <http://rashmisinha.com/2006/01/18/a-social-analysis-of-tagging/>
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265–269. doi:10.1002/asi.4630240406
- Smith, G. (2008). *Tagging: People-Powered Metadata for the Social Web*. Berkeley, CA, USA: New Riders.
- Smith, M., Barash, V., Getoor, L., & Lauw, H. W. (2008). Leveraging social context for searching social media. In I. Soboroff, E. Agichtein, & R. Kumar (Eds.): *Proceedings of the ACM Workshop on Search in Social Media (SSM)* (pp. 91–94). New York, NY, USA: ACM. doi:10.1145/1458583.1458602
- Spiteri, L. F. (2007). Structure and form of folksonomy tags: The road to the public library catalogue. *Webology*, 4(2). Retrieved from

- <http://www.webology.org/2007/v4n2/a41.html>
- Stock, W. G. (1999). Web of Science. Ein Netz wissenschaftlicher Informationen - gesponnen aus Fußnoten [Web of Science. A web of scientific information - cocooned from footnotes]. *Password*, (no. 7+8), 21–25.
- Stock, W. G., & Stock, M. (2013). *Handbook of Information Science*. Berlin/Boston: De Gruyter.
- Surowiecki, J. (2005). *The Wisdom of Crowds: Why the Many are Smarter Than the Few and how Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York, NY, USA: Anchor Books.
- Svensson, M., Laakolahti, J., Höök, K., & Waern, A. (2000). A recipe based on-line food store. In D. Riecken, D. Benyon, & H. Lieberman (Eds.): *Proceedings of the International Conference on Intelligent User Interfaces (IUI)* (pp. 260–263). New York, NY, USA: ACM. doi:10.1145/325737.325866
- Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). ArnetMiner. In Y. Li, B. Liu, & S. Sarawagi (Eds.): *Proceedings of the 14th International Conference on Knowledge Discovery and Data Mining (SIGKDD)* (pp. 990–998). New York, NY, USA: ACM. doi:10.1145/1401890.1402008
- Thomson Reuters. (2016). *Web of Science Fact Sheet*. Retrieved from http://wokinfo.com/training_support/guidesfactsheets/
- Tredinnick, L. (2006). Web 2.0 and business: A pointer to the intranets of the future? *Business Information Review*, 23(4), 228–234. doi:10.1177/0266382106072239
- Vander Wal, T. (2005a). *Explaining and showing broad and narrow folksonomies*. Retrieved from <http://www.vanderwal.net/random/entrysel.php?blog=1635>
- Vander Wal, T. (2005b). *Folksonomy Explanations*. Retrieved from <http://www.vanderwal.net/random/entrysel.php?blog=1622>
- Vander Wal, T. (2008). Welcome to the matrix! In B. Gaiser, T. Hampel, & S. Panke (Eds.): *Medien in der Wissenschaft: Bd. 47. Good Tags - Bad Tags. Social Tagging in der Wissensorganisation* (pp. 7–9). Münster, New York, NY: Waxmann.
- Weiss, A. (2005). The power of collective intelligence. *netWorker*, 9(3), 16–23. doi:10.1145/1086762.1086763
- Wenger, E. (2008). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, MA, USA: Cambridge Univ. Press.
- White, H. D., & Griffith, B. C. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32(3), 163–171. doi:10.1002/asi.4630320302
- White, H. D., & Griffith, B. C. (1982). Authors as markers of intellectual space:

- Co-citation in studies of science, technology and society. *Journal of Documentation*, 38(4), 255–272. doi:10.1108/eb026731
- Yoo, K.-H., & Gretzel, U. (2011). Creating more credible and persuasive recommender systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 455–477). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_14
- Zhao, D., & Strotmann, A. (2007). All-author vs. first author co-citation analysis of the information science field using Scopus. In *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology*. doi:10.1002/meet.1450440262
- Zhao, D., & Strotmann, A. (2011). Counting first, last, or all authors in citation analysis: A comprehensive comparison in the highly collaborative stem cell research field. *Journal of the American Society for Information Science and Technology*, 62(4), 654–676. doi:10.1002/asi.21495
- Ziegler, C.-N., & Lausen, G. (2004). Analyzing correlation between trust and user similarity in online communities. In C. Jensen, S. Poslad, & T. Dimitrakos (Eds.): *Vol. 2995. Lecture Notes in Computer Science. Proceedings of 2nd International Conference on Trust Management (iTrust)* (pp. 251–265). New York, NY, USA: ACM. doi:10.1007/978-3-540-24747-0_19

3 Introduction to Recommender Systems

The idea of recommendations is to support individuals in their decision-making. People often solicit recommendations or advice from others in order to have a basis for action. “Recommender systems assist and augment this natural social process” (Resnick & Varian, 1997, pp. 56–58). The simplest means of defining a recommender or recommendation system is to say that it recommends something. According to this description, a system is a recommender system when it recommends something, independently of the technique it uses or the source from which the recommendation is generated. A person can recommend a book to a friend, for instance, or create a movie recommendation list including all of their favorite films, assuming that others may also like these movies because he likes them. These kinds of recommendations are based on a personal assessment and valuation, and have always existed. Gärtner (2012) names them “classical” recommendations and “mouth-to-mouth recommendations”. Other terms used are “word of mouth” (Ahrens, 2011; Shardanand & Maes, 1995), a term also used in the marketing field and discussed as persuasive non-commercial consumer communication (Ahrens, 2011).

Recommendations can be personalized or non-personalized. Mouth-to-mouth recommendations may be personalized if the individual giving the recommendation is aware of the preferences the person receiving the recommendation might have. However, if the person giving the recommendations has no specific receiver in mind, the recommendations will be non-personalized. Taking the above example of movies, the person making this list may think that their list might be interesting for other people in general, not just one specific person. Other examples of non-personalized recommendations include lists of bestselling books such as the Spiegel Bestseller List¹, or a list of “most read articles” on the website of a newspaper.

¹ <http://www.spiegel.de/kultur/bestseller-buecher-belletristik-sachbuch-auf-spiegel-liste-a-458623.html>

By contrast, recommender system research is mainly concerned with personalized recommendations based on automatic or semi-automatic processes (Ricci, Rokach, & Shapira, 2011). To give personalized recommendations, it is essential that the recommender system have access to information about the user or user group receiving the recommendations. Herein lies the challenge in recommender system research.

Recommender system research derives from the field of information retrieval (Jannach, Zanker, Felfernig, & Friedrich, 2011). Both research fields aim to help users in finding relevant resources in a satisfying and best supportive way (Ricci, Rokach, Shapira, & Kantor, 2011, p. 1). Chronologically, recommender or recommendation systems developed after the first retrieval systems. According to Klahold (2009), the first work to describe the function of a recommender system is Luhn's "A business intelligence system" (Luhn, 1958). Although Luhn himself did not name his system a recommender system, he (1958, p. 316) describes the "selective dissemination of new information", which means that information should be distributed to specific points in a system. Profiles and documents are thus compared on the basis of their similarity or concordance. This process is very similar to today's recommender systems (Klahold, 2009). However, Luhn calls his system a retrieval system, and considering the current state of research, recommender and retrieval system research focus on slightly different intentions, respectively. Recommender system research has developed its own field since the 1990s, where diverse recommender systems were developed (Ricci et al., 2011). The system "Tapestry" (Goldberg, Nichols, Oki, & Terry, 1992), an early example of an information filtering system, tries to recommend only relevant documents to a target user based on the principle of collaborative filtering. "GroupLens" (Resnick, Iacovou, Suchak, Bergstrom, & Riedl, 1994) is another popular example of a collaborative filtering system for news articles, while the early content-based filtering system "The Information Lens" tries to filter mail messages (Malone, Grant, & Turbak, 1986). (Klahold, 2009) shows a timeline providing an overview of recommender systems developed between 1986 and 2004. Outside of this research field, new sub-areas have developed in

information retrieval, such as expertise retrieval, which is discussed further on. Expertise retrieval and expert recommendation share the same intentions, although the models and techniques differ.

What both retrieval as well as recommender systems have in common is that they aim to help users find relevant resources. Due to the steadily growing web and increasing amount of resources available, this goal appears to be the most relevant and important. However, recommender systems concentrate on different aspects than classical retrieval systems do in order to satisfy user needs and solve the task of overcoming information overload. While the focus of recommender systems lies on personalized recommendations, there is also a trend towards personalized search results in retrieval systems such as search engines (Das et al., 2007). Hence, retrieval systems and recommender systems might adjust to one another (compare Jannach et al., 2011, p. 63) in future.

On the one hand, there are web users who search for specific content to satisfy their user needs and are experiencing difficulties. For example, a user may want to learn about different types of strawberries, and thus searches for relevant and adequate information on this topic. Offering relevant information of this kind is a task for information retrieval research. On the other hand, there are users looking for inspiration without knowing exactly what their needs are. A user who loves reading thrillers may want to find new books in this genre for future reading, but does not have any specific book titles in mind. The user's problem might be how to formulate an appropriate search question that results in titles of previously unknown thrillers. A search for "thriller AND book" might lead to some good results, but retrieval precision will surely suffer from such an imprecise search question. Another option for the user would be to use the classification system in an online bookshop and look under the category "thriller". However, this will also retrieve previously read books.

Offering users relevant new books they have not yet read and would prefer to read, is a typical task for a recommender system. The scenarios described above are similar because they share the goal of helping users get relevant information. The difference lies in the respective approaches of retrieval and recommender

systems and the resulting consequences. That means the information a user gets from a recommender system is not general relevant content (as is the case for most search engines), but must be adjusted to an individual user's needs and expectations. To offer adjusted information to a single user, the recommender system tries to predict what this user will like. Therefore, in contrast to a retrieval system, a recommender system needs to know a priori information about its users. In return, users will be offered relevant resources without explicitly searching for them. A recommender system uses no text or terms from user queries to find relevant items, but historical data showing users' preferences. This data can include any kind of user activity online, such as information about purchased products, clicked websites, watched videos, or liked or bookmarked items. One advantage here is that the user does not need to be active, for instance by formulating a query for getting relevant resources, because recommender systems use data already available from the user's past activities.

Retrieval and recommender systems both have their benefits. Retrieval systems can respond to current user needs more appropriately because a user formulates a query corresponding to a particular information need, sends it to the system, and is given relevant results in return. Simply put, one of the bottom lines for a retrieval system is to get a good query from a user to be able to offer relevant results. In practice, of course, the retrieval process is more complicated and it is difficult for a user to find relevant resources based on his or her information need. A user is not always aware of their precise information need, and even if so, he or she may not be able to formulate a good query that corresponds to it (Belkin, 1980; Kuhlthau, 2004; Taylor, 1968). Research on information retrieval discusses such aspects and problems (see for example Cole (2012) and Stock and Stock (2013) for an overview on these aspects). Another distinctive aspect for retrieval systems is that a user must actively be becoming aware of his or her information need, formulate a question, and type in a query – an important difference to recommender systems that will be discussed in the section on recommender system acceptance and usage.

A recommender system draws on historical data from users. The critical point here is that there must be enough historical data available about a user. This can be any data about a user's activities on the web or in information services. Only if this data is available and, more importantly, relevant, can the recommender system make good recommendations. If this is the case, the user will automatically be given recommendations without having to do anything more. This issue is especially appreciated in e-commerce, as it helps shop providers sell their items and products. The idea is that if a user is recommended new and interesting products, he or she might buy them without having actively searched for them (Ricci et al., 2011). This boosts sales volumes. Besides e-commerce, recommender systems have also become established on the non-commercial web. The idea is to help users filter out information where there is too much of it. Additionally, there is the issue of the "lazy user", meaning that many users are lazy in terms of searching appropriately and using the right query terms. Furthermore, many users do not know how to search well. A recommender system supports them while recommending information based on historical user and item data.

3.1 The Task of Recommender Systems

Recommender systems filter information in order to counteract information overload and help users find relevant and new items (or resources). In general, systems focus on personalized predictions about a user's preferences. A generally accepted division between relevant and irrelevant items plays a minor role, as the definition of relevance and irrelevance is highly subjective. As one user may like a movie that another user hates, it is impossible to trace a clear line between relevant and irrelevant items that satisfies all users. A recommender system that does not focus on personalization may not fulfill the tasks named above, but is still able to make general recommendations. In this work, the focus lies on a target user, who is given recommendations that are relevant specifically for him (or her) and thus derives a benefit from using the recommender service.

A recommender system can help to satisfy diverse user needs. Depending on the system's environment and the user community active in it, these needs may differ, change over time, and not always be relevant for every single user. Nevertheless, the users may accept a recommender system supporting these needs, which guarantees the system's success. In general, recommender systems' tasks can be divided into two groups: Systems that try to find out if a user likes a specific resource (the prediction problem), and systems that try to provide a list of items which a user might possibly like (top-N recommendation problem) (Deshpande & Karypis, 2004; Jannach et al., 2011). The first takes a specific resource, such as a movie, and tries to guess the likeliness rating a user would give assign to it. If the rating is positive and indicates that the user will like the movie, a recommendation is given to the user. The top-N recommender system tries to recommend movies from the database that the user does not yet know but will probably like. Depending on the task, the system is based on diverse methods and algorithms.

Both tasks named above can be split into more concrete sub-tasks, depending on the recommender system's context and application. Herlocker, Konstan, Terveen, and Riedl (2004) describe several tasks a recommender system can support (table 3.1; see also Ricci et al., 2011). It is useful that a recommender system concentrates on specific tasks and that these tasks are based on the needs of target users who must be provided with an additional benefit from using the system.

1. Find good items	Showing the user a list of unknown resources that might be relevant (top-N recommendation); predicted ratings a user would likely assign to an item may also be showed.
2. Find all good items	Similar to task 1, with the difference that it is essential to find all relevant items for a user. This may be crucial for some recommended resources. Ricci et al. (2011) name the example of recommendations concerning medical or financial information. Here the ranking of the items in the list is also more important than in task 1.

3. Improve the profile	This task is essential for providing personalized recommendations. A user should be able to update his or her profile because the recommender system builds its recommendations on this specific user information. Different scenarios are possible: A user may actively update their profile, for example via resource ratings, or the system works with user information that is available on the web, for example.
4. Annotation in context	A recommender system is also applied in existing retrieval systems or already structured databases annotating resources that are predicted to be of interest to a target user. The early systems Tapestry (Goldberg et al., 1992) and GroupLens (Resnick et al., 1994) concentrated on this task (Herlocker et al., 2004).
5. Recommend a sequence or a bundle	Concentrating on the recommendation of a bundle of resources, not only on single resources. Herlocker et al. (2004) name the recommendation of a music playlist. Or a bundle of research papers is also recommended in order to gain an introduction to a specific topic (Ricci et al., 2011).
6. Just browsing	Focusing on target users who have no concrete information need. Here the recommender system tries to lead the user to possibly interesting resources. Brusilovsky (1996) tried to solve this task in an adaptive hypermedia system.
7. Find credible recommender	This might not be a major task, but it is important for establishing user trust: Some users challenge the system to test if the recommendations are relevant. Functions that allow the testing might be important to some target users, especially for recommender systems with critical resource recommendation, such as medical or financial information resources.
8. Express self	A task that fulfills the needs of users who are not interested in recommendations but in giving feedback to resources and expressing their opinion. One example is the feedback function on Amazon. These users may not be the main target users, but, as Herlocker et al. (2004) stress,

	with their input they provide data that improves the system's recommendations, for example in collaborative filtering systems.
9. Help others or influence others	Similarly to task 8, users provide feedback to help other users in their decision-making while not requiring any recommendations themselves. Again, such users bring new data into the system and should be encouraged to give feedback, such as resource ratings. In contrast to users who want to help others, those who try to influence others often end up harming the system. A recommender system's task should be to eliminate those malicious users or to build system defenses to prevent users from acting in this way.

Table 3.1. Tasks of recommender systems, adapted from Herlocker et al. (2004) and Ricci et al. (2011).

3.2 Types of Recommender Systems

Three main types of recommender systems can be distinguished – namely, content-based approaches, knowledge-based approaches and collaborative filtering approaches (Jannach et al., 2011). However, there are various differences between recommender systems, and some researchers do not distinguish between content- and knowledge-based approaches (Jannach et al., 2011). Ricci et al. (2011) also name demographic and community-based approaches. In addition to these, there also exist several hybrid systems that combine diverse methods. Collaborative filtering approaches are the most important among them in the context of this work. The following section introduces the three main approaches, with a focus on collaborative filtering. Other types are summarized in “hybrid system approaches”.

3.2.1 Content-Based Approaches

Content-based approaches (Lops, Gemmis, & Semeraro, 2011) attempt to solve the task of recommending items that are similar to those the user has liked in the past (figure 3.1). Content-based predictions rely on two facts: firstly,

information about some of the items' characteristics should be available, and secondly, additional information should be gathered about the user's likes and interests, for example through a user profile or list of items purchased by the user (Jannach et al., 2011). Thus, a classical content-based recommender system knows the target user and the items he or she has rated. It takes the best-rated items by a target user (all items rated positively and liked by the user in the past) and then recommends similar items the target user does not know yet. Similarity can have different meanings, as discussed in chapter 2. It also depends on the data used for similarity measurement. In a content-based approach, this data is content-based, as the name of the approach indicates. Content, in this case, also includes metadata about items, meaning that information about some of the characteristics or content of the items to be recommended must be known. Characteristics about items are either described in an item's metadata or extracted via text retrieval methods. Additionally, metadata is resource-

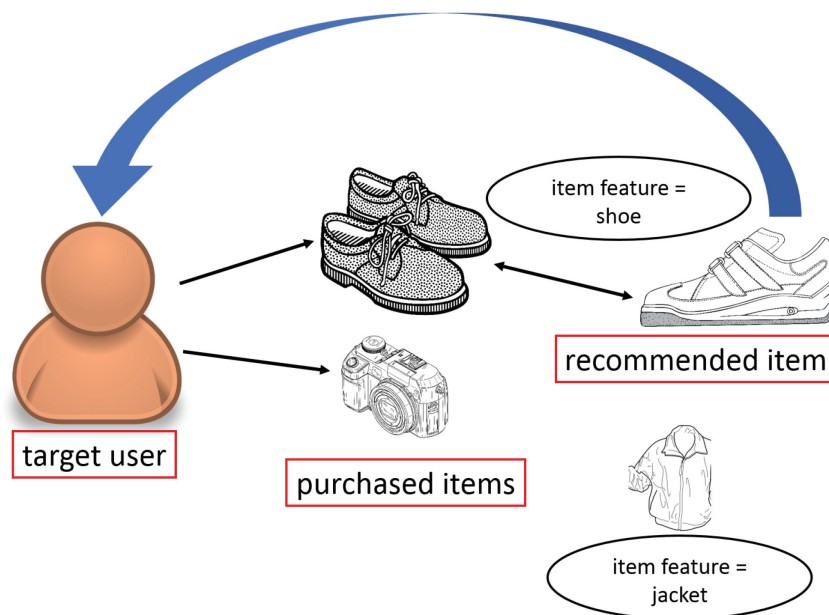


Figure 3.1. The principle of content-based filtering, where a target user gets recommendations of items that have similar content features.

dependent, meaning that metadata about a scientific article is different to metadata about any other product. Taking the example of recommending scientific papers, a target user could have saved several papers in his or her profile. The system could take keyword metadata and recommend various other papers to which the same keywords have been assigned. Hence it is important that metadata is not only available, but also complete. Incomplete metadata excludes resources from recommendation, leading resources with incomplete metadata, or none at all, to be excluded from recommendations even though they might be relevant. Thus the quality and quantity of metadata is of critical importance (Picault, Ribière, Bonnefoy, & Mercerm, 2011). If no metadata is available, text retrieval methods can be an option. Examples of text retrieval methods will be described in the following.

3.2.1.1 Examples of Content-Based Text Retrieval Methods

The vector space model (Salton, 1971; Salton, Wong, & Yang, 1975), combined with text statistics like TF*IDF (see Stock and Stock (2013) for an overview), is often used for similarity measurement in recommender systems. In a content-based recommender system, the vector space model compares terms in resources such as scientific papers with terms in a target user's profile. The terms from the profile might be terms derived from papers the user has liked in the past (indicated, for example, via a rating), or a list of keywords the user has generated as his or her interest list. The TF*IDF (term frequency*inverse document frequency) approach (there are several variants of this weight: Baeza-Yates & Ribeiro-Neto, 2011) weights the resource terms according to their appearance. Such a weighting algorithm is important because not all terms have the same significance. The idea is that significant terms have a higher relevance for resource description and thus should be assigned a higher weight. Words in a text follow Zipf's law (Zipf, 1949), which means that the distribution of the words is skewed towards a select few words that appear very often and many words that appear just a few times. According to Luhn (1958), terms with a high frequency tend to have less meaning (for example stop-words like "the", "and", etc.), while terms with a lower frequency (but not too low) carry relevant

meaning and have enough discriminative power to determine a text's content. Additionally, it can be said that terms that appear less often in a database are also discriminating (Robertson, 2004), thus inverse term frequency (IDF) weighting should be applied (Sparck Jones, 1973). The TF*IDF value (or some of its variants) gives credit to these research findings (see Stock and Stock (2013) for a further discussion). As text retrieval methods applied in content-based recommender systems are based on terms found in resources, TF*IDF is applied to give relevant terms a higher significance. An item is represented as a vector of all its TF*IDF weights for all terms. Some approaches experiment with models that only include "informative" terms in a resource's vector representation (Jannach et al., 2011). These most "informative" terms can be filtered by simply cutting off the most n "informative" words (Pazzani & Billsus, 2007). The model by Billsus and Pazzani (1999), for instance, includes a learning-based approach in which the system learns to recognize the most informative terms. Excluding stop-words and applying stemming approaches are also helpful towards improving TF*IDF measurements.

When an item's content in the vector space model is measured via TF*IDF, the similarity between the content of these resources needs to be measured. The cosine coefficient is a common measurement. If a content-based system has assessed the content of the resources and their similarity, it will need to determine whether a target user might like these resources. This decision is based on the essential information available about a user, that is to say his or her historical profile – as mentioned before, this can be resources purchased by the user, or, ideally, user ratings showing whether the user likes a certain resource or not. When the system wants to know whether the user will like a resource not featured in his or her profile, it compares its similarity to resources from the user's profile that he or she has liked. Here again, the meaning of "to like" depends on the system's definition. The choice of the resources to be compared with the resource whose relevance is to be predicted depends on the understanding of a user's ratings, which are case-dependent and described in the section concerning implicit and explicit ratings.

3.2.1.2 Limitations of Content-Based Approaches

Researchers see limitations of content-based recommender system in several aspects, the most obvious being that content-based measurements can only be performed if a text is available that characterizes an item. Content-based approaches are hard to apply in domains that include multimedia items such as images, videos, movies, music, or products (in e-commerce, for instance). Such items need to be described via metadata, that is to say by using specific keywords characterizing them. Only then does text analysis become possible. The quality of this data strongly influences recommendation. Text sequences or keywords for item descriptions are rather short, which might lead to inappropriate similarity results. Furthermore, generating additional content such as item descriptions is time-consuming, complex and expensive (Jannach et al., 2011, pp. 75–76). One way to overcome this problem is to use tags that are generated by Web 2.0 users to describe items, in which the users act as content providers (Jannach et al., 2011, p. 76). For example, a music recommender system might ask its users to complete information about an album. One way to motivate users to help complete a system's data might be to offer free songs to download. If a system does not want to have to rely on its users, the automatic extraction of metadata is another option. For example, Li, Ogihara, and Li (2003) tried to detect music genres automatically, and Shen, Cui, Shepherd, and Tan (2006) attempted to identify singers.

Outside of this essential aspect, pure text or term analysis may not be enough to capture resource quality. Balabanović and Shoham (1997, p. 67) state that content-based approaches are only able to perform a “very shallow analysis of certain kinds of content”. Concerning a webpage for example, aspects such as aesthetics or usability may also be important in allowing a user to characterize the quality of a web page (Balabanović & Shoham, 1997). This means that applying content-based analysis in order to represent a web page's content might not be sufficient to recommend qualitatively good web pages to a user. Furthermore, purely text-based methods do not differ between well-written and badly written texts. A well-written article and a poorly written article whose

respective authors both use the same words might be recommended in the same way if only words are considered for recommendation (Shardanand & Maes, 1995). Thus, no quality distinction can be made on the basis of the words used in the articles. Furthermore, pure content-based recommendations are based on similarities in the content of different items (Jannach et al., 2011, p. 75). If item content (more precisely: the texts) is not enough, as might be the case with jokes or poems, no appropriate discriminating features can be detected (Pazzani & Billsus, 2007). Thus, distinction between good and poor recommendations becomes difficult.

3.2.2 Knowledge-Based Approaches

Knowledge-based systems (Burke, 2000; Felfernig & Burke, 2008) are often distinguished from collaborative filtering and content-based systems because they can be used when no historic user information is available. More specifically, recommendation in purely knowledge-based systems does not require any a priori user information such as ratings or data about purchased items. Such systems are therefore helpful in environments where no or not enough user information is available. This is the case when a service has many one-time buyers, as, for example, an online shop that offers cameras or computers (Jannach et al., 2011, p. 81). A general user buys one camera or one computer, including the proper equipment, but does not need two similar products. In other words, a user profile in such online shops offers less information about a user's preferences and likes, meaning that recommendations might be inappropriate in such a case. Another example is when a user buys specific products infrequently (Felfernig & Burke, 2008), making it unlikely that he or she will buy a new camera each month. Additionally, user information might be too old and thus lead to false recommendations. Here Jannach et al. (2011) argue that in some cases, user preferences might change enormously. One example is a car recommender. On the one hand, a user's profile will have little data unless he or she has really bought several cars within a short time period. On the other hand, a user might currently need a family van instead of a two-seater cabriolet such as the one bought five years ago. Thus with some

products, users have the tendency to develop diverse needs over time. Content-based recommender systems that rely on the comparison of product features such as “car type”, might be less successful. Recommendations based on collaborative filtering might also be less fruitful due to the sparse user profile lacking in information (Jannach et al., 2011, pp. 81–82).

To summarize the principle, knowledge-based systems require additional user information not based on content or community filtering models. Without further information, the recommender system would be a type of retrieval system that offers the same results to all users who request common features. To get further information, such a system involves a high degree of user-system interaction. A user needs to give information to the system during his or her search process. The system itself is concerned with the task of recommending products that meet a user’s requirements (Burke, 2000) and tries to guide the user through the search process (Mandl, Felfernig, Teppan, & Schubert, 2011). One example is the car recommender service “myproductadvisor”². Here a user can choose between different car characteristics and assign preferences to them (figure 3.2). For example, he or she can state that the car type is important, while the brand is rather unimportant. The system will consequently lend more weight to the car type and adapt its recommendations. Given user information – that is, user requirements or preferences – a knowledge-based system either applies explicit recommendation rules based on product features or focuses on the measurement of similarities between user requirements and resources (Jannach et al., 2011, p. 82). Thus we can distinguish between two types of knowledge-based systems, one constraint-based and the other case-based, which concentrate on diverse tasks, respectively (Felfernig & Burke, 2008; Jannach et al., 2011, p. 82).

Constraint-based systems explicitly ask a user about his or her preferences and work with a tuple of variables or features. Here a distinction is made between

² <http://www.myproductadvisor.com/>

customer features and product features. Customer features can be selected by a user, for instance by selecting a maximum prize for a product. Product features describe a product, such as the car type or horsepower (Jannach et al., 2011). If a user wants to buy a car, a constraint-based system might ask him or her questions about specific product features: How much horsepower should the car have? How many people will sit in the car regularly? Which price does the user want to pay? Depending on these questions, the system will recommend appropriate cars. The customer and product variables need to be adapted to each other, which is done via compatibility constraints and explicit recommendation rules. Compatibility constraints depend on the recommended products. One example for such a constraint is a camera's resolution and its prize: A camera capable of making large print photos requires a price higher than €200 (Jannach et al., 2011), for instance. These constraint rules, which of course also depend on the database of the recommender system, must be defined externally, either via an expert list or via the dataset itself. In the latter case, filtering methods and

The screenshot displays the 'My Product Advisor.com' website interface. At the top, there is a navigation bar with the logo, 'FAQ, Privacy & Terms', and 'About the Advisor'. Below this is a row of feature categories: Body Type & Size, Brand, Capacity & Appearance, Fuel Hybrid Electric, Safety, Convenience Entertainment, Driving & Performance, Quality & Comfort, Price, What's Important, and Recommendation. The 'What's Important' section is active, showing a questionnaire titled 'What's Important to You?'. The questionnaire consists of five horizontal sliders, each representing a different feature: Price, Body Type & Size, Cargo Space, Fuel Type, and Comfort. Each slider has three markers: 'Less Important', 'Important', and 'More Important'. The sliders are currently positioned as follows: Price is set to 'More Important', Body Type & Size is set to 'Important', Cargo Space is set to 'Less Important', Fuel Type is set to 'Less Important', and Comfort is set to 'Important'. Each slider has a yellow checkmark at the end. A 'Continue' button is located at the bottom right of the questionnaire.

Figure 3.2. Car recommender system “myproductadvisor”, in which users are able to determine preferred characteristics as a basis for recommendations. Retrieved 06/21/2016 from <http://www.myproductadvisor.com>

thresholds are applied. If a user wants to buy a camera that enables large prints, the system will recommend cameras with a high resolution or with a price higher than a specific threshold. A formal description of the constraint-based method is shown in Jannach et al. (2011).

The advantage of constraint-based systems— besides the fact that no historic user data is needed — is that these systems can guide users that are not experts through the process of buying a product. If a user wants to buy a car for his or her family, he or she will state that the desired car must be large enough for four people. If it is also stated that the price should be below €10,000, the system can suggest that the user increase the price in order to get more recommendations, because it knows that most family vans are more expensive than that. These hints given by the system are helpful, especially when a user wants to buy complex products and does not know all the technical requirements to be considered (Felfernig & Burke, 2008).

In case-based systems, users do not have to specify their needs beforehand, but choose between specific requirements or feature options during their search process. Case-based systems concentrate on offering the best choices to a user on the basis of the situation the user is currently in (Burke, 2000). Such approaches try to adapt to user cases, which will differ relative to the purpose of the recommender system in question. The starting point is a user search, and the main task is to find resources similar to those the user searches for. Burke et al. (Burke, 2000; Burke, Hammond, & Yound, 1997) developed the restaurant recommender *Entree*, which recommends restaurants in Chicago that are similar to a given restaurant in Chicago or another US city. A Chicago visitor may enter a familiar restaurant and the system will recommend restaurants similar to the user's entry, where similarity is case-dependent. In the case of restaurant recommendation, the feature "cuisine" is quite important, meaning that when the user enters the name of an Asian restaurant, the system might recommend Asian restaurants in Chicago. Optionally, instead of entering a restaurant name, a user can set requirements for preferred restaurant types on the basis of categories like "price" and "cuisine". If the user is not satisfied with the

recommendations, a case-based system offers further choices to select. For example, it may ask the user whether the next recommendation should be a cheaper restaurant than the first. The user may click on the “less price” button and get a new recommendation. The system guides the user through the search process and the user may change his or her search requirements depending on the case-based choices offered by the system. This approach is helpful if a user wants to purchase a product and has preferences in mind, but these preferences are not specific - for example, if he or she wants to buy a camera with a high resolution but cannot name an exact number of pixels the camera should have at minimum. In a case-based approach, the user could now assign a high importance to the feature “resolution” or ask for a new recommendation with “higher resolution” without having to state explicit values for this feature (Jannach et al., 2011). Nevertheless, this approach requires the right choice of resource features that users can select, as well as the right order of importance. If a user searches for an Asian restaurant with a modern atmosphere and wishes to get a restaurant recommendation cheaper than the first one, the feature “cheaper” will be lent the highest relevance. The system might then recommend a cheaper restaurant, but this restaurant might not have a modern atmosphere. This means that the choice of the feature “cheaper” allows the user to set new preferences, which the system then adapts, while another preference, the feature “modern atmosphere”, is lost. Burke (2000, p. 186) summarizes this principle by stating that such knowledge-based systems focus on “high-level responses to particular examples, rather than on retrieval based on fine-grained details”.

3.2.2.1 Limitations of Knowledge-Based Systems

The main difference between content-based and collaborative filtering approaches on the one hand and knowledge-based systems on the other is that the latter lead to a higher user-system interaction. The system forces a user to be active during their search process and to make additional decisions. Either he or she enters a specific resource and the system tries to find similar recommendations, or the user sets requirements that the system uses in order to find appropriate recommendations. These procedures might be disturbing for

some users. If they are asked to set requirements at the beginning of the search, or are asked to change the requirements during the process, they might be overstrained (Jannach et al., 2011). For example, if a user wants to buy a camera and is asked to specify the resolution, he or she might not know which resolution fits his or her needs and which types are currently on the market. What a system can do is provide further information to help the user in making such a decision. Another option is to set defaults (Jannach et al., 2011), for example ones that take into account current market situations. If half of the cameras have a resolution of 14 million pixels, this value will be used as the default for the feature “camera resolution”. This requires the system to update its default settings regularly. Constraint rules or recommendation rules as described above may also determine default settings. If a user wants to print his or her picture in a large format, the system may automatically choose a camera resolution of 10 million pixels. The danger here is that a user can be manipulated into buying a specific product (Jannach et al., 2011; Herrmann, Heitmann, & Polak, 2007). For this reason, a knowledge-based system and its rules need to be constructed carefully.

In addition to recommendation rules, a knowledge-based system should pay close attention to the product features that are used to describe the products and to measure similarity. It is important that “similarity metrics must reflect buyers’ understanding of the product space” (Burke, 2002b). In the example of the restaurant recommender Entree, the user has the option to change his or her requirements to get a new recommendation. They can choose new features, such as “nicer”, to tell the system to recommend “a restaurant which is nicer than the first recommended one”. This “nicer” must be understood by the system in order to give a user an appropriate recommendation. In some cases, features and product data may be equal and the system can directly match both. But product data, that is to say product descriptions in a database, may also differ from general features used in the recommender system. A restaurant might be described with “has modern furnishing” or “has romantic atmosphere”. In such cases the system must be able to refer these descriptions to the feature “nice” in order to make appropriate similarity measurements, and thus good

recommendations. Furthermore, a user's concept of "nice" and the system's understanding of this term must also be equal, otherwise the application will not be successful and give false recommendations. If a user thinks a restaurant with modern furnishing is not a nice one, he or she might not be satisfied with this recommendation. The process of feature selection and similarity settings thus requires careful consideration.

3.2.3 Collaborative Filtering Approaches

The term 'collaborative filtering' was coined by Goldberg et al. (1992). Goldberg describes the principle in this way: "Collaborative filtering simply means that people collaborate to help one another perform filtering by recording their reactions to documents they read" (Goldberg et al., 1992, p. 61). The method is thus based on historical information – specifically, by users. The idea is to rely on the opinion of a user community. It can be said that the approach is based on social information filtering (Shardanand & Maes, 1995) and the behavior of a user community, which is different from content-based approaches, where the focus lies on information about resources. Therefore, collaborative filtering is also called the "word-of-mouth" recommendation (Shardanand & Maes, 1995). Systems are also named social recommender systems because recommendations rely on user content and relations between users (Victor, De Cock, & Cornelis, 2011). The principle can be summarized with the statement "more like me", which means that a user wants to be given recommendations based on other users who have similar behaviors and preferences to their own. This aspect is not so different from an everyday-life scenario: If a user needs information about a specific issue, he or she first tries to solve the problem on their own and then searches advice from friends and reliable acquaintances. These behaviors are meant to overcome not only a lack of information, but also information overload. Say, for example, that a user finishes reading a book and would like to start a new one, but does not have any suitable title in mind. Furthermore, he or she is overwhelmed by the sheer amount of possible good novels, and can thus hardly decide which to pick. Hence, he or she may ask friends who read similar novels as they do. This will

result in a limited amount of recommendations of books the user could read, reducing the difficulty of choice.

There are three requirements for a recommender system based on collaborative filtering (see e.g. Jannach et al., 2011, pp. 13–16):

1. Defining a user profile, which includes a user's likes and dislikes. These can be saved as item ratings, but also in binary form. The latter may consider, for example, a user's purchased products.
2. Finding similar pairs, based on either users, items/resources or tags. Methods are user-based, item-based or tag-based, and define similarity via relations (also called "item-to-item correlation" or "user-to-user correlation": Schafer, Konstan, & Riedl, 2001, p. 137; see also Breese, Heckerman, & Kadie, 1998).
3. Recommending users, items or tags. Diverse measurements are applied and preference predictions are given either for users, items or tags.

Collaborative filtering approaches are often applied when very little or no metadata about items is available (Jannach et al., 2011; Koren & Bell, 2011; Schafer, Frankowski, Herlocker, & Sen, 2007). In this case, the system relies on its users. 'Collaborative' means that all users of a system are involved in the recommendation process. The recommendations are shaped by the behavior of all users. The idea of working collaboratively is to filter items to make sure that only good recommendations are made. Thus the community contributes to the filtering process and each single user benefits from it by getting better recommendations (Heck & Peters, 2010). One of the first recommendation systems based on collaborative filtering was GroupLens by Resnick et al. (1994), a recommender system for news articles. Since then, collaborative filtering has become quite popular (Balby Marinho et al., 2011; Ricci et al., 2011). Web 2.0 services contribute to this success, as more and more online services arise that allow users to actively take part in and contribute to the services, such as bookmarking systems.

