Tomi Heimonen

Design and Evaluation of User Interfaces for Mobile Web Search

ACADEMIC DISSERTATION
To be presented with the permission of the Board of the School of Information Sciences of the University of Tampere, for public discussion in the Pinni auditorium B1097 on November 20th, 2012, at noon.

School of Information Sciences
University of Tampere

Dissertations in Interactive Technology, Number 14
Tampere 2012

ACADEMIC DISSERTATION IN INTERACTIVE TECHNOLOGY

Supervisor:	Professor Kari-Jouko Räihä, Ph.D. School of Information Sciences University of Tampere Finland
Opponent:	Professor Matt Jones, Ph.D. Department of Computer Science Swansea University Wales, United Kingdom
Reviewers:	Dr. George Buchanan, Reader in Human-Computer Interaction Centre for Human Computer Interaction Design City University London United Kingdom Dr. Mark D. Dunlop, Senior Lecturer in Computer Science Department of Computer and Information Sciences University of Strathclyde Scotland, United Kingdom

Dissertations in Interactive Technology, Number 14
-
School of Information Sciences
FI-33014 University of Tampere
FINLAND

ISBN 978-951-44-8945-7
ISSN 1795-9489

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2012

Abstract

Mobile Web search is a rapidly growing information seeking activity employed across different locations, situations, and activities. Current mobile search interfaces are based on the ranked result list, dominant in desktop interfaces. Research suggests that new paradigms are needed for better support of mobile searchers. For this dissertation, two such novel search interface techniques were designed, implemented, and evaluated.

The first method, a clustering search interface that presents a category-based overview of the results, was studied both in a task-based experiment in a laboratory setting and in a longitudinal field study wherein it was used to address real information needs. The results indicate that clustering can support exploratory search needs – when the searcher has trouble defining the information need, requires an overview of the search topic, or is interested in multiple results related to the same topic. The findings informed design guidelines for category-based search interfaces. How and when categorization is presented in the search interface needs to be carefully considered. Categorization methods should be improved, for better response to diverse information needs. Hybrid approaches employing contextually informed clustering, classification, and faceted browsing may offer the best match for user needs.

The second presentation method, a visualization of the occurrences of the user's query phrase in a result document, can be incorporated into the ranked result list as an additional, unobtrusive result descriptor. It allows the searcher to see how often the query phrase appears in the result document, enabling the use of various evaluation strategies to assess the relevance of the results. Several iterations of the visualization were studied with users to form an understanding of the potential of this approach. The results suggest that a novel visualization can be useful in ruling out non-relevant results and can assist when the other result descriptors do not provide for a conclusive relevance assessment. However, users' familiarity with well-established result descriptors means that users have to learn how to integrate the visualization into their search strategies and reconcile situations in which the visualization is in conflict with other metadata.

In addition, the contextual triggers and information behaviors of mobile Internet users were studied, for understanding of the role of Web search as a mobile information seeking activity. The results from this study show that mobile Web search and browsing are important information seeking activities. They are engaged in to resolve emerging information needs as they appear, whether at home, "on the go," or in social situations.

Acknowledgements

This work has been a labor of love and could not have been completed without the support of my colleagues and co-contributors. Several people have had a profound effect on this work by co-authoring research articles with me. Natalie Jhaveri, Mika Käki, and Harri Siirtola all contributed to various aspects of the research, from the design and implementation of prototypes to the evaluation, analysis, and dissemination of the results. In addition, I am deeply grateful to Anne Aula for helping me look on the other side of the fence and get a feel for what it is to work on search in an industrial setting.

I have also had the privilege of collaborating with like-minded researchers at the Tampere Unit for Computer–Human Interaction. Several of the ideas and findings that contributed to this dissertation were germinated in many a thoughtful, and often lively, discussion. In particular, Stina Boedeker has been an indefatigable source of inspiration, helping me plan and structure the work into manageable steps, and discussions with Toni Pakkanen have been extremely helpful in bouncing ideas back and forth.

Special thanks go to my supervisor, Kari-Jouko Räihä, for allowing me to find my own way and problems to tackle during the dissertation work, while always providing guidance when needed. In addition, financial support from the Finnish Doctoral Program in User-Centered Information Technology (UCIT) enabled me to focus on working primarily on the dissertation for four years.

My family and friends have been there for me throughout this journey. I would like to thank my parents for their unstinting support and for providing me with a physical and emotional refuge from the stress and pressures of the academic work. Similarly, my friends all over the world have given me outlets for venting about the problems and vagaries of a graduate student's life. Lauren, Jed, Matt, Val, Yuhri, Joanne, Aymee, Amy, Kara, Tommy, and Elio and Linda – you have demonstrated that modern information technologies can change our lives in wonderful ways.

Most importantly, I would like to thank Jessica for her patience and understanding in the final stages of the dissertation write-up. You pushed me when I was slacking and listened to my complaints tirelessly. I may have missed out on a professional football career, but here I am taking my first academic Lambeau Leap.

In Tampere on October 12, 2012

Tomi Heimonen

Contents

List of Publications

This dissertation consists of a summary and the following original publications, reproduced here by permission. The presentation of the publications is ordered by the research areas explored during the dissertation work.

I. Heimonen, T., & Jhaveri, N. (2005). Visualizing query occurrence in search result lists. In *Proceedings of the 9th International Conference on Information Visualisation, IV '05* (pp. 877–882). Washington, DC, USA: IEEE Computer Society. doi:10.1109/IV.2005.152

II. Heimonen, T., & Siirtola, H. (2009). Visualizing query occurrence in mobile Web search interfaces. In *Proceedings of the 13th International Conference on Information Visualisation, IV '09* (pp. 639–644). Washington, DC, USA: IEEE Computer Society. doi:10.1109/IV.2009.16

III. Heimonen, T., & Käki, M. (2007). Mobile Findex – Supporting mobile Web search with automatic result categories. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '07* (pp. 397–404). New York, NY, USA: ACM. doi:10.1145/1377999.1378045

IV. Heimonen, T. (2008). Mobile Findex: Facilitating information access in mobile Web search with automatic result clustering. *Advances in Human-Computer Interaction, 2008*, article ID 680640. doi:10.1155/2008/680640

V. Heimonen, T. (2012). How do users search the mobile Web with a clustering interface? A longitudinal study. *International Journal of Mobile Human–Computer Interaction*, 4(3), 44-66. doi:10.4018/jmhci.2012070103

VI. Heimonen, T. (2009). Information needs and practices of active mobile Internet users. In *Proceedings of the 6th International Conference on Mobile Technology, Applications, and Systems, Mobility '09* (Article 50). New York, NY, USA: ACM. doi:10.1145/1710035.1710085

Author's Research Contributions

This work would not have been possible without the assistance of my colleagues past and present. Three of the papers included in this thesis were co-authored.

The central ideas behind the query occurrence visualization presented in Paper I were developed in collaboration with Natalie Jhaveri as a part of the Search-In-a-Box project carried out in co-operation with the Complex Systems Computation group at Helsinki Institute of Information Technology. The present author was responsible for the implementation of the visualization algorithm and the adaptation of existing software utilized to conduct the user study, while the design and evaluation activities and authorship of the article was shared with Ms. Jhaveri.

In further work on the mobile version of the query occurrence visualization presented in Paper II, I was assisted by collaboration with Dr. Harri Siirtola. Dr. Siirtola was instrumental in helping to streamline the design of the visualization and plan the experimental procedure, as well as in providing comments on the manuscript.

The category-based user interface designs rely heavily on the dissertation work of Dr. Mika Käki on the Findex clustering algorithm and search user interface framework. Dr. Käki also assisted in the design and preparation of Paper III, and has been an invaluable source of consultation during the whole dissertation research process. Juuso Kanner developed the initial framework for the mobile application architecture utilized in the user study reported upon in papers III and IV, and the present author further improved it. The present author was responsible for developing the search interfaces utilized in two subsequent user studies (dealt with in papers II and V). All interfaces extend the functionality of the underlying Findex search framework.

The diary study reported on paper VI was the work of the present author in its entirety.

I was assisted greatly in the methodological aspects of the work by Dr. Anne Aula, Dr. Hilary Hutchinson, and Dr. Laura Granka during my internship at Google, Inc. Their assistance in developing a new experimental method for assessing user interfaces for search provided material input in implementation of the latter query occurrence visualization user study.

List of Figures

Cutrell, E., Fisher, D., Drucker, S. M., Ramos, P. A. G., & Hu, C. (2009). Visual snippets: Summarizing Web pages for search and revisitation. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI '09* (pp. 2023–2032). New York, NY, USA: ACM. doi:10.1145/1518701.1519008. Figure 3. © 2009 Association for Computing Machinery, Inc. Reprinted by permission.

Figure 12a Flamenco Fine Arts Search interface. http://orange.sims.berkeley.edu/

Figure 12b Relation Browser++ interface. Retrieved from http://ils.unc.edu/relationbrowser/index.php?page=history. Dr. Gary Marchionini.

Figure 13a Treemap search interface. Reprinted from Kules, B., & Shneiderman, B. (2005). Categorized graphical overviews for Web search results: An exploratory study using U.S. government agencies as a meaningful and stable structure. In *Proceedings of the Third Annual Workshop on HCI Research in MIS* (pp. 20–23). Figure 2. Used with permission.

Figure 13b ResultMap visualization. Reprinted from Clarkson, E., Desai, K., & Foley, J. (2009). ResultMaps: Visualization for search interfaces. *IEEE Transactions on Visualization and Computer Graphics, 15*(6), 1057–1064. Figure 2. © 2009 IEEE.

Figure 14 SearchMobil interface. Reprinted from Springer Berlin Heidelberg, *Mobile and Ubiquitous Information Access: Mobile HCI 2003 International Workshop, Udine, Italy, September 8, 2003, Revised and Invited Papers, Lecture Notes in Computer Science 2954*, 2004, pp. 158–171, SmartView and SearchMobil: Providing overview and detail in handheld browsing, Milic-Frayling, N., Sommerer, R., Rodden, K., & Blackwell, A., Figure 4a, © Springer-Verlag Berlin Heidelberg 2004, with kind permission from Springer Science and Business Media.

Figure 15 Query occurrence visualization in mobile Web search interface. Reprinted from Heimonen, T. & Siirtola, H. (2009). Visualizing query occurrence in mobile Web search interfaces. In *Proceedings of the 13th International Conference on Information Visualisation, IV '09* (pp. 639–644). Washington, DC. USA: IEEE. Figure 2. © 2009 IEEE.

Figure 16 CloudCredo tag cloud interface. Reprinted from Mizzaro, S., Sartori, L., & Strangolino, G. (2012). Tag clouds and retrieved results: The CloudCredo mobile clustering engine and its evaluation. In *Proceedings of the 3rd Italian Information Retrieval Workshop* (pp. 191–198). Figure 1b. Used with permission.

List of Tables

1 Introduction

1.1 Objective

Today, searching the World Wide Web is undoubtedly the most common way people find information online, having superseded manually maintained link repositories and portals. Web search engines are one of the most frequently used online computer applications and an essential part of most information systems. With these search engines, people are able to search for text-based information, images, news content, videos, and much more. Advances in information technology have made it possible to engage in Web search across a variety of devices, from desktop computers to mobile phones. The use of Web content on mobile devices has exploded in the recent years with the increasing availability of affordable broadband mobile Internet services. Similarly to desktop developments, mobile Internet access has become an indispensable means of information access for users around the world. It is used for communicating, gathering information, performing various transactions, and engaging in social networking interactions (Taylor et al., 2008).

Mobile Internet search is increasing in importance as a mobile information access method. According to the survey results of Kaikkonen (2011), the frequency of various mobile search activities grew significantly between 2007 and 2010. Other studies have highlighted the importance of mobile search as an "on-demand" information access tool that is used to satisfy information needs as they arise (Church & Oliver, 2011; Paper VI). For design of better mobile search services, it is also critical to understand how various contextual factors, such as the time, location, and activity, influence mobile information needs (e.g., Church & Smyth, 2009; Hinze, Chang, & Nichols, 2010; Sohn, Li, Griswold, & Hollan, 2008), and how these needs are met by means of search tools.

Mobile search services come in many forms, from Web-based keyword search to dedicated on-device applications. Although these services and products are designed for mobile devices and make use of useful features such as location sensing and voice interaction, especially Web search results are still in many cases are presented in the form of the traditional ranked result lists, comprising of information such as the page title, a brief summary extracted from the text, and the URL. Previous research on mobile Web search patterns (Church, Smyth, Bradley, & Cotter, 2008; Kamvar & Baluja, 2006) has shown that mobile search users are likely struggling to satisfy their information needs with these interfaces.

The research reported upon in this dissertation had two key objectives. The first was to design, implement and evaluate new interface solutions to support the mobile search process. This is done by introducing presentation methods that complement the ranked result list, both in terms of organizing the search results into informative overviews and by supplementing the typical search result descriptors with informative visualizations. The second objective was to study mobile information needs and the roles of search as an information seeking strategy in order to inform search interface design further. Finally, because the proposed interface solutions are grounded in previous work on desktop information access, the findings from this research will contribute to the ongoing discussion of the differences between desktop and mobile Web search, how well existing presentation and visualization techniques transfer from the desktop to the mobile context, and what adaptations are necessary for making them better address the mobile context of use.

1.2 Context of the Research

The research conducted for this dissertation is situated at the juncture of several disciplines that deal with how humans interact with information and information technology. The design and evaluation of search interfaces here draws heavily from prior research into both information retrieval and human–computer interaction. Similarly, the design and evaluation of visualization approaches for search results incorporates ideas, techniques and methodology from disciplines that address information visualization and information retrieval. Human–computer interaction analysis provides the overarching user-centered focus for all research efforts – attempting to understand how people access information when mobile; how the mobile context of use affects information search behavior, strategies and needs; and how user interface solutions could assist in fulfilling these information needs.

This dissertation summarizes previous research on the key topics related to the theme of mobile information access. First the theoretical frameworks of information access and information visualizations are described. These frameworks form the foundation for the design of the search result visualizations and interfaces introduced in the present work. The next chapter outlines the main themes in the design and evaluation of search interfaces. These provide the context for the treatment of the main theme of this thesis, mobile information access and search interfaces. These chapters consider 1) research related to the mobile context of use, mobile information needs, the role of mobile Internet access and search in information access, and mobile Web search interfaces; 2) the methods for organizing, presenting, and visualizing results in mobile search interfaces; and 3) the unique challenges presented by mobility for the evaluation of mobile interactions. The dissertation concludes with a discussion of the key findings and contributions generated by the research, and their implications with respect to how they expand our understanding of mobile information access behavior and technologies.

1.3 METHODOLOGY

The methodology covered in the research reported on this dissertation consisted of a variety of design, implementation, and evaluation methods. The basic premise was that of user-centered design, whereby the design solutions are based on identified user needs and vetted with real users via a variety of research methods, from lab-based experiment protocols to long-term field studies, in which users utilized the research prototypes in their everyday information access tasks. The research work was highly constructive in the sense that each individual study, apart from the mobile information needs diary study reported upon in Paper VI, included a functioning search prototype. This necessitated a considerable amount of iterative software development work for production of a system that would stand up to the rigors of interactive experimentation and daily use during longitudinal studies.

The most challenging aspect of the research was conducting and analyzing the user studies. The laboratory-based studies were all traditional controlled experiments that examined the effect of the user interface design changes on objective search performance and subjective response with primarily quantitative metrics, with an unaltered search interface as the baseline. Owing to the limitations of these research paradigms, the focus in the latter studies shifted towards study of search interactions *in situ* in naturalistic contexts of use. The study on mobile information access strategies of active mobile Internet users and the longitudinal evaluation of the mobile clustering search interface both took a more qualitative stance. The analysis was focused on identifying salient themes in the in-depth interview data and diary entries, and finding behavior patterns from usage logs containing real search queries and interface interactions.

1.4 Results

The results presented in this dissertation are clustered around several interlinked topics: mobile information needs and the strategies used to fulfill them, the evaluation of the search results, and how this can be supported both at the level of the full result set and when one is assessing individual results.

The findings of the study on mobile information needs and information access strategies contribute to our understanding of how active mobile Internet users approach information seeking. The results show that search is a tool typically employed to address time-sensitive information needs as they arise, whereas Web browsing and applications are more likely to be employed as tools for addressing focused information needs such as those for timetables or phone number lookup.

The evaluations of the mobile clustering search interface Mobile Findex – both in the laboratory and in the field – contribute to research on category-based search interfaces. The longitudinal field study in particular revealed the extent to which clustering can assist in realistic information access scenarios, and how category-based interfaces should be improved so as to take into better account the mobile context of use, with elements such as mobile information needs and users' search strategies.

Finally, the two studies focusing on the query occurrence visualization examined how search result evaluation could be facilitated by visualization of the locations of the query phrase within the text of a result document. As a space-saving visualization technique, it could be easily embedded in the search result list. Similarly to clustering, the query occurrence visualization was considered situationally useful, for example, for exclusion of non-relevant results from consideration, and to assist when the other result descriptors did not provide a conclusive assessment.

In summary, the studies not only sought to validate the novel interface solutions but also attempted to understand how people perceive the search process and utilize search tools to address their information needs. Web search is an important means of information seeking on mobile devices and is affected by context: location, activities, and social situations. Despite its limitations with respect to providing overviews and enabling efficient subtopic access, the ranked result list is a good fit for many information needs that the user can express well in query form, or ones that address a familiar topic. The benefits of advanced presentation and visualization methods, such as those explored in this dissertation, come into play when the traditional descriptors fail to provide a good assessment of relevance or when one needs to understand and explore an unfamiliar topic or has problems in expressing the information need. This necessitates certain design considerations for the employment of these advanced features. Learning to trust in and use the features takes time, and their role is to complement rather than replace the traditional, familiar result information.

1.5 STRUCTURE

This dissertation is structured as follows. It first introduces the research objectives, then defines the main concepts, frameworks, and theoretical constructs. Next, it provides a review of existing research into the key topics and situates the present work within this framework. After this groundwork is laid, the key research articles that comprise the bulk of this dissertation and their results are introduced. The dissertation concludes with discussion of the findings and relevance of the present research, and it charts avenues for further research on mobile information access interfaces.

2 Information Seeking and Visualization

Information behavior includes activities that people engage in when identifying their needs for information, searching for it, and using the information for some purpose (T. D. Wilson, 1999). According to Wilson's (1999) model, information seeking behavior encompasses the methods of satisfying information needs by utilizing information resources. Finally, information search behavior has to do with the interactions between users and computer-based information systems.

In the literature, terms describing these interrelated information behaviors are often used interchangeably. M. L. Wilson, Kules, schraefel, and Shneiderman (2010) provide a useful distinction between the main activities of *information retrieval* and *information seeking*. In their terminology, information retrieval refers to the "paradigm where users enter a keyword into a system, which responds by returning the results that are most relevant to the keywords used," whereas information seeking is a broader term, encompassing behaviors such as information retrieval, browsing, and navigation. In their parlance, *search* provides the overall context for the information seeking behaviors, from identifying the need for searching to fulfilling the information need. Marchionini (1989) describes information seeking as a special case of problem-solving that includes recognizing and interpreting the information problem to be solved and the associated planned search, and is influenced by experience, knowledge, and the information need.

Various definitions exist for *information needs* in the context of the information seeking literature (Campbell, 1995; Dearman, Kellar, & Truong, 2008; Shneiderman, Byrd, & Croft, 1997; T. D. Wilson, 1981). Dearman et al. (2008) make the observation that information needs exist independently of the method used to satisfy the need. By their definition, information need is "any information that is required for a task, or to satisfy the curiosity of the mind, regardless of whether the need is satisfied or not." In the context of search systems, Shneiderman et al. (1997) simply define information need as the underlying cause for use of an information retrieval system. Campbell (1995) provides a broader definition, which considers information needs to be a combination of the expected format and location of the target information and the access methods that could be used to acquire the information. Finally, T. D. Wilson (1981) frames information need as arising in a specific context (that of the person's role and the environment), with barriers that hinder engaging in information seeking or completing the search for information, and the information seeking behavior itself. This consideration of the information need, context, constraints and access methods provides a useful framework for discussion in subsequent chapters.

Information seeking behaviors always occur in a *context*, as mentioned by T. D. Wilson (1981). Järvelin and Ingwersen (2004) argue that in traditional information seeking research the role of context is poorly understood. They suggest that the pragmatic goal of improving users' information access should remain a major goal in information seeking research and that it should be studied in the context of work task situations. Therefore, they propose an extended framework of information seeking and retrieval design and evaluation that identifies several levels of context: the socio-organizational and cultural context, the work task context, the information seeking context, and the information retrieval context. The work context, or personal goals in the case of leisurely motivated information seeking, motivates the information needs that, in turn, prompt information seeking tasks (M. L. Wilson et al., 2010). The role and influence of context is discussed in more detail in Chapter 4, where the effects of the mobile context on information seeking are examined.

The following discussion presents a summary of theoretical frameworks related to information seeking behavior. First, various models of the information seeking process are discussed. This is followed by an overview of information seeking behavior with the Web. Finally, the key models of information visualization and its relationship to information seeking are discussed.

2.1 Information Seeking and Retrieval

Understanding the human information seeking processes is the foundation for the design of effective and usable search systems (Hearst, 2009). Below, we consider some of the most common information seeking models, starting with a description of the broader process of *sensemaking*, which includes both information retrieval and search, and the result analysis during which these are associated with the task at hand.

After that, several higher-level models describing the information seeking process are outlined. This material is followed by a summary of the search strategies that people employ during the information seeking process to decide which actions to purse. Lastly, the emerging field of *exploratory search* as a distinct form of information seeking is introduced.

Sensemaking

Sensemaking is the process whereby people attempt to organize information so as to understand the world they live in. The core activities of sensemaking are processes of collecting, organizing, and representing information to solve a problem that needs to be understood (Russell, Pirolli, Furnas, Card, & Stefik, 2009). Sensemaking research combines concepts from several disciplines, such as philosophy, cognitive science, sociology, and social psychology.

The seminal work by Russell, Stefik, Pirolli, and Card (1993) analyzed different sensemaking tasks and developed a model to describe the cost structure of sensemaking. The fundamental pattern in sensemaking is described as a learning loop that consists of three main processes: searching for representations, instantiating them, and shifting between them. In the end, the information assigned in line with the instantiated schema is consumed to complete the overall sensemaking task. Two main categories of sensemaking tasks relevant to information seeking were identified: "one-off" tasks and recurring tasks. For one-off tasks, the aim for the sensemaker is to optimize the process to maximize the gain with respect to a given cost. In recurring tasks, the aim is to optimize the gain over repeated task cycles. Russell and colleagues note that in many cases most of the cost, in terms of time expended, is in data extraction: finding and selecting the relevant information and transforming it into the appropriate format. In addition, the central role of representation design is identified. In the context of Web search interfaces, it appears that it is critical that the results be provided by means of the appropriate representation and that the shift from one representation to another, and the extraction of data, be supported. For example, such a shift could take place when one is switching between a ranked result list and a category overview, which represent the results at different levels of description.

Interfaces to support sensemaking with the Web have been proposed. SearchPad (Bharat, 2000) is a tool for maintaining search context across multiple search engines and multiple sessions. Users are able to mark relevant search results, which are then organized under the respective queries. The tool allows for editing and organizing the queries and marked results to support for the creation of representations. Gotz (2007) introduced ScratchPad, a browser extension designed to capture, organize, and use Web information. In order to facilitate sensemaking tasks, users are able to create snapshots of Web pages, which can be organized, annotated, modified, and linked together to create representations. SearchTogether (Morris & Horvitz, 2007) supports sensemaking collaboration by providing overviews of captured pages, rating and commenting functions for online content, and communication tools for the collaborators. All of these tools support the key sensemaking activities of data extraction and the creation of representations.

In the context of information visualization, the sensemaking process is characterized as the core of the *knowledge crystallization* process. Its goal is to provide the most compact possible representation of a data set relative to a task (Card, Mackinlay, & Shneiderman, 1999). The knowledge crystallization model includes, in effect, the same cognitive processes and operators as sensemaking, operationalized into four distinct stages: acquiring information (e.g., through searching), making sense of it (e.g., extracting information and finding schemas), creating something new (e.g., authoring a new piece of information), and acting on it (Card, 2003). Information visualization can be used to facilitate the knowledge crystallization tasks at various levels (Card et al., 1999). Card (2003) divides interactive information visualization tools into two layers: the *infosphere* and the *information workspace*. Information is retrieved from the external infosphere, such as the Internet, to the information workspace, where it is integrated and visualized via visual knowledge tools and visually enhanced objects. The purpose of information visualization is thus to lower the cost of accessing actively used information.

Standard Models of Information Retrieval and Seeking

According to Hearst (1999), the standard search-based information seeking process can be characterized as the following sequence of steps:

1. Recognize the information need.
2. Select the information repository to search.
3. Formulate a search query.
4. Send the query to the system.
5. Receive the results.
6. Evaluate and interpret the results.
7. Stop, if the information need is fulfilled, or
8. Reformulate the query and return to Step 4.

This process whereby the user narrows down the result set on the basis of successive query refinements is by nature iterative. This iterative cycle forms the basis of many other theoretical models of information seeking (e.g., Shneiderman et al., 1997; Sutcliffe & Ennis, 1998). It is also the foundation upon which the fundamental interaction model of most Web search engines is currently built.

Dynamic Models of Information Seeking

Observational studies of information seeking have found that users' information needs change during the search process as a result of interaction with the search system, and hence the iterative model does not accurately capture the richness of real information seeking processes (Hearst, 2009). The information seeking process can exhibit both *systematic* aspects, and follow a heuristic such as the standard model, and *opportunistic* aspects, dependent on how the individual factors affecting information seeking interact (Marchionini, 1995). This fact has led to the development of new models that better account for the dynamic nature of information seeking, describing how users utilize different search tactics and strategies to search and make sense of the results.

The "berrypicking" model (Bates, 1989) makes the observation that search *evolves* as the user encounters new pieces of information, and the query and search terms continuously shift to accommodate the new directions of information seeking. The query is not satisfied by some final set of results after an optimal query; instead, people engage in "bit-at-a-time" retrieval of pieces of information at each stage of the search. Marchionini (1995) notes that searches are rarely completed with a single query and result set. Marchionini mentions, that although the information seeking can be modeled as a top-down, sequential process, it is influenced by shifts between sub-processes that may run in parallel, as a result of the intermediate results gained during the process. These sub-processes include understanding, planning and execution, and evaluation and use.

Foster (2004) presented a non-linear model of information seeking behavior, based on interviews with academic information seekers. Three core processes (opening, orientation, and consolidation) and three levels of contextual interaction (external context, internal context, and cognitive approach) are identified, which interact dynamically over time. The information behavior process is cast as a holistic and flowing experience, with no fixed start or end point, whereby different processes are repeated until terminated by either the query or the context.

Information Seeking Strategies

In addition to the holistic models of information seeking, studies have identified distinct strategies that users employ to adapt their behavior within the overall information seeking process as it unfolds. Hearst (2009) divides these strategies into several categories: strategies as sequences of tactics, information foraging theory and information scent, incremental search strategies, and browsing versus search behavior.

Studies have suggested that a user's search strategies can be characterized as sequences of search tactics, which are changed on the basis of triggers motivating a shift in tactics. Bates (1979) provides a list of search tactics grouped into four categories: monitoring, file structure, search formulation, and term tactics. Monitoring tactics are methods that aimed at ensuring the efficiency of search – i.e., comparing the current state of search to the original goal – and cost–benefit assessment of current and anticipated actions. File structure tactics are ways to navigate through the information to the desired content by, for example, breaking down a complex problem into sub-problems or selecting a search method that eliminates as much of the search domain as possible (a form of filtering). Search formulation tactics are related to ways by which the search query can be modified to include or exclude elements in the query. Finally, term tactics are described as methods of selecting and adjusting search terms during query formulation in various ways, such as by using broader or more specific terms or trying other spellings. H. Xie (2002) proposes a similar breakdown of tactics, which are used to reach sub-goals within the larger search goal, called interactive intentions; these include intentions such as identifying, learning and evaluating, among others. The tactics described by Bates (1979) are complemented by the model of D. Ellis (1989), who identifies the stages in the information seeking process as starting, chaining, browsing, differentiating, monitoring, and extracting. Ellis's model has been influential in informing other characterizations of information seeking behavior, such as Bates's (1989) berry-picking model and the behavioral model of Web information seeking by Choo, Detlor, and Turnbull (2000).

Other studies have investigated why searchers switch from one tactic to another. O'Day and Jeffries (1993) identified several triggers that characterize these reasons, as well as stop conditions for ending the searching. The triggers are divided into four categories: the next activity fits the search plan, an interesting finding prompted exploration, a change arose that requires explanation, or something was missing. The stop conditions O'Day and Jeffries identified did not fall into categories as neatly. The two main cases they mention are lack of further compelling triggers and having the sense that an appropriate amount of searching had been performed. Marchionini (1995), on the other hand, divides the causes for stopping between *external* functions (e.g., the setting or features of the search system) and *internal* functions (motivation, knowledge of the task domain or expertise, etc.).

In addition, the perceived cost of utilizing a given strategy can be a trigger for changing one's approach (Bates, 1979; Russell et al., 1993). One model that describes the adaptation of information seeking behavior is *information foraging* theory (Pirolli & Card, 1999). It is an attempt to understand how information seeking is adapted to the available information – i.e., how people change their information access strategies to maximize the amount of information they gain. A key concept in information foraging is *information scent*, which describes how users' behavior is directed by the perceived value and cost of accessing information (Chi, Pirolli, Chen, & Pitkow, 2001). Although Chi et al. (2001) discuss navigation mainly in the context of Web browsing, the concept readily applies to Web search, since the search results (a form of navigation link) can be seen as cues that provide hints about the relevance of the remote content, the result page in question. J. Nielsen (2003) frames the process of evaluating information scent as a cost–benefit analysis of navigation, whereby users make a tradeoff between the potential gains of accessing a piece of information (e.g., a Web page) and the likely cost of accessing and consuming the information (e.g., time and effort or monetary expenditure). The main implication of information foraging theory for the design of search interfaces is that it is possible to influence the search process through design, by, for example, using language that is familiar to the user or showing hints about what kind of information can be found in a particular document.

Incremental Search Strategies

Another way to approach adaptations during information seeking is to consider the incremental search strategies people apply to arrive at an understanding of the search topic. Marchionini (1995) notes that usually the initial result set is the starting point that informs further queries. That is, users approach the search in increments, refining the query to get closer to the desired information (Teevan, Alvarado, Ackerman, & Karger, 2004). The decision to iterate the query is dependent on the user's understanding of the problem, the expected effort, and an assessment of how well the information retrieved matches the search task (Marchionini, 1995). O'Day and Jeffries (1993) liken this kind of incremental search behavior to the sport of orienteering. Orienteering describes the process of exploration through a series of interconnected but diverse searches on a specific theme. Results and understanding of the present query are used for the decision on how to proceed. O'Day and Jeffries identify three distinct search modes in their observation of professional information seekers: monitoring a well-known topic over time, following a plan of information gathering, and exploring a topic in an undirected fashion. These modes each feature orienteering approaches, over several stages. Each stage is followed by analysis of the acquired material, which triggers new search directions. The results of the study by O'Day and Jeffries show that even exploratory information seeking has structure and continuity that could be supported in the system design.

Work by Teevan et al. (2004) discusses orienteering in the context of personal information management and Web search. Their findings contrast the situated, step-by-step approach of orienteering with *teleporting* – that is, focused keyword search activities performed in an attempt to zero in on the desired information target directly. One key finding is that people prefer to utilize orienteering even when teleporting would be feasible. Possible reasons for favoring orienteering include decreased cognitive load (the search can be approached without a need for precise articulation of the query), ability to maintain a sense of location during the search, and gaining of better understanding of the search results because the result was approached along an understandable path (rather than in teleporting directly to a result). Teevan et al. propose several design ideas for search tools, including the inclusion of metadata and trusted sources, for more ready identification of orienteering targets; provision of more context for the results, to aid in understanding them; and support for stepping behavior via clustering and query refinement suggestions.

Browsing vs. Searching

Although search is an important information behavior, not all information seeking is explicitly oriented around keyword search and its associated tactics. The other prevalent information access paradigm, complementing search, is browsing. It too has been widely studied (e.g., Bates, 2007; Rice, McCreadie, & Chang, 2001; Toms, 2000). Bates (2007) defines browsing as a cognitive, motivational and behavioral activity "of engaging in a series of glimpses, each of which may or may not lead to closer examination of a (physical or represented) object, which examination may or may not lead to (physical and/or conceptual) acquisition of the object." Marchionini (1995) discusses the difference between searching and browsing: analytical search strategies require active planning by the user, while passive browsing strategies follow heuristics and are dependent on recognition of relevant information. This is expanded on by Aula (2005), who discusses the increased cognitive demands of searching over browsing, noting that search entails several phases, involving planning and execution of the search queries, result evaluation, and query refinement, while in browsing it is enough to identify links of interest. Hearst (2009) states a more general distinction between searching and browsing: searching produces new collections of information, whereas browsing involves navigation through predefined links or collections of items. However, both searching and browsing can occur during the course of information search, and browsing can play a significant part in the search strategies users employ (Bates, 2007). The interplay of browsing and search is also a key component in search result clustering interfaces, one key topic of this dissertation.

Exploratory Searching

Recently, a hybrid mode of information seeking, called exploratory search, has gained prominence. Distinct from the purely analytical approach to search, it blends querying and browsing strategies, with a focus on learning and investigation instead of information lookup (Marchionini, 2006). White, Kules, Drucker, and schraefel (2006) suggest that current search engines support well-defined information needs but are less suited to situations wherein the users "lack the knowledge or contextual awareness to formulate queries or navigate complex information spaces, the search task requires browsing and exploration, or system indexing of available information is inadequate." Accordingly, White, Kules, and Bederson (2005) identify three typical situations in which exploratory search occurs: 1) the user has partial or no knowledge of the search target, 2) the search moves from certainty to uncertainty as the user is exposed to new information, and 3) the user is actively seeking useful information and determining its structure.

These qualities of exploratory searching have some implications for search interface design. First, it may be possible to use contextual information about the search activity and the target documents to aid in reducing the uncertainty. Second, there is a need to support a wide variety of search strategies and the interfaces should have information-workspace-type features (e.g., note taking), similarly to sensemaking. Third, exploratory search interfaces are likely to be best evaluated in longitudinal, ethnographic, and scenario-based settings (White, Kules, et al., 2005). In many ways, exploratory searching resembles the incremental search strategies discussed above, even if the motivations might differ. Exploratory search usually begins with a tentative query, followed by exploration of the retrieved information for determination of how to proceed (White et al., 2006).

White et al. (2006) note that, since exploratory search is often motivated by the complexity of the information problem and the searcher's limited understanding of the structure of the information space and its terminology, designing interfaces for exploratory search presents unique challenges when compared to supporting search scenarios wherein the target is well known. They highlight the prevalence of features such as interactive search and browsing, visualization, and dynamic workspaces in systems that support exploratory search. These systems provide, for example, a broader range of interactive functionality, such as integrated searching/browsing (e.g., schraefel, Wilson, Russell, & Smith, 2006; Zhang & Marchionini, 2005) and results categorization and clustering (e.g., Kules & Shneiderman, 2005; Kules & Shneiderman, 2008; Käki, 2005b). However, how best to support exploratory search is a challenging problem. Kules, Capra, Banta, and Sierra (2009) point out that exploratory search can also encompass other search tasks, such as lookup and question answering. M. L. Wilson (2009) argues that exploration can also involve several activities in which keyword search is indeed appropriate – for example when the user attempts to express his or her understanding while exploring unfamiliar information – and that this freedom should be retained in exploratory interfaces and visualizations.

2.2 INFORMATION SEEKING BEHAVIOR WITH THE WEB

In addition to understanding the higher-level frameworks and strategies of information seeking, one needs to understand users' information seeking behavior in the context of their everyday information needs. This section of the chapter focuses on reviewing work to understand Web information seeking activities and, especially, the search goals and intent behind the queries. Although the studies reviewed here focus on the Web, there are similarities between Web-related behaviors and information seeking strategies identified in earlier work.

The approaches to understanding information seeking behavior and information goals in relation to the Web can be grouped by methodology. Observational studies have gathered data from fairly limited sets of users with the objective of understanding their information seeking strategies and behavior. Studies focusing on analysis of search logs, on the other hand, have attempted to classify users' search goals and intent on the basis of their queries.

Web Information Seeking Tasks and Strategies

Several studies (Choo et al., 2000; Kellar, Watters, & Shepherd, 2007; J. B. Morrison, Pirolli, & Card, 2001; Sellen, Murphy, & Shaw, 2002) define various classes of Web information seeking tasks, based on observation of actual usage, interviews as well as other subjective feedback methods, such as diaries and surveys. Three distinct categories stand out from the various information behaviors described in previous research. All studies report the users having engaged in *browsing* and *exploration* of Web content; some form of task-specific *information search*; and more complex *information gathering*. These map to the continuum of search activities described by Hearst et al. (2002), which range from directed search to informal search and browsing to knowledge discovery. In addition, several studies report tasks that are not directly related to information seeking, such as making transactions online or routinely accessing specific websites for updates. Table 1 provides a summary of the main categories cited in the previous studies (based, in part, on the classification by Kellar et al. (2007)) and how they align mutually across studies. It should be noted that there is some inherent flexibility in the boundaries of these categories because of different ways of classifying activities. For example, monitoring activities can be thought of being contained within the broader information gathering task (Kellar et al., 2007). Similarly, the increased integration of social networking features into Web search engines blurs the line between searching and communicating.

Table 1. Summary of Web information activities reported in previous research.

	Choo et al. (2000)	Kellar et al. (2007)	J. B. Morrison et al. (2001)	Sellen et al. (2002)
Browsing/ Exploration	Undirected viewing	Browsing	Exploration	Browsing
Information search	Informal search	Fact finding	Finding	Finding
Information gathering	Formal search	Information gathering	Collecting	Information gathering
Other common activities	Conditioned viewing	Transactions	Monitoring	Transacting Communicating Housekeeping

Choo et al. (2000) studied 34 Web users to identify their Web behavior from interviews, questionnaires, and click-stream data of their Web browser use. These authors propose a model that divides the user's behavioral modes into undirected viewing (keeping up with the latest news), conditioned viewing (regular visits to bookmarked sites), informal search (simple searches with search engines), and formal search (searches via several search engines for a specific purpose). Each mode is characterized by a distinct set of information seeking tactics (moves), such as starting, chaining, browsing, monitoring, and extracting. For example, starting (begin the Web session at a portal site) and chaining (follow links) characterize undirected viewing, whereas formal search consists primarily of systematic extracting (find all relevant information about a topic, using multiple search engines).

Kellar et al. (2007) identified categories of Web tasks and the role of search and browser functions in a field study with 21 participants. The task typology is similar to those proposed in previous research, consisting of browsing, fact-finding, information gathering, and transactions. The tasks were examined in light of the dominant interaction attributes, with fact-finding and information gathering being mainly search-oriented while transactions and browsing involved Web site revisits. Only information gathering included high use of browser functions such as bookmarking, copy-and-paste, and within-page search.

