Vague Query Formulation by Design

Marcus Nitsche
Faculty of Computer Science,
Otto-von-Guericke-University,
Germany
marcus.nitsche@ovgu.de

Andreas Nürnberger
Faculty of Computer Science,
Otto-von-Guericke-University,
Germany
andreas.nuernberger@ovgu.de

ABSTRACT

When users search for information in domains they are not familiar with, they usually struggle to formulate an adequate (textual) query. Often users end up with repeating re-formulations and query refinements without necessarily achieving their actual goals. In this paper we propose a user interface that is capable to offer users flexible and ergonomic interaction elements to formulate even complex queries in a simple and direct way. We call this principle *vague query formulation by design*. By this formulation we like to point out its design-driven origin. The proposed radial user interface supports phrasing and interactive visual refinement of vague queries to search and explore large document sets. The main idea is to provide an integrated view of queries and related results, where both queries and results can be interactively manipulated and influence each other. Changes will be immediately visualized. The concept was implemented on a tablet computer and the usability was stepwise evaluated during a formative and a summative evaluation process. The results reveal high usability ratings, even if the concept was completely unknown to our test users.

Keywords

Search User Interface, Query Reformulation, Query Refinement, Exploratory Search User Interface, Information Retrieval.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval.; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces.

General Terms

Design, Human Factors, Management.

1. INTRODUCTION

When users try to handle complex information needs they often end up in conducting exploratory searches [11]. One of the main characteristics of exploratory searches is that users often do not know how to formulate their information need. Often this problem coexists with an unfamiliarity with the domain they search in [17]. In this work we like to tackle this problem of formulating appropriate queries by offering dynamic user interface (UI) elements that users can manipulate directly by touch gestures to give them *a feeling* for a certain query configuration that matches a certain result set. Thereby learning and exploring aspects will be covered as well [17, 11]. This concept of interactive visual filtering of relevant information in a more natural way enables data processing in cases, where standard algorithms can not be applied since these algorithms might filter out relevant data. We introduced the concept of this paper back in 2011 [15], where we described the basic idea and did some pre-studies with a digital mock-up prototype. In this paper, we first introduce a running implementation and a more detailed user study towards this concept. Therefore we present some related work aspects in Section 2, followed by a presentation of the UI concept in 3 and the description of the implementation, evaluation concept and results of the final user study in Section 4. Finally, we conclude and discuss possible future work in 5.

2. STATE-OF-THE-ART

User-specific context aware data filtering is not a new challenge. In the following we show two tools, that can also be used for this application. The VIBE-system [10, 16] supports users in finding relevant information using magnets to attract relevant documents to specific screen points (Fig. 1).

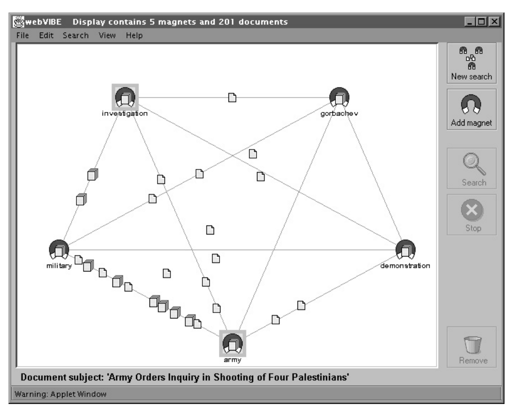

Figure 1: webVIBE, a variant of [10, 16].

This system follows the principle of *dust-and-magnet* [18]. Our proposed concept uses this principle also - as one aspect of the interaction concept. In contrast to VIBE we offer users of our system an interactive visualization without any classical WIMP-interface elements (Windows, Icons, Menus, Pointer). By this, no virtual mapping of functions is necessary and users might be able to use the interface in a more firm and reliable way. Cousins et al. [5] developed a system that follows a direct manipulation approach like done here. But in contrast to our proposed solution it is divided into different UI elements and different views. It is less integrated in a single view. Therefore user's work load might be higher since he needs to face various mode switches. Commercial systems, like the Vis4you concept[1], are more focused on visualization than on interaction via direct manipulation. Furthermore, their system is designed to be used on desktop computers with a mouse (*single point and click*-principle), no multi-touch-support. In the next section we like to present our concept in more detail.