3.2.3.1 User-Based and Item-Based Approaches

A collaborative filtering system concentrates on user-based (Breese et al., 1998;

Konstan et al., 1997; Resnick et al., 1994; Sarwar, Karypis, Konstan, & Riedl, 2000; Shardanand & Maes, 1995) or item-based approaches (Deshpande & Karypis, 2004; Linden, Smith, & York, 2003; Sarwar, Karypis, Konstan, & Riedl, 2001). In both cases, a target user is meant to be provided with appropriate recommendations. The approaches differ in that it is either the similarity between users that is measured in order to get recommendations (user-based) or the similarity between items (item-based). A user-based system works on the principle that a user might like resources that similar people also liked. It takes all rated items by a target user and finds other users who rated the same items similarly (Klahold, 2009). These similar users are also called ‘nearest neighbors’ (Desrosiers & Karypis, 2011; Herlocker, Konstan, & Riedl, 2002; Jannach et al., 2011). When the nearest neighbors are determined, the system takes their ratings to predict a rating by the target user. For example, if all nearest neighbors liked an item that the target user has not yet rated, this item will be recommended to the target user. If the nearest neighbors disliked an item, on the other hand, it will not be put on the target user’s recommendation list. In the example in figure 3.3, a neighbor shares at least two common products (shoes and camera) with the target user. The neighbors’ products that the target user has not yet purchased (here: the printer) are then recommended. An overview of user-based systems applying diverse measurement algorithms is given by Breese et al. (1998), Herlocker, Konstan, Borchers, and Riedl (1999), and Sarwar et al. (2000).

An item-based system works on the principle that a user might like resources that are similar to those he or she has liked in the past (Deshpande & Karypis, 2004). The system takes all positively rated items of a target user and then searches for similar items (Klahold, 2009). ‘Similar’, in this case, means that items have similar ratings based on all users. If no ratings are available, common purchases will be considered. In the example in figure 3.4, products similar to the target user’s purchased products (shoes or camera) share at least two users who have rated both products similarly or purchased them. Again, the collaborative aspect is taken into consideration. The printer shares two users with the shoes and is thus recommended; additionally, the jacket shares two

users with the camera and is also recommended. Comparing the simple examples in figures 3.3 and 3.4, where all relations between users and products are equal, the target user would get more recommendations with the item-based approach (top-N recommendation).

Besides the general distinction between user- and item-based approaches, collaborative filtering systems are classified according to computing techniques. A running recommender system can either use pre-processed data (model-based) or measure all similarities during an ongoing recommendation process (memory-based) (Breese et al., 1998; Jannach et al., 2011). While user-based approaches are said to be memory-based, item-based approaches are often model-based, which means that pre-computed processes are applied and the system uses a model to give recommendations (Jannach et al., 2011). Pre-processing data has the advantage that the recommendation process is faster, which is an important factor for customers' decision to use the system. Due to this time advantage, therefore, model-based recommenders are quite popular.

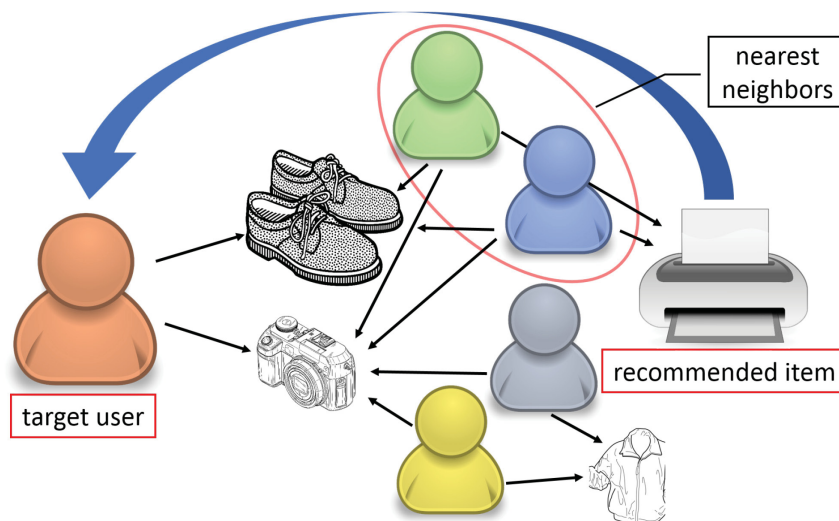


Figure 3.3. The principle of user-based collaborative filtering, where a target user gets recommendations for unknown items from nearest neighbors who have purchased the same products.

However, pre-processing data might be not as precise as data might no longer be up-to-date, which influences recommendation. Furthermore, it takes longer to apply models to a system (Deshpande & Karypis, 2004). In contrast, memory-based approaches might be more precise, but scalability is more difficult and thus leads to computational complexity (Deshpande & Karypis, 2004; Jannach et al., 2011). Services with high numbers of users and items face these problems. A user-based approach in this case might not lead to good recommendations, especially if a database has millions of users, who exceed the number of items, and neighbor networks are huge. It is said that user profiles change permanently, and this influences similarity between users. By contrast, item-to-item similarities are said to be more stable, meaning that pre-computation is more appropriate in item-based approaches. If pre-computation is applied in user-based approaches, the system will not be able to react immediately to changes in user similarity. Pre-computed item-based approaches are thus more precise (Deshpande & Karypis, 2004; Jannach et al., 2011), and they generally need less memory space if the number of items in a database is lower than the number of users. However, considering the time needed to

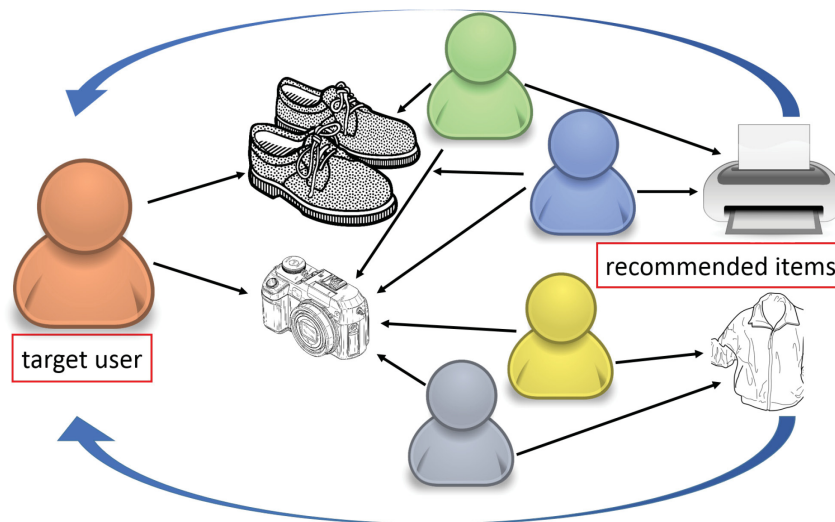


Figure 3.4. The principle of item-based collaborative filtering, where a target user gets recommendations for unknown items that are similar to those he or she already knows.

compute recommendations in a live process might be the same for both systems (Desrosiers & Karypis, 2011). Concerning user-based systems, clustering may be one option to overcome scalability problems. Users are first clustered into smaller groups before their nearest neighbors are measured (Deshpande & Karypis, 2004; Mobasher, Dai, Luo, & Nakagawa, 2002; Ungar & Foster, 1998). The danger here is that a user may get fewer personal recommendations, or none at all (Sarwar et al., 2001).

To overcome shortages within recommender systems, diverse similarity measurements are tested in addition to the classical user- and item-based approaches, such as, for example, graph-based collaborative filtering models (Aggarwal, Wolf, Wu, & Yu, 1999; Huang, Chen, & Zeng, 2004). An advantage here is that graphs, compared to classic user- and item-based approaches, do take into consideration transitive relations between users. The principle is that a user will also be given recommendations on the part of a friend's friend. Suppose that user a and user b liked the same restaurant, and user b and user c share a common preference for another restaurant. User c would not be considered as the nearest neighbor of user a in a pure user-based approach. However, as both have a common nearest neighbor, that is user b , a graph-based system regarding this indirect relation would recognize the connection between user a and user c . Thus, graph-based approaches may handle sparsity – the problem that a database has too few user-item relations to make appropriate recommendations – in a better way (Huang et al., 2004). Probabilistic approaches like Bayesian networks (Breese et al., 1998; Li, Li, Wen, & Liao, 2012) and dependency networks (Heckerman, Chickering, Meek, Rounthwaite, & Kadie, 2001), as well as association rules models (Agrawal, Imieliński, & Swami, 1993; Agrawal & Srikant, 1994; Mobasher, Dai, Luo, & Nakagawa, 2001; Srikant & Agrawal, 1995; Srikant & Agrawal, 1997), are other options to give recommendations based on user and item relationships. For an overview of these approaches, see for example Amatriain, Jaimes, Oliver, and Pujol (2011) and Jannach et al. (2011).

3.2.3.2 Limitations of Collaborative Filtering Systems

Collaborative filtering systems recommend items to a user based on other user ratings or on tags in social tagging systems. The assumption is that what other people liked will also be liked by a target user. The requirement for receiving recommendations is that the target user needs to be active. He or she must have a profile containing saved historic data, such as purchased or rated items. If a target user uses a recommender service for the first time, he or she will face the cold start problem. That means the system has to recommend possible liked items but does not know which items the target user liked in the past. Due to this data sparsity, neither nearest neighbors nor similar items can be analyzed to recommend appropriate items (Massa & Avesani, 2009; Victor et al., 2011). Cold start is a major problem in operating recommender systems:

“Considering that the average number of purchases per user in a single Internet shopping mall, even over a long period, is usually very limited and that there are always significant portion of new users or less-active users in every Internet store, the new user coldstart problem is a very serious issue for most real-world e-retailers.” (Ahn, 2008, p. 40).

A solution to cold start problems is to use more data relations than classical user- and item-based methods. For example, instead of only using user-item relations or co-occurrences in a folksonomy, item-tag relations can also be considered. Said, Wetzker, Umbrath, and Hennig (2009) apply this method and use probabilistic latent semantic indexing to recommend appropriate items. Additionally, hybrid systems exist that further apply content-based methods to overcome sparse collaborative user data (Salter & Antonopoulos, 2006). Another option to overcome sparsity is the aforementioned graph-based model by Huang et al. (2004).

A more direct way of getting more information about a user who has no profile containing historic data is to let him or her rate or bookmark items. For

CiteULike MyCiteULike

My Top Recommendations

This is a list of articles we think you might be interested in.

If you see something you like, press the '+' button to copy it into your library. If you're not interested, then 'X' will delete it and it will never show up again.

This list is only updated once every few hours so you might have to wait a while before getting a new set of recommendation. Also, you'll only see anything here once you have at least 20 [articles](#) in your library.

[\[Settings\]](#) (click to expand)

Recommendations algorithm

We have implemented several algorithms for recommendations. We suggest you try each of these settings and see which works best for you. It may take up to 24 hours for your settings to be applied and new recommendations generated. Note: changing this setting will delete the current list.

- Item-Item Co-Occurrence (IICO)** We look to see how often each article is paired with another in all other users' libraries and choose those which occur most often.
- User-Based Collaborative Filter (UBCF)** We find other libraries that are most similar to your and select articles from those other libraries.

Update

Figure 3.5. A personal recommendation page on CiteULike. A user needs 20 bookmarks as a basis for any recommendations. He or she is able to select item- or user-based recommendations. <http://www.citeulike.org/profile/tamarah/recommendations>

example, services like CiteULike inform their users that they will get recommendations when they have at least bookmarked 20 resources. Otherwise, no appropriate recommendations can be made (figure 3.5). After having bookmarked 20 resources, a user can decide which recommender algorithm he or she prefers – item- or user-based. If there is no historic data available and a collaborative filtering system wants to make recommendations without applying any of the models mentioned above, it could recommend items based on user community preferences. However, in this case a target user will not receive personal recommendations based on their interests. The system will then only recommend the most-purchased or best-rated items. As a recommender system is intended to make personal recommendations (Ricci et al., 2011), the latter solution should preferably not be applied.

3.3 Comparison of the Approaches

The application of content-based, knowledge-based and collaborative filtering approaches mainly depends on the system's tasks and the data (that is items, item metadata, user profiles, and user feedback) that is available or is meant to

be generated. However, all approaches come with advantages and disadvantages that must be considered before implementing the system. Some limitations of each of the three recommender system approaches have been mentioned above. This section compares the approaches and their shortcomings in order to summarize the main differences.

Collaborative filtering techniques are a kind of counterpart to content-based systems. These systems do not require any information about items and thus do not have to rely on appropriate content to be analyzed. However, the advantage of content-based approaches is that such systems do not need any user ratings or large user groups who actively use the system. A content-based recommender system is built on information about items that is attributes describing the items' content, which are then used for similarity measurement. Thus the system is able to recommend new items immediately, provided that item attributes are available. A collaborative filtering approach first needs at least one user who rates or purchases an item before it can recommend this item to any other user. This means that new and likely relevant items are not recommended in a collaborative filtering system. A problem here is data sparsity and cold start. In many databases, individual users only rate a few items, which lowers the system's potential of recommending items (Jannach et al., 2011).

Nevertheless, content-based models need metadata about items to measure item-to-item similarity and recommend similar items to a user. Generating and analyzing such data is expensive and time-intensive, and it is difficult to guarantee its completeness. Despite possible limitations concerning data content, these approaches are unable to recommend "surprising" items. The model's scope of recommendations is limited (Victor et al., 2011). For example, if a user liked novels by their favorite author Henry James, a content-based system would quite possibly recommend other novels by James the target user has not yet rated or purchased. Recommender system research speaks of the serendipity problem (Lops et al., 2011, McNee, Kapoor, & Konstan, 2006; Shardanand & Maes, 1995). A content-based recommender is able to recommend new or novel items a user does not know (novelty degree of a

system). However, it suffers from recommending items that are too similar to those the user already knows (degree of serendipity) (Lops et al., 2011; Victor et al., 2011). However, some users appreciate serendipitous items and want to get surprising recommendations they would not have found on their own. Herlocker et al. (2004, p. 42) speak of “recommendations that fail the obviousness test”. If too obvious items are recommended, a user may not use the recommender system because he or she does not see any additional benefit in it (Herlocker et al., 2004). Victor et al. (2011) name this shortcoming of content-based systems “over-specialization”, referring to the fact that these systems only recommend items that are very similar to those a user already knows. One way of overcoming overspecialization is to set an upper threshold, as described by Billsus and Pazzani (1999). They propose not only to exclude items whose non-similarity is too high from recommendation, but also to exclude too-similar items in order to offer a more appropriate variety in their recommendations (Billsus & Pazzani, 1999). By contrast, collaborative filtering approaches consider the opinions of other users in recommending items to a target user. Thus, these models have the “potential for serendipitous recommendations” (Herlocker et al., 2004, p. 43).

Another shortcoming of content-based approaches is that they do not consider item quality. For example, a user who liked a specific novel by Henry James in the past might not like James’ novels in general. Thus the product characteristics, here for example the item feature “author = Henry James”, may not always be significant in deciding whether a user likes an item or not. User likes might be very subjective and not subject to representation by single item features (Jannach et al., 2011). Collaborative filtering approaches rely on user communities to “measure” an item’s quality. If many users rate an item positively, the system will recommend it to a target user. In pure content-based approaches, user data is not considered. However, user ratings might also be inappropriate for clear quality statements. A target user has to trust the “wisdom of the crowd”.

In addition to an item’s quality, one question for a target user is whether he or

she can really trust recommendations. If a user does not trust a recommender system to offer appropriate suggestions, he or she will not use it. Trust and transparency are very important, even for content-based systems. As collaborative filtering approaches rely on other user ratings, however, trust-based models in recommender system research generally refer to collaborative filtering approaches. Here a target user asks if he or she can trust their nearest neighbors on whom the recommendations are based. Sinha and Swearingen (2002) found out that users want to know why a system recommends items, they want the system to be transparent and to justify its suggestions. If a system is not transparent, it may not be used (Herlocker, Konstan, & Riedl, 2000). One factor on which a user evaluates a system's trustworthiness are items a user already knows and liked before. Swearingen and Sinha (2001) distinguish between good recommendations that a user does not know and good recommendations that a user knows and has already experienced. The latter recommendations are called "trust-generating" (Swearingen & Sinha, 2001). The researchers claim that these recommendations are not useful in a traditional sense because recommender systems should suggest items a user does not know yet. However, such items may generate trust because a user directly sees that the system knows his or her preferences. Additionally, the potential of a recommender system to suggest new and surprising items (serendipity aspect) was also rated positively. Thus research studies assume that a mixture of good known and good new items is preferable for a user (Sinha & Swearingen, 2001, 2002; Swearingen & Sinha, 2001).

Additionally, Sinha and Swearingen (2001) found that users tend to trust their friends over a recommender system. Therefore, trust-based recommenders try to establish a kind of social network within the system, which shows how much a user may trust other users. In general, all systems that show any implicit or explicit relations could claim to establish a kind of trust factor for users (Victor et al., 2011). For example, collaborative filtering relies on similar users, which automatically makes it more trustworthy than systems relying on all users as similar people might lead to better recommendations for a target user. However, we must point out that similarity has different meanings for each user and that a

system's understanding of how to measure similarity might not reflect that of a user. Furthermore, a friend of a friend might also be more trustworthy to a user than any other unknown person. This means that trust can propagate among user networks. Trust propagation models apply steps similar to the models of co-citation and bibliographic coupling (Guha, Kumar, Raghavan, & Tomkins, 2004). On the other hand, trust can also aggregate, which leads to a second type of trust-based system (Victor et al., 2011). In a user network, several paths may connect two indirectly related users. Trust value between these two users is aggregated on the basis of the different paths. Trust models also attempt to solve the cold start problem, for example by referring to more implicit relations between users (Huang et al., 2004). Victor et al. (2011) give an overview of trust-based systems and their possible differentiation. Trust-based systems, which directly ask users about their trust experience, form an interesting counterpart to classic recommender types. In their studies, Massa and Avesani allowed a user to directly evaluate other users regarding their trustworthiness (Massa & Avesani, 2004, 2009). These "trust statements" are considered for recommendation while giving each user a trust weight. In addition to these trust models, research also discusses the aspect of distrust as a related counterpart to trust (Guha et al., 2004; Victor, Cornelis, De Cock, & Da Silva, 2009).

Compared to content-based systems and collaborative filtering approaches, knowledge-based recommenders rely on a higher degree of user-system interaction. A user is asked to state his or her needs more concretely, either before or during the retrieval process. Therefore, a user does not need to trust anybody else, but simply needs to be clear about his or her needs. If users are able to clearly state their requirements, a knowledge-based system will generally be able to give more precise recommendations than those generated by other recommender approaches. However, user interactions must be understandable to a user and represent his or her needs. Getting direct user feedback is helpful when a user's historic profile data is not sufficient to give good recommendations. This is the case when the items to be recommended are ones that users do not buy regularly (Jannach et al., 2011, p. 81). Furthermore, content-based and collaborative filtering models might rely on ratings by users

who do not always show accurate and consistent rating behavior (see the next section). However, user-system interaction requires a user to be active, which is disturbing to some users and might overly tax them, especially if they are no experts concerning the items to be recommended.

Hybrid Approaches

To overcome the shortcomings of the three recommender system approaches, hybrid systems are constructed that try to use the strengths of all three models (Beliakov, Calvo, & James, 2011). Research distinguishes between diverse types of hybridization regarding a system's hybrid principle. Jannach et al. (2011) name three different principle types. A monolithic hybridization approach considers aspects from diverse recommender system types and implements them in one system. Such a system might draw on content-based similarities between items and collaborative filtering user-based similarities between users to recommend appropriate items to a target user. Similarity algorithms are merged in one implementation.

Parallelized hybridization approaches (or mixed approaches (Burke, 2002a)) also consider diverse recommender system types, but the systems' specific algorithmic measurements are applied separately and in parallel to each other. In a hybridization step, further measurements are applied to give final recommendations. This means, for example, that content-based and collaborative filtering user-based similarities are measured separately. Afterwards the system may either present results from both measurements (mixed hybrid), present results from only one measurement depending on a target user's current situation (switching hybrid), or combine both result lists using a weighting approach (weighted hybrid) (Burke, 2002a). Weighting the results of different recommenders is quite common. One simple method is to rank recommendations according to their ranking scores (Jannach et al., 2011). For example, let one item have a ranking score of 0.2 in the collaborative filtering approach and a ranking score of 0.6 in the content-based approach. Giving both approaches the same relevance, an equal weighting of 0.5, the

item's scores are summed up and the result divided by two. As a result, the item will have a ranking score of 0.4 in the final recommendation list offered to a target user. Here it should be kept in mind that the ranking scores of both approaches are comparable and scales do not differ significantly. In most other cases, final recommendation rankings do not reflect the rankings of the applied approaches (Jannach et al., 2011).

The third principle of pipelined hybridization applies metrics successively, with the results from the first approach being then used by the second approach. Burke further classifies this principle in three specific methods in order to distinguish them more precisely (Burke, 2002a; Burke, 2007). One method is a cascading hybrid system, in which one approach refines the recommendation results of another approach. An example for this is the Entree restaurant recommender described above (Burke et al., 1997). Burke (2002a) improved the system's performance while refining the simple knowledge-based approach with a collaborative filtering method. The first restaurant suggestions resulting from the knowledge-based model are revised on the basis of former user choices from the system's historic database. However, cascading pipelined hybridization has the disadvantage that if the first approach does not include an item, neither will all subsequent approaches. In a basic pipelined design, successive approaches cannot introduce new items. That means that these designs decrease the recommendation set of items, which may lead to an insufficient number of recommendations (Jannach et al., 2011). A possible solution is to switch the order of the applied approaches and to set an additional threshold, which switches between the approaches when not enough recommendations are given (Zanker & Jessenitschnig, 2009).

In general, it can be said that hybridizations try to overcome the shortcomings of individual recommender system approaches. However, their tasks strongly depend on the items to be recommended and on the data available for recommendation. If a system does not have access to historic user data or item metadata, it can benefit neither from collaborative filtering nor from content-based similarities. Functioning recommender systems that aggregate data from

diverse sources in order to be able to apply several approaches are rare due to the problems associated with aggregation, data incompleteness, and data usage restrictions (Jannach et al., 2011, p. 141). However, when more and more data becomes available, especially in Web 2.0, chances increase to build better (hybrid) systems, which are capable of providing target users with better recommendations as they consider more data sources and can thus retrieve more data relations (Jannach et al., 2011, p. 301).

3.4 Explicit and Implicit User Ratings

Content-based and collaborative filtering recommender systems rely on historic user data in order to give their recommendations. Any data showing a user's interests is necessary here. The system needs statements from a user – that is, any rating about his or her item likes or dislikes – because it aims to predict likes the user has not stated yet. Thus in the following, “user rating” – also called “user-item response” (Desrosiers & Karypis, 2011) – refers to any kind of user statement a system can draw on for predicting items likely to be liked.

User ratings are either explicit or implicit (Jannach et al., 2011). If a user states an explicit rating, he or she consciously reveals an opinion about an item. It is assumed that users know when they have rated an item. Explicit ratings are, for instance, ratings on a rating scale: a user is able to give between one and five stars to an item, where one star expresses a dislike of the item and five stars a strong preference. Here a user is aware of rating an item. He or she states: “I like this item and thus rate it with five stars”. Implicit user ratings of an item are ratings where a user does not consciously show an opinion, which means that in most cases he or she is not aware of making a value judgment with regard to an item. There is not conscious statement such as “I like this item and rate it positively”. Instead, the user shows some behaviors or interactions within the system – for instance, bookmarking an article – with the system interpreting this behavior as an item rating. Schafer et al. (2007, p. 305) describe this as a system using “observations of user behavior from which preference can be inferred”.

Besides the distinction between explicit and implicit ratings, there are three

rating types that influence the options of analyzing user-item responses. A scalar rating includes a scale, which is either numerical or ordinal and which has different appearances depending on the recommender system's design. Typical numerical scalar ratings are scales that show one to five stars or points, or ordinal scales that may show options like "strongly like", "like", "neutral", and "dislike". Binary ratings only offer two options for a user, such as "like" or "dislike", "agree" or disagree". Within these types, a user explicitly expresses an opinion. By contrast, unary rating types are either implicit or explicit. In explicit unary rating systems, a user can only state whether he or she likes an item. There is no option, as there is in binary systems, of disliking an item (Schafer et al., 2007). Implicit unary ratings draw on any user-item response to identify a user's preferences (Desrosiers & Karypis, 2011). For example, if a user purchased an item, the system sees this activity as a positive user-item response, which means that the user "likes" the item. The same applies to a user who bookmarks an article in a bookmarking system, or a user who watches a movie or listens to a song on a music website. Recommender systems using unary ratings, regardless of whether they are implicit or explicit, can only analyze whether a user-item response exists or not. Scalar and binary ratings offer more options and are able to distinguish between user preferences with a finer granularity.

To distinguish user preferences in unary implicit systems more precisely, other data is analyzed. For example, a system can measure the time a user spends on a product website, or listening to a song, to make fine distinctions between higher and lower preferences (Oard & Kim, 1998; Schafer et al., 2007). Nichols (1997) gives an overview of different implicit rating types, such as purchasing a product, marking or saving a document, referring to a document (which includes citing or mentioning it in any way), or recommending an item to a friend. The more information a system gathers from a user, the more specific the implicit ratings are. However, more information might not always lead to a more accurate interpretation of user behavior. The challenge is to interpret user-item response in the right way.

Implicit ratings do not contain explicit user feedback, but must be interpreted as such. The question is, what do user-item responses mean? If a user purchases a product, does he or she really like it, or prefer it? Maybe it was meant as a gift for some relative, and he or she would never personally use the product. Such scenarios are possible, and there are many examples. Research states that if the number of ratings in a system is high, the high number of cases that are interpreted correctly outnumbers such “false” cases (Jannach et al., 2011; Schafer et al., 2007). However, the danger of misinterpreting user-item responses is real. By contrast, explicit ratings do not suffer from false interpretation as users state their preferences explicitly, which makes them more accurate. Nevertheless, some other aspects should be considered. Explicit rating systems are limited in their representation. Their scales are fixed, and the system only distinguishes user tastes based on these fixed scales. For example, a system distinguishes if a user rates a song with one star or five stars. An implicit rating system may count how often a user listens to a song and use this number to distinguish between diverse song preferences even more accurately, because the counting of played songs is not limited to a five-point scale (Schafer et al., 2007). Assuming that the system correctly interprets implicit ratings – in this example, listening to a song means that a user likes this song – the general assumption that explicit ratings are more accurate is not verified. Another question is whether a user him- or herself interprets explicit ratings correctly. Users might have diverse views on items and have different aspects in mind when they state that they like or dislike an item. For example, a user might like a movie because he or she likes the main actor. Another user might like the same movie for its plot. The latter might not like movies starring the same actor, rather disliking them. However, a content-based recommender system might suggest movies with the same actor. It is not clear, therefore, what exactly a user rates when liking an item. It also might be unclear to the user what to rate. In the online shop Amazon, many users rate delivery service. If delivery takes too long, they will rate the purchased product more negatively, although in fact they do not refer to the product itself. Furthermore, users might not be consistent in their ratings, meaning that they rate the same item differently at different points

in time, especially when rating scales offer a fine granularity (Schafer et al., 2007). In that case, user ratings are not accurate anymore because users make diverse statements with regard to their preferences. Thus, explicit ratings might not be any more accurate than implicit ratings. Advantages and shortcomings of both types of ratings strongly depend on the user community, the purpose of the recommender system, and the items which are recommended. More research is needed concerning users' intention to rate, their rating behavior, their own interpretation of ratings, and the influence of the representation of rating systems (Gena, Brogi, Cena, & Venero, 2011; Zhao, Qian, & Xie, 2016).

Concerning user ratings, we must note the aspect of item quality (Adomavicius & Tuzhilin, 2005). As mentioned above, a user rates an item on the basis of his or her individual feelings at a certain point in time. He or she might rate the quality of an item, but the ratings might also refer to other facets. Thus a user's ratings should be seen as preferences, likes, or tastes regarding an item, and not as a statement about its quality. The concept of 'quality' varies from person to person. In other words, an item's quality is not fixed simply because many users have purchased it or gave it a good rating. These factors might be an indicator of some kind of quality, but what users actually rate is their personal preference in favor of, or against, an item. Adomavicius and Tuzhilin (2005) formally describe this recommendation problem: The main goal is to predict the utility of an item in a specific context, because an item's utility is not universal; rather, it is highly related to a specific context, or the target user's current environment. Recommender systems should take care to acknowledge this aspect. Furthermore, this issue becomes important when it comes to explaining a system. Herlocker et al. (2000) emphasize the importance of a system's explanation to its users. If a user knows how a system works and why it recommends what it recommends, the user's acceptance of the system will increase. Thus, users should know what ratings mean in the system's environment and how they should be interpreted. For example, if a user thinks recommendations are content-based and expects to receive unknown movies starring their favorite actor, but instead is recommended diverse movies liked by his or her similar collaborative neighbors, the user will be disappointed and lose

trust in the system. In a people recommender system, the effect might be more intense because here, personal feelings are involved. If a user gets a recommendation of a potential collaboration partner, but does not like them, he or she will feel personally affected because the system states that the two will be a good match without explaining its decision. If the user understands on which similarity relation this recommendation is based, he or she will not judge the system for giving false recommendations but may only think that the established method the system uses might not be appropriate for his or her individual needs. To make ratings more understandable, a system can use multiple explicit ratings, such as a rating scale for services and a separate rating scale for products. However, more rating options mean more time-consuming work for the users, which might restrain many from leaving feedback. In implicit rating systems, explanation is even more important. The user needs to understand on which basis a system measures similarity and gives recommendations. This creates user trust in a system. On Amazon, users know they get recommendations because “other users who bought this product also bought the following products”. Herlocker et al. (2000) tested several explanation interfaces in their study, with each giving additional explanatory information about the recommended movies. Tested interfaces showed, for example, the settings of a user’s neighbors ratings, which explained that the recommended movie contained a user’s favorite actor, or said with what degree of confidence the system made a certain recommendation. Results showed that explanations are valuable for users, simpler explanation interfaces were more useful than more complex ones, and explanatory information – a special case in the example of this movie recommender was the aspect of the same actor – is dependent on the user community (Adomavicius & Tuzhilin, 2005). The latter outcome showed significant differences in users’ perceptions, with the authors concluding that specific content features are important for system explanation. For some users, the cast was an important criterion for evaluating movies, while others seemed not to care about this.

3.4.1 Tag-Based Approaches

In addition to explicit and implicit ratings, recommender systems use assigned user tags for similarity measurement and recommendation. Tags are assigned to items by users and have different functions (see chapter 2). As a user becomes active in explicitly assigning tags to items, tags can be seen as explicit user activity, and thus in some cases as explicit user ratings. For example, a user might tag a resource with “interesting” or “very good”, which expresses a kind of preference, and thus an explicit rating. The advantage of tags is that they show user interests more concretely than “pure” ratings because a tag term expresses the semantic concept a user has in mind when referring to a bookmarked item. The user is not bound to a limited rating scale. Additionally, tags are not subject to any restrictions and a user may choose precisely those tag terms he or she thinks represent the bookmarked resource in the best way. However, this aspect is also a shortcoming for recommender systems that apply similar tags for recommendation.

Another difference between tags and explicit scale ratings is that tag terms, when assigned appropriately, describe the content of their resource. Through these tags, topical similarity between resources is detected and used for similarity measurement. The crux is that tag terms showing an explicit rating by a user are, in most cases, not usable for these measurements because they do not show topical similarity. Hence, it is more useful for tag-based recommender systems to work with descriptions of tags rather than ratings. For example, a user may tag a bookmark of a scientific article as “interesting”. This information is useful for the current user, but not for others because they do not know why the former thinks this article is interesting. The user’s intention in assigning “interesting” is not clear. On the other hand, when a user tags a bookmarked article with “information retrieval”, this adds further useful metadata to the article, which can then be used in a recommender system to build relations between bookmarks with topical similarity. Another user, also interested in “information retrieval”, may get recommendations of all articles that are not in his or her bookmarking list and have the tag “information retrieval”.

Tags describing the content of bookmarked resources are part of folksonomies (Peters, 2009) in bookmarking systems and are used in social tagging recommender systems (Balby Marinho et al., 2011; Balby Marinho et al., 2012), also called tag-based recommender systems (Duraó & Dolog, 2009). As folksonomies include information about users, resources, tags and their relations, three different kinds of recommendations are possible (figure 3.6). Tag-based systems recommend either resources, tags, users or all three of them. Thus, the basis for recommendations also varies. For example, tags can be similar due to the number of resources, and resources can be similar due to the number of users or tags. Systems recommending resources show users new and interesting items they might like or find helpful. Systems recommending users show users new and potential partners, for example for collaboration or to discuss personal or commercial interests. There can be many different purposes. Systems that recommend tags to users also have diverse aims. They might want

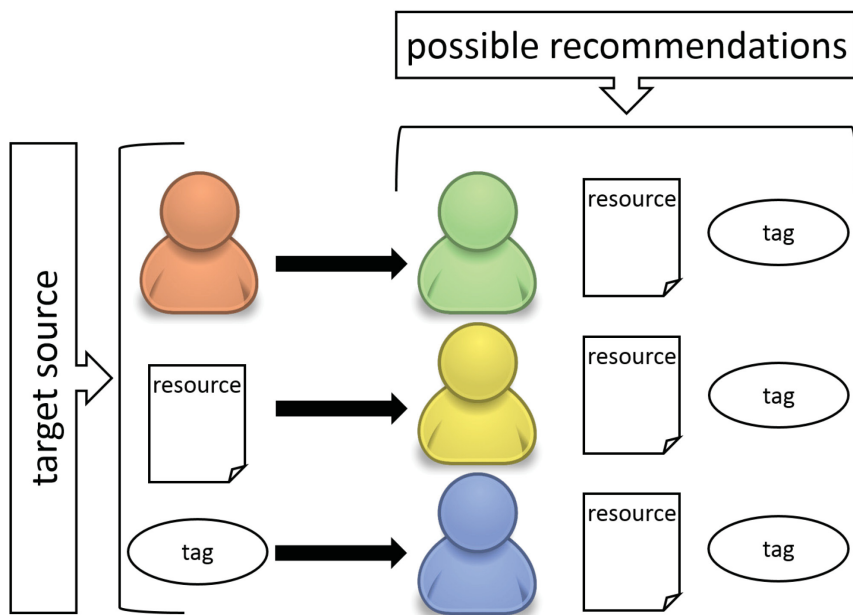


Figure 3.6. Recommendation variants in a tagging system. Figure adapted from Balby Marinho et al. (2012, p. 9).

to help a target user find appropriate tags for his or her resource. Another aspect to be noted here is that a tagging system will increase its chances to get its resources annotated. Additionally, systems want to clean folksonomies and unify tag collections so as to make better recommendations (Jäschke, Marinho, Hotho, Schmidt-Thieme, & Stumme, 2007). If a system recommends tags already used by other users, a target user will likely not assign any tag that has not yet been used in the system. Thus a tag collection includes a small and specific collection of tags instead of a broad term collection in which unique concepts are represented by several different tag terms.

This issue leads to the approach of unifying tag terms by applying specific linguistic and non-linguistic techniques. There are no rules governing tag assignment, and tag collections include linguistically diverse forms such as spelling errors, diverse tag syntactic forms such as nouns or verbs, diverse grammatical categories like singular and plural, diverse spelling conventions such as words with or without hyphen, as well as homonyms and synonyms. Recommender systems can help unify tag collections and recommend common tag forms to users who want to assign tags to their bookmarks (Jannach et al., 2011). Peters and Weller (2008) for example propose “tag gardening” to format, recommend and control tag vocabulary (see also Peters, 2009, pp. 235–247 and Stock & Stock, 2013, pp. 621–623). The idea of their system “TagCare” is to offer users the option to structure their tag collection, and further to use those collections and enhance professional knowledge organization systems. Regarding the recommendation aspect, if a user wants to use the tag “information systems”, the system might propose the singular form “information system”. Uniformed tag collections help a user order his or her bookmark collection and to find all similar bookmarks assigned via common tags. Furthermore, they help a user community by making the retrieval of new resources via tags easier. However, false tag relations and incorrect interpretations of a tag’s meaning will lead to inappropriate tag recommendations, and thus to inappropriate user and item recommendations in a tag-based system. Shepitsen, Gemell, Mobasher, and Burke (2008, pp. 259–260) summarize the problems occurring in tag-based recommender systems:

“Redundant tags can hinder algorithms that depend on identifying similarities between resources. On the other hand, tag ambiguity, in which a single tag has many meanings, can falsely give the impression that resources are similar when they are in fact unrelated”.

Besides these issues, Peters (2009) discusses the aim of tag recommender systems based on the core idea of folksonomies. She distinguishes between three sorts of tag recommendations: On the one hand, a system can make users aware of any spelling mistakes within their tags and suggest a tag variant – for example, the singular instead of plural form. Alternatively, it can suggest tags that users have previously assigned in the past in order to provide them with a consistent personal tag collection in their bookmarking management (Muller, 2007; Sinclair & Cardew-Hall, 2007). Thus a user’s personomy is their own controlled vocabulary (Peters, 2009, p. 204; see also Neal, 2007). By contrast to those tag recommenders, a system may also suggest tags assigned by other users, either the ones most frequently assigned on the basis of the system’s data or those already assigned to a specific resource (Sigurbjörnsson & Van Zwol, 2008). Here, Peters (2009) alludes to the positive feedback loop problem: “There are obvious dangers in establishing a positive feedback loop where potentially unsuitable tags may be reused due to the tag’s initial popularity and subsequent exposure as a tag recommendation” (Guy & Tonkin, 2006). In other words, tags used more frequently are recommended more often, and thence are used even more frequently. “After all, this sort of recommender system artificially generates the implicit user agreement on certain behavior. Hence we can no longer speak of a reflection of authentic user behavior” (Peters, 2009, p. 205). The idea of a folksonomy designed to reflect user behavior gets lost. Furthermore, there arises a danger of not having enough discriminating tags, which influences good recommendations (Kipp, 2006; Muller, 2007; Paolillo & Penumarthy, 2007; see also Peters (2009) for further discussion on this topic).

In addition to these issues, which should be kept in mind, there are several approaches derived from information retrieval research which generate relations among tags and try to reduce errors in tag collections (Stock & Stock, 2013,

pp. 621–628). Relations between tags can be generated either statistically or on the basis of natural language processing. Statistical approaches use tag co-occurrences in documents or in user tag collections to measure relations between tags that are used to make recommendations. In pure statistical approaches, tags are not modified linguistically. Both approaches can be combined, of course; for example, natural language processing is applied prior to statistical measures.

