Overview of the SBS 2016 Interactive Track

Gäde, M.; Hall, M.; Huurdeman, H.; Kamps, J.; Koolen, M.; Skov, M.; Bogers, T.; Walsh, D.

Published in:
CEUR Workshop Proceedings

Overview of the SBS 2016 Interactive Track

Maria Gäde[1], Mark Hall[2], Hugo Huurdeman[3], Jaap Kamps[3], Marijn Koolen[3], Mette Skov[4], Toine Bogers[4], and David Walsh[2]

[1] Humboldt University Berlin, Germany
maria.gaede@ibi.hu-berlin.de
[2] Edge Hill University, United Kingdom
{mark.hall,david.walsh}@edgehill.ac.uk
[3] University of Amsterdam, Netherlands
{huurdeman,kamps,marijn.koolen}@uva.nl
[4] Aalborg University, Denmark
{skov,toine}@hum.aau.dk

Abstract. Users looking for books online are confronted with both professional meta-data and user-generated content. The goal of the Interactive Social Book Search Track was to investigate how users used these two sources of information, when looking for books in a leisure context. To this end, participants recruited by seven teams performed two main tasks and one optional task using a user-interface that supports multiple search stages.

1 Introduction

The goal of the Interactive Social Book Search (ISBS) task is to investigate how book searchers use professional metadata and user-generated content at different stages of the search process. The purpose of this task is to gauge user interaction and user experience in social book search by observing user activity with a large collection of rich book descriptions under controlled and simulated conditions, aiming for as much "real-life" experiences intruding into the experimentation. The output will be a rich data set that includes both user profiles, selected individual differences (such as a motivation to explore), a log of user interactivity, and a structured set of questions about the experience.

The Interactive Social Book Search (ISBS) Task started in 2014 as a merge of the INEX Social Book Search (SBS, [8–10]) track and the Interactive task of CHiC [13, 16]. The aim was to augment the other Social Book Search tracks with a user-focused methodology. This will make it possible to study how the addition of opinionated descriptions and user-supplied tags allows users to search and select books based on more diverse criteria. User reviews may reveal information about plot, themes, characters, writing style, text density, comprehensiveness and other aspects that are not described by professional metadata.

This additional information is subjective and personal, and opens up opportunities to aid users in searching for books in different ways that go beyond the traditional editorial metadata based search scenarios, such as known-item and

subject search. For example, readers use many more aspects of books to help them decide which book to read next [14], such as how engaging, fun, educational or well-written a book is. In addition, readers leave a trail of rich information about themselves in the form online profiles which contain personal catalogues of the books they have read or want to read, personally assigned tags and ratings for those books and social network connections to other readers. This results in a search task that may require a different model than pure search [7] or pure recommendation.

In particular, the focus is on complex goal-oriented tasks as well as non-goal oriented tasks. For traditional tasks such as known-item search, there are effective search systems based on access points via formal metadata (i.e. book title, author name, publisher, year, etc). But even here user reviews and tags may prove to have an important role.

The long-term goal of the task is investigate user behaviour through a range of user tasks and interfaces and to identify the role of different types of metadata for different stages in the book search process. In order to take a step towards this goal, the overall experiment structure, data-set, and interface were kept the same as in 2015 [3]. Minor modifications were made to the experiment structure and interface to fix specific issues that had been identified in the ISBS 2015 track.

For the Interactive task, the main research question is:

RQ : How do searchers use professional metadata and user-generated content in book search?

This can be broken down into a few more specific questions:

RQ1 How should the UI combine professional and user-generated information?
RQ2 How should the UI adapt itself as the user progresses through their search task?

In this paper, we report on the setup and the results of the ISBS track 2016. Section 2 lists the participating teams. The experimental setup of the task is discussed in detail in Section 3 and the results in Section 4. We close in Section 5 with a summary and plans for 2017.

