Lost in Publications? How to Find Your Way in 50 Million Scientific Documents

Tuukka Ruotsalo[*]
Helsinki Institute for Information Technology HIIT, Aalto University

Jaakko Peltonen[*]
Helsinki Institute for Information Technology HIIT, Aalto University

Manuel J.A. Eugster[*]
Helsinki Institute for Information Technology HIIT, Aalto University

Dorota Głowacka
Helsinki Institute for Information Technology HIIT, University of Helsinki

Giulio Jacucci
Helsinki Institute for Information Technology HIIT, University of Helsinki

Aki Reijonen
Helsinki Institute for Information Technology HIIT, University of Helsinki

Samuel Kaski
Helsinki Institute for Information Technology HIIT, Aalto University and University of Helsinki

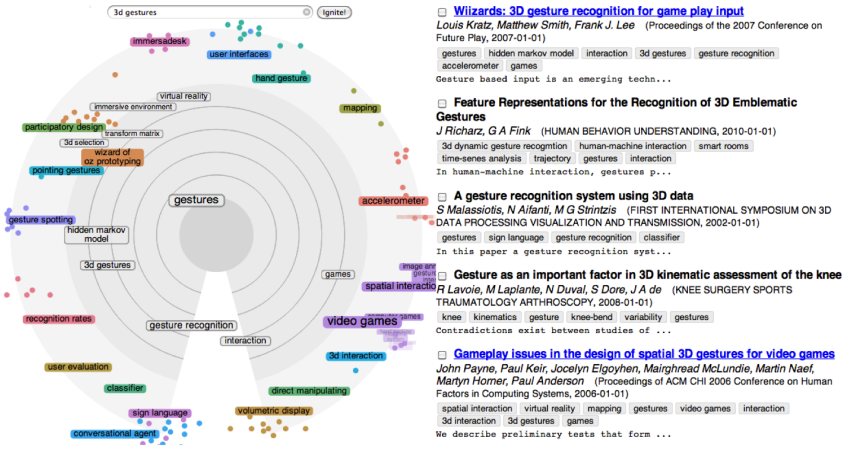
Abstract

Researchers must navigate big data. Current scientific knowledge includes 50 million published articles. How can a system help a researcher find relevant documents in her field? We introduce *IntentRadar*, an interactive search user interface and search engine that anticipates user's search intents by estimating them form user's interaction with the interface. The estimated intents are visualized on a radial layout that organizes potential intents as directions in the information space. The intent radar assists users to direct their search by allowing feedback to be targeted on keywords that represent the potential intents. Users can provide feedback by manipulating the position of the keywords on the radar. The system then learns and visualizes improved estimates and corresponding documents. *IntentRadar* has been shown to significantly improve users' task performance and the quality of retrieved information without compromising task execution time.

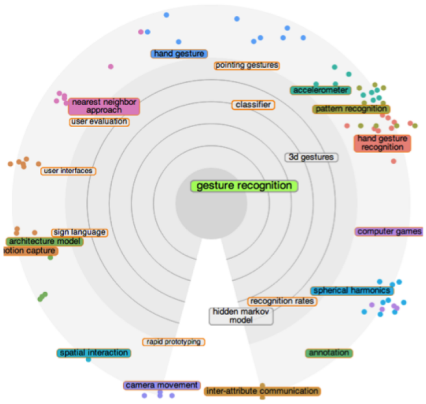
1 Introduction

Exploration and search for relevant data in the available scientific literature are main tasks of a researcher. These tasks are crucial for human analysis of big data, when strong hypotheses about the data are not yet available. Machine learning systems are needed to assist such exploration and search. In big data traditional search solutions become increasingly insufficient. One of the main problems in exploratory search is that it can be hard for users to formulate queries precisely, since information needs evolve throughout the search session as users gain more information. In a commonly observed search strategy, the information seeker issues a quick, imprecise query, hoping to get into approximately the right part of the information space, and then directs the search to obtain the information of interest around the initial entry-point in the information space [8]. Current methods to support users to explore are either based on suggesting query terms, or allowing faster access to the present search result set by faceted browsing or search result clustering [9, 4]. A disadvantage of these feedback mechanisms are that they can trap the user to the initial query context and cause cognitive burden to the user [5].