Sanderson and Croft (1999) suggest automatically detecting hierarchical relations between terms by applying “subsumption” rules based on term co-occurrences in documents. For example, term *a* is a hierarchical parent, a superordinate concept, of term *b*, if the documents including term *b* are a subset of the documents including term *a* (Sanderson & Croft, 1999). Thus, a hierarchy relation can be generated. This structure can be used for recommendation, for example by recommending the subordinate terms of a used tag to a target user in order to help him or her describe the resource more precisely. Schmitz (2006) adapted this approach to a Flickr database in order to generate parent-child relationships between tags, considering not only the co-occurrences of two tags in the image descriptions, but also the number of users who assigned both tags. Both values function as thresholds, which were adapted on the basis of experimental results. Even though the results show some noise (personal idiosyncratic tags such as abbreviations or spelling mistakes) and incorrect relations, 51% of the relations are correct (Schmitz, 2006). Most of these relations seem not to be generic or partitive hierarchical relations, but instance relations (individual concepts) (Stock & Stock, 2013, pp. 552–558), like “is-part-of” relations showing geographic belonging: “Golden Gate bridge”, for instance, is a child tag of “San Francisco” (Schmitz, 2006). A similar method is proposed by De Meo, Quattrone, and Ursino (2009), who also measured tag similarity based on the number of common resources. They consider diverse user needs and adapt the similarity techniques according to these needs. For example, if a user is no expert in a specific domain, he or she probably needs more tag recommendations to search or label resources. Therefore, tag similarity is measured more “loosely”. Tags are clustered in a neighborhood set. A more

loosely clustering strategy is the or-like strategy, where a tag is assigned to a cluster if it is close to at least one tag within the cluster. In the and-strategy, a tag is included in a cluster if it is close to all other tags (De Meo et al., 2009).

Another approach is tag clustering, where clusters with similar tags showing topical relations are generated via diverse techniques (Knautz, 2008; Shepitsen et al., 2008). Similar tags can be recommended on the basis of these clusters, but in most cases the clusters serve as a pre-step for further user or resource recommendation. Shepitsen et al. (2008) compare diverse clustering methods of recommending resources to target users. Tags are clustered on the basis of their co-appearance in resources (see also Begelman, Grigory, Keller, & Smadja, 2006). Tag frequency based on a resource as well as on a tag-based form of TF*IDF (number of times a tag was assigned to a resource multiplied by IDF: ratio between the number of all resources and the number of resources a tag was assigned) is used (Shepitsen et al., 2008). Diverse approaches are possible to generate clusters, such as maximal complete link and k-means (see examples in Gemmell, Shepitsen, Mobasher, & Burke, 2008a, 2008b) or different variants of hierarchical agglomerative clustering (Shepitsen et al., 2008). Brandes, Gaertler, and Wagner (2003) give an overview of clustering approaches.

While “subsumption” and clustering approaches try to detect related tags to give better recommendations of any resources, users or tags, natural language processing approaches generate equivalent clusters of tags designed to represent a common concept. Here tag forms are changed in order to give more appropriate recommendations. Stemming (non-linguistic) and lemmatization (linguistic) are two techniques applied to conflate terms (Galvez, De Moya-Anegón, & Solana, 2005). Lemmatization considers morphological principles to generate linguistically correct lemmas. They are formed either on the basis of defined rules or with the help of dictionaries (Stock & Stock, 2013, p. 187). One simple example is the “S-Lemmatizer” for the English language, which conflates singular and plural forms while applying four rules (Harman, 1991):

- words with 1-3 letters are not lemmatized,

- “ies” at a word’s end is replaced with “y”, except in words ending with “eies” and “aies”,
- “es” at a word’s end is replaced with “e”, except in words ending with “aes”, “ees” and “oes”,
- “s” at a word’s end is deleted, except in words ending with “us” and “ss”.

Such rules can be expanded to cover the conflation of more word forms (see for example Kuhlen, 1977). Automated approaches, which learn lemmatization rules, are also applied (Juršič, Mozetič, Erjavec, & Lavrač, 2010; Plisson, Lavrač, Mladenic, & Erjavec, 2008).

The difficulty when it comes to dictionaries lies in the immense workload involved, as a dictionary must include entries for each single term in a language, including all morphosyntactic characteristics and lemmatization forms for semantic analysis. The latter can also be used to detect paradigmatic term relations such as hierarchy or associative relations (Stock & Stock, 2013, p. 188). For the English language, the database WordNet³ is available (Fellbaum, 2005). WordNet mainly takes into account equivalent relations and group synonyms, which have the same conceptual meaning, in so-called “synsets”. Hierarchical relations, or super-subordinate relations, as well as part of speech relations that link words with the same stem, are also included. However, studies evaluating automatic query expansion via WordNet for information systems show negative results (Voorhees, 1998). It was not possible to clearly distinguish between different meanings or query terms. As tags mostly consist of single terms and not of compound terms, word sensing appears difficult.

Tag-based recommender systems use tags to analyze similarities and recommend unknown resources, users or tags to a target user. These systems use collaborative filtering methods as the basis of their recommendations. With regard to topical relationships, which tags are able to show, it can be said that

³ <https://wordnet.princeton.edu/>

tag-based approaches are similar to content-based approaches in that both try to generate an item's content for recommendation (Jannach et al., 2011). However, the method is collaborative filtering. Furthermore, the major difference between tags and metadata used in content-based approaches is that tags are generated by a user community. A tag is always related to a user. Comparable to item-based collaborative filtering approaches, tag-based approaches might not consider all items that have similar user ratings, but items that have been assigned common tags (Duraio & Dolog, 2009; Zhao et al., 2008). This means that all items with common tags are ideally similar and will thus be recommended to a target user if he or she has not yet bookmarked those items. In a scientific bookmarking system, two bookmarked articles with a common tag, for example "communities of practice", might quite possibly deal with a very similar topic. Therefore, a user who is interested in this topic and has already bookmarked one of the articles might also be interested in the second article. A tag-based model would recommend this one.

Tag-based approaches are further divided – like classical collaborative filtering methods – into user- and item-based approaches, with similarity being based in both cases on tags instead of ratings. User-based methods assume that users who use common tags to describe their resources are interested in common topics. Hence, unknown resources from similar users might be of interest to a target user. Item-based approaches assume that all items with common tags are similar. A target user gets recommendations of items he or she has not bookmarked yet and that share common tags with their already bookmarked items.

Additionally, if user ratings are available, tags are applied to predict user ratings for resources. Szomszor et al. (2007) suggest predicting a user's rating for an unknown item by comparing a user's tag clouds for rated items and the tag cloud for the unknown item in question. The user tag cloud most similar to the item's tag cloud shows the most probable ratings for the unknown item. For example, say a target user rated the movie "Emma" with five stars. The system now wants to predict the user's rating for the movie "Northanger Abbey". If the

tag cloud for “Emma”, generated from the target user’s assigned tags, is very similar to the tag cloud for “Northanger Abbey”, generated from all users’ tags, then the system predicts that the target user would rate “Northanger Abbey” with five stars and consequently make a recommendation of the movie. If the tag cloud of “Northanger Abbey” is similar to the target user’s tag cloud for “Pride and Prejudice”, and he or she rated that movie with only one star, the system would predict a rating value of one for “Northanger Abbey” and not recommend it. A more elaborate method, suggested by Szomszor et al. (2007), is the weighed approach, where the frequency of tags is considered similar to the TF*IDF weighting. The results within a Netflix dataset show that the weighted approach performs better than the unweighted approach and a simple rating prediction method based on the average rating of a movie (Szomszor et al., 2007). Therefore, if systems are able to apply tag information in addition to user, item and rating information, recommendation will improve. If additional data about users is available, user profiles will contain more information, which may work against the data sparsity problem. However, most social bookmarking systems do not contain any rating information. Studies and approaches on classic tag-based recommender systems predominate.

The assumptions in tag-based systems described above lead to different implementation techniques, including hybrid approaches that apply a combination of user- and item-based techniques. Studies apply diverse methods and algorithms, such as probabilistic models (Lee, Lee, & Kim, 2011), clustering approaches (Popescul, Flake, Lawrence, Ungar, & Giles, 2000; Ungar & Foster, 1998), graph- and network-based (Wetzker, 2012) or diffusion-based models (Zhang, Zhou, & Zhang, 2010; Zhou, Ren, Medo, & Zhang, 2007). This work will not present a comparison of these diverse models. Its focus lies on examples of expert recommendation, especially for academic purposes, which will be discussed in the next section.

3.5 Expert Recommendations

The field of expert recommendation has developed in two different directions,

namely expert recommendation systems and expert retrieval systems. Both fields have a shared goal, which is to find relevant experts for users in diverse environments, such as academic institutions (Guns & Rousseau, 2013, 2014), companies (Reichling, Veith, & Wulf, 2007), or more generally in social collaborative networks (Kautz, Selman, & Shah, 1997). The difference between the two fields lies in the application of methods to find those experts as well as in the networks where the approaches are applied. Expertise retrieval (Balog, 2012) concentrates on finding relevant people for diverse purposes, whereas expert recommendation systems have mainly developed within specific services that also focus on other areas, such as resource or tag recommendation (Deng, King, & Lyu, 2008; Liu, Curson, & Dew, 2005; Renugadevi, Geetha, Gayathiri, Prathyusha, & Kaviya, 2014). Recommendation and retrieval techniques differ, therefore. Nevertheless, both fields concentrate on defining expertise in a specific environment as well as on “translating” this definition into appropriate measurements that show this expertise and help users find their required experts.

3.5.1 General Approaches in Expert Recommendation

It seems that besides the recommendation of similar resources, recommendation of similar or relevant users is more important, especially in a social tagging system (Diederich & Iofciu, 2006). Panke and Gaiser (2008) (see also Panke & Gaiser, 2009) interviewed approximately 200 users about their tagging behavior: two thirds of them said they use tags to socialize with other users. Schaer, Mayr, and Lüke (2012) conducted a query log analysis of users in the social science database Sowiport⁴. The results showed that about one third of user queries relate to a person and that users explicitly searched for them. As interview statements show (chapter 1), scientists search for known colleagues in order to be up-to-date with current relevant literature. Thus the detection of user-user relations is sufficient to meet users’ requests. In general, recommender systems are embedded in other services such as retrieval systems, shopping

⁴ <http://sowiport.gesis.org/>

services (Amazon⁵) or movie and streaming services (Netflix⁶). The main data sources come from these services; in some cases, external data is included. Data sources include social information about persons, such as publications, citations and references of researchers or bookmarks of web users. This information is produced by the users themselves. Users may also be able to give direct feedback to recommender systems, for example in knowledge-based services. Concerning expertise knowledge, we must distinguish between two kinds of user information. On the one hand, a user produces indirect expertise information of which he or she is not aware. For example, if a researcher assigns tags to articles in a bookmarking system, they mainly do so in order to manage and index their literature list. However, the system may use these tags to analyze the researcher's interests and fields of expertise, for example to recommend them to other users. Hence, the system defines a user's expertise status and generates expertise. On the other hand, a user him- or herself may be able to state their own expertise, for example while adding free or predefined knowledge areas to their profile (Berendsen, De Rijke, Balog, Bogers, & Van den Bosch, 2013). As discussed in chapter 2, this difference is important for researcher recommendation.

There are several studies investigating expert recommendation with variable expertise definition, for example for commercial enterprises (Cai et al., 2011; Reichling & Wulf, 2009). Petry, Tedesco, Vieira, and Salgado (2008), for instance, have developed the expert recommendation system ICARE, designed to recommend experts within an organization. In this system, the focus does not lie on author publications and citations, but rather on individuals' organizational level, availability, and reputation, among other aspects. Following a field study and interviews with employees, Reichling and Wulf (2009) explored a recommender system's capacity to support knowledge management. In this system, experts are defined via their collection of written documents that have

⁵ <http://amazon.com>

⁶ <https://www.netflix.com/>

been analyzed automatically. The authors also used a post-integrated user profile with information about each person's background and job description. Using additional social information from author profiles – besides citation, reference and bookmarking information – might improve the effectiveness of author recommendation. A similar method to conduct expertise is proposed by Seo and Croft (2009), who defined expertise in the Apple discussion forum⁷ on the basis of authorship of relevant posts regarding a specific query. Similarly, academic researcher profiles can be generated from scientific publications. A researcher's expertise is then defined through his or her publications. Each publication is assigned to a specific topic, where the topics define the author expertise fields. Elaborate models learn to generate author-topic models, for example by applying Bayesian measurements (Steyvers, Smyth, Rosen-Zvi, & Griffiths, 2004). Based on these models, the system can propose relevant publications and authors for a specific academic field.

Datta, Yong, and Braghin (2014) propose a team recommender system using diverse data sources including authors, their publications, and co-authorships, where co-authors are defined as collaborating teams. Instead of using citation data, the researchers suggest determining an author's expertise via his or her publications by applying topic extraction techniques. The usage of diverse data sources appears promising, but the study only includes evaluation on a running user interface system, not any evaluation of the recommendations' relevance. Expert recommenders for academics may also concentrate on expert retrieval and are integrated in a search environment. Mutschke, Mayr, Schaer, and Sure (2011) propose showing author networks, created on the basis of author centrality, as complementary sources for finding relevant search results. For instance, a user will perform a search in a database (in this case Sowiport) and get diverse options in order to retrieve his or her results. The author network is one option, showing important authors on the basis of the user query and author centrality analysis in the retrieved results. As a result, the author network shows

⁷ <https://discussions.apple.com/>

relevant and potential partners for a specific field defined by query terms and additionally gives new relevance weights to the retrieved documents. This opportunity makes users aware of new relevant information, as generally they will only be presented with a document rank list based on term weights and are aware of the top 10 retrieved results. However, recommended authors depend on a user query, which might be inappropriate, and results are not personalized to refer to any specific information about the user except the query terms at hand. Expert recommendation based on bibliometric data – as applied in the approach outlined in this work – is a new field, but some studies show first results. For example, future collaborations can be seen to be generating co-authorship networks. Guns and Rousseau (2013; 2014) recommend this method to suggest future partners from different international cities. They generated co-authorship networks (more specifically: based on the city where an author is based professionally) from data in Web of Science. To predict good collaborations, the researchers tested diverse measurements, for example common neighbors and weighted graph distance (Egghe & Rousseau, 2003; Newman, 2001). To train the model and evaluate the results, data was split into three time periods. The later time period was used to evaluate the precision of the results. The researchers claim that good collaborations can be predicted this way and that the accuracy of recommendations is quite high (Guns & Rousseau, 2013, 2014). However, the papers do not discuss whether recommendations are serendipitous or rather obvious. Recommended collaborations often include pairs of cities from the same country. Furthermore, co-occurrence measuring on the city level (based on affiliation information in a publication's metadata) may lead to biased results as authors may change their institutions and work in different places, especially within their native country. Thus new collaborations do not arise from new interactions between two authors who did not know each other previously, but rather from colleagues who have collaborated in the past (albeit in different cities).

3.5.2 Expert Recommendation in Social Tagging Systems

Research into expert recommendation based on social information in tagging

systems applies variants of collaborative filtering methods to expert recommendation. Approaches concentrate on Web 2.0 users and academics (De Meo, Nocera, Terracina, & Ursino, 2011; Deng et al., 2008). Au Yeung, Noll, Gibbins, Meinel, and Shadbolt (2009) discuss the non-academic bookmarking system Del.icio.us. They define an expert user as someone who has deposited high-quality documents in their bookmark collection (said documents being defined by the number of their users with high-level expertise) and who tends to recognize useful documents before others do (as seen in the timestamps on users' bookmarks). Comparatively, the "high-quality documents" in the approach outlined in chapter 4 are the publications of the researcher meant to be recommended collaboration partners. Hence, it is vital for the purposes of recommendation that users bookmark at least one of the target researcher's publications. Ben Jabeur, Tamine, and Boughanem (2010) use social clues like connectivity between researchers and opportunities to meet in person, for example at scientific conferences, to improve the performance of their recommendation system. Nocera and Ursino (2011) focus on "social folksonomy", using information about user friendships and semantic information in tags for their recommendations. However, friendship relations are explicit relations between two individuals. A recommender system does not need to suggest known friends for any purpose. The researchers claim that other social relations besides friendship could be included in their system. The first approach aims at recommending tags or resources based on friends' opinion, but a focus for further research lies in the detection of unknown possible users "who share the same interests and needs" (Nocera & Ursino, 2011, p. 1280). To do so, the approach concentrates on finding reliable users (as well as relevant resources) in different social networks (De Meo, Nocera, Quattrone, Rosaci, & Ursino, 2009; De Meo, Nocera, Rosaci, & Ursino, 2011). In other words, a user is given recommendations of users from his or her own network or others. Diverse methods are applied in order to find relations based on user interactions (interactions between two users from different networks are measured by using fictitious users as representing a network's content), user comments, and resources. Two important aspects are trust and reputation. The latter is a kind of

collaborative opinion that all users share about a specific user. Trust between two users regarding a user's evaluations of another user's opinions within the network must be specified. As discussed above, trust is a kind of crucial factor for users in a recommendation service. The approaches propose good options for recommendations of new and unknown users. However, the users have to be highly active by posting opinions and resources as well as evaluating them. Recommendations are based on users' interactions. Therefore, an inactive or not especially active user will suffer from fewer and less valuable recommendations, as they are based on a small set of user information. Unfortunately, neither approach discusses the problems of user names. For example, a unique user could be a member in two or more networks using diverse nicknames. In a recommendation list, this user could appear several times, which would be inappropriate.

Another proposal is the "FolkRank" designed to recommend users, items or tags to a target user (Hotho, Jäschke, Schmitz, & Stumme, 2006b; Hotho, Jäschke, Schmitz, & Stumme, 2006a; Jäschke et al., 2007). The idea is to apply the principle of the PageRank (Brin & Page, 1998) to folksonomies. A first approach based on the PageRank principle only recommended those tags that were assigned most often in the community (Hotho et al., 2006b). This confirms the feedback loop problem discussed above (Peters, 2009). Thus, Hotho et al. (2006a) modified this "GlobalRank" (Balby Marinho et al., 2012, p. 59) while subtracting its probability distribution from a distribution starting from a fixed user-item pair. This results in a final weight: the FolkRank. The FolkRank thus considers a user's preference while starting graph computation from a specific user-item pair. Results show that a target user is recommended more personal tags, items and users (Hotho et al., 2006b).

To summarize the findings, recommender systems aim to help users find relevant and new items; these can be any items, such as products and scientific articles, tags, and even users. The focus lies on personalized recommendation, which means that recommendations should be adapted to the preferences and purposes of a unique target user. Expert recommendation or expert retrieval

focuses on finding the “right” people for a target user. Intentions in finding relevant people differ relative to a system’s purpose and, of course, a user’s needs. Models to find experts were tested successfully in diverse studies. A main concern here is the significance of the results, as these are system- and user-dependent. Thus the evaluation of those systems is a critical aspect and will be discussed in the following.

3.6 Evaluation of Recommender Systems

The evaluation of recommender systems takes into account several factors. On the one hand, quantitative metrics are applied. On the other hand, user studies are conducted. Concerning quantitative metrics to measure a system’s performance, the most important metrics are accuracy and coverage. Accuracy is a precision factor, where algorithmic variants differ slightly. The use of diverse metrics mainly depends on the purpose of the system in question. A recommender system may try to predict a rating for a target user concerning a specific item. Alternatively, it may recommend a list of new items to a target user who does not yet know these items. In the latter case, precision metrics adapted from the information retrieval field can be applied for accuracy measurement (Shani & Gunawardana, 2011; Stock & Stock, 2013, p. 114).

The precision value shows how often relevant items (positive user-item responses) are recommended relative to the number of non-relevant recommended items (equation 3.1).

$$Precision = \frac{|R_r|}{|R_r| + |R_{nr}|}$$

Equation 3.1. Precision metric, where R_r is the number of relevant recommended items, and R_{nr} the number of non-relevant recommended items.

The recall metric is another factor of a system’s effectiveness in addition to precision.. It shows how often relevant items are recommended relative to all relevant items (equation 3.2). Generally, precision decreases when recall rises. When a recommender system suggests a fixed number of items, precision value is the most important.

$$Recall = \frac{|R_r|}{|R_r| + |NR_r|}$$

Equation 3.2. Recall metric, where R_r is the number of relevant recommended items, and NR_r the number of non-recommended relevant items.

Another value is the false positive rate (Shani & Gunawardana, 2011) (or fallout ratio, also used in information retrieval (Cleverdon, 1967)), which shows the ratio between recommended non-relevant items to non-recommended non-relevant items (equation 3.3). All three metrics belong to the field of accuracy metrics for recommendation systems, which include other algorithmic variants of measuring a system's prediction performance, including mean absolute error and root mean squared error (see Shani and Gunawardana (2011) for further discussion).

$$False\ Positive = \frac{|R_{nr}|}{|R_{nr}| + |NR_{nr}|}$$

Equation 3.3. False positive metric, where R_{nr} is the number of recommended non-relevant items, and NR_{nr} the number of non-recommended non-relevant items.

The choice between the various metrics should be based on the recommender system's goals. For example, a user in a movie recommender wants to be recommended relevant movies, which means that precision should be high. This user is probably not interested in whether the number of recommended irrelevant movies (false positives) is high compared to that of irrelevant movies that are not recommended. In the case of a company that wants to recommend products and ship them to potential buyers as offers for sale, the number of all potential sales might be of interest (Shani & Gunawardana, 2011).


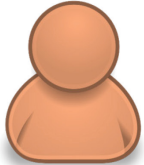
It must be noted that, in contrast to retrieval system evaluation, where user relevance feedback is available (classic scenarios use pre-determined evaluation test collections (Sanderson, 2010)) recommender system evaluation can refer to user feedback only if explicit ratings are available. If that is the case, a system will know explicitly what a user likes and does not like. Training and test sets are applied for the purposes of evaluation. Both sets are derived from a service's

historic user data, which includes information about users and their rated or bookmarked items. This data is divided into training and test sets. In a typical split, 80% of the data is used for training and 20% for testing. User ratings in the test set are deleted and the system then tries to predict them after it developed its model on the basis of the training set.

If no rating data is available, as in most tagging systems, a system will only know which items a user has bookmarked, purchased or used. However, without any user-item relation – in the absence, that is, of a clear user statement – the system will not know what a user thinks about the item. He or she might like the unknown item or not. No explicit relevance is stated unless other information, such as ratings or any other feedback data, is available. If ratings are available, a system can then define positive and negative ratings, for example via thresholds, to distinguish between a positive and negative user-item relation. In a binary system, as in bookmarking services, this option is not available. For that reason, the general assumption needs to be applied that items to which a user has no relation are not relevant, and will therefore be deemed disliked by default. This scenario leads to the following four cases (figure 3.7): All items for which a positive user-item relation exists are seen as relevant. A system is either able to recommend them (correct predictions) or fails to do so (false negatives). All items for which no user-item relation exists are seen as irrelevant. They are either recommended (false positives), or they are not (correct omissions).

The last case leads to false measurements because the pre-decided assumption is incorrect. Possibly a user was simply unaware of the item and therefore did not use or bookmark it. To generally assume that a user does not like an item because he or she has not used it is incorrect. An evaluation based on this scenario would punish a recommender system if it suggests those unknown items (false positives), although a user might, in fact, like them. Jannach et al. (2011, p. 172) conclude that an ideal evaluation of a recommender system should be based on historical user data, where all users have rated all items. Of course this is not realistic, and such a scenario would make a recommendation system unnecessary.

In real user data sets, for example in bookmarking services, users only bookmark a small number of items and a system includes many negative user-item relations. To overcome the incorrect assumption of correct omissions, all non-existing user-item relations could be deleted before any recommendations are measured and evaluated. In a social tagging system, Rendle, Marinho, Nanopoulos, and Schmidt-Thieme (2009) suggest distinguishing between positive, negative, and missing user-item relations. In this case, relations in a folksonomy are described as tensors and user-item relations defined via tags (Balby Marinho et al., 2012). Positive values are determined if a user has assigned a tag to a resource. Negative values are set if a user did not assign a specific tag to a resource, but used other tags to assign the same resource.

		 System recommends item:	
		yes	no
	yes	correct predictions (relevant items)	false negatives (relevant items)
	no	false positives ("irrelevant " items)	correct omissions ("irrelevant" items)

User likes item
(e.g. use, rating,
purchase,
bookmark):

Figure 3.7. Types of recommender system predictions in evaluation processes. Figure adapted from Jannach et al. (2011, p. 171).

Missing values are those where a user did not assign any tag to a resource. Instead of having zero user-item relations each time a user did not assign a specific tag (and only distinguishing between 0 and 1 in a binary approach (Symeonidis, Nanopoulos, & Manolopoulos, 2008)), only those relations are negative where a user did not tag a resource at all. Hence, there are positive (+), negative (-), and zero values. This also leads to a data set which is less sparse because negative values are being counted.

Other recommender evaluation metrics besides recall, precision and fallout ratio concentrate on diverse system qualities. Coverage algorithms measure the ability of a system to recommend a huge amount of stored items. For example, resources in a bookmarking system that are not bookmarked by users, or have no tags, cannot be recommended in a user- or item-based collaborative filtering approach. Other quality values are novelty and serendipity (Baeza-Yates & Ribeiro-Neto, 2011; Herlocker et al., 2004; Lops et al., 2011; Victor et al., 2011). Novel recommended items are those the user was previously unfamiliar with. This quality is one of the most important, because a recommender system would not be of any help to a user if it only suggested items he or she already knew. Serendipity focuses on surprising recommendations that a user would not have found on their own. For example, if user likes Stephen King, and a system recommends books by King the user has not purchased yet, these would be helpful recommendations. However, it is very likely that the user would have found these books without the help of the recommender system. The same holds true for a system's trust value (Victor et al., 2011). As discussed in section 4.3, a user evaluates a system's trustworthiness by reference to items he or she already knows and has liked before. Thus a system that suggests only novel unknown items would likely not be trusted. The same goes for serendipitous items if a user does not see any trust-creating sense in these recommendations⁸.

⁸ For an overview of metric algorithms and other important evaluation metrics, especially for an operating recommender system, see Shani and Gunawardana (2011), Jannach et al. (2011, pp. 166–188) and Herlocker, Konstan, Terveen,

Nevertheless, the evaluation metrics named have some shortcomings. Evaluation based on historical user data is neither capable of making any statements about missing user-item relations, nor can it analyze direct user feedback. The problem can be overcome via user evaluation, where users are asked directly whether an item is relevant or not (Jannach et al., 2011). There are several user evaluation studies available, concentrating on diverse questions (Cosley, Lam, Albert, Konstan, & Riedl, 2003; Das, Datar, Garg, & Rajaram, 2007; Herlocker et al., 2000; Lee, Smeaton, O'Connor, & Smyth, 2006; Mahmood & Ricci, 2009; Sinha & Swearingen, 2001, 2002; Swearingen & Sinha, 2001). Nevertheless, evaluation on historical data prevails because such data sets are easily available via recommender services such as MovieLens⁹, Netflix, or Entree.

Ma, Pant, and Liu Sheng (2007) propose improving searches via personalized categories that are conducted on the basis of a user's interest taxonomy, considering diverse personal information, for example deriving from a user's profile. In addition to using log-file information, they developed a questionnaire consisting of statements with which participants had to disagree or agree on a seven-point Likert scale. This allowed the authors to directly find out whether a system helped a user identify relevant documents more easily and more quickly, for example. Middleton, Shadbolt, and De Roure (2004) built a running recommender system for a computer scientist database in order to suggest relevant papers. They evaluated their system by recording users' web browsing activity and log-files. Additionally, users could give feedback to the recommendations. The authors stress that their long-time study is important for analyzing realistic user behavior and increasing the effectiveness of measuring a system's performance. They complain of a lack of experimental results with real people in recommender system research, but state that their results show system

and Riedl (2004); Herlocker et al. (2004). As the experimental approaches in this work focus on qualitative user surveys and do not include historical data from an active target user, other metric evaluations will not be considered.

⁹ <http://grouplens.org/datasets/movielens/>

performance on a more realistic level than offline experimental evaluations. Similarly, McNee, Riedl, and Konstan (2006) make researchers aware of standard evaluation metrics such as accuracy and serendipity as well as their potential pitfalls.

The evaluation model in De Meo et al. (2009) and De Meo et al. (2011) is similar to the method applied in this work in chapter 4. To evaluate their results, the authors conducted recommendation lists for each target user. These users manually decided which of the suggested persons is reliable for them (respectively, which resources are relevant to them). This information was used to create lists of reliable users (and relevant resources) for a target user, with these reliable users being further subdivided into unknown and known users (users with any relation at all to a target user, for example direct contact in a network). Novelty (recommending unknown users) and correctness (recommending reliable users) are measured according to these lists. Furthermore, there are differences in user notions concerning self-selected expertise. Berendsen et al. (2013) conducted an interesting study testing self-selected expertise against expertise generated by the system. In the first case, participants selected expert areas from a pre-defined list and added them to their profile. In the second case, participants evaluated expert areas that had been recommended by the system. The results show that system-generated profiles are more complete, and participants add more expert areas to their profile, if recommendations make them aware of these. The reason may be that users were unaware of further knowledge areas, or that stating expertise level for all knowledge areas in the system (up to 100 in the approach) simply appeared too time-consuming. Assessing one's own expertise might be difficult without any help, after all, or users simply are not in the habit of doing so. In both cases, it leads to different evaluation results. Hence, the authors suggest using system-generated profiles for system evaluation, as these are more complete (Berendsen et al., 2013). Approaches in chapter 4 will also conduct user studies in which participants evaluate the system via semi-structured interviews. Here the survey includes questions to be answered on rating scales focusing on the approaches' research questions. The results are designed to give a more detailed view of the

target users' opinions that can barely be measured via any quantitative metric. Furthermore, the results of this user evaluation are based on system-generated as well as on user-selected features.

To summarize chapter 3, recommender systems aim to help users find relevant and new items, users and tags. Good recommender systems make it easier and faster for a user to find what he or she is looking for. Additionally, systems like these can help users detect new things that they might have missed without any help. As valuable scientific collaboration comes not only from previously existing networks and properly linked researcher communities, but often derives from new and unexpected relations, and as new young researchers constantly enter the scientific community without having strong colleague relationships, the assumption is that recommendation systems are qualified to support researchers' needs. Chapter 4 shows experimental approaches toward a recommender model based on the aspects of research collaboration and social information. The question is: how can these aspects be considered in a recommendation system that has the purpose of showing relevant experts to a target scientist?

References

- Adomavicius, G., & Tuzhilin, A. (2005). Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734–749. doi:10.1109/TKDE.2005.99
- Aggarwal, C. C., Wolf, J. L., Wu, K.-L., & Yu, P. S. (1999). Horting hatches an egg: A new graph-theoretic approach to collaborative filtering. In *Proceedings of the 5th International Conference on Knowledge Discovery and Data Mining (SIGKDD)* (pp. 201–212). New York, NY, USA: ACM. doi:10.1145/312129.312230
- Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *Proceedings of the International Conference on Management of Data (SIGMOD)* (pp. 207–216). New York, NY, USA: ACM. doi:10.1145/170035.170072
- Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules in large databases. In J. Bocca, M. Jarke, & C. Zaniolo (Eds.): *Proceedings of the 20th International Conference on Very Large Databases (VLDB)*

- (pp. 487–499). San Francisco, CA, USA: Morgan Kaufmann Publishers.
Retrieved from <http://rakesh.agrawal-family.com/papers/vldb94apriori.pdf>
- Ahn, H. J. (2008). A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. *Information Sciences*, 178(1), 37–51. doi:10.1016/j.ins.2007.07.024
- Ahrens, S. C. (2011). *Recommender Systems: Relevance in the Consumer Purchasing Process*. WIM Digital Business. Berlin, Germany: Epubli.
- Amatriain, X., Jaimes, A., Oliver, N., & Pujol, J. M. (2011). Data mining methods for recommender systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 39–71). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_2
- Au Yeung, C. M., Noll, M. G., Gibbins, N., Meinel, C., & Shadbolt, N. R. (2009). On measuring expertise in collaborative tagging systems. In *Proceedings of the Web Science Conference: Society On-Line*. Retrieved from <http://www.websci09.org/proceedings/>
- Baeza-Yates, R., & Ribeiro-Neto, B. (2011). *Modern Information Retrieval: The Concepts and Technology Behind Search*. New York, NY, USA: Addison Wesley.
- Balabanović, M., & Shoham, Y. (1997). Fab: Content-based, collaborative recommendation. *Communications of the ACM*, 40(3), 66–72. doi:10.1145/245108.245124
- Balby Marinho, L., Hotho, A., Jäschke, R., Nanopoulos, A., Rendle, S., Schmidt-Thieme, L., . . . Symeonidis, P. (2012). *Recommender Systems for Social Tagging Systems*. Boston, MA, USA: Springer.
- Balby Marinho, L., Nanopoulos, A., Schmidt-Thieme, L., Jäschke, R., Hotho, A., Stumme, G., & Symeonidis, P. (2011). Social tagging recommender systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 615–644). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_19
- Balog, K. (2012). Expertise Retrieval. *Foundation and Trends in Information Retrieval*, 6(2-3), 127–256. doi:10.1561/15000000024
- Begelman, Grigory, Keller, P., & Smadja, F. (2006). Automated tag clustering: Improving search and exploration in the tag space. In *Proceedings of the Working on Collaborative Web Tagging at the 15th International World Wide Web Conference (WWW)*. New York, NY, USA: ACM. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.120.5736>
- Beliakov, G., Calvo, T., & James, S. (2011). Aggregation of preferences in recommender systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 705–734). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_22
- Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information

- retrieval. *Canadian Journal of Information Science*, 5, 133–143.
- Ben Jabeur, L., Tamine, L., & Boughanem, M. (2010). A social model for literature access: Towards a weighted social network of authors. In *Proceedings of 9th International Conference on Adaptivity, Personalization and Fusion of Heterogeneous Information (RIA0)* (pp. 32–39). Retrieved from <http://dl.acm.org/citation.cfm?id=1937063>
- Berendsen, R., De Rijke, M., Balog, K., Bogers, A. M., & Van den Bosch, A. (2013). On the assessment of expertise profiles. *Journal of the American Society for Information Science and Technology*, 64(10), 2024–2044. doi:10.1002/asi.22908
- Billsus, D., & Pazzani, M. J. (1999). A personal news agent that talks, learns and explains. In *Proceedings of the 3rd Annual Conference on Autonomous Agents* (pp. 268–275). New York, NY, USA: ACM. doi:10.1145/301136.301208
- Brandes, U., Gaertler, M., & Wagner, D. (2003). Experiments on graph clustering algorithms. In G. Goos, J. Hartmanis, J. Van Leeuwen, G. Di Battista, & U. Zwick (Eds.): *Lecture Notes in Computer Science. Algorithms - ESA 2003* (pp. 568–579). Berlin, Heidelberg, GER: Springer. doi:10.1007/978-3-540-39658-1_52
- Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence* (pp. 43–52). San Francisco, CA, USA: Morgan Kaufmann Publishers. Retrieved from <http://dl.acm.org/citation.cfm?id=2074100>
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems.*, 30(1-7), 107–117. doi:10.1016/S0169-7552(98)00110-X
- Brusilovsky, P. (1996). Methods and techniques of adaptive hypermedia. *User Modeling and User-Adapted Interaction*, 6(2-3), 87–129. doi:10.1007/BF00143964
- Burke, R. (2000). Knowledge-based recommender systems. In A. Kent (Ed.): *Encyclopedia of Library and Information Science* (69th ed., No. 32, pp. 180–201). New York, NY, USA: Marcel Dekker.
- Burke, R. (2002a). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12, 331–370. doi:10.1023/A:1021240730564
- Burke, R. (2002b). Interactive critiquing for catalog navigation in e-commerce. *Artificial Intelligence Review*, 18(3/4), 245–267. doi:10.1023/A:1020701617138
- Burke, R. (2007). Hybrid web recommender systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.): *Vol. 4321. Lecture Notes in Computer Science*.