2 Participating Teams

Table 1 shows the institutions that participated in this track and the number of participants each institution recruited. Overall the number of participants recruited is lower than in 2015 (192 participants) and the majority of institutions reported problems recruiting participants. However, the distribution of participants across institutions is more balanced than in previous years, where one or two institutions provided significantly more participants than all other institutions.

Table 1. Overview of the participating teams and number of users per team

Institution	# Participants
Aalborg University	14
Edge Hill University	12
Humboldt University, Berlin	7
Manchester Metropolitan University	13
Oslo & Akershus University College of Applied Sciences	15
Peking University, China and Stockholm University, Sweden	29
University of Duisburg-Essen	21
Total	111

3 Experimental Setup

In order to make the results comparable to the 2015 run of the track, the 2016 Interactive Social Book Search (ISBS) track used fundamentally the same methodology as in the 2015. Minor modifications were made to the experiment structure to address issues around participant time requirements and variety of tasks. A small number of bugs were also fixed in the *multi-stage* interface.

3.1 Data-Set

The track builds on the INEX Amazon/LibraryThing (A/LT) collection [1], which contains 1.5 million book descriptions from Amazon, enriched with content from LT. This collection contains both professional metadata and user-generated content.[5]

Each book is identified by an ISBN. Since different editions of the same work have different ISBNs, there can be multiple records for a single intellectual work. The records contain title information as well as a Dewey Decimal Classification (DDC) code (for 61% of the books) and category and subject information supplied by Amazon. We note that for a sample of Amazon records the subject descriptors are noisy, with a number of inappropriately assigned descriptors that seem unrelated to the books.

Each book record is originally represented as an XML file with fields like *isbn*, *title*, *author*, *publisher*, *dimensions*, *numberofpages* and *publicationdate*. Curated metadata comes in the form of a Dewey Decimal Classification in the *dewey* field, Amazon subject headings in the *subject* field, and Amazon category labels in the *browseNode* fields. The social metadata from Amazon and LT is stored in the *tag*, *rating*, and *review* fields.

[5] This collection is a subset of a larger collection of 2.8 million description. The subset contains all book descriptions which have a cover image.

The records are pre-processed and indexed into an Elastic Search instance, which the frontend interface then uses. Both the professional metadata and user-generated content are indexed. For indexing and retrieval the default parameters are used, which means stopwords are removed, but no stemming is performed. The Dewey Decimal Classification numbers are replaced by their natural language description. That is, the DDC number 573 is replaced by the descriptor *Physical anthropology*. User tags from LibraryThing are indexed both as text strings, such that complex terms are broken down into individual terms (e.g. *physical anthropology* is indexed as *physical* and *anthropology*) and as non-analyzed terms, which leaves complex terms intact and is used for faceted search.

3.2 Tasks

The 2016 ISBS track uses the same main tasks as in 2015, an open-ended *non-goal* task, and a more focused *goal-oriented* task. Additionally in 2016 participants were given the choice of undertaking an additional optional task taken from the list of tasks used in the 2016 SBS Suggestion track. The aim of this was to give future SBS Suggestion tracks some actual user behaviour data.

The training task developed for the 2015 track was used to ensure that participants are familiar with all the functions offered by the *multi-stage* interface. The queries and topics used in the training task were chosen so as not to overlap with the *goal-oriented* or additional task. However, a potential influence on the *non-goal* task cannot be ruled out. For all tasks, participants were asked to describe their motivation for particular book selections in the book-bag.

The *goal-oriented* task contains five sub-tasks ensuring that participants spend enough time on finding relevant books. While the first sub-task defines a clear goal, the other sub-tasks are more open giving the user enough room to interact with and the available content and met-data options. The following instruction text was provided to participants:

> Imagine you participate in an experiment at a desert-island for one month. There will be no people, no TV, radio or other distraction. The only things you are allowed to take with you are 5 books. Please search for and add 5 books to your book-bag that you would want to read during your stay at the desert-island:
> - Select one book about surviving on a desert island
> - Select one book that will teach you something new
> - Select one book about one of your personal hobbies or interests
> - Select one book that is highly recommended by other users (based on user ratings and reviews)
> - Select one book for fun
>
> Please add a note (in the book-bag) explaining why you selected each of the five books.

