SearchX: Empowering Collaborative Search Research

Sindunuraga Rikarno Putra
Delft University of Technology
Delft, The Netherlands
sindunuragarikarnoputra@student.tudelft.nl

Felipe Moraes*
Delft University of Technology
Delft, The Netherlands
f.moraes@tudelft.nl

Claudia Hauff
Delft University of Technology
Delft, The Netherlands
c.hauff@tudelft.nl

ABSTRACT

Collaborative search has been an active area of research within the IR community for many years. While for "single-user" research a variety of up-to-date open-source search systems exist, few "multi-user" search tools are open-source and even fewer are being maintained. In this paper, we present SearchX, an open-source collaborative search system we are currently developing—and using for our research. We designed and built SearchX using the modern Web stack (and are thus not siloed by an operating system or a particular browser type), enabling efficient research across platforms (Desktop, mobile) and with online users (e.g. crowdworkers). A video, describing the demo can be found at https://www.youtube.com/watch?v=uf24m6p3vts.

ACM Reference format:
Sindunuraga Rikarno Putra, Felipe Moraes, and Claudia Hauff. 2018. SearchX: Empowering Collaborative Search Research. In *Proceedings of The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, Ann Arbor, MI, USA, July 8–12, 2018 (SIGIR'18),* 4 pages.
https://doi.org/10.1145/3209978.3210163

1 INTRODUCTION

Collaborative search, i.e. the retrieval and sensemaking of information by multiple users with the same information need, has been an active area of research for many years—workshops focusing on collaborative search explicitly have started appear in 2008, e.g. [11] and continue to be important for the development of the field today, e.g. [3].

In contrast to single-user search and retrieval (where many up-to-date and open-source tools exist such as Terrier, Lucene or Elasticsearch), our community has few software resources available to kickstart research in collaborative search as seen in Table 1. Of the six most well-known collaborative search systems, only one—Coagmento [7]—is open-sourced and that requires users to install a browser plugin or a mobile app. In contrast, we present here SearchX, a collaborative search system we developed with the modern Web stack in mind; it does not require users to install additional software (like a plugin/app) and is accessible to anyone with a modern Web browser, no matter their platform. This in turn enables us to employ SearchX effortlessly in large-scale online experiments with for instance crowdworkers.

In order to bring down the programmatic hurdle of collaborative search research (which ideally enables quick prototyping of ideas), we open source SearchX[1] and describe here the current state of development, how it fares compared to existing software efforts and how our users (crowdworkers) experienced it in a recent deployment.

2 COMPARISONS WITH EXISTING SYSTEMS

Table 1 provides an in-depth comparison of SearchX and the most well-known existing collaborative search tools [2, 5–7, 9, 10]. We split the dimensions along which we compare the tools into three broad categories: *awareness* (features that raise searchers' awareness of each other's activities), *knowledge sharing* (features that enable searchers to relay their findings to each other), and *division of labour* (features that enable searchers to divide the work). Lastly, we also consider the *availability* of the tools to the wider community.

We find that Coagmento [7] and Querium [6] have the largest number of collaborative features implemented, with SearchX lagging in particular in the *knowledge sharing* category, which is the focus of the ongoing development effort. Interestingly, the *division of labour* category is the most sparsely populated across all systems, with only SearchTogether offering all of the features.

The information in the *availability* category provides the main motivation for our programming effort: most tools were only described in papers, but never open-sourced. Coagmento, the only tool that is open-sourced, requires an additional installation step (either a browser plugin or an iOS/Android app) that is not feasible to enable research with a large population of online users. In contrast, SearchX is a pure open-source Web application, whose architecture we describe next.

3 SEARCHX ARCHITECTURE

The high-level architecture of SearchX is shown in Figure 1. We took the single-user pienapple search framework [4][2] as a starting point, updated it, included provisions for user studies and extended it into the collaborative search realm.