- The Adaptive Web. Methods and Strategies of Web Personalization* (pp. 377–408). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-72079-9_12
- Burke, R., Hammond, K. J., & Yound, B. C. (1997). The FindMe approach to assisted browsing. *IEEE Expert*, 12(4), 32–40. doi:10.1109/64.608186
- Cai, X., Bain, M., Krzywicki, A., Wobcke, W., Kim, Y. S., Compton, P., & Mahidadia, A. (2011). Collaborative filtering for people to people recommendation in social networks. In J. Li (Ed.): *Vol. 6464. Lecture Notes in Computer Science. Proceedings of AI 2010: Advances in Artificial Intelligence* (pp. 476–485). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-642-17432-2_48
- Cleverdon, C. (1967). The Cranfield tests on index language devices. *Aslib Proceedings*, 19(6), 173–194. doi:10.1108/eb050097
- Cole, C. (2012). *Information Need: A Theory Connecting Information Search to Knowledge Formation*. Medford, NJ, USA: Information Today.
- Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., & Riedl, J. T. (2003). Is seeing believing? How recommender system interfaces affect users' opinions. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (pp. 585–592). New York, NY, USA: ACM. doi:10.1145/642611.642713
- Das, A. S., Datar, M., Garg, A., & Rajaram, S. (2007). Google news personalization: Scalable online collaborative filtering. In *Proceedings of the 16th International Conference on World Wide Web (WWW)* (pp. 271–280). New York, NY, USA: ACM. doi:10.1145/1242572.1242610
- Datta, A., Yong, J. T. T., & Braghin, S. (2014). The zen of multidisciplinary team recommendation. *Journal of the American Society for Information Science and Technology*, 65(12), 2518–2533. doi:10.1002/asi.23139
- De Meo, P., Nocera, A., Quattrone, G., Rosaci, D., & Ursino, D. (2009). Finding reliable users and social networks in a social internetworking system. In *Proceedings of the 13th International Database Engineering & Applications Symposium (IDEAS)* (pp. 173–181). New York, NY, USA: ACM. doi:10.1145/1620432.1620450
- De Meo, P., Nocera, A., Rosaci, D., & Ursino, D. (2011). Recommendation of reliable users, social networks and high-quality resources in a social internetworking system. *AI Communications*, 24(1), 31–50. doi:10.3233/AIC-2010-0484
- De Meo, P., Nocera, A., Terracina, G., & Ursino, D. (2011). Recommendation of similar users, resources and social networks in a social internetworking scenario. *Information Sciences*, 181(7), 1285–1305. doi:10.1016/j.ins.2010.12.001
- De Meo, P., Quattrone, G., & Ursino, D. (2009). Exploitation of semantic

- relationships and hierarchical data structures to support a user in his annotation and browsing activities in folksonomies. *Information Systems*, 34(6), 511–535. doi:10.1016/j.is.2009.02.004
- Deng, H., King, I., & Lyu, M. R. (2008). Formal models for expert finding on DBLP bibliography data. In *Proceedings of the 8th International Conference on Data Mining (ICDM)* (pp. 163–172). New York, NY, USA: IEEE. doi:10.1109/ICDM.2008.29
- Deshpande, M., & Karypis, G. (2004). Item-based top-n recommendation algorithms. *ACM Transactions on Information Systems*, 22(1), 143–177. doi:10.1145/963770.963776
- Desrosiers, C., & Karypis, G. (2011). A comprehensive survey of neighborhood-based recommendation methods. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 107–144). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_4
- Diederich, J., & Iofciu, T. (2006). Finding communities of practice from user profiles based on folksonomies. In *Proceedings of the 1st International Workshop on Building Technology Enhanced Learning Solutions for Communities of Practice, co-located with the 1st European Conference on Technology-Enhanced Learning*. Retrieved from <http://www.l3s.de/~diederich/Papers/TBProfile-telcops.pdf>
- Durao, F., & Dolog, P. (2009). A personalized tag-based recommendation in social web systems. In A. Dattolo, C. Tasso, R. Farzan, S. Kleanthous, D. Bueno Vallejo, & J. Vassileva (Eds.): *Proceedings of Workshop on Adaptation and Personalization for Web 2.0 at the 17th International Conference on User Modeling, Adaptation, and Personalization (UMAP)*. Retrieved from <http://ceur-ws.org/Vol-485/paper5-F.pdf>
- Egghe, L., & Rousseau, R. (2003). A measure for the cohesion of weighted networks. *Journal of the American Society for Information Science and Technology*, 54(3), 193–202. doi:10.1002/asi.10155
- Felfernig, A., & Burke, R. (2008). Constraint-based recommender systems. In D. Fensel & H. Werthner (Eds.): *Proceedings of the 10th International Conference on Electronic Commerce (ICEC)* (p. 3). New York, NY, USA: ACM. doi:10.1145/1409540.1409544
- Fellbaum, C. (2005). WordNet and wordnets. In K. Brown (Ed.): *Encyclopedia of Language and Linguistics* (pp. 665–670). Oxford, UK: Elsevier.
- Galvez, C., De Moya-Anegón, F., & Solana, V. H. (2005). Term conflation methods in information retrieval: non-linguistic and linguistic approaches. *Journal of Documentation*, 61(4), 520–547. doi:10.1108/00220410510607507
- Gärtner, K. (2012). *Analyse von Recommendersystemen in Deutschland. B.I.T. online - Innovativ: Vol. 38*. Wiesbaden, Germany: Dinges & Frick.

- Gemmell, J., Shepitsen, A., Mobasher, B., & Burke, R. (2008a). Personalization in folksonomies based on tag clustering. In *Proceedings of the 6th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems* (pp. 37–48). Menlo Park, CA, USA: AAAI. Retrieved from <http://www.aaai.org/Library/Workshops/ws08-06.php>
- Gemmell, J., Shepitsen, A., Mobasher, B., & Burke, R. (2008b). Personalizing navigation in folksonomies using hierarchical tag clustering. In I.-Y. Song, J. Eder, & T. M. Nguyen (Eds.): *Lecture Notes in Computer Science. Data Warehousing and Knowledge Discovery* (pp. 196–205). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-85836-2_19
- Gena, C., Brogi, R., Cena, F., & Venero, F. (2011). The impact of rating scales on user's rating behavior. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, . . . N. Oliver (Eds.): *Lecture Notes in Computer Science. User Modeling, Adaption and Personalization* (pp. 123–134). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-22362-4_11
- Goldberg, D., Nichols, D. M., Oki, B. M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), 61–70. doi:10.1145/138859.138867
- Guha, R., Kumar, R., Raghavan, P., & Tomkins, A. (2004). Propagation of trust and distrust. In *Proceedings of the 13th International Conference on World Wide Web (WWW)* (pp. 403–412). New York, NY, USA: ACM. doi:10.1145/988672.988727
- Guns, R., & Rousseau, R. (2013). Predicting and recommending potential research collaborations. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.): *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference (ISSI)* (pp. 1409–1418). Wien, Austria: Facultas Verlags- und Buchhandels AG. Retrieved from <http://www.issi2013.org/proceedings.html>
- Guns, R., & Rousseau, R. (2014). Recommending research collaborations using link prediction and random forest classifiers. *Scientometrics*, 101(2), 1461–1473. doi:10.1007/s11192-013-1228-9
- Guy, M., & Tonkin, E. (2006). Folksonomies: Tidying up tags? *D-Lib Magazine*, 12(1). doi:10.1045/january2006-guy
- Harman, D. (1991). How effective is suffixing? *Journal of the American Society for Information Science*, 42(1), 7–15. doi:10.1002/(SICI)1097-4571(199101)42:1<7::AID-ASI2>3.0.CO;2-P
- Heck, T., & Peters, I. (2010). Expert recommender systems: Establishing communities of practice based on social bookmarking. In *Proceedings of 10th International Conference on Knowledge Management and Knowledge Technologies (I-Know)* (pp. 458–464). Graz, Austria: J.UCS. Retrieved from <http://www.isi.hhu.de/heck>

- Heckerman, D., Chickering, D. M., Meek, C., Rounthwaite, R., & Kadie, C. (2001). Dependency networks for inference, collaborative filtering, and data visualization. *The Journal of Machine Learning Research*, 1, 49–75. doi:10.1162/153244301753344614
- Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. T. (1999). An algorithmic framework for performing collaborative filtering. In *Proceedings of the 22nd Annual International Conference on Research and Development in Information Retrieval (SIGIR)* (pp. 230–237). New York, NY, USA: ACM. doi:10.1145/312624.312682
- Herlocker, J. L., Konstan, J. A., & Riedl, J. T. (2000). Explaining collaborative filtering recommendations. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW)* (pp. 241–250). New York, NY, USA: ACM. doi:10.1145/358916.358995
- Herlocker, J. L., Konstan, J. A., & Riedl, J. T. (2002). An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Information Retrieval*, 5(4), 287–310. doi:10.1023/A:1020443909834
- Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1), 5–53. doi:10.1145/963770.963772
- Herrmann, A., Heitmann, M., & Polak, B. (2007). Die Macht der Defaults. *Absatzwirtschaft - Zeitschrift für Marketing*, 6, 46–47. Retrieved from http://printarchiv.absatzwirtschaft.de/Content/_p=1004040,an=060701027
- Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006a). FolkRank: A ranking algorithm for folksonomies. In *Proceedings of the Workshop Information Retrieval of the Special Interest Group Information Retrieval* (pp. 111–114). Retrieved from <http://www.kde.cs.uni-kassel.de/stumme/papers/2006/hotho2006folkrank.pdf>
- Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006b). Information retrieval in folksonomies: Search and ranking. In Y. Sure & J. Domingue (Eds.): *Lecture Notes in Computer Science: Vol. 4011. Proceedings of The Semantic Web: Research and Applications. 3rd European Semantic Web Conference (ESWC)* (pp. 411–426). Berlin/Heidelberg, Germany: Springer. doi:10.1007/11762256_31
- Huang, Z., Chen, H., & Zeng, D. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems*, 22(1), 116–142. doi:10.1145/963770.963775
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2011). *Recommender Systems - An Introduction*. New York, NY, USA: Cambridge Univ. Press.
- Jäschke, R., Marinho, L. B., Hotho, A., Schmidt-Thieme, L., & Stumme, G. (2007). Tag recommendations in folksonomies. In J. N. Kok, J. Koronacki,

- R. Lopez de Mantaras, S. Matwin, D. Mladenic, & A. Skowron (Eds.): *Vol. 4702. Lecture Notes in Computer Science. Knowledge Discovery in Databases: PKDD* (pp. 506–514). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-74976-9_52
- Juršič, M., Mozetič, I., Erjavec, T., & Lavrač, N. (2010). LemmaGen: Multilingual lemmatisation with induced ripple-down rules. *Journal of Universal Computer Science*, 16(9), 1190–1214. doi:10.3217/jucs-016-09-1190
- Kautz, H., Selman, B., & Shah, M. (1997). Referral Web: combining social networks and collaborative filtering. *Communications of the ACM*, 40(3), 63–65. doi:10.1145/245108.245123
- Kipp, M. E. (2006). @toread and Cool: Tagging for time, task and emotion. In *Proceedings of the 17th Annual ASIS&T SIG/CR Classification Research Workshop*. Austin, TX, USA. Retrieved from <http://eprints.rclis.org/13909/>
- Klahold, A. (2009). *Empfehlungssysteme: Recommender Systems - Grundlagen, Konzepte und Lösungen*. Wiesbaden, Germany: Vieweg + Teubner in GWV Fachverlage.
- Knautz, K. (2008). Von der Tag-Cloud zum Tag-Cluster. Statistischer Thesaurus auf der Basis syntagmatischer Relationen und seine mögliche Nutzung in Web 2.0-Diensten. In *Verfügbarkeit von Informationen. Proceedings der 30. DGI-Online-Tagung*. (Proceedings der 30. DGI-Online-Tagung 2008, pp. 269–284). Frankfurt a. M., Germany: DGI.
- Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. T. (1997). GroupLens: Applying collaborative filtering to Usenet News. *Communications of the ACM*, 40(3), 77–87. doi:10.1145/245108.245126
- Koren, Y., & Bell, R. (2011). Advances in Collaborative Filtering. In F. Ricci (Ed.): *Recommender systems handbook* (pp. 145–186). New York, London: Springer. doi:10.1007/978-0-387-85820-3_5
- Kuhlen, R. (1977). *Experimentelle Morphologie in der Informationswissenschaft*. München, Germany: Verlag Dokumentation.
- Kuhlthau, C. C. (2004). *Seeking Meaning: A Process Approach to Library and Information Services* (2nd ed.). Westport, CT, USA: Libraries Unlimited.
- Lee, H., Smeaton, A. F., O'Connor, N. E., & Smyth, B. (2006). User evaluation of fishclár-news: an automatic broadcast news delivery system. *ACM Transactions on Information Systems*, 24(2), 145–189. doi:10.1145/1148020.1148021
- Lee, J.-W., Lee, S.-G., & Kim, H.-J. (2011). A probabilistic approach to semantic collaborative filtering using world knowledge. *Journal of Information Science*, 37(1), 49–66. doi:10.1177/0165551510392318
- Li, J., Li, L., Wen, X., & Liao, J. (2012). A collaborative filtering

- recommendation system combining semantics and Bayesian reasoning. In Y. Zhao, J. Li, P. Kennedy, & P. Christen (Eds.): *Proceedings of the 10th Australasian Data Mining Conference (AusDM). Conferences in Research and Practice in Information Technology* (pp. 191–197). Darlinghurst, Australia: Australian Computer Society. Retrieved from <http://crpit.com/confpapers/CRPITV134Li.pdf>
- Li, T., Ogihara, M., & Li, Q. (2003). A comparative study on content-based music genre classification. In *Proceedings of the 26th Annual International Conference on Research and Development in Information Retrieval (SIGIR)* (pp. 282–289). New York, NY, USA: ACM. doi:10.1145/860435.860487
- Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1), 76–80. doi:10.1109/MIC.2003.1167344
- Liu, P., Curson, J., & Dew, P. (2005). Use of RDF for expertise matching within academia. *Knowledge and Information Systems*, 8(1), 103–130. doi:10.1007/s10115-004-0152-y
- Lops, P., Gemmis, M. de, & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 73–105). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_3
- Luhn, H. P. (1958). A business intelligence system. *IBM Journal of Research and Development*, 2(4), 314–319. doi:10.1147/rd.24.0314
- Ma, Z., Pant, G., & Liu Sheng, O. R. (2007). Interest-based personalized search. *ACM Transactions on Information Systems*, 25(1), 5. doi:10.1145/1198296.1198301
- Mahmood, T., & Ricci, F. (2009). Improving recommender systems with adaptive conversational strategies. In *Proceedings of the 20th ACM Conference on Hypertext and Hypermedia (HT)* (pp. 73–82). New York, NY, USA: ACM. doi:10.1145/1557914.1557930
- Malone, T. W., Grant, K. R., & Turbak, F. A. (1986). The information lens: an intelligent system for information sharing in organizations. In M. Mantei & P. Orbeton (Eds.): *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI)* (pp. 1–8). New York, NY, USA: ACM. doi:10.1145/22627.22340
- Mandl, M., Felfernig, A., Teppan, E., & Schubert, M. (2011). Consumer decision making in knowledge-based recommendation. *Journal of Intelligent Information Systems*, 37(1), 1–22. doi:10.1007/s10844-010-0134-3
- Massa, P., & Avesani, P. (2004). Trust-aware collaborative filtering for recommender systems. In R. Meersman & Z. Tari (Eds.): *Vol. 3290. Lecture Notes in Computer Science. Proceedings of OTM Confederated International Conferences, CoopIS, DOA, and ODBASE* (pp. 492–508).

- Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-30468-5_31
- Massa, P., & Avesani, P. (2009). Trust metrics in recommender systems. In J. Golbeck (Ed.): *Human-Computer Interaction Series. Computing with Social Trust* (pp. 259–285). London, UK: Springer. doi:10.1007/978-1-84800-356-9_10
- McNee, S. M., Kapoor, N., & Konstan, J. A. (2006). Don't look stupid: Avoiding pitfalls when recommending research papers. In *Proceedings of the 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW)* (pp. 171–180). New York, NY, USA: ACM. doi:10.1145/1180875.1180903
- McNee, S. M., Riedl, J. T., & Konstan, J. A. (2006). Accurate is not always good: How accuracy metrics have hurt recommender systems. In *Extended Abstracts of the Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM. Retrieved from <http://www-users.cs.umn.edu/~mcnee/mcnee-chi06-acc.pdf>
- Middleton, S. E., Shadbolt, N. R., & De Roure, D. C. (2004). Ontological user profiling in recommender systems. *ACM Transactions on Information Systems*, 22(1), 54–88. doi:10.1145/963770.963773
- Mobasher, B., Dai, H., Luo, T., & Nakagawa, M. (2001). Effective personalization based on association rule discovery from web usage data. In E.-P. Lim (Ed.): *Proceedings of the 3rd International Workshop on Web Information and Data Management (CIKM)* (pp. 9–15). New York, NY, USA: ACM. doi:10.1145/502932.502935
- Mobasher, B., Dai, H., Luo, T., & Nakagawa, M. (2002). Discovery and evaluation of aggregate usage profiles for web personalization. *Data Mining and Knowledge Discovery*, 6(1), 61–82. doi:10.1023/A:1013232803866
- Muller, M. J. (2007). Comparing tagging vocabularies among four enterprise tag-based services. In *Proceedings of the International ACM Conference on Supporting Group Work, Group'07* (pp. 341–350). New York, NY: ACM. doi:10.1145/1316624.1316676
- Mutschke, P., Mayr, P., Schaer, P., & Sure, Y. (2011). Science models as value-added services for scholarly information systems. *Scientometrics*, 89(1), 349–364. doi:10.1007/s11192-011-0430-x
- Neal, D. (2007). Folksonomies and image tagging: Seeing the future? *Bulletin of the American Society for Information Science and Technology*, 34(1), 7–11. doi:10.1002/bult.2007.1720340104
- Newman, M. E. J. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, 64(016132). doi:10.1103/PhysRevE.64.016132
- Nichols, D. M. (1997). Implicit rating and filtering. In *Proceedings of the 5th Delos Workshop on Filtering and Collaborative Filtering* (pp. 28–33). Biot,

- France: Ercim. Retrieved from <http://www.ercim.eu/publication/ws-proceedings/DELOS5/>
- Nocera, A., & Ursino, D. (2011). An approach to provide a user with recommendations of similar users and potentially interesting resources. *Knowledge-Based Systems*, 24(8), 1277–1296. doi:10.1016/j.knsys.2011.06.003
- Oard, D. W., & Kim, J. (1998). Implicit feedback for recommender systems. In *Workshop on Recommender Systems. AAAI Technical WS-98-08* (pp. 81–83). Menlo Park, CA, USA: AAAI. Retrieved from <https://www.aaai.org/Papers/Workshops/1998/WS-98-08/WS98-08-021.pdf>
- Panke, S. & Gaiser, B. (2008). *Nutzerperspektiven auf Social Tagging – Eine Online Befragung*. Retrieved from www.e-teaching.org/didaktik/recherche/goodtagsbadtags2.pdf
- Panke, S., & Gaiser, B. (2009). “With my head up in the clouds” - using social tagging to organize knowledge. *Journal of Business and Technical Communication*, 23(3), 318–349. doi:10.1177/1050651909333275
- Paolillo, J., & Penumarthy, S. (2007). The social structure of tagging internet video on del.icio.us. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences, HICSS'07* (p. 85). IEEE. doi:10.1109/HICSS.2007.555
- Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.): *Vol. 4321. Lecture Notes in Computer Science. The Adaptive Web. Methods and Strategies of Web Personalization* (pp. 325–341). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-72079-9_10
- Peters, I. (2009). *Folksonomies. Indexing and Retrieval in Web 2.0*. Berlin, Germany: De Gruyter Saur.
- Peters, I., & Weller, K. (2008). Tag gardening for folksonomy enrichment and maintenance. *Webology*, 5(3). Retrieved from <http://www.webology.org/2008/v5n3/a58.html>
- Petry, H., Tedesco, P., Vieira, V., & Salgado, A. C. (2008). ICARE. A context-sensitive expert recommendation system. In *Proceedings of the Workshop on Recommender Systems at the 18th European Conference on Artificial Intelligence (ECAI)* (pp. 53–58). Retrieved from http://www.academia.edu/14177855/ICARE_A_context-sensitive_expert_recommendation_system
- Picault, J., Ribière, M., Bonnefoy, D., & Mercerm, K. (2011). How to get the recommender out of the lab? In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 333–365). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_10
- Plisson, J., Lavrač, N., Mladenic, D., & Erjavec, T. (2008). Ripple down rule

- learning for automated word lemmatisation. *AI Communications*, 21(1), 15–26. Retrieved from <http://content.iospress.com/articles/ai-communications/aic416>
- Popescul, A., Flake, G. W., Lawrence, S., Ungar, L. H., & Giles, C. L. (2000). Clustering and identifying temporal trends in document databases. In *Proceedings of Advances in Digital Libraries* (pp. 173–182). New York, NY, USA: IEEE. doi:10.1109/ADL.2000.848380
- Reichling, T., Veith, M., & Wulf, V. (2007). Expert recommender: Designing for a network organization. *Computer Supported Cooperative Work (CSCW)*, 16(4-5), 431–465. doi:10.1007/s10606-007-9055-2
- Reichling, T., & Wulf, V. (2009). Expert recommender systems in practice: Evaluating semi-automatic profile generation. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (pp. 59–68). New York, NY, USA: ACM. doi:10.1145/1518701.1518712
- Rendle, S., Marinho, L. B., Nanopoulos, A., & Schmidt-Thieme, L. (2009). Learning optimal ranking with tensor factorization for tag recommendation. In *Proceedings of the 15th International Conference on Knowledge Discovery and Data Mining (SIGKDD)* (pp. 727–736). New York, NY, USA: ACM. doi:10.1145/1557019.1557100
- Renugadevi, S., Geetha, T. V., Gayathiri, R. L., Prathyusha, S., & Kaviya, T. (2014). Collaborative search using an implicitly formed academic network. *Aslib Journal of Information Management*, 66(5), 537–552. doi:10.1108/AJIM-12-2013-0142
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. T. (1994). GroupLens: An open architecture for collaborative filtering of Netnews. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW)* (pp. 175–186). New York, NY, USA: ACM. doi:10.1145/192844.192905
- Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56–58. doi:10.1145/245108.245121
- Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 1–35). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_1
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (Eds.). (2011). *Recommender Systems Handbook*. New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3
- Robertson, S. (2004). Understanding inverse document frequency: On theoretical arguments for IDF. *Journal of Documentation*, 60(5), 503–520. doi:10.1108/00220410410560582
- Said, A., Wetzker, R., Umbrath, W., & Hennig, L. (2009). A hybrid PLSA

- approach for warmer cold start in folksonomy recommendation. In *Proceedings of Workshop on Recommender Systems & The Social Web at the International Conference on Recommender Systems (RecSys)* (pp. 87–90). New York, NY, USA: ACM. Retrieved from <http://ls13-www.cs.tu-dortmund.de/homepage/RSWEB/papers/RSWEB-Proceedings.pdf>
- Salter, J., & Antonopoulos, N. (2006). CinemaScreen recommender agent: Combining collaborative and content-based filtering. *IEEE Intelligent Systems*, 21(1), 35–41. doi:10.1109/MIS.2006.4
- Salton, G. (Ed.). (1971). *The SMART Retrieval System—Experiments in Automatic Document Processing*. Upper Saddle River, NJ, USA: Prentice-Hall.
- Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620. doi:10.1145/361219.361220
- Sanderson, M. (2010). Test collection based evaluation of information retrieval systems. *Foundations and Trends in Information Retrieval*, 4(4), 247–375. doi:10.1561/1500000009
- Sanderson, M., & Croft, B. W. (1999). Deriving concept hierarchies from text. In *Proceedings of the 22nd Annual International Conference on Research and Development in Information Retrieval (SIGIR)* (pp. 206–213). New York, NY, USA: ACM. doi:10.1145/312624.312679
- Sarwar, B., Karypis, G., Konstan, J. A., & Riedl, J. T. (2000). Analysis of recommendation algorithms for e-commerce. In *Proceedings of the 2nd Conference on Electronic Commerce (EC)* (pp. 158–167). New York, NY, USA: ACM. doi:10.1145/352871.352887
- Sarwar, B., Karypis, G., Konstan, J. A., & Riedl, J. T. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on the World Wide Web (WWW)* (pp. 285–295). New York, NY, USA: ACM. doi:10.1145/371920.372071
- Schaer, P., Mayr, P., & Lüke, T. (2012). Extending term suggestion with author names. In P. Zaphiris, G. Buchanan, E. Rasmussen, & F. Loizides (Eds.): *Vol. 7489. Lecture Notes in Computer Science. Theory and Practice of Digital Libraries* (pp. 317–322). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-642-33290-6_34
- Schafer, J. B., Frankowski, D., Herlocker, J. L., & Sen, S. (2007). Collaborative filtering recommender systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.): *Vol. 4321. Lecture Notes in Computer Science. The Adaptive Web. Methods and Strategies of Web Personalization* (pp. 291–324). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-72079-9_9
- Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. *Data Mining and Knowledge Discovery*, 5(1/2), 115–153.

- doi:10.1023/A:1009804230409
- Schmitz, P. (2006). Inducing ontology from Flickr tags. In *Proceedings of the Collaborative Web Tagging Workshop at the 15th International World Wide Web Conference, WWW'06*. New York, NY: ACM. Retrieved from <http://libra.msra.cn/Publication/3237218/inducing-ontology-from-flickr-tags>
- Seo, J., & Croft, B. W. (2009). Thread-based expert finding. In *Proceedings of the Workshop on Search in Social Media. Annual International Conference on Research and Development in Information Retrieval (SIGIR)*. New York, NY, USA: ACM. Retrieved from <http://ciir-publications.cs.umass.edu/getpdf.php?id=893>
- Shani, G., & Gunawardana, A. (2011). Evaluating recommendation systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 257–297). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_8
- Shardanand, U., & Maes, P. (1995). Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)* (pp. 210–217). New York, NY, USA: ACM. doi:10.1145/223904.223931
- Shen, J., Cui, B., Shepherd, J., & Tan, K.-L. (2006). Towards efficient automated singer identification in large music databases. In *Proceedings of the 29th Annual International Conference on Research and Development in Information Retrieval (SIGIR)* (pp. 59–66). New York, NY, USA: ACM. doi:10.1145/1148170.1148184
- Shepitsen, A., Gemmell, J., Mobasher, B., & Burke, R. (2008). Personalized recommendation in social tagging systems using hierarchical clustering. In *Proceedings of the 2nd Conference on Recommender Systems (RecSys)* (pp. 259–266). New York, NY, USA: ACM. doi:10.1145/1454008.1454048
- Sigurbjörnsson, B., & Van Zwol, R. (2008). Flickr tag recommendation based on collective knowledge. In J. Huai (Ed.): *Proceedings of the 17th International Conference on World Wide Web (WWW)* (pp. 327–336). New York, NY, USA: ACM. doi:10.1145/1367497.1367542
- Sinclair, J., & Cardew-Hall, M. (2007). The folksonomy tag cloud: When is it useful? *Journal of Information Science*, 34(1), 15–29. doi:10.1177/0165551506078083
- Sinha, R. R., & Swearingen, K. (2001). Comparing recommendations made by online systems and friends. In *Proceedings of the 2nd DELOS Network of Excellence Workshop on Personalisation and Recommender Systems in Digital Libraries*. Biot, France: Ercim. Retrieved from <http://www.ercim.eu/publication/ws-proceedings/DelNoe02/RashmiSinha.pdf>
- Sinha, R. R., & Swearingen, K. (2002). The role of transparency in

- recommender systems. In *Extended Abstracts on Human Factors in Computing Systems (CHI)* (pp. 830–831). New York, NY, USA: ACM. doi:10.1145/506443.506619
- Sparck Jones, K. (1973). Index term weighting. *Information Storage and Retrieval*, 9(11), 619–633. doi:10.1016/0020-0271(73)90043-0
- Srikant, R., & Agrawal, R. (1995). Mining generalized association rules. In U. Dayal, P. M. D. Gray, & S. Nishio (Eds.): *Proceedings of the 21st International Conference on Very Large Data Bases (VLDB)* (pp. 407–419). San Francisco, CA, USA: Morgan Kaufmann Publishers. Retrieved from <http://www.vldb.org/conf/1995/P407.PDF>
- Srikant, R., & Agrawal, R. (1997). Mining generalized association rules. *Future Generation Computer Systems*, 13(2-3), 161–180. doi:10.1016/S0167-739X(97)00019-8
- Steyvers, M., Smyth, P., Rosen-Zvi, M., & Griffiths, T. L. (2004). Probabilistic author topic models for information discovery. In *Proceedings of the 10th Conference on Knowledge Discovery and Data Mining (KDD)* (pp. 306–315). New York, NY, USA: ACM. doi:10.1145/1014052.1014087
- Stock, W. G., & Stock, M. (2013). *Handbook of Information Science*. Berlin/Boston: De Gruyter.
- Swearingen, K., & Sinha, R. R. (2001). Beyond algorithms: An HCI perspective on recommender systems. In *Proceedings of the Workshop on Recommender Systems at the 24th Annual International ACM SIGIR Conference (SIGIR)*. New York, NY, USA: ACM. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.23.9764>
- Symeonidis, P., Nanopoulos, A., & Manolopoulos, Y. (2008). Tag recommendations based on tensor dimensionality reduction. In *Proceedings of the 2nd Conference on Recommender Systems (RecSys)* (pp. 43–50). New York, NY, USA: ACM. doi:10.1145/1454008.1454017
- Szomszor, M., Cattuto, C., Alani, H., O'Hara, K., Baldassarri, A., Loreto, V., & Servedio, V. D. (2007). Folksonomies, the semantic web, and movie recommendation. In *Proceedings of the International Workshop on Bridging the Gap between Semantic Web and Web 2.0 at the 4th European Semantic Web Conference (ESWC)*. Retrieved from <http://eprints.soton.ac.uk/264007/>
- Taylor, R. S. (1968). Question-Negotiation and information seeking in libraries. *College & Research Libraries*, 29(3), 178–194. doi:10.5860/crl_29_03_178
- Ungar, L. H., & Foster, D. P. (1998). Clustering methods for collaborative filtering. In *Proceedings of the Workshop on Recommender Systems, AAAI Technical Report WS-98-08* (pp. 114–129). Menlo Park, CA, USA: AAAI. Retrieved from <https://www.aaai.org/Papers/Workshops/1998/WS-98-08/WS98-08-029.pdf>
- Victor, P., Cornelis, C., De Cock, M., & Da Silva, P. P. (2009). Gradual trust

- and distrust in recommender systems. *Fuzzy Sets and Systems*, 160(10), 1367–1382. doi:10.1016/j.fss.2008.11.014
- Victor, P., De Cock, M., & Cornelis, C. (2011). Trust and recommendations. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 645–675). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_20
- Voorhees, E. M. (1998). Using WordNet for text retrieval. In C. Fellbaum (Ed.): *WordNet. An Electronic Lexical Database* (pp. 285–303). Cambridge/London: MIT Press.
- Wetzker, R. (2012). *Graph-Based Recommendation in Broad Folksonomies: Doctoral Thesis*. Retrieved from <https://opus4.kobv.de/opus4-tuberlin/frontdoor/index/index/docId/2557>
- Zanker, M., & Jessenitschnig, M. (2009). Case-studies on exploiting explicit customer requirements in recommender systems. *User Modeling and User-Adapted Interaction*, 19(1-2), 133–166. doi:10.1007/s11257-008-9048-y
- Zhang, Z.-K., Zhou, T., & Zhang, Y.-C. (2010). Personalized recommendation via integrated diffusion on user–item–tag tripartite graphs. *Physica A: Statistical Mechanics and its Applications*, 389(1), 179–186. doi:10.1016/j.physa.2009.08.036
- Zhao, G., Qian, X., & Xie, X. (2016). User-service rating prediction by exploring social users' rating behaviors. *IEEE Transactions on Multimedia*, 18(3), 496–506. doi:10.1109/TMM.2016.2515362
- Zhao, S., Du, N., Nauerz, A., Zhang, X., Yuan, Q., & Fu, R. (2008). Improved recommendation based on collaborative tagging behaviors. In J. Bradshaw (Ed.): *Proceedings of the 13th International Conference on Intelligent User Interfaces (IUI)* (pp. 413–416). New York, NY, USA: ACM. doi:10.1145/1378773.1378843
- Zhou, T., Ren, J., Medo, M., & Zhang, Y. C. (2007). Bipartite network projection and personal recommendation. *Physical Review E*, 76(4), 46115. doi:10.1103/PhysRevE.76.046115
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Cambridge, MA, USA: Addison Wesley.

4 Case studies on Expert Recommendation

With regard to the research questions, the following approaches conduct studies on expert recommendation for researchers while analyzing relations based on social information about researchers. Recommendation networks should make researchers aware of potential collaborators for the purposes of community building and scientific knowledge creation.

The research questions are:

1. Are scientometric and altmetric methods appropriate for determining relevant new collaboration structures for scientific purposes?
2. Are databases and services containing bibliometric and altmetric data structures appropriate sources for detecting relations between researchers?
3. Can bibliometric and altmetric methods be used to establish a recommender system for scientific purposes?

The final idea of suggesting scientific collaborators is explored in the third approach (section 4.3). Sections 4.1 and 4.2 introduce pre-studies, which regard relevant aspects for the construction of expert recommendation in section 4.3. The pre-studies (Heck, 2011; Heck & Peters, 2010a, 2010b, 2010c) show some initial results concerning bookmarking data and similarity relations. The first study analyzes the structure of bookmarking data compiled from three different services and gives an overview of the data structure of folksonomy-based bookmarking systems. Furthermore, the three standard similarity algorithms are compared in order to choose the most appropriate one for further measurements.

The second study simulates a recommender system, which recommends similar users to a target researcher on the basis of social information gleaned from a social bookmarking service. The evaluation gives initial insights into the appropriateness of bookmarking data for finding relevant research partners. Both studies served as a preparation for conducting the final experimental approach and provided relevant results for improving the proposed model

(Heck, 2012a, 2012b, 2012c, 2013; Heck, Hanraths, & Stock, 2011; Heck, Peters, & Stock, 2011).

4.1 Structure of Bookmarking Data and Similarity Metrics

This first study aims to analyze data from three social bookmarking services in light of the following questions:

- How is data structured in folksonomy-based services?
- Which similarity metric is appropriate for expert recommendations?

The results give an overview of user bookmarking behavior concerning a specific research discipline, which will also be considered for the following studies. The comparison of three diverse similarity metrics shows differences in the results of user-user relations, which should be considered for further recommendation approaches.

4.1.1 Data Set

The model analyzes bookmarking data from three services, namely BibSonomy, CiteULike and Connotea¹. In all services, users are able to bookmark scientific publications such as articles, contributions to proceedings, and books. In order to aggregate data from those services in a logical manner, data is limited to a scientific discipline. One reason for this is to obtain correct data about scientific publications, because bookmarking services include noisy data and bookmarks that do not refer to scientific publications. The second reason is that researchers from this scientific discipline participated in the further studies, and data from their field is more revealing for further discussions on the appropriateness of this data for partner recommendation.

As a result, 45 relevant journals from the field of solid-state physics were chosen (table 4.1). The choice of these journals is described in more detail in the

¹ Note: Connotea was discontinued in 2013. Mendeley was not considered at the time of this study because the service did not offer any bookmarking data.

work by Haustein (2012), who analyzed the journals' impact while including not only classical journal impact factors, but also information from bookmarking services. All three services searched the bookmarks of all articles published in the 45 periodicals between 2004 and 2008. The DOI (digital object identifier²), ISSN numbers and UT code from Web of Science were used for the search in order to determine the correct bookmark and identify unique articles.

Act Cryst A	J Phys A	Phys Rev A
Act Cryst B	J Phys Condens Matter	Phys Rev E
Ann Phys	J Phys D	Phys Scr
Appl Phys A	J Rheol	Phys Solid State
Appl Phys Let	J Stat Mech	Phys Stat Sol A
Comput Mater Sci	J Stat Phys	Phys Stat Sol B
EPL	J Vac Sci Technol A	Phys Today
Eur Phys J B	JETP Lett	Physica B
Eur Phys J E	Nanotechnol	Physica C
Hyperfine Interact	New J Phys	Pramana
IEEE Nanotechnol	Nucl Instrum Meth A	Rep Prog Phys
Int J Thermophys	Nucl Instrum Meth B	Rev Mod Phys
J Appl Phys	Phys Fluids	Soft Matter
J Low Temp Phys	Phys Lett A	Solid State Ion
J Magn Magn Mater	Phys Rep	Supercond Sci Technol

Table 4.1. Chosen journals (abbreviations) for the data set.

Vague bookmarking data was determined manually and either corrected or deleted if the bookmarks did not refer to the determined articles. For each of the correct bookmarks retrieved, all bibliographic data as well as related user and tag information was collected. Data from all three services was merged into one database and relations were specified between bookmarks from diverse services that refer to the same article. Bookmarks that refer to a unique article were identified via DOIs and verified via the UT-code. Hence, the number of false

² <http://www.doi.org/>

relations is diminished and almost entirely eliminated. Unique users were also determined manually. Upper- and lower-case letters were ignored, and user names that varied only in this regard were attributed to a unique user. In the data set, there are four cases (unique users) in which user names were the same except for their upper- and lower-case forms (mostly, the first letter differed in this regard). In four more cases, user names were adjusted manually and added to a unique user as the names differed only slightly. For example, “paul_hopkins” was adjusted to “paulhopkins”. This folksonomy-based sub-set is the basis for the following analyses.

4.1.2 Data Structure

Generating the data set leads to the following basic numbers: 2,437 unique users bookmarked 10,280 unique articles from the 45 periodicals. The total number of all bookmarks was 13,608. The analysis of user names shows that 74 users have a profile in more than one bookmarking system. Only one user is found across all three services.

	bibsonomy	citeulike	connotea	all services³
#bookmarks	10640	2028	940	13608
#unique users	145	2054	313	2437 ³
#tags	3152	27806	5772	36230

Table 4.2. Number of bookmarks and users for all three bookmarking services.

A comparison of the three bookmarking services leads to distinct results in the

³ Note: The difference in the sum of unique users between the services derives from the fact that 74 users are active in two services and one user in all three services.

case of bookmarked publications from the domain of physics (table 4.2). CiteULike has the most users, but BibSonomy users have bookmarked more articles. Of these users, a relatively small number added a huge amount of bookmarks. The ten most active users had 99-322 bookmarks, while 80 users had 30 or more bookmarks. This confirms former studies on collaborative systems, which show that a few active users are responsible for the majority of bookmarks (see the power law distribution in figure 4.1). This fact leads to sparsity and cold start problems, which might work against appropriate recommendations for less active users (Breese, Heckerman, & Kadie, 1998; Huang, Chen, & Zeng, 2004; Schein, Popescul, Ungar, & Pennock, 2002).

In the generated data set, users assigned 36,230 tags. Interestingly, the number of tags compared to the number of bookmarks in the three services differs immensely. BibSonomy, the service with the most bookmarks, has only 145 users, who assigned 3,152 tags. By contrast, CiteULike has only a small amount of bookmarks, but 2,054 users assigned 27,806 tags to them. Hence, the latter user community seems to be more tag-active. As tags are able to show topical relations, and those relations are applied in tag-based recommender systems in

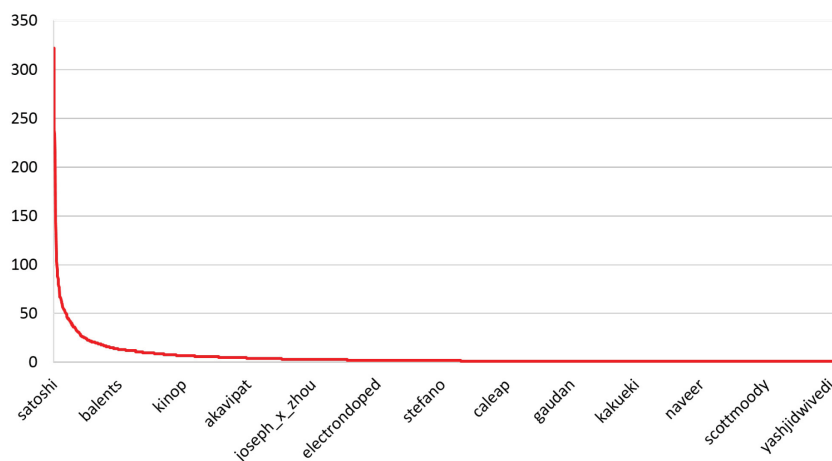


Figure 4.1. Power law distribution of bookmarks ($n=13,608$) (y-axis) per user (unique users: 2,437).

order to detect similar interests, the data distributions in CiteULike are more appropriate for measuring similarity in tag-based relations. However, only 1,949 users assigned tags to their bookmarks, while 1,579 bookmarks remained untagged.

