

Designing for Exploratory Search on Touch Devices

Khalil Klouche^{1,3}, Tuukka Ruotsalo², Diogo Cabral¹
Salvatore Andolina¹, Andrea Bellucci⁴, Giulio Jacucci^{1,2}

¹Helsinki Institute for Information Technology HIIT, Department of Computer Science,
University of Helsinki, PO Box 68, 00014 University of Helsinki, Finland

²Helsinki Institute for Information Technology HIIT, Aalto University,
PO Box 15600, 00076 Aalto, Finland

³Aalto University, School of Arts, Design and Architecture, Media Lab Helsinki,
Hämeentie 135 c, 00560 Helsinki, Finland

⁴Universidad Carlos III de Madrid, Av. Universidad 30, 28911 Leganés (Madrid), Spain

¹first.last@helsinki.fi, ²first.last@aalto.fi, ⁴abellucc@inf.uc3m.es

ABSTRACT

Exploratory search confront users with challenges in expressing search intents as the current search interfaces require investigating result listings to identify search directions, iterative typing, and reformulating queries. We present the design of Exploration Wall, a touch-based search user interface that allows incremental exploration and sense-making of large information spaces by combining entity search, flexible use of result entities as query parameters, and spatial configuration of search streams that are visualized for interaction. Entities can be flexibly reused to modify and create new search streams, and manipulated to inspect their relationships with other entities. Data comprising of task-based experiments comparing Exploration Wall with conventional search user interface indicate that Exploration Wall achieves significantly improved recall for exploratory search tasks while preserving precision. Subjective feedback supports our design choices and indicates improved user satisfaction and engagement. Our findings can help to design user interfaces that can effectively support exploratory search on touch devices.

Author Keywords

User Interfaces; Exploratory Search; Touch Devices

ACM Classification Keywords

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

INTRODUCTION

Surface computing technologies hold great potential for enhancing information retrieval activities. Devices with touch interaction capabilities make possible to design engaging direct manipulation interactions, facilitate awareness of information available for the user beyond conventional search engine result pages, and afford visualization and spatial organization of content. However, conventional search user interfaces rely exclusively on typed-query interaction and result presentation as ranked list of documents [24], and thus they present challenges when transferred to touch devices.

These design conventions are pernicious for many search scenarios, in particular, *exploratory search* scenarios, that describe a class of search activities that go beyond basic lookup, typically involving the user in a field with which she is not familiar [38]. Exploratory search patterns are very diverse, but share together common complex user-centered challenges: The need to overcome difficulties in formulating queries in unknown information spaces, ways to learn about the information space and identify possible search directions beyond the entry point specified by an initial query [28].

Causes of inadequacy of classical search interfaces for touch-enabled devices are the poor substitutes for keyboard and mouse inputs. Virtual keyboards are reported to be less performing than their physical counterpart [34], and do not provide usual text editing shortcuts (e.g., copy, cut, paste, cancel). As for mouse-based interactions, touch-based substitutes constrain natural touch interactions and prove difficult for quick and accurate text selection [10]. Also, the lack of window management on touch devices does only allow the visualization of a single query at a time, which hinders comparison and revisiting previously retrieved information.

As a consequence, the fluidity and search performance expected while searching with surfaces is hampered [38]. Therefore, it becomes crucial to design new solutions that overcome the limitations of conventional search user interfaces and bring in forefront the potential of multi-touch interaction.

We present the design of Exploration Wall, a novel search user interface for facilitating exploratory search tasks on touch devices. Exploration Wall is based on the following design principles targeting above-mentioned challenges:

1. Flexible reuse and combination of items to facilitate query formulation.
2. Result sets of not only documents but most relevant entities to foster iterative query reformulation.
3. Use of spatial configuration of multiple search streams to identify search directions and learn about the information space.

Our design was found to facilitate exploratory search behavior when compared to the conventional baseline search user interface, as indicated by measured system effectiveness. Moreover, users were found to be more engaged with the task and subjectively more satisfied by their exploratory search. Our findings suggest that our principles can be effective when designing search user interfaces for touch devices, and can overcome many limitations of the direct adaptation of conventional search user interfaces to surfaces.

BACKGROUND

Exploratory Search

Most available tools for information retrieval focus on lookup retrieval, such as looking up the address of a restaurant or reminding a historical fact, while many users search to solve more complex tasks that require exploration of the information space. White [38] describes exploratory search as activities that move beyond basic lookup retrieval. Such activities rely on learning and investigation [22]. Exploratory search activities have no predetermined goals and are described as open-ended [38]. Therefore, the absence of clear user intents leads to difficulties in formulating queries.

Exploratory search processes are considered dynamic. As the information space is unknown or unfamiliar to the user, the query formulation evolves iteratively as the user becomes more familiar with the context [7]. In addition, exploratory search tasks include cognitive and behavioral attributes [40]. Cognitive attributes can be defined as those that have the reasoning associated with conducting an exploratory search and involve learning and investigation as goals; general and ill-structured problems; uncertainty; dynamic, multi-faceted and complex search tasks and are accompanied with sensemaking, decision making or other cognition.

As stated, in the context of exploratory search users need particular support in formulating queries, learning about the information space and identifying possible search directions. Next we review how visual interfaces have been developed to address these issues.

Visual Interfaces in Search

Recently, a variety of search systems have been developed in order to enable faster relevance judgment and effective feedback [12, 17, 19, 23, 32]. Several visualization approaches have been explored including multiple linked lists, scatter

plots, graphs and their combinations [31, 19]. These types of visual search systems are distinguished from familiar query composition ones (e.g., Project Blacklight [30]) because of their emphasis on rapid filtering to reduce result sets, progressive refinement of search parameters, continuous reformulation of goals, and visual scanning to identify results [1].

Currently, visual approaches attempt to better support exploration in different ways: supporting sense making by incrementally and interactively exploring the network of data [8], showing how visualization support user involvement in the recommendation providing rationale behind suggested items [35], visualizing relations of different queries and result sets [2]. Recent work shows how to support users to view and manipulate their search intent models as viewed by the system [2, 3, 29, 11]. This work attempts to combine personalization of search with visualization approaches offering support to formulate queries and learning about the information space helping users in directing their search [29]. While these systems demonstrate the importance of investigating visual user interfaces in exploratory search, they have not been considered for multi-touch devices.