Front-end. The front-end (shown in Figure 3) is written in JavaScript and makes use of React, a popular and well-maintained JavaScript front-end library. As the front-end is a standard Web application, any user with a modern browser (to be more specific a modern browser that implements ES6 functionalities as all major browser

*The first two authors contributed equally.

[1] We open-sourced SearchX at https://github.com/felipemoraes/searchx-frontend (front-end) and https://github.com/felipemoraes/searchx-backend (back-end).

[2] The authors kindly provided us with their source code.

SIGIR'18, July 8–12, 2018, Ann Arbor, MI, USA
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-5657-2/18/07.
https://doi.org/10.1145/3209978.3210163

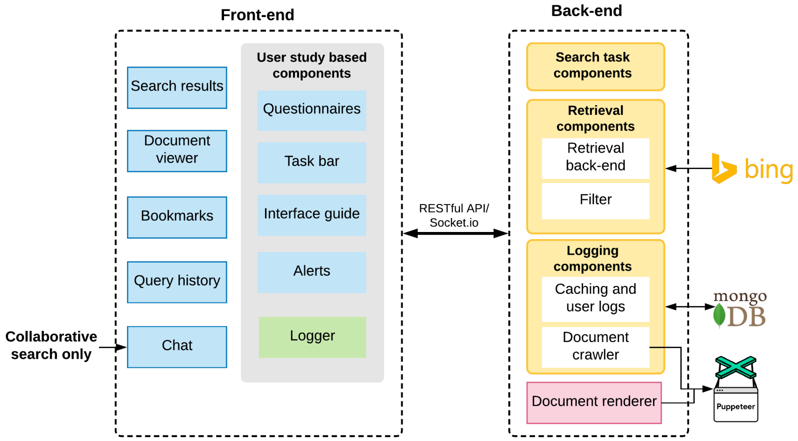
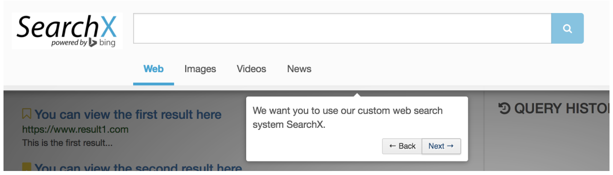


Figure 1: SearchX architecture overview.

Table 1: Feature comparison of existing collaborative search systems (ordered by publication year of the first paper describing the system) with SearchX. SearchX features marked with (✓) are currently under development. A dash – indicates that this information is not available. Language and Platform abbreviations: JS=JavaScript, BP=Browser Plugin, IE=Internet Explorer, FF=Firefox, GC=Google Chrome. †The Coagmento iOS app is only available in Apple's US app store.

	SearchTogether[9]	CoSearch[2]	CoSense[10]	Coagmento[7]	Querium[6]	ResultsSpace[5]	SearchX
Awareness							
Query History	✓		✓	✓	✓	✓	✓
Colour Coding		✓	✓				✓
Page Views				✓	✓		(✓)
Knowledge Sharing							
Bookmarks				✓			✓
Page Rating	✓				✓	✓	(✓)
Page Annotation	✓	✓	✓	✓			(✓)
Group Notes			✓	✓	✓		
Group Summary		✓	✓		✓		
Division of Labour							
Chat	✓		✓	✓	✓		✓
Results Split	✓	✓					
Recommend / Share	✓			✓	✓		
Availability							
Functioning	✗	-	-	✓	-	-	✓
Last Update	2009	-	-	2017	-	-	2018
Open Source	-	-	-	✓	-	-	✓
Language	-	-	-	PHP & JS	-	PHP	JS
Platform	BP (IE)	Windows	-	BP (FF, GC) & iOS†, Android	Web	Web	Web

do) can access it. We follow React's programming model and thus each front-end component is implemented in a modular life cycle manner and communicates with the *back-end* via the *RESTful API* or Socket.io, depending on how instantaneous the information needs to arrive at the collaborators' interfaces. As we are employing SearchX for research purposes, the front-end also contains components which are geared towards user studies and facilitate the deployment of SearchX for online experiments with crowdworkers. Those components include *questionnaires*, a *task bar*, *alerts* (to handle crowdworker compliance), an interactive *interface guide* (Figure 2), and a *logger* to enable the detailed logging (not just clicks,



You can view the first result here
https://www.result1.com
This is the first result...

