

# Designing Experts' Interactions with a Semi-Automated Document Tagging System

Sebastian Müller<sup>1</sup>, Beat Tödli<sup>1</sup>, Janine Vetsch<sup>2</sup>, Melanie Rickenmann<sup>1</sup>, Simon Haug<sup>2</sup>, Matthias Baldauf<sup>1</sup> and Peter Fröhlich<sup>3</sup>

<sup>1</sup>Eastern Switzerland University of Applied Sciences, IPM Institute for Information and Process Management, Rosenbergstrasse 59, 9001 St.Gallen, Switzerland

<sup>2</sup>Eastern Switzerland University of Applied Sciences, IPW Institute of Applied Nursing Science, Rosenbergstrasse 59, 9001 St.Gallen, Switzerland

<sup>3</sup>AIT Austrian Institute of Technology, Giefinggasse 4, 1120 Wien, Austria

## Abstract

For evidence-based care, nurses are dependent on latest scientific findings. However, screening and annotating respective publications is a time-consuming process for domain experts. In our ongoing work, we investigate how such experts can be supported by a (semi-)automated, machine-learning based recommendation system and how to design and integrate such a tool into their overall workflow. This paper outlines a current tag recommendation prototype and presents first results from an initial co-design workshop where we investigated the requirements of two nursing care researchers for such a system. The mockups created during the workshops lay the foundation for our goal of integrating a recommender system seamlessly into the researchers annotation process.

## Keywords

Tag recommendation, document annotation, expert interaction

## 1. Introduction

Currently, access to research-based knowledge in nursing practice is often difficult. Challenges include limited time resources and a lack of research skills among many nurses. There are thousands of new articles published in the broader field of health sciences every month, making it difficult to effectively and efficiently select the most relevant. *FIT-Nursing Care*<sup>1</sup>, an initiative of the Eastern Switzerland University of Applied Sciences, supports nurses in the selection process of the most relevant articles for their field of interest. Experts in the field of nursing sciences search for and annotate the most relevant publications to help nurses getting started with evidence-based nursing more quickly.

However, the manual search and annotation of comprehensive scientific publications by domain experts are lengthy and cost-intensive. In our ongoing work we investigate the application of AI techniques for the (semi-)automated identification, categorization and recommendation of relevant scientific studies. In particular, we aim at developing an assistive tagging approach (cf. [1]) for the nursing domain which supports nursing experts in annotating new publications

---

*AutomationXP22: Engaging with Automation, CHI'22, April 30, 2022, New Orleans, LA*



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

<sup>1</sup><https://www.fit-care.ch/>

in a user-friendly and engaging manner through a co-design research approach. While traditional tag recommending systems (e.g., for image or video hosting platforms) often aim at shifting efforts from domain specialists to general Web users while ensuring a decent tagging quality, our work aims at supporting domain experts' workflows by reducing their workload without questioning their expertise or even overruling their decisions. We target designing for the "sweet spot" of the recommender "that balances serving users effectively, while ensuring that the users have the control they desire" [2].

The project builds upon a large body of existing research in the field of tag recommendation (cf. [3, 4, 5, 1, 6]). While the majority of prior works focuses on tagging efficiency and accuracy, often on large sets multimedia data such as images and videos, aspects of user experience for experts and the integration in professional workflows have received less attention.

In the remainder of this position paper, we outline our current prototype of the tag recommendation system and introduce first results of an early co-design workshop on visualizing tag recommendations for domain experts.

## 2. Tag Recommendation Prototype

In our research, we design, develop and evaluate a Python-based tag recommendation system. The system calculates a probability for a pre-defined set of tags based on the title and abstract of a newly uploaded publication.

The data basis for our project consists of 1144 articles in the field of nursing care that have been manually tagged by the team that manages the platform. Each document was manually assigned with one, in rare cases two tags out of a defined list consisting of 23 tags. However, there is no equal distribution of tags among these articles. The range is between 177 articles for the most assigned tag and one article for the tag which was assigned the least. The recommender system was not trained to assign tags with less than 10 annotated documents, which was the case for four tags. Therefore the system calculates the possibility for a tag out of the remaining 19 tags.