Recent visual interfaces in search have shown the effectiveness of interacting visually with query elements and results combined with computational techniques that support exploration. Multi-touch devices could provide an opportunity for visual interactive search for their capability to encourage manipulation of visual elements and for the limits posed by the absence of mouse and keyboard. In the next section we inspect previous work on search interfaces on touch-enabled devices.

Search with Touch Devices

The workshop on exploratory search and Human-Computer Interaction at the 2007 CHI conference [37] demonstrates the interest in the research community on extending search interfaces to new kind of interactive environments. One of the discussed topics, in fact, was about the need to better understand how to design exploratory search systems for beyond-the-desktop interaction. There is a raising need for search system designs on interactive displays that take advantage of the idiosyncrasies of multi-touch interactions [39] instead of simply being directly ported from desktop-based interfaces.

One of the critical issues is to reduce the need for text entry, which is not suitable for touch-based interaction [42]. In the case of a small form-factor, FaThumb [18] explores web browsing on mobile phones by providing an interface that exploits facet navigation and limits text entry only to further narrow search results. Findings from the user evaluation demonstrated that text entry is more efficient for direct search while for open-ended search facet navigation offers better performances. Multi-touch gesture-based interaction has also been exploited as a mean to improve targeted search of specific content. For instance, Gesture Search [20] is a tool that allows users to define personalized touch gestures to quickly access data items on their mobile phone. The Questions not Answers (QnA) [16] prototype is an interesting instance of a system that exploits social interactions and context-awareness capabilities of mobile devices in order to offer reuse of previous

queries based on geographical locations. The system maps other users' queries to their physical location and provides an interface to display them on an interactive map. Results from the user study demonstrated that displaying previous query result in less need to formulate new query or to a better formulation, which is influenced by the displayed queries.

Concerning large surfaces, like in public display settings, surface computing has been examined particularly for open-ended information exploration [14, 9]. Especially interesting cases are the EMDialog [13], which provides a visual exploration environment for an artist's work in a museum via temporal and contextual dimensions, and the Bohemian Bookshelf [33], which is designed to systematically support serendipitous discoveries while searching in a book collection through different interlinked visualisations. Insights from these research works led to several design principles for search interfaces in public spaces, such as combining different search strategies, rewarding short-term and long-term exploration and making information exploration appealing through engaging multi-touch interaction and information visualizations. Research literature emphasized on the use of large interactive surfaces for collaborative search tasks. Morris et al. [24] have conducted extensive empirical research on collaborative information seeking using horizontal surfaces, providing discussions on opportunities, challenges and design principles for the development of co-located search systems on tabletops along different dimensions such as collaboration styles, search input, group size and application domains. Efficient text entry by enabling the reuse of existing text instead of typing all searches on a virtual keyboard is also a leading design goal in the case of large surfaces [25].

From the state of the art, it is manifest that current solutions are modelling search systems for small or large surface and there is a lack of investigation on touch interfaces for medium-sized display screens, such as tablets. They present different affordances if compared to smaller or larger form-factor, and therefore need a different design approach. For instance, they do not support collaborative tasks as large surfaces do—given the limited screen size—but the display dimension is still bigger than mobile phones—while supporting users mobility—thus allowing richer visualisations, arrangements of interface elements and touch-based manipulations (e.g., two-handed gestures).

DESIGN CHALLENGES

From the state of the art, we identified three main challenges:

1. Formulating queries in unknown information spaces

Activities considered as related to exploratory search are very diverse and hard to define in a consolidated way. But unlike basic lookup search, they usually take place in areas that are unfamiliar to the user and are characterized by the frequent need to reformulate the query.

2. Learning about the information space and identifying possible search directions

The spatial metaphor of exploration applied to information retrieval well describes the need for steering the exploratory evolution in the information space. Narrowing possibilities

to make steering decisions implies continuous gain of topical information.

3. Going through long lists of results with low information gain

In current search systems, users are forced to invest significant cognitive efforts in acquiring cues to formulate queries from the intermediate results, instead of focusing on collecting and learning from relevant information. Long lists of results conflict with the idea of a dynamic steering by slowing down the iterative process of query reformulation.

4. Typing and manipulating queries

Exploratory search activities performed on a traditional typed-query search interface require a lot of text entry and text manipulation. On touch devices, text manipulation is made difficult by the absence of a physical keyboard, hotkeys or shortcuts, and the lack of an accurate selection tool.

GENERAL DESIGN PRINCIPLES

From the above mentioned challenges, we came up with targeted design principles to guide the implementation of our prototype:

A. Flexible reuse and combination of information items to facilitate query formulation

To reduce the need for text entry (challenge 4), we chose to itemize information into entities of different types that can be flexibly manipulated and "dragged around" to support and facilitate all fundamental tasks like selection, duplication, grouping, deletion. Entities would be used to formulate queries, either individually or combined, to get a set of new entities as search results. An existing query could then be easily refined or reformulated by addition or removal of such entities and the results would update accordingly.

The possibility to input text is still necessary in some situations, for example when the system fails to make the proper suggestions or when specifying a first query. We decided our design should thus support it as an alternative and as a way to instantiate a search session.