[*]Equal contributions.



(a) The system makes use of a radar visualization metaphor. The current intent estimation for which the results on the right-side list are retrieved, is visualized for the user on a radar screen (inner dark grey area). The angular distance corresponds to similarity of intents and the distance to their relevance. Predicted intents, that help users to find directions on the radar from their currently estimated intent, are visualized on the outer (light gray) area. The angular distance corresponds to similar intents of the current intent estimation (position is relative to the keywords in the inner dark grey area). The user can provide feedback by dragging concepts closer to the center of the radar.



(b) The user increased the relevance of "gesture recognition" by dragging it to the center of the radar. The system has computed and visualized new estimated relevant intents, such as "pattern recognition", "pointing gestures", "recognition rates", "nearest neighbor approaches", and "hidden Markov models" to continue the exploration.

Figure 1: The *IntentRadar* interface and resulting documents for a query "3d gestures" (a), and the *IntentRadar* visualization after an increasing the importance of "gesture recognition" (b).

We propose that better support for exploration can be provided by visualizing the relevant information space using higher level representations of the data, namely keywords extracted from documents and using reinforcement learning to adjust the visualization [7, 3]. Our system improves interactive search of 50 million scientific articles from Thomson Reuters, ACM, IEEE, and Springer. [1]

2 Search User Interface

The *IntentRadar* interface is presented in Figure 1. It is designed to assist users in exploring information related to a given research topic effectively by allowing rapid feedback loops and assisting users in making sense of the available information around the initial query context. We use radial layout and optimize locations of keywords in the inner circle (representing current intent) and keywords in the outer circle (representing future intents) by probabilistic modeling-based nonlinear dimensionality reduction (see [7] for details).

[1] This work was recently presented at CIKM 2013. [7]

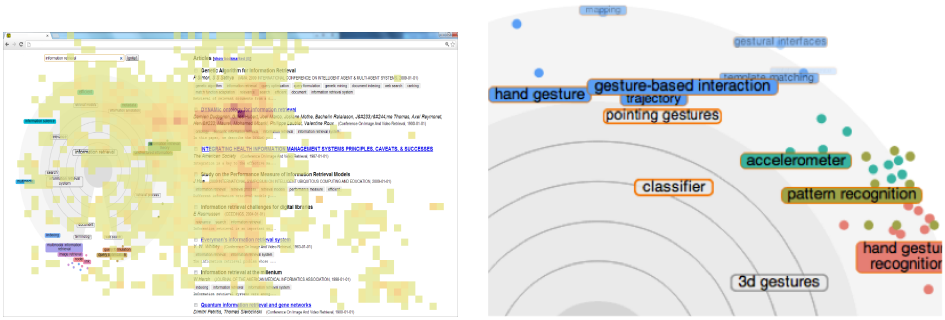


Figure 2: **Left:** Heatmap visualization of eye-tracking data of an exemplar user of the *IntentRadar*. IntentRadar is the main interface element that the user focuses to make sense of the returned information. **Right:** the *IntentRadar* visualization can be inspected in detail using a fisheye lens that follows the mouse cursor and enlarges labels (at top of this figure). This allows users to have overview and detailed view to a specific cluster representing a direction in the information space.

Figure 1 (a) presents a system response to an initial query ”3d gestures”. The system has retrieved a set of documents and visualized the potential intents on the *IntentRadar* visualization. It offers directions towards, for example, ”video games”, ”user interfaces”, ”gesture recognition” and ”virtual reality”. In Figure 1 (b) the user has first selected ”gesture recognition” and is offered further options to continue the exploration towards more specific topics, such as ”nearest neighbor approach”, ”hidden Markov models”, but also towards general topics, such as ”pointing gestures” and ”spatial interaction” that are estimated to be relevant for the interaction history of the user. The interface provides a non-intrusive relevance feedback mechanism, where the user pulls keywords closer to the center of the radar to increase their importance and pushes keywords away from the center of the radar to decrease their importance. The keywords can be enlarged with a fisheye lens that follows the mouse cursor (see Figure 2 (Right)). The radial layout has a good tradeoff between the amount of shown information and comprehensibility compared to alternative visualizations with lower or higher degrees of freedom that could make interaction with the visualization more difficult [2].

