

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



VERSTEHEN UND UNTERSTÜTZEN VON E-BOOK SUCHE UND ANNOTIERUNG MITTELS TAG RECOMMENDATION

ZUSAMMEN MIT HGV HABEN WIR TAG RECOMMENDATION ALGORITHMEN
ENTWICKELT, WELCHE TAGS VON EDITOREN UND AMAZON NUTZERN VEREINEN.

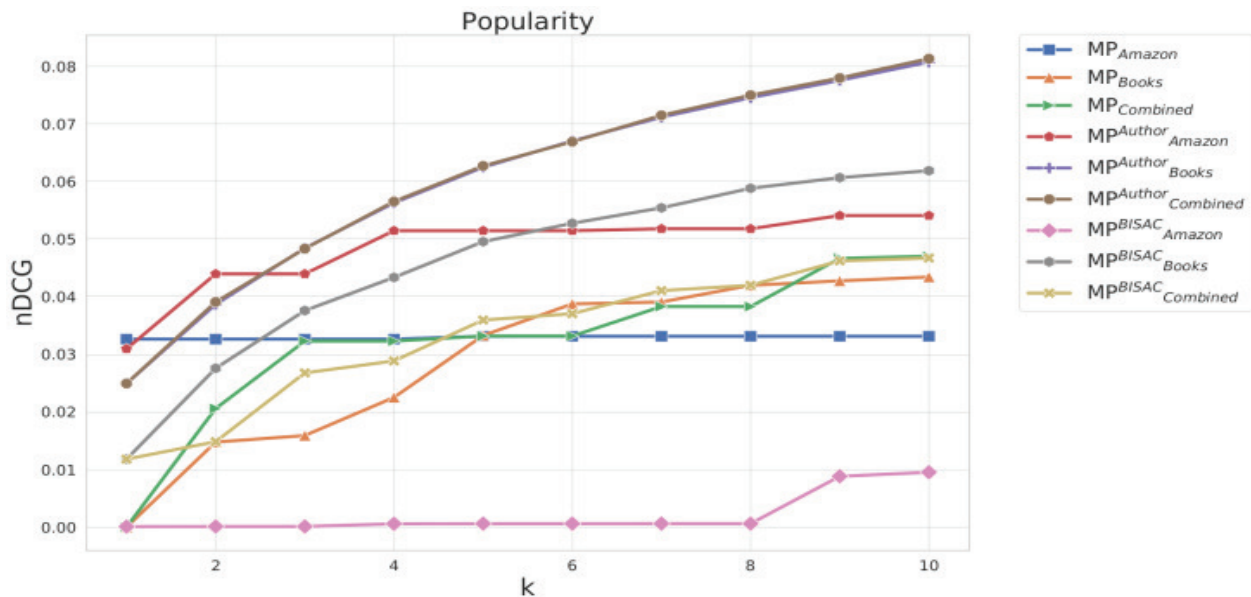
Die Annotierung von e-books mittels Tags (= Schlagwörter) ist ein wichtiger Prozess um die Metadaten-Qualität sowie die Auffindbarkeit von e-books zu erhöhen. Um Editoren bei der Findung aussagekräftiger Tags zu unterstützen, hat die Forschung Tag Recommendation Algorithmen entwickelt, welche Metadaten (z.B.: Titel und Beschreibungstext) sowie bereits vergebene Tags analysieren um Tags für ein neues e-book vorzuschlagen.

Die bisherigen Algorithmen haben allerdings eine Schwachstelle: da sie nur das Tag-Vokabular der Editoren betrachten, stimmen die Tag Recommendations oft nicht mit dem Vokabular der Nutzer von e-book Plattformen wie Amazon überein.

Um diesem Effekt entgegenzuwirken, haben wir zusammen mit unserem Partner HGV (die) ein

hybrides Tag Recommendation System entwickelt, welches auch Such-Logs von Amazon analysiert mit dem Ziel die Vokabulare der Editoren und Amazon Nutzer zu vereinen. Dazu kombinieren wir 2 Arten von Tag Recommendation Algorithmen:

- **Popularitäts-basierte Tag Recommendations:** diese schlagen die häufigsten Tags im System vor.
- **Ähnlichkeits-basierte Tag Recommendations:** diese schlagen Tags von ähnlichen Büchern vor (z.B.: basierend auf Titel und Beschreibungstext).



Beispiel unserer Evaluierungsergebnisse, hier anhand von 9 Popularitäts-basierter Algorithmen. Kontextuelle Informationen wie der Autor des Buches tragen zu einer höheren Genauigkeit (hier mit nDCG gemessen) für die kalkulierten Recommendations bei (hier ist k die Anzahl der Recommendations). Bei der Vereinigung der Vokabulare von Editoren und Amazon Nutzern (z.B., der in braun dargestellte $MP_{Author_Combined}$ Algorithmus) erzielen wir eine höhere Genauigkeit als nur bei der Betrachtung des Vokabulars der Editoren.

Wirkungen und Effekte

Wir haben unser hybrides Tag Recommendation System mittels e-book Tagging Daten von HGVS sowie mittels Daten aus Amazon Such-Logs evaluiert. Unsere ersten Ergebnisse haben gezeigt, dass wir die Genauigkeit von Tag Recommendations für e-books erhöhen können, wenn wir die Tags von Editoren mit den Suchbegriffen der Amazon Such-Logs mit Hilfe unseres hybriden Tag Recommendation Systems verbinden (siehe Graphik rechts). Insgesamt haben wir hier 19 Algorithmen-Kombinationen evaluiert. Unser Tag Recommendation System wird von HGVS in

ein Tool integriert, welches von e-book Editoren verwendet wird. Außerdem werden wir unsere Ergebnisse bei der bekannten UMAP Konferenz einreichen.

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

Projektpartner

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