# NSF Workshop on Task-Based Search

NSF Workshop on Task-Based Search
University of North Carolina at Chapel Hill
School of Information and Library Science
March 14-15, 2013

# NSF Task-Based Information Search Systems Workshop: Workshop Overview

Search engines are optimized for particular types of tasks (e.g., look-up tasks and commerce tasks such as travel and shopping), for particular types of search behaviors (i.e., enter a query, review snippets, make a transaction) and for particular types of searchers (i.e., those who want to quickly find a single piece of information). Search engines are not optimized for tasks that require sustained interaction and engagement with information, the use of multiple, diverse search approaches to finding information or for searchers who want to cultivate a deeper understanding of a problem or topic.

This workshop will gather leading researchers in interactive information retrieval to discuss research and challenges in incorporating models of tasks, task-types, and users' needs into systems/tools to support complex, multi-search and multi-session tasks. There are many challenges in creating such task-oriented search systems and the goal of this workshop is to enumerate, discuss, and document these issues into a research agenda that can help guide work in this field. Specifically, this workshop will focus on the following topics:

1. Identification, elicitation, modeling and tracking of tasks, processes and states, including the identification of frameworks for conceptualizing task and relevance models;
2. Creation of task-specific and task-aware search environments, including the development of interfaces, tools, features, indexing techniques and search algorithms;
3. Development of methods and measures for studying user behavior and evaluating task-based search systems.

Workshop Dates: March 14-15, 2013
Workshop Location: Chapel Hill, NC, USA
Organized by:

Jaime Arguello, Rob Capra, and Diane Kelly
School of Information & Library Science
University of North Carolina
100 Manning Hall, CB#3360
Chapel Hill, NC 27599-3360

# Contents

## Schedule

**Thursday, March 14:** Day One

| Time | Agenda |
|---|---|
| 7:40 a.m. | Van trips from hotel to Wilson Library, Pleasants Family Assembly Room |
| 8:00-8:30 a.m. | *Breakfast*<br>*(coffee, tea, fruit, bagels with cream cheese and assorted quiches)* |
| 8:30-9:00 a.m. | Introductions and Overview |
| 9:00-9:40 a.m. | Core Area 1: Task-based Information Seeking (Pertti Vakkari) |
| 9:40-10:20 a.m. | Core Area 2: Search Engines and Task (Sue Dumais) |
| 10:20-10:50 a.m. | *Break*<br>*(coffee, tea and snacks)* |
| 10:50-11:30 a.m. | Core Area 3: Interfaces and Task (Gene Golovchinsky) |
| 11:30-11:50 a.m. | Field Views: Session Workshop Report (Nick Belkin) |
| 11:50-12:10 p.m. | Field Views: TREC Session Track (Ben Carterette) |
| 12:10-1:10 p.m. | *Lunch*<br>*(sandwiches, salads, and cookies)* |
| 1:10-1:30 p.m. | Instructions for Break-out Groups |
| 1:30-3:00 p.m. | Break-Out Session 1 |
| 3:00-3:30 p.m. | *Coffee Break*<br>*(coffee, tea and snacks)* |
| 3:30-4:45 p.m. | Break-Out Group Presentations |
| 4:45 p.m. | Van trips back to hotel begin |
| 6:30-7:30 p.m. | *Reception at the WXYZ Lounge* |
| 7:30-9:30 p.m. | *Walk to dinner at Elements* |

Access to wireless internet is available through the UNC-Guest network. Please close your web browser before connecting.