Tag assignment is not subject to any restrictive rules, and some of the approaches discussed in chapter 3 contain methods of adjusting tags. For example, Juršič, Mozetič, Erjavec, and Lavrač (2010) propose an automatic approach toward learning lemmatization rules while measuring the similarity of words via the length of their suffixes. With regard to tag collection, such adjustments may improve user-item-tag relations in a folksonomy and thus lead to better recommendations. However, it must be noted that manipulated tag terms do not represent authentic user behavior anymore (Peters, 2009). Thus only minor adjustments were considered in order to exclude obvious tag “errors” and inappropriate tags. The original tag set includes 11,507 unique tags (not taking into consideration capitalization). In one case, tag adjustments leads to 9,661 unique tags (tag set one). Here the system considers all signs from [a-z,A-Z,0-9,-,_,], deletes all other signs and adjusts all tags to the lower case. In tag set two, 8,454 unique tags are derived by applying the following method: The system considers all signs from [a-z,A-Z,0-9,_,]. Hyphens are deleted because manual analysis shows that many tags are similar except for the presence or absence of hyphens. All signs are set to the lower case, and the underscore is replaced by a blank space (also on the basis of manual analysis). Additionally, Porter stemmer⁴ is applied, where both terms in two-term tags are considered for stemming. For both tag sets, fiugre system deletes typical stop tags, which are “import”, “imported”, “jabref” (upload from jabref reference manager) and “upload”. Bookmarking services assign these tags automatically when users upload or import bookmarking data. The differences in the number of unique tags are not too significant, but user-user relations would increase from more common tags, and the number of possible recommendations for each

⁴ <https://pypi.python.org/pypi/stemming/1.0>

target user would grow slightly.

However, earlier studies proved that relevance measurement is hardly affected at all by these adjustments. Peters, Schumann, Terliesner, and Stock (2011) analyzed tags from Del.icio.us⁵ to find out whether power tags (Peters, 2012; Peters & Stock, 2010) and Luhn tags (Luhn, 1958) enhance retrieval effectiveness. Analyses based on a retrieval test with test sets and expert relevance data showed that recall and precision are affected in different ways. The Usage of power tags improved precision for one-word queries. Moderately used tags, as favored by Luhn's thesis, did not lead to more relevant results. Instead, the combination of power and Luhn tags worked well for recall values, but not for precision values (i.e. for one-word queries). All analyses were conducted via three diverse tag sets: Original tag, unified tags with lower-case digits and without any special characters, and unified and stemmed tags. These adjustments proved inappropriate. Peters et al. (2011) conclude: "The unification and stemming of tags do not enhance precision of results."

Due to these findings and the numbers of tag sets generated, it was decided to analyze further measurements on the basis of tag set one, which considers the most common signs and deletes automatic tag assignments. Based on this data, initial user similarity measurements showed differences in the recommended ranking lists, concerning, on the one hand, variances between bookmarks and tags, and on the other hand, variances between the standard similarity metrics in information science.

4.1.3 Differences Between Similarity Models and Coefficients

Similarity proceeds on the assumption that two users behave in a similar way in a specific environment. In a social tagging system, the following assumptions can be applied (Balby Marinho et al., 2012):

⁵ <http://del.icio.us.com>: Service discontinued.

- If users bookmark common articles, they are interested in the same items. As a result, they are assumed to be similar and a relation between them is measured with regard to the bookmarks (referring to a unique article) both users have in their bookmarking list.
- If users use common tags, they are interested in the same topics. Hence, they are assumed to be similar and a relation between them is measured with regard to the tags both users have assigned to their bookmarks.

Users with only one bookmark were not considered for similarity measurement because they would highly distort the results. User-user pairs with one bookmark in common, where that bookmark is the only one both have assigned, would result in a similarity of 1, which means that both users are identical in their behavior. However, proving this assumption on only one bookmark is not appropriate. It would be important for a user recommender system to set a threshold: either a user should have a minimum of bookmarks, and/or a minimum of similar bookmarks to another user's, before this user is

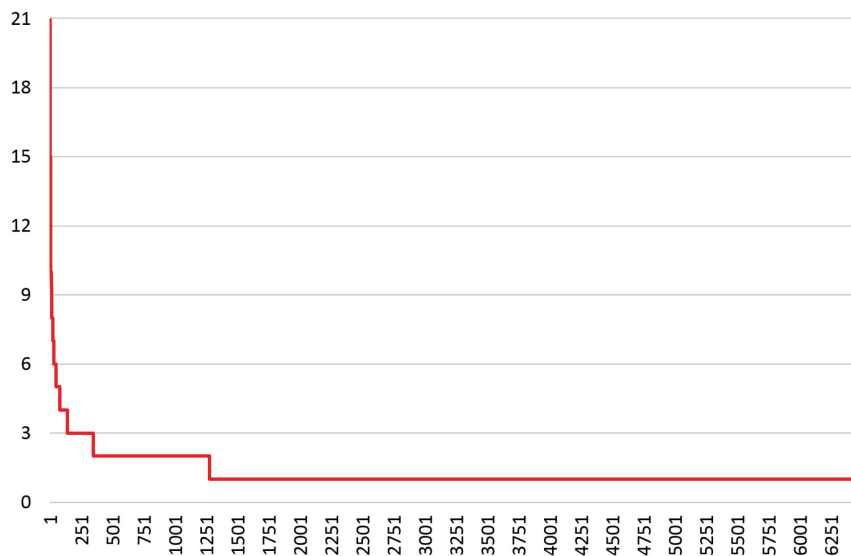


Figure 4.2. Distribution of common bookmarks (y-axis) for the 6,491 user-user pairs.

recommended to him or her. Another option is to let the user decide. For example, a user can adjust a slider to regulate which users, or what number of users, are recommended (see Knautz, Soubusta, & Stock, 2010 for resource recommendation and a system with slider functions). CiteULike has a minimum of 20 resources a user must have in his or her online library before resources are recommended (see chapter 3). Leaving out all users with fewer than two bookmarks, the data set showed 6,491 user-user pairs (897 unique users)⁶ who share at least one bookmark, instead of 11,007 pairs. Again, the number of shared bookmarks in the user-user pair data set shows a power law distribution in which only 12 user-user pairs share 10 or more bookmarks (figure 4.2).

For similarity analyses, we applied Jaccard-Sneath (Jaccard, 1901; Sneath, 1957), Dice (Dice, 1945) and the cosine coefficient (Salton, Wong, & Yang, 1975), the most common coefficients in information science (Stock & Stock, 2013, p. 116):

$$\begin{aligned} \text{sim}(a, b) &= \frac{2(|B_a \cap B_b|)}{|B_a| + |B_b|} \\ \text{sim}(a, b) &= \frac{|B_a \cap B_b|}{|B_a| + |B_b| - (|B_a \cap B_b|)} \\ \text{sim}(a, b) &= \frac{|B_a \cap B_b|}{\sqrt{|B_a| * |B_b|}} \end{aligned}$$

Equation 4.1. Similarity metrics Dice, Jaccard-Sneath, and cosine coefficient, where B_a , respectively B_b , is the number of bookmarks of user a , respectively b . The same goes for similarity based on common tags, where $|B_a|$ is replaced by $|T_a|$, respectively $|T_b|$.

Rasmussen (1992, p. 422) claims that “Dice, Jaccard and cosine coefficients have the attractions of simplicity and normalization and have often been used

⁶ Note: Unique users were determined as described above (with user names that occurred twice being merged), but user-user pairs may include double pairs relating to users who are featured in two or three services.

for document clustering.” According to Van Eck and Waltman (2008), a similarity measurement should fulfill two conditions:

1. The similarity between two users should be maximal if the “profiles differ by at most a multiplicative constant” (p.1654).
2. There should be no similarity if the authors have nothing in common, i.e. co-citations or, in our case, bookmarks or tags.

All three coefficients meet these conditions. The normalized metrics show values between 0 and 1, where 0 means that two objects are not similar at all and 1 that both have the highest similarity possible. Van Eck and Waltman (2009) claim that the most popular similarity measurements are association strength, cosine coefficient, the inclusion index and the Jaccard index. In the Pearson correlation (Pearson, 1895), which is also often used for similarity measurement, they identify some shortcomings (Van Eck & Waltman, 2008). Pearson labels objects as similar even though they do not have any features in common, and shows negative values for objects that are considered to be quite similar. To clarify, consider the following example by Van Eck and Waltman (2008): Two authors are compared, via the Pearson correlation, on the basis of co-citations they have in common with other authors (here $n=4$). If the profile of author 1 is [1 2 3 4] and that of author 2 is [10 20 30 40], both authors have a similarity value of 1 (totally similar) because author 1 has 1 co-citation with author 3 (respectively, author 2 has 10 with author 3), 2 co-citations with author 4 (respectively, author 2 has 20), 3 co-citations with author 5 (respectively, author 2 has 30), and 4 co-citations with author 6 (respectively, author 2 has 40). The total number of co-citations is not considered, but only the relative frequency with regard to the other authors. If the profile of author 1 is [11 12 13 14] and author 2's is [14 13 12 11], the resulting Pearson correlation values would be -1 (no similarity at all), although both authors have almost the same number of co-citations with the other four authors. If both authors have the profiles [10 1 0 0] and [0 0 1 10], but have no common author by whom they have been co-cited, the correlation value will still be -0.43. Thus, Van Eck and Waltman (2008) conclude that the Pearson correlation is not a satisfying

similarity measurement for their co-citation analysis. Ahlgren, Jarneving, and Rousseau (2003) also showed that the Pearson correlation, used for co-citation analysis, has shortcomings when expanding the data sample, even if only zero-vector values are added.

Similarity is based either on common bookmarks or on common tags. In both models, Dice and Jaccard-Sneath showed similar results, as both metrics are quite similar (Egghe, 2010). The latter generally shows minor similarity values. By contrast, the cosine coefficient leads to different ranking results while showing higher similarity values for user-user pairs, both for common bookmarks and common tags (compare also Hamers, Hemeryck, Herweyers, & Janssen, 1989 for an analysis of the cosine coefficient). Dice and Jaccard-Sneath tend to punish two users whose total numbers of bookmarks, respectively tags, differ largely (figure 4.3). Thus the question arises: which ranking of similar users best serves the target user and his or her needs? Consider the following example: Table 4.3 shows similar users for target user “dchen”, who has 212 bookmarks. According to Dice, the most similar users also have the most bookmarks in common with “dchen”. Hence, the most similar user “weeks”,

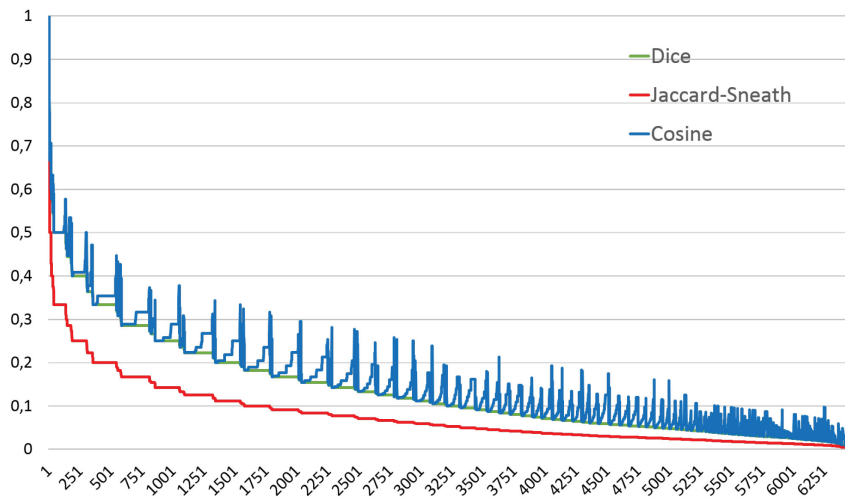


Figure 4.3. Distribution of similarity values for all 6,491 user-user pairs, ordered by Dice.

with 17 common bookmarks, is ranked first. Users “katiehumphry” and “rodney” both share five bookmarks with “dchen”. However, “katiehumphry” is on number five, whereas “rodney” comes in eighth place. For target user “dchen”, the relevance of the different rankings depends on his or her needs. If “dchen” were to search for relevant users to get more relevant literature, “rodney” would be better than “katiehumphry” because he has more bookmarks unknown to “dchen”, which offer potentially interesting literature.

However, in the case of expert search, “katiehumphry” would be a possible candidate because she shares a higher percentage of her bookmarks with “dchen” than “rodney” does. As user networks showing user-user relations based on common bookmarks or tags are undirected, it is more appropriate to lend a greater weight to the percentage of features shared by both the target user and the other. All three similarity metrics consider this aspect.

However, differences arise between Dice (or Jaccard-Sneath) and the cosine coefficient. The latter ranks users in a slightly different way. Applying the cosine coefficient, “rodney” is ranked 16th, while “jeevanjyoti” is listed higher. Both users may be potential partners, but Dice considers “rodney” better because of the five common bookmarks. The cosine coefficient takes into account the fact that “jeevanjyoti” only has 14 bookmarks in total. As two common bookmarks out of 14 are slightly better than five out of 103, “jeevanjyoti” has a better cosine value than “rodney”.

Another, more important aspect when comparing Dice and the cosine coefficient is the fact that the latter makes more obvious distinctions between user similarities, which means that cosine values have a greater range. Consider, for example, users “knordstr” and “peteryunker”. Both share two bookmarks with “dchen”, and “peteryunker” has three times more bookmarks than “knordstr”. With 0.0186 and 0.0181, the Dice values for both users only differ in the fourth decimal place. By contrast, the cosine values for both users are 0.0793 and 0.0458, respectively. In social bookmarking systems, where users generally have low numbers of shared features such as bookmarks and tags, more

distinguishable values between users are more helpful. In the data set at hand, the highest amount of common bookmarks is 21. Figure 4.3 shows the power law distribution of shared bookmarks by all 6,491 user-user pairs.

user a	user b	$ \mathbf{B}_a $	$ \mathbf{B}_b $	$ \mathbf{B}_a \cap \mathbf{B}_b $	Dice	Cosine
dchen	weeks	212	57	17	0.1264	0.1546
dchen	ghunter	212	57	16	0.1190	0.1456
dchen	kdesmond	212	50	10	0.0763	0.0971
dchen	kkims	212	66	7	0.0504	0.0592
dchen	katiehumphry	212	22	5	0.0427	0.0732
dchen	kedmond	212	25	5	0.0422	0.0687
dchen	tathabhata	212	15	4	0.0352	0.0709
dchen	rodney	212	103	5	0.0317	0.0338
dchen	waitonhill	212	8	3	0.0273	0.0728
dchen	caortiz	212	2	2	0.0187	0.0971
dchen	knordstr	212	3	2	0.0186	0.0793
dchen	peteryunker	212	9	2	0.0181	0.0458
dchen	jeevanjyoti	212	14	2	0.0177	0.0367
dchen	softsimu	212	88	2	0.0133	0.0146
dchen	lgolick	212	3	1	0.0093	0.0397
dchen	whitead	212	3	1	0.0093	0.0397
dchen	ccthomas	212	4	1	0.0093	0.0343
dchen	devries	212	5	1	0.0092	0.0307
dchen	governmentmen	212	5	1	0.0092	0.0307
dchen	LGNR	212	5	1	0.0092	0.0307
dchen	mazlans2	212	7	1	0.0091	0.0260
dchen	kopelman	212	10	1	0.0090	0.0217
dchen	kdaniel	212	14	1	0.0088	0.0184
dchen	dhbook	212	15	1	0.0088	0.0177
dchen	forgoston	212	15	1	0.0088	0.0177
dchen	kubyaddi	212	19	1	0.0087	0.0158
dchen	kaigrass	212	30	1	0.0083	0.0125
dchen	6rheology	212	42	1	0.0079	0.0106
dchen	sobolevnm	212	51	1	0.0076	0.0096
dchen	chiufanlee	212	60	1	0.0074	0.0089

Table 4.3. Similarity between target user “dchen” and 30 similar users, based on shared bookmarks, ordered by Dice coefficient.

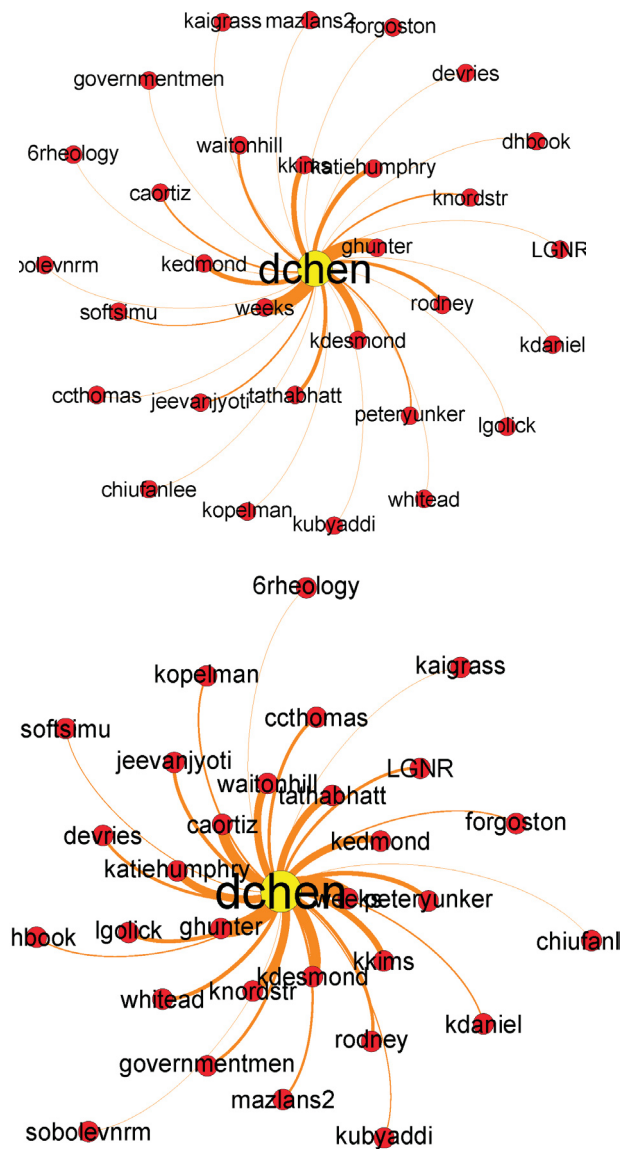


Figure 4.4. Similar users for “dchen” based on Dice (above) and on cosine coefficient (below). Graph generated with Gephi, algorithm: Force Atlas (see section 4.3).

21 users only have two or one bookmarks in common with “dchen”. The Dice values for these users range from 0.0187 to 0.0074 (with a difference of 0.0113), whereas the cosine coefficient values range from 0.0971 to 0.0089 (difference: 0.0882). More refined differences between user similarities might be more helpful for a target user. Especially in a visualized user network, more distinctive similarity values are desirable because they help a target user distinguish between similar users in a better way. With the cosine coefficient, visualized networks show distinctive user groups more clearly (figure 4.4).

Apart from similarity algorithms, there are differences regarding the basis for these measurements, which are common bookmarks and tags. User-user relations based on common tags lead to different recommendation results. Again, users with only one tag were deleted, which leads to 63,687 user-user pairs (1,543 unique users) with at least one common tag, instead of 66,307 pairs. This number is almost ten times higher than that of pairs based on common bookmarks. This is also reflected in the distribution of common tags for user-

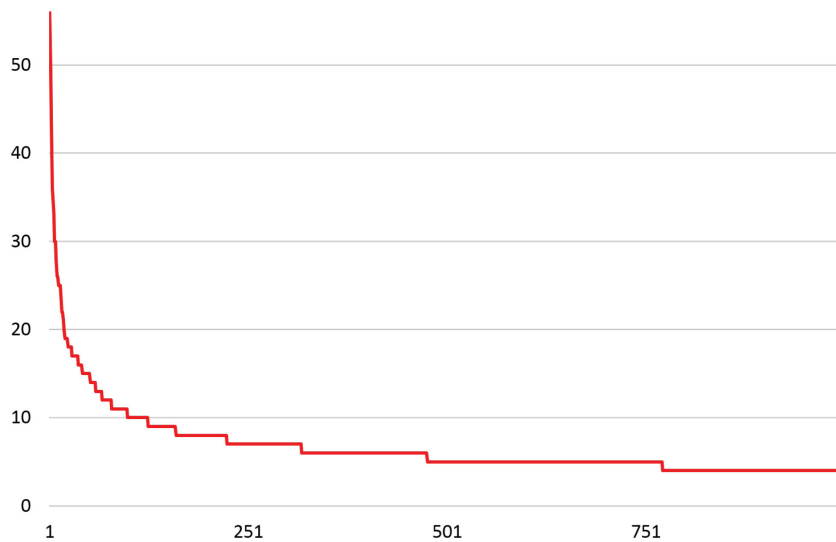


Figure 4.5. Distribution of common tags for the first 1,000 user-user pairs.

user pairs (figure 4.5), which still shows a power law, but the number of common tags is higher than the number of common bookmarks.

user a	user b	$ \mathbf{T}_a $	$ \mathbf{T}_b $	$ \mathbf{T}_a \cap \mathbf{T}_b $	Dice	Cosine
dchen	weeks	181	62	25	0,2058	0,2360
dchen	ghunter	181	69	21	0,1680	0,1879
dchen	kedmond	181	30	17	0,1611	0,2307
dchen	kkims	181	35	16	0,1481	0,2010
dchen	rodney	181	263	30	0,1351	0,1375
dchen	andreab	181	110	19	0,1306	0,1347
dchen	michaelbusmann	181	579	49	0,1289	0,1514
dchen	paulschlesinger	181	109	17	0,1172	0,1210
dchen	chiufanlee	181	113	15	0,1020	0,1049
dchen	bronckobuster	181	187	17	0,0924	0,0924
dchen	cguido	181	15	9	0,0918	0,1727
dchen	kdesmond	181	42	9	0,0807	0,1032
dchen	ddahlem	181	33	8	0,0748	0,1035
dchen	pbuczek	181	63	9	0,0738	0,0843
dchen	jaeseung	181	39	8	0,0727	0,0952
dchen	barrat	181	44	8	0,0711	0,0896
dchen	kaigrass	181	16	7	0,0711	0,1301
dchen	gdurin	181	78	9	0,0695	0,0757
dchen	Tomste	181	21	7	0,0693	0,1135
dchen	nurban	181	28	7	0,0670	0,0983
dchen	hendysh	181	29	7	0,0667	0,0966
dchen	kevina	181	120	10	0,0664	0,0679
dchen	andreapuglisi	181	31	7	0,0660	0,0934
dchen	andrewsun	181	33	7	0,0654	0,0906
dchen	6rheology	181	45	7	0,0619	0,0776
dchen	CDivin	181	50	7	0,0606	0,0736
dchen	georgwachter	181	86	8	0,0599	0,0641
dchen	kristgy	181	95	8	0,0580	0,0610
dchen	l-alex	181	77	7	0,0543	0,0593
dchen	itmeson	181	78	7	0,0541	0,0589

Table 4.4. Similarity between target user “dchen” and the 30 most similar users, based on common tags, ordered by Dice.

Whereas there are 30 users with common bookmarks for user “dchen”, he or she

shares tags with 639 users, of which 154 share more than two tags. Table 4.4 shows the 30 most similar users with tag co-occurrences, where users who are not found via common bookmarks have been marked. If more user-user relations are found, the number of recommendations may rise. However, as table 4.4 shows, users who bookmark a lot are also potential candidates for tagging. The first five users were found via common bookmarks and tags, all five having quite high similarity values with “dchen”. In fact, the Pearson correlation value between the number of bookmarks and the number of tags per user is 0.689 (with 2,035 users sharing at least one bookmark or tag with another user). However, there is a lower correlation between the number of common bookmarks and common tags of user-user pairs. Thus a user-user pair with many common bookmarks does not generally share many common tags. The Pearson correlation value is only 0.383 (between all user-user pairs that share at least one bookmark) and 0.255 (between all user-user pairs that share at least one common tag). Figure 4.6 shows the distribution of common bookmarks

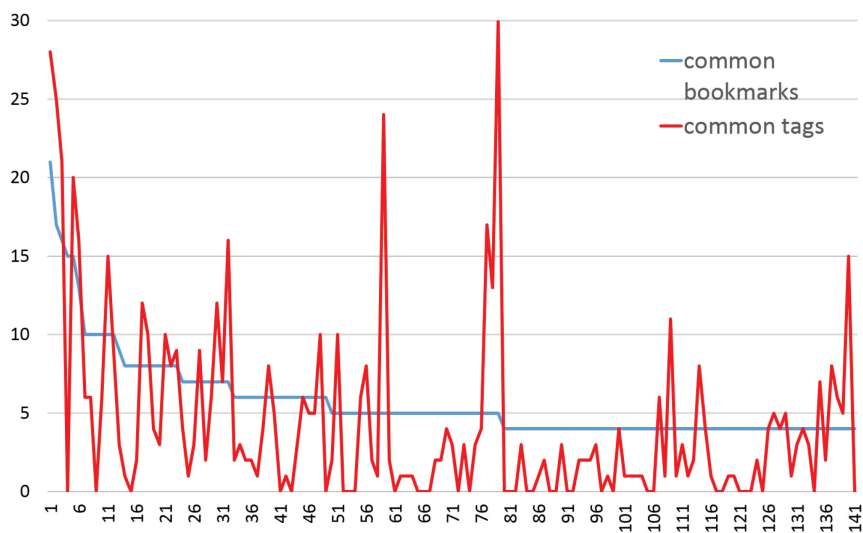


Figure 4.6. Distribution of common bookmarks and common tags for all the top 141 user-user pairs who have the most bookmarks in common.

compared to the number of common tags for the 141 user-user pairs that share at least 4 bookmarks.

Concerning the numbers of shared tags and their similarity values, which are higher than the values based on common bookmarks, common tags seem more promising than common bookmarks as a basis of generating a data set for user similarity measurement and recommendations. There are more diverse users who would not be found via shared bookmarks. Additionally, tags are able to inform a target user in which context other users assign their bookmarks (Peters, 2009; Szomszor et al., 2007). If “dchen” searches for project partners, the tags can give him or her an impression of a bookmarked article’s content. This provides a glimpse of other users’ research field. The precondition is that a recommender system must show the tags assigned by the recommended users in an adaptation of the idea of multiple features recommendation (see figure 3.6 in chapter 3 by Balby Marinho et al. (2012)).

To sum up the findings, we gain two important insights for further models regarding the two research questions about data structure and similarity metrics. Concerning the three standard similarity metrics, the cosine coefficient has some advantages over Dice and Jaccard-Sneath. On the one hand, the degree of shared features (bookmarks and tags) between users is considered in a more efficient way. User-user pairs are not punished for having unequal numbers of bookmarks or tags. On the other hand, the cosine coefficient is more precise in terms of clarifying user-user similarity, which supports data structure in social bookmarking systems, especially for visualization purposes. Furthermore, analysis of the folksonomy-based data structures showed that similarity based on common tags leads to more results, i.e. to more user-user relations. The number of common tags is also higher than the number of common bookmarks. Therefore, data set generation for recommendation purposes should be based on common tags rather than on common bookmarks. These findings are considered for expert recommendation in section 4.3. Prior to this study, a user evaluation involving researchers provides first insights into the value of a collaborator

recommender as well as the relevance of the results.

4.2 Qualitative Evaluation of Recommendations for Researchers

For initial studies on expert recommendation, it was decided to propose a simple model and to evaluate the results via qualitative interviews with researchers. Hence, direct user feedbacks should answer the question: Does social researcher information from bookmarking services lead to relevant recommendation results concerning relevant literature and users? Direct user feedback leads to first insights into the appropriateness of recommendation results and is able to show more concrete user concerns by “real” researchers (compare discussions in section 4.6, for example in Ma, Pant, and Liu Sheng (2007)).

For this case study, three scientists working at *Forschungszentrum Jülich*

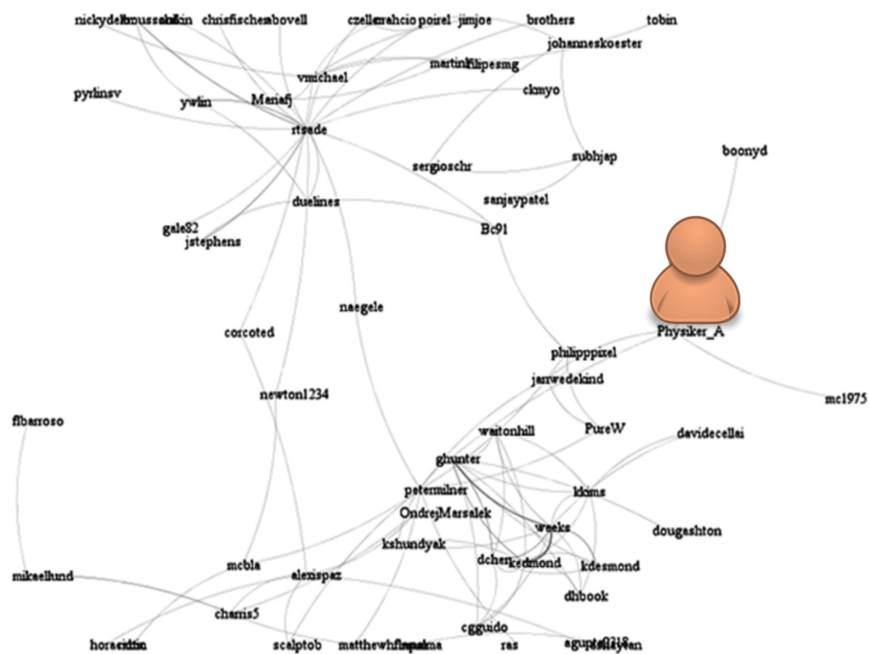


Figure 4.7. Single-link cluster for target researcher *a*, with a cosine coefficient threshold of <math><0.1</math>.

participated in the evaluation. The same scientists also took part in the survey about social bookmarking systems and communities of practice (see chapter 1). The first recommendation model aimed to gather a researcher's bookmarks from a bookmarking service and to determine their similarity to other users. As no participating target researcher used any of the considered bookmarking systems, the model was modified. A "fictive user" served as a target researcher for whom recommendations were to be made. The fictive user's profile represents a target researcher who bookmarked their own articles. The researcher's articles taken into consideration are those which were bookmarked by real users of a service. Hence, all bookmarked articles that were published by a target researcher are added to the fictive user's profile.

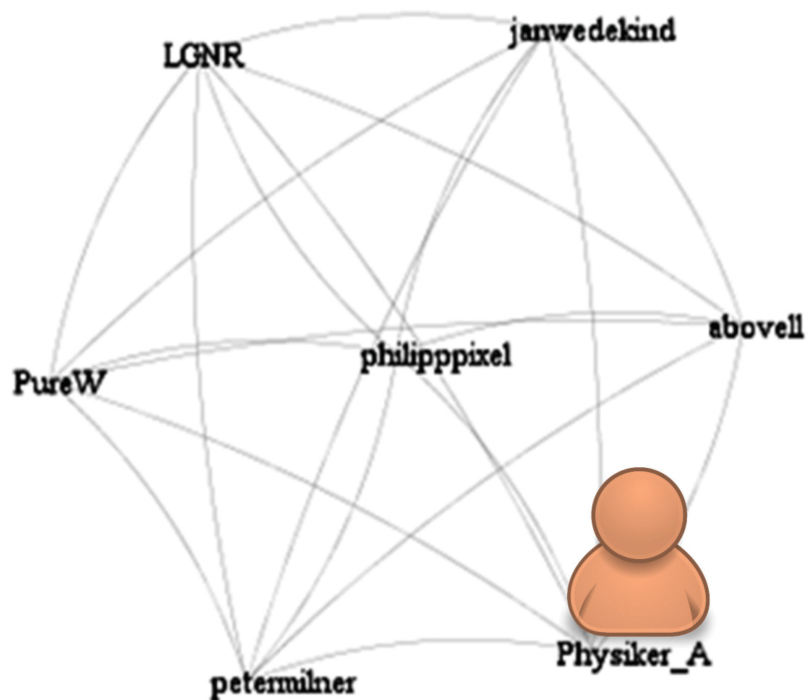


Figure 4.8. Complete-link cluster of target researcher *a*, with a cosine coefficient threshold of <0.1 .

4.2.1 Data Set and Recommendation Network Model

The bookmarking service CiteULike was chosen, for the reason that its coverage of bookmarks related to the field of physics is appropriate (Reher & Haustein, 2010). Bookmarking data from the target researchers' publications were gathered via the CiteULike website, again using DOI, ISSN, as well as author and title information to obtain the right resources. For all bookmarks, user information was gathered using CiteULike article IDs. Similarity measurement via the cosine coefficient was based on common bookmarks (equation 4.1). Tag similarity was not given preference in this case as a researcher's publications were defined as the bookmarks on the basis of which the data set was generated. The researcher himself did not assign any tags because he did not use the bookmarking service. Tags for his bookmarked articles were only available from other CiteULike users (this model is applied in the next study in section 4.3). An option for modeling tags for non-active or fictive users is to take, for example, title terms or keywords from bookmarked publications. The title terms are then defined as "fictive user tags". Landia et al. (2012) propose such a model to extend the basic model of the FolkRank (Hotho, Jäschke, Schmitz, & Stumme, 2006).

Similarity between a target researcher and CiteULike users based on common bookmarks is visualized as simple user networks showing user-user relations. Because the approach aims at detecting researcher networks, the best way of evaluating the approach was by showing the participants visualized graphs. As discussed in chapter 3, cluster methods are one way of finding appropriate recommendations. To generate expert networks in this case, single-link (nearest neighbor) and complete-link (farthest neighbor) clusters were generated (Gemmell, Shepitsen, Mobasher, & Burke, 2008; Knautz et al., 2010). The size of the clusters was regulated by cosine coefficient threshold. All users whose similarity lay above this threshold were featured in the single-link cluster, with the starting point being the target researcher. In the complete-link clusters, all users had to be related to each other and needed to have a greater similarity than

the set threshold. To give an example, the two generated clusters for a target researcher a are shown in figures 4.7 and 4.8, where the size of the edges refers to the similarity between the users based on the cosine coefficient. The search for researcher a 's publications in CiteULike resulted in 142 bookmarks posted by 197 users. The relatively small size of the data set for each scientist caused a low cosine coefficient threshold being set, at < 0.1 .

4.2.2 Evaluation and Discussion

The evaluation conducted in a semi-structured interview aimed to answer the following questions:

1. Are the representations of user networks showing relations between oneself (the target researcher) and other persons helpful? Would they be helpful for finding relevant partners?
2. How valuable are these recommendations, which are based on CiteULike users and their bookmarks?

The three target researchers stated that visualized networks are sufficient for user recommendation. They rated the usefulness of such networks for finding potential research partners positively. However, the scientists tended to prefer small clusters as recommended networks, as the complete-link clusters – which show a small part of the associated single-link clusters – were seen to be arranged more clearly. The interviewed researchers considered these smaller networks to be sufficient for an expert recommender system. The reasons for this preference were, on the one hand, that the respective numbers of recommended users and their bookmarked literature was more manageable and reduced information overload. On the other hand, a researcher can only collaborate effectively with a rather small number of colleagues. Hence, the interviewed physicists did not prefer recommendations of huge partner networks.

Besides the clusters, the target researchers evaluated the similar users' bookmarked publications that they had not known before. The resources

appeared reasonable and appropriate at first sight. The scientists were surprised to find interesting resources, which, they assumed, would not have been retrieved via their regular literature search. All three physicists said that they almost exclusively search for literature on the webpages of their preferred journals. In other words, they always limit their search to preferred sources. The participants stated that this was the easiest way of filtering for relevant literature needed for research. However, limiting searches to a few chosen journals causes the scientists to miss out on interesting publications by other sources. The researchers stated that a recommendation system for finding these publications would be helpful. This system would be a source in addition to their already-used sources and preferred search strategies.

Another advantage of a running expert recommender system applying folksonomy-based data is that the existing user-item-tag relations allow direct reference to any of these nodes. When a target researcher gets expert recommendations, existing user-item links directly lead him or her to resources of potential relevant literature; or, on the other hand, item-tag relations lead to tags describing resource content. Such a system may offer multiple recommendations of either users, resources, or tags (Bogers, 2009). Depending on which model is applied, systems can easily switch between diverse recommendations and consider up-to-date user needs. For example, Hotho et al. (2006) propose a model in which nodes on a graph-based approach are represented by all three different entities. Hence, a recommendation system is an appropriate tool for the establishment of academic communities of practice because not only are community members related, but so is their shared repertoire – that is to say, resources referring to a community's practice and tags representing a community's shared language.

To sum up the findings, using social information in bookmarking services to recommend potential researchers as well as their work leads to appropriate results. Through the user community, new relations are detected that show recommendations beyond a researcher's traditional network and search

behavior. However, a running recommendation system should offer more information about recommended authors that a target user can use to evaluate the recommendation. As an author's reputation and scientific work is defined by their publications, this information is relevant for a target user's decision of which author has the potential to be a future partner. Two problems arise in this study. Firstly, all users in a social bookmarking system are mostly anonymous. Some users choose their real name as their user name, and a system might be able to link a user name to a real author name, but not without severe difficulties concerning, for example, ambiguous names. Here, further approaches must be applied in order to obtain appropriate results (Demaine, 2011; Wooding, Wilcox-Jay, Lewison, & Grant, 2006). Nevertheless, most users have nicknames, which makes it impossible for a user to find out an author's real name. The only option for a target user is to look at another user's profile, which must then include information about this user's real name and their own publications. Recently, new services like ResearcherID⁷ and Orcid⁸ have appeared, in which scientific authors are assigned a unique identifying ID and are able to edit their profile with information about their research area and links to their publications. What is needed for a collaborative filtering recommender system are tags or bookmarked articles for each user. In order to use bibliometric data, a link to a citation database is needed. For example, Web of Science and ResearcherID are connected. To date, Thomson Reuters mainly uses researcher IDs to disambiguate author names. A closer cooperation could include an exchange of data on both sides. Then it would be possible to have a target scientist defined by ResearcherID and recommend articles or researchers to them on the basis of their publications, citations and co-citations gleaned from Web of Science. The advantage of researcher services is their generation of unique author IDs, which solves the problem of author name ambiguities.

