

Effective Metadata for Social Book Search from a User Perspective

Hugo Huurdeman^{1,2}, Jaap Kamps^{1,2,3}, and Marijn Koolen¹

¹ Institute for Logic, Language and Computation, University of Amsterdam

² Archives and Information Studies, Faculty of Humanities, University of Amsterdam

³ ISLA, Faculty of Science, University of Amsterdam

Abstract. In this extended abstract we describe our participation in the INEX 2014 Interactive Social Book Search Track. In previous work, we have looked at the impact of professional and user-generated metadata in the context of book search, and compared these different categories of metadata in terms of retrieval effectiveness. Here, we take a different approach and study the use of professional and user-generated metadata of books in an interactive setting, and the effectivity of this metadata from a user perspective.

We compare the perceived usefulness of general descriptions, publication metadata, user reviews and tags in focused and open-ended search tasks, based on data gathered in the INEX Interactive Social Book Search Track. Furthermore, we take a tentative look at the actual use of different types of metadata over time in the aggregated search tasks.

Our preliminary findings in the surveyed tasks indicate that user reviews are generally perceived to be more useful than other types of metadata, and they are frequently mentioned in users' rationales for selecting books. Furthermore, we observe a varying usage frequency of traditional and user-generated metadata across time in the aggregated search tasks, providing initial indications that these types of metadata might be useful at different stages of a search task.

1 Introduction

In 2014, the Interactive Social Book Search (iSBS) task has been introduced. The goal of this task is “to investigate how book searchers deal with professional and user-generated content at different stages of the of the search process”¹.

The iSBS track uses the Amazon/LibraryThing collection with book descriptions for 1.5 million books, as crawled by the University of Duisburg-Essen in early 2009. This dataset contains both professional metadata and user-generated descriptions from Amazon and LibraryThing.

The task has introduced two interfaces for book search: a baseline interface, including search elements, results lists, and item details, and a ‘multistage’ interface, in which different features for different stages of a search have been introduced. Both interfaces feature a “book-bag”, to which users can add books

¹ <https://inex.mmci.uni-saarland.de/tracks/books/#interact>

when completing the assigned tasks². Two kinds of tasks are employed: a goal-oriented task, and a non-goal oriented task.

The goal-oriented task was the following: *Imagine you are looking for some interesting physics and mathematics books for a layperson. You have heard about the Feynman books but you have never really read anything in this area. You would also like to find an “interesting facts” sort of book on mathematics.*

The non-goal oriented task, without a predefined information need, was the following: *Imagine you are waiting to meet a friend in a coffee shop or pub or the airport or your office. While waiting, you come across this website and explore it looking for any book that you find interesting, or engaging or relevant...*

The participants in the iSBS experiment carried out one goal, and one non-goal oriented task, and were randomly assigned the baseline or multistage interface. A total of 41 users participated in the experiment, at Aalborg University Copenhagen, Edge Hill University, Humboldt University, and the University of Amsterdam. Of these participants, 19 used the baseline interface, and 22 users utilized the multistage interface. The resulting data, in the form of usage logs and questionnaires gathered in the experiment, was shared among the participating teams.

In this extended abstract, we take an initial look at the results, and specifically focus on the usefulness and effectivity of professional and user-generated metadata for books in an interactive setting, and compare the perceived usefulness of general descriptions, publication metadata, user reviews and tags across tasks.

2 Previous work

Previous work related to the iSBS track has been carried out in the INEX Interactive Retrieval Experiments (2004-2010) [7] and the Cultural Heritage in CLEF (CHiC) Interactive Track 2013 [8]. In these tracks, a standard procedure for collecting data was being used by participating research groups, including common topics and tasks, standardized search systems, document corpora and procedures. The system used for the iSBS track is a modified version of the one used for CHiC and is based on the Interactive IR evaluation platform developed by Hall and Toms [1], where different search engines and interfaces can be plugged into fully developed IIR framework that runs the entire user study [2].

The Interactive SBS Track complements the system-oriented Social Book Search Track [6]. Both tracks use the Amazon/LibraryThing collection and investigate the value of professional metadata and user-generated content. The system-oriented evaluation of the SBS Track has shown that retrieval effectiveness increases when systems include user-generated content for a broad range of tasks [4, 5]. These findings prompted the desire to study how book searchers use professional metadata and user-generated content.

