

Coagmento: Past, Present, and Future of an Individual and Collaborative Information Seeking Platform

Matthew Mitsui

Department of Computer Science
Rutgers University
mmitsui@cs.rutgers.edu

Jiqun Liu, Chirag Shah

School of Communication & Information
Rutgers University
{jl2033,chirags}@rutgers.edu

ABSTRACT

In this demo, we present *Coagmento*, a Web-based, open-source tool for information seeking projects that collects information for individuals and groups and helps facilitate collaborative information seeking. *Coagmento* has been used in information retrieval and human-computer interaction studies to investigate individual and group information seeking behaviors in a lab or a field setting. In this demo, we discuss what *Coagmento* is, its past uses in prior studies, and its present state. We also discuss current work in progress. With *Coagmento* recently passing its 10th anniversary, we discuss our intention to make it a tool that is easy to configure for a human information behavior researcher with little programming skill.

CCS CONCEPTS

• **Information systems** → **Collaborative search**; Open source software; • **Human-centered computing** → **Computer supported cooperative work**; Open source software;

KEYWORDS

collaborative information seeking; user studies; human-computer interaction; web search behavior;

ACM Reference Format:

Matthew Mitsui and Jiqun Liu, Chirag Shah. 2018. Coagmento: Past, Present, and Future of an Individual and Collaborative Information Seeking Platform. In *CHIIR '18: 2018 Conference on Human Information Interaction Retrieval, March 11–15, 2018, New Brunswick, NJ, USA*. ACM, New York, NY, USA, Article 4, 4 pages. <https://doi.org/10.1145/3176349.3176896>

1 BACKGROUND

Over decades of interactive information retrieval research, there has been much work linking interactive information retrieval behaviors to interesting phenomena, such as the type of task a person is working on or a person's task/topic familiarity. Such works have also attempted to use a searcher's information seeking behavior to support their overall search experience, for instance by providing live suggestions. Behind each of these data-driven studies about information seeking behavior is a great data collection tool. Other data collection tools include recent works by He and Yilmaz [6] and

Huurdeman et al [8]. Such tools collect an individual's behaviors as they browse through the web and/or a custom search engine web portal on a desktop browser such as Firefox or Chrome. Such infrastructure offers the flexibility of allowing researchers to choose whether to conduct a laboratory study or a field study.

Some information retrieval researchers have focused particularly on collaborative information seeking (CIS) efforts where people work in groups. Corresponding tools to study and support CIS include Ariadne [17], Cerchiamo [3], CoSearch [1], and SearchTogether [12]. *Coagmento*, inspired by these past systems, has been in development and use since 2007 as a tool for supporting collaborative information seeking [2]. It made its public debut in 2009 [16] with subsequent versions demonstrated at various places (e.g., [5, 11]). Some of the design lessons in developing a system like this were presented by Shah [13]. *Coagmento* fundamentally differs by supporting both individual and collaborative work in both synchronous and asynchronous contexts. It has been used in several studies.

Coagmento has recently passed its 10-year anniversary, and it has since undergone several foundational shifts. We will present the current state of *Coagmento*, future work, and the implications towards future research in individual and collaborative information seeking.

2 WHAT IS COAGMENTO?

2.1 As a Researcher's Tool

While much focus on *Coagmento* has been the types of interface support it provides, at its core it is also a browser logger. It logs the pages and queries of users as they search for information. This functionality has been utilized in a variety of studies purely for the sake of recording and analyzing log data. Hendaheva and Shah[7] used these recording capabilities on top of traditional browser logging methods to forecast how well a person will perform in a session-based Web search task. In addition, components of the collection have been modified and utilized to create new interfaces. For instance, Mitsui et al.[10] incorporated *Coagmento* and other tools to record study query-level information seeking intentions in search sessions to study interactive IR. Other traditional data capture methods can of course be incorporated with *Coagmento*, such as diary studies in Shah and Leeder[15], commercial video capture tools such as Morae in Mitsui et al.[10], and also face capture in González-Ibáñez et al.[4]. Of course, *Coagmento* is intended for studying CIS, for which it has also been utilized as in Shah et al.[14] and Knight et al.[9]. Shah et al. [14] directly examined the effect of crowd size on search performance, assigning Web search tasks to synchronously working groups of dyads and triads. Knight et al.[9]

