Towards Humane Feedback Mechanisms in Exploratory Search

Esben Sørig
Goldsmiths, University of London
esben.sorig@gold.ac.uk

Nicolas Collignon
University of Edinburgh
n.collignon@ed.ac.uk

Rebecca Fiebrink
Goldsmiths, University of London

Noriko Kando
National Institute of Informatics
Tokyo, Japan

ABSTRACT

Machine learning (ML) plays a central role in modern information retrieval (IR) systems. We argue that, in IR systems for multi-session exploratory search, there are unexploited opportunities for IR document ranking models to leverage users' knowledge about the search task to better support users' search needs. Specifically, we propose a method to enable users to adapt an IR document ranking model according to their information needs, using an interface that supports search strategies and methods for engaging with documents known to be useful when people explore new or complex domains of knowledge. We also discuss the major challenges in creating human-centered machine learning models and interfaces for exploratory search.

CCS CONCEPTS

• **Information systems** → **Search interfaces**; *Personalization*; • **Human-centered computing** → **Interaction paradigms**; • **Computing methodologies** → *Machine learning*.

HCMLP Workshop, CHI'19, May 2019, Glasgow, United Kingdom

KEYWORDS

Exploration; Relevance Feedback; Interactive Search; HCI; Human-Centered Machine Learning

ACM Reference Format:

Esben Sørig, Nicolas Collignon, Rebecca Fiebrink, and Noriko Kando. 2019. Towards Humane Feedback Mechanisms in Exploratory Search. In *Proceedings of Human-Centered Machine Learning Perspectives Workshop, ACM Conference on Human Factors in Computing Systems (HCMLP Workshop, CHI'19).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.475/123_4

"Our ineptitude in getting at the record is largely caused by the artificiality of systems of indexing. When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. one has to have rules [...] and the rules are cumbersome. The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain. [...] the speed of action, the intricacy of trails, the detail of mental pictures, is awe-inspiring beyond all else in nature. Man cannot hope fully to duplicate this mental process artificially, but he certainly ought to be able to learn from it. In minor ways he may even improve it."

V. Bush, As We May Think*, July 1945 [2]*

INTRODUCTION

One of the main problems in information retrieval (IR) is to model the relevance of documents to users' information needs. Machine learning can be applied to this problem, learning a ranking function that depends on the user's query and any contextual data used for personalization of the results [8]. The ranking function is typically learned by collecting implicit feedback from users, as this type of feedback can be collected at large scales without imposing any cognitive load on search users [7]. A common form of implicit feedback is user click data from which the relevance of a document given a query is inferred. Applying machine learning to such data can lead to performance boosts compared to traditional, non-ML methods [8]. However, most work only considers the creation of ML models for single-session look-up tasks and does not consider how to enable users to interactively adapt the ranking model to their information needs as they explore new or complex information landscapes.

In 1945, Bush was perhaps the first to describe a vision for a knowledge indexing system that would support users in the exploration of information. The vision was a system that would augment the human mind by mimicking it [2]. Humane approaches to the design of tools that support human intellectual activities have a long, yet often overlooked history. Most recently, Victor described the aim of designing *Humane media* as creating "environments where users are able to express their full range of capabilities" [17]. Engelbart, in his *Conceptual Framework for Augmenting Human Intellect* presents the two clear steps to approach this [4]:

"(1) to find the factors that limit the effectiveness of the individual's basic information-handling capabilities in meeting the various needs for problem solving in its most general sense; and (2) to develop new techniques, procedures, and systems that will better match these basic capabilities to the needs' problems, and progress."

Applying this human-centered analysis to exploratory search highlights one of its main challenges and shows how machine learning provides a promising approach to addressing it when one introduces extensions to the learning-to-rank paradigm.

Expressing Information Needs

The primary constraint on the user in IR lies in effectively communicating their information needs and interests. This is particularly hard when the goals of the user may initially be unclear, as is often the case for people delving into domains of knowledge over extended periods of time, people confronted with chronic illness, work teams designing complex solutions, families making long-term plans, or scientists investigating complex phenomena. These goals may develop as the searcher acquires new knowledge, and evolve over search sessions that can last over extended periods of time.