You can view the second result here

We want you to use our custom web search system SearchX.

Figure 2: SearchX interactive guide.

but also hovers and scrolls) of users' interactions with the interface. When users form a group of collaborators (determined by the back-end), the front-end also includes a number of *shared* components (i.e. all collaborators see them in the same state): bookmarks, query history, and chat widget.

Extensions. As this layer is implemented based on the `React` library, it is easily extensible. We are currently extending the interface with page view information (which collaborator viewed which document), page ratings (how useful did a collaborator find the document) and page annotations.

Back-end. The back-end makes use of the `node.js` server framework and is written in JavaScript. It provides the core components of our collaborative search system which we broadly categorise into four types:

- *search tasks components*: examples are search task assignment (tasks are assigned at *random* or based on users' *prior knowledge* levels) and group formation (users are allocated to an available group to collaboratively search);
- *retrieval components* include retrieval back-end components (at the moment we rely mainly on the Bing search API but can also plug in an Elastic instance) and filters (e.g. to remove certain search results before forwarding them to the front-end);
- *logging components*: these include caching of search results (to serve users consistently the same results for the same issued query), a document crawler based on `Puppeteer` (a tool to access Headless Google Chrome) that crawls all documents shown in a SERP and user trace logging with all logs being stored in a MongoDB instance.
- the *document renderer* provides a document viewing experience that does not require the user to leave the `SearchX` interface (in contrast to clicking on the SERP and opening a new browser tab), thus allowing us to also log what users do (dwell-time and scrolling) within documents they view.

Extensions. Due to the modular nature of `SearchX`, the back-end can be extended to cater for a larger variety of search APIs, additional user study components and more fine-grained logs if desired.

4 USER FEEDBACK

In a recent collaborative search experiment (pairs of users, conducting searches for a learning task for 20+ minutes at a time) involving all components of `SearchX` we collected feedback from fifty crowdworkers in total.

At the end of the task we presented them with a questionnaire that included a set of questions elaborating on their satisfaction with the different collaborative features (chat, bookmarks and query history). Specifically, we asked *"Did you find the collaborative features useful?"* and expected answers on a five point Likert scale (*strongly disagree* to *strongly agree*). Figure 4 summaries the crowdworkers' feedback in this user study. More than 60% of our users considered the shared query history and bookmarks to be useful in a collaborative search setting, while only 24% said the same about the chat feature. A possible explanation can be found in our user population (crowdworkers) which have no intrinsic motivation to excel at a task.

Among the qualitative feedback (elaborations on each of the collaborative features) the following are some of the—admittedly selected—positive statements we received:

- *"We used the group chat to discuss what we had learned and the areas that we had to focus on more and what we didn't understand."*
- *"I mainly used the group chat to summarize and see if my partner saw anything wrong with my summary- to test my knowledge"*
- *"I knew what pages were already deemed useful by my partner based on which pages were already bookmarked."*
- *"I looked at what my partner had searched* [in the query history] *so that I could search different terms to broaden our search"*
- *"I knew which base information would already be covered by my partner due to the shared query history."*

5 DEMONSTRATION

Figure 3 shows the type of demonstration we plan to show off at the SIGIR conference. We will provide at least two laptops and let participants solve fun search puzzles in a collaborative manner. The search puzzles are collected from *A Google a Day*[3], following the template in Figure 5. This type of task has also been employed in prior interactive (collaborative) IR studies [1, 8]. We are able to artificially increase the difficulty of each task by playing around with the type of search results we return from the back-end (e.g. instead of results 1-10 we can also return results ranked 91-100).