B. Result sets of not only documents but most relevant entities to foster iterative query reformulation

To foster iterative query reformulation (challenge 3), we introduce the notion of *search streams* which describes an interactive structure supporting a query and related results: the query itself is formed of one or more entities and is composed by the user, while the results are shown as a vertical arrangement of entities related to the query and positioned above it. In the query area, items can be moved freely. Under a certain horizontal distance threshold, those entities are considered as a single query. The unity of a query is visualized through a network of thin lines linking the entities together. At first the query visually leads to a button that triggers the retrieval. The search engine then returns a set of entities related to the query. Those represents not only retrieved documents but also new entities, such as keywords or persons. They are vertically ordered by type and relevance. The flexibility of the search stream comes from its two level structure. It acts partly as a consolidated unit which can be moved around and considered as an almost traditional list of results, but each document or

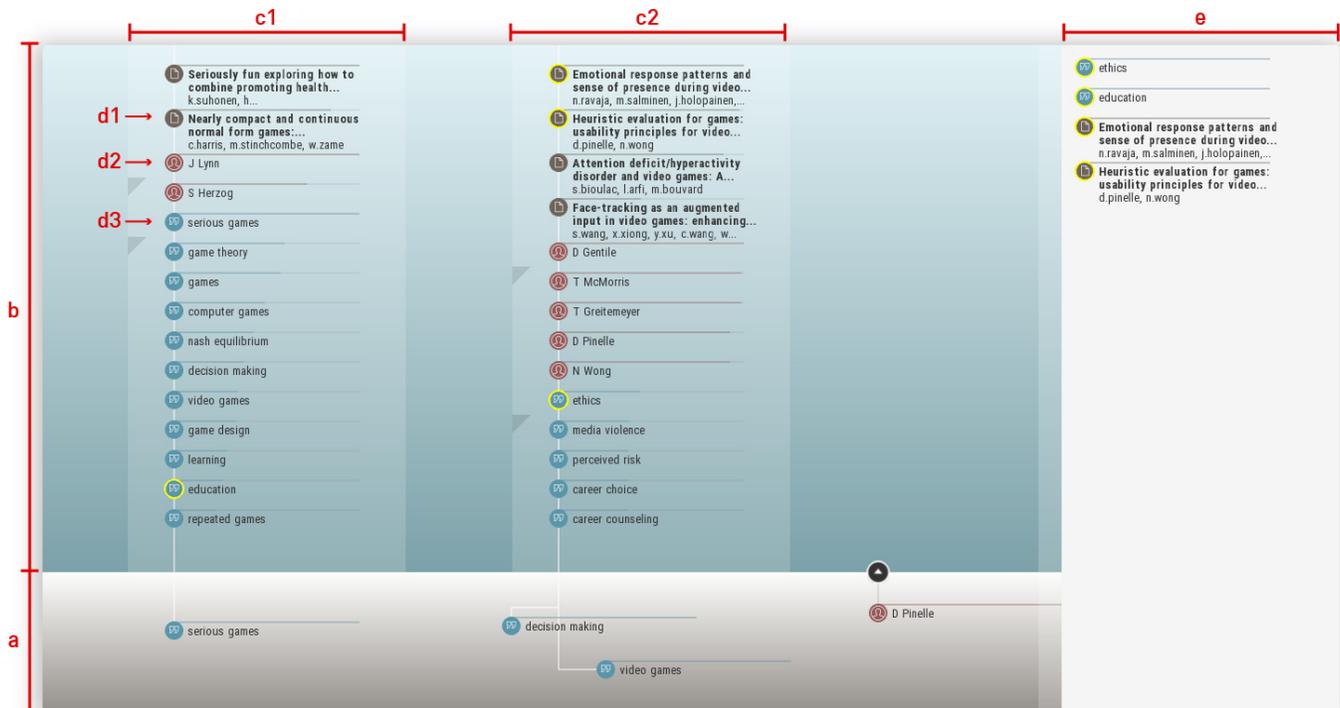


Figure 1. The Exploration Wall interface is composed of the (a) query area, (b) the results area, (c1, c2) search streams, three types of entities: (d1) documents (brown icon), (d2) authors (red icon), and (d3) keywords (blue icon), and (e) the reading-list drawer. The user can move any information item to compose queries by spatially grouping them on the query area. The whole workspace is scrollable and unlimited. Multi-touch gestures allows user to easily add or remove space between streams, or combine streams.

entity can become a new query, or part of an existing query, in the same stream or a parallel one.

C. Use of spatial configuration of multiple streams to identify search directions and learn about the information space

To facilitate steering decisions (challenge 2) and help the user formulate queries (challenge 1), our design supports search on simultaneous parallel streams. Persistency of search and context improves exploration by fostering trials without fear of losing current work, and supporting information comparison and entity association leading to quick instantiation of new queries or quick query re-formulation. It also allows the user to keep track of former queries and results while supporting unconstrained branching and revisits in the actual search process.

EXPLORATION WALL

Here we describe the interactions and implementation of the system based on the above mentioned design principles.

The User Interface

The interface of Exploration Wall is entirely dedicated to its main workspace (Figure 1), which is divided in two areas: the query area at the bottom (Figure 1-a) and the result area on top (Figure 1-b). The workspace supports information in the form of parallel search streams (Figure 1-c1 and c2) organized by taking advantage of the multi-touch ability: it can be scrolled on the horizontal axis with a simple swipe gesture on the background, horizontal space can be added or removed

at will from a specific location using a conventional pinch gesture, the same pinch gesture can also be used to dilate or contract space (e.g., to quickly improve legibility of an area cramped with information).

In current instantiation, entities are of three types (Figure 1): Documents (Figure 1-d1), Authors (Figure 1-d2) and Keywords (Figure 1-d3). Each entity is represented by a pictogram, a label and a relevance gauge. One can move an entity by dragging its pictogram. Additional interactions include: tap on the title of a document to reveal additional information like source and content, tap on the icon to store the entity. Stored entities appear highlighted and can be found in the storage drawer described below.

The storage drawer (Figure 1-e) offers an unobtrusive solution that acts as a reading list as well as an always accessible storage area for information transit. One opens and closes it by performing a swipe gesture from the right edge of the display.