The *non-goal* task is a repeat of the 2015 *non-goal* task [3] and was originally developed based on the open-ended task used in the iCHiC at CLEF 2013 [15] and ISBS at CLEF 2014 [4]. The aim of this task is to investigate how users interact with the system when they have no pre-defined goal in a more exploratory search context. It also allows the participants to bring their own goals or subtasks to the experiment in line with the "simulated work task" idea [2]. The following instruction text was provided to participants:

> 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. Explore anything you wish until you are completely and utterly bored. When you find something interesting, add it to the book-bag. Please add a note (in the book-bag) explaining why you selected each of the books.

The *additional* task allowed the the participant to undertake an additional task with a known-subject focus. Participants were shown one of the following eight tasks:

South Africa You're interested in non-fiction history books on the background to and the actual time of the Boer War in South Africa. Search the collection using any of the interface features to find at least one book that meets these criteria.

Elizabethan You enjoy books covering the Elizabethan era (English monarch, 1558-1603) and have read books by George Garrett, e.g. "The Death of the Fox" and "The Succession". You'd like to find other books in this theme that look engaging, either fiction or non-fiction. Search the collection using any of the interface features to find at least one book that meets these criteria.

Communication For a university project you're looking for books about communication on the internet, that are well-written and well-researched. The books could be about how communication in general has changed or how language has changed. Search the collection using any of the interface features to find at least one book that meets these criteria.

Painters You've read a few books about specific paintings, i.e. "Portrait of Dr. Gachet" by Cynthia Saltzman, about the painting with the same name, by Vincent Van Gogh, "Strapless" by Deborah Davis, about John Sargent's painting of Virginie Gautreau and "The Lost Painting" by Jonathan Harr, about a missing painting by Caravaggio. You'd like to find more books about specific paintings. Search the collection using any of the interface features to find at least one book that meets these criteria.

Complex Mystery You like reading mystery novels (including thriller, crime and suspense) that have very complex plots. Search the collection using any of the interface features to find at least one book that meets these criteria.

Astronomy You're interested in astronomy and astrophysics, but have only beginners knowledge on these subject. You're looking for good beginners

books on these two subjects. Search the collection using any of the interface features to find at least two books that meet these criteria and together cover both astronomy and astrophysics.

Romance Mystery You're interested in mystery and thriller novels that also have romance. Search the collection using any of the interface features to find at least one book that meets these criteria.

French Revoluation You've read "For Whom the Bell Tolls", which is set during the Spanish Civil War and would now like to read a good fiction book set during the French Revolution. Search the collection using any of the interface features to find a book that meets these criteria.

The tasks were taken from the LibraryThing discussion forums as "real" tasks and were selected in order to create a data-set of user interactions that could be used in future instances of the SBS Suggestion track, as the same tasks are used in that track as well.

3.3 Experiment Structure

The experiment was conducted using the SPIRE system[6] [5], using the flow shown in Figure 1. When a participant started the experiment, they were automatically assigned either to the *non-goal* or *goal-oriented* task. The SPIRE system automatically balances allocations to ensure an even distribution of participants to the two tasks. This is a change from the 2015 task, where participants undertook both tasks. The reason for this change is that in 2015 there was significant participant feedback indicating that overall the experiment was too long. As the 2015 data showed no ordering influences, the change is unlikely to have any significant impact on the comparability of the two data-sets.

A further change from the 2015 experiment is that each participating institution was allocated their own instance of the experiment. This ensured that the participant allocation was balanced for each institution, not only for the experiment overall.