Although they may find it difficult to express their information need, users can often identify relevant documents in returned search results. This observation has led to an interactive machine learning technique known as *relevance feedback* [11]. Using relevance feedback, users can mark documents as relevant or irrelevant to adapt the ranking model to more closely match their current information needs. Relevance feedback has been shown to improve retrieval performance of search systems [12]. However, research has pointed towards low user engagement with relevance feedback during general search engine use [16]. We hypothesise that the low user engagement with relevance feedback is due to the additional cognitive load introduced by an interaction that is not part of natural search behaviour. We argue that feedback signals for ranking models should take advantage of users' established capabilities in information exploration. In the next section, we introduce one such feedback signal based on document annotation and active reading.

HIGHLIGHT ANNOTATIONS FOR EXPLORATORY SEARCH

We propose to let users adapt the ranking model interactively by letting them highlight passages relevant to their current information need. Many researchers still prefer to print hard copies of source materials, to read and annotate them. Golovchinsky et al (1999)[5] argued that this highlights some of the limits that current search systems have. Interacting with a document can help enhance the reader's understanding, or recall, of the information (a practice sometimes refered to as *active reading* [1]). Annotating documents is therefore a natural part of the exploratory search process for many users. This usually entails highlighting key passages of text, and making notes in the margins of the document[3]. Today most of this rich annotation information is lost, either on the physical copy of a document, or simply because digital systems prevent such annotations. We argue that the role of a successful exploratory IR system is not only to make use of such rich feedback but also to encourage it. As part of our ongoing research, we have developed a ranking model that learns from annotation feedback and a user interface to support users in exploratory search.

The nature of annotation feedback differs from relevance feedback in that it refers to the relevance of specific features of a document rather than the overall relevance of a document. Annotation feedback can be modelled using *feature feedback* methods proposed in works such as [10, 15]. Feature feedback

