Dynamic Linking: Interactive Filtering Of Web Search Results Using Automatically Generated Tags

Marcus Nitsche and Andreas Nürnberg
Data & Knowledge Engineering Group, Faculty of Computer Science
Otto von Guericke University Magdeburg, Germany
{marcus.nitsche, andreas.nuernberger}@ovgu.de

ABSTRACT
Web searches often return results the user is not familiar with. To provide a more transparent view on these search results we propose to use automatically extracted tags for interactive filtering of these result sets. This content-based extraction structures an unknown information space. The concept was developed following an iterative human centred design process, evaluated with potential users testing paper mock-ups and a web-based application in a summative evaluation process. We investigated especially how well the use of this dynamic linking supports the information search process, facilitates overview generation in users’ minds and finally leads them easier to relevant results. The conducted evaluation revealed promising results.

Author Keywords
Search User Interface; Human Computer Interaction; Human Computer Information Retrieval; Tagging; Filtering.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User Interfaces.

General Terms
Human Factors; Design.

MOTIVATION
Commonly used web search engines like Google\(^1\), Bing\(^2\) or Yahoo\(^3\) try to offer the most relevant results to their users. They attempt to map the user’s information need to relevant results. These result sets are usually offered as simple lists of links and short descriptions, distributed on several pages, in most cases without supporting users in filtering these results.

To support users in getting a better overview and enabling a structured exploration of result sets we propose to use interactive tags. This concept supports users in getting better overviews about previously unknown search results, in gaining insights and in filtering result sets faster, in comparison to current forms of result representation.

RELATED WORK
Web Search User Interfaces. Common search engines offer users a horizontal search\(^4\). Search results are displayed as a classical text listing. The ranking is usually calculated by criteria the end-user can not see and in most cases is not even able to understand.

Searching & Filtering. Faceted search user interfaces (SUIs) usually support their users in providing options to filter result sets, e.g. [10] or [9]. VisGets by Dörrk et al. [3], for example, is a web-based SUI, that offers methods to filter news entries via RSS feeds. The filter criteria, called facets, can be used here in different ways: A tag cloud, a time line - represented as a bar chart - for the number of events in single time frames and a map are provided to support dynamic filtering of news documents.

Tagging Information. A tag is a non-hierarchical keyword or term assigned to a piece of information. In the proposed context of using it as a filter for search results the concrete information is a document retrieved by an IR system. Tags help to describe such a document and allow to be found again by browsing or searching. They can be generally chosen informally and personally by its viewer later on. Tags support users in categorizing, classification and re-finding of information [20] and can be part of a result representation, a so called surrogate. Tagging supports users in browsing and exploring information spaces by the compressed information the tag’s titles providing [23]. Tags can be mapped to search queries, existing categories or term frequencies in documents [6]. Therefore, information retrieval methods are applied that are often based on term frequencies per document.

A second possibility of creating tags is social tagging. Unfortunately, the quality of user generated tags strongly depends on the user who created it. Advantages of collaborative annotation processes are a kind of isolation of the describing text.

\(^1\)http://www.google.com, 02.07.2013
\(^2\)http://www.bing.com, 02.07.2013
\(^3\)http://www.yahoo.com, 02.07.2013
\(^4\)A horizontal search is a general search. It is not limited to a certain kind of data like it would be the case in the search functionality of online shopping sites or book searches in library systems.
and the fact that tagged documents have been reviewed by different users from various point of views. Thereby, a rich faceted description of the document’s content can be created. Even UI adaptations using social tagging were successfully applied in [12]. According to [8] the trend of social tagging is also a suitable technique to support (web) searching.

CONCEPT OF USING TAGS AS INTERACTIVE FILTERS
Special requirements to support filtering and result reviewing functionality are:

- Domain description - To understand the content and domain coverage of the returned results, a user is usually trying to get an overview about presented results. This can be a time-consuming process. Therefore, content based tags - generated from search results - are used to facilitate the classification and evaluation of the result information space by using them as dynamic links.

- Displaying context information - Next to a document surrogate, that includes title, snippet, URI and further information of a certain result, context information in form of corresponding tags should be offered. Clicking on a tag reveals all connected results by highlighting them and vice versa all related tags are highlighted when hovering rather clicking on a certain result.

- Presenting similar results - Users expect to receive results according to their queries. Therefore, a tag the user is hovering over highlights results covering similar content. These results are displayed when a user opens a result in preview or full screen mode (Fig. 4, from [4]).

The concept provides interaction support in two directions. In (1) user starts from choosing tag(s) (Fig. 1, from defence of [4]):

1. User chooses from a list extracted tags on top of the UI to highlight the results covering these terms.
2. By this, a user gets a rough separation in relevant and irrelevant documents. It might also support users in result familiarity if he has previously visited the page of interest, since the search results are visualized within a surrogate including title, URL and a text snippet.
3. Clicking on a single result opens a preview window that further supports visual re-finding and familiarity.

The other interaction direction of the concept allows (2) users to select specific search result(s) first (Fig. 2, from defence of [4]):

1. User chooses result from ranked list (displayed as tiles).
2. After marking the result the corresponding tags included in this result are highlighted on top.