**Friday, March 15:** Day Two

| Time | Agenda |
|---|---|
| 7:40 a.m. | Van trips from hotel to campus |
| *8:00-8:30 a.m.* | *Breakfast* <br> *(coffee, tea, fruit, bagels with cream cheese and assorted quiches)* |
| 8:30-9:30 a.m. | Plenary Discussion: Refine and Prioritize Research Questions |
| 9:30-10:30 a.m. | Break-Out Session 2 |
| *10:30-11:00 a.m.* | *Break* <br> *(coffee, tea and snacks)* |
| 11:00-12:00 p.m. | Break-Out Session 2 continues |
| *12:00-1:00 p.m.* | *Lunch* <br> *(from Mediterranean Deli)* |
| 1:10-2:30 p.m. | Reports from Break-out Sessions |
| 2:30-3:30 p.m. | Plenary Discussion: Next Steps and Research Agenda |
| 3:30 p.m. | Close |
| 3:45 p.m. | Van trips to hotel begin.  Taxi trips to airport (for departures before 7pm) |
|  | *no pre-arranged dinner* |

# Challenges and References Identified by Participants

## Eugene Agichtein

Automatically identifying and naturally supporting long-running (multi-session or multi-day) search tasks. Aspects of the problem include:

- Building a taxonomy of complex search tasks, and important components of the task, e.g., a template for the kinds of things people find when planning a trip.
- Automatically detecting early on that a user is embarking on a (potentially) long search task (e.g., as in [1]).
- Identifying the type of a task by matching to the taxonomy in 1
- Detecting whether the user has completed the task or may resume it later.
- Understanding the possible interfaces to help the searcher resume the task from the last state (e.g., by expanding on [2]).

References

1. Eugene Agichtein, Ryen W. White, Susan T. Dumais, and Paul N. Bennet. Search, interrupted: understanding and predicting search task continuation. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 315–324, New York, NY, USA, 2012. ACM.
2. Debora Donato, Francesco Bonchi, Tom Chi, and Yoelle Maarek. Do you want to take notes?: identifying research missions in yahoo! search pad. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 321–330, New York, NY, USA, 2010. ACM.

## Jae-wook Ahn

What are the limitations of visual user interfaces for task-based search and how can we overcome them? [1,2] shows when a transparent user model (or task model) can fail. Unlike [1,2] which implement an offline search system or a text-based transparent user model, [3] presents a 2-D visualization based approach, which can overcome some of the limitations of the past approaches.

What are the properties that should be considered when evaluating task-based search system user interfaces that emphasize transparency? [4] suggests a list of aims for explanatory recommender systems, which could be helpful for defining the aims of task-based search system user interfaces.

References

1. Annika Waern. User involvement in automatic filtering: An experimental study. *User Modeling and User-Adapted Interaction*, 14(2-3):201–237, 2004.
2. Jae-wook Ahn, Peter Brusilovsky, Jonathan Grady, Daqing He, and Sue Yeon Syn. Open user profiles for adaptive news systems: help or harm? In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 11–20, New York, NY, USA, 2007. ACM.
3. Jae-wook Ahn and Peter Brusilovsky. Adaptive Visualization for Exploratory Information Retrieval. To appear in *Information Processing and Management*.
4. Nava Tintarev and Judith Masthoff. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction*, 22:399–439, 2012.

## Nicholas Belkin

I think that the most fundamental problem in this respect is the ability to infer motivating task type from the searcher's past and current information-seeking behaviors. This implies having a typology of motivating search tasks to start with, which in and of itself is a significant research problem. I find it difficult to separate these two research problems, so consider them in this context as one.

References

1. Yuelin Li and Nicholas J. Belkin. A faceted approach to conceptualizing tasks in information seeking. *Information Processing and Management*, 44(6):1822–1837, 2008. This paper proposes a scheme for classification of both motivating search task types, and information searching tasks. The major contribution here is a means for classification that is not just naming different tasks, but rather a principled scheme for characterizing different task types. This means that in experimental situations, task type can be manipulated according to different values of some facets of task.
2. Chang Liu, Michael Cole, Eun Baik, and Nicholas J. Belkin. Rutgers at the TREC 2012 Session Track. In *Proceedings of 21st Text Retrieval Conference*, TREC '12. 2012. Although this paper does not investigate the issue of predicting motivating task from information-seeking behaviors, it makes a step in this direction by considering how different motivating tasks influence interpretation of information-seeking behaviors for the purpose of identifying "useful" documents during a search session.