Another challenge besides the problem of anonymous users is that a target user

⁷ <http://www.researcherid.com/>

⁸ <http://orcid.org/>

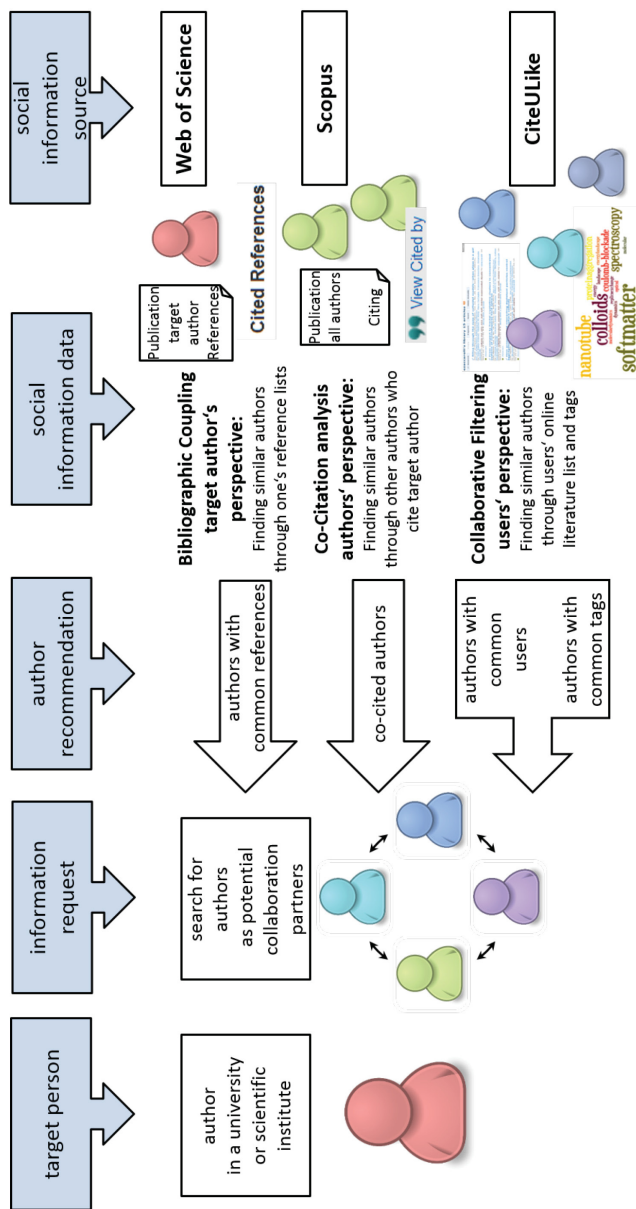
needs to be active in a bookmarking system in order to get any recommendations. The number of his or her bookmarks and tags strongly influences recommendations in item- or user-based collaborative filtering systems. Less active users encounter cold start problems in the recommendation process (Herlocker, Konstan, Terveen, & Riedl, 2004). Some solutions for cold start problems in recommender systems exist, but then further user data must be collected, for instance social information from diverse networks (Sedhain, Sanner, Braziunas, Xie, & Christensen, 2014). Both these shortcomings are overcome, at least to a great extent, within the final approach proposed for expert recommendation. By considering the Web user's perspective, any problems with inactive researchers and anonymous users can be solved.

4.3 Recommending Authors Using Multiple Kinds of Social Information

The main case study (parts are published in (Heck, 2012a, 2012b, 2012c, 2013; Heck et al., 2011)) addresses the primary research questions of this work and proposes the final model for expert recommendation. Its purpose is to detect implicit relations between researchers and to make scientists aware of them so as to foster areas of community and knowledge creation. To consider implicit new relations, it is crucial to take into account different perspectives with regard to the academic environment. Most researchers are aware of their explicit relations to other colleagues (see the survey in chapter 1). Hence, a recommender system should focus on new and unknown or hidden relations, also known as implicit relations. Implicit relations are based on diverse kinds of social information about researchers derived from different perspectives. The following approach considers three perspectives (third-party, target researcher, web user) introduced in chapter 2, and detects implicit relations on that basis.

The proposed model is shown in figure 4.9. The target user is a scientist who is meant to be made aware of colleagues belonging to the same academic community who could be potential collaborators for any academic purpose (information request). A recommender system is not interested in showing the user a network within their research field, but a network based on their personal

Figure 4.9. Concept of author recommendation based on different sources of social information.



social information. Thus, in contrast to most expert portals, a recommender system focuses on personalized recommendations. Social information is generated by people representing different perspectives on scientific work and reputation. This information is found in diverse sources. Expert recommendations are generated by applying bibliometric and altmetric measurements.

4.3.1 Data Set for Implicit Relations Based on Bibliometric Data

As the focus of the approach lies on implicit author-author relations, two methods based on co-citation and bibliographic coupling are applied. Both methods are aggregated from the document level to the author level. Thus author co-citation refers to two authors who are both cited by a third author. Bibliographic coupling refers to two authors who both cite the same references in their publications. For both methods, a database with citation data is required. The two largest scientific information services are Scopus⁹ and Web of Science. The approach first aims to use one of the two services. Both Scopus and Web of Science offer search functionalities for references and co-citations. However, neither the database in Scopus nor that in Web of Science offer sufficient options for generating both data sets for co-citation and bibliographic coupling analysis (see chapter 2).

On the basis of availability and technical feasibility, therefore, it was decided to draw on the usage levels of both information databases. One shortcoming here is that publications differ between both services, i.e. one target author may have four articles in Scopus, but three different ones in Web of Science. This issue must be noted as similarity measurements and recommendations depend on the generated data. Table 4.5 shows a summary of the target author's publications and their respective occurrences in Scopus and Web of Science. After analyzing both services, the following data was generated for co-citation and bibliographic coupling analyses in the case study.

In Scopus, all documents that cite at least one of a target author's articles were

searched. All references featured in these documents are needed to find authors who are co-cited with a target researcher. The concrete search was mainly done manually and involved the following steps:

1. Search for all publications of a target author via title and author search fields. An up-to-date publication list was requested from the target author.
2. Search the citing resources for all publications of the target author via the link button “cited by”.
3. Export all citing resources containing resource metadata, including all references, and save export in a csv file (figure 4.10).
4. Import data from csv file to database, determine author names per citing resource, and generate Excel file with author names and number of shared citing resources.
5. Manual correction: unify all unique author names from step 4 and sum up citing resources. Correct database entries.
6. Take authors with most common resources (number of authors taken depends on target author) and search their number of “times cited” in Scopus author profiles.
7. Add ‘times cited’ information to database. Format author names.

Scopus/Web of Science data	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10
# pub 2006-2011/2012	10	24	8	16	5	20	13	10	10	16
# pub in Scopus	10	21	8	16	5	16	12	7	9	13
# pub in Web of Science	8	16	7	15	5	17	12	7	10	12
# citing documents in Scopus	33	65	63	76	15	103	95	71	60	39
# citations in Scopus data set (csv)	32	65	58	64	15	94	95	64	35	36

Table 4.5. Basis for generated data sets: numbers of publications from 2006-2011/2012 per target author, numbers of their surrogates found in Scopus and Web of Science, numbers of citing resources in Scopus, and numbers of citations derived via Scopus csv files.

As author ambiguity is a great concern, the number of citations per author in

Scopus was gathered manually in step 6. This puts the emphasis on quality above quantity. As a result, not all co-cited authors were considered for analysis, and the number was limited. However, most of the similar authors were identified and it can be assumed that authors in the long tail, who only have a few citations in common with a target author, are less relevant. Those authors would have a negligible cosine value and would not be ranked high up in the recommendation list for a target author. However, with regard to author networks represented in a graph, which is discussed further on, the missing authors would change those network structures.

Author limitation was conducted as follows: initially, authors were ranked

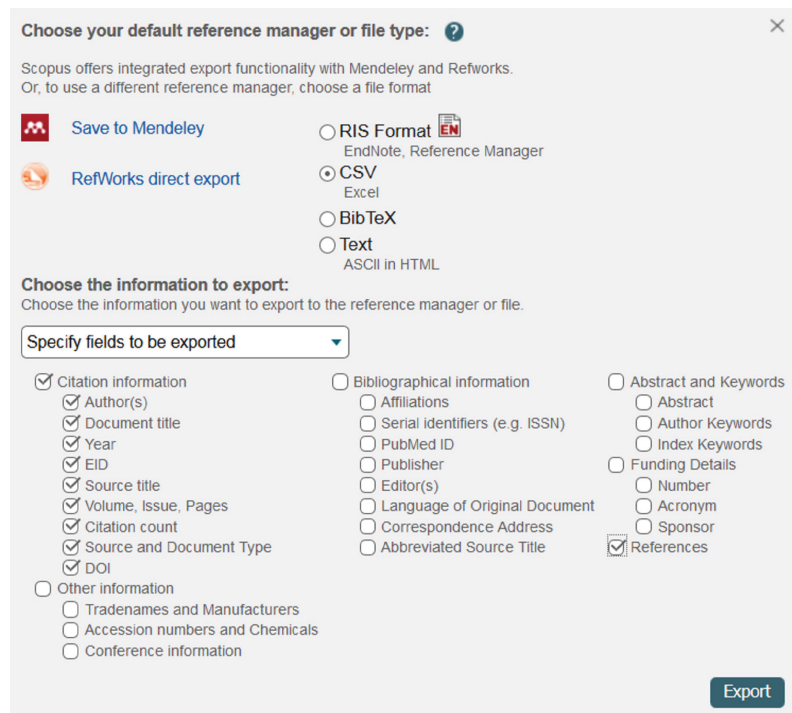


Figure 4.10. Scopus export of citing resources, including references that cite at least one of the target author’s publications. www.scopus.com

according to their total number of co-citations with a target researcher. In the next step, those authors with a minimum number of x co-citations were taken, where x determined the cut-off point of the ranking. This number of co-citations may vary for each target scientist, but it will be at least $x \geq 3$. Using this procedure, different numbers of authors are considered for similarity measurement for each target researcher. The minimum of authors considered per target researcher is 34, the maximum is 69.

In the best case, authors are identified via the Scopus unique author ID, which was used in our approach. Nevertheless, it must be noted that author details in Scopus are not always correct. If incorrect data was detected, for example when Scopus included two author IDs belonging to the same author, these were corrected manually. Where two identical author names were found, the correct unique author was identified via their co-authors and the document titles collected from the csv file.

In Web of Science, a search was conducted for all related documents that share at least n references with any of the publications of a target author, where n may vary for each target scientist. The authors of these documents were considered for bibliographic coupling analysis. These were the concrete steps taken to generate the data set:

1. Search for all publications of a target author via the title and author search fields, again using the author's current publication list.
2. For each individual publication, determine the list of "related records" via the Web of Science link button.
3. All publications that have n shared references are considered. Export those publications in a csv file and add the number of 'shared references' to the individual publications.
4. Aggregate lists from step three – one list per publication by target author –and determine unique authors via author names.
5. Take authors with most shared references (number of authors considered depends on target author) and search their publications manually in Web of Science.
6. Count all references from the publications in step 5 and insert number into database. Format author names.

To improve the quality of the data set, authors with references in common were searched manually in Web of Science, as was also done for Scopus. First and foremost, this is designed to address challenges posed by author names. All authors from the publication list exported in step 3 were considered. Author disambiguation and the merging of unique authors was performed manually. As there is no author ID in Web of Science, author names were verified by checking their document list and, if necessary, correct it on the basis of the articles' subject area (compare Persson (2001)). The number of shared references per unique authors was then summed up (step 4). This led to a list of authors with a minimum number of n common references, where n determined the cut-off point of the ranking. This number of common references may vary for each target scientist, where in all cases $n \geq 4$. By applying this procedure, different numbers of authors were considered for similarity measurement for each of the target researchers, where the minimum was 22 authors and the maximum was 53 authors.

In terms of editing the author lists, both for co-citation and bibliographic coupling analysis, two options were possible. In both cases the number of citations of similar authors ($|C_b|$), respectively the number of references ($|R_b|$) – each of which is needed for similarity measurement – had to be determined from the full database. If only citations, respectively references, of authors from the downloaded data sets had been used, cosine values would have been biased toward greater similarity with the target author. Therefore, $|C_b|$ and $|R_b|$ were retrieved manually, which led to a limitation of the number of authors taken into consideration. Two options were possible: either an equal number of authors could be considered for each target author, or author lists could be ranked according to the number of co-citations or common references, respectively, and cut at n , where n is the amount of instances of both similarity features. The latter option was applied, as it seemed more logical. If the ranking lists had an equal number of ranks, for example the first 20 authors with the most shared references, then the lists would be cut between two rankings, each with the same amount of instances of a common feature. For example, author *Maier* has 6

common references, which puts him at position 20 of the Web of Science list for a target author. Author *Schneider* also has 6 common references with the target author, but he ranked 21st and thus will not be considered for further recommendation measurements. One critical aspect is the way such a list is arranged: if *Schneider* were instead called *Köhler*, *Köhler* would come before *Maier* in an alphabetical ranking. This would lead to a final choice, which does not depend on the similarity feature “common references” or “co-citations”, but on ranking position and author names. Hence, the choice was made to cut the lists at n common similarity features. It follows that a different number of authors were selected for recommendation for each target author. Additionally, the variable n was not the same for each target scientist because numbers of co-citations and common references varied strongly. Depending on the researcher’s scientific work, i.e. their number of publications over the last five years, the number of references and citations is higher or lower and influences their similarity features with other authors. Thus n was chosen according to a target author’s characteristic numbers and relative to the determination of relatively similar numbers of possible recommendations for all authors. In sum, this means that n for all researchers meant at least three co-citations and four common references, where the maximum of n varied.

Deficiencies in Scopus and Web of Science Data

A recommendation system highly depends upon the source data set. Various problems arise when generating data in Scopus and Web of Science. The difficulty of identifying unique authors via author names has already been addressed. In this study, authors could best be identified via unique author IDs in Scopus and via ResearcherIDs in Web of Science. Nevertheless, author details in Scopus are not always correct. The service may run into problems when two or more authors with the same name are allocated to the same research field and change their workplace several times. Unfortunately, only a few potential collaborators in Web of Science had a ResearcherID. Hence, the

greatest difficulties concerning author identification arose in Web of Science. Although author search is available and suggests unique authors, including information about their affiliation and research field, it is difficult to make distinctions between two authors with the same name. Asian author names in particular lead to problems. Thus reference data (not common references, as these were gathered via explicit lists of “shared references”, including titles and co-authors) for these authors might include mistakes.

In Scopus, as mentioned above, a unique resource may appear in several formats in other resource reference lists. Thus some references are incomplete, which also concerns author names. For this study, data generated via Scopus was incomplete. Some authors are not included in the csv files that formed the basis for co-citation analysis. Those authors, then, could not be considered for co-citation similarity measurement. Table 4.5 shows the data generated via Scopus. The number of citing articles per target author could easily be generated. These articles (including their references) define the data set, thus a target author should have at least as many citations as there are articles that cite them⁹. However, citations per author are missing (see table 4.5 “citations in Scopus data set”). The completeness of reference data in Scopus varies considerably. In a random sample, where test data from the website was adjusted manually so as to exclude general errors in the Scopus database, 5 of 14 authors have complete coverage, 3 have a coverage between 70% and 90%, 5 between 55% and 70%, and one author is only covered at about 33%. The Scopus database includes the correct number of author citations (it must be assumed to be correct because it cannot be proven so), but the file including exactly the same data shows different citation values. One explanation may be that Scopus generates its references directly from a resource, adopting any citation style. This would also explain the phenomenon of different reference strings for the same resource. If a

⁹ Note that for cosine similarity measurement, one citation is counted per resource regardless of whether a resource might have cited the same document twice.

citation style cuts author names in a publication by multiple authors, these authors might not be found in the reference lists (at least not in the export list). Concerning the target researchers participating in the approach, the differences in citation numbers are not too high, except for authors $a3$ and $a9$, where the difference is higher than 10. As the number of lost citations of the authors to be compared could not be determined, wrong citation numbers were not corrected. All author citation numbers were derived from the exported data in step three (except the numbers of $|C_b|$).

4.3.2 Data Set for Implicit Relations Based on User-Generated Data

The goal of this case study was to provide scientific target authors with recommendations of potential partners, with the similarities based on the Web users' perspective. The importance of this perspective is discussed in chapter 2. The Web user perspective also solves two recommender problems. Firstly, the problem of anonymous users in a bookmarking system is solved because a target researcher will not be given user recommendations but author recommendations. Web users only generate relations between the authors of bookmarked articles. Secondly, the cold start problem of an inactive researcher is solved. A researcher does not have to use a bookmarking system to get recommendations based on its users. Of course, if there are no users who bookmark a researcher's articles, cold start issues do exist. But in this study, where the participants were physicists from Jülich who do not use bookmarking systems, the approach does not suffer from their inactivity.

In contrast to other collaborative filtering approaches, the following method does not recommend any users, resources, or tags, but authors from publications that are bookmarked in CiteULike. It is thus not CiteULike users themselves who are the target users, as they only deliver social information about the "external" target researchers while they bookmark their publications and assign tags to them. This model is also different to the "fictive user" model in section 4.2 and leads to different results. It considers the perspective of all CiteULike users. In the previous model, collaborative user opinion is not considered to the

same extent as in the following approach. In section 4.2, similar users to a fictive user are those who share a lot of common bookmarks with the former. Thus a user needs to bookmark many articles by the target researcher in order for their own bookmarks to be recommended. The collaborative opinion of other users is not considered unless other users also bookmark a researcher's articles. In the following model, recommendations are based on common users, meaning that if a researcher's bookmarked article shares many users with another article, the authors of both articles are considered to be similar. In this case, author similarity is determined by all users, as their number is counted and weighed. However, to generate the data set, common tags are crucial.

The basis for the following steps is the folksonomy of CiteULike, defined as a tuple $F = (U, T, R, Y)$, where U , T and R are finite sets consisting of the elements 'user name', 'tag' and 'resource', and Y is a ternary relation between them: $Y \subseteq U \times T \times R$ with the elements being 'tag actions' and 'assignments' (Balby Marinho et al., 2011; Peters, 2009). To use this information for the following study, the folksonomy is expanded to $FE = (U, T, R, A, Y)$, where A is added as the finite set with the element 'authors' and $Y \subseteq U \times T \times R \times A$ is their relation. Two options for setting the database for author similarity measurement are possible:

1. Searching for all users $u \in U$ who have at least one article $s \in R$ by target author $a \in A$ in their CiteULike bookmark list.
2. Searching for all resources $r \in R$ who have tags $t \in T$ in common with one bookmarked article $s \in R$ by target author $a \in A$.

The disadvantage of option one in this case lies in the small number of users who bookmark a resource compared to the number of tags, which is generally higher. Relying only on users may not be enough to identify similarity (Lee & Brusilovsky, 2010). Data analysis in the study in section 4.2 showed that the number of tags is quite high compared to the numbers of bookmarks and users, and the number of common tags is higher than the number of common bookmarks. Additionally, the data included many users who did not share

common bookmarks or tags with others. Out of 2,437 users, only 897 shared at least two bookmarks, and only 1,543 users shared at least two tags. All other users without these relations are not valuable for similarity measurement because in the following approach this would lead to similarity based on only one user (namely the user without common features). Such a user would have bookmarked a researcher's article and other publications, but all these publications would not have any other users. Otherwise, this user would share common bookmarks with any other person.

However, it must be noted that applying the second option might lead to users who have bookmarked a resource by a target author but not tagged it might get lost and not be in the data set. Table 4.6 shows the bookmarks of the target authors' publications found in CiteULike as well as the number of tags assigned to them (all tags and unique tags) as well as the number of users (all users and number of unique users) who bookmarked them. As the number of tags is sufficiently higher, the second option was applied to generate a data set for further similarity measurement. This means that resources (here: scientific papers by a target researcher) are deemed to be similar if common tags have been assigned to them. This further leads to the assumption that the authors of these resources are also similar. As tags point to topical relations, authors connected via such relations regarding their research fields can be potential collaboration partners. Additionally, the more tags are shared by two resources, the more similar they are. In some cases, target author resources were labeled with very general tags, such as "nanotube" and "spectroscopy". Therefore, a minimum of 2 unique tags was set: a resource must have two tags in common with at least one of a target author's resources in order to be considered in the data set. This data set forms the basis for further measurements distinguishing two collaborative filtering models. Again, to assure better data quality, one subset from CiteULike was derived and author name corrections performed manually. This step was important, especially for CiteULike data, because the service contains diverse spellings of the same names.

The options chosen lead to the following steps for generating the data set in CiteULike:

1. Request database dump from CiteULike (for each part of the study the current dump was used).
2. Find publications by target author via search fields “DOI” (if DOI of publication was available), “author”, “title”. Determine all CiteULike bookmark IDs for each publication retrieved.
3. Search for bookmark IDs within database dump and determine all tags assigned to IDs.
4. Get all bookmarked resources that have at least two shared tags with a target author resource. Determine the resources with at least n common tags, where the number of tags depends on a target author.
5. For all resources determined in step 4, derive all authors from the CiteULike website via the unique CiteULike bookmark ID (example: www.citeulike.org/article/3182948) (because author information was not included in database dump).
6. Prove and unify unique author names and correct database entries. Format author names.

CiteULike data	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10
# pub found	5	13	2	1	5	1	6	4	2	10
# pub bookmarked by users	4	11	2	0	3	1	6	1	1	4
# tags assigned	17	73	11	0	12	3	25	0	4	12
# unique tags assigned	15	45	11	0	13	3	21	0	4	12
# users who bookmarked	5	17	4	0	4	1	8	1	1	4
# unique users who bookmarked	5	6	4	0	4	1	5	1	1	4

Table 4.6. Basis for generated data set in CiteULike: number of publications found, number of bookmarked publications, number of tags (all tags and unique tags) assigned to target authors’ publications, number of users (all users and unique users) who bookmarked.

The number of tags a resource should have in common with any target researcher’s publication was adapted via methods similar to those already discussed for the Scopus and Web of Science data sets. The aim was to have

nearly equally large data sets for all researchers in order to make recommendations comparable. Furthermore, scientists claimed to prefer rather small author networks (section 4.2). Hence, when a target researcher's publications shared two or more common tags with many other publications, this list was cut at the number of n tags. In three cases, 2 common tags were considered, in four cases 3 and in one case 4.

Despite these settings, exceptions were made in generating the CiteULike data set. One of the 10 authors (*a4*) did not have any users who bookmarked his articles in CiteULike. Some resources were found but they had been adjusted to the system by the CiteULike operators (authenticated articles), and not by any user. Thus no CiteULike data set could be generated for this scientist. Another researcher's resources were bookmarked but not tagged. In this case, option two for data set generation could not be applied. As the researcher had one bookmarked resource by a single user – an article published in 2000 – this bookmark was considered, even despite the general publication date restriction. In the latter case, the data set was not generated via tags but via resources bookmarked by the same single user. This means that a kind of user-based collaborative filtering was applied for this researcher (*a8*). All authors who were represented at least six times in the single user's bookmark list were considered. Author similarity in this case was measured via the total number of author occurrences. Applying the cosine coefficient would not make sense here as the common feature would only be one single user. Furthermore, researcher *a10* was given different restrictions regarding the publications considered in CiteULike. As only one bookmarked resource was found referring to an article from 2007, the search was expanded to bookmarked resources published earlier than 2006. Ten additional publications were found, and three of them were bookmarked. In the end, the four bookmarked resources were considered.

Generating the data sets is a critical precondition for recommendation purposes. Data completeness, limitations, and thresholds of values and data correctness influence further measurements and results. The focus in the current approach

lies on the quality of the data, which meant that many steps were processed manually. A running recommender system should therefore apply further techniques, aiming for example to lower author ambiguity or incorrectness in resource data. The challenges in a running system will not be discussed further at this point, as the current model concentrates on the first results of the applied models and their evaluation. Picault, Ribière, Bonnefoy, and Mercerm (2011) discuss a possible model for establishing a running recommender system and indicate critical aspects to be considered.

4.3.3 Recommendation Models and Evaluation Method

Models and Similarity Algorithm

There are many different similarity algorithms and models for generating recommendations. Their application is based on recommender system types such as content-based recommenders or collaborative filtering recommenders (Adomavicius & Tuzhilin, 2005). Some studies show that simpler measurement techniques are not always the worst (see for example Linden, Smith, & York, 2003). Jannach, Zanker, Felfernig, and Friedrich (2011) point out that more complex techniques are currently in favor (pp. 47-49). However, practical examples show that complex algorithmic models are probably not required, as seen in the example of Amazon, which uses a simple item-based collaborative filtering technique. Furthermore, as the focus of this approach shows, research in model combinations is promising: “[...] recommendation algorithms that exploit additional information about items or users and combine different techniques can achieve significantly better recommendation results than purely collaborative filtering approaches can” (Jannach et al., 2011, p. 48). Researchers refer to new hybrid models (see chapter 3) or fusion models (Tso-Sutter, Marinho, & Schmidt-Thieme, 2008; Wang, de Vries, & Reinders, 2006). The model described in this study also questions whether combined social information from different sources leads to better recommendations. However,

different models are not combined in a fusion approach. A similarity algorithm applies the cosine coefficient, which is an established standard in recommender system research (Jannach et al., 2011, p. 19) and is also common in information science research (Stock & Stock, 2013, p. 115).

As showed in the previous chapter, the cosine coefficient best reflects recommendation results based on bookmarking data. Furthermore, it is more appropriate for visualizing user networks. Similarity determined via the cosine coefficient directly reflects any of the three proposed principles, which are author co-citation, bibliographic coupling and collaborative filtering. As the data sets themselves were generated according to these principles and data was filtered on the basis of these assumptions, further filtering methods were not applied. For evaluation purposes, the k -nearest neighbors based on the cosine values are considered. The cosine coefficient (equation 4.1) was applied for all data sets and similarity features, inserting $|C_a|$ as the number of citations author a received, $|R_a|$ as the number of all references author a used, $|T_a|$ as the number of tags assigned to bookmarks with a as the author, and $|U_a|$ as the number of users who bookmarked resources where a is the author. Respectively, the same was applied for author b .

The applied collaborative filtering models should be distinguished from the classic ones described in chapter 3. Firstly, no user from a bookmarking service is considered as the target user or recommended to another target user in this case. The focus lies on authors of bookmarked publications. Secondly, collaborative filtering models mostly pre-process data to determine the similarity of items or users. In a further step, these similarity measures are applied to a specific case, i.e. a target user and their specific preferences. Because the target researchers in this study do not have preferences in CiteULike, due to the fact that they are not users of the service, collaborative filtering is adapted to the current model. Assumptions about Web user similarities are aggregated to authors of bookmarked articles. Compared to collaborative filtering in tagging systems, the core of the similarity assumptions

is the same:

- 1a) Collaborative filtering, user-based: users (nearest neighbors) who share “common items” or “common tags” are similar and appropriate for recommendation of further resources or tags. Likewise, items with many “common users” or “common tags” are similar (Balby Marinho et al., 2012).
- 2a) Adapted model, user-based: authors who share many “common users” are similar to each other and thus potential collaboration partners.

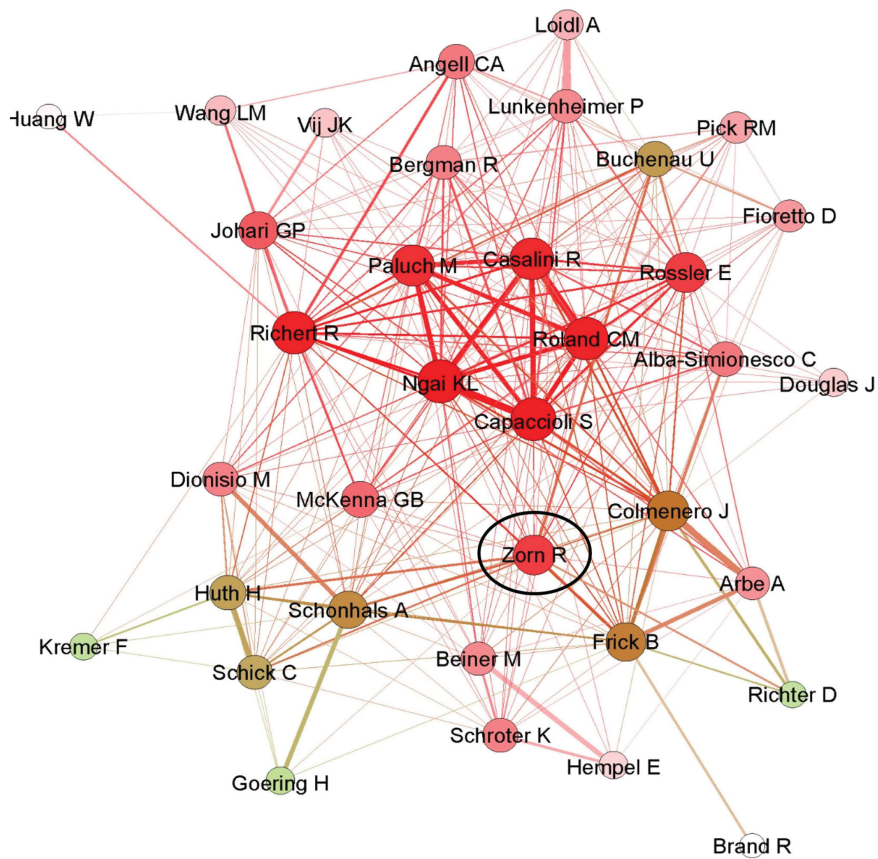


Figure 4.11. Author network based on bibliographic coupling (BICO) in Web of Science for scientist “R. Zorn” (a10), cosine coefficient threshold 0.2.

- 1b) Collaborative filtering, tag-based: Items and users that share “common tags” are similar and appropriate for recommendation of further items or tags (Duro & Dolog, 2009; Hung, Huang, Hsu, & Wu, 2008).
- 2b) Adapted model, tag-based: Authors whose bookmarked publications share “common tags” are similar and potential collaboration partners.

Author Networks and Visualization Graphs

Using the three generated data sets, and applying the proposed four similarity models, each target scientist has four different networks showing potentially

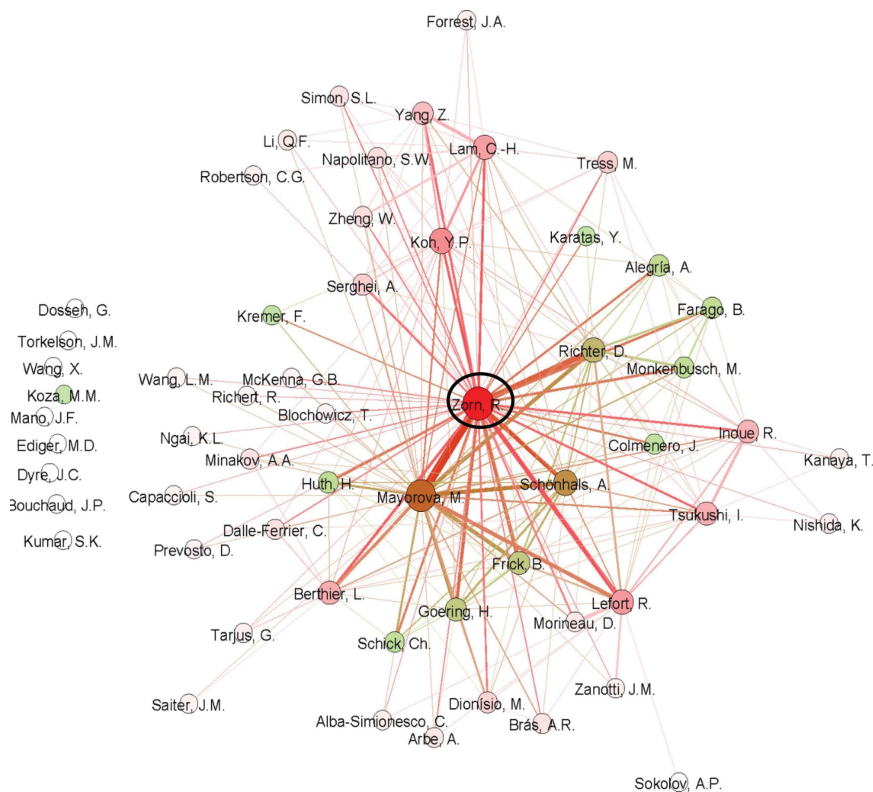


Figure 4.12. Author network based on author co-citation (ACC) network of author “Zorn”, cosine coefficient threshold 0.05.

similar authors. One network is based on bibliographic coupling (BICO network) generated from Web of Science, one is based on author co-citation (ACC network) generated from Scopus, one is based on user-based collaborative filtering in CiteULike (CULU-network), and the fourth is based on tag-based collaborative filtering in CiteULike (CULT-network). In all models, author similarity is measured via the cosine coefficient. The most similar authors are considered for evaluation. Additionally, cosine values for all author-author pairs are measured for the four similarity approaches in order to create visualized graphs of those author networks¹⁰. No clustering methods (as in 4.2) were used.

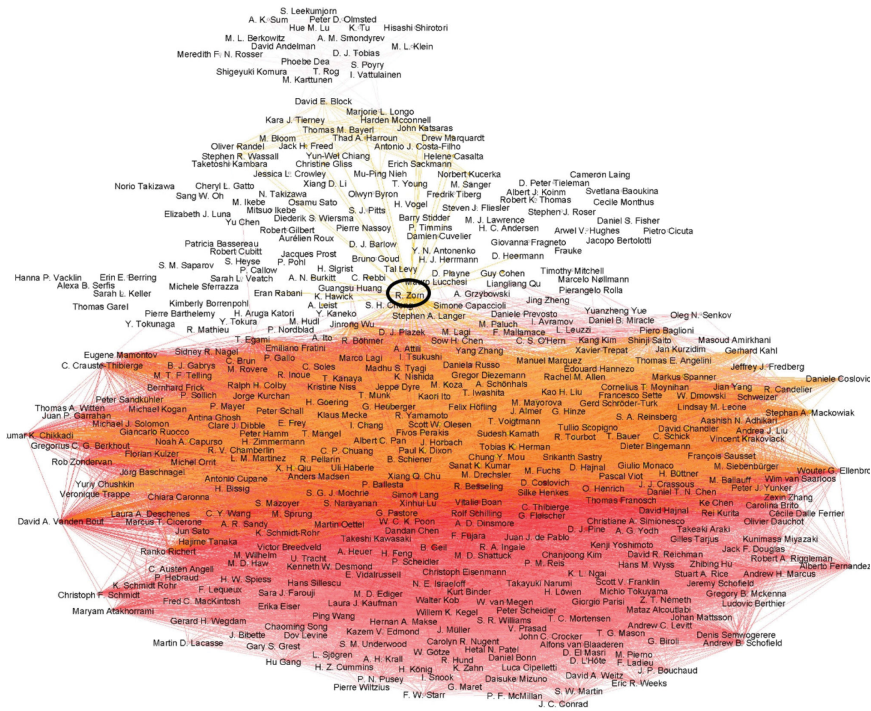


Figure 4.13. Author network based on CiteULike common users (CULU) for researcher “Zorn”, cosine coefficient threshold 0.49-0.99.

Single-linkage and complete linkage clustering would have led to very small clusters as no relation data – based on common features (like co-citations) for all authors except the target authors – was collected from the databases in Scopus and Web of Science. Figures 4.11-4.14 show the four author networks that were created for researcher “R. Zorn” (*a10*) based on his publications¹¹. The software Gephi¹² (Bastian, Heymann, & Jacomy, 2009) was used to create the graphs.

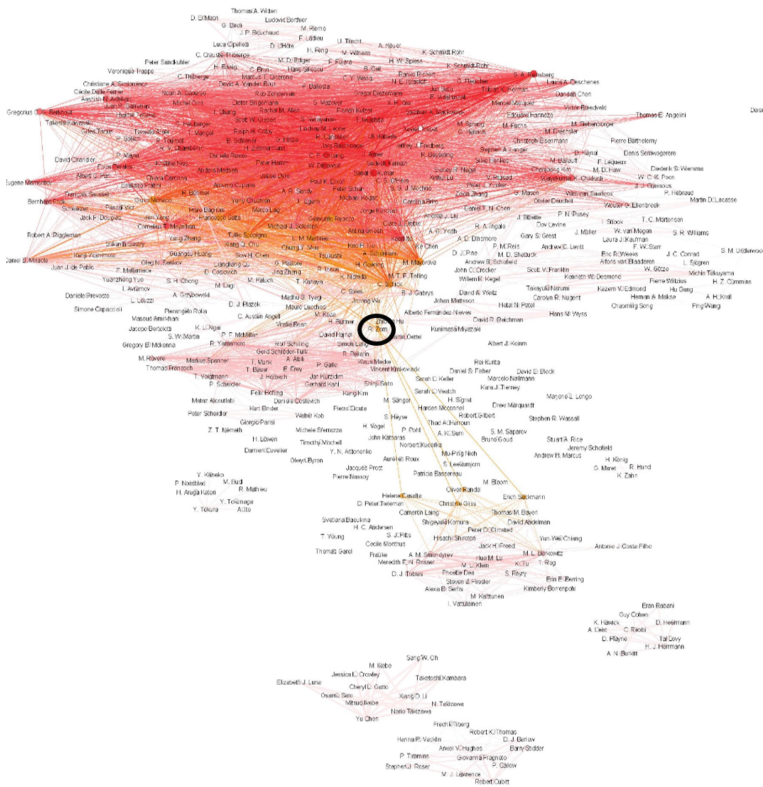


Figure 4.14. Author network based on common tags in CiteULike (CULT) for researcher “R. Zorn” (*a10*), cosine coefficient threshold 0.49-0.99.

