FFG
Promoting Innovation.

Know-Center | Know-Center GmbH – Research Center for Data-Driven Business & Big Data Analytics

Program-line: COMET – Competence Centers for Excellent Technologies: K1-Zentren

SC-05, 08/2018 – 11/2018, single-firm

hgv
publishing services

COMET

UNDERSTANDING AND SUPPORTING E-BOOK SEARCH AND ANNOTATION VIA TAG RECOMMENDATION

TOGETHER WITH HGV, WE DEVELOPPED HYBRID TAG RECOMMENDATION SERVICES FOR E-BOOKS BY COMBINING TAGS FROM EDITORS AND AMAZON USERS.

The annotation of e-books with tags (i.e., keywords) is an important process for increasing the metadata quality and findability of e-books. In order to support editors in finding descriptive tags for e-books, tag recommendation algorithms have been developed by the research community. These algorithms analyze e-book metadata (e.g., title and description text) as well as past tag assignments for suggesting tags for a new e-book.
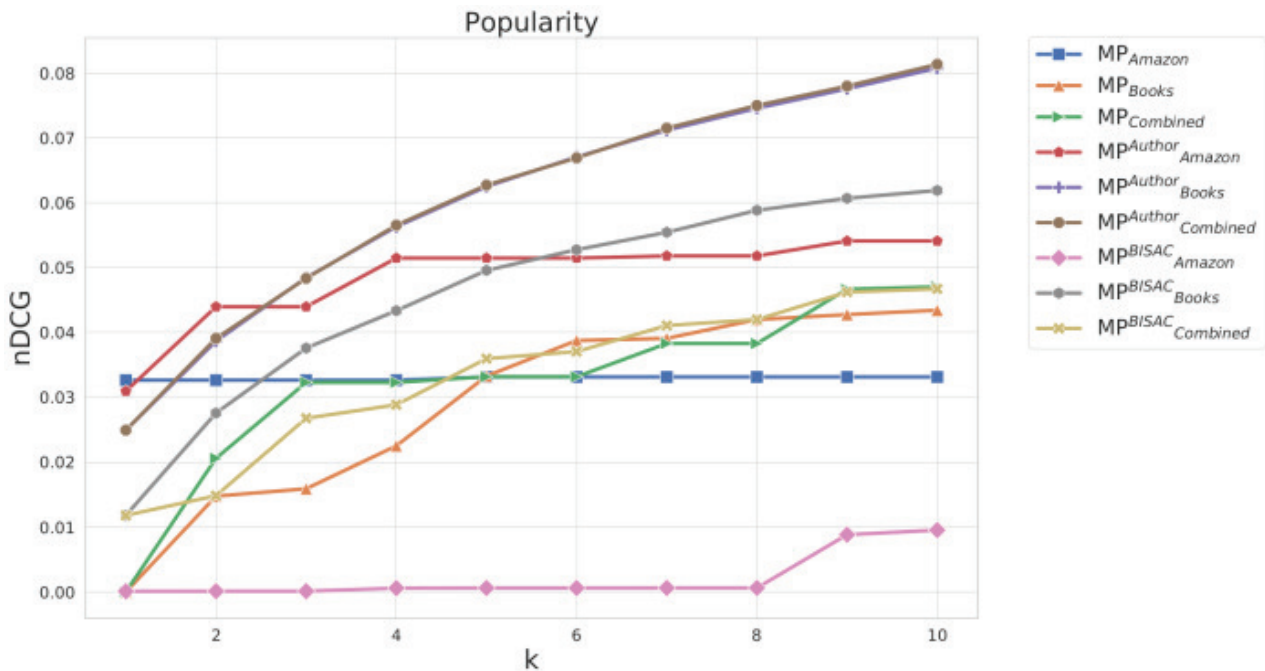
However, one issue of current tag recommendation strategies in the e-book sector is that they are solely based on the tag-vocabulary of the editors, which does not reflect the search-behavior of users in e-book systems such as Amazon.

To overcome this, together with our partner HGV (i.e., the Hanseatische Gesellschaft für Verlagsservice

mbH), we have developed a hybrid tag recommendation system for e-books, which analyzes search query logs from Amazon with the aim to bring together the vocabulary of the editors with that of the readers. Therefore, we combine two types of tag recommendation algorithms:

- **Popularity-based tag recommendations:** recommend the most frequently used tags in the system.
- **Similarity-based tag recommendations:** recommend the tags of similar books based on title and/or description text.

Example of our evaluation results, here for 9 popularity-based tag recommendation algorithms. Contextual information such as the author of the book contributes to a higher accuracy (here measured with nDCG) for the calculated recommendations (here, k is the number of recommended tags). By combining the vocabulary of the editors with that of Amazon users (e.g., the $MP^{Author}_{Combined}$ algorithm, brown line), we achieve an even higher accuracy than by solely relying on the vocabulary of the editors.

Impact and effects

We have evaluated our hybrid tag recommendation system using real-work e-book annotation data from HGV as well as Amazon search logs. We find that we can enhance tag recommendation accuracy by combining editor and Amazon tags using our hybrid tag recommendation system (see above Figure). In total, we have compared the performance of 19 tag recommendation approaches for e-books

Our tag recommendation system will be integrated by HGV into a live "keyword tool" and for future work, we also plan to extend our system with features of controllability, explainability and transparency. Furthermore, our results will be submitted to the well-known UMAP conference.

Business Development
Dipl. Ing. Matthias Traub
Business Area Head – Social Computing
Know-Center GmbH

T +43 (316) 873 - 30898
mtraub@know-center.at

Project coordination
DI Dr. Dominik Kowald
Senior Research – Social Computing
Know-Center GmbH

T +43 (316) 873 - 30846
dkowald@know-center.at

Project partner

- HGV Hanseatische Gesellschaft für Verlagsservices mbH
  Peter Langs peter.langs@hgv-online.de