² More information about the experiment and different interfaces is available in [3]

3 Results

The results described in this extended abstract are preliminary, and based on a relatively small dataset. Nevertheless, they can provide basic insights into book search from a user perspective.

We focus on four different types of displayed metadata in the baseline and multistage interface:

- general book *descriptions* which contain publisher-provided product descriptions and editorial reviews (available for 78% of all books).
- *publication* metadata, including information on publisher, price, number of pages and binding (available for all books)
- *reviews* metadata, containing up to 10 Amazon user reviews and include ratings (available for 45% of all books)
- user-provided *tags* from LibraryThing (LT), the displayed size based on how many LT users assigned each tag to the book (available for 83% of all books).

usefulness First of all, we look at the perceived usefulness of the general descriptions, publication metadata, user reviews and user tags.

Table 1: average perceived usefulness of metadata features

	<i>description</i>	<i>publication</i>	<i>reviews</i>	<i>tags</i>
baseline	4.0	3.0	3.8	2.8
multistage	4.3	2.6	3.9	2.9
<i>total</i>	<i>4.2</i>	<i>2.8</i>	<i>3.9</i>	<i>2.7</i>

In the post-questionnaire following the focused and open task of the experiment, users had to indicate how useful the metadata items about a book were on a rating scale of 1 (not at all useful) to 5 (extremely useful). Table 1 contains the results, differentiated by interface. It shows that for both interfaces, the general descriptions were deemed most useful, with an average rating of 4.2 out of 5. This is followed by the reviews, that were almost as useful, with an average rating of 3.9. Finally, both formal publication metadata, and tags were not considered highly useful by the participants.

Table 2: percentage of users not using metadata elements

	<i>description</i>	<i>publication</i>	<i>reviews</i>	<i>tags</i>
baseline	10.5%	39.5%	29.0%	44.7%
multistage	4.6%	34.1%	20.5%	18.2%
<i>total</i>	<i>7.5%</i>	<i>36.8%</i>	<i>24.7%</i>	<i>31.5%</i>

Secondly, the usage of each feature is of importance, since in the same survey question participants could also indicate that they did not use an element. Table 2 indicates that only a small minority (7.5%) of participants did not use the

general book descriptions. Again, it is followed by the reviews, that are not used by 25% of all participants. Finally, the tags and publication metadata are unused by more than 30% of all users.

So despite the fact that user reviews and ratings are only available for 45% of all books, they are actually considered almost as useful as general descriptive metadata by the participants in the experiment, and used more often than publication metadata and user tags. To contextualize this finding, we now look at the rationales of users for adding books to their bookbag selections.

contextualizing perceived usefulness In the post-task questionnaire, users could review their bookbag immediately after finishing their task, and were asked to describe why they selected these books. Users describe various reasons for adding books to their bookbags. A commonly mentioned selection criterion is related to the book’s title, subject and description. One user, for example, points out that “the first book, due to the title, seemed like a simpler read than the Feynman volumes”. Other users, surprisingly, mention that they chose a result because of it’s high position in the results list (being “closer to the top of the results”). Another reason for choosing particular books, in line with the results discussed in the previous section, is based on their ratings and reviews. For the open task, 7 out of 41 participants mention ratings and reviews as being important in their final book selections, while for the focused task another 7 participants explicitly mention ratings and reviews. For example, one user reports that based on the number of reviews, (s)he decided that a book is engaging: “I tried to look at that in the comments, if books had no comments I didn’t select them”. Reviews are also frequently used to choose among alternatives: “the first book has also a very good review”. Hence, the user-generated ratings and reviews seem to play an important role in the selection process in the context of the employed tasks.

usage statistics of metadata elements over time Based on the available usage logs, we can take a tentative look at the usage of the metadata available in the multistage experimental user interface. Here, we do not yet have statistics on the number of times the general book descriptions were seen, since they appear by default, but we can get insights on the usage of the publication, review and tag metadata.