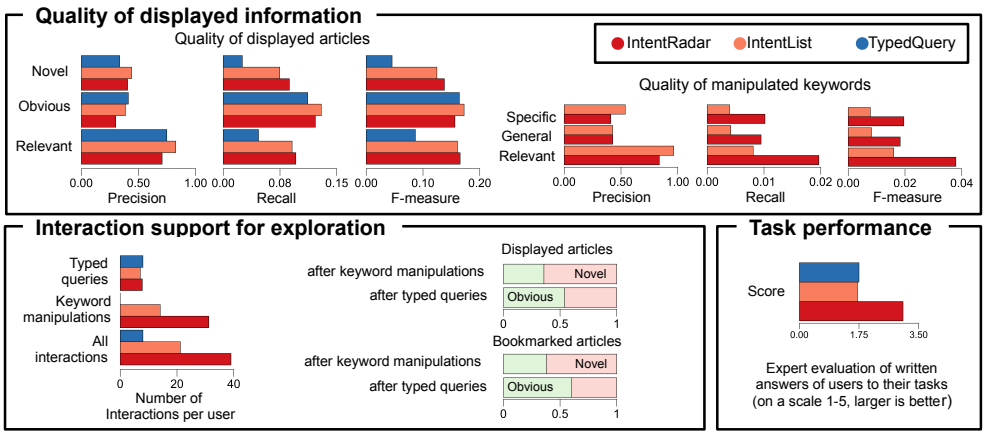
3 Learning Search Intents

The learning of user’s search intents during the interactive search is based on two models: retrieval model and intent model. The retrieval model estimates the probability of relevant documents based on the estimates of the intent model. The intent model estimates the present and potential future intents of the user based on the interaction history.

For the retrieval model, we use the language modeling approach of information retrieval [10]. The estimation is done by a unigram language model with Bayesian Dirichlet Smoothing, evaluating probability to generate the user’s desired keywords given by the intent model. To expose the user to more novel documents we sample a set of documents from the ranked list by Dirichlet Sampling.

For the intent model, we use the *LinRel* algorithm [1]. In each search iteration, LinRel yields an estimate of keyword weights. The simplest strategy would be to select keywords with highest weights given by the regression model, but as the interaction history of the user may provide only limited evidence, this exploitative choice could be suboptimal. Instead, *LinRel* exploratively picks keywords via controlling the exploration-exploitation tradeoff of the estimation. We select keywords with the largest upper confidence bound for the score to be visualized, i.e. maximizing the relevance estimate for the keywords and the uncertainty of the system of the relevance estimate simultaneously. This allows users to benefit from the intent predictions, while at the same time reducing the system’s uncertainty of the estimates. As a result, users can continuously direct their search without getting trapped into the initial query context.

Visualization of search intents is done by nonlinear dimensionality reduction: high-dimensional features of keywords are their predicted future relevance under different user feedbacks, and they are reduced by neighbor embedding to angles of keywords on the radar interface.



Comparison methods above are IntentList (a simplified interface representing intents only as a list of top keywords) and TypedQuery (a traditional search interface based on typing queries). **Improved task performance:** *IntentRadar* improves users' task performance (answers submitted for research tasks) compared to state-of-the-art retrieval methods and the prominent commercial search engine Google Scholar [7, 6]. **Quality of retrieved information:** *IntentRadar* helps users to move away from the initial query context, thus allowing to substantially increase recall while preserving precision in particularly for novel information [7, 3]. **Enhanced interaction:** Despite more complex visualization, users interacted with the *IntentRadar* interface twice to nearly four times more than the comparison systems without compromising the task execution time. Users also used the visualization as their main user interface component during the search sessions (see Figure 2 (left)) [7].

