

# Navigation Using Annotations in Electronic Documents

Jakub ŠEVCECH\*

*Slovak University of Technology in Bratislava*

*Faculty of Informatics and Information Technologies*

*Ilkovičova 3, 842 16 Bratislava, Slovakia*

`sevo_jakub@yahoo.fr`

We are facing the rise of services for sharing various kinds of content, whether in the form of links to interesting web pages, images, comments or multimedia information. Many applications provide tools for creation of annotations into web pages or various electronic documents, but few of them use these annotations to provide additional value to their authors. If any added value is provided, it takes effect sometime in the future, when the user wants to return to once read documents, or when the amount of created annotations reaches some critical value.

We aim to provide a reward for inserting annotations at the time of their creation. Annotations can be used in different ways to improve information retrieval [1, 2]. We work on supporting navigation between documents using annotations. Namely we use annotations as an input for the process of creation of query, that is used in search engine to search for related documents.

The task of search for documents relevant to source document is very similar to the task of recommending citations to academic papers. Several authors deal with this task [3] and they are searching citation using different methods taking into account context of the document, other characteristics of the document, authors and citations graph. The task of search for documents relevant to annotated document is in several ways similar to citation recommendation:

- source for the process of query creation is another document,
- relevant documents are searched using source document,
- relevant documents are searched for specific parts of source document.

In search using annotations, annotations represent important source information, that can enrich generated query. Annotations highlight the most important parts of documents and they determine the specific topics, user want to get more information for.

---

\* Supervisor: Mária Bieliková, Institute of Informatics and Software Engineering

By annotations we mean comments, highlights in the text, bookmarks or tags that users are often inserting into documents while reading them. They are actually electronic equivalents of marginalia, that we are creating while reading books or newspapers. Annotations inserted into the document, describe exactly the part of the document, that user is the most interested in. That's why we use them to find documents, that provide additional sources of information to information referred in the original document.

Annotations describe users' interests in two ways:

- They describe the interest of user for whole document and they suggest certain level of quality of document. A typical representatives of annotations describing users interest in whole document are bookmarks, where the user explicitly saves documents for later use.
- They specify parts of the document, that are the most relevant for user. Using these annotations we can extract specific topics, user is the most interested in.

Not only annotations inserted by user can be used as source information to the process of query generation. We use the content of the document, document metadata and annotations of other users. In addition to information related to the source document, user annotations attached to different documents can be used as well. Annotations in other users documents form important information about users interests. By means of these annotations we can provide personalized search results for a particular user.

To collect annotations we create a simple tool for manual creation of annotations to web pages by their visitors. Users can add tags, highlight part of web pages, attach comments and bookmark these pages. When user is reading and annotating the document, query is generated and retrieved documents are constantly displayed. When user finishes reading document, he can adjust generated query to better match his requirements and he is able to retrieve related documents using this query.

The main contribution is a method for creating queries to retrieve relevant documents using content of the source document, annotations attached to the document and other users annotations.

*Acknowledgement.* This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/0971/11.

## References

- [1] S. Bao, G. Xue, X. Wu, Y. Yu, B. Fei, and Z. Su, "Optimizing web search using social annotations," in *Proceedings of the 16th international conference on World Wide Web*, 2007, pp. 501.
- [2] S. Xu, S. Bao, Y. Cao, and Y. Yu, "Using social annotations to improve language model for information retrieval," in *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, 2007, pp. 1003.
- [3] Q. He, J. Pei, D. Kifer, P. Mitra, and L. Giles, "Context-aware citation recommendation," in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 421.