FEATURES OF THE PROTOTYPE
While implementing a prototype we considered Shneiderman et al.’s requirements towards SUIs [17]. The ideas for interactive visualizing relations between tags and results (Fig. 1 & 2) enhance standard SUIs to support users in information search. Examples are brushing and linking techniques [5] that support visual search of users [18]. The design of the user interface should be consistent [17]. A subjectively evaluated criteria like aesthetics of design influences its use. Ben-Bassat et al. [1] showed attractive aesthetics lets users perceive a higher quality of usability than it actually is. Parush et al. [15] even demonstrated that aesthetics has actually an influence on processing time.

Well-established is a central alignment of the input before presenting results and a consistent alignment above presented results. After an initial input, users find on top a possibility to type in or to refine a certain query. Below, a field of 3 x 7 interactive tags is offered to the user. The overview of tags is geared to the reading direction of the western world (left to right, top to bottom). The tag overview works as an assistance for information search. In other applications like PubCloud [11] a comparable tag cloud has been positioned at the margin. Below the tag field the relevance ranked results are offered to the user. Thereby users do not need to skip pages. By an infinite scrolling support users just need to scroll down to load further results.

The proposed SUI (Fig. 3, from [4]) offers users common interaction support for navigation, selecting and searching. Addressing user’s expectations and conventions [7] support user’s acceptance and first encounter with novel interaction methods like the presented here. Common known methods like mouse and keyboard interaction are supported. This is important since approximately 73% of our test users are use such devices. Tags are generated from analysing the result snippets (for performance reasons) provided by the Bing Search API, which we used as back-end technology. We applied standard IR algorithms based on term frequency together with stop-word-elimination and word stemming to ex-
tract a relevance ranked list of potential describing terms. For visualizing the content connected to specific tags, we use features that are easy to interpret, such as relative font size and the order of tags can be mapped to their relevance. By this, users receive more information about landing pages to facilitate more informed decisions.

The proposed visual features are based on the laws of Gestalt by Wertheimer [21, 22]. A problem is that the length of a certain term influences the user’s perception of its relevance. An irregular alignment of terms / tags would affect positively a feeling of exploration, but would also detracts the optical flow [6]. This approach tackles the disadvantage of currently used tag clouds, since they just support searching for a single term [6]. Either the creation of intersections nor the support for visualizing relations is provided usually.

**EVALUATION OF DESIGN & RESULTS**

According to a UCD process a formative and a summative evaluation process has been conducted. Thereby, a total number of 26 test users were asked in four scenarios to test the concept. A first pre test was conducted using a paper mock-up [16, 19]. 12 of the 26 users were male, the rest female. The average age was 29 years. In average, the test users spend half of their life in using computers (in average for 15.48 years) and for ten years in average they used the WWW.

Users used results more than the generated tags. Only 25% worked frequently with tags. If they used them they mentioned they would try to filter or to refine the search results, which successfully addresses the core concept. Furthermore, our potential users said they would try to compress the presented information and to classify them. Users found the presentation of tags surprising since they are placed between input and output of the search engine. The users seemed not to read a lot of the tags and worked just with a few of them. This can be explained by terms that already were similar to each other (e.g., same word stems) and the used stop words we hard coded. Working with just a few of them is also confirmed by some users’ statement the number of tags, provided by the system (21), might be to many. Overall, 56% think the number of tags is appropriate. Most users wished to have an ordered tag listing (81%) and the highlighting of the tags was looked upon favourably (43% very good, 37% adequate).

The bi-directional mapping of tags (see Fig. 5, from [4]) and search results was liked by 87% of our test users (50% found it useful, 25% very good and 12% helpful). This perception might be caused by the visual feedback the system is supporting. We also observed that users are getting a quicker awareness of the tagging functionality through this visual feedback. 87% of the users we asked found this feature supports them in their information search process (31% found it very supporting, 56% said yes or said that it would support them a little bit).

**CONCLUSION & FUTURE WORK**

The developed SUI for interactive filtering web search results was evaluated twice by using paper mock-ups and interactive test systems with a total number of 26 users. Tag overviews can be seen as a facet by which users can categorize, filter and get a feeling of the search result’s context. Usability is not only depending on supported functionality, but also on an aesthetic design. Here, features like color, alignment and contrast, have been taken into account to support an effective and efficient use of the SUI. Brushing and linking of tags and search results are supported by bluring, a simple effect based
on contrast and brightness. Overall, the unconventional use of tags and interactive filtering, this approach seems to be a promising concept to support exploration of result sets.

Future work covers a better generation of tags using enhanced algorithms, which we not focused on in this work. Dynamic contextual result sets should be supported, e.g., as described in [13]. Tags could also be derived from other sources, such as social tagging platforms [8, 2]. Alternative search result grouping and visualization techniques can be discussed. Furthermore, the creation of subsets and intersections of results, generated via tags, should be enhanced. Advancements towards supporting Noël’s [14] view or interaction in the 3rd dimension of exploratory search should be addressed. E.g., by using related result sets to offer users a jump into another domain.

ACKNOWLEDGEMENTS
We thank Tom Grope for implementing the concept and evaluating it as part of his Bachelor’s Thesis. Part of the work is funded by the German Ministry of Education and Science (BMBF) within the ViERforES II project (no. 01IM10002B).

REFERENCES