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHIIR '18, March 11–15, 2018, New Brunswick, NJ, USA

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-4925-3/18/03...\$15.00

<https://doi.org/10.1145/3176349.3176896>

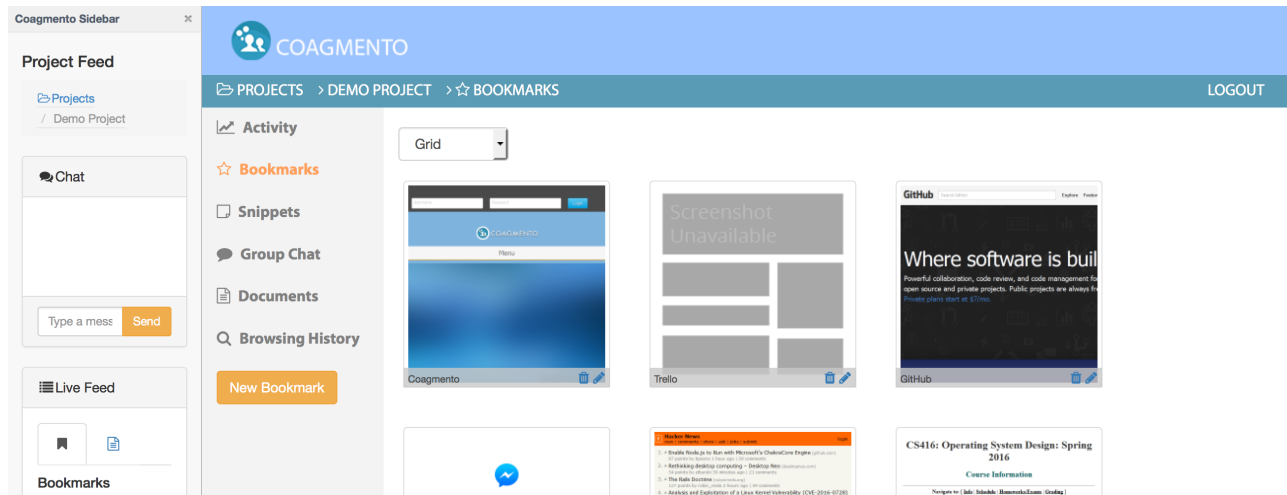


Figure 1: Coagmento web interface, including sidebar (left) and project page with thumbnails of collected bookmarks (center).

examined groups' assessments of expertise and document trust in a collaborative task. Moreover, it has also been used for several laboratory studies, such as in Hendaheewa and Shah[7]. In such studies, a traditional linear flow of "stages" is incorporated into the base version of Coagmento. Stages are a linear flow of pages typical to an IR-based laboratory study, such as a demographic survey, a pre-task questionnaire, tutorial videos, and a post-task questionnaire. While Coagmento has largely been used for small-scale studies, it has also been used for larger ones. Knight et al.[9] and Hendaheewa and Shah[7] examined the logs of over 1000 and 200 participants, respectively, engaged in a controlled study, with browser log data collected through Coagmento.