¹¹ The networks of all other target researchers are available on request.

¹² www.gephi.org

The network visualization software offers multiple functions to analyze and configure networks. Author networks were created using the algorithm “Force Atlas”¹³. The force-based algorithm (Tutte, 1963) focuses on showing complementary nodes. Repulsion strength and gravity (both influence network size) were adapted for each network in order to create clear graph arrangements for the evaluation process, as each graph was printed on a DIN A3 sheet of paper. This includes limitations for some networks where the graph was restricted to edges with a specific cosine coefficient threshold. Most of these cases concerned the larger CiteULike networks. As the entire CiteULike database was available for use, author limitations – as was the case with Web of Science and Scopus’ data sets when limiting the list of authors considered for similarity measurement – were not needed. Thus all CiteULike networks include many more authors than ACC or BICO networks. In most cases, a minor cosine coefficient threshold was set. This makes the network smaller and author names readable on a DIN A3 sheet. Additionally, for CiteULike networks all edges with similarity values of 1 were left out because analyses showed that most of these author-author relations are based on only one common feature (bookmark or tag). If both authors only have one bookmark or tag each and they share it, the similarity between these persons is 1, which leads to biased results equal to those discussed in the pre-studies (section 4.1). Besides adaptations designed to make the network and its author names printable and readable, all graphs were left unmodified for evaluation.

Evaluation Method

The evaluation consisted of three parts: a) questions concerning a researcher’s work and information management behavior (results in chapter 1), b) evaluation of recommended authors, c) evaluation of the visualized author networks. The survey was conducted in semi-structured interviews with the target scientists,

¹³ <http://de.slideshare.net/gephi/gephi-tutorial-layouts>

each of whom were guided through the questions and received assistance if any parts of the evaluation were unclear. Thus the evaluators were able to ask questions concerning the correct meaning of individual parts of the evaluation. The conversation with each interviewee was noted and is considered for the analysis of the results. Each interview took about two to three hours.

The focus of the evaluation lies on the researcher's perceived usefulness (McNee, 2006; Stock & Stock, 2013, pp. 485–486). The main question is whether a recommended author is relevant for the evaluator's current research and if he or she could be a good collaborator. One pre-condition here is that the target researcher must know the recommended author. In the case of scientific researchers, this means that the evaluator knows an author's reputation and work in their scientific community and thus is able to estimate the author's relevance. This aspect means that a target researcher can only make statements about known recommended authors. Consequently, if a target researcher is unable to evaluate an author because he or she does not know them, this does not mean that this author is not relevant, as was discussed in the prior chapter (see also figure 3.7). In this study, there are two main reasons for collecting relevance feedback for known authors and not for unknown authors, as was done for unknown users in the pre-study (section 4.2). Relevance feedback for unknown users based on their bookmarked articles, respectively of unknown authors based on their publications, is quite difficult. An evaluating person cannot immediately decide whether a person would be a good collaborator as long as he or she has no concrete information about this person. An author's publications are first pointers to their scientific reputation. However, evaluators in the pre-study only indicated a tendency regarding their estimated relevance; clear statements about author relevance were preferred. This relevance can only be given for already known authors. Thus we adapted the evaluation model of De Meo, Nocera, Quattrone, Rosaci, and Ursino (2009) and De Meo, Nocera, Rosaci, and Ursino (2011). Evaluators are asked to state which of the recommended authors they know and which of them are relevant.

#	Questions for evaluation part b): top recommended authors
1.	Do you know this author? Please mark <y> for know author name, <n> for do not know author or <p> for know author personally.
2.	If you know the author: Have you ever collaborated with him or her?
3.	Do you think the known authors are similar to you with regard to your current research?
4.	Please state the relevance of each known author regarding your current research needs on a scale from 1 (not relevant) to 10 (highly relevant).
5.	Would you collaborate with any of these authors? Please mark <y> for yes or <n> for no.
6.	Please state reasons for your non-collaboration decisions.
7.	Do you miss any relevant author in the recommendation list? Please write down all relevant author names.

Table 4.7. Questions concerning the top recommended authors in evaluation part b) (original questions in German).

Even with known authors, it is sometimes difficult to give relevance feedback, as indicated during the evaluation process. Some interviewees had difficulties answering question #5 (table 4.7), stating that they do not know the considered author well enough to clearly estimate their current relevance. A further issue was that some participants knew that the considered author's research focus had changed, leaving them unsure whether this new focus fits their current needs. Such vague cases were noted and left out of consideration for the subsequent results. Another reason for considering the evaluation of known authors is that these results allow for conclusions about the relevance of the conducted models

and their data sets.

The evaluation gives the number of relevant authors who are potential collaboration partners. With this list, clear statements can be made concerning the appropriateness of the recommendations and a precision value can be measured. Furthermore, these results are not based on historic user data, but on direct user feedback (see Ma et al. (2007) for a statement on this issue).

#	Questions for evaluation part c): author networks
1.	According to your personal valuation, does the author distribution accurately reflect the reality of author collaboration in the research community? Please state any peculiarities.
2.	Do the network graphs show researcher/author communities in the right way? Please compare the networks according to your personal valuation.
3.	Would the network recommending similar authors be of use to you, for example in organizing a workshop or finding collaboration partners?
4.	Relating to question 3: how relevant do you consider the shown networks? Please rate on a scale from 1 (not relevant) to 10 (highly relevant) and explain your decision.

Table 4.8. Questions concerning the visualized author graphs in evaluation part c) (original questions in German).

For feedback on recommended authors, the most similar authors (author names) from all four applied models were taken and ranked in alphabetical order. Co-authors identified via a researcher's publication list were not considered as it was assumed that co-authors are automatically relevant and do not need additional relevance feedback by the target scientist. The number of authors to be recommended was limited in order to keep the interview time acceptable for the participants. For the first part of the study, in 2011, the top ten authors in

each network (ACC, BICO, CULU and CULT) based on cosine values were considered. As co-authors were deleted, the number of recommended authors differs for each target scientist. Proceeding in 2012, co-authors in each set were “replaced” by non-co-authors until the list featured 40 recommended authors. Two exceptions were made: in the case where no tag information was available and the data set from CiteULike was compiled on the basis of a single user (*a8*, see section 4.3.2), all 26 derived authors were considered. As 20 authors would thus be taken for the CiteULike data set (as is the case for all other target researchers that have ten authors from CULU and ten from CULT), the latter six below the cut off point were also considered because the number of author co-occurrences was the same. In another case (*a9*), the author list gathered via the CULU network contained similar users with a cosine value of 1. As the first 20 authors had the same similarity, all these persons were put on the recommendation list. The cosine values for each recommended author were not shown, as they were likely to influence an interviewee’s decision regarding the relevance of the recommended authors. Table 4.7 shows the evaluation questions concerning this recommended author list. The questions were asked successively for each author on the list, except question 7, which was answered at the end. The rating scale from 1 to 10 was chosen to give the evaluators a chance to make more granulated distinctions between author relevance. In general, seven point rating scales are common.

Evaluation part c) was concerned with the visualized author networks. The networks needed to be examined according to the correct distribution of authors and their relations to each other. Therefore, all networks were printed on a DIN A3 sheet of paper. Table 4.8 shows the questions regarding these aspects. Each network was first shown separately, in the following order: ACC, BICO, CULU, CULT. The target researchers were able to mark the networks on the printouts. For the ratings, the interviewees saw all four printed networks and were able to compare them.

4.3.4 Results and Discussion I: Relevance of Recommended Authors

Before discussing the results of part b) in detail, we need to address some aspects with regard to the evaluation method. As already discussed above, some target researchers had difficulties judging some authors' appropriateness as collaborators. Similarly, it was not easy for some interviewees to rate recommended colleagues. Some researchers did make clear distinctions, using almost all rating values from 1 to 10, while other scientists preferred a binary distinction, rating an author as either relevant or irrelevant. In the latter case, the line separating relevance and irrelevance was drawn at the rating score of 5. Asked about the meaning of rating score 5, the researchers stated that this score is positive, while 4 means that an author is not very relevant. The evaluation process showed that a binary rating seems more appropriate for person ratings, at least to some researchers. The decision to consider a person relevant (independently of the purposes for which the person is evaluated) is quite complex, as shown by the interviews. Distinguishing the relevance of colleagues and assigning a kind of rating to them might feel wrong to some people. In that case, a simple binary rating would likely be preferable. Assigning proper weight to the target researchers' rating behavior, all recommended authors rated 5 or higher are seen as relevant authors.

The evaluation leads to the following author groups that can be distinguished: unknown authors, relevant known authors, and irrelevant known authors. This allows us to make a statement about the novelty factor of the approach with regard to the recommendation list. In almost all cases (except for three target researchers), the number of unknown authors is higher than the number of known authors. Counting the number for all interviewees, they knew 212 recommended authors, while 165 recommended authors were unknown to the target researchers. Experienced scientists (based on their active research years, for example *a8* and *a10*) already have a huge collaboration network because they know many authors personally. Figure 4.15 shows these numbers for each researcher, distinguishing between known relevant and known non-relevant

collaborators who were asked for in question #5 (excluding relevant authors determined via question #7, table 4.7). The average novelty factor based on all researchers is 0.46, where novelty is defined by the simple relation of unknown recommended authors to all recommended authors. In a slight deviation from this definition, De Meo et al. (2011) considered all users deemed reliable and divided them into known (user-user contact within a network) and unknown. Their evaluators appear to make statements about user reliability even for unknown individuals. Baeza-Yates and Ribeiro-Neto (2011) proposes an equal novelty ratio. However, in the case at hand, data about relevant unknown researchers is not available with regard to the aspects discussed above.

The relatively high number of unknown authors seems to be an appropriate result for recommendations of new potential collaborators. However, researchers are skeptical concerning the meaning of novelty (or serendipity) values, as it is difficult to design metrics for measuring novelty and serendipity

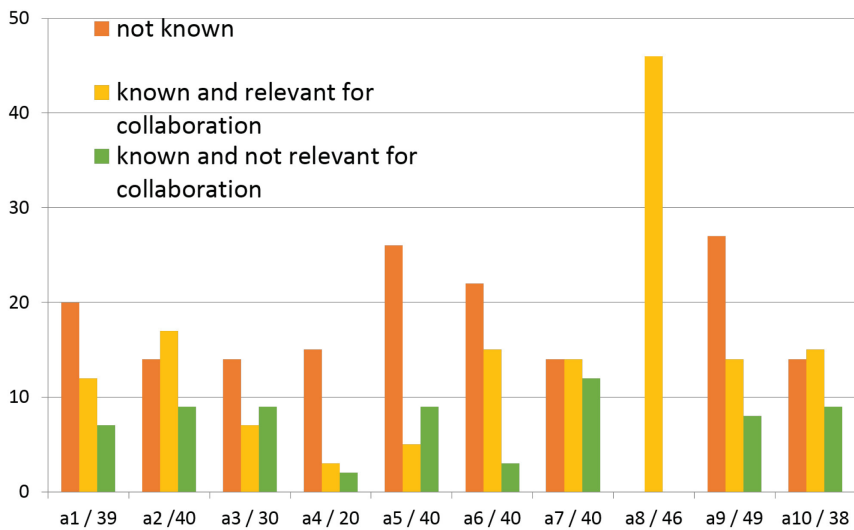


Figure 4.15. Number of recommended authors and their valuation of each target researcher, where the number shown along the x-axis legend is the total number of recommended authors per target researcher.

(Herlocker et al., 2004). Both aspects involve finding not only new items, but those which users would not have found without the help of a recommendation system (Lops, Gemmis, & Semeraro, 2011). However, in this scenario the novelty factor allows us to measure the respective proportions of new (unknown) and known people, which is important for the “credible recommender task”.

The participating physicists stated that they are interested in new collaboration partners, but would prefer, at the very least, people whose scientific work and community affiliation they know. Personal familiarity has a great influence on the decision of whether a person is interesting as a collaboration partner or not. This aspect was expected, as discussed above. Therefore, a network only showing unknown authors would recommend new authors, but would not be helpful to the scientists. Furthermore, the presence of known individuals creates trust. A system recommending known users or items fails to fulfill its task of detecting new resources, but familiar resources “increase user confidence” (Herlocker et al., 2004, p. 42), as is shown in diverse studies (Sinha & Swearingen, 2001, 2002). While showing participants their author recommendation lists, the researchers explained a known author’s relevance with reference to his or her research focus and ties or position within the community. This means that research colleagues are arranged into research groups and sub-groups within a scientific community, an arrangement that represents one factor of evaluating an author’s relevance. Hence, it is assumed that identified relevant authors can serve as a reference for estimating the relevance of unknown authors. This aspect becomes clearer when we analyze the results for the visualized networks in the next section.

Concerning questions five (“Would you collaborate with any of these authors?”) and six (what are the reasons for non-collaboration?”), the results are distinctive. 138 of the 162 recommended relevant authors are appropriate for potential collaboration. For those not deemed appropriate, diverse reasons were stated. In most cases, it was either a case of low topical overlap between the researcher’s

and the recommended author's work, or the recommended author was a competitor rather than a collaborator. In the first case, the interviewees regarded an author as relevant because his or her research results were important for their own studies, but only a small part of this research was required. To name one example, a target researcher who conducts experimental studies only referred to theorists for theoretical background and would rather collaborate with another practical physicist than with a theorist. Additionally, he knew that some colleagues doing theoretical work are themselves not interested in practical research. However, it must be noted that this is only one example. Other researchers need colleagues who complement their own skills. The case study shows that such people can be found by using social information from various sources.

Regarding the other reason for non-collaboration – where an author is seen as a competitor – the target scientists stated that some institutional conventions sometimes limit the potential for collaboration. However, this fact did not always include a negative facet from the target researchers' point of view. They could not (or would not) give clear reasons for non-collaboration with other institutes. Some guessed that these situations just developed naturally over time. As the survey in chapter 1 showed, collaborations often develop through the network of one's superior or colleagues. Thus in some cases, there is simply no tangible reason for non-collaboration. However, the five recommended authors who are seen as collaborators are highly relevant (one received a rating of 7, the others 10). An expert recommendation system does not focus exclusively on showing potential partners, but is also there to provide insight into the scientific community. Furthermore, collaboration within a community of practice is diverse and a community does not only consist of exactly like-minded people (remember Wenger's (2008) characterization of diversity as a fruitful factor in mutual engagement). For example, if a researcher is planning a workshop, it is not devious to invite research competitors to gain new insights for one's own research field.

Besides these justifications for non-collaboration, the researchers knew that some recommended authors had already retired or changed their research field. In only two cases, scientists named personal biases as their reason for non-collaboration. Two further scientists saw two, respectively three, recommended authors as competitors with whom they would not collaborate closely.

In addition to the relevant authors who were rated at five and higher in question #5 (table 4.8), the interviewees were able to name relevant authors who were not on the recommendation list (question #7 in table 4.8). These authors are also considered as relevant for the following analyses. This evaluation method is comparable with those of Berendsen, De Rijke, Balog, Bogers, and Van den Bosch (2013), who analyzed self-selected and system-generated data for evaluation. In their study they looked at recommended expertise generated by the system (such as the author recommendation list applied in this study) as well as expertise added by the users themselves. Here the target researchers also added additional expertise information concerning colleagues. In many cases, co-authors were named. Counting these numbers, the target authors named between 18 and 55 people they regarded as relevant for their current scientific work. The term “current” is important and was pointed out to the interviewees. For example, one target author had changed his research field. Some known recommended authors would have been relevant collaborators in the past, but were not appropriate for current research. Hence, a recommendation system considering an author’s last publications might suggest persons who are not currently relevant to the target researcher. A scientist could help correct this flaw in the choices he or she makes regarding those of his or her publications that a recommender system should consider for personalized expert suggestions. In other words, a researcher needs to deposit a publication list in their profile. Such a system could apply knowledge-based models in order to adapt the recommendations according to a user’s current needs (Burke, 2000). Here, aspects concerning user decision-making are important for building a system that satisfies user needs and is accepted by the users (Felfernig, Isak, & Russ, 2006; Mandl, Felfernig, Teppan, & Schubert, 2011).

The general distribution of relevant authors relative to the data sets derived from the three services is shown in figure 4.16. All three data sets yielded almost the same number of relevant authors, but the overlap between them is low, as only 35 authors are found across all three services. This means that different relevant scientists are found using the diverse methods of gathering social information data in Web of Science, Scopus, and CiteULike, respectively. Furthermore, it is interesting to take a look at the important authors who were only found via data from CiteULike: For example, 7 out of 35 important authors for target author *a2* are only found in CiteULike, and 6 out of 29 important authors for *a3*. Thus a system can recommend more relevant authors by analyzing different social

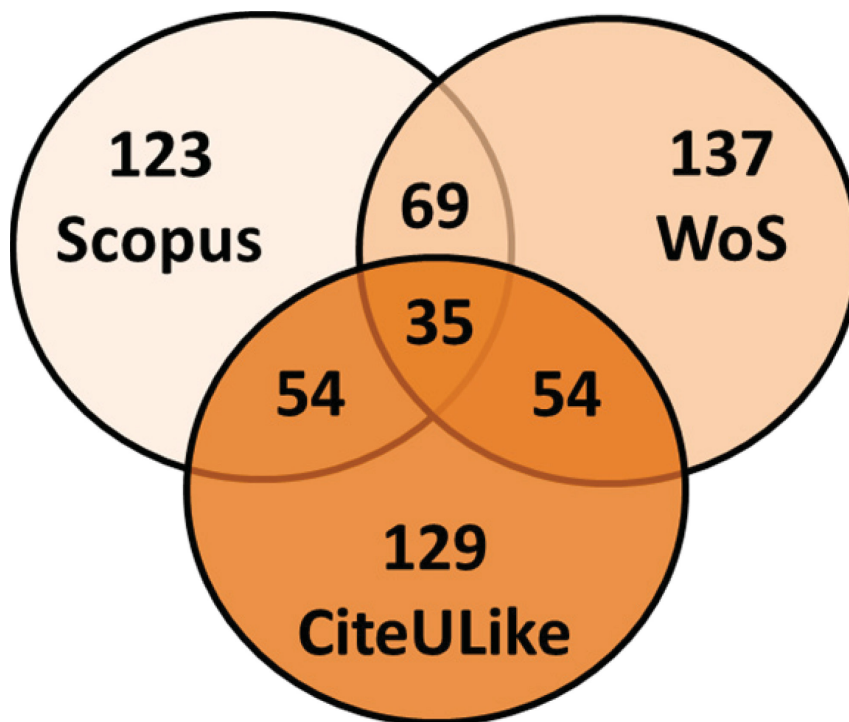


Figure 4.16. Distribution of relevant authors for all scientists for the three services Web of Science (WoS), Scopus and CiteULike, respectively.

information sources. Figure 4.17 shows the coverage of found relevant authors relative to the number of all named relevant authors. Note that author *a8* did not have any assigned tags and may have received biased results via a different method used to compile the CiteULike data set. Furthermore, author *a4* did not have any publications in CiteULike at all. In figure 4.17, both CiteULike networks – that based on common users and that based on common tags – are distinguished. In four cases, the network based on bibliographic coupling was deemed best, leading to coverage of 55.56% (*a4*), 57.78% (*a1*), 70% (*a5*) and 75% (*a6*), respectively. Target author *a6* in particular had a good BICO network compared to his other networks. All other networks were better in two cases each, which means that a CiteULike network (CULU or CULT) is best in a total of four cases. The results are positive for CiteULike because in this service, not all of the target authors' publications were bookmarked. In Scopus and Web of Science, almost all articles were retrieved (see table 4.6). Despite this fact, a high number of relevant authors was found in the bookmarking system for

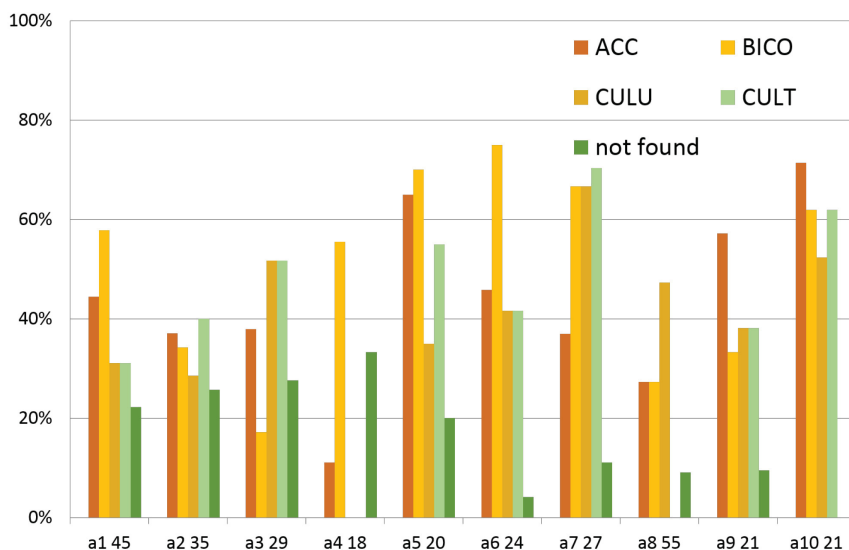


Figure 4.17. Coverage of relevant authors per target scientist (a1-a10), where the number shown on the x-axis is the total number of relevant authors named by the target researchers.

which a tag-based approach was applied in order to compile the data set.

In addition to the coverage of relevant authors for a whole network, the accuracy of the top 20 rankings is analyzed. The top 20 lists are taken (instead of the top 10) so as not only to consider those authors listed for evaluation. Concerning the coverage of the recommended authors in the top 20 list, ranked by the cosine coefficient, the BICO networks have the best results once again. For eight target researchers, the most relevant authors are found in the top 20 BICO list (figure 4.18). For the other two scientists, the CULU networks proved best, with author *a8* showing exceptional results. Remember that he was recommended 26 authors from the CiteULike network based on author occurrences in one user’s bookmarking list. He claimed that all authors were relevant, leading to 100% coverage. It is remarkable that the sparse data set delivered such results. However, these results might be biased due to the personal valuations of the target researcher, who preferred to give positive relevance ratings to his colleagues.

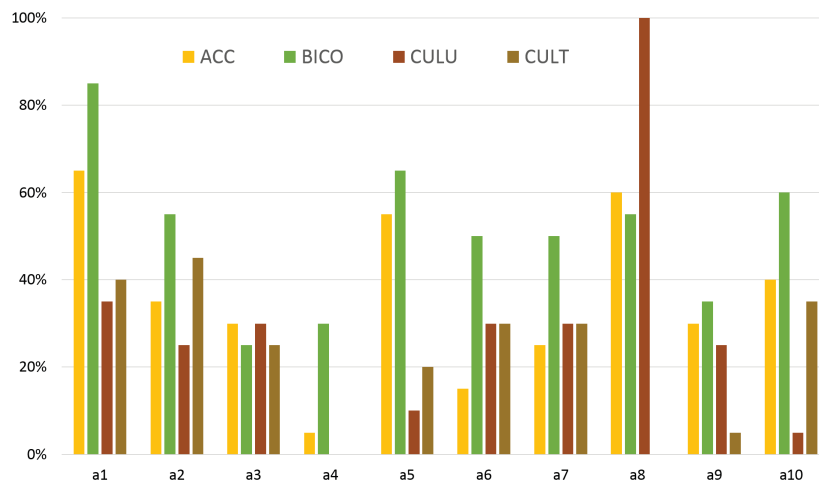


Figure 4.18. Coverage (similar to precision) of relevant authors per target scientist (a1-a10) for the top 20 recommended authors according to the cosine coefficient.

For the coverage of relevant authors for the top 20 lists, recall (similar to the coverage ratio of Baeza-Yates and Ribeiro-Neto (2011, p. 145)) and precision values can also be measured because the number of relevant authors (system-generated and user-selected) is known. The values show a recommender system's appropriateness in terms of suggesting relevant partners. In the current study, they offer a kind of counter-value to the novelty factor, as all relevant authors taken into consideration are also known and not novel. Regarding user trust aspects, it is useful for all values to show positive, but not too positive results. If recall and precision were high, a system would only recommend researchers the target scientists are already aware of. When the novelty value is high, the system recommends new partners, but the target scientists neither trust the system nor are able to evaluate the suggestions.

		a1	a2	a3	a4	a5	a6	a7	a8	a9	a10
	average novelty										
	0.46	0.51	0.35	0.54	0.75	0.65	0.55	0.35	0.00	0.54	0.37
	average recall top 20										
ACC	0.56	0.65	0.54	0.55	0.50	0.85	0.27	0.50	0.80	0.42	0.53
BICO	0.79	0.65	0.92	1.00	0.60	0.93	0.56	0.56	0.73	1.00	0.92
CULU	0.41	0.50	0.50	0.40	0.00	0.29	0.60	0.33	0.73	0.63	0.09
CULT	0.35	0.57	0.64	0.33	0.00	0.36	0.60	0.32	0.00	0.13	0.54
	average precision top 20										
ACC	0.36	0.65	0.35	0.30	0.05	0.55	0.15	0.25	0.60	0.25	0.40
BICO	0.51	0.85	0.55	0.25	0.30	0.65	0.50	0.50	0.55	0.35	0.60
CULU	0.29	0.35	0.25	0.30	0.00	0.10	0.30	0.30	1.00	0.25	0.05
CULT	0.23	0.40	0.45	0.25	0.00	0.20	0.30	0.30	0.00	0.05	0.35

Table 4.9. Evaluation values: novelty based on recommended author lists. Recall and precision values (based on named relevant authors who are appropriate collaborators) are measured according to the top 20 list ranked by cosine coefficient.

Table 4.9 shows the metric values per target researcher. The average novelty

value based on the recommended lists is 0.46. Average recall (equation 3.2.) ranges from 0.35 to 0.79, whereas average precision (equation 3.1) ranges from 0.23 to 0.51¹⁴. The average values are better for BICO and ACC networks, with BICO being the best. However, values for the CiteULike networks of some target researchers are better. Author *a8* has a precision of 1 and recall of 0.73. Both of author *a6*'s recall values, CULT and CULU, are better than the values for ACC and BICO, thus there is no distinct proof that any of the applied approaches is better than the others. Bibliographic coupling leads to slightly better results. Nevertheless, the number of found relevant authors per data set and the high recall and precision values of single target researchers support the assumption that bibliometric and altmetric approaches complement each other. Comparing the values for recall, precision and novelty, it can be said that all networks show relatively average results, with values being neither too low nor too high. Hence, the ratio between new unknown authors and relevant known authors is balanced. The appropriateness of these results for an expert recommendation system is dependent on a target researcher's needs and preferences.

The analyses of relevant authors allow us to draw some first conclusions. More diverse relevant authors are found by using different approaches (and social information) to measure similarity between scientists. Bibliographic coupling analysis retrieves the most relevant authors. However, collaborative filtering in CiteULike delivers the second-best results. Figure 4.16 shows that the number of relevant authors in all three data sets is similar. However, the crucial factor is the number of relevant authors found across all three data sets, which is quite low. It seems that the diverse methods and data sets complement each other. The next section will discuss the relevance of the visualized networks.

4.3.5 Results and Discussion II: Relevance of Visualized Graphs

Regarding the statements made by the evaluators, it is assumed that a network

¹⁴ Note: Both metrics consider relevant authors named in question #7, table 4.7.

that shows relations between known relevant authors and new unknown authors would help a target researcher estimate the relevance of the unknown individuals. This assumption will become clearer when looking at the evaluation results for the visualized networks for each applied model, i.e. ACC, BICO, CULU and CULT. The evaluation questions (table 4.8) focus on the representation of correct author communities and the appropriateness of the networks for the purposes of collaboration. The average ratings (1 to 10) for the visualized networks are:

- BICO network: 7.95
- ACC network: 6.35
- CULT network: 4.81
- CULU network: 3.75

Six of the target authors claimed that the BICO network is the best one, while author *a2* rated both BICO and ACC with nine (figure 4.19). This ranking is not surprising, since the interviewees stated that they found good collaborators among the recommended known authors. Additionally, bibliographic coupling in Web of Science retrieved the most important authors. Two authors preferred the ACC network, another one the CULT network. The CULU networks were

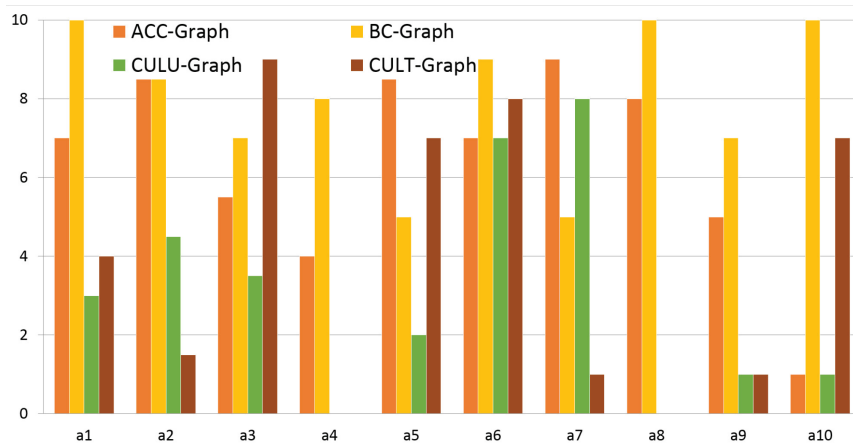


Figure 4.19. Relevance feedback (rating of 1-10) of visualized author networks.

not preferred by any of the interviewees. In many cases, the authors complained of the size of the network and its unclear arrangement. Both the CULU and CULT networks are larger because the size of the data sets is larger. As the numbers for $|U_b|$ and $|T_b|$ did not have to be searched and verified manually (as was the case for the Scopus and Web of Science data sets), there was no need to limit the author sets considered for recommendation. Another critical concern regarding the CiteULike networks are the low numbers of common users and tags – an aspect already discussed in the pre-studies (section 4.1). It means that similarity values between author-author pairs are quite high and only minor distinctions can be made between these authors. Here, user-based similarity is lower because the target researchers share between 1 and 6 common users. The amount of common tags is higher, ranging from 3 to 19, except for 41 common tags for target researcher *a2*. Nevertheless, author *a2* rated the CULT network at a mere 1.5, stating that the distribution of authors was not meaningful to him and that no relevant author was clustered around his own node. Author *a6*, who rated the CULT network at 8, only shares a maximum three common tags with 29 other authors. In his case, the CiteULike networks are not too large, which might have influenced the rating. Some CULU and CULT networks were already narrowed down on the basis of cosine coefficient values. Nevertheless, the target scientists had difficulties orienting themselves in large networks. This explains why the target scientists preferred the CULT networks instead of the CULU networks, as the number of common tags is higher than the number of common users.

In addition to users and tags, the number of bookmarked publications by researchers also influences the network structure, and thus the network rating. We assumed that novice researchers suffer from the cold start problem, meaning that none of their publications are found in the bookmarking service. In fact, this is also a problem for senior researchers. Coverage of the researchers' articles differs between Web of Science, Scopus, and CiteULike, and this data influences the construction of network graphs. However, it is surprising that CiteULike seems to feature more and more recent articles by novice researchers

(such as *a2* and *a5*, who were 36 and 30 years old at the time of evaluation). For a well-structured network visualization, the target authors' articles need to be bookmarked by users and to have appropriate tags assigned to them. If this is not the case, large CiteULike networks become very unstructured and are not useful for detecting collaborators and relationships among authors. Where a network had a clear structure showing author communities and separate small groups, the researchers deemed the graph helpful for detecting author "hubs" as well as finding new known and unknown collaborators. As a result, the target scientists were able to identify relevant authors. Figure 4.20 shows the CULT network of author *a3*, who rated this cluster at 9. The researcher suggested combining the CULT network with either the BICO or the ACC network. Where the graph was not well-structured, and only a few tags were assigned to the target researcher's articles in CiteULike, it was difficult to identify relevant authors and groups. In such cases, the interviewees regarded the graph as unhelpful in finding collaborators. This was the case for author *a7*, who rated the CULT network at 1 (figure 4.21).

Apart from the difficulties with regard to clear network arrangement, the scientists found author networks more helpful than pure lists of author names. They stated that "it is helpful because I can see whom I know and which relation I could intensify" (*a3*). Furthermore, networks help out with the tasks of finding out if "I overlooked an important research area" or "whom I could cite if I have to familiarize myself with a new topic" (*a3*). The researchers suggested combining the BICO and the CULT networks in order to be yielded all relevant authors and research communities. The reason for this is that the BICO network was deemed to show the main relevant authors "whom you find on your own" and "the usual suspects, [making it] a rather conservative network" (*a10*). Instead, the CULT network was deemed to reveal relations between authors of whom the target scientists were not aware beforehand, or "whose names were not available anymore" (*a3*). In the networks, the subjects also detected further important authors – relevant for possible collaboration – whom they forgot to name in the evaluation part b). This last point is quite crucial for the relevance

of expert recommendation systems: they help make people aware of not only implicit unknown relations, but also of explicit relevant relations whom they had forgotten.

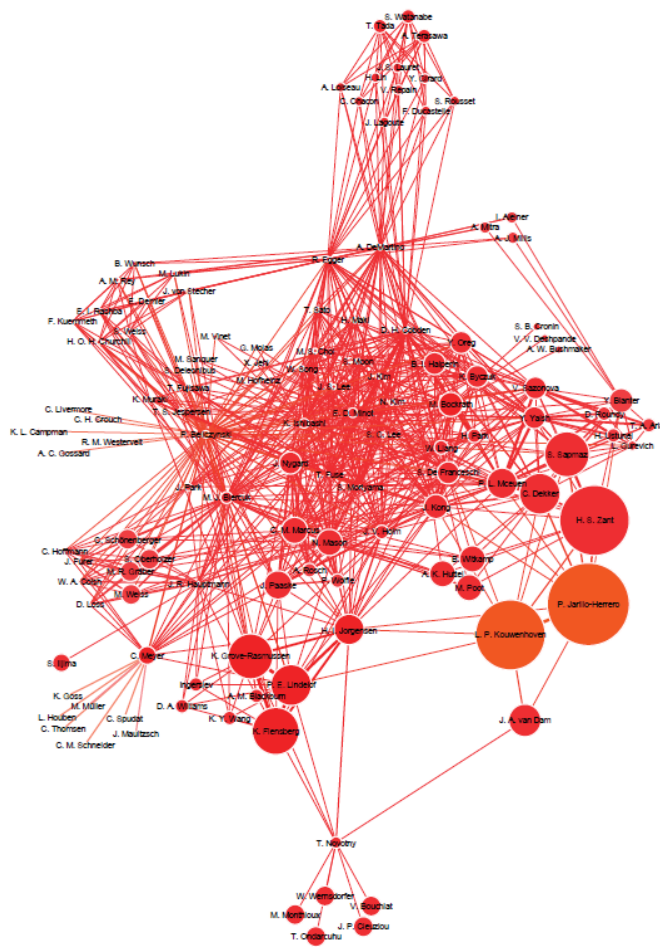
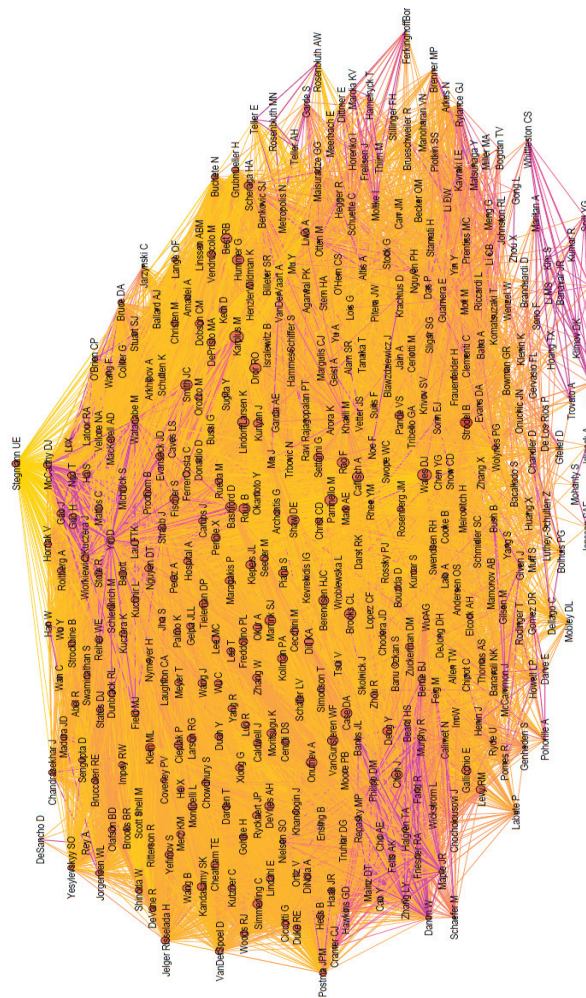


Figure 4.20. Author network of author a_3 based on common tags in CiteULike (CULT) (extract of main cluster), cosine coefficient threshold 0.45-0.99, network rating = 9.

According to the statements of the researchers, two main aspects influence the relevance of a visualized network. One is the need for author recommendation. To organize a conference or workshop, the scientists preferred bigger networks with more unknown people. Hence, CiteULike networks were better at helping

Figure 4.21. Author network of author *a7* on the basis of common tags in CiteULike (CULT), network rating = 1.



them find unrecognized scientists to invite. Another argument in favor of preferring CiteULike networks over BICO or ACC networks was that the former cover more diverse research fields: “It is more appropriate for finding people who belong to your field in a wider sense” (*a1*, who preferred his BICO network in this case). Two other researchers stated that they are working in multidisciplinary research fields, and that none of the BICO and ACC networks covered all of them. These scientists preferred the CiteULike networks.