The Search Engine

The search engine was designed to support multi-touch interaction design of Exploration Wall and is based on two design rationale. First, the *entity ranking* where entities that are returned for the user to manipulate and use to formulate queries should be as central to the topic as possible. For example, if the user searches for "information retrieval", she is not expecting back only entities that occur in the top ranked

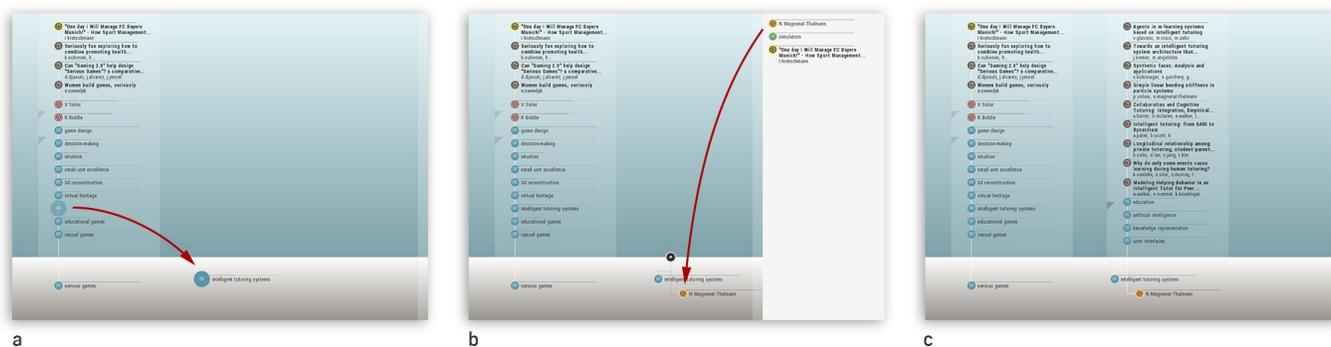


Figure 2. An exemplary search scenario illustrating the functionality of the Exploration Wall system. (a) After initiating a query the user receives a set of results, the user notices a keyword-entity that tackles her interest and drags it to the query area. (b) The user investigates the selected keyword in relation to an author entity that has been formerly saved in to the reading list by dragging the author entity to the query area as well, close enough so they become visually associated. (c) The user taps on the trigger to retrieve a new set of documents and entities (authors and keywords), that can be further manipulated and used to combine with an existing search stream or to create a new search stream.

documents, but that are central for the field of information retrieval. Second, the *document ranking* where the documents that are returned for the user as results after making some query, say "information retrieval" and "relevance feedback" should be not the most central entities, but the most relevant documents matching the query.

Entity Ranking

We represent the data as an undirected graph, where each document, keyword, and author are represented as vertices and the edges represent their occurrence in the document data.

The centrality ranking is based on the user's relevance feedback on vertices determined by dragging them into the query area. Each cluster in a query area represents a separate query that consists of a set of vertices. We use the personalized PageRank method [15] to compute the ranking of the vertices. The set of nodes that the user has chosen to be part of an individual query form the personalization vector that is set to be the prior for the PageRank computation [15]. We compute the steady distribution by using the power iteration method with 50 iterations. The top $k=10$ nodes from each entity category (keyword, author) are selected for presentation for the user.

Document Ranking

The document ranking is based on language modelling approach of information retrieval [41], where a unigram language model is built for each document and the maximum likelihood of the document generating the query is used to compute the ranking. We use Jelinek-Mercer smoothing to avoid zero probabilities in the estimation.

Intuitively, separating the entity ranking and document ranking approaches makes it possible to compute a limited set of entities that are likely to be the most important in the graph given the user interactions and allows users to target their feedback on a subset of the most central nodes given the interaction history of the user in any subsequent iteration. At the same time, the document ranking enables accurate and well-established methodology for ensuring relevance of the documents.

EVALUATION

The main purpose of the evaluation was to observe the effects and implications of the design of Exploration Wall on search performance and search behavior. Therefore, Exploration Wall was compared to a conventional search interface which was used as a baseline. The experiment concerned the following factors: effectiveness, expert rating, search behavior, usability and user engagement. The evaluation was composed of two tasks, a short one (5 minutes) and a long one (20 minutes).

Dataset

We used a document set including over 50 million scientific documents from the following data sources: the Web of Science prepared by Thomson Reuters, Inc., the Digital Library of the Association of Computing Machinery (ACM), the Digital Library of Institute of Electrical and Electronics Engineers (IEEE), and the Digital Library of Springer. The information about each document consists of: title, abstract, author names, and publication venue. Both the baseline and Exploration Wall used the same document set.

We decided to limit the data to scientific literature for two reasons. First, the data should allow retrieval tasks that result in exploration, and scientific search tasks are suitable for scenarios where users' goals are uncertain and require exploratory search behavior. Second, experts were available for providing high quality relevance assessments for task outcomes.

Baseline

The baseline, shown on Figure 3, was implemented following the interface principles of traditional search tools: typed query and resulting list of returned documents presented by title, with authors and keywords. The system uses the same dataset used by Exploration Wall to permit comparability. Also, the ranking is based on the same document retrieval model as in Exploration Wall, but to mimic traditional search engines it ranks only documents, while authors and keywords are only shown as additional information associated to each document. Last, our system did not allow dynamic updates of the search result when typing the query. All these factors

aimed to create a baseline allowing us to focus the evaluation solely on the user interface design of Exploration Wall.

Tasks

The evaluation was composed of two tasks, a short one and a long one. We chose 6 possible different topics for the two tasks: *crowdsourcing*, *smartphones energy efficiency*, *diagrams*, *semantic web*, *lie detection* and *digital audio effects*. In order to ensure that participants were not experts in the topics and could perform a real exploratory search, they pre-rated their familiarity with the topics on a 1 (less familiar) to 5 (most familiar) scale. The four less familiar topics were used in the tasks. Both tasks were performed with different topics, so the participants did not know the results from the previous task.

Short Task

For this task, we asked the users: "Search and list 5 relevant authors, documents and keywords that you consider relevant in topic Y." The time limit for this task was 5 minutes.

Long Task

For this task, we asked the users: "Imagine that you are writing a scientific essay on the topic X. Search and collect as many relevant scientific documents as possible that you find useful for this essay. During the task, please, list what you think are the top five key technologies, persons, documents and research areas and write five bullet lines, which would work as the core content of the essay." The time limit for this task was 20 minutes.