2.2 As a System

Coagmento comprises of two parts. The first and main component is the browser plugin. This is installed on individual machines where a researcher wants to collect a study participant's information seeking behavior. Installation may occur on a person's private machine in a field study or on a lab machine for a lab study. The public version of the plugin is available in several flavors, including Firefox, Chrome, and even Android and iOS¹. It includes several buttons, a toolbar and a sidebar. The toolbar is typically used to house buttons for bookmarking pages or saving snippets of text. The sidebar is used to display these things as well as to interact with the chat. More importantly, these are overlaid on the browser window so that a participant may seamlessly interact with Coagmento while searching for information in parallel. While a person may engage with the web service as in Figure 1, they may alternatively engage with the sidebar in parallel searching on Google or browsing the news. The plugin is also the main component that collects browsing and searching behavior of users. The sidebar is given in Figure 2. An example of the main web-based interface - programmed in PHP - is also shown in Figure 1, where a user can edit their browsing history.

¹<http://coagmento.org/download.php>

This front-end functionality is supported by several back-end services. First is the core functionality for storing and retrieving data. This was once programmed in vanilla PHP but is now built on a framework called Laravel². It is a Model-View-Controller PHP framework that offers several constructs that make further Coagmento development more manageable. Laravel follows modern web development conventions, such as defining routes and following RESTful APIs. It also provides support for several common functions useful to any web service and also Coagmento, such as authentication middleware, administrator privileges, and form validation. While Coagmento is traditionally built on MySQL, Laravel supports the use of several other types of SQL databases. Coagmento is additionally supported by real-time functionality using Node.js³. Real-time functionality is necessary for any interactive study, especially one where there is synchronous collaboration such as in Shah et al.[14] and Knight et al.[9]. Node.js provides this through "push notifications", which are common in mobile applications like Twitter where real-time reactivity is desirable but computational resources are scarce. Push notifications push live updates to project collaborators, such as when a collaborator creates a new bookmark.

The expected installation workflow of Coagmento is given in Figure 3. To summarize, any computer with PHP, Laravel, Node.js and a SQL database could be a *Coagmento* server.

3 WORK IN PROGRESS

The largest hurdle in accessibility to Coagmento is that it still requires programming experience to be modified. Coagmento comes with many features out of the box, but not all researchers are interested in using the full suite of tools. In such a case Coagmento needs to be modified, which currently requires at least some navigation of code. Various data collection processes and visualization components all need to be modified or omitted for new studies. Other options that need configuration include database credentials

²<https://laravel.com/>

³<https://nodejs.org/en/>

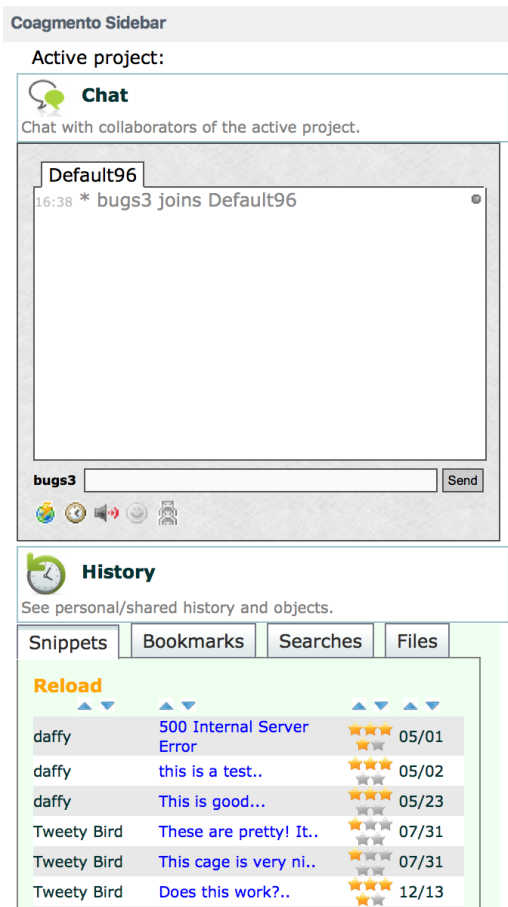


Figure 2: Sidebar.

to which data is committed and administrative credentials for researchers. If a researcher wants to design a questionnaire for a human-computer interaction study (e.g. "How was your experience using this tool?", "How satisfied were you in finding information?"), this is an extra step that needs deliberate programming effort.