## Pia Borlund

The research problem that I would like to address is in line with the fourth mentioned example: "The need to develop IR evaluation methods that operate across multiple queries and even multiple search sessions". To me the objective is to be able to evaluate the IR interaction of the user as realistically as possible, that is, to handle multiple queries and even multiple search sessions – or in other words, to understand and evaluate IIR as it takes place in real life, including multi-facetted information needs and multi-tasking/task-switching. E.g., see the papers by Belkin (2008; 2010) and Spink (2004).

Also I would like to bring attention to the need for focus on research on searching of work tasks. That is, information searching as part of work task solving, as briefly addressed in the paper by Borlund, Dreier & Byström (2012).

References

1. Nicholas J. Belkin. Some(what) grand challenges for information retrieval. *SIGIR Forum*, 42(1):47–54, 2008. This paper explicitly points out a number of issues we ought to address, not the least with reference to evaluation of IIR systems.
2. Nicholas J. Belkin. On the Evaluation of Interactive Information Retrieval Systems. In: Larsen, B., Schneider, J.W., & Åström (Eds.). *The Janus Faced Scholar: A Festschrift in Honour of Peter Ingwersen*. Copenhagen: Royal School of Library and Information Science. 13–21, 2010. This paper contributes with a new framework for IIR evaluation focusing on usefulness.
3. Pia Borlund, Sabine Dreier, and Katriina Byström. What does time spent on searching indicate? In *Proceedings of the 4th Information Interaction in Context Symposium*, IIIX '12, pages 184–193, New York, NY, USA, 2012. ACM. The reported information seeking work task study reminds us that information searching takes place also in information intensive work task performance settings. The impression we have, is that the majority of current IIR research centres on Internet searching and everyday-life information needs – including the two IIR studies reported in this paper. However, there remains a need for IIR research on information searching in relation to information intensive work task performance with respect to optimise information searching, the various platforms used for information searching, and understanding of the conditions under which work task performance takes place.
4. Amanda Spink. Multitasking information behavior and information task switching: An exploratory study. *Journal of Documentation*, 60(4), 336–351. 2004. Emerald. This paper is an early example of the addressing of multi-tasking and task switching, hence not including the seamless IT and information environment of today, which also have to be taken into account.

## Katriina Byström

I think the following issue appears as fruitful to be addressed in order to design search systems that are more task-aware:  Contextualizing task properties and search behavior, and the relationship them between into relevant information practices/behaviour.

Taylor's (1991) article discusses how different professional groups are formed around information use environments that in themselves include traits for what information is valued and consequently sought for as well as through what channels and sources this information is searched/distributed. Byström & Lloyd (2012) pushes the idea further by suggesting that each information use environment creates pervasive information practices with time sensitive professional and local influences. Work tasks fit into these environments as concrete instances where explicit and tacit knowledge culminates, which is why they provide useful base to study information search behavior and understand the role of IR systems. For the field of task-based information search this may provide a possibility to explain search behavior and design/evaluate IR systems not only from a user-oriented perspective, but also acknowledging the sociocultural aspects of search.

References

1. Katriina Byström and Annemaree Lloyd. Practice theory and work task performance: How are they related and how can they contribute to a study of information practices. *Proceedings of the American Society for Information Science and Technology*, 49(1):1–5, 2012.
2. Robert S. Taylor. Information use environments. In B. Dervin (Ed.) *Progress in Communication Sciences*, 10: 217–225,1991

## Ben Carterette

Whole-session evaluation: being able to evaluate the utility of a search system over the course of a user's interaction with it, ideally from task commencement to task completion. I'm envisioning a "task-aware" system as being one that attempts to determine a user's task from their interactions and adapt accordingly; if nothing else, it seems like some kind of sessiony evaluation would be necessary for use in objective functions. For example, Liu et al. [1] use task type prediction to select a feedback model during the course of a section.

While there are probably many ways to do whole-session evaluation (user studies, log analysis, etc), I am particularly interested in batch-style evaluations with reusable test collections. Batch evaluations allow researchers and developers to quickly perform tests of many possible combinations of features, models, and inputs while maintaining high statistical power. Reusability allows them to go back to any point in that search space and reliably get the same performance.