Participants and Procedure

We recruited 10 researchers from the computer science departments of two universities with a range of research experience. The 20% of them were females, which matched the gender ratio of both departments, and the mean age was $M=30.5$, $SD=5.52$. For the experiment participants used an iPad Air Wi-Fi tablet, as shown on Figure 4.

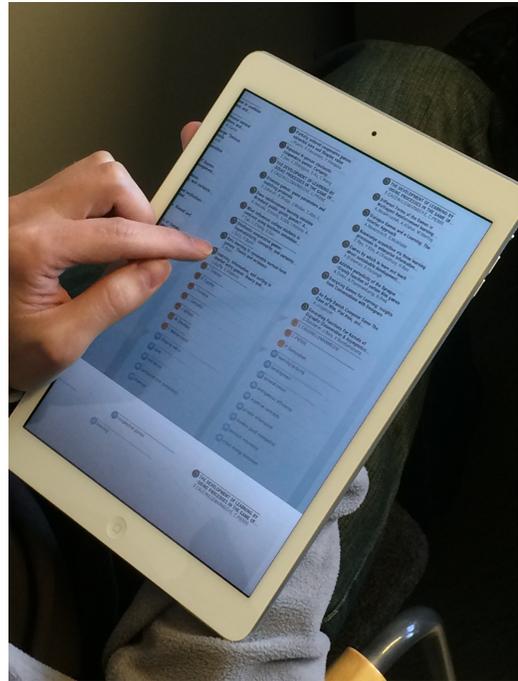


Figure 4. Exploration Wall was evaluated using using the iPad Air Wi-Fi tablet.

In this study, we followed a within-subjects experiment design, counter-balanced by changing the order of the two tested interfaces, as well as the order of the two tasks. Before starting the main tasks, users received detailed instructions on how to use the interface and performed a 5 minutes training task on each interface. For text entry, we relied on the native virtual keyboard of the tablet. At the end of the sessions participants were asked to answer the UES and SUS questionnaires for each interface via on-line forms (Google Forms). We used the API and service of logentries.com to log all actions and data.

MEASURES

The experiment considered the following factors: effectiveness, expert rating, search behavior, usability and user engagement which were measured as follows.

Effectiveness

The effectiveness refers to the quality of the information retrieved and displayed by a system. Since our baseline system returns lists of documents while Exploration Wall returns lists of mixed-type entities, we chose to solely measure the quality of the displayed documents. We created ground truth by pooling the retrieved documents from the system logs. Domain experts were then asked to assess the relevance of the retrieved documents on a binary scale (relevant or irrelevant). Effectiveness was measured by precision, recall and F-measure at two levels [21]. First, we measured the average retrieval effectiveness at a query level as an average quality of the documents returned in response to a user interaction. Second, we measured the retrieval effectiveness at task level as an cumulative quality of documents retrieved within the whole search session.

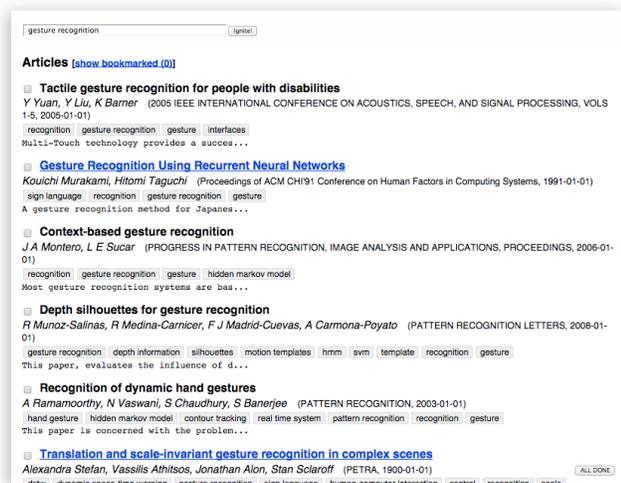


Figure 3. A screenshot of the baseline system that uses the same underlying document set and ranking model, and allows typed-query interaction.

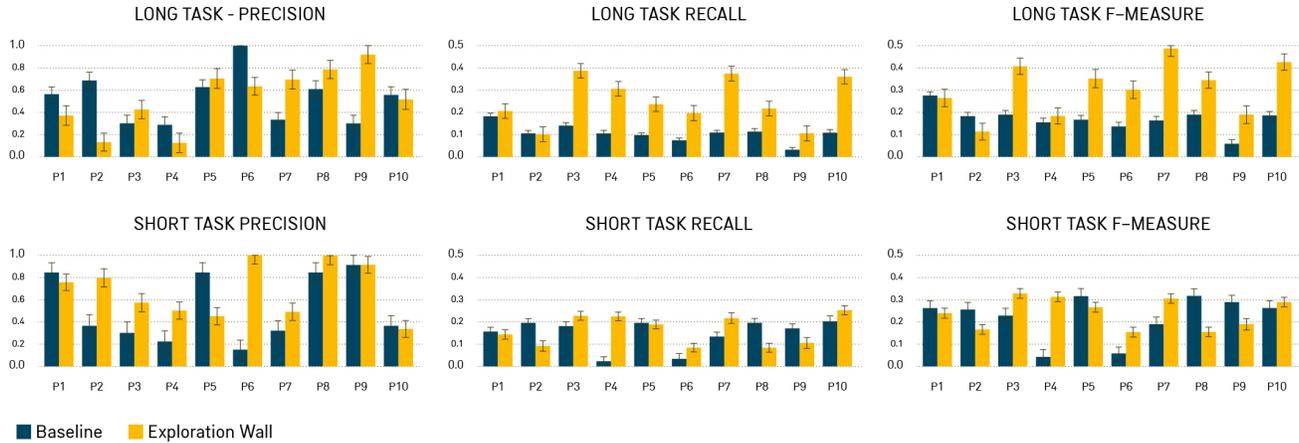


Figure 5. Effectiveness results for the short and long tasks split by participants. Results are reported as the mean of every query-response of each participant during the task.

