# Need Help? Recommending Social Care Institutions

Matthias Traub*
Know-Center
Graz, Austria
mtraub@know-center.at

Emanuel Lacic*
Graz University of Technology
Graz, Austria
elacic@know-center.at

Dominik Kowald
Know-Center
Graz, Austria
dkowald@know-center.at

Martin Kahr
NEKOM GmbH
Voitsberg, Austria
kahr.m@nekom.com

Elisabeth Lex
Graz University of Technology
Graz, Austria
elisabeth.lex@tugraz.at

## ABSTRACT

In this paper, we present work-in-progress on a recommender system designed to help people in need find the best suited social care institution for their personal issues. A key requirement in such a domain is to assure and to guarantee the person's privacy and anonymity in order to reduce inhibitions and to establish trust. We present how we aim to tackle this barely studied domain using a hybrid content-based recommendation approach. Our approach leverages three data sources containing textual content, namely (i) metadata from social care institutions, (ii) institution specific FAQs, and (iii) questions that a specific institution has already resolved. Additionally, our approach considers the time context of user questions as well as negative user feedback to previously provided recommendations. Finally, we demonstrate an application scenario of our recommender system in the form of a real-world Web system deployed in Austria.

## Keywords

content-based recommender; hybrid recommender; time context; negative feedback; BLL equation; social care institutions

## 1. INTRODUCTION

Social care institutions play a key role in many societies. Such institutions are not just buildings or places, but structures of trust, support, and social engagement. Each institution has different responsibilities and tasks assigned, which contribute to the overall function and stability of the social care system and our whole society.

For example, refugees and asylum seekers need to be provided with a range of social care services (e.g., health care, mental health, housing help, etc.), but their personal experiences, background and marginalized positions present challenges in providing good quality social care [8].

**Problem.** Social care institutions often have to assist a large number of people who search for help with respect to personal and other issues. As specific social services cannot effectively be delivered by one standalone institution, it can be challenging to identify an appropriate institution from the vast amount of providers. Often, inquiries need to be urgently resolved or at least timely re-directed to the most appropriate institution (e.g., to help with domestic violence). Many social workers additionally face a high stress level due to an increasing number of inquiries [3].

Thus, reducing the workload of social workers by means of a recommender system, which automatically assigns people searching for help to social care institutions, would give social workers more time to focus on an individual's needs.

**Objectives & contributions.** Together with several partnered social care institutions, we identified the need to provide a recommendation system, which can support (i) people in identifying the appropriate institution, and (ii) social care providers by reducing their workload when looking to help people with their issues. Two key requirements however need to be met in that process. Firstly, a person's privacy and anonymity need to be assured [1]. This is especially important as reducing the restraints that we feel that others may impose on us (e.g., being judged) results in more honest questions [10].

Secondly, the recommendation system needs to consider the way people state successfully resolved questions and adapt as questions change over time. This suggests that considering the time factor as well as user feedback should be of special importance as word meanings do change over time [15].

Therefore, in this paper, we present a recommender system that aims to suggest social care institutions to people searching for help, which utilizes three data sources containing textual content, namely (i) metadata from social care institutions, (ii) institution specific FAQs, and (iii) questions that a specific institution resolved. Additionally, we consider the context of time a user questions has been asked as well as negative feedback to previously provided recommendations. This recommender system is than demonstrated in an application scenario in the form of a real-world Web system deployed in Austria.

**Structure of this paper.** The remainder of this paper is organized as follows: In Section 2, we present related work in the field, which is followed by a detailed description of our recommender system design in Section 3. Next, we demon-

*Both authors contributed equally to this work.

[1]Defined in the EU Data Protection Directive 95/46/EC and the upcoming General Data Protection Regulation (GDPR).

strate our application scenario in Section 4 and close the paper with our conclusions and plans for the future in Section 5.

# 2. RELATED WORK

With respect to research related to our work, the area of expert recommendations [11] seems to be the most relevant one. Here, a ranked list of domain experts is recommended for a given search query.

Existing work on expert recommendations can be categorized into two methods: The first method is based on link analysis. For instance, the authors of [13] apply a link analysis approach combined with a community-aware algorithm to identify and rank experts.

However, as shown in [7], experts tend to answer questions where they have a higher chance to make a valuable contribution. This means that there is a selection bias when experts answer questions. One explanation for this could be that "normal" users do not have enough expertise to give a complete solution or because they do not want to put much effort into answering. Based on that, the authors use a Gaussian classification model to distinguish between experts and "normal users".

The second method to identify experts is topic-oriented and is based on latent topic modeling techniques (e.g., [4]). Although such an approach could be suitable for the task of finding the right social institution, techniques like Latent Dirichlet Allocation (LDA) are computationally expensive and do not support the requirement of providing real-time recommendations without the need of retraining the model at every data update.