3. CONCEPT & DESIGN

Due to the increasing amount of data and complexity, it is necessary to apply and improve the concepts of visual information filtering and retrieval. This goes along with the underlying methods and tools. Considering clustering algorithms (e.g., k-means [3]), we thought about the concept of *vague query formulation*: Since users sometimes do not know what they are searching for, we like to support them by the opportunity to formulate vague queries. Here the user is asked to narrow the search results by dragging user interface (UI) elements, so called *widgets*, with query terms, see also Fig. 2.

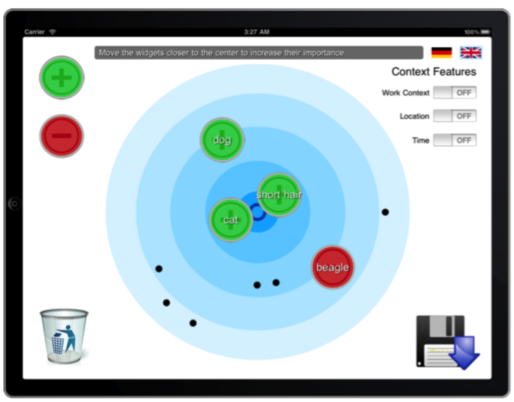

Figure 2: Radial design of the implemented UI.

The concept follows the idea that more relevant data are centred. Note, this is equivalent to filtering an overcrowded desktop, cf. Fig. 3 (left picture)[2], where the more centralized documents are possibly more important (highlighted in the right picture).

The system was designed to be a multi-user system. Therefore a number of multiple users need to be supported at the same time, also considering security aspects [14]. To offer each user the same possibility to interact with the system we use a radial form for the interface layout. Furthermore, an underlying multi-touch device is a hardware requirement, that enhances the combination of tool and application domain significantly. Another appealing advantage is, that multi-touch also supports users in a more natural way of interaction [9]. Other radial user interfaces for selecting or filtering often offers fixed places for items. In contrast to this our system is supposed to be more flexible since users are allowed to position their query widgets where they like.

We offer users a dimension merging according specified weights, similar to the result listing of search engines, where also different

Figure 3: Desktop: more relevant documents are centred.

weights can be linked to specific query terms (Fig. 2). Data points represent the data space. Query objects (widgets) can be entered via a virtual keyboard and can also be dragged by the user to formulate more complex or vague queries. Selecting a specific data point supports the user with additional information on this data point and highlights all further related data points.

The distance of a certain term is directly connected to its importance for the user. In other words, if a user thinks a specific term is more relevant to its actual filter/search task, he positions the corresponding UI element nearer to the center, which influences the weight of this term when computing its Term Frequency / Inverted Document Frequency (TF/IDF)-value [2], which in fact is a calculated weight to influence the ranking of the data space and this in return effects the visualization (Fig. 6). Thereby, users do not need to specify a concrete position of UI elements on the screen, we support this by a non-determined precision. The widget-induced relevance of a query term is calculated according to the formula in Fig. 4.

$$R_{Widget} = \frac{d_{Center\ Point}}{r_{Search\ Area}}$$

Figure 4: Widget-induced relevance of a query term.

Result elements are placed near to corresponding query elements. The formula for calculating the relevance of a SearchResult object (result dot) is shown in Fig. 5.

$$R_{Search\ Result} = \frac{\sum_{i=0}^{\#\ relevant\ widgets} R_{Relevant\ Widget_i}}{\#\ relevant\ widgets} - \left(\frac{\sum_{i=0}^{\#\ non\ relevant\ widgets} R_{Non\ Relevant\ Widget_i} + \sum_{i=0}^{\#\ relevant\ exclusion\ widgets} R_{Relevant\ Exclusion\ Widget_i}}{\#\ non\ relevant\ widgets + \#\ relevant\ exclusion\ widgets} \right)$$

Figure 5: Relevance of a SearchResult object.