6 CONCLUSIONS

We have presented `SearchX`, a system for collaborative search research, recently deployed in a crowdsourcing study with up to 60% of our users rating the collaborative features as useful to their search needs. `SearchX` is meant to a be a continuous effort; we plan to include more user interfaces components, in particular shared workspaces at scale (i.e. those supporting many collaborators at once). We open source `SearchX` (cf. footnote 1) and will continue developing it for at least three years.

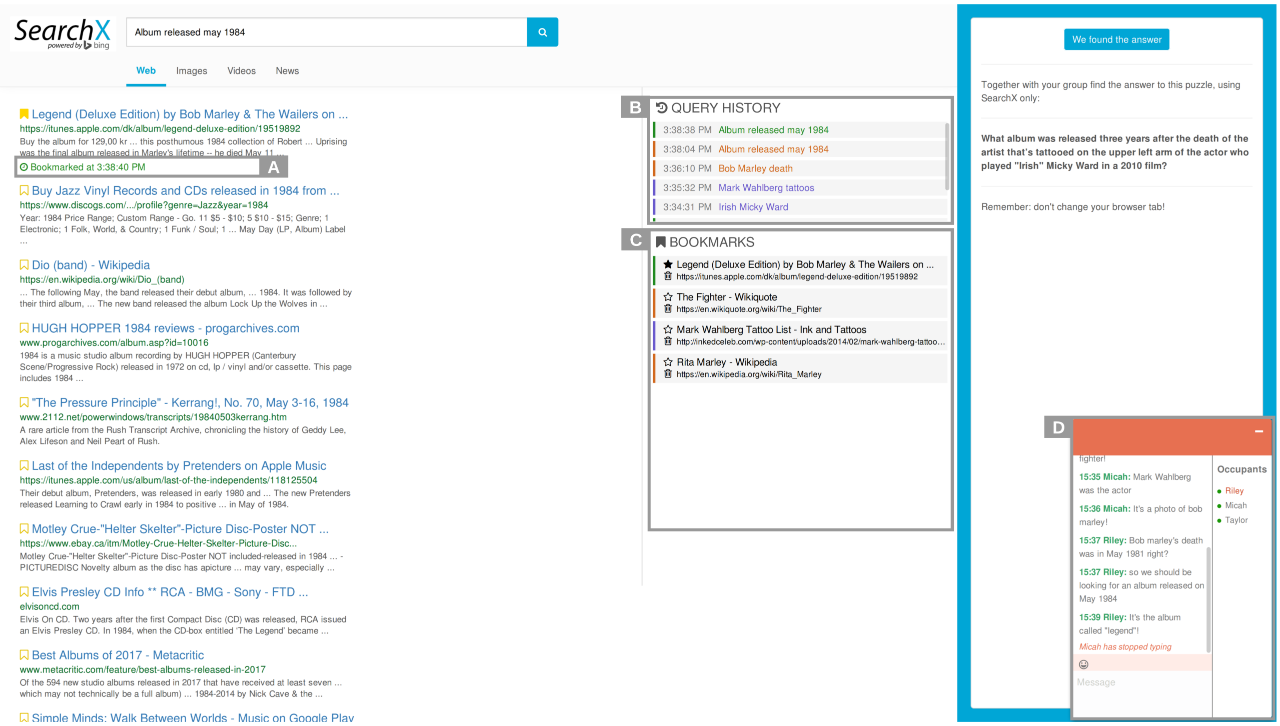
ACKNOWLEDGEMENTS

This research has been supported by NWO projects LACrOSSE (612.001.605) and SearchX (639.022.722).

REFERENCES

[1] Elena Agapie, Gene Golovchinsky, and Pernilla Qvarfordt. 2013. Leading people to longer queries. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3019–3022.