Although expert recommendation is already a well defined and studied research area, to the best of our knowledge, our proposed recommender system is the first work tackling the problems and requirements of recommending social care institutions.

# 3. RECOMMENDER SYSTEM DESIGN

Our recommendation system is based on a Hybrid Content-Based Filtering strategy. As shown in Figure 1, we not only propose an approach that leverages different data sources but also considers the time context of asked question as well as any negative feedback that has been given to provided recommendations.

## 3.1 Data Sources

In collaboration with partnered social care institutions, we identified three common data sources which can be utilized to generate relevant recommendations:

**SCI - Social Care Institutions.** *SCI* data contains textual content about a specific institution. An institution's specific description of what the institution is doing and to whom it is aimed for can be changed over time. Such a change hover could result that previous issues resolved by the particular institutions would not be resolved anymore.

**FAQ - Frequently Asked Questions.** *FAQ* data contains predefined textual questions and answers related to a specific institution. This is a potentially rich data source as institutions usually have examples of real issues they successfully resolved.

The *SCI* and *FAQ* data sources are especially helpful in

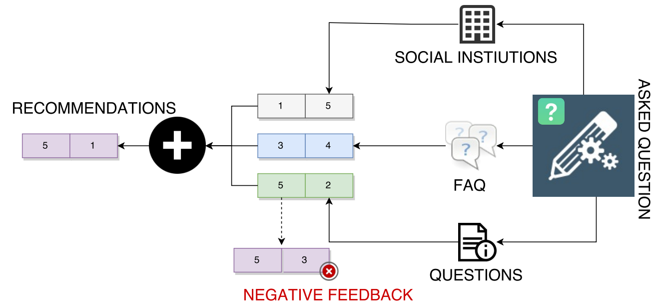


**Figure 1: Schematic illustration of our recommender system for social care institutions. The various data sources are combined via a hybrid content-based approach. Negative feedback on already resolved questions is also dynamically incorporated to boost the relevancy of the recommended social instiutions.**

tackling the problem of cold-start items (i.e., there are only few or none users with interaction data to a specific institution available). Since they ensure that already relevant content is available, it is not only possible to calculate recommendations at the initial stage of the project, but also recommend relevant institutions to uncommon issues or even when the work description of the institution changes.

**RQT - Resolved Questions.** *RQT* data contains the information which institution could finally help with the asked question. In addition to the textual content of the question and the relation to a specific institution, the time when a specific question (i.e., personal issue) has been resolved is also tracked.

There are several ways that questions are resolved: either directly in the system or when the person who originally posed the question then visits the recommended institution. Due to the fact that we cannot store who posed a question (i.e., because of the anonymity constraints), there is no direct way to trace back a question to a user. We tackle this by enabling the social worker to search for unresolved questions that deal with the particular topic, together with the client.

In case no suitable question is retrieved (e.g., when during their talk a completely new question was formulated), the social worker can post a new question and either resolve it, or forward the person to other, more appropriate institutions. To reduce the number of unresolved questions, any unresolved question will get invalidated after a defined period of time (e.g., 1 month).

## 3.2 Hybrid Approach

With respect to the requirements, the core method to find relevant institutions utilizes Content-Based Filtering (CBF) [9]. Each data source entry (i.e., *SCI*, *FAQ* or *RQT*) is represented by means of a Vector Space Model [12]. In order to find data source entries linked to an institution that are similar to the posed question, we apply TF-IDF [12] on the pre-processed question text (e.g., removing stop-words, stemming, etc.). To assure real-time recommendations and to handle a high load of requests, we utilized and extended our recently recently presented ScaR framework [6] which