Creating test collections for whole-session evaluation is a difficult problem. We have been attempting to tackle it through the TREC Session track for the last three years [2, 3], and while we are happy with what we have accomplished, we still have a long way to go. The main problem is that it is difficult to model the fact that user interactions at time t+1 can depend on what the system does at time t; if the same test collection is going to be used to evaluate n different systems, it has to be able to model up to n different possible user actions at each time step. A direction we are considering is to use user simulation; while it is not likely that we will be able to accurately simulate users, we may be able to produce interactions that are at least useful for improving task-aware search systems.

The two Session track papers describe our efforts towards creating test collections for session evaluation. The second paper on the 2012 track is more specifically related to task-aware search, as our topics were categorized into four different broad task types. The Liu et al. paper describes the participation of Rutgers in the track. They built different feedback models for different task types and showed substantial improvements on some task types. This suggests that such a test collection can actually be useful for training task-aware systems.

References

1. Chang Liu, Michael Cole, Eun Baik, and Nicholas J. Belkin. Rutgers at the TREC 2012 Session Track. In *Proceedings of 21st Text Retrieval Conference*, TREC '12. 2012.
2. Evangelos Kanoulas, Ben Carterette, Mark Hall, Paul Clough, and Mark Sanderson. Overview of the 2011 Session Track. In *Proceedings of 20th Text Retrieval Conference*, TREC '11. 2011.
3. Evangelos Kanoulas, Ben Carterette, Mark Hall, Paul Clough, and Mark Sanderson. Overview of the 2012 Session Track. In *Proceedings of 21st Text Retrieval Conference*, TREC '12. 2012.

The notion of "relevance", which is so important to batch-style system-based evaluation, strikes me as limited in its ability to capture what users need from systems in order to actually complete tasks. If we instead talk about "utility"---as in the utility of a document to aid task completion---we can model utility not just by relevance but also by other important criteria such

as timeliness, readability, truthfulness and trustfulness, completeness, novelty, obtainability, and more. Test collections in which documents are judged for utility given a specific task and context would allow researchers and developers to build and train systems that are more aware of tasks and user needs.

This idea is not new; it goes back to the late 60s and especially a number of papers by Cooper through the 70s (Stefano Mizzaro's review of the concept of relevance briefly describes much of this work [1]). But it hasn't been applied much, possibly because there are so many dimensions on which one can discuss "utility" that only looking at one or two at a time is even feasible. A few recent TREC tracks have done this: the Contextual Suggestion track, the Web track's diversity task.

Mark Rorvig argued that utility can be sufficiently modeled with preference judgments [2]: give an assessor two documents and a context, and ask which document they would prefer in that context. These preference judgments capture utility without needing to enumerate and judge against every possible aspect of utility. We have been applying this idea to building large collections of preferences that capture novelty and diversity along with relevance and other aspects of utility [3, 4].

References

1. Stefano Mizzaro. Relevance: the whole history. Journal of the American Society for Information Science, 48(9):810–832, 1997.
2. Mark E. Rorvig. The simple scalability of documents. *Journal of the American Society for Information Science*, 41(8):590–598, 1990.
3. Ben Carterette, Paul N. Bennett, David Maxwell Chickering, and Susan T. Dumais. Here or there: preference judgments for relevance. In *Proceedings of the IR research, 30th European Conference on Advances in Information Retrieval*, ECIR'08, pages 16–27, Berlin, Heidelberg, 2008. Springer-Verlag.
4. Praveen Chandar and Ben Carterette. Using preference judgments for novel document retrieval. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 861–870, New York, NY, USA, 2012. ACM.

## Fernando Diaz

What are appropriate auxiliary tools for different types of search tasks? Previously studied tools include query and URL history. However, it may be that finer-grained specialization of tools may be helpful. For example, when a user is researching a product, supplying a simple spreadsheet for price or review information may be useful; when a user is planning a trip, decomposing an interface into trip subtasks (e.g. accommodation, plane tickets) may be useful.

Can we adaptively augment traditional search interfaces with these auxiliary tools?