Participant responses were collected in the following five steps using a selection of questionnaires:

- *Consent* – all participants had to confirm that they understood the tasks they would be asked to undertake and the types of data collected in the experiment. Participants also specified who had recruited them;
- *Demographics* – the following factors were acquired in order to characterise the participants: gender, age, achieved education level, current education level, and employment status;
- *Culture* – to quantify language and cultural influences, the following factors were collected: country of birth, country of residence, mother tongue, primary language spoken at home, languages used to search the web;

[6] Based on the Experiment Support System – https://bitbucket.org/mhall/experiment-support-system

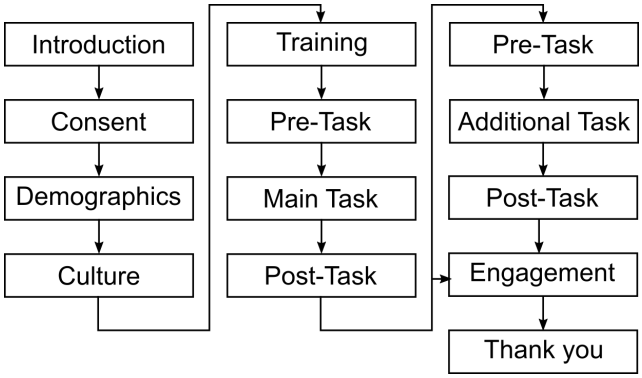

Fig. 1. The path participants took through the experiment. The SPIRE system automatically balanced allocation to either the *non-goal* or *goal-oriented* task. After the first *Post-Task* stage, participants were asked whether they had time to do another task and if not, were taken directly to the *Engagement* stage. No data was acquired in the *Introduction*, *Pre-Task*, and *Thank you* steps.

- *Post-Task* – in the post task questions, participants were asked to judge how useful each of the interface components and meta-data parts that they had used in the task were, using 5-point Likert-like scales;
- *Engagement* – after participants had completed both tasks, they were asked to complete O'Brien et al.'s [12] engagement scale.

3.4 System and Interfaces

As stated earlier, in 2016 only the *multi-stage* interface developed for ISBS 2015 [3] was used. It is built using the PyIRE[7] workbench, which provides the required functionality for creating interactive IR interfaces and logging all interactions between the participants and the system. This includes any queries they enter, the books shown for the queries, pagination, facets selected, books viewed in detail, metadata facets viewed, books added to the book-bag, and books removed from the book-bag. All log-data is automatically timestamped and linked to the participant and task.

The *multi-stage* interface used an IR backend implemented using ElasticSearch[8], which provided free-text search, faceted search, and access to the individual books complete metadata.

The aim of the *multi-stage* interface is to support users by taking the different stages of the search process into account. The idea behind the *multi-stage* interface design is supported by two theoretical components.

[7] Python interactive Information Retrieval Evaluation workbench – https://bitbucket.org/mhall/pyire

[8] ElasticSearch – http://www.elasticsearch.org/

Firstly, several information search process models look at stages in the search process. A well-known example is Kuhlthau [11], who discovered "common patterns in users' experience" during task performance. She developed a model consisting of six stages, which describe users' evolving thoughts, feelings and actions in the context of complex tasks. Vakkari [17] later summarized Kuhlthau's stages into three categories (pre-focus, focus formulation, and post-focus), and points to the types of information searched for in the different stages.

The multi-stage search interface constructed for iSBS was inspired by [17]. It includes three distinct panels, potentially supporting different stages: *browse*, in which users can explore categories of books, *search*, supporting in-depth searching, and *book-bag*, in which users can review and refine their book-bag selections.

Secondly, when designing a new search interface for social book search it has also been relevant to look more specifically at the process of choosing a book to read. A model of decision stages in book selection [14] identifies the following decision stages: browse category, selecting, judging, sampling, and sustained reading. This work supports the need for a user interface that takes the different search and decision stages into account. However, the different stages in [14] closely relate to a specific full text digital library, and therefore the model was not applicable to the present collection.