J. B. Morrison et al. (2001) analyzed the results of a large-scale Web usage survey. Web search activities were categorized by the purpose of the search, the methods used, and the content of the information sought. The methods included exploration (general searching for information without a particular goal), monitoring (making repeated visits to specific Web sites), finding (purpose-triggered searching for a particular piece of information), and collecting (purpose-driven seeking of multiple pieces of information). The methods suggested above clearly align with the information seeking modes suggested by Choo et al. (2000).

Sellen et al. (2002) studied the Web use of 24 knowledge workers by interviewing them with reference to their Web history. The authors sorted participants' activities into six categories, three of which are similar to the categories discussed above: browsing, finding, and information gathering. In addition, they identified activities such as transacting (using the Web to execute a transaction), communicating, and housekeeping (using the Web to check up on Web resources).

The Effect of Expertise on Search Strategies

The effect of expertise on information search strategies has been studied also (Aula, Jhaveri, & Käki, 2005; Hölscher & Strube, 2000; Navarro-Prieto, Schaife, & Rogers, 1999). The findings from the survey study by Aula et al. (2005) point to there being certain expert strategies for both search and repeat access. In search activities, experienced users utilized several tabs and browser windows to manage the search process and appreciated the benefits of category-based search engines for getting an overview of the search topic and query term suggestions. In information re-access, experienced users utilized search engines for revisiting previously found material and made extensive use of bookmarks to organize the information they found.

Navarro-Prieto et al. (1999) developed a framework for Web searching based on observational studies of novice and expert Web searchers engaging in fact-finding and exploratory search. They identified three search strategies: 1) top-down strategy, which involves finding a general site that provides information organized into categories, or starting with a very general query that is subsequently narrowed down; 2) bottom-up strategy, used primarily by experienced searchers for specific fact-finding searches, wherein users typed specific keywords into the search engine and the information seeking consisted of selectively following the links in search results until the information need was satisfied; and 3) mixed strategy, used only by experienced searchers, which entailed utilizing multiple windows to search for the required information – similarly to what Aula et al. described (2005). According to Navarro-Prieto et al., the key difference in the search models between novice and expert searchers is that experts engage in searching that follows a plan, which accounts for the search goal and prior understanding of the form in which the information is likely to be organized on the Web, both of which influence the choice of search strategy and direction. Novices, in contrast, exhibited a very directed strategy that was driven primarily by the external representations (i.e., how the information was presented).

Hölscher and Strube (2000) developed probabilistic models of Web information seeking that describe both the higher-level browsing vs. searching strategies and the steps in the process of using a search engine. Their findings indicate that "double experts" (i.e., people possessing both domain and search expertise) were the most successful in finding the information needed, whereas "double novices" were the least successful and engaged significantly more in inefficient query reformulation. When facing difficulties in finding the desired information from the search results, novices tended to resort to backward-oriented behavior (going back to previous pages or search results), while experts were more likely to attempt more creative solutions such as forward browsing or switching search engines.

It is evident on the basis of previous findings that novice searchers require more structured assistance from the search interface – both to execute more successful queries and to understand their search topic better and make sense of the information provided by the search engine.

Web Search Goals and Intents

Since this dissertation focuses on Web search, it makes good sense to focus on the operationalization of information needs in the form of search queries. Here the central concepts are defined in terms of search *goals* and search *intent*. Jansen, Booth, and Spink (2008) define user intent in Web search in terms of the "affective, cognitive, or situational goal" that is expressed by interacting with the search engine. They emphasize the notion that intent determines the type of resource desired by the user and the central role of search queries as the expression of intent. It follows that it is possible to identify the characteristics of different user intent categories by analyzing the queries.

In one of the seminal works on Web search, Broder (2002) introduced a classification of Web search intent types and drew a distinction between searches motivated by an information need and those driven by other intents. Searches are classified into three categories, by intent: informational (the user wishes to find information that he or she assumes to be available on the Web), navigational (the user is looking for the URL of a site he or she wants to reach), and transactional (the user is looking for a site on which to perform an online transaction – e.g., download a game or shop for products). According to an analysis of AltaVista search logs for English-language queries, 20% were navigational, 48% informational, and 30% transactional. Broder notes that, while search engines generally support informational and navigational queries, satisfying transactional queries requires semantic analysis of the query intent and integration of external databases (resources outside the set of Web pages indexed by the search engine). It should be noted that modern search engines do this increasingly, by, for example, including advertisements and content relevant to the information need (e.g., in local searches for services).

Rose and Levinson (2004) explain the search goal as the answer to the question "[W]hy are you performing that search?" They provide a conceptual framework of user goals that extends Broder's (2002) taxonomy. The main difference lies in their more detailed classification of queries under informational and transactional (resource) goals; subtypes of informational goals include, for example, directed and undirected search, advice (the user seeks advice or instructions), locate (the user wants to find out where something can be obtained in the real world), and list (the user wishes to get a list of web sites to investigate further). Similarly, resource goals are divided into categories on the basis of whether the user wishes to download, interact or obtain a resource, or is looking for entertaining content. Jansen et al. (2008) extended the framework of Rose and Levinson through synthesis of previous studies. Their analysis adds a third level, one that specifies further subtypes for directed informational and transactional queries (e.g., online/offline and open/closed). In addition, they define two additional third level subtypes for transactional queries, which describe searches aimed at generating the desired information directly in the presentation of the search results.

The framework offered by Broder (2002) and that of Rose and Levinson (2002) both were developed from fairly small sets of manually classified searches. More recent work on the topic has focused on automatically classifying the intent behind search queries. This has met with varying levels of success (Jansen & Booth, 2010; Jansen, Booth, & Spink, 2007; Lee, Liu, & Cho, 2005). For example, Lee, Liu and, Cho (2005) focused on identifying informational and navigational queries and had a success rate of 54%. Jansen et al. (2007) classified four million queries against the categorization proposed by Rose and Levinson and achieved a 74% success rate. Their results further indicate that approximately 75% of queries can be classified with confidence under a single top-level intent category (i.e., informational, navigational, or transactional). Later work (Jansen & Booth, 2010) extended the classification to deeper levels of the search goal framework (Jansen et al., 2008), identifying pertinent differences among the top-level query goals and their subtypes. The implications of accurate real-time intent classification for search interface design are clear. Not only could searchers be provided with more accurate results, but also the interface functionality could be adapted to suit the information need better.

2.3 INFORMATION VISUALIZATION

There are close links between information visualization and information access, for these fields address the same problem: assisting human beings in the acquisition of insight. Information visualization specifically deals with the selection, processing, and presentation of information. Searching can be one way of selecting which information is to be presented.

The literature offers several definitions for information visualization (e.g., Card et al., 1999; C. Chen, 2005; Purchase, Andrienko, Jankun-Kelly, & Ward, 2008). For example, Purchase et al. (2008) define information visualization as utilizing "computer graphics and interaction to assist humans in solving problems." The core activity in this problem-solving is *visualization*, the human cognitive activity of forming mental models of the phenomena being investigated. In information visualization, the computational features of modern information technology, including processing power, interaction capabilities, and high-resolution displays, support the process of visualization (Spence, 2007). In effect, information visualizations are aids that enable *external cognition*, the human information processing activity that combines internal human cognitive processing with perception and manipulation of external representations (PARC, 2012). When cast in the context of information visualization, search user interfaces are clearly external aids of this nature, and it is of fundamental concern how search results are represented (encoded into visual form), presented (laid out on the display), and interacted with for purposes of aiding in external cognition.

Because of its multidisciplinary nature, information visualization lacks unified, clearly defined theories (Purchase et al., 2008). Instead, it makes use of theories and frameworks from other disciplines, such as cognitive psychology (see e.g., Ware (2008) for a thorough introduction to human visual perception) and statistical data graphics (e.g., Tufte's (1983) principles for design of data graphics), to ground visualization design practice. In addition, many useful models describing the information visualization process "pipeline" from raw data to visualization have been proposed in the literature. Reference models of the information visualization process describe the path from data to visualization in varying levels of detail, corresponding to the models of information seeking. Similarly to classifications of information seeking strategies, taxonomies of information visualizations describe data types, visualization operators, and tasks that users can perform with the visualization.

Reference Models for Visualization

Reference models for visualization describe the information visualization in terms of a pipeline of stages wherein data are turned into interactive visualizations (Card et al., 1999; Spence, 2001). The model proposed by Card et al. (1999, p. 17) labels these stages as data transformations, visual mappings, view transformations, and human interaction. Spence's model (2001, p. 13) is broadly similar; the stages are called selection, encoding, presentation, and interactive control by viewer. According to the model of Card and colleagues, data transformations turn raw, idiosyncratic data into organized data tables, which, in turn, are mapped into visual structures (e.g., on-screen objects with graphical attributes, such as tables or graphs). Visual structures are further transformed into views by means of various techniques. Location probes reveal additional information about visual objects, based on the location of the user's selection (e.g., showing more details about the selected object); viewpoint controls transform the viewpoint through affine transformations (e.g., zooming and panning the view); and distortion techniques modify the geometry of the visual structure (e.g., a hyperbolic transformation maps a 2D plane to a circle). Human interaction can take place in every stage of the process, with, for example, the choice of which sets of raw data are included, selection of the mappings into visual structures, and manipulation of the view transformations. In practice, these steps form a complex loop of interrelated actions (Card et al., 1999). Interaction with the visualization yields insights that lead to new selections and changes in encoding and presentation, resembling the process described by the dynamic models of information seeking (Bates, 1989; Foster, 2004) introduced previously.

The data-state model of Chi and Riedl (1998) extends the pipeline model by including multiple values and views and applying a state model, which models the process as a series of data states and operators that modify them. Because of their expressiveness in breaking the visualization design down into a series of data- and view-oriented operations, these models have served as the basis for information visualization toolkits such as Flare (http://flare.prefuse.org/) and prefuse (Heer, Card, & Landay, 2005), which facilitate creation of interactive visualizations. Also, information seeking frameworks have been adapted to the context of visualization design. Mann (1999) utilized the information seeking model proposed by Shneiderman et al. (1997) to describe how visualization could assist in the formulation, initiation, result assessment, and refinement stages of search.

2.4 SUMMARY

This chapter has reviewed existing research on models of information seeking processes and strategies, Web information seeking tasks, and search goals. The previous research provides a context in which the work presented in this dissertation is situated. It also defines a vocabulary of information seeking activities that enables systematic discussion of the proposed search interface techniques with respect to users' core information needs and behaviors.

Previous research also highlights the interrelatedness of information seeking and information visualization in the overall context of sensemaking. Shneiderman (1996) notes that the distinction between these is subtle, since they share common goals. He further argues (2008) that there is a need to examine how interactive information visualizations could be integrated with exploratory search technologies, since provision of visual overviews of Web search results has potential to improve human search performance.

3 Search User Interfaces

The user interface is the user-facing aspect of an information search system. It is critical to the success of the information seeking process that the user interface be usable and aid rather than hinder the fulfillment of the information need. The way the search interface is designed can significantly affect search behavior, especially in the case of inexperienced searchers (Navarro-Prieto et al., 1999). For this reason, Web search engine user interfaces should be simple enough to be used by the general population with minimal assistance from information retrieval experts (Resnick & Vaughan, 2006). This presents several challenges for the design of search interfaces.

This chapter summarizes research on the design of user interfaces for search, from general guidelines to techniques used for organizing and presenting search results. Thorough reviews of existing research on search user interface design practice are provided by Hearst (2009) and M. L. Wilson et al. (2010). The organization applied by M. L. Wilson et al. was adopted as the basis for the discussion in sections 3.2 and 3.3.

3.1 Search User Interface Design Guidelines

While general guidelines for good user interface design (e.g., Shneiderman, Plaisant, Cohen, & Jacobs, 2009) are also applicable to the design of search interfaces, guidelines specific to the design of search interfaces have been proposed. These vary in nature from general guidelines (e.g., Aula & Käki, 2003; Hearst et al., 2002; Rose, 2006; Shneiderman et al., 1997) to domain-specific guidelines (e.g., Hearst, 2006b; Kules & Shneiderman, 2008).

General Design Guidance

The oft-cited guidelines laid down by Shneiderman et al. (1997) propose the following general rules for the design of keyword-based search interfaces. These are grounded in the four-phase "formulation – action – review results – refinement" framework and "eight golden rules of interface" design proposed by Shneiderman (1992):

1. Strive for consistency across search user interfaces.
2. Provide shortcuts for skilled users.
3. Offer informative feedback.
4. Design for closure.
5. Offer simple error-handling.
6. Permit easy reversal of actions.
7. Support user control.
8. Reduce the load on short-term memory.

The authors also provide practical examples of how to realize the guidelines in interface design, and they illustrate this application with two search interface design case studies.

Whereas the guidelines of Shneiderman et al. (1997) were based on general user interface guidelines adapted to the context of search interfaces, Rose (2006) proposes a concise set of research-based design implications. First, the richness of information needs identified in previous studies (e.g., Broder, 2002; Rose & Levinson, 2004) suggests that different interfaces or forms of interaction should be available for different search goals. Second, the search engine interfaces should be sensitive to the varying cultural and situational contexts of search. Finally, the iterative nature of the search task should be accounted for through provision of tools that support exploration and search refinement.

Hearst et al. (2002) consider the problem of supporting both browsing and searching in the interface by reflecting on the results of usability studies on search systems. Their findings support the notion of Rose (2006) as to differing functionality for different goals: users prefer browsing-oriented interfaces for browsing tasks and direct search interfaces when they know precisely what they are looking for. Useful features identified in previous studies include highlighting search terms in the results, sorting the results by relevant criteria (e.g., date or author), and grouping the results into well-organized categories. Hearst and colleagues highlight helpful features that do, however, require robust algorithmic solutions if they are to be useful, such as spelling correction, query expansion, and relevance feedback. The role of visualization and search result clustering is discussed also, with the conclusion being that these technologies are likely to be most useful for knowledge discovery tasks – akin to the exploratory scenarios addressed by Shneiderman (1997).

Resnick and Vaughan (2006) describe best practice for user interface design for search, gathered from leading researchers in the field and grouped into five domains: content, search algorithms, user and task context, the search interface, and mobility. Of interest here is the design guidance on the content characteristics, context, and search interface. Best practice related to content emphasizes the need for providing mechanisms to structure large corpora that are being searched by inexperienced users. In an echo of Hearst et al.'s (2002) suggestions, the use of structuring methods such as faceted metadata (for controlled corpora) and content clustering are recommended. These methods should consider user requirements, which may feature domain-specific foci such as geographic searches or searches targeting a specific language. With respect to search task types, user, and environment, several issues are discussed. For example, both search and browsing should be supported, to address different tasks; the user should retain control of the use of contextual information; and the search user interface should be customizable, to account for the domain or search system expertise of the user. Resnick and Vaughan also address issues specific to the design of the querying and result presentation features, such as presenting the query keywords in context, organizing large sets of results into categories, and supporting iterative searching by allowing query modification and search within the existing results.

Aula and Käki (2003) studied the search strategies of experienced Web searchers and suggest four guidelines informed by the results, including search term suggestions, to facilitate more accurate queries; explaining the effects of Boolean operators in natural language; providing a search history; and facilitating the evaluation of search results by, for example, using clustering to provide easy access to subtopics within the result set.

Domain- and Application-Specific Guidelines

Domain-specific guidelines focus on particular aspects of the information seeking context, such as the user or setting, or the design of the presentation and organization of the search results. For example, children and elderly users are using search engines increasingly and require support that addresses them specifically. Hutchinson (2005) suggests several guidelines for searching and browsing interfaces for children that are related to the design of category structures and navigation systems, and she discusses the pros and cons of flat and hierarchical category structures for organizing content. Aula and Käki (2005) developed a search engine for older users, which offers several design improvements, including interface simplicity and icons that show the type of the result document (e.g., PDF, Word, or HTML document). They also underline the importance of observing the intended user group so as to understand what functionality to exclude and include when simplifying the interface design.

Several studies have focused on providing design recommendations for category-based search interfaces. Hearst (2006b) offers recommendations for interfaces that utilize hierarchical faceted metadata. In the design of the navigation hierarchy, approaches that combine the presentation of the currently selected level, the trail to the higher level, and the options directly below the selected level into a single compact level appeared useful. The layout of labels within the facets poses another challenge, with users preferring a familiar ordering in facets (e.g., alphabetical, numerical, or by number of results). Deciding on the number of facets to show is another challenge. This is affected by the search domain. Directed tasks may benefit from a smaller number of facets relevant to the search goal, while exploratory searches could benefit from exposure to the full structure. Maintaining consistency in the facet display was also considered essential from a usability perspective (i.e., showing facets with no content instead of suppressing them). Hearst also discussed the importance of graphic design, especially in terms of visually suggesting to the user what to do next and how to accomplish it (e.g., use of whitespace and item layout to indicate relationships), incorporating keyword search (broad matches against the content of items are usually preferred over a focus on facets' labels), and designing query history well (each facet should be contained within a separate visual component).

Kules and Shneiderman (2008) provide the following list of design guidelines for categorized overviews, which are intended for supporting exploratory search scenarios:

1. Provide overviews of large sets of results.
2. Organize overviews around meaningful categories.
3. Clarify and visualize category structure.
4. Tightly couple category labels to result list.
5. Ensure that the full category information is available.
6. Support multiple types of categories and visual presentations.
7. Use a separate facet for each type of category.
8. Arrange text for scanning/skimming.

For my work, the items of interest in the guidelines are those related to the construction of the categories. Kules and Shneiderman (2008) advocate use of stable and meaningful (e.g., topic-, geography-, or language-based) categories, to facilitate reuse of category knowledge between searches. Shallow category structures are suggested, with tight coupling between the category labels and search result list (i.e., clicking on a category label filters the results). However, the authors also acknowledge that different types of categories (faceted categories and clusters) and presentation styles may be more appropriate in different situations. This brings with it an inherent tradeoff: should the designer decide on the appropriate category and presentation method, or should this be left to the user? The latter provides more freedom to explore the results from different viewpoints but may also introduce unnecessary complexity to the search interface.

3.2 Search Result Organization

Application of organization structures to search results is one of the main fields in research on search interfaces for users. Organizing results into category-based structures provides several benefits – for example, in allowing searchers to interact and control the presentation of search results in new ways (M. L. Wilson et al., 2010) and providing overviews that present the dominant themes and concepts of the result set (Hearst, 2009). Category overviews facilitate browsing and navigation of the search results generated via querying, either through series of intentional scan and result selection operations or via undirected, casual exploration of the categories (Hearst, 2009). Several classifications of result organization methods exist. Venkatsubramanyan and Perez-Carballo (2007) divide the organization methods into text categorization, text clustering, and location indexing. The key differences among these are that categorization classifies the results against an existing set of concepts, clustering computes these concepts from the results themselves, and location indexing uses location metadata from the results in performing the assignment. Similarly, M. L. Wilson et al. (2010) divide the common approaches to categorization into four techniques: hierarchical, faceted and social classifications, and automated clustering. The latter division is used below as we summarize the research on search result organization.

Hierarchical Classification

Hierarchical classifications organize results into groups in line with a specific, predefined structure such as the Dewey Decimal System (Dewey et al., 2011), used in library informatics, or by using other human-generated classification hierarchies – for example, the Open Directory Project (http://dmoz.org/) taxonomy of Web content (see Figure 1). The organization of documents for the classification can be done manually or, alternatively, automatically by means of various classification algorithms.

The benefit of fixed category systems is that the resulting structures are typically logical, consistent, and familiar to users (Hearst, 2009). It can be argued, however, that the usefulness of classification in practical Web search scenarios is somewhat limited because the classifications need manual organization and maintenance, although automated methods are being developed.

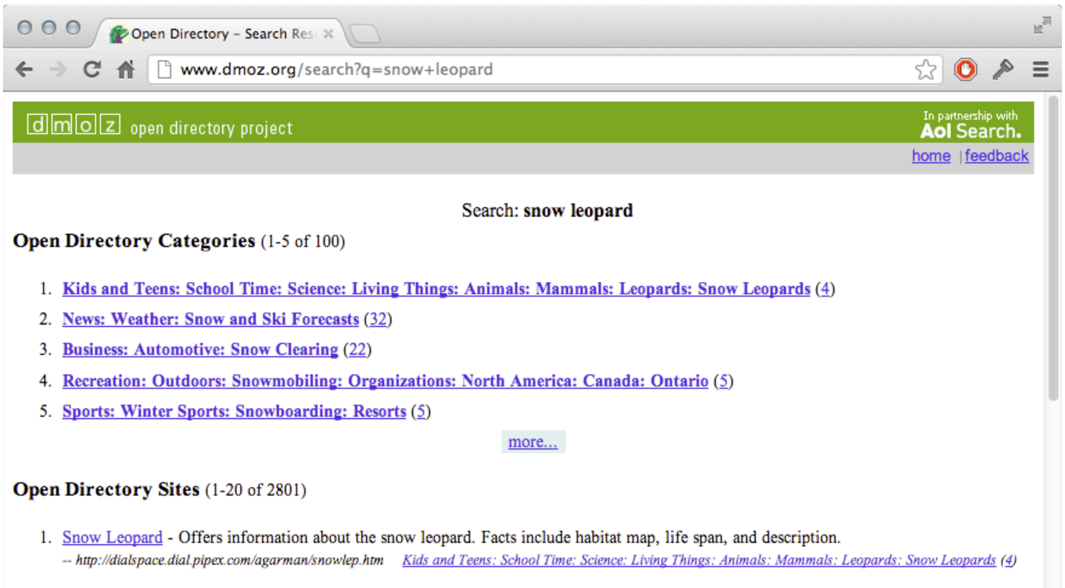


Figure 1. Open Directory Project interface displaying results for the query "snow leopard."

Several noteworthy classification interfaces have been suggested in the literature. M. Chen, Hearst, Hong, and Lin (1999) introduced Cha-Cha, an interface for grouping results in intranet searches around the structure of the site. It is based on recording the hyperlink paths to pages from the server root and combining these paths into an outline to show the structure within which the search results are found. For example, in this system, the home pages of the various departments of a university may form the top-level categories under which the node pages are hierarchically organized. While user experiments did not provide tangible support for the proposed interface over the ranked result list alternative, a follow-up survey indicated that about two thirds of respondents preferred Cha-Cha's outline view. Anecdotal evidence suggested that the interface might be useful in situations wherein information is hard to find.

Hao Chen and Dumais (2000) proposed a system called SWISH for automatically grouping Web search results into hierarchical categories. A support vector machine classification approach was used to order arbitrary search results into a human-generated two-level hierarchy. A user study showed the benefits of the proposed approach over the ranked result list both subjectively and objectively. For example, the participants were less likely to give up on searching and were faster in completing the task when using the proposed category interface. A follow-up study (Dumais, Cutrell, & Chen, 2001) focused on experimenting with different presentation styles for the category information and found the inclusion of the category information to be effective, whether part of the ranked result list or when results were organized into a visual category structure. The best results were seen when both the category titles and page titles were shown. Designs that eliminated either the category titles or page titles were still more efficient, but users preferred the interface that showed category titles to the version that removed them altogether.

Similar objective and subjective benefits of category use were reported by Drori and Alon (2003). In their user study, the interface that displayed page and category titles along with relevant lines from the documents was found superior to alternative designs in terms of search time, accuracy, and user satisfaction. Interestingly, the addition of categories was found to improve search time more than did the addition of relevant lines from the document, which suggests that the inclusion of categories may have a larger effect on the effectiveness of search than the various strategies for showing the actual result descriptors do.

Faceted Classification

Faceted classifications organize documents by means of multiple, orthogonal categorizations, in effect assigning multiple labels to each item, unlike hierarchical classification, which places each item in a single category. Selecting a label in a faceted interface filters the document collection to produce all items assigned that label and its subcategories, and selecting labels from different parts of the faceted hierarchy creates a subset among from those items that are assigned to all of the selected labels (Hearst, 2008).

Faceted browsing is commonly used on e-commerce Web sites as a means of searching for and filtering products to relevant constraints, such as price, format, genre/type, availability, and user reviews (see Figure 2 for an example of the faceted browsing interface at Amazon.com).

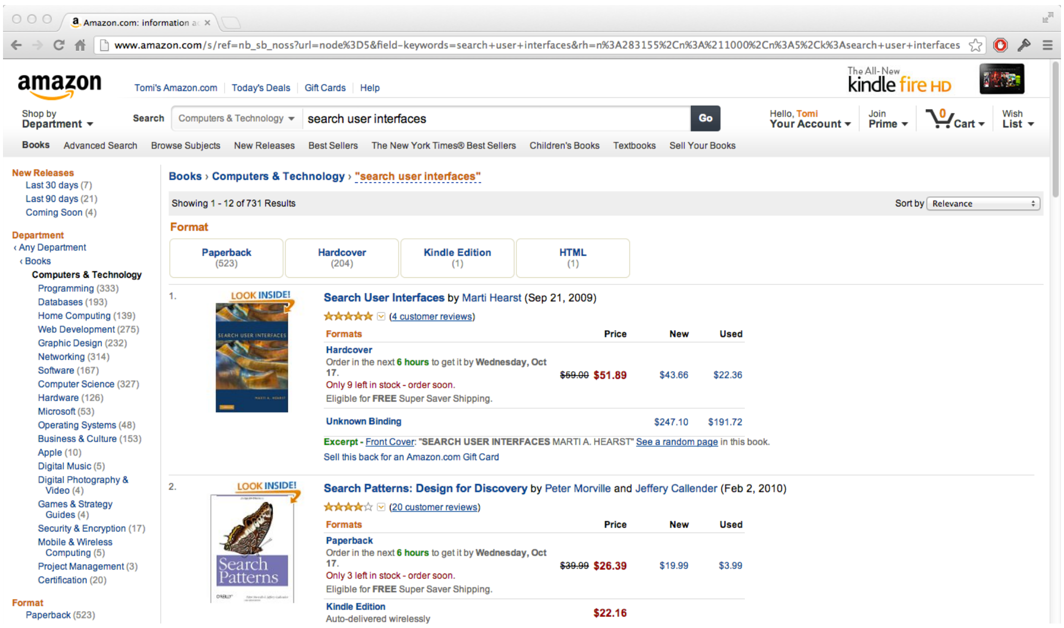


Figure 2. Amazon.com book search results for the query "search user interfaces," which can be browsed along several facets.

Faceted classification systems vary in the type of faceted organization that is used (M. L. Wilson et al., 2010). Bounded systems are based on fixed collections of facets that are applied to a known document collection, while unbounded systems create the facets automatically for unrestricted document sets such as the Web.

Several systems with fixed facet structures have been proposed. The Flamenco interface (Yee, Swearingen, Li, & Hearst, 2003) utilizes multiple faceted hierarchies for browsing image collections. In a user study, the faceted navigation approach was well received by participants, who considered it to help them learn more about the image collection and to be more flexible and easy to use than a baseline keyword search system resembling the Google Image search. The participants felt significantly more confident in having found all task-relevant images from the collection and were more satisfied with their results than with the baseline. Participant remarks suggest that the organization guidance afforded by the facets is one of the key benefits because they suggest ideas about what to search for and guide the search. One shortcoming of the Flamenco interface is that it can be difficult to change selections inside a facet, since all facets are filtered when the user makes a selection (M. L. Wilson et al., 2010). Alternative selection methods have been proposed that overcome this limitation. For example, Huynh, Karger, and Miller (2007) proposed a Web publishing framework called Exhibit, which includes a faceted browsing component that allows multiple selections within a facet. The mSpace interface (schraefel, Wilson, Russell, & Smith, 2006) arranges the facets as left-to-right columns similarly to the Apple iTunes interface, through which any facet can be used for making selections (see Figure 3 for a screenshot of an online demo of mSpace). However, only the facets to the right filter the results, which allows the user to retain a search context by being able to see all the other labels in the leftward columns. Also, the columns can be rearranged, which facilitates flexible exploration of relationships in the data.

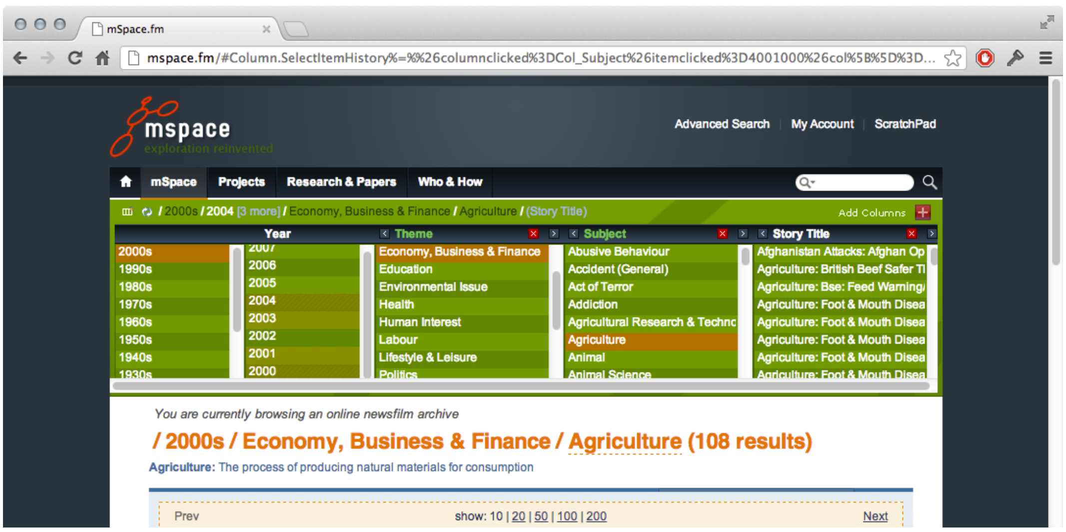


Figure 3. mSpace interface showing the content of an online news film archive filtered according to several facets.

While fixed facets are useful in discrete domains such as product and media search, it is more challenging to come up with suitable facet hierarchies for general Web content. Kules, Kustanowitz, and Shneiderman (2006) investigated different approaches to categorizing Web search results into stable and meaningful categories. This research is particularly interesting, because it encourages the use of fast-feature techniques that increase the practicality of classification for online searching. For example, a distinction is drawn between "rich" categories, based on existing human-generated taxonomies, and "lean" categories that can be inferred from document attributes such as file format, domain, and file date and size. Distinction is made also between fast-feature and full-feature classification, wherein the former uses information found in the search results to build lean or rich category structures and the latter considers the whole result document when doing so. A third distinction is between offline and online classification. Initial experiments showed that online, fast-feature categorization proved feasible at query time; it was able to classify 76% to 90% of top-100 results successfully. A follow-up user study (Kules & Shneiderman, 2008) showed that the fast-feature categories helped users feel more organized and aided in result evaluation during exploratory searches. Categories also enabled search tactics such as making broad queries followed by narrowing with categories, organizing examination of results by category, and assessing the success of the query in view of the category overview.

Significant challenges remain in creating faceted search interfaces for less well-structured or unknown domains. One issue is how to automate the creation of facet structures and the assignment of labels. Several methods have been proposed for automated facet construction. These utilize external resources such as WordNet (Stoica, Hearst, & Richardson, 2007), Wikipedia (Li, Yan, Roy, Lisham, & Das, 2010), or both (Dakka & Ipeirotis, 2008). Human ratings in user studies have shown that the resulting facets are useful for searching and navigating content on specific Web sites. For example, Capra, Marchionini, Oh, Stutzman, and Zhang (2007) found that automated facet extraction provided for a feasible alternative to a manually created classification hierarchy in the context of search of a governmental Web site. It is unclear, however, whether these methods would function as well with general Web content.

Another relevant issue for faceted interface design is how to automate the selection of facets and labels to show to the user, especially in large domains, where manual selection of optimal facets may not feasible and displaying all possible facets and labels could overwhelm users. Koren, Zhang, and Liu (2008) suggest the use of explicit user and collaborative feedback to personalize the facet structure on the basis of how users interact with the facets. Hearst (2008) describes how the commercial search engine Yelp (http://www.yelp.com/) addresses this issue by reordering the facets and adapting the labels within facets in view of the content of the keyword search over the document collection and automatically eliminates facets that are not applicable once the user makes a category selection.

Automated Clustering

While using hierarchical and faceted classifications is feasible with limited collections, their utility for generic Web search tasks may be lacking. Even if automated categorization of the whole Web were possible, classification categories may not be suitable for large-scale Web search because some topics associated with a query may not be represented in the classification (Kummamuru, Lotlikar, Roy, Singal, & Krishnapuram, 2004). An alternative is to use automated clustering techniques, which form collections of interrelated documents on the basis of similarity. As such, clustering is based on the cluster hypothesis (van Rijsbergen, 1979), which states that mutually similar documents tend to be relevant for the same requests. A comprehensive review of clustering algorithms for Web search results is presented by Carpineto, Osínski, Romano, and Weiss (2009). In addition to research prototypes, clustering has been used in some commercial search engines – e.g., Yippy (http://search.yippy.com/), which is based on Vivisimo clustering technology (see Figure 4), and iBoogie (http://www.iboogie.com/).

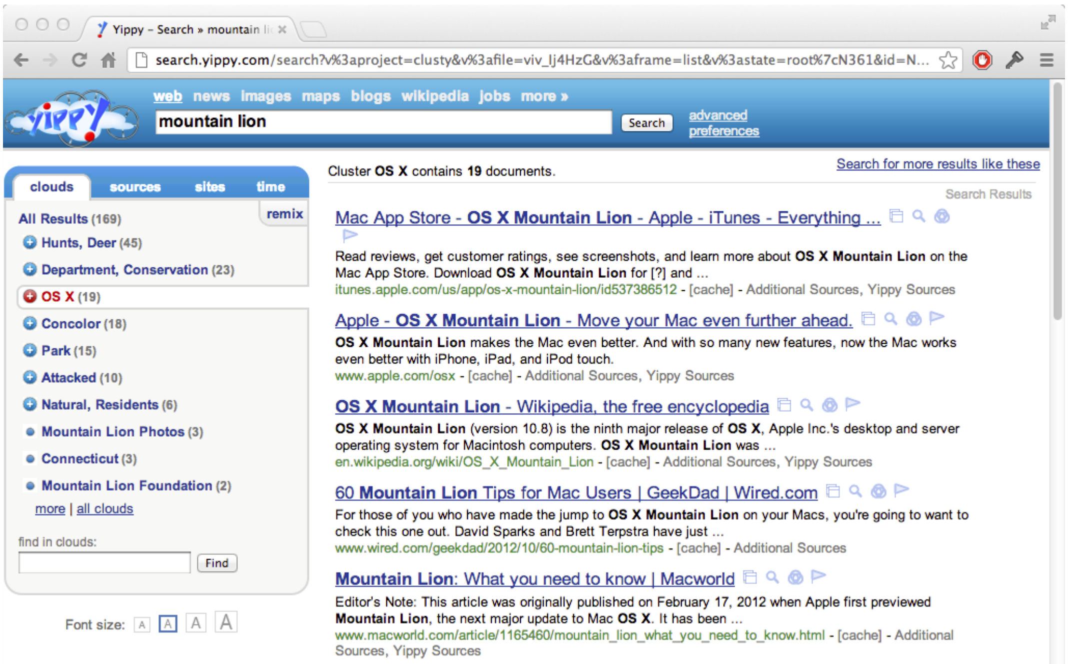


Figure 4. The Yippy search engine showing results for the query “mountain lion,” with cluster titled “OS X” selected.

Search engines that otherwise utilize the ranked result list as the presentation interface may engage in an implicit for of clustering (Carpineto, Osínski, et al., 2009), whereby the results are diversified through integration of results for several topics into the results page.

Clustering combines the benefits of keyword search and classification-based browsing: the cluster can provide highly specific topics in response to ambiguous queries, into which the user can drill down to focus on results of interest (Carpineto, Osínski, et al., 2009). One key benefit of clustering here is that the clusters emerge organically from the content of the documents and so are specific to the results retrieved for the query in question. For example, a clustering algorithm that utilizes the most common words in the documents could produce clusters titled "panthera leo," "king," "os x," and "detroit lions" as a response to the query "lion," reflecting different subtopics, such as animal species, movie, operating system, and sports team. Clustering can also be sensitive to timely topics that organically emerge in the results, as in extraction of results related to Hurricane Katrina in a search for "new orleans" (Hearst, 2009).

According to Kummamuru et al. (2004), clustering can be based on the overall similarity of the documents (polythetic clustering) or on shared features such as words and phrases appearing frequently in the documents (monothetic clustering). Another parameter in clustering is whether the resulting cluster structure is flat (no explicit structure to relate the clusters to one another) or hierarchical (arrangement of clusters into a hierarchy). A third distinction can be made, between soft and hard clustering: in soft clustering, a document can be a member of several clusters, while hard clustering assigns each document to exactly one cluster. For example, the Findex algorithm (Käki, 2005b) used in the studies reported upon in this dissertation is a monothetic soft clustering algorithm that produces flat clusters.

Early systems based on overall inter-document similarity such as Scatter/Gather by Cutting, Karger, Pedersen, and Tukey (1992) clustered documents into groups by applying weighted document term vectors. A follow-up experiment (Hearst & Pedersen, 1996) suggested that clustering could be an effective method for organizing retrieved search results because relevant results tended to fall into a few key clusters. By looking into these clusters, the user could focus on relevant documents and bypass the non-relevant ones.

However, among the drawbacks of clustering by overall similarity is that documents can be similar or differ from one another in several ways, which makes the results unpredictable and hard to understand, especially if hard clustering is used (Hearst, 2009). For example, a document about concussions in football could conceivably be placed within a medical cluster related to health effects or a legal cluster about lawsuits filed by former players. Carpineto, Osínski, et al. (2009) also note the difficulties that polythetic clustering techniques have with producing understandable cluster labels, arising from the need for the labels to be extracted from the cluster content after clustering.

Monothetic clustering methods using a single shared feature can overcome the labeling issues by utilizing immediately recognizable and meaningful features (Carpineto, Osínski, et al., 2009), such as frequent words and phrases (Ferragina & Gulli, 2005; Käki, 2005b; Zamir & Etzioni, 1999). Methods based on frequent words and phrases are also able to place results under several concepts, allowing a result to be reached through several paths and thus avoiding some of the topic conflation issues that arise with polythetic methods. Carpineto, Osínski, et al. (2009) distinguish between methods that are description-aware (attempt to maximize the descriptiveness of the cluster labels) and description-centric (prioritize labels over document allocation). Techniques of the latter sort attempt to ensure the labels' conciseness and comprehensibility in various ways – e.g., by utilizing external resources such as previous queries (Wang & Zhai, 2007), Wikipedia content (Carmel, Roitman, & Zwerdling, 2009), human ratings (Zeng, He, Z. Chen, Ma, & Ma, 2004), and the co-occurrence of concepts within the results (Kummamuru et al., 2004).

User studies have shown that clustering can be useful in supporting Web search. For example, when the query is vague, displaying the dominant themes in the results helps users determine whether the search has provided useful results on the basis of the cluster labels (Hearst, 2009) or eliminate groups of documents from consideration (Hearst & Pedersen, 1996). Also, clustering has been shown to provide tangible benefits in search efficiency. For example, its use resulted in fewer sessions wherein no results are followed and in multiple results being selected more often than from a ranked result list (Käki, 2005b; Zamir & Etzioni, 1999), which suggests that clusters were particularly useful for exploratory search sessions. Clusters also allow users to reach potentially relevant results further down the result index than they normally would with a ranked result list (Käki, 2005b; Y. B. Wu, Shankar, & Chen, 2003). When Carpineto, D'Amico, and Romano (2012) compared result diversification (i.e., mixing results with different topics in the ranked result list) and clustering, they found that clustering is better than result diversification in cases wherein full subtopic retrieval is desired. They also note that it is difficult to gain improvements over the baseline ranked result list when full subtopic retrieval is not the aim. However, the evaluations did not involve users, and the authors note that user studies are needed for correlation of their findings with real user experience.