[3] http://www.agoogleaday.com/



Album released may 1984

Legend (Deluxe Edition) by Bob Marley & The Wailers on ...
https://itunes.apple.com/dk/album/legend-deluxe-edition/19519892
Buy the album for 129,00 kr ... this posthumous 1984 collection of Robert ... Uprising was the final album released in Marley's lifetime -- he died May 11 ...
Bookmarked at 3:38:40 PM

Buy Jazz Vinyl Records and CDs released in 1984 from ...
https://www.discogs.com/.../profile?genre=Jazz&year=1984
Year: 1984 Price Range; Custom Range - Go. 11 $5 - $10; 5 $10 - $15; Genre; 1 Electronic; 1 Folk, World, & Country; 1 Funk / Soul; 1 ... May Day (LP, Album) Label ...

Dio (band) - Wikipedia
https://en.wikipedia.org/wiki/Dio_(band)
... The following May, the band released their debut album, ... 1984. It was followed by their third album, ... The new band released the album Lock Up the Wolves in ...

HUGH HOPPER 1984 reviews - progarchives.com
www.progarchives.com/album.asp?id=10016
1984 is a music studio album recording by HUGH HOPPER (Canterbury Scene/Progressive Rock) released in 1972 on cd, lp / vinyl and/or cassette. This page includes 1984 ...

"The Pressure Principle" - Kerrang!, No. 70, May 3-16, 1984
www.2112.net/powerwindows/transcripts/19840503kerrang.htm
A rare article from the Rush Transcript Archive, chronicling the history of Geddy Lee, Alex Lifeson and Neil Peart of Rush.

Last of the Independents by Pretenders on Apple Music
https://itunes.apple.com/us/album/last-of-the-independents/118125504
Their debut album, Pretenders, was released in early 1980 and ... The new Pretenders released Learning to Crawl early in 1984 to positive ... in May of 1984.

Motley Crue-"Helter Skelter"-Picture Disc-Poster NOT ...
https://www.ebay.ca/itm/Motley-Crue-Helter-Skelter-Picture-Disc...
Motley Crue-"Helter Skelter"-Picture Disc-Poster NOT included-released in 1984 ... - PICTUREDISC Novelty album as the disc has apicture ... may vary, especially ...

Elvis Presley CD Info ** RCA - BMG - Sony - FTD ...
elvisoncd.com
Elvis On CD. Two years after the first Compact Disc (CD) was released, RCA issued an Elvis Presley CD. In 1984, when the CD-box entitled 'The Legend' became ...

Best Albums of 2017 - Metacritic
www.metacritic.com/feature/best-albums-released-in-2017
Of the 594 new studio albums released in 2017 that have received at least seven ... which may not technically be a full album) ... 1984-2014 by Nick Cave & the ...

Simple Minds: Walk Between Worlds - Music on Google Play

QUERY HISTORY

3:38:38 PM	Album released may 1984
3:38:04 PM	Album released may 1984
3:36:10 PM	Bob Marley death
3:35:32 PM	Mark Wahlberg tattoos
3:34:31 PM	Irish Micky Ward

BOOKMARKS

★ Legend (Deluxe Edition) by Bob Marley & The Wailers on ...
https://itunes.apple.com/dk/album/legend-deluxe-edition/19519892

☆ The Fighter - Wikiquote
https://en.wikiquote.org/wiki/The_Fighter

☆ Mark Wahlberg Tattoo List - Ink and Tattoos
http://inkedceleb.com/wp-content/uploads/2014/02/mark-wahlberg-tattoo…

☆ Rita Marley - Wikipedia
https://en.wikipedia.org/wiki/Rita_Marley

Together with your group find the answer to this puzzle, using SearchX only:

What album was released three years after the death of the artist that's tattooed on the upper left arm of the actor who played "Irish" Micky Ward in a 2010 film?

Remember: don't change your browser tab!

fighter!
15:35 Micah: Mark Wahlberg was the actor
15:36 Micah: It's a photo of bob marley!
15:37 Riley: Bob marley's death was in May 1981 right?
15:37 Riley: so we should be looking for an album released on May 1984
15:39 Riley: It's the album called "legend"!
Micah has stopped typing

● Riley
● Micah
● Taylor

Figure 3: SearchX Collaborative search interface. [A] bookmarking information including who bookmarked, [B] shared query history, [C] shared bookmarks, [D] chat. All features employ colour coding to visually represent the searcher who performed the search, added the bookmark and so on.