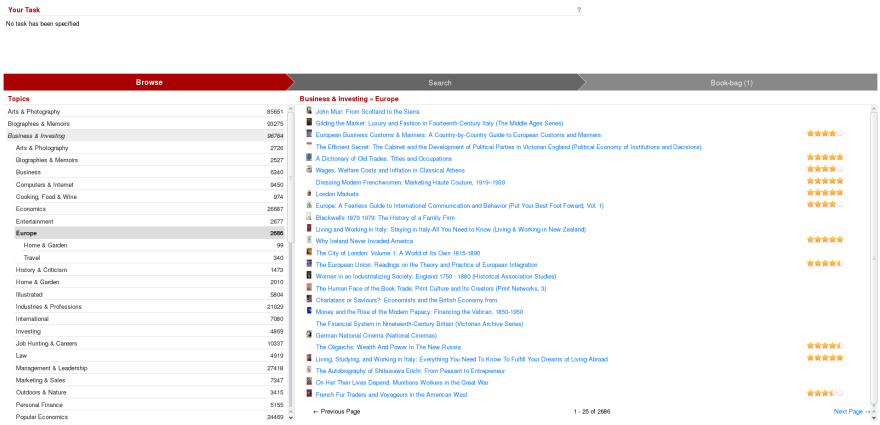

Fig. 2. *Multistage* interface – *Browse* view with the topic tree on the left and the dense list of books for the currently selected on the right.

When the *multi-stage* interface first loads, participants are shown the *browse* stage (fig. 2), which is aimed at supporting the initial exploration of the dataset. The main feature to support the free exploration is the hierarchy browsing component on the left, which shows a hierarchical tree of Amazon subject classifications. This was generated using the algorithm described in [6], which uses the relative frequencies of the subjects to arrange them into the tree-structure with the most-frequent subjects at the top of the tree. The search result list

is designed to be more compact to allow the user to browse books quickly and shows only the book's title and aggregate ratings (if available). Clicking on the book title showed a popup window with the book's full meta-data using the same layout and content as used in the *baseline* interface's search result list.

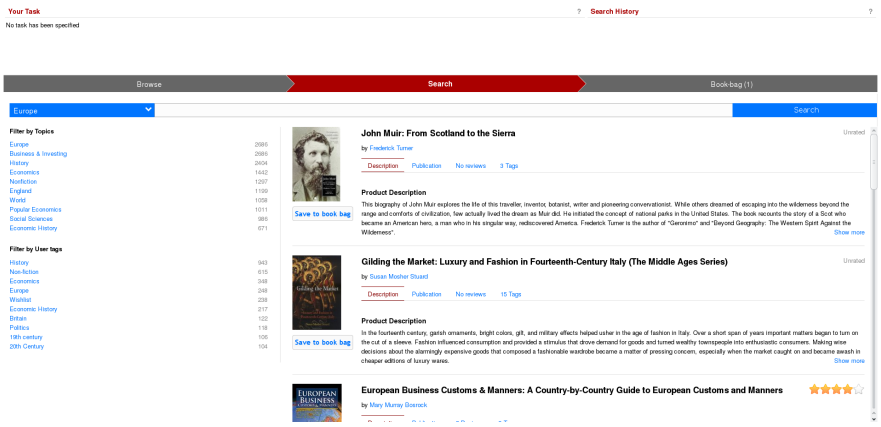

Fig. 3. *Multistage* interface – *Search* view showing a standard faceted-search interface.

Participants switched to the *search* stage by clicking on the "Search" section in the gray bar at the top. The *search* stage (fig. 3 uses the same interface as the *baseline* with only two differences. The first is that as the book-bag is a separate stage, it is not shown on the *search* stage interface itself. The second is that if the participants select a topic in the *browse* stage, this topic is pre-selected as a filter for any queries in the blow box to the left of the search box. Participants can click on that box to see a drop-down menu of the selected topic and its parent topics. Participants can select a higher-level topic to widen their search.