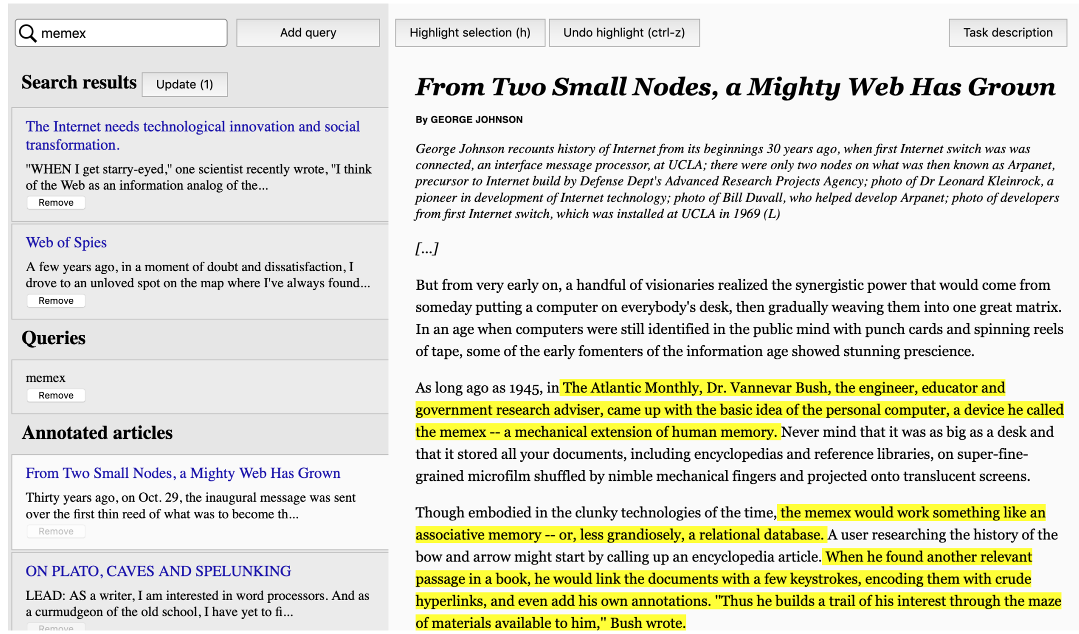

Figure 1: The user interface of our search system that supports annotation feedback. Upon receiving their initial list of search results based on a query, the user can highlight passages they find relevant. The user can update their search results at any point based on the annotations or on additional queries they enter. The user interface shows the complete search context that is used to compute search results including previous queries and annotated articles. The user may edit the search context at any point.

approaches use labels on the features that are used to represent each training example. In general, such features may not be interpretable for the labeler to give feedback on. In our search system, we use bag-of-words features to represent documents and interpret highlights as relevance statements about the features (words) that make a document relevant. Our system classifies documents as relevant or irrelevant using a Multinomial Naive Bayes model and follow the approach of [15] by using relevance feedback to modify the priors of features highlighted by the user.

User Interface

The user interface of our search system (Figure 1) is designed to support users in interactively adapting and actively engaging with search results. Users may add any number of queries to their search

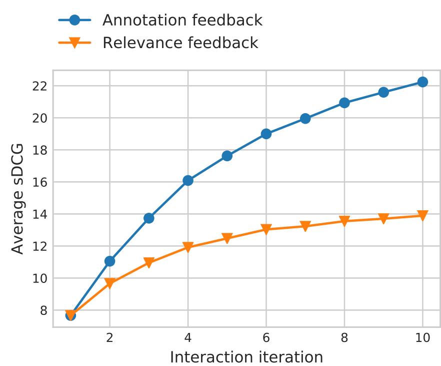

Figure 2: Performance of annotation feedback compared to document-level relevance feedback using a simulated user performing search tasks on the New York Times Annotated Corpus [13]. An interaction iteration consists of the simulated user reviewing and annotating all 5 search results returned by the system [18]. Session Discounted Cumulative Gain (sDCG) is a retrieval performance metric that takes the order of returned results and their relevance into account [18]. The results are averaged over 60 search tasks.

context, annotate relevant passages in returned documents, and remove documents such that they do not appear in future results. Users can undo and edit all these aspects of their search context at any point and the current context is always displayed in the interface. Users can request the result list to be updated when they have modified the search context. Users steer the document relevance model in this way. For example, if a user adds the initial query "apple" and highlights passages that are relevant to them in the results, the system may then return articles about the fruit or the technology company depending on the information need they have clarified through their annotations.

Preliminary Results

A previous study has showed that annotation feedback can improve retrieval performance compared to relevance feedback [5]. However, the study was done in a static setting with a single iteration of feedback, with predefined queries, and without participants seeing the updated results. To investigate the effect of performance in an interactive setting with multiple feedback iteration, we have evaluated the feedback type with a user simulator (see Figure 2).

We are currently working on an experiment that will compare the two feedback types when used by human subjects using our search system. We aim to understand not only the retrieval performance but also the user experience and information engagement of human subjects under each condition.

DISCUSSION AND FUTURE WORK

Current approaches to ML in IR are not well matched to human abilities for exploration, knowledge discovery, and learning. We have argued that feedback signals for ranking models should take advantage of users' full range of capabilities. With this aim, we presented an explicit learning signal based on document annotations. We argued that this was a step towards a more humane approach for search in two ways. First, making complex search goals explicit can be difficult for people during the exploratory search process. Using annotation feedback, a user can interactively navigate a complex information landscape, steering the model towards new relevant documents to be explored. Second, encouraging users to annotate documents through the search interface removes disruptive barriers between the different stages of exploratory search. Annotating documents is an essential part of actively engaging with documents, and a system for knowledge exploration should encourage such practices. Previous research, and initial results with simulated users, both carry evidence that this type of feedback leads to better retrieval performance. Our future work will aim to evaluate the feedback approach with human subjects. We believe that this type of interface is an initial step towards more humane tools for exploratory search.