Expert Rating

Experts were asked to rate the task outcome. For the short task, the outcome was a list of documents, and two types of entities: authors and keywords. The relevance of each item was evaluated on a 5-point scale (1 less relevant - 5 most relevant). The outcome of the long task was an essay, a set of documents, and a set of entities: keywords representing technologies and research areas, and persons. The sets of documents and entities were evaluated in the same way as in the short task, while the essay was evaluated on a different 5-point scale (5=Excellent, 4=Good, 3=Satisfactory, 2=Deficient, 1=Failing).

Search Trail Analysis

In order to understand and compare users' search behavior, we logged user actions and extracted corresponding search trails using a method resembling White's [36]. In a similar manner, we then looked for descriptive statistics of the search trails by selecting six parameters relevant to both interfaces.

- Number of queries: the total number of queries that were submitted during each task on both interface.
- Number of text entries per query
- Number of revisits: The number of revisits to a query or stream consulted earlier in the current trail.
- Number of branches: The number of times a subject revisited a query or stream on the current trail and then proceeded with formulation of a new query.
- Number of queries/min: the number of queries per minute that were submitted during each task on both interface.
- Number of parallel queries: Number of parallel streams produced with Exploration Wall or number of tabs opened with the baseline.

Usability and Engagement

As usability assessment questionnaires we used the standard System Usability Scale (SUS) [6] and the User Engagement

Scale (UES) for exploratory search [26]. SUS consists of a ten item questionnaire and is a widely used and validated for measuring perceptions of usability. Since the degree of user engagement is a strong indicator of exploratory search performance [38], we chose to use UES for exploratory search. The User Engagement Scale (UES) questionnaire include 27 questions considering six different dimensions: Aesthetics (AE), Focused Attention (FA), Felt Involvement (FI), Perceived Usability (PUs), Novelty (NO) and Endurability (EN) aspects of the experience.

RESULTS

In this section, we present results from the user experiments divided according to the different factors: effectiveness, expert rating, search trail analysis, and usability and engagement.

Effectiveness

The effectiveness results are given in Table 1. The results show that Exploration Wall shows substantial improvement in the long task. The improvement was found to hold for task-level measurement, but also for averaged interaction-level measurement for which the recall and the F-measure were found to be significantly higher compared to the baseline. On average at the query level, the F-measure for the Exploration Wall was improved ($M=0.136$, $SD=0.122$). This im-

	Long Task			Short Task		
	BL	EW	p	BL	EW	p
P (Task)	0.40	0.42	0.85	0.52	0.58	0.67
R (Task)	0.13	0.38	<0.01	0.18	0.21	0.59
F (Task)	0.17	0.34	<0.01	0.25	0.26	0.90
P (Query)	0.53	0.53	0.96	0.52	0.69	0.16
R (Query)	0.11	0.25	<0.01	0.15	0.16	0.69
F (Query)	0.17	0.31	<0.01	0.22	0.24	0.41

Table 1. Effectiveness results for the short and long tasks. Results are reported cumulatively for the whole duration of the task and as a mean of every query-response during the task. P =Precision, R =Recall, $F=F_1$ measure, EW=Exploration Wall, BL=Baseline.

	Long Task						
Search Trail Features	BL			EW			BL vs EW
	M	SD	Median	M	SD	Median	Wilcoxon Test
No. of queries	4.30	3.09	4.50	12.10	6.97	13.50	Z = -2.76, p < 0.01
No. of text entries/query	1.00	0.00	1.00	0.36	0.35	0.27	Z = 2.67, p < 0.01
No. of branches	0.10	0.31	0.00	5.70	4.55	6.00	Z = -2.68, p < 0.01
No. of revisits	0.70	1.64	0.00	7.00	6.09	6.00	Z = -2.67, p < 0.01
No. of queries/min	0.26	0.17	0.26	0.63	0.36	0.70	Z = -2.70, p < 0.01
No. parallel queries	1.70	1.06	1.00	8.50	5.89	7.00	Z = -2.76, p < 0.01
	Short Task						
Search Trail Feature	BL			EW			BL vs EW
	M	SD	Median	M	SD	Median	Wilcoxon Test
No. of queries	2.50	1.58	2.00	3.50	2.12	4.00	Z = -1.46, p > 0.05
No. of text entries/query	1.00	0.00	1.00	0.55	0.35	0.47	Z = 2.55, p < 0.05
No. of branches	0.00	0.00	0.00	0.8	1.03	0.5	Z = -2.21, p < 0.05
No. of revisits	0.20	0.42	0.00	1.1	1.10	1.0	Z = -1.81, p > 0.05
No. of queries/min	0.59	0.33	0.45	0.86	0.36	0.93	Z = -2.24, p > 0.05
No. of parallel queries	1.30	0.67	1.00	2.70	2.00	2.00	Z = -2.40, p < 0.05

Table 2. Results of the search trail analysis for the short and long tasks. Means, Standard Deviation, Median (used in the Wilcoxon Matched-Pairs test) as well as Significant differences of search trail feature considering both interfaces. The values in bold show the significant differences. BL=baseline, EW=Exploration Wall.

provement was statistically significant, $t(9)=3.519, p < 0.01$. This is a direct consequence of the improvement in the recall ($M=0.142, SD=0.094, t(9)=4.790, p < 0.001$). The difference in precision was not significant ($M=0.005, SD=0.366$) which indicates that while Exploration Wall improves recall it retains precision. In terms of effectiveness, no statistically significant differences between the systems were found in the short task.

Figure 5 shows the query-level effectiveness for the long tasks and the short task split by participants. Exploration Wall constantly outperforms the baseline system in terms of recall and F-measure in the long task. The effect is steady across participants. No significant differences between the systems were found in the short task.

Expert Rating

Unlike the effectiveness, the expert rating showed no significant differences between the Exploration Wall and the Baseline. Regarding the relevance of selected items, the mean values for the long task were $M=3.54, SD=0.67$ for Exploration Wall and $M=3.45, SD=0.82$ for Baseline, while for short task they were $M=3.60, SD=1.23$ for Exploration Wall and $M=3.83, SD=0.99$ for Baseline. Regarding the the relevance of the essays produced in the end of the long task the mean values were $M=3.90, SD=0.75$ for Exploration Wall and $M=4.05, SD=0.69$ for Baseline.