The final stage is the *book-bag* shown in Figure 4, where participants review the books they have collected and can provide the notes for each book. For each book, four buttons were provided that allowed the user to search for similar books by title, author, topic, and user tags. The similar books are shown on the right using the same compact layout as in the *browse* stage. As in the *browse* stage, clicking on a book in that list shows a popup window with the book's details.

3.5 Participants

A total of 111 participants were recruited (see Table 1), 51 female and 60 male. 65 were between 18 and 25, 29 between 26 and 35, 16 between 36 and 45, 1 between 46 and 55. 31 were in employment, 2 unemployed, 77 were students and 1 selected *other*. Participants came from 15 different countries (country of

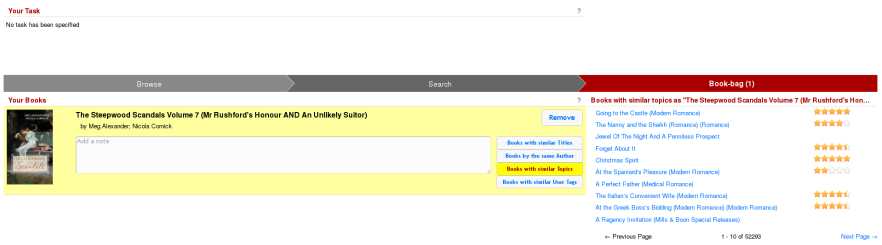

Fig. 4. *Multistage* interface – *Book-bag* view with the books the user has selected in the main area. For each book a free-text annotation area is provided. Selecting one of the buttons shows a list of similar books for the selected aspect (title, authors, subject, user-tags).

birth), with a wide geographical spread including China (27), UK (25), Norway (14), Germany (13), India (11), and Denmark (10). Participants were residents of 8 different countries that mirrored the participating team's locations (China, UK, Germany, Norway, Denmark, India, Taiwan). Participants mother tongues included Chines, English, German, Norwegian, Danish, and 8 others.

56 participants were allocated to the *goal-oriented* task, while 55 undertook the *non-goal* task. 22 participants only undertook the required task, while 89 undertook both the required an *additional* tasks. For the *additional* task, Table 2 shows the distribution of participants to the 8 additional tasks.

Table 2. Distribution of participants to the *additional* task.

Task	**# Participants**
Astronomy	12
Communication	9
Complex Mystery	9
Elizabethan	12
French Revoluation	11
Painters	11
Romance Mystery	13
South Africa	12

3.6 Procedure

Participants were invited by the individual teams, either using e-mail or by recruiting students from a lecture or lab. Where participants were invited by e-mail, the e-mail contained a link to the online experiment, which would open in the participant's browser. Where participants were recruited in a lecture or lab, the experiment URL was distributed using e-learning platforms. The following

browsers and operating systems had been tested: Windows, OS X, Linux using Internet Explorer, Chrome, Mozilla Firefox, and Safari. The only difference between browsers was that some of the graphical refinements such as shadows are not supported on Internet Explorer and fall back to a simpler line-based display.

After participants had completed the experiment as outlined above (3.3), they were provided with additional information on the tasks they had completed and with contact information, should they wish to learn more about the experiment. Where participants that completed the experiment in a lab, teams were able to conduct their own post-experiment process, which mostly focused on gathering additional feedback on the system from the participants.

4 Results

Based on the participant responses and log data we have aggregated summary statistics for a number of basic performance metrics.

Session length was measured automatically using JavaScript and stored with the participants' responses. Table 3 shows median and inter-quartile ranges for all tasks. With the exception of the *french revolution* additional task, it is clear that the known-subject additional tasks are significantly easier and faster to complete. Why the *french revolution* tasks has such a higher time requirement requires further study. For the *non-goal* and *goal-oriented* tasks the task duration matches the results from the 2015 data, indicating that the experiment is stable and the data comparable.

Table 3. Session lengths for the tasks. Times are in minutes:seconds and are reported median (inter-quartile range).