The results of large-scale study of the commercial Vivísimo clustering search engine (Koshman, Spink, & Jansen, 2006), which appears to operate along monothetic principles (per Hearst (2009)), suggest that clusters are actively used by searchers (in about 48% of searches), cluster use does not significantly change search behavior, and clusters may provide a more direct approach to results of interest than the ranked result list does. The results of Koshman et al. (2006) also suggest that flat cluster lists may be informative enough for most information needs, since the cluster hierarchies were rarely expanded.

Social Classification

The ascendance of social media Web services has given rise to social classifications on the Web. For example, Delicious (http://delicious.com/ or http://del.icio.us/) allows users to save bookmarks of Web pages to a personal collection and tag them with labels, which can be subsequently used for filtering of the content (see Figure 5). These shared vocabularies of tags are known as *folksonomies*, collaboratively created and managed flat social classification systems that can be utilized as one browses and searches for content. The success of social bookmarking can be attributed to its relative ease when compared to the use of controlled vocabularies and bookmarking; however, search interfaces with social classifications have thus far been limited in their search options (Hearst, 2009). Social classifications are usually used to aid in keyword search rather than as the main mechanism for browsing, unlike in other category-based search interfaces.

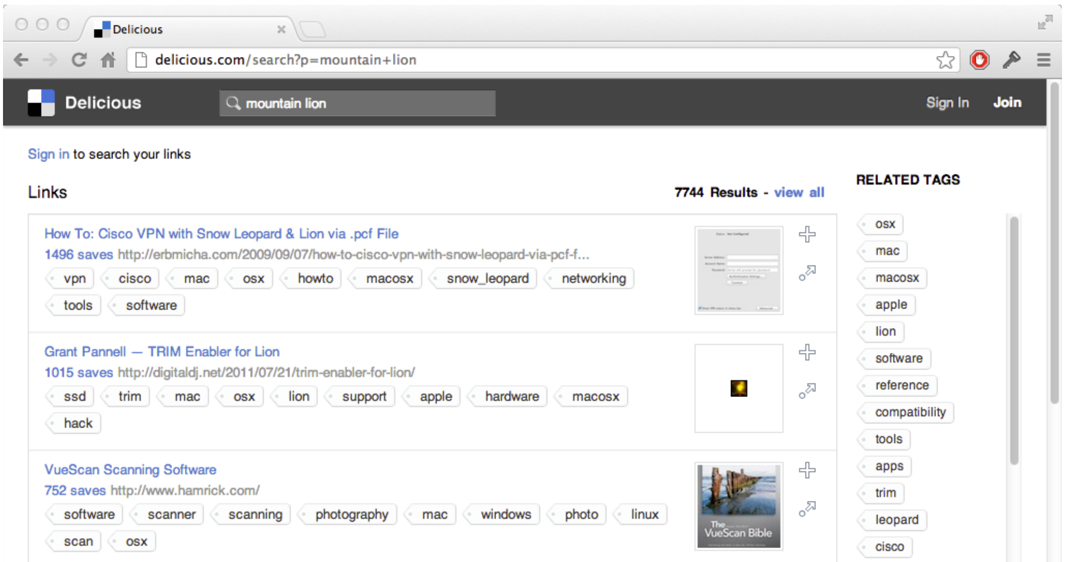


Figure 5. Delicious search results for the query "mountain lion," including the tags associated with the results as well as related tags.

Several studies have investigated the benefits of social bookmarking for exploratory search. Millen, Feinberg, and Kerr (2006) have proposed a tag-based social bookmarking system called dogear for intranets wherein the user-assigned tags can be used in searching and browsing other users' bookmarks. A field study of the dogear system (Millen, Yang, Whittaker, & Feinberg, 2007) showed that users engaged in three types of search activities. The most popular methods were related to community browsing, such as viewing recent and popular bookmarks of other users, browsing other users' bookmark collections, and browsing topics via the tags (e.g., to familiarize oneself with a new concept). The second most common method was keyword search over the bookmark collection. The last activity is concerned with browsing one's own bookmark collection – for example, to revisit previously collected information sources. The results suggest that social bookmarking is a promising method for supporting exploratory search, from personal lookup activities to learning and investigation through community browsing. This conclusion is supported by the findings of Kammerer, Nairn, Pirolli, and Chi (2009), who studied the use of a social tagging system called Mr Taggy. Their findings suggest that tags can be helpful to exploratory search in various ways, such as by disambiguating ambiguous queries, supporting learning and investigation activities, and assisting in gaining of domain knowledge.

Tag clouds are a popular presentation and navigation method for flat social tag classifications. Sinclair and Cardew-Hall (2008) studied the benefits and drawbacks of tag clouds as an interface for information seeking in folksonomies. Their results indicate that tag clouds are useful for browsing and forming general understanding of a topic, with the tag cloud providing a visual summary of the content of the information space. Conversely, tag clouds were found unsuitable for specific information seeking, partly because the tags may not be precise enough to answer the given questions and also because not all of the articles may have been sufficiently tagged by the users.

Other studies have investigated whether social classifications could be used to improve specific Web search functionality, such as indexing and ranking of pages and recommendations. Heymann, Koutrika, and Garcia-Molina (2008) found that social bookmarking data (in their case, harvested from Delicious) could inform Web content crawling (since tagged URLs tend to be recently updated) and ordering of results, because of the large overlap between tags and query terms and the frequent occurrence of tagged URLs in top search results. However, the impact of social bookmarking is likely to be limited by the small size of the social bookmarking data sets when compared to Web content and by redundancy between the tags and information already used for indexing (e.g., domain, and page title and content). Hotho, Jäschke, Schmitz, and Stumme (2006) proposed a new personalization algorithm, FolkRank, which could be used to recommend to users tags, documents of potential interest, and user communities of interest on the basis of the structure of the folksonomy. A similar approach is advocated by H. Wu, Zubair, and Maly (2006), who also utilized social classifications as input to clustering. Their results suggest that tags could inform the creation of useful hierarchical structures.

Summary

All of the search result organization methods discussed in this section are potentially useful in various search scenarios, but relatively few comparisons of the efficacy of the proposed methods exist. Clustering by visual similarity has been compared to text-based classifications in the context of image browsing (Rodden, Basalaj, Sinclair, & Wood, 2001), and dynamic clustering to domain-knowledge-based classification in the medical domain (Pratt, Hearst, & Fagan, 1999). In both studies, the participants received the classifications more favorably; however, the generalizability of these results to general Web search scenarios is arguably limited. For example, the DynaCat system proposed by Pratt et al. (1999) requires a heavyweight terminology model of the domain.

Hearst (2006a) provides a review of studies of clustering and faceted classifications, outlining the various techniques' strengths and benefits. This is extended by Hearst (2009), who discusses the advantages and disadvantages of clustering versus other categorization methods (mainly hierarchical and faceted classifications). The main arguments by Hearst (2009) and the findings from studies focusing on the individual structure construction methods discussed in this section are summarized in Table 2.

Table 2. Summary of advantages and disadvantages of different content structuring methods.

	Advantages	Disadvantages
Classification and faceted browsing	- The units of meaning are understandable and consistent - Irrelevant results can be ruled out - Facets enable navigation by several topics concurrently	- The categories require manual management - They may not fully reflect all topics in the result set - Automated online classification may not be fully accurate
Clustering	- The system can be fully automated for Web search results - It reveals dominant themes in the results - It is sensitive to idiosyncrasies in results and aids in disambiguation of vague queries - It enables elimination of irrelevant results from consideration	- It can lack consistency and comprehensibility in cluster labeling - Cluster content can be difficult to understand (polythetic clusters) - The usefulness of hierarchical categories is unclear
Social classification	- It can help disambiguate vague queries - It is helpful in learning to understand new domains - Support is given for social information seeking and personal information revisit	- The tags may not cover all information needs - Not all documents are necessarily tagged, which limits access via tags

3.3 Presentation and Visualization of Search Results

Whereas the previous section focused on approaches for organizing the search results into meaningful categories, this section summarizes the research on visual representation of search results. Traditionally, search results are presented in a vertical list using document surrogates (Hearst, 2009), which contain human-readable metadata related to each search result – e.g., the document title, a short summary of or extract from the document content ("a snippet"), the uniform resource locator (URL), and other information items such as document length. Marchionini and White (2008) note that document surrogates are enforcing a pull information seeking strategy: searchers must assess the document surrogate first, then potentially access and assess the document. Conversely, in a push approach, potentially useful information is extracted from the documents and presented in the result listing as-is or through other representations of the content.

In the rest of this section and those that follow, the focus is on reviewing systems that attempt to facilitate the assessment of the result in various ways that resemble the push approach in its broadest sense – by shifting the onus of extracting useful information from the search results and result documents from the user to the system. In essence, this takes place on two levels. First, the system can help the user assess the relevance of the result set returned from the query and help formulate a better query. At the same time, it can aid in assessing the usefulness of individual result documents. As discussed above, classification structures are of use for the former goal. The emphasis in this section is more on individual results' presentation, although the challenge of representing classifications is discussed also.

Regardless of the method of result presentation and content representation, this process often involves a tradeoff between the simplicity of the visualization and the number of document features that are included. Several distinct approaches exist for selecting, processing, and visualizing the selected document features and their relationships. The following subsections summarize the research into the main trends in search result presentation and visualization: presentation of textual summaries in search result listings, visualization of query terms (e.g., locations and frequency), visual result previews (with or without query terms), visualization of classification structures, and other graphical approaches to visualizing relationships among results. The section concludes with a discussion of the overall benefits of complex information visualizations with respect to the efficiency and effectiveness of search.

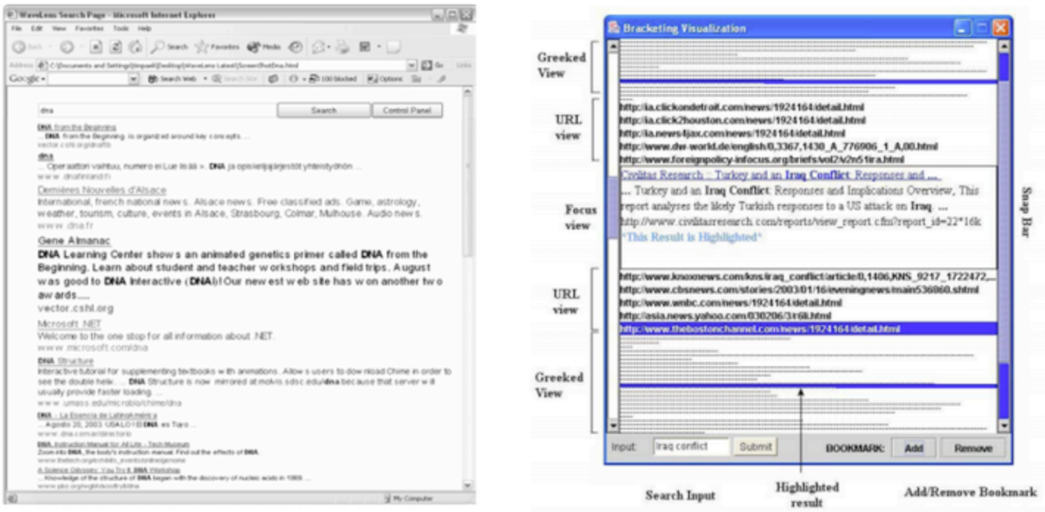
Improvement of Search Result Lists

One of the key pieces of data shown for each search result is the brief snippet extracted from the retrieved document. A significant amount of research has examined how these summaries should be constructed (e.g., which sentences to select from the document content), how the summary should be presented, and at what length.

By far the most common method of constructing the document extracts is to take the query terms into account and produce a query-biased summary of the result (Tombros & Sanderson, 1998), which contains sentences in which the query terms appear. Tombros and Sanderson (1998) found in their user study that such query-biased summaries are superior to showing the first few sentences of document content, the latter being a typical approach for document retrieval. White, Jose, and Ruthven (2003) compared query-biased summary extraction techniques to the methods Google and AltaVista used to create their snippets, reporting similar results. Query-biased summaries helped users search more effectively, with preference for the query-biased approach being reported. Interestingly, the query-biased technique did not prove to be as useful with Google as with AltaVista, possibly because Google already utilized some query-based methods in its summary construction. In later work, White, Jose, and Ruthven (2005) experimented with showing top-ranking sentences selected from the top matching documents. The top-ranking sentences approach led to more effective and efficient searching than that seen with a baseline interface using traditional ranked result lists.

Another issue in result presentation is the length of the summaries. This involves a tradeoff between minimizing the amount of screen space taken and the informativeness of the summary (Hearst, 2009). For some search goals, such as informational queries and question answering, showing longer summaries may be more beneficial than it is for others. Cutrell and Guan (2007) studied summaries of varying length: short (one line of words), medium (two to three lines, the search engine default), and long (six or seven lines) across different information seeking tasks (navigational and informational). Their results show that increasing the length of the query-biased snippet from medium to long improved performance for informational queries but degraded it for navigational ones (possibly because of users focusing more on the textual content and ignoring the URL). Although users had to scroll more, they were more accurate, proved faster, and looked at fewer items when using the extended snippets. In a study of question-answering systems, Lin et al. (2003) found that users overall prefer paragraph-length summaries to other presentation methods (exact answer, answer in the context of a sentence from the text, and answer in the context of the full document). Presentation method did not have an effect on task completion time; however, the extended summaries resulted in progressively fewer questions being posed in scenarios that involved answering multiple questions on the same topic (likely because the extended extracts could be used to answer multiple questions without the need arising to pose new questions).

One alternative in choosing between short and long extracts is to provide more details on demand. Paek, Dumais, and Logan (2004) performed an experiment that compared the traditional search result listing to the WaveLens technique, which showed an extended document summary (containing additional query-biased sentences from the document) either instantly as the user clicked on the result or in increments as the user hovered the mouse over the result (see Figure 6a). The key finding in their work is that users preferred the novel presentation methods to the static summary, and presenting additional query-relevant text from the result document on demand is a viable approach. In another study, Roberts and Suvanaphen (2003) proposed a presentation method called visual bracketing, which presents the selected result in full detail, bracketed on both sides by results displayed with a lower level of semantic detail. In a two-level design, the URLs of the documents were shown in the context views, and in the three-level design (see Figure 6b), they were shown both as readable text (close to the selection) and in greeked text (further from the selection). A small user study suggests that the visual bracketing approach is easy to understand and manipulate. For example, the greeked text aided in visual search by providing visual landmarks.



A. B.

Figure 6. a) WaveLens layout technique displaying search results with a fisheye view distortion (Paek et al., 2004, Figure 1, © 2004 Association for Computing Machinery, Inc.) and b) search results presented with visual bracketing (Roberts & Suvanaphen, 2003, Figure 4, © 2003 IEEE).

Although the results returned by the search engine in the ideal case are selected in such a manner that they satisfy the searcher's information need, it is not always obvious why certain results have been returned. Coyle and Smyth (2005) propose the use of interaction data from a community of searchers (so-called collaborative Web search) to illustrate the relevance of the results. They suggest three explanation types and combinations thereof: relevance (e.g., "85% of searchers with the same query selected this result"), related queries (e.g., "this result was also selected for queries …"), and timing (e.g., "the result was last selected 5 minutes ago"). A user study contrasting the explanations with the standard search result presentation found that relevance was clearly preferred (over 80% of users), because it helped the respondents better understand the usefulness of the results. Timing and related queries were preferred less (about 50%). Another option is to explain how the query affects the results. For example, Aula and Käki (2003) studied expert search strategies and found that even experienced searchers had problems understanding the functioning of Boolean operators in keyword search (i.e., which operator is used by default when the query contains several terms). Aula and Käki suggest that explaining the query in natural language might help to alleviate some of these misconceptions.

It should be noted that much of the research summarized above compared novel content representations and result visualizations to the baseline search result listings as they were nearly 10 years ago. Current search interfaces include many useful features that supplement the textual summaries, such as related query suggestions, and spelling correction. One typical addition is to weave in results from search verticals (specific segments of the content corpus such as blogs or medical information) among the search results. Another form is question answering, which provides the exact answer to the information need (e.g., the query "normal body temperature" could result in a top result providing the answer as 37 degrees Celsius). Yet another additional feature is to include, under the other metadata, query-relevant links to pages on the site returned as the search result (also called "sitelinks").

Visualization of Query Terms

The success of query-biased summaries evidences the benefits of being able to see the presence and distribution of the query terms in the retrieved document (Hearst, 2009), with several visualization techniques having been proposed to address this issue. Two seminal examples of this approach are TileBars (Hearst, 1995) and the visualization tool proposed by Veerasamy and Belkin (1996). TileBars represents the retrieved document with a horizontal rectangle whose width indicates document length (Figure 7). The rectangle is divided into squares that represent document segments, each square representing a specific query term set. The depth on the gray scale within the squares shows the combined frequency of query term instances within the document segment.

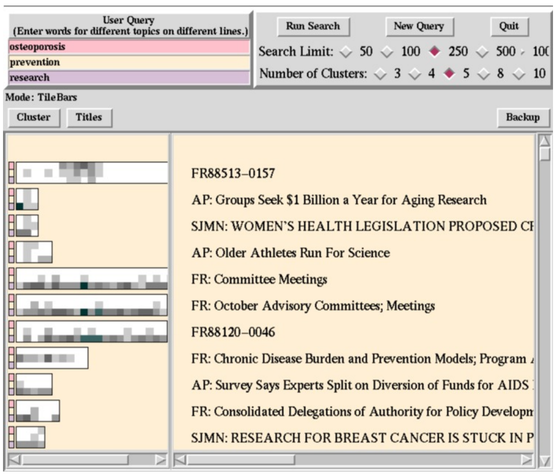


Figure 7. TileBars search interface displaying the results for a search with three query term sets (courtesy of Dr. Marti Hearst).

Several variations of the TileBars approach have been proposed since (Hoeber & Yang, 2006; Mann & Reiterer, 2000; Paper I). The query occurrence visualization approach suggested by the present author as part of this dissertation work (Paper I) adopts a more minimalist design than TileBars, whereby the document is divided into four segments, of equal size (see Figure 8a). The number of glyphs on each of the four rows shown, where a row resembles a line of words on a page, depicts how many times the full query occurs in the same context (within a certain set of words) in the corresponding quarter of the document. In a user study with 18 participants, this visualization was found useful primarily for eliminating poor results. Interestingly, 11 of the 18 subjects did not initially notice the visualization during task completion (in a condition wherein the visualization was not specifically introduced to the participants). This suggests that integrating small visualizations into the search result list does not unduly disturb users. Hoeber and Yang (2006) proposed HotMap, another simplified version of TileBars, which shows one square per query term, colored from yellow to red according to the number of hits in the document (see Figure 8b). Users can also re-rank the results on the basis of the frequencies of the query terms, and a scrollable overview window to the right of the results provides a compressed representation of the top search results. The results of a user study indicate preference for this design over the ranked result list interface, and the use of HotMap resulted in most participants accessing documents of low relevance less often than with the baseline Google-based interface.

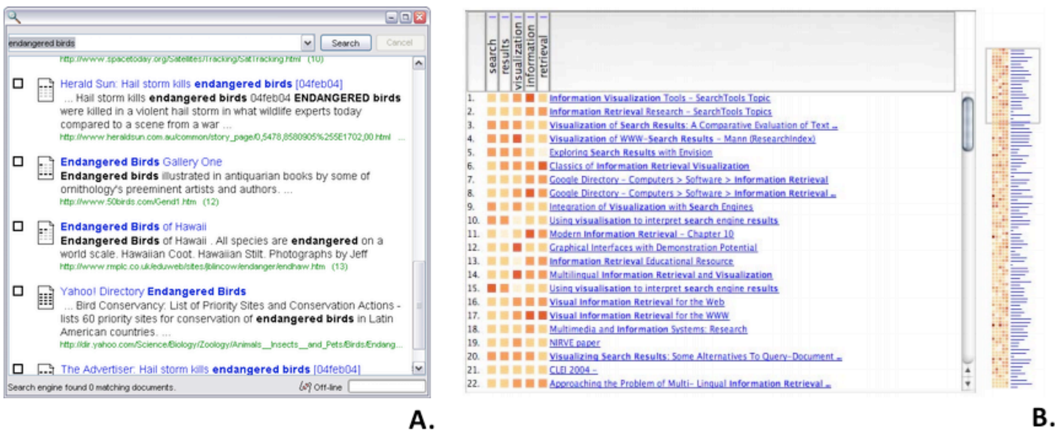


Figure 8. a) The query occurrence visualization icons shown alongside search results for the query "endangered birds" and b) the HotMap interface showing search results for the query "search results visualization information retrieval" (Hoeber & Yang, 2006, Figure 1a, © 2006 IEEE).

The visualization tool by Veerasamy and Belkin (1996) takes a different approach by organizing the result documents in columns and keywords in rows, showing up to 150 documents in a single view. As the authors note, this has the benefit of providing an overview of the query word distribution for the whole set of documents. At the intersection of row and column, a bar showed by its height the contribution of the query word to the retrieval of the given document. Moving the mouse pointer over the column displayed the document title. In addition to the query terms input by the user, the system includes terms that it uses itself to compute the results. A subsequent evaluation (Veerasamy & Heikes, 1997) found that the visualization tool enabled users to assess documents' usefulness about 20% more quickly when the precision of returned results is low. Visualization tool users were also more accurate in their relevance assessments. Veerasamy and Heikes (1997) conclude the reason to be that users utilized the visualization to dismiss clearly non-relevant results from further perusal.

Mann and Reiterer (2000) introduced a system called INSYDER that combined several techniques inspired by existing visualizations such as TileBars (called SegmentView), the approach of Veerasamy and Belkin (1996), a 2D scatterplot display, and a sortable table view (titled ResultTable). Reiterer, Tullius, and Mann (2005) report the results of an extensive user study of INSYDER. In the tasks wherein the participants were given the opportunity to use either the ResultTable or one of the other three visualizations, the users utilized the ResultTable more frequently (it accounted for over 50% of the time used for the tasks). ResultTable was the preferred interface, followed by SegmentView and the ranked result list. However, all visualization conditions, including the ResultTable, had significantly higher overall task times than the ranked result list did.

Some approaches have attempted to visualize the frequency and relationships of the query terms in alternative forms. For example, Anderson, Hussam, Plummer, and Jacobs (2002) introduced pie chart visualization in a prototype system called Semantic Highlighting. Two distinct views are proposed: in the "Pie and Text" view, the pie chart is embedded next to the traditional descriptors, and only the pie charts are shown in the "All Pies" view, which shows more documents (up to 30 at a time), with the URL of the document being visible on mouseover. The documents in both views are ordered by the total number of query terms matched, which the authors argue anticipates user expectations. A limited evaluation suggests that the pie chart visualizations can reduce the time to locate the desired information in fact-finding tasks, although the extent to which this was due to the pie chart visualizations is somewhat unclear, since other helpful features were provided in the full text view (e.g., term highlighting and keyword-to-keyword jumping). However, subjective feedback does indicate that nearly 90% of participants understood the meaning of both visualizations, and all "Pie and Text" users found this presentation to be helpful in locating relevant documents. Also very simple visualizations of query term frequency can be useful. Hoeber and Yang (2008) introduced WordBars, an interface that displayed the most frequent terms extracted from the result documents, which could be used for re-ranking the result list interactively. In their user study, participants' subjective assessments of WordBars were positive, and in most cases also improvement in search performance was found over the baseline ranked result list interface.

In addition, some interfaces have included features wherein the query term location is shown in the context of the full text of the document by means of either the document scrollbar (Byrd, 1999) or creation of a thumbnail representation of the document (Ogden, Davis, & Rice, 1998). Byrd (1999) argues that, while showing the locations of query terms with an iconic display such as TileBars is feasible for helping users decide on the documents to view in the result list, highlighting the term locations in an enhanced scrollbar widget better facilitates deciding which passages within the document to view. The proposed scrollbar "contains a miniature view of highlighted words in the entire document" (Byrd, 1999). Two user studies did not find a significant effect on search performance; however the participants in both studies expressed subjective preference for the visualization over a version that only highlighted the query terms in the text. The Chrome Web browser combines both approaches by showing the instances of the query within the scrollbar while also highlighting them in the textual content (see Figure 9).

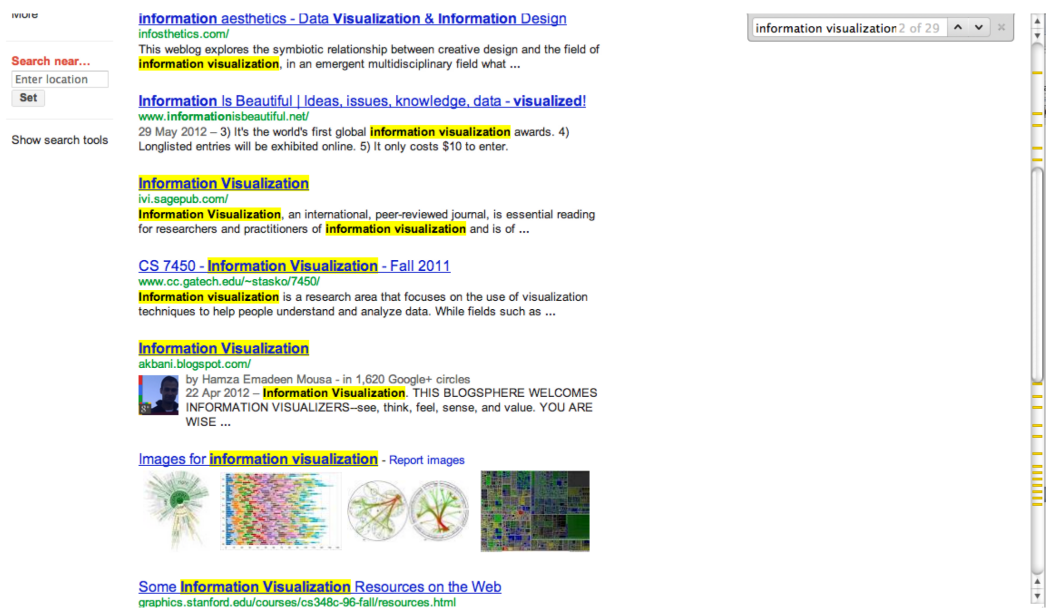


Figure 9. Results of the within page search for the query "information visualization" highlighted inside the scrollbar and in the document text in the Chrome Web browser.

Ogden et al. (1998) studied the advantages of document thumbnails, which show an outline of the content of the document with color-highlighted query term locations. Their small-scale user study, based on the TREC-7 interactive track methodology, did not reveal significant differences between the document thumbnail condition and a baseline interface in terms of performance time, precision or recall. The authors conclude that the effectiveness of interface elements such as the proposed thumbnails is affected by the users' decision-making strategy (e.g., possibly aiding in rejection of a non-relevant document on grounds of the thumbnail not containing a keyword). Accordingly, when the decision-making strategy does not play a significant role in overall task performance, the effect of interface features is minor.

In summary, while it appears that users generally like visualizations that show the location and frequency of the query terms in the result documents, the main benefit of these visualizations lies not so much in their ability to point to useful documents as in the highlighting of non-relevant documents. The benefits of search result visualization in general are discussed in more detail in the final subsection, below.

Thumbnails and Visual Summaries of Web Pages

One drawback of textual summaries is that their assessment is slow when compared to scanning of images (Hearst, 2009). Several researchers have proposed interfaces that incorporate visual summaries either by using accurate thumbnail-sized renditions of the Web page (Aula, Khan, Hong, Guan, & Fontes, 2010; Czerwinski, van Dantzich, Robertson, & Hoffman, 1999; Dziadosz & Chandrasekar, 2002) or applying more complex representations (Teevan et al., 2009; Woodruff, Rosenholtz, Morrison, Faulring, & Pirolli, 2002).

Czerwinski et al. (1999) investigated the effects of individual features (thumbnails, spatial location, and mouseover text) of the Data Mountain 3D environment on the cued retrieval of previously stored Web pages. With the thumbnails, the participants were no slower in retrieving the pages than in the collection phase four months earlier and preferred the thumbnails to mouseover and spatial location. However, thumbnails were only more important in terms of retrieval speed at the beginning of the study, and with time the participants began utilizing other recall methods so effectively that even the removal of the thumbnails no longer disrupted performance. Dziadosz and Chandrasekar (2002) utilized thumbnails in a more traditional Web search scenario, investigating the utility of thumbnail previews alone and in combination with textual summaries. The results from their user study suggest that the combination interface led to more accurate relevance decisions than text-only or thumbnails-only interfaces did; however, no statistical significance testing of the differences is reported. Aula et al. (2010) cite similar results from a series of user studies wherein they investigated the effect of several thumbnail zoom levels, thumbnail dimensions, and combinations of thumbnails and textual summaries on users' assessment of Web page helpfulness. They found that combining the textual summaries and thumbnails results in more accurate relevance predictions – using thumbnails alone resulted in users underestimating the usefulness of the Web page, while textual snippets alone led to overestimation. They also found that a size of 200×200 pixels appears to be optimal with respect to accuracy and that thumbnails with a smaller zoom factor (20% zoom level) result in more accurate helpfulness predictions from users than zoomed-in thumbnails do (38% zoom level).

Figure 10. Enhanced thumbnail of a textual Web page in which the query terms are highlighted and displayed with larger font size (Woodruff et al., 2002, Figure 1h, © 2002 John Wiley & Sons, Inc.).

Woodruff et al. (2002) suggested the use of hybrid thumbnails, which combine the features of visual summaries with those of plain thumbnails. These "enhanced thumbnails" were created by changing the font size, text and background color of text elements and adding query terms in highlighted callouts to the rendered thumbnail (see Figure 10). The enhanced thumbnails were compared to plain thumbnails and visual summaries in a user study in which the participants engaged in a series of navigational, transactional, and information gathering tasks. The results indicate that the usefulness of each of the presentation methods depended on question category. Use of enhanced thumbnails had the most consistent performance across question categories, which the authors suggest is due to its high information scent, which allows for fast and accurate judgments as to which results are likely to be relevant. The majority of participants were reported to have utilized features present in the enhanced thumbnails when assessing the results, such as relationships between the search terms, their location, and frequency in the documents.

Finally, Teevan et al. (2009) introduced a compact representation called visual snippet that consists of the Web page title, a salient image from the page, and the logo (see Figure 11). The visual snippet approach was compared to other representations (textual snippets and thumbnails) in Web content finding and re-finding across navigational, transactional, and fact-finding tasks. In finding tasks, the representation did not have a significant effect on task completion time, but visual snippets received fewer clicks during tasks than did thumbnails and were also preferred to thumbnails. It was reported also that a number of participants mentioned that the value of various representations was task-dependent, with visual snippets being particularly useful for shopping tasks. The results suggest that participants used different result access strategies for the different representations. Visual representations were faster than text for re-finding previously assessed content, although performance with plain thumbnails suffered if the participant had not been exposed to the thumbnail in the finding phase. The overall findings support the conclusions of Woodruff et al. (2002) in suggesting that a combination of text and images captures the best features of both representation methods. The resulting improvements to visual snippets moved the designs closer to the enhanced thumbnails design space, although the representation is query-independent (i.e., it emphasizes salient text from the page rather than query terms).

Figure 11. Visual snippets, consisting of logo, salient content image, and title, alongside the full Web pages (Teevan et al., 2009, Figure 3, © 2009 Association for Computing Machinery, Inc.).

Previous research outlines two uses for Web page previews: retrieving previously seen content (re-finding) and identifying potentially relevant results. It appears that, unlike query term visualizations, which are useful for removing non-relevant results from consideration, page previews help users identify relevant results (e.g., by recognizing a familiar site logo). From a presentation angle, it appears that, while plain thumbnails are of use for re-visiting, they are not recommendable for finding new content.

Visualization of Classification Structures

Classification structures act as both navigation interfaces and overviews of results, which has been reflected in their presentation in search interfaces. List- and column-type textual layouts have been commonly used for faceted classifications (Kules & Shneiderman, 2008; schraefel et al., 2006; Yee et al., 2003; Zhang & Marchionini, 2005).

One of the key functions of faceted browsing is the iterative search and filtering of the results. For this purpose, many faceted systems use query previews, information about how many results, of what kind, are covered by a facet and its labels. A typical approach is to show this information by using numbers (as seen on many Web sites such as Amazon.com). In addition to showing the number of items contained within a facet, the Flamenco browser (see Figure 12a) allows the perusal of the subcategories of a facet by mouseover.

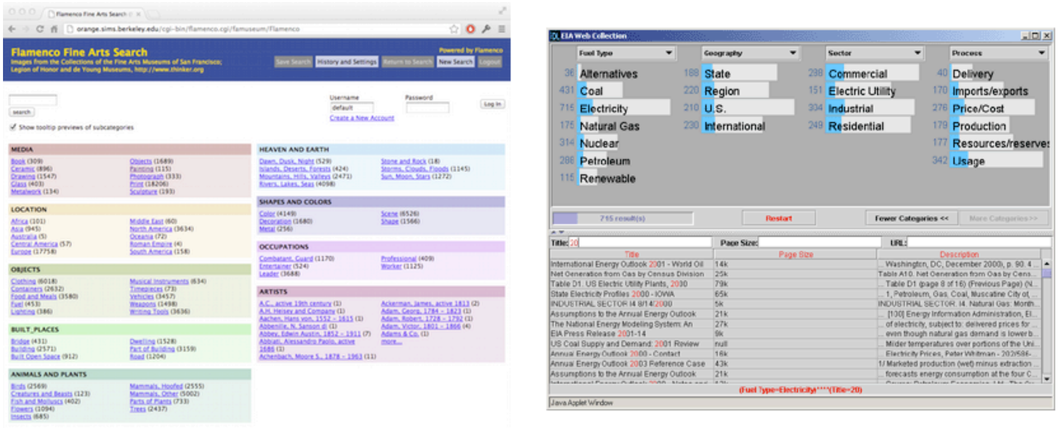


Figure 12. Faceted browsing interfaces: a) Flamenco Fine Arts Search interface displaying first- and second-level facets and b) the Relation Browser++ interface representing the results of filtering operations in the category overview (courtesy of Dr. Gary Marchionini).

Other systems have presented solutions for showing the effects of facet selection, in various ways. The Relation Browser++ (RB++) faceted browsing interface by Zhang and Marchionini (2005) displays the counts of items within facets, using varying-width graphical bars (see Figure 12b). The user is able to explore the effects of facet selection by moving the mouse pointer over a facet. The other bars are proportionally highlighted to show the effects of the selection. In addition, users can dynamically filter the result set by specifying additional keywords, the results of which are immediately reflected in the facet overview and results table. A user study showed that the faceted interface was more effective than a standard keyword search interface for exploratory tasks while showing no effect in simple search tasks. However, the graphical representation was not explicitly compared to faceted browsing without graphical previews. The latest version of the Relation Browser interface also provides an alternative tag-cloud-oriented display called Facet Cloud, wherein the number of items under a facet is mapped to the font size for the label (Capra & Marchionini, 2008). In an early version of the mSpace browser the user was presented with a multimodal preview of documents included within a particular facet item when hovering the mouse over an item (schraefel, Wilson, & Karam, 2003). In the user study of music browsing reported by schraefel et al. (2003), a higher percentage of users (50%) preferred being able to hear early audio preview at every level of facet selection (e.g., composer or arrangement) to an approach where the preview was played upon selecting a particular piece (40%). However, many of the participants who preferred the “late” previews commented that they would have preferred early previews if they could have better controlled the playback of the cues.

Flat lists and visual tree hierarchies of category labels are also typical representation approaches for clustering and hierarchical classification structures in research prototypes (e.g., M. Chen et al., 1999; Dumais et al., 2001; Ferragina & Gulli, 2005; Käki, 2005b) and commercial interfaces, such as the Yippy search engine (see Figure 4). In contrast to the relatively restrained visualizations of facet hierarchies discussed above, clusters and hierarchical category structures have also been visualized via various 2D and 3D techniques that exploit the spatial and organizational aspects of the collection, such as the distance and connections between clusters (e.g., Hsinchun Chen, Houston, Sewell, & Schatz, 1998; Andrews, Sabol, Lackner, Gütl, & Moser, 2001). Below, we summarize some of the key findings from studies that have advocated graphical overviews.

Overall, graphical representations have shown mixed effectiveness. For example, Hsinchun Chen et al. (1998) found that, while their Self-Organizing Map –based approach is suitable for browsing, it does not adequately support searching. Rivadeneira and Bederson (2003) compared the zoomable graphical clustering interface of the now defunct Grokker search engine with two textual clustering representations (Vivísimo's hierarchical clusters and the column-based Grokker Text interface). The findings of their user study indicate no differences between the interfaces in terms of the effectiveness or efficiency of search, and the participants were significantly more satisfied with and preferred Vivísimo's text-based interface. Unlike in many other studies comparing textual and graphical search interfaces, an effect of task on the relative performance was not observed. Finally, Turetken and Sharda (2005) compared the ranked result list to two graphical clustering interfaces based on a treemap layout. The results show that the participants were able to complete tasks sooner with the graphical cluster interfaces, but differences in accuracy or subjective satisfaction were not significant.

Visualizations of classification hierarchies have been suggested for both the Web search domain (Kules & Shneiderman, 2005) and searching of digital libraries (Clarkson, Desai, & Foley, 2009; Kampanya, Shen, Kim, North, & Fox, 2004). Kules and Shneiderman (2005) report the results of an exploratory study that compared a ranked result list interface to two overview interfaces: an expandable hierarchical textual tree overview and a graphical treemap overview (pictured in Figure 13a). The results show higher performance for the overview interfaces in one of the tasks (identifying the government agency with the most pages). The overviews were also subjectively preferred. Most participants expressed preference for the expanding tree view, although several considered the treemap presentation more appealing. Kules and Shneiderman note the importance of user control over the presentation method and categorization method.

A. B.

Figure 13. Treemap-based visualizations of classification hierarchies: a) Treemap search interface displaying the first and second levels of hierarchical result categories (Kules & Shneiderman, 2005, Figure 2) and b) the ResultMap interface showing an overview of a digital library classification with highlighting based on content type (Clarkson et al., 2009, Figure 2, © 2009 IEEE).

In a similar fashion, Clarkson et al. (2009) proposed ResultMaps, a treemap-based visualization that uses existing hierarchical subject classifications to map the documents to parts of the treemap (see Figure 13b). The approach was evaluated in two lab studies. The first compared ResultMaps to a text-only interface (with a tree widget for navigating the hierarchy) in outlier detection tasks. The results did not show a significant difference in task time or accuracy, although the participants rated the text-only interface less favorably in terms of task difficulty and ResultMaps users were more aware of the characteristics of the subject hierarchy. In a follow-up study, the size of the document repository was changed (small vs. large) and the tasks varied from item search to open-ended. The interface did not have a significant effect on task time or query characteristics although ResultMaps users reported significantly greater enjoyment. Clarkson et al. noted that many of the benefits of ResultMaps appear to support analytical interests (matching interest related to the structure and characteristics of the classification metadata) rather than directed needs (related to the documents themselves). Finally, Kampanya et al. (2004) proposed Citiviz, an interface that uses a hyperbolic tree to visualize the relationships between documents (either membership in a classification hierarchy or document cluster), and a 2D scatterplot that arranges the documents by publication date and rank. In the scatter plot, documents are represented via a stacked glyph that visually indicates membership in certain subcategories or clusters. A small usability test was conducted to compare Citiviz to a traditional text-based search interface. As many other studies do, this one showed task-dependent benefits of individual interfaces. For example, when the task is related to finding a publication by its title, the text-based interface was faster because it displays the titles in the search results. Conversely, in searches by topic and publication date, Citiviz was faster, since representing these characteristics forms the basis for its interface design. These findings suggest that graphical overviews, such as those in ResultMaps and Citiviz, might be more suitable for analytical, exploratory search tasks that focus on learning and investigation rather than directed search.