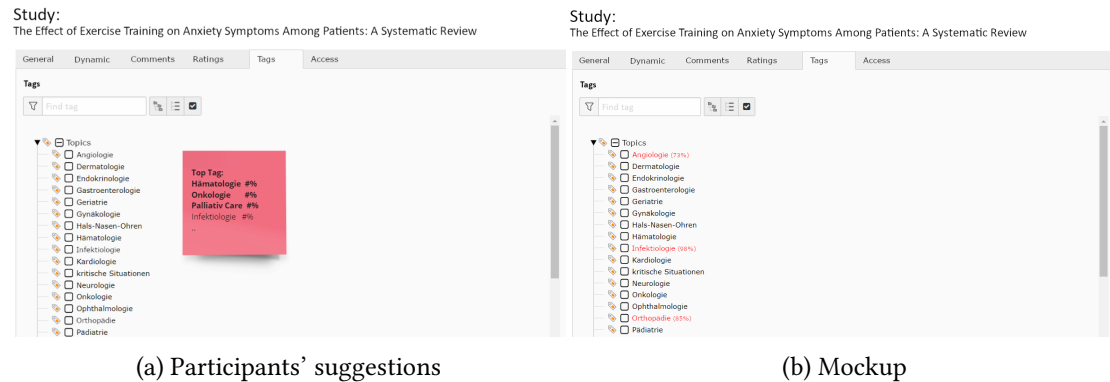
The recommendation system consists of a Python application that implements *scikit-learn tools*<sup>2</sup> and models from *Hugging Face*<sup>3</sup> Core of the system is the machine learning text classification tool that has been trained with 564 manually tagged articles. The validation data set consists of 172 documents, the remaining 408 documents belong to the test data set. Our classification tool finally calculates a percentage probability of how well each tag matches a newly uploaded document.

New articles are uploaded and annotated on a TYPO3-based website. The recommendation system interacts via a REST-API with the database, enabling it to retrieve newly uploaded documents and to store the recommended tags. Furthermore, we store the tags selected by the experts to allow evaluating and retraining the recommendation tool. The nursing care experts are able to override the tag recommendations of the system at any time.

---

<sup>2</sup><https://scikit-learn.org/stable/>

<sup>3</sup><https://huggingface.co/>



**Figure 1:** Examples of the visualization ideas developed during the workshop.

### 3. Initial Requirement and Design Workshop

While traditional and familiar recommendation systems such as the ones used by Amazon or similar online-shops aim on an increase in sales and recommend similar or complementary products to the ones in a users shopping history, we aim on increasing the efficiency of the user within a well-known process. The users are experts in annotating nursing related publications.

#### 3.1. Method

To co-design suitable user interfaces for the recommendation system with the actual users, we conducted a workshop with the two nursing experts responsible for tagging the documents. The workshop was conducted by a researcher who moderated and documented the workshop. During the workshop, ideas for visualizing the tag recommendations were created in a co-design approach. The workshop was conducted online utilizing MS Teams and the online whiteboard tool Miro for collaborative working. As materials, we prepared screenshots of the underlying TYPO3-based website and integrated them into the collaboration platform for design tasks. The workshop took about one hour. Based on the participants' ideas and suggestions, the involved researcher created more formal mockups. These were discussed with the two experts to resolve potential misunderstandings.

#### 3.2. First Workshop Results

Overall, the two participants appreciated a respective tagging support and expressed their interest in the performance of the final recommender. Both emphasized the importance of displaying the estimated matching accuracy for recommended tags and suggested to integrate this information by annotating the screenshot provided (Figure 1 (a)). In addition, they proposed an additional tag ordered by their possibility from high to low, including the matching accuracy calculated by the recommender for all tags.