More generally, we believe that the application of ML in IR and text domains will have an important role in matching interfaces with human capabilities for exploration. The user interfaces of Web search systems are still limited to single-session query-based interaction. While ML algorithms have significantly optimized the ranking of results, we believe such interfaces do not provide the user with an adequate overview of the information landscape, and remain far from supporting basic human capabilities. With a better understanding of how people represent their environments during exploration, and the strategies they follow, ML methods could improve the design of interfaces to match more natural representations, and support intuitive strategies for information gathering (see [6, 9, 14]), by e.g. allowing for explicit control over how exploratory results should be when compared to the user's previously attended sources.

ACKNOWLEDGMENTS

This work was supported by Microsoft Research through its PhD Scholarship Programme and by the National Institute of Informatics (Japan). JSPS KAKENHI Grant Numbers JP16H01756.

REFERENCES

[1] Mortimer J Adler and Charles Van Doren. 1972. How to Read a Book, rev. ed. *New York* (1972).
[2] Vannevar Bush. 1945. As we may think. *The atlantic monthly* 176, 1 (1945), 101–108.
[3] Sally Jo Cunningham and Chris Knowles. 2005. Annotations in an academic digital library: the case of conference note-taking and annotation. In *International Conference on Asian Digital Libraries*. Springer, 62–71.
[4] D. C. Engelbart. 1962. AUGMENTING HUMAN INTELLECT: A CONCEPTUAL FRAMEWORK.
[5] G. Golovchinsky, M. N. Price, and B. N. Schilit. 1999. From reading to retrieval: freeform ink annotations as queries. In *Proceedings of ACM SIGIR 99*. ACM, 19–25.
[6] Thomas T Hills, Peter M Todd, David Lazer, A David Redish, Iain D Couzin, Cognitive Search Research Group, et al. 2015. Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences* 19, 1 (2015), 46–54.
[7] D. Kelly and J. Teevan. 2003. Implicit feedback for inferring user preference: a bibliography. In *SIGIR Forum*, Vol. 37. 18–28.
[8] Tie-Yan Liu et al. 2009. Learning to rank for information retrieval. *FTIR* 3, 3 (2009), 225–331.
[9] Jaimie Murdock, Colin Allen, and Simon DeDeo. 2017. Exploration and exploitation of Victorian science in Darwin's reading notebooks. *Cognition* 159 (2017), 117–126.
[10] S. Poulis and S. Dasgupta. 2017. Learning with feature feedback: from theory to practice. In *AISTATS*. 1104–1113.
[11] I. Ruthven and M. Lalmas. 2003. A Survey on the Use of Relevance Feedback for Information Access Systems. *Knowl. Eng. Rev.* 18, 2 (June 2003), 95–145.
[12] G. Salton and C. Buckley. 1990. Improving retrieval performance by relevance feedback. *JAIST* 41, 4 (1990), 288–297.
[13] Evan Sandhaus. 2008. The new york times annotated corpus. *Linguistic Data Consortium, Philadelphia* 6, 12 (2008), e26752.
[14] Eric Schulz, Rahul Bhui, Bradley C Love, Bastien Brier, Michael T Todd, and Samuel J Gershman. 2018. Exploration in the wild. *BioRxiv* (2018), 492058.
[15] Burr Settles. 2011. Closing the Loop: Fast, Interactive Semi-supervised Annotation with Queries on Features and Instances. In *Proceedings of EMNLP '11*. 1467–1478.
[16] A. Spink, B. J. Jansen, and H. Cenk Ozmultu. 2000. Use of query reformulation and relevance feedback by Excite users. *Internet research* 10, 4 (2000), 317–328.
[17] Bret Victor. 2014. Humane Representation of Thought: A Trail Map for the 21st Century. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. 699–699. https://doi.org/10.1145/2642918.2642920
[18] Grace Hui Yang, Zhiwen Tang, and Ian Soboroff. 2017. TREC 2017 Dynamic Domain Track Overview.. In *TREC '17*.