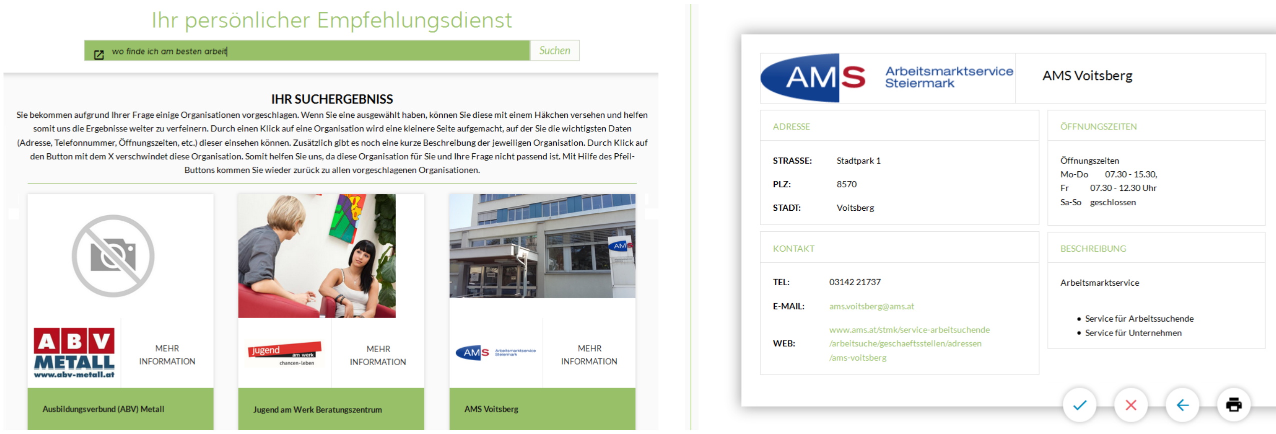


**Figure 2: The current state of the portal in which people can find the right social institution for their needs. After entering a search query, the users are provided with a list of recommendations (left). When clicking on one of the recommended institutions, detailed information are shown with the possibility to rate the recommendation (right). This rating is used to adjust future recommendations ("feedback loop").**

exploits the Apache Solr[2] enterprise search engine and its built-in TF-IDF ranking formula[3].

In order to find the most relevant institutions for the asked question, we combine the CBF approach using each data source as seen in Figure 1. To be specific, we adapt, modify and extend the Cross-Source Hybrid defined by Bostandjiev et al. [2] using the following formula:

$$W_{rec_i} = \sum_{s_j \in S} (W_{rec_i,s_j} \cdot BLL_{rec_i,s_j}) \cdot NF_{rec_i} \cdot |S_{rec_i}| \quad (1)$$

, where $W_{rec_i,s_j}$ is the similarity value of a data source entry calculated using the posed question, $BLL_{rec_i,s_j}$ is a time-context component and $NF_{rec_i}$ is the dynamically adaptable negative feedback component. The final weight of a recommendable institution, $W_{rec_i}$, is then given by the sum of the component multiplications for each data source. The number of data sources where the recommended institution $rec_i$ appears (i.e., $|S_{rec_i}|$) is also used to strongly favor social institutions that have been identified by more than one data source.

**Time context.** Considering the time context is especially important as institution responsibilities do change over time (e.g., adding or removing responsibilities, an institution gets completely shut down, etc.). In such cases, the available SCI and FAQ data sources, as well as the corresponding resolved questions would provide false positives in the list of recommended institutions. To consider the time context (i.e., favor more recent data) between an asked question and a similar data source entry tied to an institution, we make use of the Base-Level Learning ($BLL$) equation proposed by Anderson et al. [1]. As shown by related work [5, 14], the $BLL$ equation can be used in terms of time-dependent recommender systems and thus, should be applicable to determine a relevance value for an institution based on the institution's assignment time:

$$BLL_{rec_i,s_j} = ln((t_{now} - t_{rec_i,s_j})^{-d}) \quad (2)$$

, where $t_{now}$ is the current time of the recommendation and $t_{rec_i,s_j}$ is the time of the institution assignment to a particular data source entry (i.e., the time when a question was resolved). The exponent $d$ is used to model the power law of forgetting and is usually set to .5 (see [1]). To map these $BLL$ values to a range of 0 - 1, they are normalized according to [5].

**Negative Feedback.** To lower the relevance of candidate institutions that were found less helpful, as seen in Figure 2 (right) it is possible to provide negative feedback for every recommendation to a posed question. As such, every candidate $RQT$ data entry contains in addition to the relevant social institution also a list of recommendations assigned to it, including the information which ones were deemed as not relevant (also seen in Figure 1). By considering every extracted recommendation list, we formulate the negative feedback function as follows:

$$NF_{rec_i} = 1 - n^{-position_{rec_i}} \quad (3)$$

, where $n$ is the size of the recommendation list where the institution $rec_i$ occurred, and $position_{rec_i}$ is the index ($position_{rec_i} \in [1, n]$) of the institution which was found not useful. This means, the lower $NF_{rec_i}$ gets, the less relevant the candidate institution will be. However, this will not affect the relevance of institutions with no negative feedback, as in this case $NF_{rec_i}$ is set to 1.

## 4. APPLICATION SCENARIO

Our proposed recommender system is currently deployed for one district in southern Austria in the course of a field testing phase. It should help the local government to track the activities with all the different social care institutions and help people to find the right institutions more quickly. In the current phase already 85 institutions are registered among which 63 are active and thus, will be considered in the recommendation process. Apart from that, 80 FAQ questions and 73 FAQ answers are recorded to help the system

[2]http://lucene.apache.org/solr/
[3]https://lucene.apache.org/core/5_0_0/core/org/apache/lucene/search/similarities/TFIDFSimilarity.html

with institutions without related user questions (i.e., cold-start institutions) and to relate some very specific questions to this institution. Since its launch at the beginning of 2016, in only 8 months, we already recorded 2,875 user questions and 2,877 user feedbacks.