The calculated relevance determines the distance to the center, considering further result objects.

To address various types of end devices such as multi-touch desktops or mobile interfaces with large displays, we use direct manipulation as a central interaction paradigm. Only the relative distance of an UI element to the center is relevant for the system. Thus, we provide users with a direct linking to the data they like to filter. By this interaction concept, we propose to achieve more precise results. Additionally, we support users with the concept of *What-if*-queries, which supports a fault-tolerant interaction system, using a ghosting technique: Dragging an element and holding it on a specific position triggers the system to show the user how many items are in the center point of interest (POI) after releasing the element. Thereby, users are able to explore the impact of possible next steps.

[1] http://www.vis4you.com/vis4you/ (accessed on 04.07.2012)
[2] http://lawprofessors.typepad.com/ (accessed on 04.07.2012)

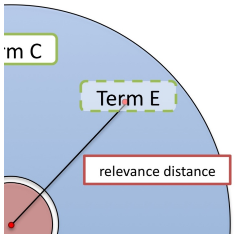

Figure 6: Concept of relevance mapping.

Changes of the query configuration also effect the data points to provide the user with a direct link to the data (interactive visualization). By the underlying metaphor of magnets, we offer an integrated feedback, comparable to *Dust-and-Magnet* [18]: When users drag a specific UI element to a certain point, relevant data points follow this UI element. Data points that have the same TF-IDF value (equal relevance to a query configuration) are drafted with a minimal distance to each other to minimize the possibility of occlusions.

3.1 Features

The UI supports direct feedback since the relevance value is simultaneously shown while users interact with the widget (Fig. 7).

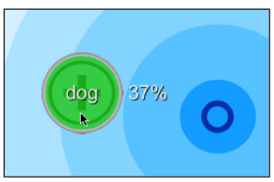

Figure 7: Direct feedback: relevance value next to the widget.

Results, corresponding to a specific query object are visually highlighted and grouped to each other (Fig. 8).

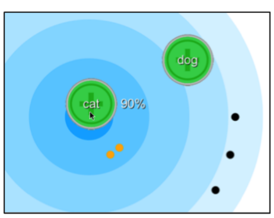

Figure 8: Corresponding results are visually highlighted to group them (e.g. highlighted results for the search term 'cat').

Detailed information on particular result objects, like a website preview, is provided after clicking on the result dot (Fig. 9).

4. IMPLEMENTATION, EVALUATION & RESULTS

Since this contribution is basically driven by fields of human factors and user interface design, we are using common methods from these research areas. Such as user centred design (UCD) processes [7], formative evaluation methods [12], questionnaires [6], think-aloud-protocols [8], and cognitive walkthroughs [4].

To proof the concept of the proposed user interface, a prototype was implemented. This was done by using an Apple iPad. Therefore the application was written in ObjectiveC using the xCode environment[3]. The backend architecture is the CARSA system [1], an information retrieval framework for research purposes. For a detailed overview about the system's architecture see Fig. 10.

Figure 9: Prototypical search result popover as a website preview feature, here a result for 'Labrador Retriever'.