Benefits of Visualization in Search Interfaces

While researchers have frequently suggested graphical visualizations for representing search results, both individual results and classifications, they have not been shown to be generally more effective than textual representations. In discussing the results of their evaluation of different visualization methods, Sutcliffe, Ennis, and Hu (2000) note, ""[W]hile visual user interfaces for information searching might seem to be usable, they may not actually improve performance." However, visualizations may offer advantages in particular types of tasks that draw on the strengths of the graphical representations, and they tend to generate positive feedback from participants.

Summarizing several previous studies, C. Chen and Yu (2000) carried out a meta-analysis of 35 studies of tree and network visualization tools for information retrieval. They concluded that users tend to perform better, in terms of accuracy and efficiency, with simpler visual-spatial interfaces (i.e., visualizations encoding information into structures that exist in dimensional space). Though the meta-analysis was limited to visual-spatial interfaces, the findings also suggest that for search tasks the benefit of complicated visualizations is limited. Julien, Leide, and Bouthillier (2008) carried out a subsequent meta-analysis of 31 controlled user studies of information visualization search tools. They report that visualization tools did not show a significant improvement over text-only interfaces in the eight studies that could be compared on the basis of their quantitative results. However, they note that this may be due to other sources of variability, such as users, tasks, or the measures used, and the diversity of methodological approaches affects the reliability of meta-analytical comparison. Overall effects of visualization aside, advanced visual interfaces may provide benefits with the right combination of task, user and user interface. Sebrechts, Cugini, Laskowski, Vasilakis, and Miller (1999) carried out a comparative evaluation of text, 2D, and 3D interfaces for the visualization of search results. The results showed high interface costs for controlling the 3D visualization, which decreased as the participants gained experience with the prototype. However, the 3D visualization resulted in performance comparable to that of textual and 2D tools in some situations, such as when used by experienced computer users, with input devices suited to controlling 3D views, in tasks that required comparison of document features.

Hearst (2009) argues there are two reasons for the relative lack of utility of visualizations for improving search over text collections. First, assessing search results is primarily a reading activity, and the visual perception activity required by use of visualizations conflicts with this primary task. Second, it is difficult to convert text successfully into meaningful visual analogues. Relatively simple visualization methods seem to be an exception here, either because they incur very little perceptional overhead or because they integrate the visual perception task with the reading activity (as in visualizations of query term location and frequency or in enhanced thumbnails). These techniques can help by pointing out the absence or presence of query terms in the documents or helping people recognize previously visited or potentially relevant results with thumbnails. Hoeber and Yang (2008) succinctly summarize the appeal of simple visualization techniques when discussing their WordBars interface: "[A]lthough the techniques used in this work are rather simple, there is a benefit to the users for this simplicity. The interface is uncluttered, easy to learn, and simple to use."

Alternative interpretations exist for the lack of convincing utility of visualization, however. It is possible that we simply have not yet been able to find the best ways to apply visualization in search interfaces (Hearst, 2009). It also may be that the methods that have been traditionally used for evaluating visualizations are not ideal for this task (G. Ellis and Dix, 2006). Information seeking is fundamentally an activity concerned with satisfying information needs, which are always specific to the user, task, and context, and this should be reflected in the evaluation methodology applied. The issue is examined further in the following sections.

3.4 EVALUATION OF SEARCH USER INTERFACES

Many of the evaluation approaches taken in interactive information retrieval research can be considered to fall on a continuum from system-focused research to human-focused research. Evaluations with end users lie on the human half of the continuum, starting with, closest to the center, studies employing the TREC interactive track approach (Dumais & Belkin, 2005). These studies usually consider the performance, interaction, and usability of the whole system or interface feature. They are followed by more controlled experiments that isolate and study specific aspects of the search process. The next group of studies focuses on studying and understanding naturalistic search behavior, with or without using an experimental search system. Finally, studies at the human extreme look at information needs and behaviors without particular regard for a specific type of information seeking system. (Kelly, 2009)

The exact purpose of the user evaluations may vary, but general distinction can be drawn among exploratory, predictive, formative, and summative evaluation. Exploratory evaluation is focused on how the interface is used and what it is used for. Predictive evaluation, in contrast, produces an estimate of user performance with the interface, based on an analysis of the time and effort required to carry out the various actions that constitute the user's task. Formative evaluation typically occurs within the framework of system development, producing design guidance in the form of problem lists and recommendations for improvement. Finally, summative evaluation is concerned with summarizing the overall effectiveness and impact of the final system. The evaluation activities can either involve experts inspecting the system with reference to various guidelines/checklists or utilize representative end users engaging with the system. Several methods for the individual stages of the process are available for evaluating interfaces with users. (Andrews, 2008)

In her review of search interface research, Hearst (2009) summarizes the types of studies that have been carried out to evaluate search systems, placing them into the following overall categories:

- Standard information retrieval evaluations
- Informal usability testing
- Formal user studies and controlled experiments
- Longitudinal studies
- Analysis of search engine server logs
- Large-scale log-based usability testing

This classification of studies, aside from large-scale testing, is used in the discussion below, which examines the suitability of certain of these methods that have been popular in the evaluation of search interfaces in the literature.

Standard Information Retrieval Evaluations

In classic information retrieval studies, referred to as the Cranfield experiments (see, for example, Cleverdon, 1970), the focus was on evaluating the performance of retrieval algorithms in automated studies without involving users. The work utilized metrics emphasizing retrieved documents' relevance: *precision* (the proportion of relevant documents among the results) and *recall* (the proportion of relevant documents among the items retrieved). The same measurements, along with other performance measures, such as task duration and time to select results, have been applied in studies of interactive Web search to evaluate searcher performance with a search interface (e.g., Hao Chen & Dumais, 2000; Dumais et al., 2001; Käki, 2005b; Zamir & Etzioni, 1999).

The problem with purely objective performance measurements is that they do not accurately reflect the user's search process and the inherent variability in human judgment and evaluation processes. User interactions with the search system are influenced by cognitive, behavioral, and domain factors that are not particularly easy to observe or measure (Kelly, 2009). O'Day and Jeffries (1993) note that the characteristics of successful search interactions are not proxied well by precision, recall, and task performance measures, because search problems are loosely defined and often unique. They suggest that the best indicator of system success appears to be user satisfaction. Subjective feedback on search system use has been gathered in several studies for understanding how well the proposed system meets users' needs. In fact, it is important to consider all three facets of system usability (effectiveness, efficiency, and satisfaction) when evaluating search interfaces, because their correlation tends to be low (Hornbæk & Law, 2007). Therefore, focusing just on performance or the user experience may yield a biased view of the overall usefulness of the system.

Subjective feedback collection has been used in all of the user studies included in this dissertation. In addition to use of interviews and questionnaires focused on the key functionality of the systems, several novel methods have been experimented with. The SUXES user experience assessment method (Turunen et al., 2009) was applied in the study reported on in Paper V to collect expectations and experiences of the interface's use to produce understanding of how the search experience evolved during the study. The product reaction cards approach proposed by Benedek and Miner (2002) was utilized in the studies described in papers II and V both as a quantitative metric and to elicit subjective feedback during the interviews. Finally, the forced-choice paradigm suggested by the present author and colleagues (Heimonen, Aula, Hutchinson, & Granka, 2008) was used in the study reported on in Paper II as a form of implicit feedback alongside explicit preference ratings.

It is possible to make provisions to increase the representativeness of quantitative measurements – e.g., adopting criteria for assessment of the precision of results that is more sensitive to how real users go about choosing which results to view in the search result list. Käki (2004) proposed two such measures, called immediate accuracy and qualified search speed. The former involves the proportion of queries for which at least one relevant result has been selected by the time the user has looked at a specified number of results. The measure is grounded in the reality that users typically consider only a few results per query; hence, immediate accuracy provides an assessment of how well the system is able to support real-world search behavior. Qualified speed combines relevance and time into a new metric, which measures the time taken by the user to select results with a given relevance rating (e.g., relevant vs. non-relevant). The latter measure was used in Paper II's work for studying query term visualizations in a mobile search interface and for Paper III to study a mobile clustering search interface.

An alternative to employing explicit relevance measures is to focus on collecting *implicit* user behavior measurements (Fox, Karnawat, Mydland, Dumais, & White, 2005). These include result-level measures, such as time spent on the page, time to first click, and exit type, and session-level measures, such as query count and number of results visited. Fox et al. (2005) analyzed the relationship between implicit and explicit measures of user satisfaction and found that it is possible to predict user satisfaction through a combination of implicit measures, with result click-through, time, and exit type proving to be the best predictors. That work also explored the use of behavior patterns to predict satisfaction. In addition to using patterns for prediction, Fox et al. suggest that these could be used in an exploratory fashion. The present author utilized a similar interaction pattern approach for Paper V to understand search and interaction behavior with a mobile clustering search interface.

Formal User Studies and Controlled Experiments

Similarly to information retrieval evaluations, formal user studies and controlled experiments are designed to increase the understanding of how search interfaces are used and how different features of the search interface affect the process. In addition to speed, success and satisfaction measurements, the observations gathered during the experiments form the basis for refinement and redesign of the interface, guidelines, and theories (Plaisant, 2004). One of the issues with controlled experiments is the friction between the *internal* and *external* validity of the evaluation setting. Internal validity is related to controlling the experimental setting in such a way that it is possible to state with confidence how the features studied (independent variables) have affected what was measured (dependent variables). External validity refers to how well the experimental setting resembles real usage situations, tasks, and information needs of users. In order to retain internal validity and allow comparison between interface alternatives in formal experiments, the features of the interface and the context of use sometimes have to be artificially constrained. It is possible to aim for a compromise between internal and external validity, by, for example, providing pre-formulated queries, using balanced tasks sets, setting task time limits, pre-caching the results returned by the search engine, and limiting access to the result documents (M. L. Wilson et al., 2010). Another problem with controlled experiments is that it is often difficult to find potential users of the proposed system to participate in experiments, which leads to the use of surrogates, such as students and colleagues (Hearst, 2009). This, coupled with the predefined task scenarios and queries, may adversely affect the external validity of the results obtained from the evaluation. It is possible for the findings not to be necessarily representative of real-world use, or interesting findings can end up being confounded by the experimental constraints, such as task completion time, or ambiguously worded task descriptions. In addition, one is left wondering how changes in the constraints would have affected the outcomes of the study (Plaisant, 2004).

Plaisant (2004) suggests several improvements to specific aspects of empirical studies of information visualization tools that are also applicable to search interface evaluations. Similarly to search interface evaluations that require users to select results relevant to relatively straightforward information needs (e.g., fact-finding), visualization evaluations generally involve simple tasks that may not be representative of real-world needs. Including more varied tasks would improve evaluation. Another problem is in reporting on results across tasks: a tool may in reality work differently for different tasks. Reporting on performance of individual tasks has potential to help highlight how the interfaces benefit various types of tasks. Finally, Plaisant notes that data and task selection is something of an *ad hoc* process, which makes it difficult to compare results across studies. Repositories of data sets and tasks could be useful here.

G. Ellis and Dix (2006) discuss the empirical evaluation of information visualization applications, highlighting concerns that are equally relevant to some forms of search interface, especially exploratory search systems. Their selection of case studies also includes evaluations of search systems. One concern is that sometimes a formal experiment is not necessary for understanding the benefits of the proposed system, or experiments may be asking the wrong questions. For example, G. Ellis and Dix discuss the WaveLens study by Paek et al. (2004) and note that the most relevant insight stemming from it arose from user comments rather than measurement data. In another example, G. Ellis and Dix note that, instead of attempting to validate the efficacy of the Scatter/Gather interface as a browsing interface (Pirolli et al., 1996), the authors might have been more successful had they asked what sorts of tasks the interface is good for. G. Ellis and Dix argue that the problems with empirical evaluation arise from the *generative* nature of visualization interfaces and lack of clarity as to the purpose of the evaluation. With respect to the generative nature, they note that the interfaces are not of value in themselves: they provide value only in a specific context when users use a particular implementation of a design for some purpose. This makes it impractical to validate visualizations via empirical evaluations alone. Validations of ideas should, therefore, consist of justifications originating from existing research, common sense and empirical results, and evaluations that are attempts to verify the questionable aspects of the justification. G. Ellis and Dix call for explorative evaluation approaches, which help researchers see new things about the ideas and concepts being investigated – for example, to understand what kinds of tasks they are good for rather than attempt to force validation of the design through summative or formative evaluations.

Longitudinal Studies

Given the many challenges in organizing controlled experiments and their inability to address naturalistic use, there has been a trend recently to move towards more ecologically valid longitudinal evaluation procedures when evaluating information-centric tools. In these studies, participants use the system as a part of their daily information work, over a longer time, such as weeks or months (e.g., Hoeber, Schroeder, & Brooks, 2009; Käki, 2005b; Shneiderman & Plaisant, 2006; M. L. Wilson & schraefel, 2008). Kellar et al. (2007) note that "the primary strength of field studies is the increase in realism, as participants are observed working in their own environment(s), with their own tools (e.g., bookmarks, history, choice of browser), and completing tasks that are motivated by the participant and not the researcher." According to Kules (2006), longitudinal study protocols that integrate ethnographic methods, such as observation and interviews, with quantitative ones (e.g., surveys and usage logs) may be beneficial for studying search interfaces. Most importantly, the results from field studies can help in evaluating the changes in search strategies and subjective preferences (M. L. Wilson et al., 2010). Such approaches closely resemble the exploratory evaluation paradigm suggested by G. Ellis and Dix (2006). The user studies reported upon in Paper V and Paper VI were undertaken in this fashion to build a richer understanding of the mobile Web information access phenomena under investigation.

The limiting factor in the use of longitudinal studies is the relatively large resource investment required. If realistic user experiences are to be gathered, the participants have to be able to use the system productively; i.e., it needs to function well enough to help its users accomplish their tasks (Shneiderman & Plaisant, 2006). This can be a significant challenge, with negative effects on the results of the study if the participants do not find the prototype reliable enough for everyday use. Often, significant implementation work is required if one is to provide the participants with a working environment. For example, in the longitudinal study of a category-based mobile search interface (Paper V), the system was functioning to the extent that participants could use it for practical information access through the underlying commercial search engines. Developing the system to this level of stability, even with the access to an existing search middleware platform, took several months.

In addition, the unavoidable lack of refinement of the prototypes and the learning curve often associated with novel user interface solutions may affect the participants' motivation and retention. Participants might resort to the existing services they are adept at using when faced with situations wherein their information need has to be met rapidly or cannot be easily addressed with the prototype. In the longitudinal study reported on in Paper V, I utilized several forms of user feedback collection in an attempt to identify these situations. The participants were also asked to report any technical issues they encountered during use. In the final interview, they were presented with questions focusing on the experienced level of satisfaction with the prototype. These addressed frustrations caused by service breaks or other issues the users encountered during the study, the degree to which they were able to satisfy their information needs with the prototype, and the extent to which they utilized the prototype to address their information needs (as opposed to using other tools, such as their favorite commercial search engine).

While it is necessary to recruit representative users if the system is to be evaluated realistically, it is also needful to understand their motivations, information needs, and usage practices beyond what we can infer from the usage logs. Diary studies are a promising method of data collection for the purpose of studying mobile information needs (e.g., Church & Smyth; 2009; Hinze et al., 2010; Sohn et al., 2008; see also Paper VI) and mobile information access behavior (Amin, Townsend, Ossenbruggen, & Hardman, 2009; Nylander, Lundquist, Brännström, & Karlson, 2009). However, even when studies examine information access practices, challenges remain in capturing salient, timely information from the users without unduly burdening them with data entry. In the study reported upon in Paper VI, a simple mobile-device-friendly Web form was utilized in an attempt to minimize completion effort, but this approach was only partially successful. The number of entries submitted per participant varied greatly, as did the level of detail of the entries. It becomes clear that several interrelated factors have to be considered when one is conducting diary studies. These include prompting frequency, the threshold set for the smallest reportable incident, and the level of detail required (Brandt, Weiss, & Klemmer, 2007).

Search Log Analysis

Search engines, and Web server software systems in general, record information about searches and interaction via the search interface in log files. In addition to improving the search algorithms and providing data for personalization features (e.g., query history and query recommendations), these log files are a rich data source for analysis aimed at improving the search interface (Hearst, 2009). Search log studies have been used extensively for understanding of both the search behavior and query topics (e.g., Silverstein, Henzinger, Marais, & Moricz, 1999; Spink, Wolfram, Jansen, & Saracevic, 2001) and the search goals and intent of users (Broder, 2002; Jansen et al., 2007, 2008; Jansen & Booth, 2010; Rose & Levinson, 2004).

Kamvar and Baluja (2007a) note that one of the weaknesses of search log analysis as a research method is that search logs do not really tell us the full story behind the search experience, such as the contextual factors that inspired the search. Bar-Ilan (2007) suggests that the best approach for understanding how people search the Web is to combine search log analysis with qualitative studies and surveys. However, even in the latter use, search logs present some practical and ethical concerns for academic research. Academic researchers who are not actively collaborating with search engine companies rarely have access to search logs on a large scale, and it is not clear how such researchers could gain access to them (Bar-Ilan, 2007). When such logs have been made available, their use has been problematic in terms of privacy concerns (Hafner, 2006). Bar-Ilan suggests the use of institutional review boards, analogous to the oversight bodies in medical and social sciences research, to ensure that users' privacy and rights are respected. Regardless of such concerns, search log analysis remains the method of opportunity for industry practitioners.

3.5 Summary

The design of search interfaces focuses on understanding and facilitation of searchers' access to the information they need for satisfying their information needs. Previous research has targeted all stages of the users' search process, from query formulation to the evaluation of the results. The emphasis of this dissertation is on supporting result evaluation, which has two key stages: evaluating the success of the query in light of the results returned and assessing the relevance of individual results. Previous research has shown that the evaluation of the entire result set can be assisted with result categorization. Judgment of individual results can be successfully supported not only by extracting salient content from the result documents but also by enhancing this content with simple visualizations.

Unlike text-based category overviews and simple query-biased visualizations, complex visual search interfaces have been found lacking. On one hand, this may be due to the task they are attempting to support – many general Web search goals do not require the use of advanced filtering, overview, or sorting functions. On the other hand, an issue may be found in how we evaluate search interactions. Longitudinal mixed-method studies may be more successful in uncovering the benefits of alternative search result visualizations and the extent of their applicability in naturalistic settings.

4 Mobile Information Access

Existing frameworks for information seeking recognize the importance of context. For example, Järvelin and Ingwersen (2004) various contexts affecting the information seeking process. Further, Foster (2004) considers the effects of external context, internal context, and the user's cognitive approach to information seeking. Accordingly, context information, such as the text of the page the user is reading, has been utilized, for example, for query augmentation, leading to greater search success (Finkelstein et al., 2002; Kraft, Maghoul, & Chang, 2005).

Mobility is an additional source of contextual information and a context of use in itself. This needs to be accounted for in the design and evaluation of mobile search interfaces. Mobility changes the dynamics of internal and external contexts in several ways. The user's internal context (e.g., tasks and goals) is different in mobile situations, and the external factors, including social situations, are dynamic and unpredictable, giving rise to rich variation in users' actions (Tamminen, Oulasvirta, Toiskallio, & Kankainen, 2004). For example, elements of the technical infrastructure, such as network coverage, may change as the user moves from one location to another; the availability of mobile services may differ; physical context (e.g., illumination, temperature, and weather) can affect the interaction; and the use of the mobile device can affect the social situation, and vice versa (Kaasinen, 2003).

Dey (2001) provides the following general definition for context: "Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves." In the discussion that follows, the examination of context focuses primarily on research that seeks to understand the situations giving rise to mobile information needs, and how people resolve their information needs in mobile situations by means of mobile Internet access, particularly by using search tools.

4.1 MOBILE INFORMATION NEEDS

Several researchers have carried out user-centered examinations of mobile information needs, and the methods of and contexts in which mobile information behavior takes place. Studying information needs, activities, and contexts is particularly critical for mobile information access because mobile users need applications tailored to the requirements of the mobile context of use (Sohn et al., 2008). As an example, mobile users need to divide their attentional resources between mobility tasks (e.g., attending to the environment, processing sensor data, and handling other cognitive calculations) and the use of the mobile application or service (Oulasvirta, Tamminen, Roto, & Kuorelahti, 2005). The relative priority of these tasks influences how the cognitive resources are allocated (Oulasvirta et al., 2005), which, in turn, affects whether and how people decide to access information services to resolve their information needs (Sohn et al., 2008). It is argued that mobile information access applications and services that are sensitive to the users' context could thus better support mobile users.

The sections below summarize various studies of mobile information needs in terms of several dimensions. One of the challenges in summarizing research into mobile information needs is the variability in the research questions and the organization of the findings, since authors vary in their classifications of topics and intents. One typical approach has been to identify the topics and intentions behind the information needs and the locations in which they emerge. Another approach is to consider the effect of context on the type of information need and access method. Finally, some studies have investigated if and when the information needs are addressed. This division is used in the following discussion.

Topics and Locations of Information Needs

Kaasinen (2003) studied user needs for location-aware mobile services in a series of user interviews, laboratory studies, and field evaluations. In terms of the type of content desired, the users identified dynamically updated information (e.g., weather forecasts and traffic information) as important, especially since such information was not widely available on the Web at the time of the study. Such information sources are now commonplace and geographical information needs identified in virtually all of the recent studies focusing on mobile information needs.

Sohn et al. (2008) carried out a two-week diary study of mobile information needs while users were on the go. Their results include classification of the topics, of which the most frequent were trivia (19%), directions (13%), and points of interest (12%). Trivia needs are particularly interesting, because they were typically prompted by the social situation or location-based cues, covering a wide spectrum of information needs, from fact-finding with lookup of specific information (e.g., the date and cause of a celebrity's death) to more open knowledge discovery (e.g., health benefits of a specific foodstuff).

Other studies have considered the spread of information needs across mobile and non-mobile situations. This author carried out a four-week diary study of the information needs and access strategies of active mobile Internet users (Paper VI). About 33% of the information needs occurred when the participants were mobile (on the move, incl. utilizing public transportation), and about 35% arose in the home. The most frequent subjects of information needs were trivia (27%); information related to work, studies, and interests (16%); and public transportation (12%).

Dearman et al. (2008) conducted a four-week diary study of everyday information needs and sharing opportunities. They considered all location contexts (home, work, and mobile). Here too, only about 33% of the needs occurred when the participants were mobile, with home being the most common non-mobile location (43%). The most common information needs across all location contexts included finding (35%), availability (22%), and guidance (11%). Finding covered needs that were related to identifying an entity (e.g., a person, event, or establishment) and locating it. Availability needs arose in relation to both scheduled availability (e.g., of a person or on opening hours) and circumstantial availability of people and services. Guidance was related to the knowledge required to perform an action or inform a decision. Although most of the needs were related to a task, about 14% of needs had to do with acquiring knowledge or satisfying curiosity independent of a task.

Church and Smyth (2009) carried out a four-week diary study involving mobile information needs, focusing specifically on the intent behind the needs. The majority of needs (67%) arose when the user was mobile (e.g., away from the desk, commuting, or traveling abroad), with about 34% occurring when the user was specifically on the move and about a third arising at home or at work. The most frequent topics were local services (24%), travel and commuting (20%), general information (16%), and entertainment (13%). The intent behind the information needs was classified as informational (58%), finding information about a topic; geographical (31%), finding an answer to a question that is dependent on location (e.g., the nearest café); or personal information management (11%). Most topics were associated with a single intent (e.g., trivia and general information needs were all informational), while some covered several intents (e.g., entertainment, travel/commuting, and local information).

Finally, Hinze et al. (2010) carried out a diary study over one week, focusing on the relationship between information needs and their location. About one third of the information needs occurred when the participants were mobile (going out or in a car). Two main types of information needs were identified: problem-solving and geographical. Problem-solving needs involved finding published information, facts, and advice to further a task or activity. Geographical needs were focused on getting directions to or an estimate of the time needed to reach a particular place, quick ways to get to a location, and information about nearby places.

Effects of Context on Information Needs

Several studies have shown that there is a strong influence of context on the emergence of information needs. For example, in one of the earliest studies of mobile information needs, Kaasinen (2003) discovered that information needs and mobile service needs varied not only by location but also with the user and the usage situation. For example, when looking for accommodation information via a mobile application, some users would have preferred information about room availability and price, while others desired information about the location, and still others would have wished to find out more about the quality and facilities of the hotel.

In more recent diary studies the effects of context have been investigated in greater detail. Context influences mobile information needs in some way in the majority of situations. Sohn et al. (2008) found that context was highly influential as the trigger for information needs expressed when on the move (72% of diary entries). Hinze et al. (2010) and the present author (Paper VI) found similar results when examining information needs in both mobile and non-mobile settings, with context influencing 78% and 67% of needs, respectively. Location, both the user's current location and his or her destination, is one of the most influential factors, alone and in combination with other factors. Sohn et al. report location as the most frequent contextual trigger for information needs (35% of diary entries). Similarly, Church and Smyth (2009) found that about 30% of reported information needs were location-dependent (related to the user's location either explicitly or implicitly).

Hinze et al. (2010) report that the questions participants needed answers to appeared to differ with location. At home the focus was on problem solving advice (e.g., "what is this" or "how to do this"), whereas in mobile situations geographical questions predominated. The findings suggest a close link between moving and location-specific questions. Another important category is traveling outside the user's familiar surroundings. Church and Smyth (2009) report that about 13% of users' information needs occurred during travel abroad and 20% of needs focused on travel and commuting information, with participants who traveled away from their place of residence requesting information that helped them navigate and familiarize themselves with the area they were visiting. L. Chen and Qi (2010) investigated information needs emerging during leisure travel. Unsurprisingly, geographical intents influenced the majority of information needs (61%), and most of the topics were related to information that supported the travel-related activity (e.g., restaurants, sightseeing, shopping, hotels, and transportation). Correspondingly, when one is at home, at work, in school, or in some other social situation, the information needs cover a wider spectrum and are influenced by factors such as social interactions. For example, Church and Smyth report greater prevalence of information and personal information management needs in non-mobile situations than geographical needs.

In addition to location, the influence of other factors has been identified. Sohn et al. (2008) also found time, conversations, and activities to be contextual factors nearly as influential as location. This author (Paper VI) found that activities (23%), social situations (16%), and time (17%) were stronger contributing factors to the emergence of information needs than location was. In contrast to findings of previous studies that highlight the importance of geographical needs, few of the on-the-move information needs were related to the immediate surroundings. Rather, the location was only relevant as the starting point or frame of reference. For example, this was the case for public transportation timetable lookups and route guidance queries. In addition, the present author found that mobile information needs were both activity- and time-dependent. For example, public transportation users had information needs related to the route they were taking or the activity that would follow (e.g., the time and place of a meeting). Church and Smyth (2009) found that most geographical information needs were temporally dependent even if the diary entries did not refer to explicit temporal cues. Finally, social aspects of information needs are clearly important. The present author (Paper VI) found that information needs emerging in social situations are typically tied in with the person's engagement in the social situation and are likely to be triggered by either discussion or artifacts in the immediate environment. Similarly, Church and Smyth (2009) and Sohn et al. (2008) describe how information needs arise as a result of conversations.

These findings have implications for the design of tools that support mobile information seeking. Sohn et al. (2008) suggest that mobile technology should better account for the user's current task and its effects on recording information needs and utilizing the results. Several researchers highlight the need to account for variety of contextual factors in the design of mobile information access interfaces (Amin et al., 2009; Church and Smyth, 2009; Hinze et al., 2010; Teevan, Karlson, Amini, Bernheim Brush, & Krumm, 2011). For example, Hinze et al. (2010) judged about 60% of the queries expressed by their participants to require some contextual information to be answered, with roughly 20% requiring additional information from two or more contexts. The present author (Paper VI) has suggested that mobile technology could proactively seek information in view of context information (e.g., discussion), allowing users to focus on their primary activity.

Addressing the Information Needs

Addressing the information needs that emerge in everyday situations is the other side of the equation when one is considering mobile information access. Various studies have attempted to identify the strategies and rationale people exercise in satisfying their information needs. Sohn et al. (2008) identified factors that affecting their participants' decisions about reacting to the information needs, including the importance of addressing the need, its urgency, perceived cost, and the situational context. These factors interacted in the participants' cost–benefit assessments, influencing whether the information need was to be addressed right away or later, as well as which method was utilized. In about 45% of cases, the information need was addressed at the time it arose. In their study, the Internet was used 40% of the time; however, calling an information source (23%) or calling someone to access the information on the person's behalf (16%) were also frequent strategies. It should be noted that only five of the 18 participants in the study had mobile Internet access on their phones, and the lack of mobile Internet access was the top reason (32%) for addressing a need later and the second most important reason for not addressing the information need at all (23%). Dearman et al. (2008) also report that people were less able to address their information needs when mobile than when at work or at home, although a larger proportion of the mobile information needs were time-critical when the participants were mobile. When information needs were addressed immediately, the Web was used more frequently at home (33%) and at work (49%) than when on the go (18%). The low incidence of Web use when one was mobile is likely partially explained by the fact that none of the participants had Internet access on their mobile devices. In fact, the participants commented that it would have been useful for addressing information needs.

This potential for usefulness of mobile Internet access to address information needs has been confirmed in other studies. For example, Sohn et al. (2008) found that mobile Internet users utilized this option in 73% of the cases to address their information needs. The results indicate that ability to address one's information needs without having to resort to outside assistance, and rapid turnaround in using mobile services, increased the appeal of mobile Internet access. The active mobile Internet users in the present author's study (Paper VI) attempted to address virtually all of their information needs (145 out of 147) at the time they emerged, utilizing mobile Internet access and mobile applications in 94% of these cases. The smartphone users in the study by Matthews, Pierce and Tang (2009) reported appreciation for the usefulness of mobile Internet access for finding information, especially in situations wherein the participants needed contextually relevant information (e.g., information about places or answers to questions arising in a social situation). The monetary cost of mobile Internet use appears not to be a major concern for people, presumably because many users have a fixed-rate mobile broadband subscription or have access to wireless Internet connections at school, libraries, or work (Chua, Balkunje, & Goh, 2011; Church & Oliver, 2011; Cui & Roto, 2008; Paper VI). When a fixed-rate subscription is not available, people may restrict their Internet access to addressing urgent needs, out of cost concerns (Chua et al., 2011).

A recent survey by Kaikkonen (2011) confirms the importance of mobile Internet access and a good browsing experience. The results show that 92% of respondents reported it as important to have a phone with a good mobile browser. Some mobile Internet activities are even reported as being more common on cellular phones than on desktop computers (e.g., information search, reading e-mail and news, downloading applications, finding contact information, and using maps). The findings also show that mobile Internet use is increasing; the frequency of every mobile Internet activity studied grew between 2007 and 2011, several of them significantly.

4.2 Mobile Internet Use for Information Access

A key implication of the studies that were aimed at understanding the emergence and types of mobile information needs is that, although not entirely unproblematic, mobile Internet use is an important tool for addressing users' emerging mobile information needs. However, mobile Internet access differs from desktop Internet access in terms of three important factors: the user, the environment, and the technology. Chae and Kim (2003) note that, firstly, mobile Internet devices are personal and rarely shared with other users as computers are. Second, Internet-enabled phones provide for an "always-on" connection to the Internet that enables information access anywhere and at any time. Third, mobile devices are still limited in technological terms when compared to computers. The screens are smaller, and the input and output features are more limited, as is the processing power (although this gap is rapidly diminishing).

These individual characteristics map to different aspects of mobile Internet use. The personal nature of mobile devices may affect the type of content people utilize. Technological handicaps can affect how well people are able to access information via mobile Internet access and require design solutions that take these limitations into account. But perhaps most importantly, the widespread availability of mobile Internet use, regardless of the situation and activity, has transformed how people attempt to access information.

The following discussion summarizes research on three related strands of inquiry examining mobile Internet use: characterizations of mobile Internet activities, the effect of context on mobile Internet use, and search versus browsing.

Types of Mobile Internet Activities

Several studies have focused on use of the Internet on mobile devices, with attempts to characterize the types of activities and the motivations for this Internet use. Taylor et al. (2008) interviewed early mobile Internet adopters in order to understand the motivations and behaviors associated with mobile Internet use. Motivations were characterized as utilitarian (awareness and time management) or hedonic (satisfaction of curiosity, diversion, and social connection and social avoidance). The behaviors were classified into information seeking (status checking for dynamic information, browsing, information gathering, and fact checking), action support (in-the-moment actions and planning), and information exchange (transactions and communication). These breakdowns match the general mobile phone activities reported by Matthews et al. (2009), who identified general types of interactions people engage in with their smartphones: in decreasing order of prevalence, contextual information seeking (which was also identified by many participants as qualitatively the most valuable activity), accessing entertainment or infotainment, completing concrete tasks, maintaining social ties, and maintaining awareness of information. In terms of patterns of motivation and mobile Internet behaviors, Taylor et al. (2008) report that information seeking behaviors were employed primarily to satisfy curiosity (e.g., interest in unfamiliar topics), while various status check behaviors (e.g., checking e-mail or Facebook updates) and (undirected) browsing were used as diversions. Church and Oliver (2011) report similar results with respect to the use of different tools to satisfy different motivations. For example, curiosity was often satisfied via a mobile search engine or searching within a mobile application. The present author (Paper VI) found a similar connection between intents and the choice of access method. Informational, often hedonic needs were approached via Web search, and focused, pragmatic needs were satisfied by means of a known Web site or an application likely to provide the desired information.

Nylander et al. (2009) studied mobile Internet usage motivations with a diary study. The most common activities in their study were reading news (20%), passing time (19%), checking e-mail (17%), situated information search (16%), and general information search (15%). Interestingly, the top three motivations were not particularly connected with mobility or the situation at the time, which runs counter to the findings from other studies examining information needs, which suggest that most information needs are influenced by context. One possible explanation is that activities such as browsing the latest news headlines and checking e-mail are not typically classified as "serious" information needs by the strictest definition. Another interpretation is that such diversionary activities arise from a *lack of context* – bridging the free moments between activities that people experience in the course of the day (Matthews et al., 2009). Nylander et al. note that situated information search was focused on finding information about the current situation and activity, reflecting information needs arising from the users' context. However, general information searches without a specific connection to the location or activity were nearly as prevalent.

Finally, Cui and Roto (2008) studied use of the Web on mobile devices by using contextual inquiries and phone usage log analysis. They identified four general categories of mobile Web activities: information seeking, communication, transactions, and personal space extension. Information seeking activities followed themes similar to those reported by Matthews et al. (2009), including fact-finding (especially from Wikipedia), information gathering, and casual browsing. Information gathering was a relatively rare task, occurring only when demanded by user goals and supported by the user's context (e.g., to support physical transactions, or evolving from fact-finding sessions). The authors note the challenges mobile Web users experience with information gathering tasks: they must internally process and retain the information, because most current mobile tools do not sufficiently support strategies such as use of multiple windows or tabs, which are commonly used in desktop environments (Aula & Käki, 2003). Casual browsing was identified as a relatively common task, especially when users had a fixed-rate plan.

It is clear that, although useful in many situations, mobile Internet access is not a panacea for resolving mobile information needs. Participants in many studies have reported difficulties with finding information by means of their mobile device. For example, the present author (Paper VI) found that in 20% of cases, the participants were unable to locate the information they required or were able to satisfy the need only partially. Similarly, Sohn et al. (2008) report that the mobile Internet users felt mobile Internet access to be inadequate for addressing their needs. The analysis by Hinze et al. (2010) offers some insights into the reasons. For the questions posed in their study, 68% were likely to have answers available on the Internet (including queries requiring enhancement with context information) and 31% in digital libraries. However, 29% of the questions were deemed not to be answerable by existing online services, since they targeted personal assistance in decision-making.

Effects of Context on Mobile Internet Use

Several studies have found that mobile information needs emerge across several physical locations (e.g., Church & Smyth, 2009; Dearman et al., 2008; Paper VI). In response to information needs, the mobile Internet is utilized correspondingly, to address needs in various locations (Cui & Roto, 2008; Nylander et al., 2009; Paper VI). The present author (Paper VI) has identified common themes in diary entries that reflected the participants' use of mobile technology in familiar environments. In most cases, use was due to proximity and convenience – the mobile phone was instantly available for fulfilling the information need. In some cases, the mobile phone was a natural means of information access (e.g., checking for missed calls or looking up the weather forecast while in bed when waking up). Several other researchers (e.g., Amin et al., 2009; Church & Oliver, 2011; Cui & Roto, 2008; Nylander et al., 2009; Taylor et al., 2008) note also that people often used their phone even when computers were available, because of comfort and convenience. For example, Nylander et al. (2009) found that in one third of cases, the phone was chosen over the computer out of convenience and laziness (i.e., not wanting to disrupt one's current activity to utilize the computer). They also note how mobile Internet access provided people with a new type of micro-level mobility – being able to combine moving around the home environment and carrying out various activities without being tied to a specific place for using the Internet.

With respect to other locations where mobile Internet use is common, Nylander et al. (2009) found that outside the home (31%), the most frequent locations were outdoors (23%), in transit situations (23%), and other indoor locations (16%). Location also had an effect on the type of Internet activity. At home, the most common activities were reading news and accessing e-mail, whereas outdoor mobile Internet use was applied most often for situated information search (which is consistent with the findings from information needs studies that highlight the importance of contextually emergent questions). In transit situations (e.g., on a bus, subway, or train), the most commonplace activities included the same awareness activities as in home settings, but another notable category here is passing time. When participants were indoors in locations other than home or work, mobile Internet use was motivated by awareness activities such as checking e-mail and passing the time. The results of Church and Oliver (2011) confirm the importance of stationary situations. Over 70% of the diary entries in their study were related to familiar contexts such as the home or work.

Other contexts too have been found to be influential in triggering mobile Internet use. Cui and Roto (2008) report that social mobile Web use occurred quite often in their data set, acting as a conversation enhancer. Mobile Web access was used to start discussions on new topics, expand an ongoing discussion, or settle disputes. Church and Oliver (2011) also note that in about 65% of the cases, mobile Internet use was a social activity. Congruent with the findings from other information needs studies is Church and Oliver's report that the urgency of mobile Web access increases when the users are mobile as compared to being at home, with search used especially to satisfy more immediate information needs.