Search Trail Analysis

Table 2 shows the results of the search trail analysis. The Shapiro-Wilk test indicated that the search trail data did not follow a normal distribution, and the Wilcoxon Matched-Pairs test was used for significance testing. The users in the Exploration Wall condition were found to use all of the measured interaction features significantly more than the users in the baseline condition in the long task. Differences were also found in the short task. The users in the Exploration Wall

condition typed less, branched more, and used more parallel queries.

Usability and Engagement

The results for the mean of answers of the SUS questionnaire, i.e., for usability, were $M=78.85, SD=12.43$ for Exploration Wall and $M=62.25, SD=15.65$ for the baseline. A paired t-test showed a significant difference ($t(9)=2.36, p < 0.05$) between the two systems, revealing higher usability for Exploration Wall. The results of the UES questionnaires are also favorable for Exploration Wall. Wilcoxon Matched-Pairs test shows that in 70% of the questions there is a significant difference between the interfaces, all in favour of Exploration Wall.

DISCUSSION AND CONCLUSIONS

Challenges in supporting exploratory search include providing resources for formulating queries in unknown areas [11], learning about possible directions in the information space [29], and going through long list of results with low information gain [38, 5]. In particular on keyboard-less touch devices the challenges are aggravated by typing efforts. We introduced Exploration Wall a novel user interface that addresses these challenges with a principled design. The founding principle is to transform results into entities that can be flexibly manipulated and used for creating queries and search streams. The wall is a canvas where parallel and previous search streams are juxtaposed and provide a spatial exploration of the information space and possible exploration directions. The manipulation includes inspecting relationship between entities and facilitating the creation of new search streams.

The study shows how Exploration Wall is an effective tool for exploratory search on touch surfaces. Participants using Exploration Wall were able to exploit parallel search streams to iteratively refine their queries and deeply explore the search

tree. The difference in recall proves that more relevant documents were retrieved when using Exploration Wall.

Exploration Wall also led to a more active search behavior, with more queries per minute and more branches. In addition, if we consider the fact that participants used more parallel queries with Exploration Wall (parallel streams) than with the baseline (parallel tabs), we can conclude that the participants took advantage of parallel streams with consequent avoidance of text input.

Results from the UES questionnaire also show a better user engagement, a factor that is likely to have contributed to the more active search behavior. In addition, the SUS scale shows that Exploration Wall presents a better usability than conventional search interfaces on tablets.

The study confirms how our design approach facilitates query formulation, by directing exploration in unknown areas, and providing alternatives to text inputs. While little or no differences were appreciated in short tasks, Exploration Wall proved to be an effective tool for long tasks by showing improved recall while preserving precision, as well as improved user engagement and satisfaction.

In addition to the positive results, this work is adaptable to many applications and setups that would enable new possibilities to be found through deeper study of user behavior (e.g. search strategies and nature of composed queries). It has important implications for future development of exploratory search systems in particular considering multimodal interaction and user interface for entity oriented search [27]. The principles are applicable to other datasets such as for example news search [4] as well as other devices and sizes (e.g. large multi-touch screen for collaborative work, mobile devices for mobility and privacy, combinations of devices, desktop).

Considering this, as well as the growing popularity of touch devices, our work offers a powerful and flexible template to be considered when designing user interfaces supporting exploratory search.

ACKNOWLEDGMENTS

This research was partially funded by TEKES (through the Re:KnoW project), the European Commission through the FP7 Project MindSee 611570, the Academy of Finland (Multivire, 255725 and 278090), TIPEX Project (TIN2010-19859-C03-01) and "Postdoc Mobility Scholarship Programme" of Universidad Carlos III de Madrid. 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 We would like to thank Kumari Athukorala for her valuable suggestions.

REFERENCES

1. Ahlberg, C., and Shneiderman, B. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In *Proc. CHI*, ACM (New York, NY, USA, 1994), 313–317.
2. Ahn, J., and Brusilovsky, P. Adaptive visualization for exploratory information retrieval. *Information Processing & Management* 49, 5 (2013), 1139–1164.
3. Ahn, J.-w., and Brusilovsky, P. Adaptive visualization of search results: Bringing user models to visual analytics. *Information Visualization* 8, 3 (2009), 167–179.
4. Andolina, S., Klouche, K., Peltonen, J., Hoque, M., Ruotsalo, T., Cabral, D., Klami, A., Głowacka, D., Floréen, P., and Jacucci, G. Intentstreams: smart parallel search streams for branching exploratory search. In *Proc. IUI*, ACM (2015).
5. Athukorala, K., Oulasvirta, A., Głowacka, D., Vreeken, J., and Jacucci, G. Narrow or broad?: Estimating subjective specificity in exploratory search. In *Proc. CIKM*, ACM (New York, NY, USA, 2014), 819–828.
6. Brooke, J. Sus-a quick and dirty usability scale. *Usability evaluation in industry* 189 (1996), 194.
7. Capra, R., Marchionini, G., Oh, J. S., Stutzman, F., and Zhang, Y. Effects of structure and interaction style on distinct search tasks. In *Proc. JCDL*, ACM (2007), 442–451.
8. Chau, D. H., Kittur, A., Hong, J. I., and Faloutsos, C. Apolo: Making sense of large network data by combining rich user interaction and machine learning. In *Proc. CHI*, ACM, ACM (New York, NY, USA, 2011), 167–176.
9. Coutrix, C., Kuikkaniemi, K., Kurvinen, E., Jacucci, G., Avdouevski, I., and Mäkelä, R. Fizzyvis: designing for playful information browsing on a multitouch public display. In *Proc. DPPI*, ACM (2011), 27.
10. Esenther, A., and Ryall, K. Fluid dtmouse: Better mouse support for touch-based interactions. In *Proc. AVI*, ACM (New York, NY, USA, 2006), 112–115.
11. Głowacka, D., Ruotsalo, T., Konyushkova, K., Athukorala, K., Jacucci, G., and Kaski, S. Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Proc. IUI*, ACM (New York, NY, USA, 2013), 117–128.
12. Havre, S., Hetzler, E., Perrine, K., Jurrus, E., and Miller, N. Interactive visualization of multiple query results. In *Proc. InfoVis*, IEEE Computer Society (Washington, DC, USA, 2001), 105.
13. Hinrichs, U., Schmidt, H., and Carpendale, S. Emdialog: Bringing information visualization into the museum. *Visualization and Computer Graphics, IEEE Transactions on* 14, 6 (2008), 1181–1188.
14. Jacucci, G., Morrison, A., Richard, G. T., Kleimola, J., Peltonen, P., Parisi, L., and Laitinen, T. Worlds of information: designing for engagement at a public multi-touch display. In *Proc. CHI*, ACM (2010), 2267–2276.
15. Jeh, G., and Widom, J. Scaling personalized web search. In *Proc. WWW*, ACM (New York, NY, USA, 2003), 271–279.