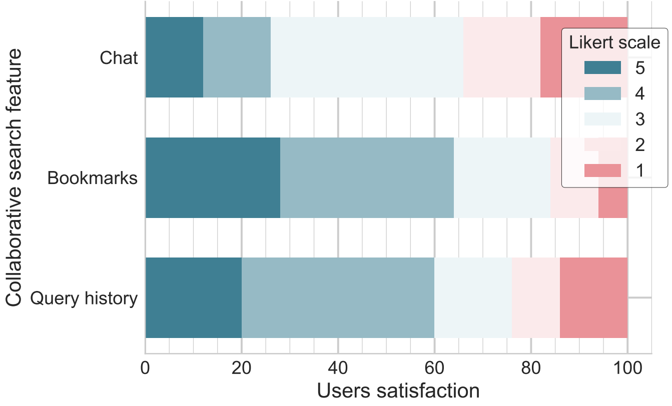


Figure 4: User feedback overview.

Together with your group find the answer to this puzzle, using SearchX only:
What album was released three years after the death of the artist that's tattooed on the upper left arm of the actor who played "Irish" Micky Ward in a 2010 film?

Figure 5: Task template for our demonstration.

[2] Saleema Amershi and Meredith Ringel Morris. 2008. CoSearch: a system for co-located collaborative web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1647–1656.

[3] Leif Azzopardi, Jeremy Pickens, Chirag Shah, Laure Soulier, and Lynda Tamine. 2017. Second International Workshop On the Evaluation of Collaborative Information Seeking and Retrieval (Ecol'17). In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval (CHIIR '17)*. ACM, New York, NY, USA, 429–431. DOI: http://dx.doi.org/10.1145/3020165.3022164

[4] Martynas Buivys and Leif Azzopardi. 2016. Pienapple search: an integrated search interface to support finding, refinding and sharing. *Proceedings of the Association for Information Science and Technology* 53, 1 (2016), 1–5. http://dx.doi.org/10.1002/pra2.2016.14505301122

[5] Robert Capra, Annie T. Chen, Katie Hawthorne, Jaime Arguello, Lee Shaw, and Gary Marchionini. 2012. Design and evaluation of a system to support collaborative search. *Proceedings of the American Society for Information Science and Technology* 49, 1 (2012), 1–10. DOI: http://dx.doi.org/10.1002/meet.14504901181

[6] Abdigani Diriye and Gene Golovchinsky. 2012. Querium: a session-based collaborative search system. In *European Conference on Information Retrieval*. Springer, 583–584.

[7] Roberto González-Ibáñez and Chirag Shah. 2011. Coagmento: A system for supporting collaborative information seeking. *Proceedings of the Association for Information Science and Technology* 48, 1 (2011), 1–4.

[8] Roberto González-Ibáñez and Chirag Shah. 2012. Investigating positive and negative affects in collaborative information seeking: A pilot study report. *Proceedings of the Association for Information Science and Technology* 49, 1 (2012), 1–4.

[9] Meredith Ringel Morris and Eric Horvitz. 2007. SearchTogether: An Interface for Collaborative Web Search. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST '07)*. ACM, New York, NY, USA, 3–12. DOI: http://dx.doi.org/10.1145/1294211.1294215

[10] Sharoda A. Paul and Meredith Ringel Morris. 2009. CoSense: Enhancing Sensemaking for Collaborative Web Search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1771–1780. DOI: http://dx.doi.org/10.1145/1518701.1518974

[11] Jeremy Pickens, Gene Golovchinsky, and Meredith Ringel Morris. 2009. Proceedings of 1st International Workshop on Collaborative Information Seeking. *CoRR* abs/0908.0583 (2009). arXiv:0908.0583 http://arxiv.org/abs/0908.0583