One of the challenges facing Internet use in the mobile context is how to integrate device use with the primary mobility task, such as navigation. Matthews et al. (2009) divide the strategies people used in their study into making time for application use and filling time between tasks. Their findings show a clear division between activities: information seeking and accomplishing tasks would interrupt the current activity, whereas accessing entertainment content and social networking were carried out as time-filling. Counter-pressures influence the time spent on using the phone before subjects returned their attention to other activities (Matthews et al., 2009). Counter-pressures arise out of the contextual activities the user engages in and are either internal (phone use interrupts a social interaction) or external (phone use is forbidden). Matthews et al. note that if the desire or need for information is strong enough, people make time amidst other activities and then resist the counter-pressures until either the information is found or the counter-pressure grows to be significant enough to terminate the attempt (e.g., if one cannot find information within a certain time).

Browsing vs. Searching in the Mobile Internet Context

Given the amount of research on mobile Internet use, it is surprising that relatively few studies address the relationship between browsing for content on the Web and searching for it. Although many of the diary studies allude to the use of search engines for satisfying information needs, it is somewhat unclear to what extent these various methods are used and in which situations. Several studies do point to the importance of mobile search as having steadily increased over the years, as evidenced in the findings of large-scale search log studies and user surveys.

The results from the large-scale study by Church, Smyth, Cotter, and Bradley (2007) show that the majority (94%) of mobile sessions utilizing the Web are focused on browsing, containing no search-related activity. While the proportion of search sessions was small, searching contributed significantly to mobile Internet usage. For example, duration, the quantity of data transferred, and the number of content requests all were higher for sessions that featured search activities. There are several possible explanations for these differences. First, search sessions often also include browsing whereby people locate and access relevant content. Second, searching is a more challenging information access task than browsing, since the user needs to come up with query terms rather than navigate more familiar information structures. Another likely interpretation is that search users are early adopters who have information needs they feel are best addressed by search rather than browsing (e.g., with use of operator portals). Subsequent studies have shown increased search activity. Church et al. (2008) report that 8–10% of mobile Internet users engaged in search activities. Survey studies have reported higher percentages, perhaps due to the magnitude of difference in sampling, and likely represent uptake among early-adopter demographics. For example Kaikkonen (2011) found that 58% of respondents reported using search engines to search for information, up from 39% in 2007. A Nielsen Mobile white paper (Nielsen, 2008) reports that 40% of mobile Internet users said that they access the Web via search engines.

Some studies have investigated the frequency of use of various mobile Internet tools with diary study approaches (Church & Oliver; 2011; Paper VI). The participants in the present author's study (Paper VI) reported using Web search in 60%, Web browsing in 31%, and mobile applications in 3% of entries. The results also revealed decision-making strategies for addressing the information needs. The participants used multiple methods to address urgent needs, such as consulting offline resources or colleagues at work if they could not find the information themselves from the Web. Important but non-time-sensitive needs would be addressed later if the answer could not be found immediately by means of mobile Web search. Church and Oliver (2011) report a much lower frequency of Web search (10%) and browsing (15%), with an increased role of social tools and applications (51%), such as Twitter, Facebook, games, streaming music, and Internet-enabled native applications. One reason for the difference in the figures might lie in the slightly differing research questions. The present author's focus was on information needs and how they are addressed, whereas Church and Oliver collected data on mobile Internet use in general.

4.3 Web Search as a Mobile Information Access Method

The following sections summarize the findings from large-scale mobile search studies along the key dimensions: search behavior (queries and result click-through), search topics and goals, and effects of contextual factors (e.g., culture, region, and search interface).

Mobile Search Behavior

In the earliest large-scale examination of mobile Web search behavior, Kamvar and Baluja (2006) studied one million search queries to determine how people search the Web on mobile phones and personal digital assistants (PDAs). The length of queries was similar to that seen in desktop use (between two and three words per query), which the authors note may suggest that people optimize the number of query terms to gain a suitable number of results for action, regardless of the medium. Mobile search queries had significantly less variation than desktop queries did, suggesting that people target sites that are usable with the mobile browser. Click-through rates were reported as being low, with fewer than 10% of queries receiving at least one result selection, and just 1.7 result selections being made per query, on average. Kamvar and Baluja suggest that this may indicate mobile users' reliance on the result summaries to make their relevance judgments. The results also show that mobile searchers do not explore the results actively, since in only about 9% of queries do the users go beyond the first result page. Mobile searchers remained focused on their initial search topic; about 75% of consecutive queries were related to one another. The results of Kamvar and Baluja suggest that results fulfilling the users' information needs can readily remain unseen simply because of ambiguous queries that do not produce relevant results within the first result page, leading to lower click-through. The lack of result exploration can be explained by the cost of interactions being higher, in relative terms, in the mobile environment, with slow loading and the overhead of browsing through result pages.

A subsequent study (Kamvar & Baluja, 2007a) of one million queries indicates that developments in mobile technology have likely improved result exploration to some extent. The query click-through rate rose to over 50%, although result-page exploration remained low (more result pages were requested in only 10% of queries). Querying, however, became more varied (with more queries per session and more diverse queries) and the query content less homogenous. A third study (Kamvar, Kellar, Patel, & Xu, 2009) further underlined the differences between mobile and desktop search. On mobile devices, search queries are still less diverse both in terms of the queries and in the user information needs, and they are more focused on local information. Users of mobile phones also make fewer queries per session, which the authors suggest is due to different foci of information needs with the different devices and difficulties with text entry. Mobile phone users are, therefore, likely to search quickly for factual information and browse multiple results rather than make query refinements.

Another early large-scale study of mobile search by Church et al. (2007) analyzed over 400,000 queries, from 50,000 users. The results for mobile Web search behavior confirm the earlier findings by Kamvar and Baluja (2006) with respect to the high number of repeat queries and lack of variation in query vocabulary, but the study also found that mobile queries use less advanced features, which may also be a result of the limited text-input capabilities of the phones of the era. A subsequent study by Church et al. (2008), which considered six million search requests, made by 260,000 users, highlights significant problems with mobile-search user experience: almost 90% of queries and nearly 60% of search sessions did not lead to any result selections by the user. Church and colleagues note that, although it may be possible that in some cases the content of the result snippets may satisfy information needs, it is unlikely that the majority of mobile searches can be answered without accessing of the results. Therefore, it is more likely that in many cases the conventional approach to mobile search adopted by search engines results in users failing to find relevant information with the result lists.

Search Topics and Goals

Comparison of search topics is challenging across studies, similarly to that of information needs, because of differences between the classifications that individual authors have used. However, it is clear that entertainment is a significant driver of mobile search, especially in the realm of adult content. The most popular query categories across the various log-based analyses that provide comparable data (Church et al., 2007; Church et al., 2008; Kamvar & Baluja, 2006; Kamvar & Baluja, 2007a; Yi, Maghoul, & Pedersen, 2008) are collected in Table 3.

Table 3. The top search query categories in log analysis studies.

Kamvar & Baluja (2006)	Kamvar & Baluja (2007a)	Church et al. (2007)	Church et al. (2008)	Yi et al. (2008)
Adult (> 20%)	Adult (> 25%)	Adult (53%)	Adult (61%)	Entertainment (42%)
Entertainment (> 10%)	Entertainment (> 10%)	Multimedia (10%)	Email, messaging & chat (9%)	Technology (5%)
Internet & telecoms (> 5%)	Internet & telecoms (> 4%)	Email, messaging & chat (8%)	Search & finding things (7%)	People (3%)
Local services (> 5%)	Lifestyle/ online communities (> 4%)	Search & finding things (8%)	Entertainment (5%)	Retail (3%)
Games (> 2%)	Local (> 4%)	Entertainment (8%)	Multimedia (5%)	Travel (3%)

Although the classifications differ in their labeling of what falls within each topic area, it is clear that mobile search is dominated by niche interests (Church et al., 2008), with most queries targeting adult and entertainment content, and media consumption. Researchers note, however, that this may be a sign of the nascent state of mobile search and that the diversity of searches may increase as the user base grows, following the trends seen in desktop Web search (Church et al., 2008; Yi et al., 2008). Indeed, the results of Kamvar et al. (2009) clearly indicate that search on high-end devices, such as the Apple iPhone, resembles desktop search to a remarkable extent in terms of the search topics.

In addition to identifying specific topics, it is of interest to identify the search intent behind the queries. Church et al. (2008) classified queries according to their intent (using the taxonomy by Broder (2002)), with 10% of queries being informational, 29% navigational, and 60% transactional. The results highlight an interesting departure from desktop search (e.g., Jansen et al., 2007), in which informational queries dominate. One likely explanation is the high incidence of adult queries in the mobile data set (61%), which were classified as transactional. The large number of navigational queries both in the data set of Church et al. (2008) and in that of Kamvar and Baluja (2006) suggests, on the other hand, the use of search as a replacement for bookmarking, with search used, in practice, to initiate browsing sessions on the desired Web sites.

Although diary studies have shown that location-dependent information needs are frequent and that people frequently carry out local searches, it does not consistently stand out in the query topic classifications as a distinct category. Yi et al. (2008) offer an explanation that may also apply to other studies, noting that, although local search intent did not explicitly appear in their categorization of search topics, it nevertheless was present in about 9–10% of queries. Church and Oliver (2011) report higher figures for local search intents (about 30%) in their diary study; however, drawing parallels between a large-scale log study and a diary study is not feasible.

Effects of Context on Web Search Behavior

As noted earlier, both internal and external contexts can affect search behavior. The role of cultural differences, the mobile devices used for search, and the contexts of use have been studied to some extent in previous research.

While most of the studies have targeted English-speaking searchers, Baeza-Yates, Dupret, and Velasco (2007) examined one million queries issued from mobile phones in Japan. The results are consistent with the other mobile search studies with respect to the number of query terms; however, language differences (involving use of ideogram-based language) were noted as the reason for there being fewer characters per query, and cultural differences as the reason for difference in query topics in comparison to the breakdown provided by Kamvar and Baluja (2006). Yi et al. (2008) carried out one of the largest studies to date, examining 20 million English-language queries that had been submitted, internationally. They found meaningful quantitative variations in regional query patterns. For example, although topics of interest are similar in general, users in the United States issued longer and more homogenous queries, with more query terms, than international users did.

Some studies have also investigated the effect of the mobile interface on search behavior. Kamvar and Baluja (2006) compared the search behavior of those who use mobile phones and PDAs. The PDA users issued slightly longer queries, in terms of the number of both words and characters, and it took longer for mobile phone users to enter their query (again likely as a result of the ease of text entry). Searches also varied in the query content – 17% of queries from mobile phones contained URLs, whereas such queries constitute a small minority (2%) with the PDAs, suggesting that mobile phone users utilized the search engine in lieu of bookmarks for content revisit. The topic categories also differed, with adult content being most popular on cell phones and local content on the PDAs. Entertainment content was popular on both platforms. Kamvar and Baluja suggest that the differences are due to demographics and privacy issues. Mobile phones are typically personal devices and allow for privacy (unlike most computers when shared between users), whereas PDAs are typically used for business-oriented purposes. Yi et al. (2008) compared the search behavior of users across a variety of interfaces (SMS, Java application, and XHTML Web interface) and found distinct differences in the breakdown of the queries by topic. Browser and application users searched primarily for entertainment content (55% and 34%, respectively), whereas for SMS users travel-related searches were most frequent (17%), although entertainment was popular too (14%). The authors conjecture that some variations may be due to differences in the capacities of the devices, whereas others arise from demographic differences. Finally, Kamvar et al. (2009) compared Google search patterns in the United States across three device types: computers, iPhone smartphones, and conventional mobile phones. One of the key findings was that search patterns of iPhone users consistently mimicked those of desktop searchers in terms of query length, distribution of query categories, and the lower prevalence of adult content. Interestingly, local searches were less popular on iPhones than on conventional mobile phones, suggesting dedicated application use to meet such needs. Feature-phone and iPhone users had a lower rate of return to the search engine than did desktop searchers, which the authors hypothesize means that mobile search is still a secondary mode of searching. The study's take-home message for high-end phones is that, given the similarities with desktop search, knowledge from this domain could likely be applied to mobile search to improve the user experience.

The effects of the context of use on mobile search have been studied primarily in the context of mobile local search (e.g., Amin et al., 2009; Teevan et al., 2011). Unsurprisingly, these studies indicate the importance of the same contextual factors as the studies of information needs do: location, activities, time, and social situations. Location serves as a general frame of reference for search but also as a target of information needs. For example, Teevan et al. (2011) report that, while 40% of their survey respondents reported looking for information related to their current location, equally many searches were related to the destination and route. Similar findings are reported by Amin et al. (2009), who found that the majority of searches were related to familiar locations (home and work) rather than the user's current location. Church and Oliver (2011) suggest that the user's exact location has less relevance for local searches than when one is using map applications, in that location information and markers in queries are used more to access the properties of a specific place, event, or establishment (e.g., the journey planner of the local public transportation company).

With respect to the temporal context of mobile search, Church and Oliver (2011) found that mobile search use was considered a more urgent activity and was not used to support repetitive patterns as much as general mobile Web use was. Mobile search needs thus are related to particular moments in time and heavily influenced by the user's current activity. Finally, all of the aforementioned studies discuss the importance of social context for search, indicating that most searches were carried out with another person or as part of a group, with the needs often having been triggered by conversations or the needs of the group. Church and Oliver also note from their data that social search activities had a strong influence on fact-checking, trivia-finding, and general informational searches, which refers back to the role of mobile Internet use as a social enhancer (Cui & Roto, 2008).

4.4 SUMMARY

Mobile information access is distinguished from desktop search by the mobile context of use and the influence context has on information-related behaviors. Contextual factors, such as time, location, and social interactions, affect what kind of information people look for and which methods they will use to satisfy their information needs. Availability of and experience with mobile Internet access has dramatically shifted the focus in Web and application-based browsing and search, with particular emphasis being placed on location-sensitive information.

This dissertation focuses especially on understanding and supporting mobile Web search. Previous research has shown that users may struggle with the existing search interfaces, creating a need to consider alternative and complementary methods of result representation and interaction. Aside from addressing the technological limitations posed by mobile devices (e.g., related to connectivity, interaction methods, and display size), a key challenge in supporting mobile Web search lies in interpreting contextual information and integrating it into the search process. Identifying the user's location, his or her current activity, or the social situation may be undertaken, to personalize the search experience, a topic that is explored further in Section 5.3.

5 User Interfaces for Mobile Web Search

Design guidance for mobile search interfaces has been suggested in the literature (e.g., M. Jones, Buchanan, & Thimbleby, 2002; M. Jones & Marsden, 2006), but guidelines specifically addressing mobile search are not prevalent. In general terms, M. Jones and Marsden (2006) suggest that mobile search interfaces should support two main goals: helping the user in assessing the result set (to decide quickly whether a new query is needed) and supporting the user in making good search result choices. These two themes will be explored in subsequent sections. More specific guidance on interface design is provided by M. Jones et al. (2002), who propose the following guidelines for designing small-screen search interfaces:

- Page-to-page navigation should be reduced.
- More rather than less information should be provided for each search result.
- Indicate whether the result points to a normal Web page or a small-screen-optimized page.
- Preprocess normal Web pages to fit better on the small screen.
- Adapt the content for vertical scrolling.

Clearly, advances in mobile technology have rendered some of these guidelines somewhat obsolete, such as the emphasis on preprocessing and flagging content suitable for mobile screens, given that modern touchscreen devices provide much better interaction and presentation capabilities than did early Internet-enabled devices. In a similar fashion, the discussion in the following sections omits the early work on mobile

Web interfaces, which was primarily focused on optimizing the content of the Web page to fit the limited displays of the time (e.g., Björk et al., 2000; Buyukkokten, Garcia-Molina, Paepcke, & Winograd, 2000; M. Jones, Buchanan, & Mohd-Nasir, 1999; Lam & Baudisch, 2005; Wobbrock, Forlizzi, Hudson, & Myers, 2002; X. Xie, Miao, Song, Wen & Ma, 2005).

Despite the considerable amount of research on Web content adaptation, current mobile search interfaces differ little from their desktop counterparts. Search results are presented in what amounts to a scaled-down version of the desktop interface. Even though it is no longer entirely necessary to process Web content specifically to fit better on smaller displays, it can be argued that the available space could be used more efficiently with respect to the presentation of the search results prior to the actual Web access and browsing. It has been argued that the traditional snippet-based result presentation (Church, Smyth, & Keane, 2006), and the scaled-down approach in general (Church et al., 2008), may be inappropriate in the mobile search context. Similarly, clustering interfaces have been suggested as a particularly suitable method for mobile search, since they help to reduce the need for keyword entry, scrolling, and use of display space (Carpineto, Della Pietra, Mizzaro, & Romano, 2006; M. Jones et al., 2002). It should be noted that commercial search engines have recently introduced many useful features that can help overcome the display space, text entry and interaction issues of mobile devices and anticipate user needs. For example, the Google Mobile Search interface (http://www.google.com/xhtml) integrates thumbnail previews of the result Web pages that can be easily browsed with touch gestures. Also, similarly to desktop search, the search result list contains result “cards,” which attempt to answer the user’s query directly (e.g., typing in “weather tampere” shows the current weather information as the top result).

Keeping with the above themes, the following sections summarize research that has been carried out on improving mobile search result presentation and on organizing the results in category-based interfaces. In addition, research aimed at accounting for contextual information needs is reviewed, along with approaches for evaluating mobile interfaces.

5.1 PRESENTING SEARCH RESULTS

Previous research on mobile search result presentation has focused, on one hand, on enhancing the construction of the search result summaries, by, for example, utilizing additional metadata, and, on the other hand, on complementing or replacing the ranked-result-list-driven approach with novel visualizations.

Enhancing Search Result Summaries

Much of the work on interfaces for mobile search result display has focused on determining the optimal amount of result information to display (Hearst, 2009). For example, Sweeney and Crestani (2004, 2006) studied how summaries of different length (only the title or up to 7%, 15%, or 30% of the document content) affect information access performance. These studies' findings point to the maxim that less is more, even regardless of display size (Sweeney & Crestani, 2006): the participants performed better in terms of precision and recall with the shorter summaries. Removing redundancy from the summary content had minimal effect on the correctness of subjects' relevance assessment, precision, and recall, and the differences were not significant (Sweeney, Crestani, & Losada, 2008). However, the evaluations utilized TREC queries and data sets, which are likely not to be representative of realistic mobile search information needs (Hearst, 2009).

Rather than attempt to find optimal length for the summaries, other efforts have explored efficient presentation alternatives for mobile devices. S. Jones, Jones, and Deo (2004) proposed the use of keyphrases automatically generated from result document content as an alternative form of result surrogate. Their user study compared the keyphrases to document titles in a document categorization task on a PDA device. The results of that study indicate that when the title provided for a search result is missing or poor, keyphrases could aid the user in making sense of the results. Similarly, the authors hypothesize that complementing the traditional surrogates with categorizations and keyphrases might help users make better use of search results. As with the summarization experiments discussed above, the results may be influenced by the artificiality of the categorization task (versus an information seeking task having to do with finding results that meet an information need). Church, Keane, and Smyth (2005) suggested use of related queries in a similar fashion to automatically extracted keywords, as a more economical alternative to snippets. A subsequent user study (Church et al., 2006) compared a title-only presentation of results with the traditional ranked result list (title + snippet) and an enhanced result list that showed the title and queries related to the result. The traditional snippet-based interface was judged to provide the most information about the results by 60% of the participants; however, 75% of participants preferred the titles + queries presentation when asked to select the interface that struck the best balance between informativeness and use of display space.

Search Result Visualization

Some approaches have attempted to apply information visualization techniques for displaying search results more effectively on a small display, or to augment the search result display with additional metadata, such as the number or distribution of instances of the query terms.

Milic-Frayling, Sommerer, Rodden, and Blackwell (2004) proposed a system called SearchMobil to support search on PDA devices. This was based on the notion that exposing relevant parts of the document in the search results might make searching easier. Instead of using the ranked result approach, SearchMobil shows a thumbnail overview of the result document for each result, which can be accessed via tabs at the side of the page. The occurrences of query terms are shown visually on top of the salient regions of the page, with the region with the most hits being highlighted in red (see Figure 14). The interface also allows the user to zoom in to the sub-regions of the page to access the document content and perform local searches within the search results. The interface was compared to the unmodified presentation of a Web page in an evaluation wherein the participants were required to find an answer to various information-oriented fact-finding questions. A significant effect of interface on task time was not found; however, a significant interaction effect between task type and interface was observed. Accordingly, SearchMobil was faster than the standard view in five out of the six tasks wherein the answer could be found in the region highlighted as most relevant in view of the number of query term instances.

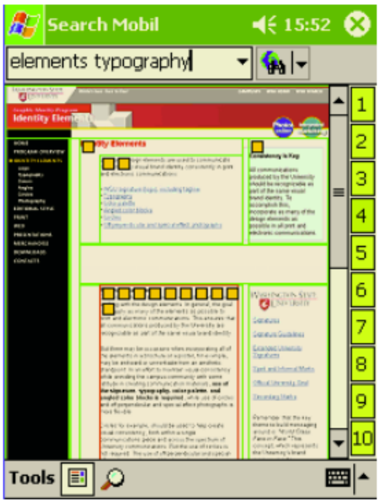


Figure 14. The SearchMobil interface showing the query term hits inside page regions for the selected result document, with the region with most hits highlighted in red (Milic-Frayling et al., 2004, Figure 4a, © 2004 Springer-Verlag).

The work reported in Paper II in this dissertation is an attempt to support the search result evaluation process, annotating the traditional result descriptors with a visualization that depicts the occurrences of the searcher's query within the result document. Two alternative forms of the visualization, graphical and numerical, were compared in a user study wherein the participants carried out informational search tasks with their own queries. The visualizations were calculated and displayed on the fly in the mobile search interface (see Figure 15).

In a user study, most of the participants preferred the visualizations to the standard interface, and the preference also had a significant effect on which interface they elected to use when given the choice. The majority of participants had mixed reactions to the visualizations, with the visualizations being judged useful in some situations, such as when the search engine failed to provide a text snippet for a result. The visualizations were used to disambiguate relevance when the textual descriptors were not clear enough. Subjective feedback suggests that the number-based visualization may be slightly easier to interpret and better supports the identification of non-relevant results; however the type of visualization did not affect the perceived difficulty of task completion or finding of useful results.

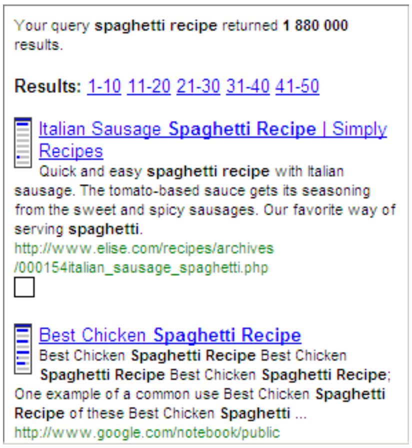


Figure 15. Graphical query occurrence visualizations displayed alongside mobile search results (Heimonen & Siirtola, 2009, Figure 2, © 2009 IEEE).

Shtykh, Chen, and Jin (2008) proposed a method for presenting the search results as a "slidefilm" interface rather than as a long ranked result list. Each search result is presented as a self-contained page ("slide"), and the user can navigate through the result list by using the left and right softkeys of the mobile phone. This presentation would have the advantage of being able to show more details about a result than the ranked result list does, and thereby reduce the amount of scrolling. A formative user study with predefined queries and cached results compared the proposed interface to a baseline ranked result list. Although the proposed interface resulted in faster task completion, the number of successfully completed tasks was smaller. Eight out of nine participants preferred the proposed interface to the baseline, and the subjective satisfaction rating was higher, yet whether these differences were statistically significant is not reported.

Improvements have also been suggested to the visualization of search result clusters in mobile search interfaces. Mizzaro, Sartori, and Strangolino (2012) proposed a tag cloud interface called CloudCredo for representing clustered search results as an alternative to the traditional tree view representation (see Figure 16). The cluster hierarchies can be expanded in the tag cloud visualization to reveal subcategories. The tag cloud visualization was compared to the tree-based cluster interface in a user evaluation on iPhone and iPad. The tag cloud was found to be more effective than the tree interface in one task out of four on the iPhone, and it was subjectively preferred to the tree view on the iPad.

Figure 16. The CloudCredo tag cloud interface displaying the search result clusters (Mizzaro et al., 2012; Figure 1b).

5.2 Organizing Search Results

In terms of organizing the search results, category-based interfaces have been suggested as one key approach. With respect to the design of category-based mobile search interfaces, Chan, Luk, Leong, and Ho (2009) carried out an *in situ* contextual inquiry study to solicit requirements for category-based mobile search. Their results conform to the lessons learned from utilization of category-based interaction in desktop search interfaces as discussed in Section 3.2. For example, they found that participants would prefer selection-based input to typing keywords. Hierarchical access was considered especially convenient if the user either knows which part of the category hierarchy the result likely belongs to or does not have a clear objective for the search. Most importantly, the categories should make sense to the users. In terms of interaction design, analysis of search behavior suggested that the hierarchy, search results and content should be provided on separate pages while enabling quick navigation between the different views.

The following sections review different category-based approaches that have been suggested for mobile search: clustering interfaces and classification-based interfaces.

Clustering Interfaces

Carpineto et al. (2006) introduced Credino, a clustering search engine for mobile devices based on concept lattices, which is a form of hierarchical clustering. The hierarchical categories are arranged as an expanding tree, where the cluster labels serve as links to result pages. A small-scale user study suggested that clustering can provide better search performance than ranked result lists. Figure 17 presents a more recent version of Credino adapted for the iPhone (Mizzaro et al., 2012).

Figure 17. The Credino interface showing search results clusters in an expanding outline view (Mizzaro et al., 2012, Figure 1a).

A follow-up study by Carpineto, Mizzaro, Romano, and Snidero (2009) compared desktop, PDA, and mobile phone search interfaces, with and without clustering, in a controlled experiment. Participants carried out four tasks, representing different information needs (e.g., informational, transactional, and navigational). The results show that clustering is more effective than the ranked result list on PDAs and mobile phones. Closer examination of individual tasks identified some benefits and drawbacks of clustering. Clustering failed to provide relevant category labels in the fourth task, leading to low performance. In contrast, clustering performed well in the first task because the cluster hierarchy provided a good overview of an otherwise unfamiliar search topic and helped narrow the search down to the correct result. Carpineto et al. noted the need for collecting more evidence about clustering performance, with more diversified tasks specific to real-world mobile search scenarios.

The present author presented Mobile Findex (see Figure 18a), a mobile Web search interface that provides a flat list of clusters computed on the basis of the most common words and phrases within the search result captions (Paper III). Our 16-participant laboratory user study compared the ranked result list to the clustering interface for 12 information seeking tasks with topics such as trivia and shopping. The results suggest that the clustering interface could offer search performance comparable to that of a traditional mobile Web search interface, with participants preferring Mobile Findex because of its perceived efficiency, suitability for the type of tasks used in the study, and ease of finding results. The overview and filtering capabilities provided by the clusters were judged to be more essential to the user experience than pure search performance. Further analysis of the results showed that clustering aided with ambiguous queries because the cluster labels could be used for disambiguation and drilling down into relevant sets of results (Paper IV). A subsequent re-implementation of the interface was evaluated in a longitudinal field study (Paper V). The touchscreen-optimized version of this interface is depicted in Figure 18b. The results confirmed those from the laboratory study but also revealed issues in the interface and clustering designs that affected user experience, such as the quality of cluster labeling, presentation of clusters in the interface, and type of category structure generation algorithm used. These results are discussed in more detail in Chapter 6.

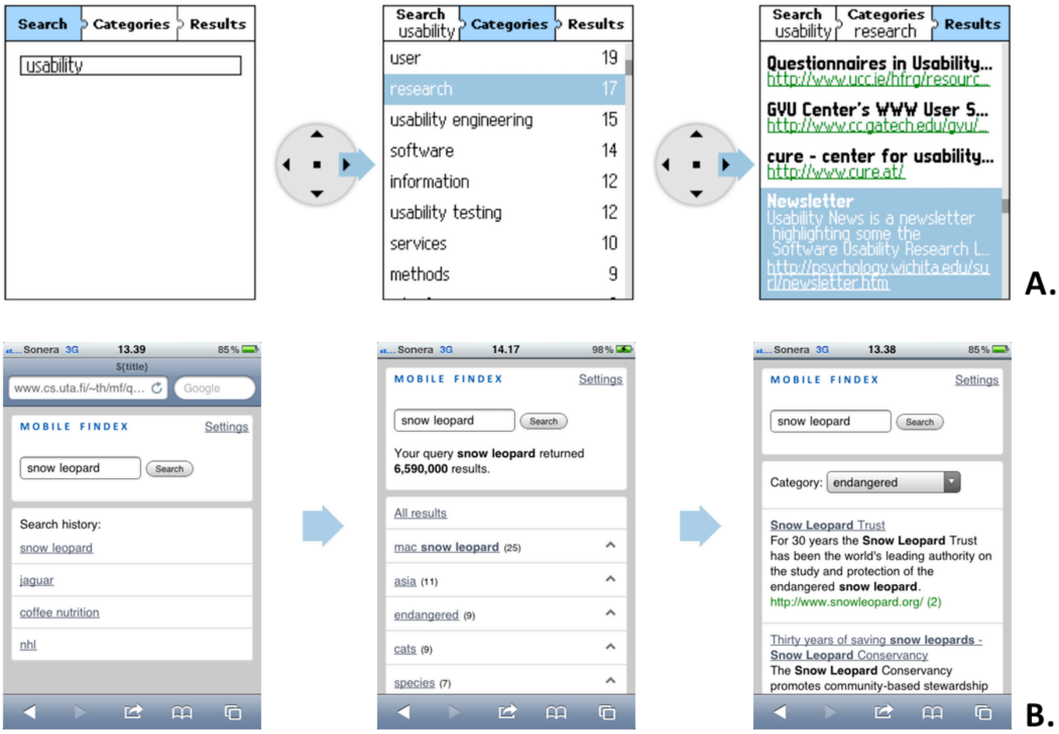


Figure 18. Query, category, and results views in different Mobile Findex versions: a) J2ME interface (Heimonen & Käki, 2007, Figure 1, © 2007 Association for Computing Machinery, Inc.) and b) touch-screen optimized HTML interface.

Hierarchical and Faceted Classification Interfaces

In addition to clustering interfaces, classification and faceted search interfaces have been suggested for mobile search. Buchanan, Jones, and Marsden (2002) introduced LibTwig, a category-based overview interface for mobile digital libraries. Similarly to Credino (Carpineto et al., 2006), the LibTwig user interface organizes results into an expanding outline tree, which the user can explore by selecting tree nodes until the actual result documents are reached. Evaluations of LibTwig suggest that non-expert Web users prefer the outline approach because it provides them with a good overview of the result set. De Luca and Nürnberger (2005) proposed an approach that combines several classification methods, such as ontology-based classification of the documents, and use of the user's bookmark structure to train a document classifier, which is further fed by the addition of new bookmarks. Evaluation of the concept is not reported upon, so it is difficult to assess the efficacy of the proposed approach in practical search scenarios.

A few examples of mobile faceted search and browse interfaces have also been proposed in addition. Karlson, Robertson, Robbins, Czerwinski, and Smith (2006) introduced FaThumb, a mobile search interface based on browsing of faceted hierarchical metadata and incremental text entry. The facets are browsed with the mobile phone's keypad, whereby each key is spatially mapped to the facet zones arranged into a similar grid on the screen (see Figure 19a). Pressing the corresponding key navigates into the content of the facet, arranging the items in the next-level hierarchy into the zones. The center zone provides a spatial overview of the navigation and the left and right menu buttons (below the facets) are used to show context-dependent commands. Additionally, freeform text can be entered to filter the results within a facet. The facet interface was compared to keyword search in a formative user study. While an overall difference in search speed between the interfaces was not observed, faceted browsing was rated higher for overall satisfaction than keyword search was. The results indicate that keyword entry is faster that for known item search, while browsing for unknown targets is faster with facets (which was also explicitly stated by participants).

M. L. Wilson, Russell, Smith, and schraefel (2006) proposed a mobile version of the mSpace faceted search interface. The design addresses the problem that multiple queries are often needed to address compound information needs, which require the user to satisfy several constraints across information domains (e.g., finding a restaurant for lunch within short walk from a movie theater that is showing a film of interest, which is readily accessible via public transportation). The interface enables its users to make multiple concurrent queries by using columns, each of which encapsulates a facet of the underlying information space. The facet panes are arranged into a grid, where the user can progressively zoom into a specific pane to show information in more detail (see Figure 19b). Elements found via searching can be added into the "Favourites" view, allowing items of interest to be kept persistently available. A formative field trial indicates that this interface performs better than Google Local search does in location discovery tasks, both when the user is stationary and when he or she is moving. The authors also argue that breaking away from Web-page-oriented designs enables the creation of better interfaces for mobile information access and exploration. In the case of mSpace Mobile, this is suggested to have been thanks to the persistent display of information and the focus + context views, which reduce the need for network access and decrease cognitive load by emphasizing recognition of relevant information over recall.

Figure 19. Approaches to faceted search and browsing: a) The FaThumb interface (Karlson et al., 2006, Figure 2a, © 2006 Association for Computing Machinery, Inc.), and b) the mSpace Mobile interface (courtesy of Max L. Wilson), and c) the Mambo music browser interface (Dachselt & Frisch, 2007, Figure 1, © 2007 Association for Computing Machinery, Inc.).

Finally, Dachselt and Frisch (2007) introduced Mambo, which is a facet-based music browser for mobile devices such as PDAs and ultra-mobile PCs. The browser allows for accessing the music data along several metadata facets, such as genre and name. It is based on a zooming widget called FacetZoom, which organizes the underlying facet hierarchy into an interactive, visual display. Each level in the hierarchy of the facet structure is mapped to a horizontal bar, which is divided into as many columns as there are nodes (see Figure 19c). The widget supports multiple interaction modes: horizontal panning is for moving seamlessly between nodes on the same level; vertical panning moves between levels in the hierarchy; and tapping a node will center it in the display, which allows quick transition between nodes on different levels. Results from a formative user study wherein the participants engaged in search, comparison, and filtering tasks suggest that the zoomable user interface design scales well across various mobile displays in terms of search performance. Subjective satisfaction feedback indicates that the faceted browsing and hierarchical refinement of the query received a positive reception.

5.3 ADDRESSING THE MOBILE CONTEXT

Some mobile search interface designers have attempted to account for the contextual nature of mobile search by emphasizing certain features, such as the social and location-dependent nature of mobile search.

Using Context Information for Query Refinement

Researchers have sought to improve mobile searchers' query formulation by using contextual information. Arias et al. (2008) proposed a mobile search interface based on context-based personalization of query suggestions. When the user begins to type a query, the system suggests relevant concepts from a domain-specific thesaurus on the basis of the user's context (e.g., weather, location, and time). The prototype implementation covers transportation, leisure, and public services domains. The user can select from among the concepts to construct the query, which is then executed by means of a normal mobile search engine. In its essence, this approach resembles search result categorization wherein the classification is performed prior to execution of the query, and the categories (or concepts) are presented to the user to inform query formulation. A formative user study revealed that the proposed system is intuitive and reduces the effort of typing queries. Observations of system use indicate that the participants would engage in refinement of the query with the concepts rather than write a new query from scratch, which is consistent with the findings from large-scale studies that indicate mobile searchers' engagement in significant amount of query reformulation (e.g., Church et al., 2008). Because the evaluation compared the context-based concept recommendations to standard keyword search, it is difficult to assess what benefit, if any, the contextual recommendations provide over query recommendations based on popular queries by other users.

Results from extensive experiments reported by Kamvar and Baluja (2007b) suggest that it is also possible to predict users' queries on the basis of the application (SMS vs. Web search), location and time of day, potentially reducing the need for text entry by as much as 46%. Kamvar and Baluja also propose that context could be used to suggest queries even before the user begins typing. Query generation and results' ranking can also be augmented automatically on the basis of the inferred contextual model. To this end, Yndurain, Bernhardt, and Campo (2012) introduced a system architecture for context modeling, based on signals acquired from device use and sensors (e.g., location, time, and application) to identify the user's state, which can be used to enhance the queries by means of heuristic rules.

Social Mobile Searching

M. Jones, Buchanan, Harper, and Xech (2007) investigated how capturing and utilizing the incidental search activity of people in certain physical locations could be used to support mobile search. Queries were captured in urban locations from nearly 400 searchers. In a subsequent study, these location-dependent queries were presented in the search interface as participants performed queries in the same locations. The results suggest that providing other people's queries can positively influence the mobile information seeking. The participants were divided into two groups on the basis of their behavior: *Searchers* and *Clickers*. The Searchers made verbatim use of the queries presented when searching (about 15% of their queries) and reported an improved perception of search effectiveness (in comparison to the standard search interface without location queries). The Clickers, who used only the queries submitted by other users, found the search experience less enjoyable than Searchers did but similar in terms of performance and mental effort. These findings informed the design of the Questions Not Answers interface (Arter, Buchanan, Jones, & Harper, 2007), which shows location-specific queries from other users in a map-based interface, overlaid on a map display (see Figure 20). Clicking on a query automatically retrieves the first 10 search results associated with the query and shows them in a list format. One of the key takeaway messages from the field study of the prototype is that location-based search interfaces would benefit from a clear distinction between functional information about a location and less familiar, playful information, with the latter being supported with exploratory interfaces such as Questions Not Answers. The results also show that the insights and sense of place that can be gained and reflected by queries vary by location type. The map-based prototype gave the participants a clearer sense of place in locations where queries were distinct (i.e., related to the qualities of the place rather than generic properties such as nearby pubs or coffee shops).

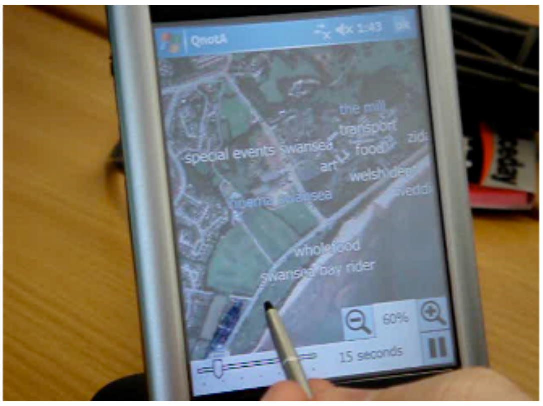


Figure 20. The Questions Not Answers interface showing queries over a map display (courtesy of Dr. Matt Jones).

Church, Neumann, Cherubini, and Oliver (2010b) introduced Social Search Browser (see Figure 21), which is similar to the Questions Not Answers interface but provides additional filtering options based on time, relationship and query similarity, and social collaboration features (i.e., users can pose queries to each other and, in turn, provide answers). The results from a longitudinal evaluation show that the participants were not overly concerned about sharing their queries or answers to questions posed by other users. The majority of users also preferred the human-generated content to traditional search engine results, especially for personal and time-sensitive information needs.

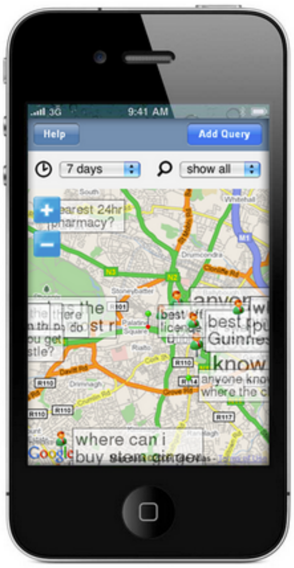


Figure 21. The Social Search Browser interface showing queries over a map display, with filtering controls (courtesy of Dr. Karen Church).

