

Navigation Leads for Exploratory Search and Navigation in Digital Libraries

Róbert MÓRO*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
robert.moro@stuba.sk*

One of the fundamental problems in search in general is how to support the users in formulating their initial as well as any of the subsequent queries. This is even more pronounced in exploratory search where the users can have only a very vague idea of what it is what they want to find in order to solve their problem at hand. Therefore, an exploratory search system needs to support *query formulation* by providing (i) *global overview* of the domain that enables users to quickly comprehend the basic structure of the underlying information space and the possible directions that they can take, and (ii) *means of following different paths* employing the browsing techniques.

In order to support exploratory search and navigation we proposed an approach of exploratory navigation using the *navigation leads*, which represent important words automatically extracted from the documents present in the information space. We distinguish two types of navigation leads:

1. *View navigation leads* which provide a global overview of the domain—what the most important concepts or the most promising leads (paths) to follow are—before the initial query and an overview of the generated view of the domain consisting of the documents of the filtered information subspace at the later stages of search. They are usually placed in a cloud and play a crucial role in formulating an initial query.
2. *Document navigation leads* which highlight terms (keywords) relevant in the context of a single document (search result) as well as the terms with the highest navigational value. They are placed directly in a summary (or abstract) of a document or under it in order for the users to perceive them in their context. They represent navigation starting points allowing to follow a specific path by filtering only documents pertinent to the selected lead. They play crucial role in query

* Supervisor: Mária Bielíková, Institute of Informatics, Information Systems and Software Engineering

refinement and add an aspect of serendipity by allowing the users to follow also unexpected new paths.

The navigation leads augment the result set retrieved by the search engine. The process of augmentation of the search results with the navigation leads consists of three main steps [4]:

1. *Identification of the navigation lead candidates* – in this step, keywords are automatically extracted from the documents and their document relevance is assessed [4].
2. *Selection of the navigation* – besides the document relevance of the individual keywords, their navigational value is computed. It reflects how relevant the lead candidates are for the whole information subspace [3].
3. *Presentation of the leads with the search results* – the selected navigation leads are presented with the search results list, either under the summaries (abstracts) of the search results or in a cloud of terms next to the whole list (in case of the view navigation leads) [2].

The selection of a navigation lead (both view-related and document-related) by a user results in a query refinement and subsequently in a construction of a new partial view of the information space.

Although our proposed approach can be applied in any domain, we focused on its application in the domain of digital libraries, namely in the *researcher novice scenario*. The goal of a researcher novice is not to find specific facts, but to *learn* about the given domain and *investigate* the topics and existing approaches as well as the gaps in the current state of knowledge. We evaluated our approach in the web-based bookmarking service *Annota*¹ [1], which allows the users to organize, collaboratively annotate, and share their bookmarked Web resources with the special focus on research articles. We conducted synthetic experiments on its dataset as well as live experiment with its users as participants.

References

- [1] Holub, M. et al.: Annota: Towards Enriching Scientific Publications with Semantics and User Annotations. *D-Lib Magazine*, vol. 20, no. 11/12, (2014).
- [2] Molnar, S., Moro, R., Bielikova, M.: Trending Words in Digital Library for Term Cloud-Based Navigation. In: *SMAP '13: Proc. of 8th Int. Workshop on Semantic and Social Media Adaptation and Personalization*. Washington, DC, IEEE CS, (2013), pp. 53–58.
- [3] Moro, R., Bielikova, M.: Navigation Leads Selection Considering Navigational Value of Keywords. In: *WWW '15 Companion: Proc. of the 24th Int. Conf. on World Wide Web Companion*. Geneva, IW3C2, (2015), pp. 79–80.
- [4] Moro, R., Vangel, M., Bielikova, M.: Identification of Navigation Lead Candidates Using Citation and Co-Citation Analysis. In: *SOFSEM '16: Proc. of the 42nd Int. Conf. on Current Trends in Theory and Practice of Computer Science*. Springer Berlin Heidelberg, (2016), pp. 556–568.

¹ <http://annota.fiit.stuba.sk/>