HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY

Aalto University · UNIVERSITY OF HELSINKI · TAMPERE UNIVERSITY OF TECHNOLOGY

Interactive Visualization of Search Intent for Exploratory Information Retrieval

Antti Kangasrääsiö, Dorota Głowacka, Tuukka Ruotsalo, Jaakko Peltonen, Manuel J. A. Eugster, Ksenia Konyushkova, Kumaripaba Athukorala, Ilkka Kosunen, Aki Reijonen, Petri Myllymäki, Giulio Jacucci, Samuel Kaski
firstname.lastname@hiit.fi

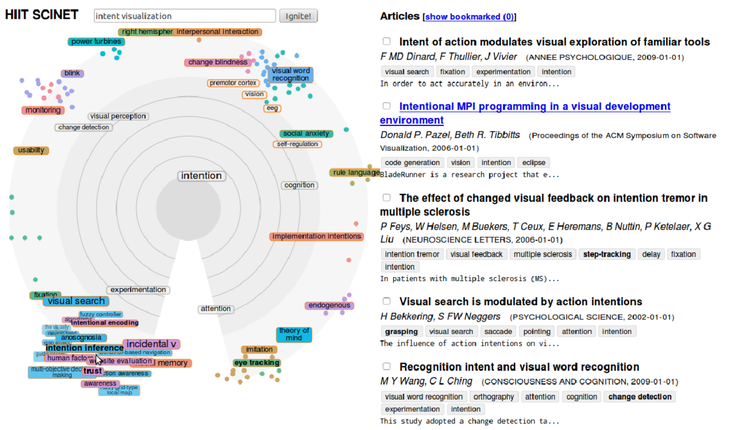
Motivation

In exploratory search a user is at the same time searching and learning. For example, a scientist who is searching for information about a new research area that she is not very familiar with is doing exploratory search.

With current search interfaces, such as Google Scholar, exploratory search requires iteration between examining search results, observing new information and reformulating the search query. However, this requires considerable cognitive effort from the user.

Some interfaces help the user by suggesting popular related query terms. However, this may trap the user to the initial query context and hinder exploration.

Intent Radar



With the Intent Radar search interface, the user search intent model is visualized directly to the user, and the user can direct the search by moving the keywords on the radar. The closer to the center a keyword is moved the more relevant it is. The Intent Radar is pictured above, acting as the user interface to the HIIT SciNet scientific search engine.

The user is able to see the topics that exist in the nearby information space. Because of this, the user does not need to browse through the retrieved documents to find relevant search keywords.

Intuitively directing the search by moving keywords closer or further from the center reduces the cognitive effort required from the user. The ease of use also makes the user more likely to take advantage of the Intent Radar.

Intent Visualization

The most relevant current search keywords are visualized in the inner part of the Intent Radar. The closer to the center they are, the larger the relevance. The most relevant future search keywords are visualized on the outer part of the Intent Radar in a similar manner.

The angles of the keywords are calculated so that keywords with more similar feature vectors are located closer to each other. This makes the visualization a more intuitive representation of the information space.

Interactive Intent Modeling

Our search engine relies on two underlying models: the user intent model and the document retrieval model. The user intent model assigns a relevance score to the current and future search keywords and the document retrieval model uses these probabilities for retrieving the most relevant documents.

The calculation of the relevance scores of the current search keywords is formulated as a multi-armed bandit optimization problem.

- We have observed that the user has given certain feedback to certain keywords. These represent measurements from the multi-armed bandit system.
- We construct a binary feature vector for each keyword based on the documents it appears in.
- Based on this knowledge, we can calculate the probability of each keyword having certain relevance score by using the LinRel algorithm.
- We then approximate the keyword relevance by the upper confidence bound of the corresponding probability distribution. This allows the system to balance between exploration and exploitation.

As on the first iteration there is no actual feedback from the user, the relevance scores are calculated based on pseudo-feedback acquired from the top-ranked documents retrieved using just the initial keyword query.

The relevance scores of future keywords are calculated using imaginary feedbacks to simulate multiple likely future states. The most highly ranked keywords in all of these possible future states are then used to calculate the approximate relevance scores for future search keywords.

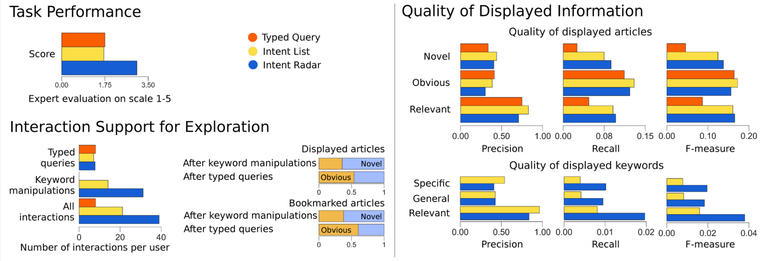
The most relevant documents are found by calculating the document's probability of generating the current keywords in the user intent model. The calculation is based on a unigram language model with Bayesian Dirichlet Smoothing. To expose the user to more novel documents, the documents displayed to the user are chosen by Dirichlet sampling the full ranked list of documents.

Experimental Results

The performance of the Intent Radar has been studied with user experiments (N=30). The task of the users was to bookmark scientific articles relevant to a certain topic and answer to a set of predefined questions related to the same topic. Each user was randomly assigned to one of two different topics and one of three different interfaces.

The three interfaces were

- Typed Query: only keyword search
- Intent List: most relevant keywords were displayed as a list
- Intent Radar: most relevant keywords were displayed with the radar interface



The task performance was measured by the grades given by experts to the predefined questions. Task performance with Intent Radar was clearly superior to the other interfaces (image top left).

The users preferred directing the search with the Intent Radar interface over making keyword queries (bottom left). Furthermore, the documents retrieved after manipulating keywords were more likely to be novel than obvious documents, compared to documents retrieved after making keyword searches (bottom center).

In general, the users found more novel and relevant documents by using the Intent List or the Intent Radar (top right). However, the quality of the keywords displayed to the user was much better in the Intent Radar interface (bottom right).

Further Directions of Research

There is still room for improvement in making the users feel more in control of the search session when using the Intent Radar, as moving a keyword in the radar can cause large changes in the visualized state.

Also, implicit physiological signals could be used to augment the explicit user feedback.

Głowacka, D., Ruotsalo, T., Konyushkova, K., Athukorala, K., Kaski, S., and Jacucci, G. **Directing exploratory search: Reinforcement learning from user interactions with keywords.** In Proceedings of IUI 2013, International Conference on Intelligent User Interfaces, pp. 117–128, New York, NY, 2013. ACM.

Ruotsalo, T., Athukorala, K., Głowacka, D., Konyushkova, K., Oulasvirta, A., Kaipiainen, S., Kaski, S., and Jacucci, G. **Supporting exploratory search tasks with interactive user modeling.** In Proceedings of ASIST 2013, the 76th ASIS&T Annual Meeting, Silver Spring, MD, 2013a. Association for Information Science and Technology.

Ruotsalo, T., Peltonen, J., Eugster, M. J. A., Głowacka, D., Konyushkova, K., Athukorala, K., Kosunen, I., Reijonen, A., Myllymäki, P., Jacucci, G., and Kaski, S. **Directing exploratory search with interactive intent modeling.** In Proceedings of CIKM 2013, the ACM International Conference of Information and Knowledge Management, pp. 1759–1764, New York, NY, 2013b. ACM.