A subsequent longitudinal user study (Church, Neumann, Cherubini, and Oliver, 2010a) compared a text-based interface to a map-based version. The results indicate that the choice of interface depends on personal preferences, the information need, and the situational context. The authors suggest that it may be possible to determine automatically which interface to show by analyzing the query intent, usage patterns, and contextual factors. For example, the participants preferred the text-based interface when browsing the queries submitted by other users, and when under time pressure. Church et al. (2010a) suggest as a general take-away message that location-based search tools should support both text and map interfaces in a hybrid fashion whereby the users can switch between the two presentation modes seamlessly.

Reis, Church, and Oliver (2012) reflect on the findings from previous research as they focus on the importance of social interactions, curiosity, and boredom as motivators for mobile Internet use. Their survey of mobile Internet users showed that curiosity and boredom were the main motivators for social mobile search, that most of the popular information needs were related to trivia and pop culture, and that mobile searchers share results by speaking and showing their devices to others. The key take-home message from the study was that social, casual mobile search experiences could be improved by providing better facilities for sharing search results with one's peers.

New Paradigms for Mobile Search

Researchers have also suggested completely new paradigms for mobile search; emerging from the understanding that keyword-based searching centered on important information needs is not always necessary or desirable in mobile settings. M. Jones, Buchanan, Cheng, and Jain (2006) suggested "background" information seeking as a way to complement traditional information seeking with a less cognitively demanding and intrusive approach. They developed a tool for users to capture queries on a PDA device while offline. Queries can be entered manually or selected from keywords extracted from the content on the device (e.g., notes, calendar items, and task entries). When the PDA is connected to a PC, the queries are processed and fed to Web searches, and the resulting Web pages are cached for access on the PDA and the desktop computer both.

Subsequently, M. Jones (2011) reviewed earlier research trends in mobile search and suggested that a need will arise to explore user needs that are not met with current systems. Such future systems could move from the directed, immediate, and discrete mode of search towards more indirect and delayed information seeking wherein the results are delivered continuously rather than in neat chunks. He also notes that there are many challenges and opportunities in addressing mobile search in emerging markets, where interface paradigms other than the predominant text-based keyword search may be more effective.

5.4 Evaluation of Mobile Interactions

Earlier discussion has established that evaluating search interfaces is challenging on account of numerous factors, many of them related to the internal and external validity of the evaluation. Mobility heightens this tension because of the effects of the mobile context of use in general, and the contextual nature of mobile information needs in particular. Nakhimovsky, Eckles, and Riegelsberger (2009) describe the three major goals of existing mobile user experience evaluation approaches: capturing the user's interaction with the system and its state, capturing and recognizing the user's context, and gathering and managing self-reported feedback from users. While many of the standard methods in mobile user experience evaluation involve lab-based evaluations, these authors point out that the demands of mobile contexts have motivated the development of methods that enable the study of situated interactions.

Laboratory vs. Field Studies

Several researchers have addressed the tension between laboratory-based evaluation and field studies of mobile applications. For example, Kjeldskov, Skov, Als, and Høegh (2004) argue that field-based usability studies of mobile systems are difficult to conduct, time-consuming, and of unclear added value. They conducted laboratory and field-based evaluations of a context-aware mobile system and found that the field evaluations provided little added value. The main argument is that by recreating the central aspects of the context, the laboratory allows for the identification of the same usability problems. Also, Kaikkonen, Kallio, Kekäläinen, Kankainen, and Cankar (2005) report findings indicating that the same problems were encountered in both environments, with variation only in the frequency of occurrences.

C.M. Nielsen, Overgaard, Pedersen, Stage, and Stenild (2006) and Rogers et al. (2007) have provided critique of the use of lab studies for evaluating mobile systems. C.M. Nielsen et al. report findings from identical evaluations carried out in the field and in the lab, where they found that the field context resulted in more usability problems being identified. The field setting also revealed problems with the interaction style of the application and the users' cognitive load that were not identified in the laboratory. However, the authors note that extending the usability evaluation approach to the field may interfere with the realism of the situation – e.g., if participants are recorded and their tasks are constrained on account of desire for comparability with results in laboratory settings. One factor that could confound the results of studies comparing laboratory and field studies is that the research questions are focused rather more on finding differences between the evaluation settings than on understanding the use of the system in question. Rogers et al. note that it is possible to design *in situ* field studies in such a way that important information about the use of various functions and difficulties encountered by the participants in different contexts are captured, even though this can be costly in terms of time and effort. For example, they captured the interaction data from the device, recorded vignettes of the system use on video, and gathered opportunistic observations during the use of the system. They argue that it is not possible, or even desirable, to capture everything the users do *in situ*. Instead, they suggest using a mix of methods to reveal the merits of various system features and why they were used or not used.

Lab-based evaluation can, however, be appropriate for some contexts and applications, especially if the environmental conditions are of relevance to the use of the proposed system. For example, Lumsden, Kondratova, and Durling (2007) studied how to support mobile speech-based data entry effectively in a lab setting where the physical surroundings and soundscape of a city street were simulated. This appears to be highly appropriate when the primary interest is in providing an environmentally realistic context of use. Realistically simulating an environment that stimulates the naturalistic use of mobile search interfaces is somewhat more challenging in the lab. The research on mobile information discussed above clearly indicates that mobile information needs, and subsequently use of mobile search, are related to various temporal, social, and activity contexts that are relevant to the user's everyday schedules and interactions.

Multi-method Field Studies

As already mentioned, it is unlikely that all aspects of mobile search interfaces could be successfully evaluated in laboratory settings, given the contextual richness of mobile information needs and the effect of the context of use on the interactions. This author has carried out a number of evaluations of mobile search interfaces in the lab (papers II and III) by utilizing approaches similar to those used for controlled experiments with desktop search interfaces. Other researchers (e.g., Carpineto, Mizzaro, et al., 2009; Church et al., 2006; S. Jones et al., 2004; Karlson et al., 2006) too have used formal lab-based user studies in their evaluations. It can be argued that without the influence of context, the results of these studies reflect primarily the qualities of the search interface presentation and interaction. In contrast, multi-method field studies of mobile search interfaces, recently increasing in popularity, use several distinct data collection methods in concert in an attempt to understand the use of the system under evaluation from different perspectives. For example, Riegelsberger and Nakhimovsky (2008) report their experiences from utilizing a multi-method approach to study Google Maps for Mobile, which included recorded usage and its analysis, group session briefings, field trials, telephone interviews, and in-person debriefing. The use of multiple approaches balanced out the weaknesses of the individual methods and provided insights not only into the use of the system but also about product adoption and the mobile ecosystem.

Longitudinal studies using combined methods have also been utilized in exploration of mobile search interfaces. In the Questions Not Answers study (Arter et al., 2007), the researchers used usage logs and diary entries to collect data about the use of the application over a five-day period. These were combined with telephone interviews to canvass initial impressions and conduct post-study interviews. Arter et al. (2007) report that this approach was instrumental in uncovering unexpected behavior patterns that would have been difficult, if not impossible, to discern in a short lab session. Amin et al. (2009) investigated location-based mobile search behavior in a 12-day study that incorporated Web-based diary entry collection, search interaction logging, and interviews. This method enabled the researchers to capture explicit search behavior (queries), intentions (motivations behind search), and the context of the search. Church et al. (2010a, 2010b) studied the Social Search Browser prototype and its two interface modes in two field studies, collecting interaction data from the usage. At the end of the study period, the participants were presented with a post-study survey for collection of subjective feedback on application use. In the latter study, also contextual experience sampling (Church & Cherubini, 2010) was used to investigate the reasons for *extreme behavior* (e.g., non-use of the application or switching of the interface type). The experience samples revealed visual and location-dependent aspects of the use of the Social Search Browser interfaces – the map was better for visual overviews and pinpointing local queries, while text was better for accessing the questions and answers submitted by other users. The present author used a longitudinal approach in a study of a mobile clustering search interface carried out in 2009 (Paper V). The selection of research methodology was motivated by the tradeoff between increased richness of data collection and longer exposure to the search interface with a decreased burden of attendance for the participants. A semi-structured interview conducted at the end of the study allowed for exploration of the main themes in the search interface's use, highlighting contextual factors related to the decision to use or not use the clustering interface.

Capturing the Context of Mobile Interactions

One of the challenges with using diary study protocols for capturing contextual information is that the participants may be unable or unwilling to provide rich, thorough entries when mobile or engaged in an activity (Brandt et al., 2007). This could lead to deferring their submission to more appropriate situations, which can affect the frequency of reporting and recollection. Brandt et al. (2007) suggest an SMS-based approach to resolving this issue, which entails asking the participants to record only brief snippets of data when active, then later fill in the details by means of a Web-based diary tool. Such an approach shares the potential limitation that later recollection of events and details may be difficult; however, preliminary results suggested that the snippets can help mitigate such concerns.

Another approach is to use an experience sampling method (ESM), whereby the participants are sampled at specific points in the day over a longer period of time. The departure from the diary study approach is that the participants are not required to recall anything; instead, they are asked about their current activities and feelings (Consolvo & Walker, 2003). Consolvo and Walker (2003) note that a potential problem with experience sampling is in the scheduling of the sampling. Scheduled sampling and sampling based on user-generated alerts may introduce a cognitive bias: with the former, the participants may anticipate the sampling and modify their behavior, and with the latter they may pick and choose when to trigger the alerts. To overcome these biases, Cherubini and Oliver (2009) propose a refinement of the experience sampling method, which relies on recording contextual data on the device (e.g., user location and the use of device features) and using it as a way of triggering contextually relevant requests to the user. Further, Church and Cherubini (2010) suggest a combination of diary studies and contextual experience sampling, wherein participants are sampled as to their information needs at appropriate times and in a non-intrusive manner via SMS. As discussed above, the method was utilized in the evaluation of a social mobile search interface and was able to reveal interesting context-dependent behavioral patterns.

Contextual experience sampling offers possibilities for enhancing field studies, especially in capturing salient user feedback related to relevant usage patterns. The challenge with mobile search interface evaluation is the identification of such patterns of interest. For example, Church et al. (2010a) employed non-use and switching of the interface used as triggers. In the context of category-based interfaces, similar criteria could prompt sampling for experiences if users cease using the category features or use them in interesting ways. For example, one such pattern could include users engaging in a significant amount of exploration of results within a session or making several result selections. These interactions may point to either an in-depth sensemaking session or problems with the quality of category overview. Experience sampling could aid in understanding which type of situation has occurred.

Research "in the Wild" and Large-scale Trials

Longitudinal *in situ* studies and evaluations that attempt to capture rich contextual data from realistic usage situations are becoming increasingly "standard" in current mobile HCI research practice. However, this kind of research is not free of its own methodological problems. Brown, Reeves, and Sherwood (2011) examined the challenges of field trial methods by analyzing the practices of both researchers and participants. They discuss their findings with respect to three issues that are rarely highlighted: the effects of demand characteristics (participants' desire to contribute to the success of the study); influence of lead participants (a subset of participants who provide key insights or foster activity in others); and differences introduced by trial design choices (even when the same system is under evaluation). Brown, Reeves, and Sherwood conclude that the nature of "in the wild" trials has certain implications for how such methods should be developed. They suggest not only that one should reject the strict notion of reproducibility inherently associated with traditional HCI evaluation paradigms but that the goals of evaluations should move away from the normative notion of success and consider more how technology interacts with users and their practices.

One of the key issues with field studies is resourcing, since it is seldom feasible to recruit more than a few dozen users for any one study, given the time-consuming nature of data capture and feedback elicitation. Therefore, researchers have embraced application stores and other channels for wide distribution to distribute the prototype applications to be studied. Application stores are a great opportunity for researchers because they allow for large-scale deployment of mobile applications, the naturalistic use of which can then be observed in longitudinal settings (Church & Cherubini, 2010). However, as any other evaluation approach does, application stores also have their own challenges to account for and overcome. Michahelles (2010) reports experiences from several case studies and highlights three main concerns with the adoption of the applications: the actual initial adoption is unclear, unlike in user studies whose participants are explicitly recruited; external factors outside the researchers' control (e.g., ratings and competing applications) may inhibit adoption; and user demographics are often unknown beyond broad geographical classifications. This clearly introduces a certain degree of ambiguity to the interpretation of the results. Church and Cherubini (2010) share these concerns with user demographics, noting that it may be difficult to generalize beyond the early adopter group of high-end smartphone users who frequent application stores. For example, A. Morrison, Reeves, McMillan, and Chalmers (2010) report that users aged 15 to 35, usually male, have been typical adopters of their applications. A. Morrison et al. also discuss the nature of participation in application store trials. For example, how is one to choose the activity threshold for flagging a user as a participant in the study? They provide an example of their own research, wherein a mobile game was downloaded over 180,000 times but was actively used over a more sustained time by only 3,080 people.

Some guidelines have been suggested for carrying out research with application stores. Michahelles (2010) suggests several guidelines, which share qualities with other experience-based recommendations for longitudinal field studies (e.g., Shneiderman & Plaisant, 2006; Church & Cherubini, 2010; A. Morrison et al., 2010). According to these, the application should provide user benefits; the objective of the research should be transparent to the users; application functionality should not be restricted for non-participating users; collection of feedback should be minimized, in favor of interaction logging; and remote configuration mechanisms should be used to update the content of the application without requiring installation of a new version. A. Morrison et al. (2010) offer additional recommendations, advocating the integration of simple and expedient data collection and evaluation into the context of the application (e.g., providing a feedback section from which users can respond with text or multiple-choice selections).

Application stores offer an interesting avenue for conducting global user trials of mobile search interfaces. The challenges of demographics and feedback collection notwithstanding, it would be extremely interesting to evaluate the clustering search interface concept reported upon in this dissertation on such a large scale. One major research question would be whether the interaction patterns identified in the smaller-scale longitudinal trial are also exhibited on a larger scale, and how the adoption evolves over a longer time. Also, one of the issues with prior research on desktop and mobile category-based search interfaces (with the exception of Koshman et al.'s (2006) work) is that the evaluations have included a few dozen users – far too few to argue convincingly for the adoption of such methods in commercial search engines. Large-scale evaluation would also allow for the integration of implicit feedback collection techniques, to improve the quality of the clustering, which was identified in the participant feedback as a pain point.

5.5 SUMMARY

Much of the early research into mobile search interfaces was dominated by the desire to improve their usability – for example, by making it easier to access the results by decreasing scrolling or using the available display space more effectively to represent the results. Several studies have attempted to adapt methods from desktop search interfaces, such as visualizing the locations and frequency of query terms, creating effective thumbnail representations, or using categories to provide easier access to subtopics among the results. The aim of such activities is primarily to reduce the searcher's uncertainty with respect to which result or results to access. The motivation is that the cost of erroneous selections in terms of access time and convenience can still be high.

With the advances in mobile technology, these device-oriented concerns have become somewhat secondary to ability to support contextually emergent mobile information behaviors. In addition to providing better query suggestions and supplying contextually relevant information (e.g., on nearby services), context information can be put to *indirect* use, as in exposing the searchers to other users' queries and questions. This leverages the prevalent social aspects of mobile information access. Similarly, it may be possible to improve the search result visualizations and result organization methods by considering the context. Identifying the intent and possible location-based or social motivations behind the queries could inform the choice of result representation method. For example, clustering-based presentation of salient discussion topics and related information could support the addressing of needs that emerge out of social interactions, whereas consistent and understandable classifications or faceted interfaces would likely better facilitate search for, and comparison of, local services.

The evaluation of mobile interactions provides an additional challenge for search interface research. Although controlled experiments and usability studies are still necessary, as they are for evaluation of search interfaces in general, longitudinal *in situ* studies are best situated to address issues that arise over time in the mobile context of use. Application stores and other wide distribution channels appear to be particularly suitable for large-scale evaluations in naturalistic settings to complement smaller-scale studies.

6 Introduction to the Publications

The research done for this dissertation consists of the design and development of search result visualization and organization interface concepts, user studies for these, and respective research articles that have targeted two distinct topics: presentation of individual search result elements (papers I and II) and presentation and interaction involving search result sets (papers III, IV, and V). This fundamental division was selected on the basis of the guidance for mobile search interface design given by M. Jones and Marsden (2006), who remind that it is essential to provide an overview of the goodness of the search (at the level of search results) and also support the evaluation of individual results. In addition, the mobile information needs and information access methods of mobile Internet users were studied, for understanding of the role of search in everyday information practices (Paper VI).

These topics, and their evolution, took shape over several years in the course of the research process. Insights gained during the work on the space-saving, compact visualization introduced in Paper I suggested that the proposed visualization would especially suit mobile search interfaces, owing to their lack of available screen space. The visualization was then further refined and evaluated in mobile search scenarios, as reported in Paper II. Papers III through V cover a similar series of research activities, culminating in a long-term user study wherein the proposed interface was evaluated in the context of real search tasks. The need for the study of information needs reported on in Paper VI came about from the understanding that, in addition to studying *how* users interact with mobile search interfaces, one must understand *why* mobile search is used, and *what* kinds of needs it is used to address.

The starting point for the research process was the work reported on in the first publication titled "Visualizing query occurrence in search result lists." This study reports the design process for a small document-shaped icon that can be embedded in the search result list to visualize the occurrence of search terms in a result document. Included are the findings from an 18-participant user study in which the design was compared to the traditional search result list in a controlled experiment. Although we did not observe statistically significant differences in objective search performance, the participants' subjective ratings and opinions of the visualization's utility were positive. Our findings also pointed to a need to study such interfaces in a longitudinal setting, given the limitations of the rigidly controlled experimental methods when it comes to assessing utility in realistic search scenarios.

In the follow-up study, described in "Visualizing query occurrence in mobile Web search interfaces," we further developed the query occurrence visualization to suit mobile Web search interfaces. The visual design of the document-shaped icon was simplified, but the underlying term extraction algorithm and data model were kept relatively unchanged. In the user study, carried out with 18 participants, we introduced a text-based variant of the design and compared the two to the traditional search result list. The results from the study suggest that the key benefit of the proposed visualization in the mobile context is its unobtrusiveness. It does not distract from normal use, yet it offers situational utility. The results of studies I and II indicate, however, that the utility of the proposed query-occurrence-based visualization approach is dependent on how well it matches the user's mental model and existing search strategies.

In the third publication, "Mobile Findex – Supporting mobile Web search with automatic result categories," the focus shifted to search result presentation at the level of search result sets. We presented a native J2ME mobile-search user interface called Mobile Findex, which uses search result clustering to present the user with an initial, navigable overview of the search result set. We carried out a lab-based user study with 16 participants to evaluate Mobile Findex, and the subjective evaluations show that the users clearly preferred it in terms of ease of finding relevant results, suitability for the tasks, and perceived efficiency. The use of categories showed a marginal objective benefit in traditional search evaluation, although a significant difference between Mobile Findex and the traditional, Google-like mobile Web search interface was not observed. The fourth publication – "Mobile Findex: Facilitating information access in mobile Web search with automatic result clustering" – extends the discussion in the third paper by providing a more thorough literature review, more detailed analysis of the results of the user study than possible in a conference paper, and an expanded discussion of the results.

The fifth publication (titled "How do users search the mobile Web with a clustering interface? A longitudinal study") reports the results from a longitudinal user study with 17 participants, in which a Web-based version of the Mobile Findex clustering interface was studied for a period of four weeks. We used a variety of qualitative and quantitative subjective feedback collection methods to assess user experience with Mobile Findex. In addition to collection of interaction data from use of the system, user-experience-oriented research methods provided feedback on the perceived usefulness and utility of the clustering interface. The results reveal the situational benefits provided by clustering, realized in situations in which an overview of the results is needed to inform focusing on a subset, or when the user is unsure of how to approach the search. Several insights were also gained with respect to the design of category-based interfaces, related to the creation of the categories, their presentation, and interaction with them.

The sixth and final publication, "Information needs and practices of active mobile Internet users," focuses on the contexts of mobile information seeking activities and the information needs of users that could potentially be supported with mobile Web search activities involving the various presentation methods developed in conjunction with papers I–V. We conducted a four-week diary study with experienced and active mobile Internet users, focusing on the physical and situational contexts of their mobile information needs and the information access practices employed to fulfill them. Our results complement earlier studies and indicate that unrestricted mobile Internet access shapes information access behavior. Web browsing and search were found to be the dominant information access methods, and they were used to address the needs as they emerged. Our results suggest that for everyday information needs, the role of the activity and social setting is more pronounced than is location as a trigger for information needs, and this should be taken into account in the design of mobile information search interfaces.

6.1 Visualizing Query Occurrence in Search Result Lists

Reference

Heimonen, T., & Jhaveri, N. (2005). Visualizing query occurrence in search result lists. *In Proceedings of the 9th International Conference on Information Visualisation, IV '05* (pp. 877–882). Washington, DC, USA: IEEE Computer Society. doi:10.1109/IV.2005.152 (Paper I)

Objective

The objective of this study was to investigate the applicability of query-biased visualization in augmenting individual search results. The work was carried out in conjunction with a research project focusing on next-generation user interfaces for search and the use of visualization techniques to aid in the search result evaluation process.

The idea behind the query occurrence visualization builds on the TileBars approach proposed by Hearst (1995). Our aim was to reduce the complexity of the TileBars visualization while retaining the underlying design principle of showing keyword frequencies in different sections of the document (see Figure 8a). The visualization design went through several iterations, in which the number of document sections and the data to visual structure mappings were experimented with. Early on, we decided on the full query phrase as the atomic visualization unit, instead of – as is more typical – the individual query terms, with the reasoning that the query phrase encapsulates the user's information need and that the occurrence of all the query terms in the same context would signify an item of interest. The validity of this approach and the related questions of benefits for search performance were then examined in a laboratory-based user study with 18 participants.

We implemented a desktop search user interface prototype on top of the Findex platform (Käki, 2005b) for the purposes of the user study. A significant amount of work was done on the back-end systems that enabled the harvesting of Web pages and subsequent parsing to provide the material for the visualization generation algorithms. The visualization generation algorithm itself was also challenging to design and implement, given the "noisiness" of parsed Web content. For example, we had to forgo sentence-based calculation of query occurrence because we could not reliably chunk the content into full sentences. Instead, an approach was adopted whereby the occurrences of the other query terms were computed around the instances of the most frequently occurring query term by means of a fixed 20-word boundary.

Results and Discussion

We obtained methodology-related findings as well as results connected with the design of the query occurrence visualization. The user study was designed in the longstanding tradition of experiments that compare a novel interface design to a baseline interface by employing objective and subjective metrics in a laboratory environment. The objective and subjective results gained from the user study were mixed. We attempted to simulate a mobile context of use by introducing a 60-second cap on the task completion time. This was, in practical terms, reached in all conditions (with an average of around 56–57 seconds). It therefore is likely that the participants utilized all of the available time for making decisions, rendering time-based performance metrics relatively useless. Additionally, no significant differences were found in metrics targeting the users' ability to discriminate among non-relevant, related, and relevant search results.

One way to interpret this result is to deduce that the visualization neither hindered nor helped in result evaluation. Another, perhaps more plausible interpretation is that a method wherein users are instructed to select relevant search results on the basis of "canned" search tasks and search results sets is too coarse an instrument to tease out the potential benefits offered by an interface technique such as our visualization. The study setting was too artificial to provide insights into the users' real-world search strategies and the added value our visualization could provide. A more ecologically valid approach would have been to carry out a longitudinal study wherein users utilize the visualization in their real search tasks over several weeks, to gain an appreciation for the frequency of incidents in which the visualization could potentially provide benefit (e.g., when the search fails to provide result descriptors that on their own form a good enough basis for the result selection) and the users' perception of whether the visualization was useful or not.

6.2 Visualizing Query Occurrence in Mobile Web Search Interfaces

Reference

Heimonen, T., & Siirtola, H. (2009). Visualizing query occurrence in mobile Web search interfaces. In *Proceedings of the 13th International Conference on Information Visualisation, IV '09* (pp. 639–644). Washington, DC, USA: IEEE Computer Society. doi:10.1109/IV.2009.16 (Paper II)

Objective

The aim of the next research was to study the applicability of the query occurrence visualization approach for augmenting mobile Web search interfaces. In addition, we examined different variations of the visualization, to ascertain which design the users preferred.

Mobile Web browsing and search differs from its desktop counterpart in many ways, from the types of search topics to the length of queries and the strategies used. In mobile Web search, it is not as easy to utilize efficient strategies, such as opening multiple results in tabs, as it is on the desktop. Also, interaction is slower, because of restricted data transfer rates and lack of direct manipulation, although today's touchscreen devices are bridging this gap. Our initial assessment was that the query occurrence visualization could be useful in this context as an additional piece of information on which users can build their result evaluations, especially when the other metadata are inconclusive or lacking.

We implemented a Web-based mobile search interface on top of the Findex platform that incorporated a refined version of the query occurrence visualization proposed in Paper I. In the design of the visualization, we retained the four-part division of the result document but increased its granularity. Two distinct versions were designed, to address different search goals. The graphical version (see Figure 15) caters for getting an overview of the document and could assist with exploratory, informational search goals by, for example, pointing out documents that feature several instances of the query. The text version allows spotting the exact number of query occurrences, potentially enabling quick identification of documents that provide a specific piece of information (as in question answering).

The design alternatives were evaluated in an 18-participant laboratory study with users. We strove as much as possible to retain the ecological validity of the experiment, with the participants using the search interface of a smartphone over a wireless network connection instead of a desktop-based emulator. Also, they were able to utilize their own query terms and search strategies. Instead of using typical performance metrics such as precision and recall, we measured our participants' use of the search interface through a forced-choice paradigm (Heimonen et al., 2008), in which they were made to choose before each task which search interface to use (i.e., a traditional interface or one with the visualization). The main reasons we opted for this approach were the concerns that have been raised in previous studies as to the suitability of objective metrics for evaluation of mobile search interfaces and the lack of correlation between objective and subjective metrics observed in many user studies in other domains (e.g., Hornbæk & Law, 2007).

Results and Discussion

Our results were primarily subjective and focused on the preferences and choices expressed by users presented with the forced-choice paradigm, their responses to interview questions, and other subjective assessment metrics. Our subjects preferred the visualization designs to the traditional search user interface. This preference had a significant effect on interface selections in the forced-choice scenarios, with users who preferred the visualization design also consistently selecting it in the forced-choice tasks when given the choice. Responses to questions eliciting subjective feedback such as "How useful do you consider the visualization?" did not display any significant differences between the visualization designs. Finally, when asked to select representative words to describe their experience with the visualization designs, the participants made more positive selections when describing the graphical design. Both designs garnered primarily positive response from the participants, with no negative words in the top-ranking selections.

The majority of the participants had, however, mixed reactions to the visualization designs. In general, they found the visualizations to be useful in some situations – for example, when the search engine did not provide a text snippet for a result. Unsurprisingly, the participants used the textual result descriptors as the key relevance information when evaluating the results. The visualizations were used in a supporting role to disambiguate the relevance when the title or URL failed to provide a clear enough assessment of relevance. Several participants noted that they would likely find the visualization more useful after getting used to using it in results' evaluation and after gaining confidence in the visualization. Another concern expressed by the participants was the occasional mismatch between the textual descriptors and the visualization. Six participants noted that they found it misleading when the visualization showed no or very few query occurrences for a result that seemed otherwise like a good fit for the query in view of the snippet, title, or URL. This is directly linked to the design premises of the visualization – the query occurrence approach is unlikely to satisfy information needs where the link between the query and results is weak (i.e., the query keywords are a method of acquiring relevant results and are not of interest in themselves).

We evaluated the query occurrence visualization approach in two separate user studies, first to prove the concept and then to integrate it into a realistic mobile Web search interface. While the results of these two studies did not show the approach to be universally useful, there are circumstances wherein it provides utility. It is possible that prolonged use of the visualization would help users to incorporate it better into their result evaluation strategies. Subjective feedback from our subjects helped us identify areas for improvement in both the data extraction algorithm and the visualization interface. For example, one solution for the mismatches between the visualization and the textual descriptors is to show the visualization only when we can identify a query for which added value would be provided by the visualization. Nevertheless, making full analytical use of the query occurrence data requires the user to adopt more elaborate search strategies, which undoubtedly is relevant for only a subset of mobile Web searches. Also, the ability to sort the results on the basis of the query occurrences might provide additional utility (here we echo the work of Hoeber & Yang (2008)).

6.3 Facilitating Mobile Web Search with Automatic Result Categories

References

Heimonen, T., & Käki, M. (2007). Mobile Findex – Supporting mobile Web search with automatic result categories. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '07* (pp. 397–404). New York, NY, USA: ACM. doi:10.1145/1377999.1378045 (Paper III)

Heimonen, T. (2008). Mobile Findex: Facilitating information access in mobile Web search with automatic result clustering. *Advances in Human-Computer Interaction, 2008*, article ID 680640. doi:10.1155/2008/680640 (Paper IV)

Objective

The objective of the study was to evaluate a novel mobile-search user interface concept, Mobile Findex, which utilizes a particular type of result clustering to arrange the search results into overlapping categories. The motivation for the study originated in earlier research carried out in our unit on search result clustering (Käki, 2005a). At the time of the study, in 2006, mobile Web search engines had started becoming more prevalent yet continued to utilize the flat, ranked result list paradigm familiar from desktop search interfaces.

We wanted to combine the clustering paradigm shown to be useful in desktop search interfaces with a simple navigation-based model suitable for mobile use. The goal was to overcome the problems typically associated with browsing the Web on mobile devices, such as excessive vertical scrolling and page-to-page navigation using links. One solution is the use of a category-based interface. Categories are used for quickly drilling down to smaller, focused result sets that are likely to be of interest to the user. Categories also present an overview of the most common keywords within the results, allowing the user to form an estimate of the success of the query at a glance.

We carried out a 16-participant user study with the Mobile Findex prototype to investigate how the proposed concept compared to the ranked result list presentation paradigm. First, we sought to find out whether the clustering interface could facilitate more efficient information seeking than the traditional interface does. Toward that end, we benchmarked Mobile Findex by using metrics such as precision and recall. The search tasks completed by our participants were information seeking tasks with the overall goal of finding results pointing to Web pages that fulfill a specific information need – for example, finding images of celestial bodies or further details about a news event. Additionally, the participants utilized an actual mobile device to carry out their search tasks, in order to increase the ecological validity of the test setting. Second, we wanted to establish whether there were differences in user experience between the category-driven and ranked result list interfaces, actual performance differences notwithstanding. This was done by systematically collecting participants' subjective feedback during the study.

Results and Discussion

Data were collected on the use of the Mobile Findex interface via several, interrelated measurements, among them the speed of task completion, accuracy of result selection, and qualified search speed – or the rate of acquiring relevant results. In this context, speed measures reflect the efficiency of use and accuracy measures the effectiveness of use, whereas subjective measures are related to perceived user experience and satisfaction.

We did not observe statistically significant differences in task completion times. One factor affecting this is the participants effectively hitting a performance ceiling, because the maximum task completion time was limited to three minutes, for simulation of a time-constrained mobile usage scenario. Overall, 53% of tasks were completed in time with Mobile Findex and 64% with the traditional user interface. In terms of accuracy, we were able to observe a significant effect of user interface for four tasks, with the Mobile Findex interface providing higher precision in three tasks and the traditional interface in one. However, a significant difference in recall of relevant results was not observed. Also, a difference in the rate of non-relevant result selections was observed, indicating that our subjects made erroneous selections at a lower rate than when using the traditional search interface.

Subjective feedback was facilitated by two questionnaires and semi-structured interviewing after task completion. The participants' perceived experiences indicate that with Mobile Findex, search results were easier to find, the interface was better suited to the tasks, and things felt more efficient. It is noteworthy that these differences were observed both when the participants rated the interfaces individually and when they were asked to contrast the interfaces against one another.

The conclusions we can draw from the results are not as cut and dried as those established in studies of clustering approaches in desktop Web search. We found that in the mobile context, categories improve search performance in certain situations. We attribute the lack of difference by objective measurements to differing interaction styles. Mobile Findex, being a novel solution, may have encouraged the participants to explore the result set in more detail than the familiar reference interface did. The ease of exploration that categories provide came at the expense of time and overall task performance. Also, in this particular design, the benefits afforded by the proposed category interface were not great enough to overcome the performance penalty incurred through the multiple-view navigation. In terms of overall effectiveness in finding relevant results, it is also possible that seemingly relevant cluster titles may in some cases mislead the users into expecting to find relevant results within. We examined this issue further in a longitudinal study to find out whether and to what extent it negatively affects use when people use the application in their daily information seeking tasks (see Section 6.4, below).

Although the performance measurements are inconclusive, subjective feedback showed that participants preferred the category-based interface for its perceived effectiveness. Despite its design tradeoffs, the proposed interface provides a more convenient and engaging way to browse search results than the page-by-page navigation in the ranked result list. This suggests that the ability to get an overview of the results and being able to filter and narrow the result set actively are more essential elements of user experience than the actual level of search performance is.

Unfortunately, we cannot discuss the utility of categories for the query formulation and reformulation stages of search, on account of the constraints we placed on the experimental procedure. Although categories in the present study did not actively support query formulation, category labels can suggest new query terms. In the follow-up, longitudinal study, we examined whether categories can support query formulation, along with the strategies users adopted for using the categories in their own search tasks.

6.4 How Do Users Search the Mobile Web with a Clustering Interface? A Longitudinal Study

Reference

Heimonen, T. (2012). How do users search the mobile Web with a clustering interface? A longitudinal study. *International Journal of Mobile Human–Computer Interaction*, 4(3), 44–66. doi:10.4018/jmhci.2012070103 (Paper V)

Objective

The aim of the study was to investigate the use of the Mobile Findex clustering interface in naturalistic settings, with participants employing it in their daily information access tasks for a period of four weeks.

Related research on category-based mobile search interfaces, including the previous study investigating the Mobile Findex interface, has been done in laboratory environments with predefined tasks. The objective of this user study was to answer three research questions related to efficacy and use of the clustering interface "in the wild": How is the clustering interface used for information search and result access, what are its benefits and limitations, and how do these findings inform the design of category-based interfaces?

I conducted two longitudinal user studies to these ends: a limited pilot study with eight participants and a follow-up study with 17 participants. The participants' use of the clustering search interface was captured in interaction logs, which were analyzed for the behavior patterns they revealed. The log files were complemented with subjective feedback, gathered through interviews and questionnaires aimed at understanding the participants' user experience and perceptions of the usefulness of the category-based interface.

Results and Discussion

One of the key results of the study is a descriptive model of category-based search interface use for mobile information access. The typical strategy displayed by participants was to utilize the categories first, then switch to the ranked result list if the categories did not satisfy the need. In about a fifth of the cases, the user went straight to the ranked result list, likely judging by the content of the category labels, or from an *a priori* understanding of their information need, that they would find the result there more quickly. About one query in 10 resulted in a reformulation subsequent to presentation of the category list, suggesting that the labels may support narrowing of the query both directly and indirectly.

Subjective feedback directly addressed rationale for switching to other search tools (primarily Google), including a need to find the information quickly (category use imposes some overhead) or the need for content not supported by the prototype. In performance terms, the use of the categories resembles that in desktop search with respect to result click-through and support for explorative search, while retaining essential features of mobile search (short queries and relatively low frequency of searching).

As for the benefits of the clustering approach, our results show that it can be helpful in some situations, such as when one is unsure of how to frame the information need and requires guidance in how to narrow the search, or when the information need itself cannot be expressed with a focused query that provides the desired item within the first few search results. These results confirm the benefits of category-based interfaces cited in earlier research (e.g., Käki, 2005b). One must remember also that the usefulness of the clustering interface is limited and influenced by the information needs and the descriptiveness of the cluster labels. In the present study, the response to clusters produced by the content-based algorithm was mixed.

While the clustering approach based on frequent content keywords was easily understandable, it did not mesh with the needs of several participants, which indicates a necessity to consider alternative methods of category construction. Several participants stated a preference for more intelligent category representations based on contextual understanding of the query. This kind of category creation should account for the human-generated concepts related to the user's intent. Hybrid approaches that include both clustering and human-generated taxonomies, and that consider other contextual cues (e.g., cues inferred from the query or the user's location), are likely to be beneficial across a wider variety of mobile information needs than are clustering methods focused only on the textual content of the result documents.

The results, especially subjective feedback from the participants, also suggest several design considerations for the presentation and interaction in category-based interfaces. Different strategies for presenting the categories and search results should be considered, moving away from the categories-first design rationale exhibited in earlier research. Providing the categories on demand, or in an interface that combines top-ranking results and categories, would facilitate both navigation- and lookup-oriented query intents and explorative search. Additionally, the presentation of the categories should be aimed at providing more descriptive category representations, through, for example, inclusion of content previews.

6.5 Information Needs and Practices of Active Mobile Internet Users

Reference

Heimonen, T. (2009). Information needs and practices of active mobile Internet users. In *Proceedings of the 6th International Conference on Mobile Technology, Applications, and Systems, Mobility '09* (Article 50). New York, NY, USA: ACM. doi:10.1145/1710035.1710085 (Paper VI)

Objective

The focus of the study was on examining the information needs of active mobile Internet users, people who utilize mobile Internet services on their mobile devices on a daily basis. Our work was inspired by similar studies focusing on mobile information needs that have pointed to significant variation in the contexts of mobile Internet use with dependence on the users' needs and activities. We wanted to focus on the situational and behavioral contextual factors affecting information needs and the information access strategies employed by mobile Internet users. Understanding the needs, practices, and expectations of mobile Internet users is essential for solid design of mobile information services.

We conducted a diary study with eight experienced and active mobile Internet users, who documented their mobile information needs over a four-week period in November–December 2008 by using a Web-based data collection tool either directly on the cell phone or by means of a desktop Web browser. The participants were instructed to capture information needs whose satisfaction was important enough for dedication of a non-trivial amount of time in a mobile context. However, the participants were also asked to report their mobile information needs irrespective of whether they sought to address them as they occurred. The information we gathered included the time, location, and content of the information need, as well as details about the information access method utilized to satisfy the information need and whether the information seeking activity was successful.

Our primary interest lay in studying the effect of a constantly available mobile Internet connection on mobile information needs and practices, and, speaking to a goal derived from that, we attempted to identify characteristics of users' mobile information access activity, or inactivity, that could inform design.

Results and Discussion

Overall, our results underline the importance of mobile Internet services for modern, everyday information access. The participants used mobile Web access and personal information management applications as the main tools when approaching their mobile information needs, regardless of location or context of use.

The findings of our study provide insights into two aspects of mobile information access. First, they indicate that, while experienced mobile Internet users have, in terms of topic, information needs similar to those of less active users, they make greater use of search and mobile Web services to meet their information needs *when they emerge*. One significant factor influencing this proclivity for mobile access is the availability of a persistent Internet connection. Having such always-on access enables immediate information lookup, with the majority of needs being addressed as soon as they emerged. Web access strategies were selected that matched the user's intent – informational, hedonic needs were approached via use of Web search, and focused, pragmatic needs were satisfied by means of a known Web site that was likely to provide the desired information efficiently.