The current implementation of the system is shown in the two screenshots of Figure 2. In this example, we entered the search query "where can i find work?" (**left**). Please note that although this example is in German, the system is capable of any language when provided with the correct data.

For this search query several institutions have been found and listed in the user interface (**left**). If the user clicks on one of the suggested institutions, detailed information, including contact details, are shown in another screen (**right**). Here, the user also has the possibility to provide feedback for the suggested institution (in our example the local job office). The user feedback can either be positively (using the tick) or negatively (using the cross).

## 5. CONCLUSION & FUTURE WORK

In this paper, we presented work-in-progress on a recommender system, which aims to help people in need find the most appropriate social care institution online. With respect to the domain requirements, we propose a hybrid content-based approach, which utilizes the description of the institutions, frequently asked questions and already resolved questions.

Furthermore, in our approach we suggest to dynamically incorporate the time context of posed questions as well as gathered negative feedback to previously provided recommendations. These two factors are especially important to consider as responsibilities, locations and availability of social care institutions do change over time.

We demonstrated an application scenario of our recommender system in the form of a real-world Web system deployed in a district in southern Austria. Since its launch at the beginning of 2016, in only 8 months, we already recorded 2,875 user questions and 2,877 user feedbacks.

We believe that our work can serve as a baseline of how an effective recommender system for social care institutions can be realized and deployed.

**Future work.** In the near future, in cooperation with the social workers, we want to analyze the sentiment of the posed questions and incorporate it into our recommendation approach (e.g., [16]). Furthermore, we plan to evaluate our system in our presented live setting.

In this respect, we are especially interested in determining the impact of incorporating the time context and the negative feedback. Besides that, it would be interesting to include a dynamic weighting for the different data sources in our hybrid approach based on the received user feedback.

**Acknowledgements.** This work is supported by the Know-Center and the EU-funded project Learning Layers (GA: 318209). The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency (FFG).

## 6. REFERENCES

[1] J. R. Anderson, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin. An integrated theory of the mind. *Psychological Review*, 2004.

[2] S. Bostandjiev, J. O'Donovan, and T. Höllerer. Tasteweights: a visual interactive hybrid recommender system. In *Proc.*, RecSys '12. ACM.

[3] R. Hardy. How can high social work caseloads be tackled? https://www.theguardian.com/social-care-network/2015/jan/21/how-can-high-social-work-caseloads-be-tackled, 2015. [Online; accessed 20-July-2016].

[4] S. H. Hashemi, M. Neshati, and H. Beigy. Expertise retrieval in bibliographic network: a topic dominance learning approach. In *Proc. of CIKM '13*.

[5] D. Kowald, P. Seitlinger, C. Trattner, and T. Ley. Long time no see: The probability of reusing tags as a function of frequency and recency. In *Proc. WWW '14*. ACM.

[6] E. Lacic, M. Traub, D. Kowald, and E. Lex. Scar: Towards a real-time recommender framework following the microservices architecture.

[7] A. Pal and J. A. Konstan. Expert identification in community question answering: exploring question selection bias. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 1505–1508. ACM, 2010.

[8] B. Patel and N. Kelley. The social care needs of refugees and asylum seekers. *Social Care Institute for Excellence Race equality discussion paper*, 2, 2006.

[9] M. J. Pazzani and D. Billsus. Content-based recommendation systems. In *The adaptive web*. Springer, 2007.

[10] T. Postmes and R. Spears. Behavior online: Does anonymous computer communication reduce gender inequality? *Personality and Social Psychology Bulletin '02*.

[11] F. Riahi, Z. Zolaktaf, M. Shafiei, and E. Milios. Finding expert users in community question answering. In *Proceedings of the 21st International Conference on World Wide Web*, pages 791–798. ACM, 2012.

[12] G. Salton and M. J. McGill. *Introduction to Modern Information Retrieval*. 1986.

[13] M. Shahriari, S. Parekodi, and R. Klamma. Community-aware ranking algorithms for expert identification in question-answer forums. In *Proceedings of i-KNOW'15*, page 8. ACM, 2015.

[14] C. Trattner, D. Kowald, P. Seitlinger, T. Ley, and S. Kopeinik. Modeling activation processes in human memory to predict the use of tags in social bookmarking systems. *The Journal of Web Science*, 2(1):1–18, 2016.

[15] S. Wang, S. Schlobach, and M. Klein. Concept drift and how to identify it. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(3):247–265, 2011.

[16] W. Zhang, G. Ding, L. Chen, C. Li, and C. Zhang. Generating virtual ratings from chinese reviews to augment online recommendations. *ACM Transactions on intelligent systems and technology (TIST)*, 4(1):9, 2013.