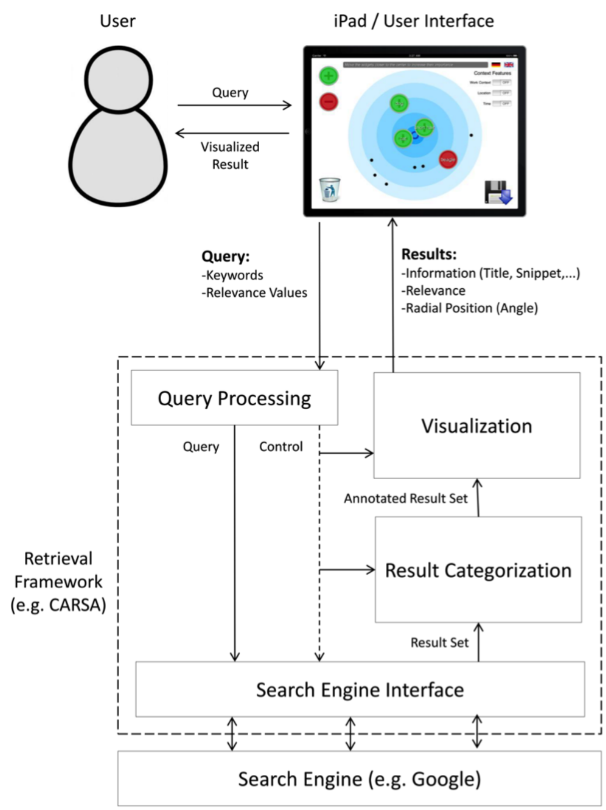

Figure 10: System architecture & UI interaction, cf. [1].

The evaluation concept followed a formative evaluation concept where several usability testings were conducted. Also in parallel to the development process: To identify at least 85% of all usability issues this mock-up was evaluated according to Nielsen and Landauer [13] with only a small number of test users since most usability issues will be mentioned repeatedly by users. The sixth tested user would report new usability issues in only 15% of all cases.

[3]developer.apple.com/xcode/ (accessed on 04.07.2012)

Therefore we decided to ask only eight users. The results of this first user test seem to be promising that this concept works as desired. Users were introduced in the main features and were asked afterwards to formulate a filter query consisting of three terms to find all relevant documents while visualizing most important relations to other potential interesting data. After going through a cognitive walk-through of a movie filtering task our eight test users (six male, two female, average age: 23.4) answered seven usability questions by filling out a 7-step Likert scale from 1 (very bad) to 7 (very good). Next to cognitive walk-throughs, we used think-aloud-protocols and questionnaires. The usefulness of the prototype was rated high, the functionality was praised by test users, performing tasks were rated as *very easy* and test users were satisfied with this prototype. Terminology, attractiveness, and consistency were rated lower. Our final evaluation revealed the results you can see in Fig. 11. Even if there is room for improvement the results reveal overall a good usability, several test users mentioned that it was fun to use it, which might is reflected by a high rating of *joy of use* measurings.

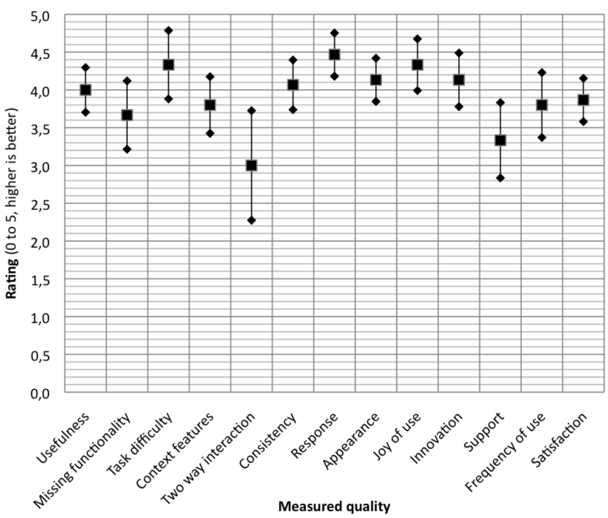

Figure 11: Results of final usability testing.

5. DISCUSSION & OUTLOOK

We described a newly designed user interface concept for filtering, exploring and managing data via direct manipulation, supporting multiple reference systems and context sensitive interaction techniques. We proposed a prototype for visual filtering, that is

- flexible: parameters can be adapted or enhanced by users
- context-sensitive: initial parameters are extracted from the current use case
- easy to learn: through work environment metaphor and direct manipulation

In the near future, a more detailed and larger user study will be conducted to identify further room for improvements of our tool and the overall concept. We also plan to re-design it slightly.