We identified two possible approaches to supporting both types of strategies. First, we could use experienced mobile Internet users' knowledge about good mobile services and applications to provide new adaptive services for other users that are based on shared recommendations and usage trends. Other social features, such as question answering and sharing of location-based search queries, could respond to many of the location-related information needs our subjects faced – for example, finding out about services thanks to the searches and recommendations of other users. Second, the information available on the mobile devices and the context of use could be used to facilitate information access. Several informational needs emerged in social contexts during the study, and addressing these needs while one is engaged in social interaction requires a shift of attention away from the primary task. One solution might be integration of multimodal input and output options, to allow for more automated and laid-back background information retrieval. For example, the application could track ongoing discussion in order to identify upcoming information needs and anticipate them through automated content retrieval, either in the form of suggested query topics or through actual Web content.

7 Discussion

Mobile information access is going through an exciting evolutionary process from "classic" search paradigms that were concerned with optimizing the use of the limited screen space and interactions to new approaches aimed at tackling the contextually rich information needs of today's mobile users. These needs do not comfortably conform to the task-based information seeking frameworks that have been constructed in the desktop domain, often through observation of knowledge workers accessing non-Web information services. Nevertheless, certain concepts and notions can be translated into the mobile domain. Mobile search continues to be a sensemaking activity (Russell et al., 1993), but it requires new kinds of representations that enable searchers to extract the information they require quickly and cost-effectively. The key challenge for mobile search in this respect is in anticipating users' information needs and providing the appropriate representations in varied context. Local search activities are likely to require a certain kind of representation that differs from representations supporting general information search, which, in turn, differ from representations for casual and social search. The need for representations may even differ within the broader sensemaking task – finding out about local, place-specific queries requires different representations than interacting with the questions and answers provided by other users. Similarly, simple navigational Web search is often well served by the ranked result list, but that presentation may fail when used in attempts to address a vague, ill-defined information need.

Many of the information seeking strategies identified in previous research are surprisingly effective in characterizing how people manage their search. For example, the present author identified what can be construed as orienteering activities (Teevan et al., 2004) when people searched the Web via a clustering interface – leveraging the ability of the clusters to suggest query terms and disambiguate results for incrementally approaching the search goal (Paper V). Browsing and searching are characteristic too of much interaction on the mobile Web. In mobile search, the browsing and searching activities intertwine, with search being used to identify sites of interest, which are subsequently browsed (Church et al., 2007). However, searching appears to be more difficult when the user is mobile, as evidenced by the high degree of query modification (Church et al., 2007; Kamvar & Baluja, 2006). Whereas browsing may be preferred in the desktop environment on cognitive grounds – it is a less demanding task than search – the difficulty of entering text in mobile information access adds an additional hurdle. This may encourage users to engage in browsing to relocate familiar resources rather than attempt search. Several category-based methods could be useful in this context, providing powerful browsing and filtering methods to access results and thereby minimizing the need to enter keyword queries. Results from naturalistic clustering interface evaluations suggest that users can learn to utilize categories in this fashion (Käki, 2005b; Paper V). Also, new methods based on voice input may alleviate the concerns with text entry, enabling people to describe their information needs in spoken form; however, the limitations of automated speech recognition in public spaces pose some challenges here. Additionally, advances in context-aware query recommendation and expansion obviate the problems in the case of popular queries, and there is evidence that context can be successfully used to support query construction (Arias et al., 2008; Kamvar & Baluja, 2007b). The improved usability of text entry with newer touchscreen devices may also make querying easier, helping bridge the gap in search behavior between non-mobile computers and cell phones.

These findings present interesting possibilities for advanced forms of search result presentation. In its simplest form, integrating browsing with search by providing category overviews can facilitate information access because users can navigate and filter the information space rather than attempting to find the perfect query for zeroing in on the desired information. On the other hand, improved interaction capabilities and increased display resolution allow for the creation of powerful category-based interfaces that can be used for more complex information gathering. For example, local search activities that require the searcher to investigate, filter, and compare information are currently lacking interface support (Amin et al., 2009). Various category-based approaches, especially faceted browsing, could be useful here. Selection of the appropriate category structures remains a key challenge (Paper V), and more research is needed to investigate whether this could be done automatically through classification of the query's intent (Carpineto et al., 2012).

Indeed, another dimension of information seeking is related to the informational intents of the searchers. Existing classifications of search goals, such as the informational–navigational–transactional distinction of Broder (2002), have been used to characterize mobile searches (e.g., Church et al., 2008); however, the contextual nature of mobile information needs suggests that additional layers are needed if one is to capture the richness of mobile information seeking. There is convincing evidence from a number of studies, the present author's work included, that context shapes the emergence of information needs and also the methods people use to address their needs. Mobile Internet use, involving both direct Web interfaces and other applications, is increasingly used to find information online in response to habits, for social inclusion, to fill time, and to respond to time-critical information needs. Mobile Web search is a key activity in the last category especially, although the role of applications here appears to be increasing. For niche information needs and the like, applications with semi-curated content could well be more useful than general information search.

The Future of Mobile Search

From the research, two distinctions that will shape the future of mobile search stand out: mobile search has a strong geographical focus (related to the users' location, route, or destination), and it is often a social activity, with information needs arising from shared activities and discussions. Enabling focused search related to the current location of the user is already relatively well covered by existing search engines and applications. For example, on the Apple iPhone, a Google Places search for cafés in Helsinki brings up a map of local services, which can be used to launch the on-device Maps application, for positioning and route guidance. Similarly, the Finnish Eat.fi local search application, targeted at finding restaurants in major cities, provides search, user positioning, and filtering with faceted categories. In addition to GPS-based navigation applications that enable directions for driving, various journey planner applications and Web services exist for planning one's route on public transportation. Perhaps the next step in this direction is to study how these local search activities could be supported in shared search settings, given that much of local searching takes places in social situations (Teevan et al., 2011).

Information needs arising from social situations and the user's activities are harder to address with a specific application or service. Much of the research into social mobile search has focused on facilitating information exchange between users by sharing other searchers' queries or answers to questions in various ways (e.g., Arter et al., 2007; Church et al., 2010b). Commercial search engines have also embraced social searching, and both Microsoft's Bing and Google can now provide results from a user's social networks. Another alternative is to reverse the situation in a sense and provide search results based on the social interactions of the users (e.g., information needs extracted from discussions in instant messaging and social media conversations). For example, Hecht, Teevan, Morris, and Liebling (2012) introduced the SearchBuddies search agent, which responds to questions posed in Facebook status updates. It is able to provide the users with useful information in the context of social media discussions. Activity-based information needs, on the other hand, possibly could be supported by anticipation of information needs with Internet content pre-caching and searching based on contextual information, such as personal content stored on devices (Jang, Kim, Shin, & Myaeng, 2010; Komninos & Dunlop, 2008) and various sensor values recorded by the mobile device (Coppola et al., 2010).

In addition to considering information access from the Web and dedicated applications, future mobile search will likely have to concern itself increasingly with how we can manage the myriad interactive, Internet-enabled devices embedded in our everyday environments. Jenson (2011) argues that a discovery service allowing users to open and interact with smart devices in their surroundings will be "the next Google in a few years['] time."

Reflections on the Contributions of the Dissertation

This dissertation provides contributions in three categories. First, papers I and II present the query occurrence visualization concept and its iterations, along with the results from user studies that describe their benefits and drawbacks for search result visualization, and situations wherein it is likely to be of use (e.g., for elimination of non-relevant results from consideration). Second, papers III, IV, and V present the Mobile Findex search clustering concept, its iterations, and evaluations in both laboratory and longitudinal field studies. The benefits and drawbacks of category-based searching were identified, which led, in turn, to practical design guidelines. The guidelines suggest that we need to consider how and when to integrate the categories into the search engine. There is also a need to improve the category construction process so as to account well for both the need for intelligent topical understanding of the information need and the contextual nature of the information needs. Finally, Paper VI reported on the results of a diary study that investigated mobile information needs and the methods active Internet users utilize to address them. The results contribute to our understanding of mobile information access, and they underline the importance of context as a trigger for information needs and as affecting how the needs are addressed.

The user evaluations reported upon in papers I–IV all utilized an empirical evaluation paradigm wherein the ecological validity of the setting was to a lesser or greater extent constrained to enable the comparison of the novel interface solutions to baseline systems. In retrospect, it has become clear that such study methods have not been very useful in terms of forming an understanding of how the proposed interface solutions match real users' information needs or fit into the larger search process. It is unrealistic to expect, unless one has come up with a revolutionary design breakthrough, that people would immediately respond to the novel system with adoption, given their familiarity with existing solutions such as the ranked result list. Indeed, I very much share the perspective of G. Ellis and Dix (2006), who argue for more explorative evaluations of information tools, which focus on finding the gaps in the design rationale of the proposed system. Often you can gain more insight from examining the failures of a proposed system than when fixated on validation of the design over an alternative. Further, in examination of activities such as casual and exploratory searching, it is unlikely that existing performance, relevance, and satisfaction metrics fully capture the user's experience. For example, it may be argued that in casual searching spending more time on the search process could indicate increased immersion, rather than problems.

Additionally, the search interface or search result visualization is always a particular implementation of the underlying concept or idea. The justification for the concepts, such as the notion that clusters can assist in sharpening vague queries and understanding the topics present in the results, may be entirely valid while problems with the implementation confound the results of the study. In view of the experience gained from conducting several controlled experiments; laboratory-based usability studies of mobile applications; and most recently the longitudinal, ethnographic field studies, I would argue that any serious attempts at understanding mobile searching and the impact of new interfaces should be conducted *in situ*. On the other hand, laboratory studies, expert reviews, and controlled experiments are instrumental in ensuring that the usability of interfaces and the parameters of the underlying system (e.g., data-to-visualization mappings and clustering algorithms) are good enough for a real deployment.

Limitations

All of the user studies reported upon herein have been relatively limited in their number of participants. This is a fundamental limitation of smaller-scale studies that are carried out in resource-constrained settings. Thus, generalizing the results beyond the particular users and tasks is understandably challenging, which is a typical concern with such studies. However, most of the results obtained in the studies appear to confirm the salient results from previous research, while also addressing some new research questions. Indeed, the number of citations of the research included in this dissertation and the treatment of that research indicate that subsequent work has found the results informative.

It would be interesting to address many of the research questions further that have been studied with small user studies here, in larger-scale, longitudinal evaluations. For example, would continued exposure and learning increase the usefulness of the query occurrence visualization? Would the addition of different, context-sensitive categorization methods alongside clustering improve category-based mobile search? Therefore, my future research endeavors in the field of mobile information access will aim at evaluating novel search interface features "in the wild" – with specific focus on the understanding and use of contextual information.

8 Conclusion

The research reported upon in this thesis had two objectives. The first objective was concerned with designing, implementing, and evaluating new interface solutions to support the mobile search process. This was done by developing new search interface solutions grounded in the findings of previous research, and by evaluating them in a systematic manner, to understand their benefits and drawbacks.

A new presentation method called query occurrence visualization was introduced to complement the ranked result list. The visualization was evaluated in two controlled experiments and found to be beneficial for results' evaluation in lieu of the traditional result descriptors. In addition, a clustering search interface called Mobile Findex was introduced to assist in the evaluation of the whole result set returned by the search engine. The interface was evaluated in two user studies, both in the laboratory and in the field. The clustering interface was found to be situationally useful in situations wherein users had problems coming up with good query terms or required a better overview of the search results than is afforded by the ranked result list. The user studies revealed new research questions for both approaches and, in the case of the clustering interface, also design implications for category-based mobile search interfaces. In addition to considering how and when the categories are presented in the search user interface, there is need for context-aware category construction methods that better support mobile searchers.

The second objective was to study how mobile Internet access is used for information seeking, in order to situate the findings from the constructive research in the context of naturalistic mobile information access. A longitudinal diary study was carried out to understand the information needs and the role of search and browsing as methods for satisfying these needs. The results showed that active mobile Internet users rely heavily on

Internet access to resolve their information needs as they arise, with mobile Web search being the dominant method. The results of the study also suggest new avenues for further research – for example, how to take into account the context of search: the social situations and situated activities in which people engage.

The interface solutions presented in this dissertation represent "classic" mobile search in the sense that they are primarily focused on usability and performance: on making the presentation and interaction with search results more efficient within the constraints of the limited display space and interaction modalities of mobile devices. Future mobile search interfaces must also embrace and support the diversity of mobile information needs. These needs range from contextual, often location-dependent and urgent queries to informational, sometimes casual needs that arise from activities and social interactions. More research is needed for understanding how context affects information needs in realistic information access scenarios, and how users' understanding of context could be used to support the process of finding relevant information. It would be too easy to subscribe to the notion that rich capture of context from device sensors is a fix for these problems. Identifying the context (e.g., the user is running because a meeting ran late or is sitting on a bus) does not solve anything in itself, because the context has to be meaningfully mapped into some kind of system activity that can support the user's information needs in the particular situation faced. This is, however, a general challenge in context-aware computing and not specific to mobile information seeking. Research in this field will undoubtedly advance in the coming years.

Future mobile search interfaces must also support new kinds of exploratory and casual search experiences that emphasize insight over finding, and the process of (and user experience of) searching rather than the end results. Category-based interfaces may be useful in bridging the old and the new. They provide effective overviews of search results by showing the themes emergent in the results and enable their exploration. This could be beneficial in several context-awareness scenarios, such as for understanding the nature of places and topics of local queries in social search, or facilitating serendipitous browsing of the search results when one is engaged in casual information seeking.

9 Bibliography

Amin, A., Townsend, S., Ossenbruggen, J., & Hardman, L. (2009). Fancy a drink in Canary Wharf?: A user study on location-based mobile search. In T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R. O. Prates, & M. Winckler (Eds.), *Human-Computer Interaction – Interact 2009: Proceedings of the 12th IFIP TC 13 International Conference, Part I, LNCS 5726* (pp. 736–749). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-642-03655-2_80

Anderson, T. J., Hussam, A., Plummer, B., & Jacobs, N. (2002). Pie charts for visualizing query term frequency in search results. In E.-P. Lim, S. Foo, C. Khoo, H. Chen, E. Fox, S. Urs, & T. Costantino (Eds.), *Digital Libraries: People, Knowledge, and Technology: Proceedings of the 5th International Conference on Asian Digital Libraries, ICADL 2002, LNCS 2555* (pp. 440–451). Berlin Heidelberg, Germany: Springer. doi:10.1007/3-540-36227-4_52

Andrews, K., Sabol, V., Lackner, W., Gütl, C., & Moser, J. (2001). Search result visualisation with xFIND. In *Proceedings of the Second International Workshop on User Interfaces to Data Intensive Systems, UIDIS 2001* (pp. 50–58). Washington, DC, USA: IEEE Computer Society. doi:10.1109/UIDIS.2001.929925

Andrews, K. (2008, April). Evaluation comes in many guises. In *Beyond Time and Errors: Novel Evaluation Methods for Information Visualization.* Workshop organized in conjunction with CHI '08, Florence, Italy. Retrieved from http://www.dis.uniroma1.it/beliv08/pospap/andrews.pdf

Arias, M., Cantera, J. M., Vegas, J., de la Fuente, P., Alonso, J. C., Bernardo, G. G., Llamas, C., & Zubizarreta, Á. (2008). Context-based personalization for mobile Web search. *Proceedings of PersDB2008 – 2nd International Workshop on Personalized Access, Profile Management, and Context Awareness: Databases* (pp. 33–39). Retrieved from http://persdb08.stanford.edu/PersDB08_proc.pdf

Arter, D., Buchanan, G., Jones, M., & Harper, R. (2007). Incidental information and mobile search. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '07* (pp. 413–420). New York, NY, USA: ACM. doi:10.1145/1377999.1378047

Aula, A. (2005). *Studying user strategies and characteristics for developing web search interfaces* (Doctoral dissertation, Dissertations in Interactive Technology, Number 3, University of Tampere, Finland). Retrieved from http://urn.fi/urn:isbn:951-44-6488-5

Aula, A., Jhaveri, N., & Käki, M. (2005). Information search and re-access strategies of experienced web users. In *Proceedings of the 14th International Conference on World Wide Web, WWW '05* (pp. 583–592). New York, NY, USA: ACM. doi:10.1145/1060745.1060831

Aula, A., Khan, R., Hong, P., Guan, Z., & Fontes, P. (2010). A comparison of visual and textual page previews in judging the helpfulness of web pages. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10* (pp. 51–60). New York, NY, USA: ACM. doi:10.1145/1772690.1772697

Aula, A., & Käki, M. (2003). Understanding expert search strategies for designing user-friendly search interfaces. In P. Isaías & N. Karmakar (Eds.), *Proceedings of IADIS International Conference WWW/Internet 2003, Volume II* (pp. 759–762). IADIS Press.

Aula, A., & Käki, M. (2005). Less is more in web search interfaces for older adults. *First Monday, 10*(7). Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1254/1174

Baeza-Yates, R., Dupret, G., Velasco, J. (2007, May). A study of mobile search queries in Japan. In *Query Log Analysis: Social and Technological Challenges*. Workshop organized in conjunction with WWW '07, Banff, Alberta, Canada. Retrieved from http://www2007.org/workshops/paper_50.pdf

Bar-Ilan, J. (2007, May). Position paper: Access to query Logs – An academic researcher's point of view. In *Query Log Analysis: Social and Technological Challenges*. Workshop organized in conjunction with WWW '07, Banff, Alberta, Canada. Retrieved from http://www2007.org/workshops/paper_39.pdf

Bates, M. J. (1979). Information search tactics. *Journal of the American Society for Information Science, 30*(4), 205–214. doi:10.1002/asi.4630300406

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407–424. doi:10.1108/eb024320

Bates, M. J. (2007). What is browsing – Really? A model drawing from behavioural science research. *Information Research*, *12*(4), paper 330. Retrieved from http://informationr.net/ir/12-4/paper330.html

Benedek, J., & Miner, T. (2002, July). Measuring desirability: New methods for evaluating desirability in a usability lab setting. In *Usability Professionals Association 2002 Annual Conference*. Retrieved from http://www.microsoft.com/usability/uepostings/desirabilitytoolkit.doc

Bharat, K. (2000). SearchPad: Explicit capture of search context to support Web search. *Computer Networks: The International Journal of Computer and Telecommunications Networking, 33*(1–6), 493–501. doi:10.1016/S1389-1286(00)00047-5

Björk, S., Holmquist, L. E., Redström, J., Bretan, I., Danielsson, R., Karlgren, J., & Franzén, K. (1999). WEST: A Web browser for small terminals. In *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology, UIST '99* (pp. 187–196). New York, NY, USA: ACM. doi:10.1145/320719.322601

Brandt, J., Weiss, N., & Klemmer, S. R. (2007). txt 4 l8r: Lowering the burden for diary studies under mobile conditions. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems* (pp. 2303–2308). doi:10.1145/1240866.1240998

Broder, A. (2002). A taxonomy of Web search. *SIGIR Forum, 36*(2), 3–10. doi:10.1145/792550.792552

Brown, B., Reeves, S., & Sherwood, S. (2011). Into the wild: Challenges and opportunities for field trial methods. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, CHI '11* (pp. 1657–1666). New York, NY, USA: ACM. doi:10.1145/1978942.1979185

Buchanan, G., Jones, M., & Marsden, G. (2002). Exploring small screen digital library access with the Greenstone digital library. In M. Agosti, & C. Thanos (Eds.), *Research and Advanced Technology for Digital Libraries: Proceedings of the 6th European Conference on, ECDL 2002, LNCS 2458* (pp. 583–596). Berlin Heidelberg, Germany: Springer. doi:10.1007/3-540-45747-X_44

Buyukkokten, O., Garcia-Molina, H., Paepcke, A., & Winograd, T. (2000). Power Browser: Efficient Web browsing for PDAs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '00* (pp. 430–437). New York, NY, USA: ACM. doi:10.1145/332040.332470

Byrd, D. (1999). A scrollbar-based visualization for document navigation. In *Proceedings of the Fourth ACM International Conference on Digital Libraries, DL '99* (pp. 122–129). New York, NY, USA: ACM. doi:10.1145/313238.313283

Campbell, I. (1995). Supporting information needs by ostensive definition in an adaptive information space. In I. Ruthven (Ed.), *Proceedings of the Final Conference on Multimedia Information Retrieval* (pp. 1–25). British Computer Society. Retrieved from http://www.bcs.org/upload/pdf/ewic_mi95_paper3.pdf

Capra, R. G., & Marchionini, G. (2008). The relation browser tool for faceted exploratory search. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '08* (pp. 420–420). New York, NY, USA: ACM. doi:10.1145/1378889.1378967

Capra, R., Marchionini, G., Oh, J. S., Stutzman, F., & Zhang, Y. (2007). Effects of structure and interaction style on distinct search tasks. In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '07* (pp. 442–451). New York, NY, USA: ACM. doi:10.1145/1255175.1255267

Card, S. (2003). Information visualization. In A. Sears & J. A. Jacko (Eds.), *The human–computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Card, S., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. Morgan Kaufmann.

Carmel, D., Roitman, H., & Zwerdling, N. (2009). Enhancing cluster labeling using Wikipedia. In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '09* (pp. 139–146). New York, NY, USA: ACM. doi:10.1145/1571941.1571967

Carpineto, C., D'Amico, M., & Romano, G. (2012). Evaluating subtopic retrieval methods: Clustering versus diversification of search results. *Information Processing & Management*, 48(2), 358–373. doi:10.1016/j.ipm.2011.08.004

Carpineto, C., Della Pietra, A., Mizzaro, S., & Romano, G. (2006). Mobile clustering engine. In M. Lalmas, A. MacFarlane, S. Rüger, A. Tombros, T. Tsikrika, & A. Yavlinsky (Eds.), *Advances in Information Retrieval:*

Proceedings of the 28th European Conference on IR Research, ECIR 2006, LNCS 3936 (pp. 155–166). Berlin Heidelberg, Germany: Springer. doi:10.1007/11735106_15

Carpineto, C., Mizzaro, S., Romano, G., & Snidero, M. (2009). Mobile information retrieval with search results clustering: Prototypes and evaluations. *Journal of the American Society for Information Science and Technology, 60*(5), 877–895. doi:10.1002/asi.21036

Carpineto, C., Osínski, S., Romano, G., & Weiss, D. (2009). A survey of Web clustering engines. *ACM Computing Surveys, 41*(3), Article 17. doi:10.1145/1541880.1541884

Chae, M., & Kim, J. (2003). An empirical study on the breadth and depth tradeoffs in very small screens: Focusing on mobile Internet phones. *Proceedings of 9th Americas Conference on Information Systems, AMCIS 2009* (Paper 268). Association for Information Systems.

Chan, D. L., Luk, R. W., Leong, H. V., & Ho, E. K. (2009). Discovering user interface requirements of search results for mobile clients by contextual inquiry. In G. Salvendy & M. J. Smith (Eds.), *Human Interface and the Management of Information. Information and Interaction: Proceedings of the Symposium on Human Interface 2009, Part II, LNCS 5618* (pp. 365–374). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-642-02559-4_40

Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications, 25*(4), 12–16. doi:10.1109/MCG.2005.91

Chen, C., & Yu, Y. (2000). Empirical studies of information visualization: A meta-analysis. *International Journal of Human-Computer Studies, 53*(5), 851–866. doi:10.1006/ijhc.2000.0422

Chen, H. [Hao], & Dumais, S. (2000). Bringing order to the Web: Automatically categorizing search results. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems, CHI '00* (pp. 145–152). New York, NY, USA: ACM. doi:10.1145/332040.332418

Chen, H. [Hsinchun], Houston, A. L., Sewell, R. R., & Schatz, B. R. (1998). Internet browsing and searching: User evaluations of category map and concept space techniques. *Journal of the American Society for Information Science, 49*(7), 582–608. doi:10.1002/(SICI)1097-4571(19980515)49:7<582::AID-ASI2>3.0.CO;2-X

Chen, L., & Qi, L. (2010). A diary study of understanding contextual information needs during leisure traveling. In *Proceedings of the Third Symposium on Information Interaction in Context, IIiX '10* (pp. 265–270). New York, NY, USA: ACM. doi:10.1145/1840784.1840823

Chen, M., Hearst, M., Hong, J., & Lin, J. (1999). Cha-Cha: A system for organizing intranet search results. In *Proceedings of the 2nd Conference on USENIX Symposium on Internet Technologies and Systems – Volume 2*. Berkeley, CA: USENIX Association.

Cherubini, M., & Oliver, N. (2009, April). A refined experience sampling method to capture mobile user experience. In *Mobile User Experience Research: Challenges, Methods & Tools*. Workshop organized in conjunction with CHI '09, Boston, MA. Retrieved from https://sites.google.com/site/chi09mobileworkshop/papers-upload/06_ExperienceSampling.pdf

Chi, E. H., Pirolli, P., Chen, K., & Pitkow, J. (2001). Using information scent to model user information needs and actions and the Web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '01* (pp. 490–497). New York, NY, USA: ACM. doi:10.1145/365024.365325

Chi, E. H., & Riedl, J. (1998). An operator interaction framework for visualization systems. In *Proceedings of the 1998 IEEE Symposium on Information Visualization, INFOVIS '98* (pp. 63–70). Washington, DC, USA: IEEE Computer Society. doi:10.1109/INFVIS.1998.729560

Choo, C. W., Detlor, B., & Turnbull, D. (2000). Information seeking on the Web: An integrated model of browsing and searching. *First Monday, 5*(2). Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/729/638

Chua, A. Y. K., Balkunje, R. S., & Goh, D. H.-L. (2011). Fulfilling mobile information needs: a study on the use of mobile phones. In *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication, ICUIMC '11* (Article 92). New York, NY, USA: ACM. doi:10.1145/1968613.1968721

Church, K., & Cherubini, M. (2010). Evaluating mobile user experience in-the-wild: Prototypes, playgrounds and contextual experience sampling. In *Research in the Large: Using App Stores, Markets and Other Wide Distribution Channels in Ubiquitous Computing Research*. Workshop organized in conjunction with UbiComp 2010, Copenhagen, Denmark. Retrieved from http://large.mobilelifecentre.org/dotclear/public/Church-in-the-large.pdf

Church, K., Keane, M. T., & Smyth, B. (2005). Towards more intelligent mobile search. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI'05* (pp. 1675–1677). San Francisco, CA: Morgan Kaufmann.

Church, K., Neumann, J., Cherubini, M., & Oliver, N. (2010a). The "Map Trap"?: An evaluation of map versus text-based interfaces for location-based mobile search services. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10* (pp. 261–270). New York, NY, USA: ACM. doi:10.1145/1772690.1772718

Church, K., Neumann, J., Cherubini M., & Oliver, N. (2010b). SocialSearchBrowser: A novel mobile search and information discovery tool. In *Proceedings of the 14th International Conference on Intelligent User Interfaces, IUI '10* (pp. 101–110). New York, NY, USA: ACM. doi:10.1145/1719970.1719985

Church, K., & Oliver, N. (2011). Understanding mobile Web and mobile search use in today's dynamic mobile landscape. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '11* (pp. 67–76). New York, NY, USA: ACM. doi:10.1145/2037373.2037385

Church, K., & Smyth, B. (2009). Understanding the intent behind mobile information needs. In *Proceedings of the 13th International Conference on Intelligent User Interfaces, IUI '09* (pp. 247–256). New York, NY, USA: ACM. doi:10.1145/1502650.1502686

Church, K., Smyth, B., Bradley, K., & Cotter, P. (2008). A large scale study of European mobile search behaviour. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '08* (pp. 13–22). New York, NY, USA: ACM. doi:10.1145/1409240.1409243

Church, K., Smyth, B., Cotter, P., & Bradley, K. (2007). Mobile information access: A study of emerging search behavior on the mobile Internet. *ACM Transactions on the Web, 1*(1), Article 4. doi:10.1145/1232722.1232726

Church, K., Smyth, B., & Keane, M. T. (2006). Evaluating interfaces for intelligent mobile search. In *Proceedings of the 2006 International Cross-disciplinary Workshop on Web Accessibility: Building the Mobile Web: Rediscovering Accessibility?* (pp. 69–79). New York, NY, USA: ACM. doi:10.1145/1133219.1133232

Clarkson, E., Desai, K., & Foley, J. (2009). ResultMaps: Visualization for search interfaces. *IEEE Transactions on Visualization and Computer Graphics, (15)*6, 1057–1064. doi:10.1109/TVCG.2009.176

Cleverdon, C. W. (1970). The effect of variations in relevance assessments in comparative experimental tests of index languages. *Cranfield Library Report No. 3*. Cranfield, UK: Cranfield Institute of Technology.

Consolvo, S., & Walker, M. (2003). Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing, 2*(2), 24–31. doi:10.1109/MPRV.2003.1203750

Coppola, P., Della Mea, V., Di Gaspero, L., Menegon, D., Mischis, D., Mizzaro, S., Scagnetto, I., & Vassena, L. (2010). The context-aware browser. *IEEE Intelligent Systems, 25*(1), 38-47. doi:10.1109/MIS.2010.26

Coyle, M., & Smyth, B. (2005). Explaining search results. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI'05* (pp. 1553–1555). San Francisco, CA: Morgan Kaufmann.

Cui, Y., & Roto, V. (2008). How people use the Web on mobile devices. In *Proceedings of the 17th International Conference on World Wide Web, WWW '08* (pp. 905–914). New York, NY, USA: ACM. doi:10.1145/1367497.1367619

Cutrell, E., & Guan, Z. (2007). What are you looking for?: An eye-tracking study of information usage in Web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07* (pp. 407–416). New York, NY, USA: ACM. doi:10.1145/1240624.1240690

Cutting, D. R., Karger, D. R., Pedersen, J. O., & Tukey, J. W. (1992). Scatter/Gather: a cluster-based approach to browsing large document collections. In N. Belkin, P. Ingwersen, & A. M. Pejtersen (Eds.), *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '92* (pp. 318–329). New York, NY, USA: ACM. doi:10.1145/133160.133214

Czerwinski, M. P., van Dantzich, M., Robertson, G.G., & Hoffman, H. (1999). The contribution of thumbnail image, mouse-over text and spatial location memory to Web page retrieval in 3D. In M. A. Sasse, & C. Johnson (Eds.), *Human-Computer Interaction, INTERACT '99: IFIP TC.13 International Conference on Human-Computer Interaction* (pp. 163–170). Amsterdam, The Netherlands: IOS Press.

Dachselt, R., & Frisch, M. (2007). Mambo: A facet-based zoomable music browser. In *Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia, MUM '07* (pp. 110–117). New York, NY, USA: ACM. doi:10.1145/1329469.1329484

Dakka, W, & Ipeirotis, P. G. (2008). Automatic extraction of useful facet hierarchies from text databases. In *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering, ICDE '08* (pp. 466–475). Washington, DC, USA: ACM. doi:10.1109/ICDE.2008.4497455

Dearman, D., Kellar, M., & Truong, K. N. (2008). An examination of daily information needs and sharing opportunities. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, CSCW '08* (pp. 679–688). New York, NY, USA: ACM. doi:10.1145/1460563.1460668

De Luca, E. W., & Nürnberger, A. (2005). Supporting information retrieval on mobile devices. In *Proceedings of the 7th International Conference on*

Human Computer Interaction with Mobile Devices & Services, MobileHCI '05 (pp. 347–348). New York, NY, USA: ACM. doi:10.1145/1085777.1085859

Dewey, M., Mitchell, J. S., Beall, J., Green, R., Martin, G., & Panzer, M. (2011). *Dewey decimal classification and relative index*. Dublin, OH: Online Computer Library Center.

Dey, A. K. (2001). Understanding and using context. *Personal and Ubiquitous Computing, 5*(1), 4–7. doi:10.1007/s007790170019

Drori, O., & Alon, N. (2003). Using document classification for displaying search results list. *Journal of Information Science, 29*(2), 97–106. doi:10.1177/016555150302900202

Dumais, S. T., & Belkin, N. J. (2005). The TREC interactive tracks: Putting the user into search. In E. M. Voorhees, & D. K. Harman (Eds.), *TREC: Experiment and Evaluation in Information Retrieval* (pp. 123–153). Cambridge, MA, USA: MIT Press.

Dumais, S., Cutrell, E., & Chen, H. (2001). Optimizing search by showing results in context. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '01* (pp. 277–284). New York, NY, USA: ACM. doi:10.1145/365024.365116

Dziadosz, S., & Chandrasekar, R. (2002). Do thumbnail previews help users make better relevance decisions about Web search results? In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '02* (pp. 365–366). New York, NY, USA: ACM. doi:10.1145/564376.564446

Ellis, D. (1989). A behavioural model for information retrieval system design. *Journal of Information Science, 15*(4-5), 237–247. doi:10.1177/016555158901500406

Ellis, G., & Dix, A. (2006). An explorative analysis of user evaluation studies in information visualisation. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization* (pp. 1–7). New York, NY, USA: ACM. doi:10.1145/1168149.1168152

Ferragina, P., & Gulli, A. (2005). A personalized search engine based on Web-snippet hierarchical clustering. In *Special Interest Tracks and Posters of the 14th International Conference on World Wide Web, WWW '05* (pp. 801–810). New York, NY, USA: ACM. doi:10.1145/1062745.1062760

Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G., & Ruppin, E. (2002). Placing search in context: The concept revisited. *ACM Transactions on Information Systems, 20*(1), 116–131. doi:10.1145/503104.503110

Foster, A. (2004). A nonlinear model of information-seeking behavior. *Journal of the American Society for Information Science and Technology, 55*(3), 228–237. doi:10.1002/asi.10359

Fox, S., Karnawat, K., Mydland, M., Dumais, S., & White, T. (2005). Evaluating implicit measures to improve Web search. *ACM Transactions on Information Systems, 23*(2), 147–168. doi:10.1145/1059981.1059982

Gotz, D. (2007). The ScratchPad: Sensemaking support for the Web. In *Proceedings of the 16th International Conference on World Wide Web, WWW '07* (pp. 1329–1330). New York, NY, USA: ACM. doi:10.1145/1242572.1242834

Hafner, K. (2006, August 23). Researchers yearn to use AOL logs but they hesitate. *The New York Times.* Retrieved from http://www.nytimes.com/2006/08/23/technology/23search.html

Hearst, M. A. (1995). TileBars: Visualization of term distribution information in full text information access. In I. R. Katz, R. Mack, L. Marks, M. B. Rosson, & J. Nielsen (Eds.), *Proceedings of the SIGCHI conference on Human Factors in Computing Systems, CHI '95* (pp. 59–66). New York, NY, USA: ACM Press/Addison-Wesley. doi:10.1145/223904.223912

Hearst, M. A. (1999). User interfaces and visualization. In R. Baeza-Yates, & B. Ribeiro-Neto (Eds.), *Modern information retrieval* (pp. 257–323). New York, NY: ACM Press.

Hearst, M. A. (2006a). Clustering versus faceted categories for information exploration. *Communications of the ACM, 49*(4), 59–61. doi:10.1145/1121949.1121983

Hearst, M. A. (2006b, August). Design recommendations for hierarchical faceted search interfaces. In *SIGIR'2006 Faceted Search Workshop*. Workshop organized in conjunction with SIGIR 2006, Seattle, WA. Retrieved from http://flamenco.berkeley.edu/papers/faceted-workshop06.pdf

Hearst, M. A. (2008, October). UIs for faceted navigation: Recent advances and remaining open problems. In *HCIR'2008: Second Workshop on Human–Computer Interaction and Information Retrieval.* Workshop hosted by Microsoft Research, Redmond, WA. Retrieved from http://flamenco.berkeley.edu/papers/hcir08.pdf

Hearst, M. A. (2009). *Search user interfaces*. New York, NY: Cambridge University Press. Retrieved from http://searchuserinterfaces.com

Hearst, M., Elliott, A., English, J., Sinha, R., Swearingen, K., & Yee, K.-P. (2002). Finding the flow in Web site search. *Communications of the ACM, 45*(9), 42–49. doi:10.1145/567498.567525

Hearst, M. A., & Pedersen, J. O. (1996). Reexamining the cluster hypothesis: Scatter/Gather on retrieval results. In *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '96* (pp. 76–84). New York, NY, USA: ACM. doi:10.1145/243199.243216

Hecht, B., Teevan, J., Morris, M. R., & Liebling, D. (2012). SearchBuddies: Bringing search engines into the conversation. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media, ICWSM-12*. Palo Alto, CA: AAAI Press.

Heer, J., Card, S. K., & Landay, J. A. (2005). prefuse: A toolkit for interactive information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05* (pp. 421–430). New York, NY, USA: ACM. doi:10.1145/1054972.1055031

Heimonen, T., Aula, A., Hutchinson, H., & Granka, L. (2008, April). Comparing the user experience of search user interface designs. In *Now Let's Do it in Practice: User Experience Evaluation Methods in Product Development*. Workshop organized in conjunction with CHI '08, Florence, Italy. Retrieved from http://www.cs.tut.fi/ihte/CHI08_workshop/papers/Heimonen_etal_UXEM_CHI08_06April08.pdf

Heymann, P., Koutrika, G., & Garcia-Molina, H. (2008). Can social bookmarking improve Web search? In *Proceedings of the International Conference on Web Search and Web Data Mining, WSDM '08* (pp. 195–206). New York, NY, USA: ACM. doi:10.1145/1341531.1341558

Hinze, A. M., Chang, C., & Nichols, D. M. (2010). Contextual queries express mobile information needs. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '10* (pp. 327–336). New York, NY, USA: ACM. doi:10.1145/1851600.1851658

Hoeber, O., Schroeder, D., & Brooks, M. (2009). Real-world user evaluations of a visual and interactive Web search interface. In *Proceedings of the 13th International Conference Information Visualisation, IV '09* (pp. 119–126). Washington, DC, USA: IEEE Computer Society. doi:10.1109/IV.2009.20

Hoeber, O., & Yang, X. D. (2006). The visual exploration of Web search results using HotMap. In *Proceedings of the Tenth International Conference on Information Visualization, IV '06* (pp. 157–165). Washington, DC, USA: IEEE Computer Society. doi:10.1109/IV.2006.108

Hoeber, O., & Yang, X. D. (2008). Evaluating WordBars in exploratory Web search scenarios. *Information Processing & Management, 44*(2), 485–510. doi:10.1016/j.ipm.2007.07.003

Hornbæk, K., & Law, E. L. C. (2007). Meta-analysis of correlations among usability measures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07* (pp. 617–626). New York, NY, USA: ACM. doi:10.1145/1240624.1240722

Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006). Information retrieval in folksonomies: Search and ranking. In Y. Sure, & J. Domingue (Eds.), *The Semantic Web: Research and Applications: Proceedings of the 3th European Semantic Web Conference, ESWC 2006, LNCS 4011* (pp. 411–426). Berlin Heidelberg, Germany: Springer. doi:10.1007/11762256_31

Hutchinson H. B. (2005). *Children's interface design for searching and browsing* (Doctoral dissertation, University of Maryland, College Park). Retrieved from http://hcil2.cs.umd.edu/trs/2005-32/2005-32.htm

Huynh, D. F., Karger, D. R., & Miller, R. C. (2007). Exhibit: Lightweight structured data publishing. In *Proceedings of the 16th International Conference on World Wide Web, WWW '07* (pp. 737–746). New York, NY, USA: ACM. doi:10.1145/1242572.1242672

Hölscher, C., & Strube, G. (2000). Web search behavior of Internet experts and newbies. *Computer Networks: The International Journal of Computer and Telecommunications Networking, 33*(1–6), 337–346. doi:10.1016/S1389-1286(00)00031-1

Jang, H., Kim, S., Shin, W., & Myaeng, S.-H. (2010). Personal information access using proactive search and mobile hypertext. *IEEE Intelligent Systems, 25*(1), 27–36. doi:10.1109/MIS.2010.24

Jansen, B. J., & Booth, D. (2010). Classifying Web queries by topic and user intent. In *CHI 2010 Extended Abstracts on Human factors in Computing Systems* (pp. 4285–4290). New York, NY, USA: ACM. doi:10.1145/1753846.1754140

Jansen, B. J., Booth, D. L., & Spink, A. (2007). Determining the user intent of Web search engine queries. In *Proceedings of the 16th International Conference on World Wide Web, WWW '07* (pp. 1149–1150). New York, NY, USA: ACM. doi:10.1145/1242572.1242739

Jansen, B. J., Booth, D. L., & Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Information Processing & Management*, 44(3), 1251–1266. doi:10.1016/j.ipm.2007.07.015

Jenson, S. (2011). *Mobile apps must die*. Retrieved from http://designmind.frogdesign.com/blog/mobile-apps-must-die.html

Jones, M. (2011). Classic and alternative mobile search: A review and agenda. *International Journal of Mobile Human Computer Interaction, 3*(1), 22–36. doi:10.4018/jmhci.2011010102

Jones, M., Buchanan, G., Cheng, T.-C., & Jain, P. (2006). Changing the pace of search: Supporting "background" information seeking. *Journal of the American Society for Information Science and Technology*, *57*(6), 838–842. doi:10.1002/asi.20304

Jones, M., Buchanan, G., Harper, R., & Xech, P.-L. (2007). Questions not answers: A novel mobile search technique. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07* (pp. 155–158). New York, NY, USA: ACM. doi:10.1145/1240624.1240648

Jones, M., Buchanan, G., & Mohd-Nasir, N. (1999). An evaluation of WebTwig — A site outliner for handheld Web access. In H.-W. Gellerson (Ed.), *Handheld and Ubiquitous Computing: Proceedings of the First International Symposium, HUC'99, LNCS 1707* (pp. 343–345). Berlin Heidelberg, Germany: Springer. doi:10.1007/3-540-48157-5_41

Jones, M., Buchanan, G., & Thimbleby, H. W. (2002). Sorting out searching on small screen devices. In F. Paternó (Ed.), *Human Computer Interaction with Mobile Devices: Proceedings of the 4th International Symposium, Mobile HCI 2002, LNCS 2411* (pp. 81–94). Berlin Heidelberg, Germany: Springer. doi:10.1007/3-540-45756-9_8

Jones, M., & Marsden, G. (2006). *Mobile interaction design*. John Wiley & Sons.