Figure 3: Key benefits of the IntentRadar search user interface

4 Findings

The effectiveness of *IntentRadar* has been studied in task-based experiments where users (30 graduate students from two universities) were asked to solve research tasks using a database of over 50 million scientific articles. The comparisons were conducted against 1) within-system baselines [7, 3] of list-based visualization and only typed-query interaction, and 2) Google Scholar [6]. Experts conducted double-blind relevance assessments of articles and keywords presented by any of the sysystems, on binary scales: *relevance*—is this article relevant to the search topic, *obviousness*—is it a well-known overview article, *novelty*—is it uncommon yet relevant to a given topic/subtopic. The assessments were used as ground truth for evaluations of *user task performance* (assessment of their answers to tasks), *quality of displayed information* (precision, recall, F-measure), *interaction support for directing exploration* (numbers of interactions, information received in response). Full details of the procedures are in [7, 3, 6]. The benefits along with references to the original articles are summarized in Figure 1. The system with *IntentRadar* interface improved user's task performance. The answers that users provided in response to the given search tasks were graded higher by experts. The interface also enhanced interaction. The users of the *IntentRadar* interface initiated up to three times more interaction and the interface reduces users' scanning time with respect to the available option space. Most importantly, interactions with the *IntentRadar* resulted in improved quality of retrieved information (precision and recall of novel information returned by the search engine in response to user interactions during search sessions).

Acknowledgements

This work has been partly supported by the Academy of Finland (Multivire and the COIN Center of Excellence, and 252845) and TEKES (D2I and Re:KnoW). Certain data included herein are derived from the Web of Science prepared by THOMSON REUTERS, Inc., Philadelphia, Pennsylvania, USA: Copyright THOMSON REUTERS, 2011. All rights reserved. Data is also included from the Digital Libraries of the ACM, IEEE, and Springer.

References

[1] P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *J. Mach. Learn. Res.*, 3:397 – 422, 2002.

[2] G.M. Draper, Y. Livnat, and R.F. Riesenfeld. A survey of radial methods for information visualization. *IEEE T. Vis. Comput. Gr.*, 15(5):759 –776, 2009.

[3] Dorota Glowacka, Tuukka Ruotsalo, Ksenia Konuyshkova, Kumaripaba Athukorala, Samuel Kaski, and Giulio Jacucci. Directing exploratory search: reinforcement learning from user interactions with keywords. In *Proc. IUI'13*, pages 117–128. ACM, 2013.

[4] Marti A. Hearst and Jan O. Pedersen. Reexamining the cluster hypothesis: scatter/gather on retrieval results. In *Proc. SIGIR'96*, pages 76–84. ACM, 1996.

[5] Diane Kelly and Xin Fu. Elicitation of term relevance feedback: an investigation of term source and context. In *Proc. SIGIR'06*, pages 453–460. ACM, 2006.

[6] T. Ruotsalo, K. Athukorala, D. Głowacka, K. Konuyshkova, A. Oulasvirta, S. Kaipiainen, S. Kaski, and G. Jacucci. Supporting exploratory search tasks with interactive user modeling. In *Proc. ASIST' 13*, 2013. To appear.

[7] Tuukka Ruotsalo, Jaakko Peltonen, Manuel Eugster, Dorota Glowacka, Ksenia Konyushkova, Kumaripaba Athukorala, Ilkka Kosunen, Aki Reijonen, Petri Myllymäki, Giulio Jacucci, and Samuel Kaski. Directing exploratory search with interactive intent modeling. In *Proc. CIKM'13*, pages nn–nn. ACM, 2013. To appear.

[8] Jaime Teevan, Christine Alvarado, Mark S. Ackerman, and David R. Karger. The perfect search engine is not enough: a study of orienteering behavior in directed search. In *Proc. of SIGCHI*, pages 415–422, 2004.

[9] Ka-Ping Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *Proc. CHI'03*, pages 401–408. ACM, 2003.

[10] Chengxiang Zhai and John Lafferty. A study of smoothing methods for language models applied to information retrieval. *ACM Trans. Inf. Syst.*, 22(2):179–214, 2004.