6. ACKNOWLEDGEMENT

Part of the work is funded by the German Ministry of Education and Science (BMBF) within the ViERforES II project (no. 01IM10002B). We also thank Martin Schemmer for the implementation of the presented concept during his diploma thesis.

7. REFERENCES

[1] Bade, K., De Luca, E. W., Nürnberger, A., Stober, S.: CARSA - an architecture for the development of context adaptive retrieval systems. In: Proceedings of Adaptive Multimedia Retrieval: User, Context, and Feedback, Volume 3877/2006, Lecture notes in computer science, pp. 91-101, Springer Berlin / Heidelberg (2006).

[2] Baeza-Yates, R. and Ribeiro-Neto, B.: Modern Information Retrieval, pp. 29-30. Addison Wesley / ACM Press, NY (1999).

[3] Bradski, G. and Kaehler, A.: Learning OpenCV Computer Vision with the OpenCV Library. O'Reilly, p. 479 (2001).

[4] Busemeyer, J. R.: Choice behavior in a sequential decision-making task. In: Organizational Behavior and Human Performance 29 (2), pp. 175-207(1982).

[5] Cousins, S. B., Paepcke, A., Winograd, T., Bier, E. A., Pier, K.: The digital library integrated task environment (DLITE). In: Proceedings of the second ACM international conference on Digital libraries (DL '97). ACM, New York, NY, USA, pp. 142-151 (1997).

[6] Czaja, R. and Blair, J.: Designing Surveys. Pine Forge Press. A useful resource for factual-style surveys, including material on interviews as well as mail surveys (1996).

[7] Eason, K. D.: User centred design for information technology systems. In: Physics in Technology 14 (5), p. 219 (1983).

[8] Ericsson, K. A. and Simon, H. A.: Verbal reports as data. In: Psychological Review 87 (3), pp. 215-241 (1980).

[9] Han, J. Y.: Multi-touch interaction wall. In: Proceedings of ACM SIGGRAPH 2006 Emerging technologies (2006).

[10] Koshman, S. L.: VIBE User Study Technical Report LS062/IS97001, University of Pittsburgh (1997).

[11] Marchionini, G.: Exploratory search: from finding to understanding. In: Communications of the ACM 49 (4), pp. 41-46 (2006).

[12] Moxley Jr., R. A.: Formative and non-formative evaluation. In Instructional Science 3(3), pp. 243-283 (1974).

[13] Nielsen, J. and Landauer, T. K.: A mathematical model of the finding of usability problems. In: Proceedings of ACM INTERCHI'93 Conference, pp. 206-213, Amsterdam, The Netherlands (1993).

[14] Nitsche, M., Dittmann, J., Nürnberger, A., Vielhauer, C., Buchholz, R.: Security-relevant Challenges of selected Systems for Multi-User Interaction. In Proceedings of the 7th International Workshop on Adaptive Multimedia Retrieval (2011).

[15] Nitsche, M. and Nürnberger, A.: Supporting vague query formulation by using visual filtering. In Proceedings of Lernen, Wissen, Adaption (2011).

[16] Olsen, K. A., Korfhage, R. R., Sochats, K. M., Spring, M. B., and Williams, J. G.: Visualization of a Document Collection: the VIBE System. In: Information Processing & Management, 29(1), pp. 69-81 (1993).

[17] White, R. W. and Roth, R. A.: Exploratory search: Beyond the Query-Response paradigm. In: Synthesis Lectures on Information Concepts, Retrieval, and Services, Ed.: G. Marchionini, Morgan & Claypool Publishers (2009).

[18] Yi, L. S., Melton, R., Stasko, J. and Jacko, L.: Dust & Magnet: multivariate information visualization using a magnet metaphor. In: Information Visualization, pp. 239–256 (2005).