16. Jones, M., Buchanan, G., Harper, R., and Xech, P.-L. Questions not answers: A novel mobile search technique. In *Proc. CHI*, ACM (New York, NY, USA, 2007), 155–158.
17. Käki, M. Findex: Search result categories help users when document ranking fails. In *Proc. CHI*, ACM (New York, NY, USA, 2005), 131–140.
18. Karlson, A. K., Robertson, G. G., Robbins, D. C., Czerwinski, M. P., and Smith, G. R. Fathumb: A facet-based interface for mobile search. In *Proc. CHI*, ACM (New York, NY, USA, 2006), 711–720.
19. Kules, W., Wilson, M. L., Schraefel, M. C., and Shneiderman, B. From keyword search to exploration: How result visualization aids discovery on the web. Tech. rep., University of Southampton, 2008.
20. Li, Y. Gesture search: a tool for fast mobile data access. In *Proc. UIST*, ACM (2010), 87–96.
21. Manning, C. D., Raghavan, P., and Schütze, H. *Introduction to information retrieval*, vol. 1. Cambridge university press Cambridge, 2008.
22. Marchionini, G. Exploratory search: from finding to understanding. *Commun. ACM* 49, 4 (Apr. 2006), 41–46.
23. Matejka, J., Grossman, T., and Fitzmaurice, G. Citeology: Visualizing paper genealogy. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, ACM (New York, NY, USA, 2012), 181–190.
24. Morris, M. R., Fisher, D., and Wigdor, D. Search on surfaces: Exploring the potential of interactive tabletops for collaborative search tasks. *Information processing & management* 46, 6 (2010), 703–717.
25. Morris, M. R., Lombardo, J., and Wigdor, D. Wesearch: supporting collaborative search and sensemaking on a tabletop display. In *Proc. CSCW*, ACM (2010), 401–410.
26. O'Brien, H. L., and Toms, E. G. Examining the generalizability of the user engagement scale (ues) in exploratory search. *Information Processing & Management* 49, 5 (2013), 1092 – 1107.
27. Ruotsalo, T., and Hyvönen, E. Exploiting semantic annotations for domain-specific entity search. In *Proc. ECIR*, Springer (2015).
28. Ruotsalo, T., Jacucci, G., Myllymäki, P., and Kaski, S. Interactive intent modeling: Information discovery beyond search. *Communications of the ACM* 58, 1 (Jan. 2015), 86–92.
29. Ruotsalo, T., Peltonen, J., Eugster, M., Glowacka, D., Konyushkova, K., Athukorala, K., Kosunen, I., Reijonen, A., Myllymäki, P., Jacucci, G., et al. Directing exploratory search with interactive intent modeling. In *Proc. CIKM*, ACM (2013), 1759–1764.
30. Sadler, E. B. Project blacklight: a next generation library catalog at a first generation university. *Library Hi Tech* 27, 1 (2009), 57–67.
31. Stasko, J., Görg, C., and Liu, Z. Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization* 7, 2 (2008), 118–132.
32. Terveen, L., Hill, W., and Amento, B. Constructing, organizing, and visualizing collections of topically related web resources. *ACM Transactions on Computer-Human Interaction* 6, 1 (1999), 67–94.
33. Thudt, A., Hinrichs, U., and Carpendale, S. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proc. CHI*, ACM (2012), 1461–1470.
34. Varcholik, P. D., LaViola, J. J., and Hughes, C. E. Establishing a baseline for text entry for a multi-touch virtual keyboard. *Int. J. of Human-Computer Studies* 70, 10 (2012), 657 – 672.
35. Verbert, K., Parra, D., Brusilovsky, P., and Duval, E. Visualizing recommendations to support exploration, transparency and controllability. In *Proc. IUI*, ACM, ACM (New York, NY, USA, 2013), 351–362.
36. White, R. W., and Drucker, S. M. Investigating behavioral variability in web search. In *Proc. WWW*, ACM (2007), 21–30.
37. White, R. W., Drucker, S. M., Marchionini, G., Hearst, M., and schraefel, m. c. Exploratory search and hci: Designing and evaluating interfaces to support exploratory search interaction. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, ACM (New York, NY, USA, 2007), 2877–2880.
38. White, R. W., and Roth, R. A. Exploratory search: Beyond the query-response paradigm. *Synth. Lec. on Inf. Conc., Retr., and Serv. 1*, 1 (2009), 1–98.
39. Wigdor, D., and Wixon, D. *Brave NUI world: designing natural user interfaces for touch and gesture*. Elsevier, 2011.
40. Wildemuth, B. M., and Freund, L. Assigning search tasks designed to elicit exploratory search behaviors. In *Proc. HCIR*, ACM (New York, NY, USA, 2012), 4:1–4:10.
41. Zhai, C., and Lafferty, J. A study of smoothing methods for language models applied to information retrieval. *ACM Trans. Inf. Syst.* 22, 2 (2004), 179–214.
42. Zhai, S., Kristensson, P.-O., and Smith, B. A. In search of effective text input interfaces for off the desktop computing. *Interacting with computers* 17, 3 (2005), 229–250.