Understandably, not everyone who has used Coagmento (or will ever want to use it) has a background in Computer Science, let alone computer programming. The main goal of human behavior research is to generate and investigate questions and to publish the answers. It is not to generate, investigate, and publish code. Therefore, any extra steps from question generation to answer are a barrier to research. Our current work is in breaking down this barrier so that human behavior researchers with little to no programming experience can quickly investigate their research questions. Future work in this regard includes:

- Questionnaire generation through a GUI
- Stage generation through a GUI, for a study that contains stages such as consent forms, pre-task questionnaires and post-task questionnaires
- Toggling existing components in a configuration file
- A (official) Chrome extension

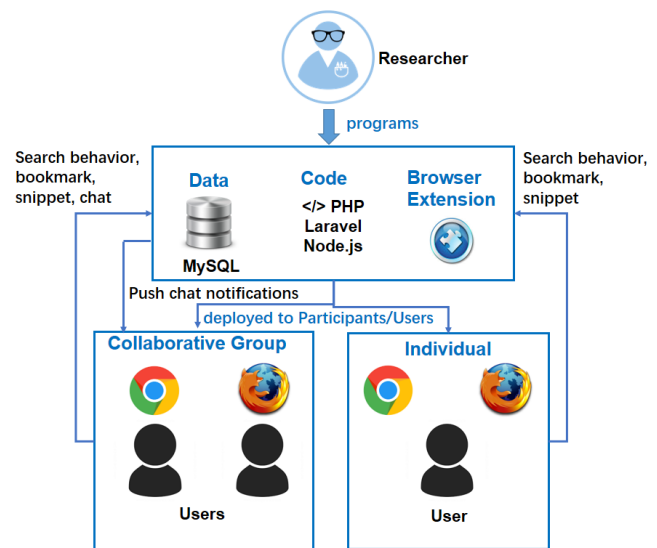


Figure 3: Overview of the flow of data and software for Coagmento. Researchers program the main code and distribute plugin instances to participants.

4 WHAT'S IN THE DEMO?

In this demo, we will show some of the basic front-end functionality for out-of-the-box Coagmento. This includes the functionality of the web-based service and also the desktop browser plugin. We can show interested parties how to download and install Coagmento on their machines. We can also walk interested programmers through the basics of the code base, and we can help interested parties create a fresh installation of Coagmento on their machines. The tool is currently available for public use at the footnote below.⁴

ACKNOWLEDGMENTS

Work on Coagmento is supported through the Institute of Museum and Library Services (IMLS) grant LG-81-16-0025-16 and the National Science Foundation (NSF) grant IIS-1717488.

REFERENCES

- [1] Saleema Amershi and Meredith Ringel Morris. 2008. CoSearch: A System for Co-located Collaborative Web Search. In *Proceedings of CHI '08*. ACM, New York, NY, USA, 1647–1656. <https://doi.org/10.1145/1357054.1357311>
- [2] Xin Fu, Diane Kelly, and Chirag Shah. 2007. Using Collaborative Queries to Improve Retrieval for Difficult Topics. In *Proceedings of ACM SIGIR '07*. ACM, New York, NY, USA, 2. <https://doi.org/10.1145/1277741.1277955>
- [3] Gene Golovchinsky, John Adcock, Jeremy Pickens, Pernilla Qvarfordt, and Mari-beth Back. 2008. Cerchiamo: a collaborative exploratory search tool. *Proceedings of CSCW* (2008), 8–12.
- [4] Roberto González-Ibáñez, Chirag Shah, and Natalia Cordova-Rubio. 2011. Smile! Studying expressivity of happiness as a synergic factor in collaborative information seeking. *Proceedings of the American Society for Information Science and Technology* 48, 1 (2011), 1–10. <https://doi.org/10.1002/meet.2011.14504801171>
- [5] Roberto González-Ibáñez and Chirag Shah. 2011. Coagmento: A System for Supporting Collaborative Information Seeking. *The 74th Annual Meeting of the American Society for Information Science and Technology ASIST 2011* (2011).
- [6] Jiyin He and Emine Yilmaz. 2017. User Behaviour and Task Characteristics: A Field Study of Daily Information Behaviour. In *Proceedings of the 2017 Conference*