For BICO and ACC networks, most of the interviewees stated that the authors in these two networks were too obvious to be similar and said that they were interested in bigger networks including more potential collaborators. Networks that only show relevant but previously known authors might not be appropriate for recommendation if there is a need for more unknown researchers in order to find new collaborators. However, some interviewees said that a small network would be sufficient for them because they preferred to have a clear structure. This answer was given when the researcher did a lot of work on their own and currently did not need many new collaborators, or when the researcher preferred known scientists as collaborators and thought that he or she had already built up a sufficiently large network. A running recommendation system could enable users to choose whether they want to be shown smaller or larger networks. Users could then decide either to stick to a smaller group of more familiar people or to extend their network to find new collaborators. Again, such systems can apply knowledge-based scenarios that allow for user interference. However, as studies show (see chapter 3), users might be lazy and not want to actively seek any recommendations. This fact should also be considered.

In any case, the advantage of a visualized network is that it shows the relations between authors, which helps a target scientist recognize the relevance of a person (“many authors are unknown, but there are also many familiar ones in the network” (*a2*)). A simple list of recommended authors cannot fulfill this task. Hence, it is important for a visualized network to include known persons, such as co-authors, and not exclude them. The researchers stated that these

people work as a kind of mediator between a target scientist and his or her unknown colleagues. In other words, the relevance of an unknown person can be estimated according to their relationships with known persons. The researchers said that the distribution of researchers and researcher communities was shown correctly in almost all networks. It is likely, albeit not explicitly proven, that any unknown scientist would also be allocated correctly within the graph. Hence, it is assumed that if an unknown person is shown to be clearly

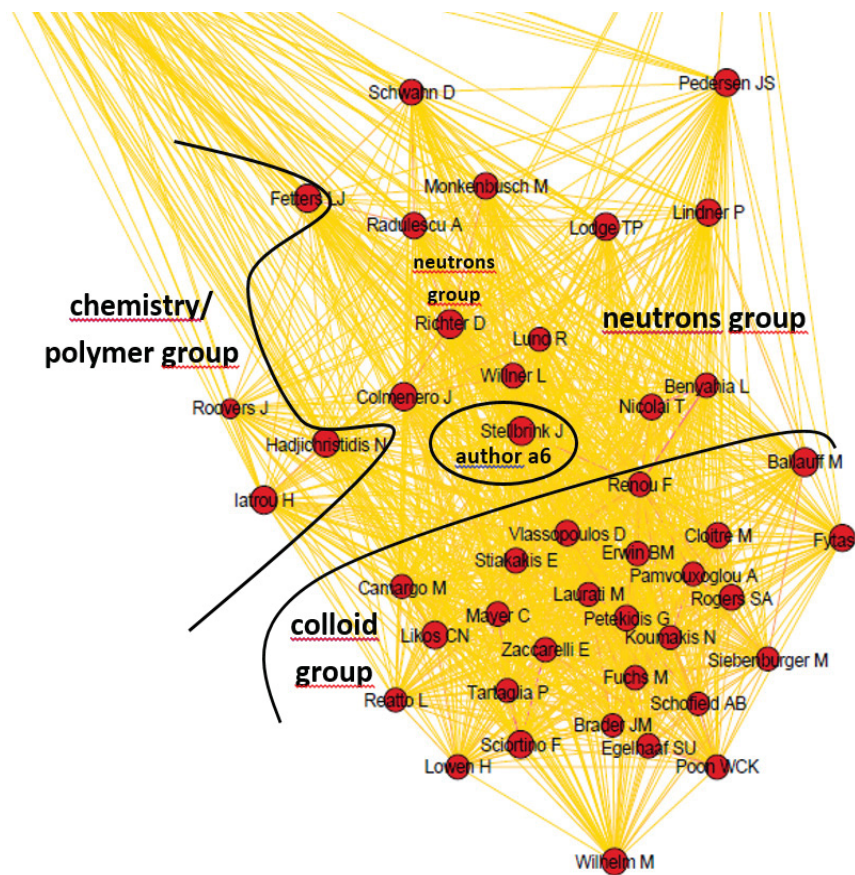


Figure 4.22. Author network for author *a6* on the basis of author co-citations (ACC) (extract of main cluster), with keywords assigned by *a6* during the evaluation.

connected to a known relevant research group, he or she probably does similarly relevant work.

Of course, in order to make a clear statement about the appropriateness of unknown authors, a target scientist needs to know more about these persons. Besides showing author-author relations, a folksonomy-based structure offers more information to be used for relevance valuation. Further categorization via tags helps to classify the scientists' work and gives less-than-clear user distributions in a graph more meaning. During the network evaluation, the researchers already arranged author groups within the networks with the help of keywords (see figure 4.22 for *a6*). Author *a6* identified different researcher groups, tagging them with the terms “chemistry/polymer group”, “neutrons group”, and “colloid group”. The users in CiteULike assigned the tags “colloids”, “md_p02_015_soft”, and “soft_matter” (to wit, polymer research belongs to soft matter) to his bookmarked publications. This shows that the tags assigned to the target authors' articles are quite appropriate and helpful for

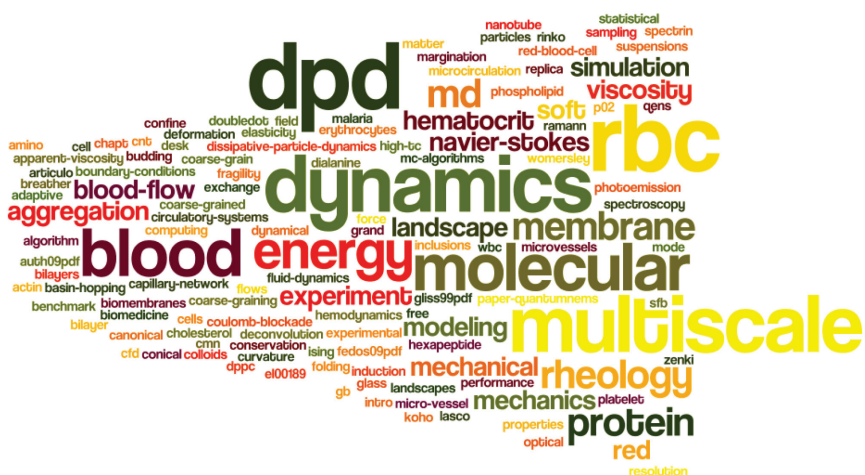


Figure 4.23. Tag cloud containing all tags assigned to the target researchers' bookmarked publications in CiteULike.

detecting such author groups. Figure 4.23 shows a tag cloud of the terms assigned to the target authors' articles in CiteULike (where tag size corresponds to frequency). For example, physicist *a2* described his current work with the terms “modeling blood flow”, “viscosity”, and “simulation”. The users assigned the same tags “blood-flow”, “modeling”, and “viscosity” to his articles.

To sum up the findings from the visualized network evaluation, BICO and ACC networks show a high number of potential collaborators, which the target researchers found helpful. The scientists stated that these graphs cover the core of their research fields and showed many familiar people. However, the larger CiteULike networks offered a better overview of related research fields and their authors. CULU and CULT networks were considered more helpful for a researcher searching for many new collaborators, for example to organize a conference, or for a researcher who has diverse research fields that should be covered in a network. Two target authors stated that they are working in multidisciplinary research fields and that neither the BICO nor the ACC network covered all of them. These latter preferred the CiteULike networks.

4.3.6 Combination of Applied Models

The applied approaches discussed above are shown to complement one another with regard to the task of finding potential collaborators. Hence, a recommendation system combining diverse methods could be designed on the basis of these findings. In chapter 3, different hybrid approaches are discussed. Monolithic hybridization systems work with a single implementation that considers aspects from diverse models. For example, the similarity between two users is measured on the basis of a common purchasing history and their navigation across item pages. Another method is to regard information about item content in order to recommend items to a target user in addition to opinions by neighboring users in a collaborative filtering system (Jannach et al., 2011). Jannach et al. (2011) further suggest exploiting constraint user feedback familiar from knowledge-based systems. In the suggested model, such a monolithic hybridization system is difficult, as author similarity is based on four

diverse models regarding data from diverse services. The diverse similarity assumptions must be combined before any measurement can be performed. The relevant aspects concerning similarity statements discussed in chapter 2 would get lost otherwise. Similarly, a pipelined design (Burke, 2002, 2007) would not make sense.

Important authors	mod Cos	rank mod Cos	cos ACC	rank ACC	cos BC	rank BC	cos CULU	rank CULU	cos CULT	rank CULT
Wales DJ	3.20	1	0.39	2	0.37	1	0.46	36	0.59	1
Wenzel W	1.92	3			0.02	47	0.87	2	0.51	3
Klenin K	1.86	4					0.87	1	0.51	2
Carr JM	1.83	5			0.22	2	0.71	3	0.24	89
Stock G	1.62	7	0.09	26	0.18	3	0.33	61	0.36	17
Derreumaux P	1.51	8	0.15	9	0.15	4	0.29	67	0.30	28
Klimov DK	1.48	9			0.09	30	0.50	11	0.39	7
Johnston RL	1.32	10					0.71	4	0.30	38
Nguyen PH	1.31	14	0.09	22			0.38	42	0.42	5
Miller MA	1.25	16					0.71	5	0.25	66
Caflisch A	1.15	25	0.05	59	0.12	9	0.22	124	0.29	44
Pande VS	1.12	30			0.13	7	0.32	64	0.24	93
Scheraga HA	1.00	44			0.09	32	0.26	99	0.27	47
Mu Y	0.84	62					0.35	47	0.25	67
Karplus M	0.82	64			0.10	21	0.17	191	0.21	143
Brooks CL	0.77	72			0.10	17	0.11	249	0.22	133

Table 4.10. Comparison: modified cosine and cosine values and ranks for the 16 important authors for a_7 .

To wit, a system would consider all authors with high similarity on the basis of common references, and then measure the similarity of these authors based on common citations, and so on. New authors would not be included in further steps. Hence, all relevant authors who were found via author co-citation and collaborative filtering analyses would be lost and the complementary aspect of the approaches could not be used anymore.

Then there is parallelized hybridization, which considers diverse approaches side by side and combines them in a final step. In the proposed study, this

design would call for unified author names and author disambiguation across all services, after which author similarity could be measured according to the cosine coefficients. The assumption is that authors who have high similarity values with a target researcher across all four models (common citation and references as well as common users and tags) are most alike, and thus highly relevant. More concretely, if the perspectives of a target author, other authors, and Web users bring together two authors via implicit relations, this indicates a high similarity between both. Thus the more implicit relations two authors have, the more similar they are.

A simple model involves summing up the values for each author-author pair from all four networks. However, there are great differences concerning the cosine values. In general, cosine values in the BICO network are very low compared to ACC as well as to similarity measurements based on collaborative filtering methods. The highest cosine value for similarity based on bibliographic coupling is 0.68, where value intervals between 0.3 and 0.01 are the most common. The high number of references the authors assigned to their publications minimizes similarity values between two authors. Additionally, similarity is comparatively high for collaborative filtering measurements in CiteULike because the number of common users and assigned tags is quite low. Here, the maximum cosine value is 1. If results are meant to be combined into one recommendation list for a target researcher, any summation and ranking regarding these cosine values will not be appropriate. Therefore, cosine values must be normalized. In the example in table 4.10, the important authors received relatively good ranking lists via the normalized and summarized cosine values. However, many relevant authors in the current study are found in only one network. The disadvantage of combining the data sets in a hybrid system is that these relevant authors would have a low chance of reaching a good ranking. If a researcher gets an author recommendation list, he or she will probably not pay any attention to the lower-ranked authors.

Concerning a visualized network for a hybrid approach, combined cosine values

would lead to different author relations and thus to enormously different networks. This would influence the perception of a target author concerning relevant persons and researcher groups. Another great issue that should be considered when combining the networks is the problem of author ambiguity, which would be tough to handle when using different services with diverse spelling preferences. Thus it is doubtful whether the combination of the methods leads to any new advantages for a target author. Evaluation showed that target researchers have clear preferences for BICO and ACC networks on the one hand (search for partners in one's close network), and larger CULU and CULT networks on the other (detection of new fields or finding people to invite to a conference). In a combined network, these preferences are not considered.

References

- Adomavicius, G., & Tuzhilin, A. (2005). Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734–749. doi:10.1109/TKDE.2005.99
- Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 54(6), 550–560. doi:10.1002/asi.10242
- Baeza-Yates, R., & Ribeiro-Neto, B. (2011). *Modern Information Retrieval: The Concepts and Technology Behind Search*. New York, NY, USA: Addison Wesley.
- Balby Marinho, L., Hotho, A., Jäschke, R., Nanopoulos, A., Rendle, S., Schmidt-Thieme, L., . . . Symeonidis, P. (2012). *Recommender Systems for Social Tagging Systems*. Boston, MA, USA: Springer.
- Balby Marinho, L., Nanopoulos, A., Schmidt-Thieme, L., Jäschke, R., Hotho, A., Stumme, G., & Symeonidis, P. (2011). Social tagging recommender systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 615–644). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_19
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi : An open source software for exploring and manipulating networks. In *Proceedings of the 3rd International AAAI Conference on Weblogs and Social Media*. Menlo Park, CA, USA: AAAI. Retrieved from

- <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>
- Berendsen, R., De Rijke, M., Balog, K., Bogers, A. M., & Van den Bosch, A. (2013). On the assessment of expertise profiles. *Journal of the American Society for Information Science and Technology*, 64(10), 2024–2044. doi:10.1002/asi.22908
- Bogers, A. M. (2009). *Recommender Systems for Social Bookmarking: Doctoral dissertation*. Retrieved from <http://ilk.uvt.nl/~toine/phd-thesis/>
- Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence* (pp. 43–52). San Francisco, CA, USA: Morgan Kaufmann Publishers. Retrieved from <http://dl.acm.org/citation.cfm?id=2074100>
- Burke, R. (2000). Knowledge-based recommender systems. In A. Kent (Ed.): *Encyclopedia of Library and Information Science* (69th ed., No. 32, pp. 180–201). New York, NY, USA: Marcel Dekker.
- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12, 331–370. doi:10.1023/A:1021240730564
- Burke, R. (2007). Hybrid web recommender systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.): *Vol. 4321. Lecture Notes in Computer Science. The Adaptive Web. Methods and Strategies of Web Personalization* (pp. 377–408). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-72079-9_12
- De Meo, P., Nocera, A., Quattrone, G., Rosaci, D., & Ursino, D. (2009). Finding reliable users and social networks in a social internetworking system. In *Proceedings of the 13th International Database Engineering & Applications Symposium (IDEAS)* (pp. 173–181). New York, NY, USA: ACM. doi:10.1145/1620432.1620450
- De Meo, P., Nocera, A., Rosaci, D., & Ursino, D. (2011). Recommendation of reliable users, social networks and high-quality resources in a social internetworking system. *AI Communications*, 24(1), 31–50. doi:10.3233/AIC-2010-0484
- Demaine, J. (2011). Author disambiguation for enhanced Science-2.0 services. In *Proceedings of the Information Science and Social Media - International Conference (ISSOME)* (pp. 29–40). Retrieved from <http://issuu.com/informationsvetenskap/docs/issome2011>
- Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3), 297–302. doi:10.2307/1932409
- Durao, F., & Dolog, P. (2009). A personalized tag-based recommendation in social web systems. In A. Dattolo, C. Tasso, R. Farzan, S. Kleanthous, D.

- Bueno Vallejo, & J. Vassileva (Eds.): *Proceedings of Workshop on Adaptation and Personalization for Web 2.0 at the 17th International Conference on User Modeling, Adaptation, and Personalization (UMAP)*. Retrieved from <http://ceur-ws.org/Vol-485/paper5-F.pdf>
- Egghe, L. (2010). Good properties of similarity measures and their complementarity. *Journal of the American Society for Information Science and Technology*, 61(10), 2151–2160. doi:10.1002/asi.21380
- Felfernig, A., Isak, K., & Russ, C. (2006). Knowledge-based recommendation: Technologies and experiences from projects. In G. Brewka (Ed.): *Vol. 141. Frontiers in artificial intelligence and applications. Proceedings of the 17th European Conference on Artificial Intelligence (ECAI), including Prestigious Applications of Intelligent Systems* (pp. 632–636). Amsterdam/Washington, D.C.: IOS Press.
- Gemmell, J., Shepitsen, A., Mobasher, B., & Burke, R. (2008). Personalizing navigation in folksonomies using hierarchical tag clustering. In I.-Y. Song, J. Eder, & T. M. Nguyen (Eds.): *Lecture Notes in Computer Science. Data Warehousing and Knowledge Discovery* (pp. 196–205). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-540-85836-2_19
- Hamers, L., Hemeryck, Y., Herweyers, G., & Janssen, M. (1989). Similarity measures in scientometric research: The Jaccard index versus Salton's cosine formular. *Information Processing & Management*, 25(3), 315–318. doi:10.1016/0306-4573(89)90048-4
- Haustein, S. (2012). *Multidimensional Journal Evaluation: Analyzing Scientific Periodicals beyond the Impact Factor* (1. Aufl.). *Knowledge and Information*. Berlin/Boston: De Gruyter Saur.
- Heck, T. (2011). A comparison of different user-similarity measures as basis for research and scientific cooperation. In *Proceedings of the Information Science and Social Media - International Conference (ISSOME)* (pp. 117–127). Retrieved from <http://issuu.com/informationsvetenskap/docs/issome2011>
- Heck, T. (2012a). Analyse von Folksonomy basierten Netzwerken als komplementärer Ansatz für Autorenempfehlungen in der Wissenschaft. In *Proceedings of the 2. DGI Conference: Social Media und Web Science - Das Web als Lebensraum* (pp. 179–193). Frankfurt a. M., Germany: DGI. Retrieved from <http://www.isi.hhu.de/heck>
- Heck, T. (2012b). Analyse von sozialen Informationen für Autorenempfehlungen. *Information - Wissenschaft & Praxis*, 63(4), 261–272. doi:10.1515/iwp-2012-0048
- Heck, T. (2012c). Recommendation for social networking in academia. In G. Widén & K. Holmberg (Eds.): *Library and Information Science Volume 5:*

- Social Information Research* (pp. 239–265). Bingley, UK: Emerald Group Pub. doi:10.1108/S1876-0562(2012)0000005013
- Heck, T. (2013). Combining social information for academic networking. In A. Bruckman, S. Counts, C. Lampe, & L. G. Terveen (Eds.): *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW)* (pp. 1387–1398). New York, NY, USA: ACM. doi:10.1145/2441776.2441932
- Heck, T., Hanraths, O., & Stock, W. G. (2011). Expert recommendation for knowledge management in academia. In A. Grove (Ed.): *Proceedings of the Annual Meeting of the American Society for Information Science and Technology*. Silver Spring, MD: ASIS&T. doi:10.1002/meet.2011.14504801365
- Heck, T., & Peters, I. (2010a). Expert recommender systems: Establishing communities of practice based on social bookmarking. In *Proceedings of 10th International Conference on Knowledge Management and Knowledge Technologies (I-Know)* (pp. 458–464). Graz, Austria: J.UCS. Retrieved from <http://www.isi.hhu.de/heck>
- Heck, T., & Peters, I. (2010b). Experten-Empfehlungen mit Social Bookmarking-Services. *i-com -Zeitschrift für interaktive und kooperative Medien*, 9(3), 7–11. doi:10.1524/icom.2010.0027
- Heck, T., & Peters, I. (2010c). Implizite Digitale Soziale Netze als Basis für Expertenempfehlungssysteme. In K. P. Fähnrich & B. Franczyk (Eds.): *Informatik 2010, Service Science – Neue Perspektiven für die Informatik (1). Lecture Notes in Informatics* (pp. 613–618). Bonn, Germany: Köllen Druck+Verlag. Retrieved from <https://www.gi.de/service/publikationen/lni.html>
- Heck, T., Peters, I., & Stock, W. G. (2011). Testing collaborative filtering against co-citation analysis and bibliographic coupling for academic author recommendation. In J. Freyne, S. S. Anand, I. Guy, & A. Hotho (Eds.): *Proceedings of the 3rd Workshop on Recommender Systems and the Social Web. 5th Conference on Recommender Systems (RecSys)*. New York, NY, USA: ACM. Retrieved from <http://www.dcs.warwick.ac.uk/~ssanand/RSWeb11/>
- Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1), 5–53. doi:10.1145/963770.963772
- Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006). FolkRank: A ranking algorithm for folksonomies. In *Proceedings of the Workshop Information Retrieval of the Special Interest Group Information Retrieval* (pp. 111–114). Retrieved from <http://www.kde.cs.uni-kassel.de/stumme/papers/2006/hotho2006folkrank.pdf>

- Huang, Z., Chen, H., & Zeng, D. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems*, 22(1), 116–142. doi:10.1145/963770.963775
- Hung, C. C., Huang, Y. C., Hsu, J. Y., & Wu, D. K. C. (2008). Tag-based user profiling for social media recommendation. In *Proceedings of the Workshop on Intelligent Techniques for Web Personalization and Recommender Systems at AAAI2008* (pp. 49–55). AAAI. Retrieved from <https://www.aaai.org/Papers/Workshops/2008/WS-08-06/WS08-06-006.pdf>
- Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin del la Société Vaudoise des Sciences Naturelles*, 37, 547–579. doi:10.5169/seals-266450
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2011). *Recommender Systems - An Introduction*. New York, NY, USA: Cambridge Univ. Press.
- Juršič, M., Mozetič, I., Erjavec, T., & Lavrač, N. (2010). LemmaGen: Multilingual lemmatisation with induced ripple-down rules. *Journal of Universal Computer Science*, 16(9), 1190–1214. doi:10.3217/jucs-016-09-1190
- Knautz, K., Soubusta, S., & Stock, W. G. (2010). Tag clusters as information retrieval interfaces. In *Proceedings of the 43th Annual Hawaii International Conference on System Sciences (HICSS)*. New York, NY, USA: IEEE. doi:10.1109/HICSS.2010.360
- Landia, N., Anand, S. S., Hotho, A., Jäschke, R., Doerfel, S., & Mitzlaff, F. (2012). Extending FolkRank with content data. In B. Mobasher, D. Jannach, W. Geyer, & A. Hotho (Eds.): *Proceedings of the 4th Workshop on Recommender Systems and the Social Web. 6th Conference on Recommender Systems (RecSys)* (pp. 1–8). New York, NY, USA: ACM. doi:10.1145/2365934.2365936
- Lee, D. H., & Brusilovsky, P. (2010). Social networks and interest similarity: The case of CiteULike. In *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia* (pp. 151–156). New York, NY, USA: ACM. doi:10.1145/1810617.1810643
- Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1), 76–80. doi:10.1109/MIC.2003.1167344
- Lops, P., Gemmis, M. de, & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 73–105). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_3
- Luhn, H. P. (1958). A business intelligence system. *IBM Journal of Research*

- and Development*, 2(4), 314–319. doi:10.1147/rd.24.0314
- Ma, Z., Pant, G., & Liu Sheng, O. R. (2007). Interest-based personalized search. *ACM Transactions on Information Systems*, 25(1), 5. doi:10.1145/1198296.1198301
- Mandl, M., Felfernig, A., Teppan, E., & Schubert, M. (2011). Consumer decision making in knowledge-based recommendation. *Journal of Intelligent Information Systems*, 37(1), 1–22. doi:10.1007/s10844-010-0134-3
- McNee, S. M. (2006). *Meeting User Information Needs in Recommender Systems: Doctoral dissertation*. Minnesota, MN, USA: University of Minnesota. Retrieved from <http://www-users.cs.umn.edu/~mcnee/mcnee-thesis-preprint.pdf>
- Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London (1854-1905)*, 58, 240–242. doi:10.1098/rsp1.1895.0041
- Persson, O. (2001). All author citations versus first author citations. *Scientometrics*, 50(2), 339–344. doi:10.1023/A:1010534009428
- Peters, I. (2009). *Folksonomies. Indexing and Retrieval in Web 2.0*. Berlin, Germany: De Gruyter Saur.
- Peters, I. (2012). Power tags as tools for social knowledge organization systems. In W. A. Gaul, A. Geyer-Schulz, L. Schmidt-Thieme, & J. Kunze (Eds.): *Studies in Classification, Data Analysis, and Knowledge Organization. Challenges at the Interface of Data Analysis, Computer Science, and Optimization* (pp. 281–289). Berlin/Heidelberg, Germany: Springer. doi:10.1007/978-3-642-24466-7_29
- Peters, I., Schumann, L., Terliesner, J., & Stock, W. G. (2011). Retrieval effectiveness of tagging systems. In A. Grove (Ed.): *Proceedings of the Annual Meeting of the American Society for Information Science and Technology*. Silver Spring, MD: ASIS&T. doi:10.1002/meet.2011.14504801338
- Peters, I., & Stock, W. G. (2010). “Power tags” in information retrieval. *Library Hi Tech*, 28(1), 81–93. doi:10.1108/07378831011026706
- Picault, J., Ribière, M., Bonnefoy, D., & Mercier, K. (2011). How to get the recommender out of the lab? In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.): *Recommender Systems Handbook* (pp. 333–365). New York, NY, USA: Springer. doi:10.1007/978-0-387-85820-3_10
- Rasmussen, E. M. (1992). Clustering algorithms. In W. Frakes & R. Baeza-Yates (Eds.): *Information Retrieval. Data Structures & Algorithms* (pp. 419–442). Prentice Hall, NJ, USA: Englewood Cliffs.
- Reher, S., & Hausteijn, S. (2010). Social bookmarking in STM. Putting services to the acid test. *Online*, 34(6), 34–42. Retrieved from

- <https://www.questia.com/library/p3702/online>
- Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620. doi:10.1145/361219.361220
- Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. In *SIGIR '02. Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 253–260). New York, NY, USA: ACM. doi:10.1145/564376.564421
- Sedhain, S., Sanner, S., Brazianus, D., Xie, L., & Christensen, J. (2014). Social collaborative filtering for cold-start recommendations. In A. Kobsa, M. Zhou, M. Ester, & Y. Koren (Eds.): *Proceedings of the 8th Conference on Recommender Systems (RecSys)* (pp. 345–348). New York, NY, USA: ACM. doi:10.1145/2645710.2645772
- Sinha, R. R., & Swearingen, K. (2001). Comparing recommendations made by online systems and friends. In *Proceedings of the 2nd DELOS Network of Excellence Workshop on Personalisation and Recommender Systems in Digital Libraries*. Biot, France: Ercim. Retrieved from <http://www.ercim.eu/publication/ws-proceedings/DelNoe02/RashmiSinha.pdf>
- Sinha, R. R., & Swearingen, K. (2002). The role of transparency in recommender systems. In *Extended Abstracts on Human Factors in Computing Systems (CHI)* (pp. 830–831). New York, NY, USA: ACM. doi:10.1145/506443.506619
- Sneath, P. H. A. (1957). Some thoughts on bacterial classification. *Journal of General Microbiology*, 17(1), 184–200. doi:10.1099/00221287-17-1-184
- Stock, W. G., & Stock, M. (2013). *Handbook of Information Science*. Berlin/Boston: De Gruyter.
- Szomszor, M., Cattuto, C., Alani, H., O'Hara, K., Baldassarri, A., Loreto, V., & Servedio, V. D. (2007). Folksonomies, the semantic web, and movie recommendation. In *Proceedings of the International Workshop on Bridging the Gap between Semantic Web and Web 2.0 at the 4th European Semantic Web Conference (ESWC)*. Retrieved from <http://eprints.soton.ac.uk/264007/>
- Tso-Sutter, K. H. L., Marinho, L. B., & Schmidt-Thieme, L. (2008). Tagaware recommender systems by fusion of collaborative filtering algorithms. In *Proceedings of the Symposium on Applied Computing (SAC)* (pp. 1995–1999). New York, NY, USA: ACM. doi:10.1145/1363686.1364171
- Tutte, W. T. (1963). How to draw a graph. In *Proceedings of the London Mathematical Society* (pp. 743–767). doi:10.1112/plms/s3-13.1.743
- Van Eck, N. J., & Waltman, L. (2008). Appropriate similarity measures for

- author co-citation analysis. *Journal of the American Society for Information Science and Technology*, 59(10), 1653–1661. doi:10.1002/asi.20872
- Van Eck, N. J., & Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology*, 60(8), 1635–1651. doi:10.1002/asi.21075
- Wang, J., de Vries, A. P., & Reinders, M. J. T. (2006). Unifying user-based and item-based collaborative filtering approaches by similarity fusion. In *Proceedings of the 29th Annual International Conference on Research and Development in Information Retrieval (SIGIR)* (pp. 501–508). New York, NY, USA. doi:10.1145/1148170.1148257
- Wenger, E. (2008). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, MA, USA: Cambridge Univ. Press.
- Wooding, S., Wilcox-Jay, K., Lewison, G., & Grant, J. (2006). Co-author inclusion: A novel recursive algorithmic method for dealing with homonyms in bibliometric analysis. *Scientometrics*, 66(1), 11–21. doi:10.1007/s11192-006-0002-7

Conclusion

The aim of this work was to analyze and evaluate a proposal for an expert recommender system for researchers, while concurrently looking at scientometric and alternative approaches. The focus was placed on the usefulness of the models and data applied with regard to the opinions of physicists who were given personalized recommendations based on their current publications.

There is a need for collaboration and the expansion of researchers' networks. These issues are important, on the one hand, because of extrinsic intentions, as collaborations with research colleagues lead to higher levels of reputation for institutions as well as for individual scientists. On the other hand, they are also crucial because of intrinsic needs – for novices as well as for senior researchers – because scientific knowledge creation and the development of new knowledge require further interaction between researchers. New collaborations are possible if environments are established in which academic communities of practice can evolve. These communities and their nature of mutual engagement and interactive participation support the process of academic knowledge creation by researchers. Thus we concluded that support should be provided to foster the detection of such communities. There are new ways of detecting relations between researchers that are not found directly and of which most researchers are unaware. On the one hand, scientometric models that look at indirect author relations are appropriate because their data directly refers to a researcher's reputation. On the other hand, further social information from user-generated data is relevant because it introduces a new perspective and further considers the opinions of readers in a social bookmarking system. Relations within these systems are comparable to co-citation relations and thus may be used in conjunction with scientometric data. Considering the concept of collective intelligence, this type of data can expand a researcher's perspective. Hence, a model is proposed which considers three diverse perspectives – target researcher, third author, users – in order to derive multiple views on scientific

relations. The intention of an expert recommender system is to consider those views in order to help researchers find new collaborators. The idea of the “wisdom of crowds” is applied via collaborative filtering models, which gather historic user data to detect similarities between elements and to give personalized recommendations. Collaborative filtering models in bookmarking services are adapted to the academic field and used as alternative metrics in conjunction with scientometric models so as to establish expert recommendations for researchers. The case studies in chapter 5 conducted experiments for these recommendations and answered the work’s research questions:

1. Can researchers be helped to find relevant experts for collaboration via expert recommendation on the basis of scientometric and alternative approaches?

The evaluation results showed that relevant researchers for collaboration can be found via the four applied models. Novelty as well as recall and precision factors allow for the calculation of mean values. In other words, the applied models retrieve known and relevant authors as well as unknown authors. This mixture seems appropriate. A recommendation system should be able to suggest novel elements; otherwise it is no help to a user. However, research shows that systems suggesting known elements are trusted, which is a crucial factor for a person’s decision to use a system. These findings were also confirmed in the case study. The evaluation of the author lists showed that researchers tend to “think in networks”. They evaluated their peers by referring to their reputations and connections to other researchers. The evaluation of the author networks painted a clearer picture. Researchers who found known colleagues in the author networks referred to unknown authors via their connections to known authors. The participants stated that known colleagues can act as a mediator, and that the relevance of unknown researchers can be estimated via their connections to known peers. These connections provide a first hint toward the appropriateness of a colleague with regard to future collaboration. However, it must be noted that the correct relevance of an unknown author cannot not be stated *ad hoc*

without further investigation. In a running recommender system, though, links to an author's papers can be made that facilitate relevance estimates. Furthermore, aboutness tags (Peters, 2009) can deliver information because they show topical relations between researchers. The evaluation shows that terms used by the interviewees and tags assigned by users are quite similar to one another.

Concerning the representation of author recommendations, there is a clear preference for visualized networks because they have one great advantage. Known authors within the networks are connectors between a target researcher and unknown authors. Additionally, the participants stated that these networks include more valuable information because researcher groups and their connections could be detected. In these networks, potential collaborators were found whom researchers were not aware of before.

2. Is professionally indexed data from information services, based on scientometric approaches, appropriate for use in expert recommendation?

Retrieving data from Scopus and Web of Science is sometimes difficult because it must be conducted manually. McNee (2006) applied automatic approaches with data from an online service that also collects citation data based on automatic approaches. His results show that co-citation does not deliver good results compared to other approaches. Although the work at hand did not compare diverse algorithmic measurements, we can say that data based on co-citation analysis does lead to appropriate results concerning the relevance of recommended authors. Similar approaches using co-citation analysis also lead to positive results (Guns & Rousseau, 2013, 2014). The number of found publications by the participants was quite high for both services, namely Scopus and Web of Science. These numbers appear sufficient. Improvements have to be made with regard to author disambiguation.

3. Is user-generated data from social bookmarking services, based on collaborative filtering models, appropriate for use in expert

recommendation?

Despite low coverage of target researcher publications, data from CiteULike delivered a high number of relevant authors, even for novice researchers (figure 5.16). Furthermore, different relevant authors were found compared to those retrieved via co-citation and bibliographic analysis. As other studies previously suggested (Lee & Brusilovsky, 2010), similarity based on tags is more appropriate because the number of common tags is higher than the number of common users. Participants preferred the CULT network to the CULU network. User-generated data contains more information. Besides only showing author-author relations, a folksonomy-based structure delivers tags that show topical relations between researchers. As this study and other works (Heckner, Mühlbacher, & Wolff, 2008) show, tags are useful for navigating a researcher's field. This advantage can be used in a running recommender system and helps users to evaluate the suggested recommendations.

4. Are there any differences in the outcomes concerning the approaches and the data sets?

The comparison of the diverse models leads to two major findings. Scientometric as well as collaborative filtering approaches lead to similar numbers of relevant authors being found, whereas bibliographic coupling leads to the highest-quality results. The crucial factor is that only a few authors are found by all three services, although all models refer to a target researcher's current publications. Hence, if only one model or source is chosen for author recommendation, relevant authors will be lost. The case study shows clearly that diverse models should be applied, using different sources. In this case, user-generated data complements classic scientometric models. Despite a lower coverage of researcher publications, an appropriate number of relevant authors was found in CiteULike. Researchers also stated that the BICO and ACC networks did not show surprising new authors. Authors in these networks were the usual suspects. Participants who were more open to other research fields preferred CiteULike networks because they show more diverse author groups

from research fields connected to the participants'. Interviewees said that CULT networks in particular include researcher groups concerned with topics that do not belong to their immediate purview. Therefore, one argument for preferring CiteULike networks instead of BICO or ACC networks was that these networks cover more diverse research fields: "It is more appropriate for finding people who belong to your field in a wider sense". Depending on the purpose of collaboration, for example planning a conference or learning about neighboring research fields, participants deemed those networks to be more appropriate. Other researchers preferred the BICO networks, which were rated the best.

To summarize the findings, expert recommendations for the academic sphere should consider diverse social information about researchers and apply diverse models in order to give relevant recommendations. User-generated data and collaborative filtering confirm that the user perspective can expand a scientist's network and suggest further possible collaboration. However, a target author's need is crucial in this respect. To implement a running recommender system, it must be kept in mind that researcher needs and preferences differ. Hence, a system should be able to adapt to these circumstances. Jannach, Zanker, Felfernig, and Friedrich (2011) suggest further exploiting constraint user feedback known from knowledge-based systems. "Knowledge-based recommendation paradigms have a clear advantage in utilizing explicit user input and transforming it into recommendations" (Zanker & Jessenitschnig, 2009, p. 161). However, a recommendation process that overcomes the narrowness of a researcher's perspective is highly desirable. Therefore, systems that react to current researcher needs and simultaneously suggest new and previously unknown elements are suitable for the academic field. A good example for such a model is the database Sowiport (Mutschke, Mayr, Schaer, & Sure, 2011) (see also chapter 4), which does not favor a combination of diverse models, but aims to represent them simultaneously. A user gets different ranking results based on diverse metrics and is free to choose based on his or her preference. He or she is encouraged to look at the results from different perspectives, but is not forced to stick to a choice pre-defined by the system.

The researcher has to trust the system, which should not make any strict prescriptions concerning ideal collaborators. Only then can a system foster community building and knowledge creation beyond restricted perspectives.

References

- Guns, R., & Rousseau, R. (2013). Predicting and recommending potential research collaborations. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.): *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference (ISSI)* (pp. 1409–1418). Wien, Austria: Facultas Verlags- und Buchhandels AG. Retrieved from <http://www.issi2013.org/proceedings.html>
- Guns, R., & Rousseau, R. (2014). Recommending research collaborations using link prediction and random forest classifiers. *Scientometrics*, *101*(2), 1461–1473. doi:10.1007/s11192-013-1228-9
- Heckner, M., Mühlbacher, S., & Wolff, C. (2008). Tagging tagging. Analysing user keywords in scientific bibliography management systems. *Journal of Digital Information*, *9*(2). Retrieved from <https://journals.tdl.org/jodi/index.php/jodi/article/view/246>
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2011). *Recommender Systems - An Introduction*. New York, NY, USA: Cambridge Univ. Press.
- Lee, D. H., & Brusilovsky, P. (2010). Social networks and interest similarity: The case of CiteULike. In *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia* (pp. 151–156). New York, NY, USA: ACM. doi:10.1145/1810617.1810643
- McNee, S. M. (2006). *Meeting User Information Needs in Recommender Systems: Doctoral dissertation*. Minnesota, MN, USA: University of Minnesota. Retrieved from <http://www-users.cs.umn.edu/~mcnee/mcnee-thesis-preprint.pdf>
- Mutschke, P., Mayr, P., Schaer, P., & Sure, Y. (2011). Science models as value-added services for scholarly information systems. *Scientometrics*, *89*(1), 349–364. doi:10.1007/s11192-011-0430-x
- Peters, I. (2009). *Folksonomies. Indexing and Retrieval in Web 2.0*. Berlin, Germany: De Gruyter Saur.
- Zanker, M., & Jessenitschnig, M. (2009). Case-studies on exploiting explicit customer requirements in recommender systems. *User Modeling and User-Adapted Interaction*, *19*(1-2), 133–166. doi:10.1007/s11257-008-9048-y