Subsequently, the workshop discussion focused on the potential influence of such a presentation on the user. Since the users are nursing experts, the design of the recommender UI should support their decision-making process, however, not affect their decision. Experts should not

be misled into choosing a wrong tag if they get used to just selecting the one the system listed on top. Therefore the decision was made to keep the tags alphabetically ordered. User guidance could be provided by coloring the three best matching tags in red. The calculated matching accuracy is displayed next to the three colored tags (Figure 1 (b)). Other user interface discussions included different fonts, font markup (bold, italic, underlined or a combination of the three) and various colors.

Overall, the participants emphasized the importance of a user-friendly and efficient integration of the recommendation system into their familiar process. Both had already gained experience with command-line tool for automatically deriving document annotations. However, due to the complicated handling of the tool, it did not become part of their daily work routine. Therefore, both of them appreciate the implementation of the recommendation results directly into their already known TYPO3-based website.

The design and continuous evaluation of our recommender system are still ongoing. Nevertheless, we initiated and discussed a multistage process to build user trust in the system and optimize the tag assignment process. The ultimate goal would be to inform the user after uploading a new publication, that the system is sure about the tags to be assigned and there is no need for the expert to read and manually annotate the document. With measures of statistical dispersion, we would like to determine how accurately a tag can be recommended. The unequal distribution of documents over all tags also has to be considered. We imagine a notification after the upload, that the system is sufficiently confident about the annotation that the tag could be assigned automatically. Within this process too, all available information on why the system is confident enough should be made visible to the user. Starting with tags where the system has a high accuracy, this approach could be tested after the evaluation has been finished. Continuous retraining with the documents for which the system recommended a wrong tag would allow to extend the number of documents which the system can confidently annotate by itself.

## 4. Conclusion and Outlook

In this position paper, we presented our ongoing work on a semi-automated document tagging system for nursing experts. We introduced the current state of our tag recommendation prototype, our data basis and the system architecture. Furthermore, we outlined results of an initial co-design workshop, where two nursing care experts described their requirements for integrating the recommender into their current processes and developed ideas for visualizing tag recommendations.

This project will advance evidence-based nursing with modern data-driven text analysis methods and provides the foundation for better literature suggestions for nurse practitioners. To ensure the continuous use of the classification tool, ease of use and the seamless implementation into the current process is crucial.

## References

- [1] M. Wang, B. Ni, X.-S. Hua, T.-S. Chua, Assistive tagging, *ACM Computing Surveys* 44 (2012) 1–24. URL: <https://doi.org/10.1145/2333112.2333120>. doi:10.1145/2333112.2333120.

- [2] J. A. Konstan, J. Riedl, Recommender systems: from algorithms to user experience, *User Modeling and User-Adapted Interaction* 22 (2012) 101–123. URL: <https://doi.org/10.1007/s11257-011-9112-x>. doi:10.1007/s11257-011-9112-x.
- [3] F. M. Belém, J. M. Almeida, M. A. Gonçalves, A survey on tag recommendation methods, *Journal of the Association for Information Science and Technology* 68 (2016) 830–844. URL: <https://doi.org/10.1002/asi.23736>. doi:10.1002/asi.23736.
- [4] A. Dattolo, F. Ferrara, C. Tasso, The role of tags for recommendation: A survey, in: *3rd International Conference on Human System Interaction*, IEEE, 2010. URL: <https://doi.org/10.1109/hsi.2010.5514515>. doi:10.1109/hsi.2010.5514515.
- [5] S. Vairavasundaram, V. Varadharajan, I. Vairavasundaram, L. Ravi, Data mining-based tag recommendation system: an overview, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 5 (2015) 87–112. URL: <https://doi.org/10.1002/widm.1149>. doi:10.1002/widm.1149.
- [6] Z.-K. Zhang, T. Zhou, Y.-C. Zhang, Tag-aware recommender systems: A state-of-the-art survey, *Journal of Computer Science and Technology* 26 (2011) 767–777. URL: <https://doi.org/10.1007/s11390-011-0176-1>. doi:10.1007/s11390-011-0176-1.