Jones, S., Jones, M., & Deo, S. (2004). Using keyphrases as search result surrogates on small screen devices. *Personal and Ubiquitous Computing, 8*(1), 55–68. doi:10.1007/s00779-004-0258-y

Julien, C.-A., Leide, J. E., & Bouthillier, F. (2008). Controlled user evaluations of information visualization interfaces for text retrieval: Literature review and meta-analysis. *Journal of the American Society for Information Science and Technology, 59*(6), 1012–1024. doi:10.1002/asi.20786

Järvelin, K., & Ingwersen, P. (2004). Information seeking research needs extension towards tasks and technology. *Information Research, 10*(1), paper 212. Retrieved from http://informationr.net/ir/10-1/paper212.html

Kaasinen, E. (2003). User needs for location-aware mobile services. *Personal and Ubiquitous Computing, 7*(1), 70–79. doi:10.1007/s00779-002-0214-7

Kaikkonen, A. (2011). Mobile Internet, Internet on mobiles or just Internet you access with variety of devices? In *Proceedings of the 23rd Australian Computer-Human Interaction Conference, OzCHI '11* (pp. 173–176). New York, NY, USA: ACM. doi:10.1145/2071536.2071563

Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A., & Cankar, M. (2005). Usability testing of mobile applications: A comparison between laboratory and field testing. *Journal of Usability Studies, 1*(1), 4–16. Retrieved from http://www.upassoc.org/upa_publications/jus/2005_november/mobile.html

Kammerer, Y., Nairn, R., Pirolli, P., & Chi, E. H. (2009). Signpost from the masses: Learning effects in an exploratory social tag search browser. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI '09* (pp. 625–634). New York, NY, USA: ACM. doi:10.1145/1518701.1518797

Kampanya, N., Shen, R., Kim, S., North, C., & Fox, E. A. (2004). Citiviz: A visual user interface to the CITIDEL system. In R. Heery & L. Lyon (Eds.), *Research and Advanced Technology for Digital Libraries: Proceedings of 8th European Conference on Digital Libraries, ECDL 2004, LNCS 3232* (pp. 122–133). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-540-30230-8_12

Kamvar, M., & Baluja, S. (2006). A large scale study of wireless search behavior: Google mobile search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06* (pp. 701–709). New York, NY, USA: doi:10.1145/1124772.1124877

Kamvar, M., & Baluja, S. (2007a). Deciphering trends in mobile search. *Computer, 40*(8), 58–62. doi:10.1109/MC.2007.270

Kamvar, M., & Baluja, S. (2007b). The role of context in query input: Using contextual signals to complete queries on mobile devices. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '07* (pp. 405–412). New York, NY, USA: ACM. doi:10.1145/1377999.1378046

Kamvar, M., Kellar, M., Patel, R., & Xu, Y. (2009). Computers and iPhones and mobile phones, oh my!: A logs-based comparison of search users on different devices. In *Proceedings of the 18th International Conference* *on World Wide Web, WWW '09* (pp. 801–810). New York, NY, USA: ACM. doi:10.1145/1526709.1526817

Karlson, A. K., Robertson, G. G., Robbins, D. C., Czerwinski, M. P., & Smith, G. R. (2006). FaThumb: A facet-based interface for mobile search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06* (pp. 711–720). New York, NY, USA: ACM. doi:10.1145/1124772.1124878

Kellar, M., Watters, C., & Shepherd, M. (2007). A field study characterizing Web-based information-seeking tasks. *Journal of the American Society for Information Science and Technology*, *58*(7), 999–1018. doi:10.1002/asi.20590

Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, *3*(1–2), 1–224. doi:10.1561/1500000012

Kjeldskov, J., Skov, M. B., Als, B. S., & Høegh, R. T. (2004). Is it worth the hassle? Exploring the added value of evaluating the usability of context-aware mobile systems in the field. In S. Brewster & M. Dunlop (Eds.), *Mobile Human-Computer Interaction – MobileHCI 2004: Proceedings of the 6th International Symposium, MobileHCI, LNCS 3160* (pp. 61–73). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-540-28637-0_6

Komninos, A., & Dunlop, M. D. (2008). A calendar based Internet content pre-caching agent for small computing devices. *Personal and Ubiquitous Computing, 12(*7), 495–512. doi:10.1007/s00779-007-0153-4

Koren, J., Zhang, Y., & Liu, X. (2008). Personalized interactive faceted search. In *Proceedings of the 17th International Conference on World Wide Web, WWW '08* (pp. 477–486). New York, NY, USA: ACM. doi:10.1145/1367497.1367562

Koshman, S., Spink, A., & Jansen, B. J. (2006). Web searching on the Vivisimo search engine. *Journal of the American Society for Information Science and Technology*, *57*(14), 1875–1887. doi:10.1002/asi.20408

Kraft, R., Maghoul, F., & Chang, C. C. (2005). Y!Q: Contextual search at the point of inspiration. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management, CIKM '05* (pp. 816–823). New York, NY, USA: ACM. doi:10.1145/1099554.1099746

Kules, B. (2006, May). Methods for evaluating changes in search tactics induced by exploratory search systems. In *Evaluating Exploratory Search Systems*. Workshop organized in conjunction with ACM SIGIR 2006, Seattle, WA. Retrieved from http://www.takomasoftware.com/techreports/KulesESSEval-20060630b.pdf

Kules, B., Capra, R., Banta, M., & Sierra, T. (2009). What do exploratory searchers look at in a faceted search interface? In *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '09* (pp. 313–322). New York, NY, USA: ACM. doi:10.1145/1555400.1555452

Kules, B., Kustanowitz, J., & Shneiderman, B. (2006). Categorizing Web search results into meaningful and stable categories using fast-feature techniques. In *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '06* (pp. 210–219). New York, NY, USA: ACM. doi:10.1145/1141753.1141801

Kules, B., & Shneiderman, B. (2005, January). Categorized graphical overviews for Web search results: An exploratory study using U.S. government agencies as a meaningful and stable structure. In *Proceedings of the Third Annual Workshop on HCI Research in MIS* (pp. 20–23). Retrieved from http://hcil2.cs.umd.edu/trs/2004-38/2004-38.html

Kules, B., & Shneiderman, B. (2008). Users can change their Web search tactics: Design guidelines for categorized overviews. *Information Processing & Management, 44*(2), 463–484. doi:10.1016/j.ipm.2007.07.014

Kummamuru, K., Lotlikar, R., Roy, S., Singal, K., & Krishnapuram, R. (2004). A hierarchical monothetic document clustering algorithm for summarization and browsing search results. In *Proceedings of the 13th International Conference on World Wide Web, WWW '04* (pp. 658–665). New York, NY, USA: ACM. doi:10.1145/988672.988762

Käki, M. (2004). Proportional search interface usability measures. In *Proceedings of the Third Nordic Conference on Human-Computer Interaction, NordiCHI '04* (pp. 365–372). New York, NY, USA: ACM. doi:10.1145/1028014.1028072

Käki, M. (2005a). *Enhancing Web search result access with automatic categorization* (Doctoral dissertation, Dissertations in Interactive Technology, Number 2, University of Tampere, Finland). Retrieved from http://urn.fi/urn:isbn:951-44-6490-7

Käki, M. (2005b). Findex: Search result categories help users when document ranking fails. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05* (pp. 131–140). New York, NY, USA: ACM. doi:10.1145/1054972.1054991

Lam, H., & Baudisch, P. (2005). Summary thumbnails: readable overviews for small screen Web browsers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05* (pp. 681–690). New York, NY, USA: ACM. doi:10.1145/1054972.1055066

Lee, U., Liu, Z., & Cho, J. (2005). Automatic identification of user goals in Web search. In *Proceedings of the 14th International Conference on World Wide Web, WWW '05* (pp. 391–400). New York, NY, USA: ACM. doi:10.1145/1060745.1060804

Li, C., Yan, N., Roy, S. B., Lisham, L., & Das, G. (2010). Facetedpedia: Dynamic generation of query-dependent faceted interfaces for Wikipedia. In *Proceedings of the 19th International Conference on World Wide Web, WWW '10* (pp. 651–660). New York, NY, USA: ACM. doi:10.1145/1772690.1772757

Lin, J., Quan, D., Sinha, V., Bakshi, K., Huynh, D., Katz, B., & Karger, D. R. (2003). What makes a good answer? The role of context in question answering. In M. Rauterberg, M. Menozzi, & J. Wesson (Eds.), *Human-Computer Interaction – INTERACT '03: IFIP TC13 International Conference on Human-Computer Interaction* (pp. 25–32). Amsterdam, The Netherlands: IOS Press.

Lumsden, J., Kondratova, I., & Durling, S. (2007). Investigating microphone efficacy for facilitation of mobile speech-based data entry. In *People and Computers XXI – HCI … But Not As We Know It: Proceedings of HCI 2007* (pp. 89–97). Swinton, UK: British Computer Society. Retrieved from http://www.bcs.org/upload/pdf/ewic_hc07_lppaper9.pdf

Mann, T. M. (1999). Visualization of WWW-search results. In *Proceedings of the 10th International Workshop on Database and Expert Systems Applications, DEXA '99* (pp. 264–268). Washington, DC, USA: IEEE Computer Society. doi:10.1109/DEXA.1999.795176

Mann, T. M., & Reiterer, H. (2000). Evaluation of different visualizations of Web search results. In *Proceedings of the 11th International Workshop on Database and Expert Systems Applications, DEXA '00* (pp. 586–590). Washington, DC, USA: IEEE Computer Society. doi:10.1109/DEXA.2000.875084

Marchionini, G. (1989). Information-seeking strategies of novices using a full-text electronic encyclopedia. *Journal of the American Society for Information Science, 40*(1), 54–66. doi:10.1002/(SICI)1097-4571(198901)40:1<54::AID-ASI6>3.0.CO;2-R

Marchionini, G. (1995). *Information seeking in electronic environments*. New York, NY: Cambridge University Press.

Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM, 49*(4), 41–46. doi:10.1145/1121949.1121979

Marchionini, G., & White, R. W. (2008). Find what you need, understand what you find. *Journal of Human-Computer Interaction, 23*(3), 205–237. doi:10.1080/10447310701702352

Matthews, T., Pierce, J., & Tang, J. (2009). *No smartphone is an island: The impact of places, situations, and other devices on smartphone use* (IBM Technical Report RJ10452). Retrieved from http://domino.watson.ibm.com/library/Cyberdig.nsf/papers/F5FD878B5B062ACA85257635004EC3F5/$File/rj10452.pdf

Michahelles, F. (2010). Getting closer to reality by evaluating released apps? In *Research in the Large: Using App Stores, Markets and Other Wide Distribution Channels in Ubiquitous Computing Research*. Workshop organized in conjunction with UbiComp 2010, Copenhagen, Denmark. Retrieved from http://large.mobilelifecentre.org/dotclear/public/Michahelles-in-the-large.pdf

Milic-Frayling, N., Sommerer, R., Rodden, K., & Blackwell, A. (2004). SmartView and SearchMobil: Providing overview and detail in handheld browsing. In F. Crestani, M. Dunlop, & S. Mizzaro (Eds.), *Mobile and Ubiquitous Information Access: Mobile HCI 2003 International Workshop, Udine, Italy, September 8, 2003, Revised and Invited Papers, LNCS 2954* (pp. 158–171). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-540-24641-1_12

Millen, D. R., Feinberg, J., & Kerr, B. (2006). Dogear: Social bookmarking in the enterprise. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06* (pp. 111–120). New York, NY, USA: ACM. doi:10.1145/1124772.1124792

Millen, D. R., Yang, M., Whittaker, S., & Feinberg, J. (2007). Social bookmarking and exploratory search. In L. Bannon, I. Wagner, C. Gutwin, R. Harper, & K. Schmidt (Eds.), *ECSCW'07: Proceedings of the Tenth European Conference on Computer Supported Cooperative Work* (pp. 21–40). Retrieved from http://www.ecscw.org/2007/02 paper 108 Millen et al.pdf

Mizzaro, S., Sartori, L., & Strangolino, G. (2012). Tag clouds and retrieved results: The CloudCredo mobile clustering engine and its evaluation. In *Proceedings of the 3rd Italian Information Retrieval Workshop* (pp. 191-198). Retrieved from http://ceur-ws.org/Vol-835/paper22.pdf

Morris, M. R., & Horvitz, E. (2007). SearchTogether: An interface for collaborative Web search. In *Proceedings of the 20th Annual ACM Symposium on User interface Software and Technology, UIST '07* (pp. 3–12). New York, NY, USA: ACM. doi:10.1145/1294211.1294215

Morrison, A., Reeves, S., McMillan, D., & Chalmers, M. (2010). Experiences of mass participation in ubicomp research. In *Research in* *the Large: Using App Stores, Markets and Other Wide Distribution Channels in Ubiquitous Computing Research*. Workshop organized in conjunction with UbiComp 2010, Copenhagen, Denmark. Retrieved from http://large.mobilelifecentre.org/dotclear/public/Morrison-Mass-in-the-large.pdf

Morrison, J. B., Pirolli, P., & Card, S. K. (2001). A taxonomic analysis of what World Wide Web activities significantly impact people's decisions and actions. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (pp. 163–164). New York, NY, USA: ACM. doi:10.1145/634067.634167

Nakhimovsky, Y., Eckles, D., & Riegelsberger, J. (2009). Mobile user experience research: challenges, methods & tools. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (pp. 4795–4798). New York, NY, USA: ACM. doi:10.1145/1520340.1520743

Navarro-Prieto, R., Scaife, M., & Rogers, Y. (1999). Cognitive strategies in Web searching. In *Proceedings of the 5th Conference on Human Factors & the Web*. Retrieved from http://zing.ncsl.nist.gov/hfweb/proceedings/navarro-prieto/index.html

Nielsen, C. M., Overgaard, M., Pedersen, M. B., Stage, J., & Stenild, S. (2006). It's worth the hassle!: The added value of evaluating the usability of mobile systems in the field. In A. Mørch, K. Morgan, T. Bratteteig, G. Ghosh, & D. Svanaes (Eds.), *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles, NordiCHI '06* (pp. 272–280). New York, NY, USA. doi:10.1145/1182475.1182504

Nielsen, J. (2003). *Information foraging: Why Google makes people leave your site faster*. Retrieved from http://www.useit.com/alertbox/20030630.html

Nielsen Mobile. (2008). *Critical mass: Worldwide state of the mobile Web*. Retrieved from http://nielsen.com/us/en/insights/reports-downloads/2008/critical-mass-worldwide-state-of-the-mobile-web.html

Nylander, S., Lundquist, T., Brännström, A., & Karlson, B. (2009). "It's just easier with the phone" – A diary study of Internet access from cell phones. In H. Tokuda, M. Beigl, A. Friday, A. J. Bernheim Brush, & Y. Tobe (Eds.), *Pervasive Computing: Proceedings of the 7th International Conference on Pervasive Computing, Pervasive 2009, LNCS 5538* (pp. 354–371). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-642-01516-8_24

O'Day, V. L., & Jeffries, R. (1993). Orienteering in an information landscape: How information seekers get from here to there. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human*

Factors in Computing Systems, CHI '93 (pp. 438–445). New York, NY, USA: ACM. doi:10.1145/169059.169365

Ogden, W. C., Davis, M. W., & Rice, S. (1998). Document thumbnail visualization for rapid relevance judgments: When do they pay off? In *Proceedings of The Seventh Text REtrieval Conference* (pp. 528–534). Retrieved from http://trec.nist.gov/pubs/trec7/papers/nmsu.pdf.gz

Oulasvirta, A., Tamminen, S., Roto, V., & Kuorelahti, J. (2005). Interaction in 4-second bursts: The fragmented nature of attentional resources in mobile HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05* (pp. 919–928). New York, NY, USA: ACM. doi:10.1145/1054972.1055101

Paek, T., Dumais, S., & Logan, R. (2004). WaveLens: A new view onto Internet search results. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04* (pp. 727–734). New York, NY, USA: ACM. doi:10.1145/985692.985784

PARC. (2012). *Sensemaking glossary*. Retrieved from: http://www2.parc.com/istl/groups/hdi/sensemaking/glossary.htm

Pirolli, P., & Card, S. K. (1999). Information foraging. *Psychological Review, 106*(4), 643–675. doi:10.1037/0033-295X.106.4.643

Pirolli, P., Schank, P., Hearst, M., & Diehl, C. (1996). Scatter/Gather browsing communicates the topic structure of a very large text collection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '96* (pp. 213–220). New York, NY, USA: ACM. doi:10.1145/238386.238489

Plaisant, C. (2004). The challenge of information visualization evaluation. In *Proceedings of the Working Conference on Advanced Visual Interfaces, AVI '04* (pp. 109–116). New York, NY, USA: ACM. doi:10.1145/989863.989880

Pratt, W., Hearst, M, & Fagan, L. (1999). A knowledge-based approach to organizing retrieved documents. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence, AAAI '99/IAAI '99* (pp. 80–85). Menlo Park, CA: AAAI Press.

Purchase, H. C., Andrienko, N., Jankun-Kelly, T. J., & Ward, M. (2008). Theoretical foundations of information visualization. In A. Kerren, J. T. Stasko, J.-D. Fekete, & C. North (Eds.), *Information Visualization: Human-centered Issues and Perspectives, Part I, LNCS* 4950 (pp. 46–64). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-540-70956-5_3

Reis, S., Church, K., & Oliver, N. (2012). Rethinking mobile search: Towards casual, shared, social mobile search experiences. In D. Elsweiler, M. L. Wilson, & M. Harvey (Eds.), *Proceedings of the "Searching4Fun!" workshop* (pp. 1–4). Retrieved from http://ceur-ws.org/Vol-836/paper1.pdf

Reiterer, H., Tullius, G., & Mann, T. M. (2005). INSYDER: A content-based visual-information-seeking system for the Web. *International Journal on Digital Libraries, 5*(1), 25–41. doi:10.1007/s00799-004-0111-y

Resnick, M. L., & Vaughan, M. W. (2006). Best practices and future visions for search user interfaces. *Journal of the American Society for Information Science and Technology, 57*(6), 781–787. doi:10.1002/asi.20292

Rice, R. E., McCreadie, M., & Chang, S. L. (2001). *Accessing and browsing information and communication*. Cambridge, MA: MIT Press.

Riegelsberger, J., & Nakhimovsky, Y. (2008). Seeing the bigger picture: A multi-method field trial of Google Maps for mobile. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (pp. 2221–2228). New York, NY, USA: ACM. doi:10.1145/1358628.1358655

Rivadeneira, W., & Bederson, B. B. (2003). *A study of search result clustering interfaces: Comparing textual and zoomable user interfaces* (Technical Report HCIL-2003-36, CS-TR-4682). University of Maryland. Retrieved from http://hcil2.cs.umd.edu/trs/2003-36/2003-36.pdf

Roberts, J.C., & Suvanaphen, E. (2003). Visual bracketing for Web search result visualization. In *Proceedings of 7th International Conference on Information Visualisation, IV '03* (pp. 264–269). Washington, DC, USA: IEEE Computer Society. doi:10.1109/IV.2003.1217989

Rodden, K., Basalaj, W., Sinclair, D., & Wood, K. R. (2001). Does organisation by similarity assist image browsing? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '01* (pp. 190–197). New York, NY, USA: ACM. doi:10.1145/365024.365097

Rogers, Y., Connelly, K., Tedesco, L., Hazlewood, W., Kurtz, A., Hall, R. E., Hursey, J., & Toscos, T. (2007). Why it's worth the hassle: the value of in-situ studies when designing ubicomp. In J. Krumm, G. D. Abowd, A. Seneviratne, & T. Strang (Eds.), *Ubicomp 2007: Ubiquitous Computing, Proceedings of the 9th International Conference, LNCS 4717* (pp. 336–353). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-540-74853-3_20

Rose, D. E. (2006). Reconciling information-seeking behavior with search user interfaces for the Web. *Journal of the American Society for Information Science and Technology*, *57*(6), 797–799. doi:10.1002/asi.20295

Rose, D. E., & Levinson, D. (2004). Understanding user goals in Web search. In *Proceedings of the 13th International Conference on the World Wide Web, WWW '04* (pp. 13–19). New York, NY, USA: ACM. doi:10.1145/988672.988675

Russell, D. M., Pirolli, P., Furnas, G., Card, S. K., & Stefik, M. (2009). Sensemaking workshop CHI 2009. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (pp. 4751–4754). New York, NY, USA. doi:10.1145/1520340.1520732

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The cost structure of sensemaking. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, CHI '93* (269–276). New York, NY, USA: ACM. doi:10.1145/169059.169209

schraefel, m. c., Karam, M., & Zhao, S. (2003). Listen to the music: Audio preview cues for exploration of online music. In M. Rauterberg, M. Menozzi, & J. Wesson (Eds.), *Human-Computer Interaction – INTERACT '03, IFIP TC13 International Conference on Human-Computer Interaction* (pp. 192–195). Amsterdam, The Netherlands: IOS Press.

schraefel, m. c., Wilson, M. L., Russell, A., & Smith D. A. (2006). mSpace: Improving information access to multimedia domains with multimodal exploratory search. *Communications of the ACM, 49*(4), 47–49. doi:10.1145/1121949.1121980

Sebrechts, M. M., Cugini, J., Laskowski, S. J., Vasilakis, J., & Miller, M. S. (1999). Visualization of search results: A comparative evaluation of text, 2D, and 3D interfaces. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '99* (pp. 3–10). New York, NY, USA: ACM. doi:10.1145/312624.312634

Sellen, A. J., Murphy, R., & Shaw, K. L. (2002). How knowledge workers use the Web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '02* (pp. 227–234). New York, NY, USA: ACM. doi:10.1145/503376.503418

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualization. In *Proceedings of the 1996 IEEE Symposium on Visual Languages, VL '96* (pp. 336–343). Washington, DC, USA: IEEE Computer Society. doi:10.1109/VL.1996.545307

Shneiderman, B. (1997). A framework for search interfaces. *IEEE Software, 14*(2), 18–20. doi:10.1109/52.582969

Shneiderman, B. (2008, June). Research agenda: Visual overviews for exploratory search. In *Information Seeking Support Systems Workshop*. Sponsored by the National Science Foundation, Chapel Hill, NC, USA.

Retrieved from http://ils.unc.edu/ISSS/papers/papers/shneiderman.pdf

Shneiderman, B., Byrd, D., Croft, W. B. (1997). Clarifying search: A user-interface framework for text searches. *DLib Magazine*, *3*(1). Retrieved from http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html

Shneiderman, B., & Plaisant, C. (2006). Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization* (pp. 1–7). New York, NY, USA: ACM. doi:10.1145/1168149.1168158

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2009). *Designing the user interface: Strategies for effective human-computer interaction (5th ed.)*. Boston, MA, USA: Addison-Wesley Publishing Company.

Shtykh, R. Y., Chen, J., & Jin, Q. (2008). Slide-film interface: Overcoming small screen limitations in mobile Web search. In C. Macdonald, I. Ounis, V. Plachouras, I. Ruthven, & R. W. White (Eds.), *Advances in Information Retrieval*: *Proceedings of the 30th European Conference on IR Research, ECIR 2008, LNCS 4956* (pp. 622–626). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-540-78646-7_67

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web search engine query log. *ACM SIGIR Forum, 33*(3), 6–12. doi:10.1145/331403.331405

Sinclair, J., & Cardew-Hall, M. (2008). The folksonomy tag cloud: When is it useful? *Journal of Information Science, 34*(1), 15–29. doi:10.1177/0165551506078083

Sohn, T., Li, K. A., Griswold, W. G., & Hollan, J. D. (2008). A diary study of mobile information needs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08* (pp. 433–442). New York, NY, USA: ACM. doi:10.1145/1357054.1357125

Spence, R. (2001). *Information visualization (1st ed.)*. Harlow, UK: ACM Press Books, Addison-Wesley.

Spence, R. (2007). *Information visualization – Design for interaction (2nd ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall.

Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology, 52*(3), 226–234. doi:10.1002/1097-4571(2000)9999:9999<::AID-ASI1591>3.0.CO;2-R

Stoica, E., Hearst, M., & Richardson, M. (2007). Automating creation of hierarchical faceted metadata structures. In C. L. Sidner, T. Schultz, M. Stone, & C. Zhai (Eds.), *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, NAACL-HLT '07* (pp. 244–251). The Association for Computational Linguistics.

Sutcliffe, A. G., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers, 10*(3), 321–351. doi:10.1016/S0953-5438(98)00013-7

Sutcliffe, A. G., Ennis, M., & Hu, J. (2000). Evaluating the effectiveness of visual user interfaces for information retrieval. *International Journal of Human-Computer Studies, 53*(5), 741–763. doi:10.1006/ijhc.2000.0416

Sweeney, S., & Crestani, F. (2004) Supporting searching on small screen devices using summarisation. In F. Crestani, M. Dunlop, & S. Mizzaro (Eds.), *Mobile and Ubiquitous Information Access: Mobile HCI 2003 International Workshop, Udine, Italy, September 8, 2003, Revised and Invited Papers, LNCS 2954* (pp. 187–201). Berlin Heidelberg, Germany: Springer. doi:10.1007/978-3-540-24641-1_14

Sweeney, S., & Crestani, F. (2006). Effective search results summary size and device screen size: Is there a relationship? *Information Processing & Management, 42*(4), 1056–1074. doi:10.1016/j.ipm.2005.06.007

Sweeney, S., Crestani, F., & Losada, D. E. (2008). 'Show me more': Incremental length summarisation using novelty detection. *Information Processing & Management, 44*(2), 663–686. doi:10.1016/j.ipm.2007.03.012

Tamminen, S., Oulasvirta, A., Toiskallio, K., & Kankainen, A. (2004). Understanding mobile contexts. *Personal and Ubiquitous Computing, 8*(2), 135–143. doi:10.1007/s00779-004-0263-1

Taylor, C. A., Anicello, O., Somohano, S., Samuels, N., Whitaker, L., & Ramey, J. A. (2008). A framework for understanding mobile Internet motivations and behaviors. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (pp. 2679–2684). New York, NY, USA: ACM. doi:10.1145/1358628.1358744

Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04* (pp. 415–422). New York, NY, USA: ACM. doi:10.1145/985692.985745

Teevan, J., Cutrell, E., Fisher, D., Drucker, S. M., Ramos, P. A. G., & Hu, C. (2009). Visual snippets: Summarizing web pages for search and revisitation. In *Proceedings of the 27th International Conference on Human*

Factors in Computing Systems, CHI '09 (pp. 2023–2032). New York, NY, USA: ACM. doi:10.1145/1518701.1519008

Teevan, J., Karlson, A., Amini, S., Bernheim Brush, A. J., & Krumm, J. (2011). Understanding the importance of location, time, and people in mobile local search behavior. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '11* (pp. 77–80). New York, NY, USA: ACM. doi:10.1145/2037373.2037386

Tombros, A., & Sanderson, M. (1998). Advantages of query biased summaries in information retrieval. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '98* (pp. 2–10). New York, NY, USA: ACM. doi:10.1145/290941.290947

Toms, E. G. (2000). Understanding and facilitating the browsing of electronic text. *International Journal of Human-Computer Studies, 52*(3), 423–452. doi:10.1006/ijhc.1999.0345

Tufte, E. (1983). *The visual display of quantitative information (2nd ed.)*. Cheshire, CT, USA: Graphics Press.

Turetken, O., & Sharda, R. (2005). Clustering-based visual interfaces for presentation of Web search results: An empirical investigation. *Information Systems Frontiers, 7*(3), 273–297. doi:10.1007/s10796-005-2770-7

Turunen, M., Hakulinen, J., Melto, A., Heimonen, T., Laivo, T., & Hella, J. (2009). SUXES – User experience evaluation method for spoken and multimodal interaction. In *Proceedings of Interspeech 2009* (pp. 2567–2570). International Speech Communication Association.

van Rijsbergen, C. J. (1979). *Information retrieval*. London, UK: Butterworth-Heinemann.

Veerasamy, A., & Belkin, N. J. (1996). Evaluation of a tool for visualization of information retrieval results. In *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '96* (pp. 85–92). New York, NY, USA: ACM. doi:10.1145/243199.243218

Veerasamy, A., & Heikes, R. (1997). Effectiveness of a graphical display of retrieval results. In *Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '97* (pp. 236–245). New York, NY, USA: ACM. doi:10.1145/278459.258580

Venkatsubramanyan, S., & Perez-Carballo, J. (2007). Techniques for organizing and presenting search results: A survey. *Journal of Information Science and Technology, (4)*2, 27 pages. Retrieved from http://ssrn.com/abstract=1132817

Wang, X., & Zhai, C. (2007). Learn from Web search logs to organize search results. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval 2007, SIGIR '07* (pp. 87–94). New York, NY, USA: ACM. doi:10.1145/1277741.1277759

Ware, C. (2008). *Visual thinking for design*. San Francisco, CA, USA: Morgan Kaufmann.

White, R. W., Jose, J., & Ruthven, I. (2003). A task-oriented study on the influencing effects of query-biased summarisation in Web searching. *Information Processing & Management, 39*(5), 707–733. doi:10.1016/S0306-4573(02)00033-X

White, R. W., Jose, J. M., & Ruthven, I. (2005). Using top-ranking sentences to facilitate effective information access. *Journal of the American Society for Information Science and Technology, 56*(10), 1113–1125. doi:10.1002/asi.20203

White, R. W., Kules, B., & Bederson, B. (2005). Exploratory search interfaces: categorization, clustering and beyond: Report on the XSI 2005 workshop at the Human-Computer Interaction Laboratory, University of Maryland. *SIGIR Forum*, *39*(2), 52–56. doi:10.1145/1113343.1113356

White, R. W., Kules, B., Drucker, S. M., & schraefel, m.c. (2006). Supporting exploratory search, Introduction. *Communications of the ACM, 49*(4), 36–39. doi:10.1145/1121949.1121978

Wilson, M. L. (2009, October). Keyword search: Quite exploratory actually. In *3rd International Workshop on Human-Computer Interaction and Information Retrieval*. Retrieved from http://eprints.soton.ac.uk/267951/1/hcir09_keyword_CR_mlw.pdf

Wilson, M. L., Kules, B., schraefel, m. c., & Shneiderman, B. (2010). From keyword search to exploration: Designing future search interfaces for the Web. *Foundations and Trends in Web Science*, 2(1), 1–97. doi:10.1561/1800000003

Wilson, M. L., Russell, A., Smith, D. A., & schraefel, m. c. (2006). mSpace Mobile: Exploring support for mobile tasks. In N. Bryan-Kinns, A. Blanford, P. Curzon, & L. Nigay (Eds.), *People and Computers XX – Engage: Proceedings HCI 2006* (pp. 193–202). London, UK: Springer. doi:10.1007/978-1-84628-664-3_15

Wilson, M. L., & schraefel, m. c. (2008). A longitudinal study of exploratory and keyword search. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '08* (pp. 52–56). New York, NY, USA: ACM. doi:10.1145/1378889.1378899

Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation*, *37*(1), 3–15. doi:10.1108/eb026702

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation, 55*(3), 249–270. doi:10.1108/EUM0000000007145

Wobbrock, J. O., Forlizzi, J., Hudson, S. E., & Myers, B. A. (2002). WebThumb: Interaction techniques for small-screen browsers. In *Proceedings of the ACM Symposium on User Interface Software and Technology, UIST '02* (pp. 205–208). New York, NY, USA: ACM. doi:10.1145/571985.572014

Woodruff, A., Rosenholtz, R., Morrison, J. B., Faulring, A., & Pirolli, P. (2002). A comparison of the use of text summaries, plain thumbnails, and enhanced thumbnails for Web search tasks. *Journal of the American Society for Information Science and Technology, 53*(2), 172–185. doi:10.1002/asi.10029

Wu, H., Zubair, M., & Maly, K. (2006). Harvesting social knowledge from folksonomies. In *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, HYPERTEXT '06* (pp. 111–114). New York, NY, USA: ACM. doi:10.1145/1149941.1149962

Wu, Y. B., Shankar, L., & Chen, X. (2003). Finding more useful information faster from Web search results. In *Proceedings of the Twelfth International Conference on Information and Knowledge Management, CIKM '03* (pp. 568–571). New York, NY, USA: ACM. doi:10.1145/956863.956975

Xie, H. (2002). Patterns between interactive intentions and information-seeking strategies. *Information Processing & Management, 38*(1), 55–77. doi:10.1016/S0306-4573(01)00018-8

Xie, X., Miao, G., Song, R., Wen, J.-R., & Ma, W.-Y. (2005). Efficient browsing of Web search results on mobile devices based on block importance model. In *Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications, PERCOM '05* (pp. 17–26). Washington, DC, USA: IEEE Computer Society. doi:10.1109/PERCOM.2005.16

Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '03* (pp. 401–408). New York, NY, USA: ACM. doi:10.1145/642611.642681

Yi, J., Maghoul, F., & Pedersen, J. (2008). Deciphering mobile search patterns: A study of Yahoo! mobile search queries. In *Proceedings of the 17th International Conference on World Wide Web, WWW '08* (pp. 257–266). New York, NY, USA: ACM. doi:10.1145/1367497.1367533

Yndurain, E., Bernhardt, D., & Campo, C. (2012). Augmenting mobile search engines to leverage context awareness. *IEEE Internet Computing, 16*(2), 17–25. doi:10.1109/MIC.2012.17

Zamir, O., & Etzioni, O. (1999). Grouper: A dynamic clustering interface to Web search results. *Computer Networks: The International Journal of Computer and Telecommunications Networking, 31(*11-16), 1361–1374. doi:10.1016/S1389-1286(99)00054-7

Zeng, H. J., He, Q. C., Chen, Z., Ma, W.-Y., & Ma, J. (2004). Learning to cluster Web search results. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '04* (pp. 210–217). New York, NY, USA: ACM. doi:10.1145/1008992.1009030

Zhang, J., & Marchionini, G. (2005). Evaluation and evolution of a browse and search interface: Relation Browser++. In *Proceedings of the 2005 National Conference on Digital Government Research, d.go '05* (pp. 179–188). Digital Government Society of North America.

Paper I

Heimonen, T., & Jhaveri, N. (2005). Visualizing query occurrence in search result lists. *Proceedings of the 9th International Conference on Information Visualisation, IV '05* (877–882). Washington, DC: IEEE Computer Society.

Original article available online at:
http://dx.doi.org/10.1109/IV.2005.152

Paper II

Heimonen, T., & Siirtola, H. (2009). Visualizing query occurrence in mobile Web search interfaces. *Proceedings of the 13th International Conference on Information Visualisation, IV '09* (pp. 639–644). Washington, DC: IEEE Computer Society.

Original article available online at:
http://dx.doi.org/10.1109/IV.2009.16

Paper III

Heimonen, T., & Käki, M. (2007). Mobile Findex – Supporting mobile web search with automatic result categories. *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '07* (pp. 397–404). New York, NY: ACM.

Original article available online at:
http://doi.acm.org/10.1145/1377999.1378045

Paper IV

Heimonen, T. (2008). Mobile Findex: Facilitating information access in mobile Web search with automatic result clustering. *Advances in Human-Computer Interaction*, *2008*, article ID 680640.

Original article available online at:
http://dx.doi.org/10.1155/2008/680640

Paper V

Heimonen, T. (2012). How do users search the mobile Web with a clustering interface? A longitudinal study. *International Journal of Mobile Human-Computer Interaction, 4*(3), 44–66.

Original article available online at:
doi:10.4018/jmhci.2012070103

Paper VI

Heimonen, T. (2009). Information needs and practices of active mobile Internet users. *Proceedings of the 6th International Conference on Mobile Technology, Applications, and Systems* (Article 50). New York, NY: ACM.

Original article available online at:
http://doi.acm.org/10.1145/1710035.1710085

Publications in the **Dissertations in Interactive Technology** series

1. **Timo Partala:** Affective Information in Human-Computer Interaction
2. **Mika Käki:** Enhancing Web Search Result Access with Automatic Categorization
3. **Anne Aula:** Studying User Strategies and Characteristics for Developing Web Search Interfaces
4. **Aulikki Hyrskykari:** Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid
5. **Johanna Höysniemi:** Design and Evaluation of Physically Interactive Games
6. **Jaakko Hakulinen:** Software Tutoring in Speech User Interfaces
7. **Harri Siirtola:** Interactive Visualization of Multidimensional Data
8. **Erno Mäkinen:** Face Analysis Techniques for Human-Computer Interaction
9. **Oleg Špakov:** iComponent – Device-Independent Platform for Analyzing Eye Movement Data and Developing Eye-Based Applications
10. **Yulia Gizatdinova:** Automatic Detection of Face and Facial Features from Images of Neutral and Expressive Faces
11. **Päivi Majaranta:** Text Entry by Eye Gaze
12. **Ying Liu:** Chinese Text Entry with Mobile Devices
13. **Toni Vanhala:** Towards Computer-Assisted Regulation of Emotions
14. **Tomi Heimonen:** Design and Evaluation of User Interfaces for Mobile Web Search