Task	Time
Non-goal	7:38min (9:38min)
Goal-oriented	12:20min (14:28min)
South Africa	4:51min (3:51min)
Elizabethan	5:48min (3min)
Communication	4:58min (2:47min)
Painters	4:42min (4:07min)
Complex Mystery	3:36min (4:22min)
Astronomy	5:12min (5:54min)
Romance Mystery	2:21min (3:15min)
French Revolution	6:36min (7:16min)

Number of queries was extracted from the log-data. Queries could be issued by typing keywords into the search box or by clicking on a meta-data field to

search for other books with that meta-data field value. Both types of query have been aggregated and Table 4 shows the number of queries for each task. There is a clear difference between the *non-goal* and the *goal-oriented* task. On the *additional* tasks, more analysis is needed to investigate why the *south africa*, *complex mystery*, and *romance mystery* tasks have such low values for the number of queries. However, for the other *additional* tasks, it is clear that as far as complexity of the task and number of queries required, they lie between the *non-goal* and *goal-oriented* tasks.

Table 4. Number of queries executed. Numbers are reported median (inter-quartile range).

Task	**# Queries**
Non-goal	1 (5)
Goal-oriented	5 (9)
South Africa	1 (1.5)
Elizabethan	3.5 (3.25)
Communication	4 (4)
Painters	4 (4)
Complex Mystery	0 (0)
Astronomy	3 (3.25)
Romance Mystery	1 (2)
French Revolution	4 (4.5)

Number of books collected was extracted from the log-data. Participants collected those books that they felt were of use to them. The numbers reported in Table 5 are based on the number of books participants had in their book-bag when they completed the session, not the total number of books collected over the course of their session, as participants could always remove books from their book-bag in the course of the session.

The number of books collected is clearly determined by the task, although the *elizabethan* and *south africa*, and *communication* tasks have different potential interpretations on how many books are needed to satisfy the task. As is to be expected, the *non-goal* task shows the highest variation in the number of books collected, as participants were completely free to define what "success" meant for them in that task.

5 Conclusions and Plans

This was the third year of the SBS Lab Interactive Track. Our goal remains to investigate how users deal with professional metadata and user-generated content when searching for books. The track makes use of a large collection of

Table 5. Number of books collected. Numbers are reported median (interquartile range).

Task	# Books
Non-goal	3 (3)
Goal-oriented	5 (0)
South Africa	2.5 (2)
Elizabethan	2.25 (2.25)
Communication	2 (2)
Painters	1 (2)
Complex Mystery	1 (0)
Astronomy	2 (1)
Romance Mystery	2 (1)
French Revolution	1 (0)

book descriptions from Amazon and LibraryThing with a mixture of professional metadata in the form of subject descriptors and classification codes and user-generated content in the form of user reviews, tags and ratings. Because search processes often consist of multiple stages, we developed an interface to identify and analyse these different stages. The *multistage* interface provides three components. The first provides a broad overview of the collection, the second allows the user to look at search results in a more detailed view, and the final part allows the user to directly compare selected books in great detail.

The overview results clearly indicate that the results are in-line with the 2015 results, making it possible to handle both data-sets as one meta-data set. The inclusion of the optional additional tasks has also shown how the kind of real-world tasks differ in behaviour from the *goal-oriented* task used as the baseline.

Plans for the next year are to update the data-set to be more current and to investigate more tasks and interfaces as well as interface components.

Bibliography

1. T. Beckers, N. Fuhr, N. Pharo, R. Nordlie, and K. N. Fachry. Overview and results of the inex 2009 interactive track. In M. Lalmas, J. M. Jose, A. Rauber, F. Sebastiani, and I. Frommholz, editors, *ECDL*, volume 6273 of *Lecture Notes in Computer Science*, pages 409–412. Springer, 2010. ISBN 978-3-642-15463-8.
2. P. Borlund and P. Ingwersen. The development of a method for the evaluation of interactive information retrieval systems. *Journal of documentation*, 53(3):225–250, 1997.
3. M. Gde, M. M. Hall, H. Huurdeman, J. Kamps, M. Koolen, M. Skov, E. Toms, and D. Walsh. Overview of the sbs 2015 interactive track. In L. Cappellato, N. Ferro, G. J. F. Jones, and E. San Juan, editors, *CLEF2015 Working Notes*. CEUR Workshop Proceedings, 2015. URL http://ceur-ws.org/Vol-1391/78-CR.pdf.