Table 3: use of metadata features in the multistage interface (n=22)

	<i>begin</i>	<i>middle</i>	<i>end</i>	<i>total</i>
publication	5	17	10	32
reviews	8	20	24	52
tags	6	14	20	40

Table 3 shows the number of times publication metadata, reviews and tags were explicitly selected in the focused and open-ended search task. While these numbers, upon a total number of 22 users across two tasks are quite low, they

do indicate a higher utilization of user-generated reviews in the combined tasks, used in total 52 times by 22 users.

Furthermore, the use of these metadata features over time can be surveyed. Therefore we divided the total task time into three segments, *begin*, *middle* and *end*, each taking up one third of the task process. From this data, we see that the number of times these data elements are used varies over time: the publication metadata is used more in the middle, and the reviews and tags are used more frequently in the end. Due to the low number of data points, we cannot derive strong conclusions, but the use of metadata features over time could be interesting to assess in future studies.

4 Conclusion

The preliminary results documented in this extended abstract provide indications that besides professional metadata, also user-generated metadata might be useful for book search, from the perspective of users. In terms of perceived usefulness, and usage, especially user reviews seem to aid users in determining their book selections for different types of tasks, even though they are not available for all books. While the findings here are based on the first iteration of the Interactive Social Book Search task, they provide handles for future research, and we intend to extend these findings in future work in the context of Social Book Search.

Acknowledgments This research was supported by the Netherlands Organization for Scientific Research (NWO projects # 612.066.513, 639.072.601, and 640.005.001) and by the European Community's Seventh Framework Program (FP7 2007/2013, Grant Agreement 270404).

Bibliography

1. M. Hall and E. Toms. Building a common framework for iir evaluation. In P. Forner, H. Müller, R. Paredes, P. Rosso, and B. Stein, editors, *Information Access Evaluation. Multilinguality, Multimodality, and Visualization*, volume 8138 of *Lecture Notes in Computer Science*, pages 17–28. Springer Berlin Heidelberg, 2013. ISBN 978-3-642-40801-4. doi: 10.1007/978-3-642-40802-1_3.
2. M. Hall, S. Katsaris, and E. Toms. A pluggable interactive ir evaluation work-bench. In *European Workshop on Human-Computer Interaction and Information Retrieval*, pages 35–38, 2013.
3. M. Hall, H. Huurdeman, M. Koolen, M. Skov, and D. Walsh. Overview of the INEX 2014 interactive social book search track. In L. Cappellato, N. Ferro, M. Halvey, and W. Kraaij, editors, *CLEF 2014 Labs and Workshops, Notebook Papers*, CEUR Workshop Proceedings (CEUR-WS.org), 2014.

4. M. Koolen. "user reviews in the search index? that'll never work!". In M. de Rijke, T. Kenter, A. P. de Vries, C. Zhai, F. de Jong, K. Radinsky, and K. Hofmann, editors, *ECIR*, volume 8416 of *Lecture Notes in Computer Science*, pages 323–334. Springer, 2014. ISBN 978-3-319-06027-9.
5. M. Koolen, J. Kamps, and G. Kazai. Social Book Search: The Impact of Professional and User-Generated Content on Book Suggestions. In *CIKM 2012*. ACM, 2012.
6. M. Koolen, G. Kazai, M. Preminger, and A. Doucet. Overview of the INEX 2013 social book search track. In *CLEF 2013 Evaluation Labs and Workshop, Online Working Notes*, 2013.
7. R. Nordlie and N. Pharo. Seven years of inex interactive retrieval experiments - lessons and challenges. In T. Catarci, P. Forner, D. Hiemstra, A. Peñas, and G. Santucci, editors, *CLEF*, volume 7488 of *Lecture Notes in Computer Science*, pages 13–23. Springer, 2012. ISBN 978-3-642-33246-3.
8. V. Petras, T. Bogers, E. Toms, M. Hall, J. Savoy, P. Malak, A. Pawowski, N. Ferro, and I. Masiero. Cultural heritage in clef (chic) 2013. In P. Forner, H. Müller, R. Paredes, P. Rosso, and B. Stein, editors, *Information Access Evaluation. Multilinguality, Multimodality, and Visualization*, volume 8138 of *Lecture Notes in Computer Science*, pages 192–211. Springer Berlin Heidelberg, 2013. ISBN 978-3-642-40801-4. doi: 10.1007/978-3-642-40802-1_23.