⁴<https://github.com/InfoSeeking/Coagmento>

- on *Conference Human Information Interaction and Retrieval (CHIIR '17)*. ACM, New York, NY, USA, 67–76. <https://doi.org/10.1145/3020165.3020188>
- [7] Chathra Hendaheewa and Chirag Shah. 2017. Evaluating User Search Trails in Exploratory Search Tasks. *Inf. Process. Manage.* 53, 4 (July 2017), 905–922. <https://doi.org/10.1016/j.ipm.2017.04.001>
- [8] Hugo C. Huurdeman, Max L. Wilson, and Jaap Kamps. 2016. Active and Passive Utility of Search Interface Features in Different Information Seeking Task Stages. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval (CHIIR '16)*. ACM, New York, NY, USA, 3–12. <https://doi.org/10.1145/2854946.2854957>
- [9] Simon Knight, Bart Rienties, Karen Littleton, Matthew Mitsui, Dirk Tempelaar, and Chirag Shah. 2017. The relationship of (perceived) epistemic cognition to interaction with resources on the internet. *Computers in Human Behavior* 73, Supplement C (2017), 507 – 518. <https://doi.org/10.1016/j.chb.2017.04.014>
- [10] Matthew Mitsui, Jiquan Liu, Nicholas J. Belkin, and Chirag Shah. 2017. Predicting Information Seeking Intentions from Search Behaviors. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17)*. ACM, New York, NY, USA, 1121–1124. <https://doi.org/10.1145/3077136.3080737>
- [11] M. Mitsui and C. Shah. 2016. Coagmento 2.0: A system for capturing individual and group information seeking behavior. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, Vol. 2016-Sept. <https://doi.org/10.1145/2910896.2925447>
- [12] Meredith Ringel Morris and Eric Horvitz. 2007. SearchTogether: An Interface for Collaborative Web Search. In *Proceedings of the 20th ACM UIST Symposium (UIST '07)*. ACM, New York, NY, USA, 3–12. <https://doi.org/10.1145/1294211.1294215>
- [13] Chirag Shah. 2012. Coagmento – A Case Study in Designing User-Centric Collaborative Information Seeking System. In *System Science and Collaborative Information Systems: Theories, Practices and New Research*, Emilia Currás and Nuria Lloret (Eds.). IGI Global, 242–257.
- [14] Chirag Shah, Chathra Hendaheewa, and Roberto González-Ibáñez. 2015. Two's company, but three's no crowd: Evaluating exploratory web search for individuals and teams. *Aslib Journal of Information Management* 67, 6 (2015), 636–662. <https://doi.org/10.1108/AJIM-05-2015-0082> arXiv:<https://doi.org/10.1108/AJIM-05-2015-0082>
- [15] Chirag Shah and Chris Leeder. 2016. Exploring collaborative work among graduate students through the C5 model of collaboration: A diary study. *Journal of Information Science* 42, 5 (2016), 609–629. <https://doi.org/10.1177/0165551515603322> arXiv:<https://doi.org/10.1177/0165551515603322>
- [16] Chirag Shah, Gary Marchionini, and Diane Kelly. 2009. Learning Design Principles for a Collaborative Information Seeking System. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 3419–3424. <https://doi.org/10.1145/1520340.1520496>
- [17] Michael B. Twidale, David M. Nichols, and Chris D. Paice. 1997. Browsing is a collaborative process. *Information Processing Management* 33, 6 (1997). [https://doi.org/10.1016/S0306-4573\(97\)00040-X](https://doi.org/10.1016/S0306-4573(97)00040-X)