4. M. Hall, H. Huurdeman, M. Koolen, M. Skov, and D. Walsh. Overview of the INEX 2014 interactive social book search track. In E. Kanoulas, M. Lupu, P. Clough, M. Sanderson, M. Hall, A. Hanbury, and E. Toms, editors, *Information Access Evaluation. Multilinguality, Multimodality, and Interaction*, volume 8685 of *Lecture Notes in Computer Science*, pages 212–228. Springer International Publishing, 2014. ISBN 978-3-319-11381-4. doi: 10.1007/978-3-319-11382-1_19. URL http://dx.doi.org/10.1007/978-3-319-11382-1_19.
5. M. M. Hall and E. Toms. Building a common framework for iir evaluation. In *CLEF 2013 - Information Access Evaluation. Multilinguality, Multimodality, and Visualization*, pages 17–28, 2013. doi: 10.1007/978-3-642-40802-1_3.
6. M. M. Hall, S. Fernando, P. Clough, A. Soroa, E. Agirre, and M. Stevenson. Evaluating hierarchical organisation structures for exploring digital libraries. *Information Retrieval*, 17(4):351–379, 2014. ISSN 1386-4564. doi: 10.1007/s10791-014-9242-y. URL http://dx.doi.org/10.1007/s10791-014-9242-y.
7. M. Koolen, J. Kamps, and G. Kazai. Social Book Search: The Impact of Professional and User-Generated Content on Book Suggestions. In *Proceedings of the International Conference on Information and Knowledge Management (CIKM 2012)*. ACM, 2012.
8. M. Koolen, G. Kazai, J. Kamps, A. Doucet, and M. Landoni. Overview of the INEX 2011 books and social search track. In S. Geva, J. Kamps, and R. Schenkel, editors, *Focused Retrieval of Content and Structure: 10th International Workshop of the Initiative for the Evaluation of XML Retrieval (INEX 2011)*, volume 7424 of *LNCS*. Springer, 2012.
9. M. Koolen, G. Kazai, J. Kamps, M. Preminger, A. Doucet, and M. Landoni. Overview of the INEX 2012 social book search track. In S. Geva, J. Kamps, and R. Schenkel, editors, *Focused Access to Content, Structure and Context: 11th International Workshop of the Initiative for the Evaluation of XML Retrieval (INEX'12)*, LNCS. Springer, 2013.
10. 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.
11. C. C. Kuhlthau. Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5):361–371, 1991. ISSN 1097-4571. doi: 10.1002/(SICI)1097-4571(199106)42:5⟨361::AID-ASI6⟩3.0.CO;2-#. URL http://dx.doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-#.
12. H. L. O'Brien and E. G. Toms. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50–69, 2009.
13. 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. Mller, 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. URL http://dx.doi.org/10.1007/978-3-642-40802-1_23.

14. K. Reuter. Assessing aesthetic relevance: Children's book selection in a digital library. *JASIST*, 58(12):1745–1763, 2007.
15. E. Toms and M. M. Hall. The chic interactive task (chici) at clef2013. http://www.clef-initiative.eu/documents/71612/1713e643-27c3-4d76-9a6f-926cdb1db0f4, 2013.
16. E. G. Toms and M. M. Hall. The CHiC Interactive Task (CHiCi) at CLEF2013. In *CLEF 2013 Evaluation Labs and Workshop, Online Working Notes*, 2013.
17. P. Vakkari. A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of documentation*, 57(1):44–60, 2001.