References

1. Dan Morris, Meredith Ringel Morris, and Gina Venolia. Searchbar: a search-centric web history for task resumption and information refinding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1207–1216, New York, NY, USA, 2008. ACM.
2. Henry Allen Feild and James Allan. Task-aware search assistant. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 1015–1015, New York, NY, USA, 2012. ACM.
3. Debora Donato, Francesco Bonchi, Tom Chi, and Yoelle Maarek. Do you want to take notes?: identifying research missions in yahoo! search pad. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 321–330, New York, NY, USA, 2010. ACM.

## Abdigani Diriye

One of the challenges stifling work on task-aware systems is identifying and mapping out the kind of search support and features needed to help users during different search tasks. The challenge here is identifying the inherent search activities the user might be engaged in, and the set of features and functionality that would best support them.

References

1. Gene Golovchinsky, Abdigani Diriye, and Tony Dunnigan. The future is in the past: designing for exploratory search. In *Proceedings of the 4th Information Interaction in Context Symposium*, IIIX '12, pages 52–61, New York, NY, USA, 2012. ACM.  The above paper provides a good introduction on how to design for more complex and exploratory search tasks and some of the factors that need to be kept in mind.

## Susan T. Dumais

Identifying tasks using implicit interactions. This is especially important for tasks that extend across time and devices. The references below provide examples of techniques for identifying queries related to tasks, for predicting whether a task will be resumed, and looking at tasks over a longer time scale.

References

1. Alexander Kotov, Paul N. Bennett, Ryen W. White, Susan T. Dumais, and Jaime Teevan. Modeling and analysis of cross-session search tasks. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '11, pages 5–14, New York, NY, USA, 2011. ACM.
2. Debora Donato, Francesco Bonchi, Tom Chi, and Yoelle Maarek. Do you want to take notes?: identifying research missions in yahoo! search pad. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 321–330, New York, NY, USA, 2010. ACM.

Thinking broadly about what support for search tasks looks like. The references below provide examples from simple "answers" seen in web search engines, to apps for specific tasks, to richer environments for exploratory search.

References

1. Lydia B. Chilton and Jaime Teevan. Addressing people's information needs directly in a web search result page. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, pages 27–36, New York, NY, USA, 2011. ACM.
2. 50 ultimate travel apps ... so far

## Luanne Freund

We still do not know very much about how tasks influence search, or more specifically: what are the task-based requirements of IR systems? Much of the research on task-based IR has focused on behavioural analyses of searchers in different task contexts, which informs our understanding of task as a contextual variable that influences behaviour, but does not necessarily have design implications for search.

This is a multifaceted problem, as it involves the relationships between task characteristics, document characteristics and characteristics of retrieval systems. We have descriptive models of each of these components that can help up identify key characteristics, but we are lacking in theoretical and empirical models that identify the relationships between them that are most likely to influence search outcomes. The empirical studies that we do have are of limited value due to the lack of a standard nomenclature for tasks and the idiosyncratic operationalization of task characteristics in assigned search tasks.

References

1. Norbert Fuhr. Salton award lecture information retrieval as engineering science. *SIGIR Forum*, 46(2):19–28, 2012.
    My thinking about this problem has been influenced by the 2012 SIGIR Salton Award keynote delivered by Norbert Fuhr, in which he discusses the need for an engineering approach in IR that would allow us to predict the kinds of systems and features needed in response to particular domain and task scenarios. The paper points us towards to importance of developing theoretical models of task-based IR as well as conducting more carefully controlled and systematic empirical studies to test and further develop these models.
2. Robert Capra, Gary Marchionini, Jung Sun Oh, Fred Stutzman, and Yan Zhang. 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, pages 442–451, New York, NY, USA, 2007. ACM.
    There is very little published research that predicts and tests for task-based effects of retrieval system features on retrieval outcomes rather than user behaviours. The Capra et al. (2007) study comes close, as it examines relationships between task types, interaction styles and information architecture.
3. Wildemuth, B.M., & Freund, L. (2009). Search tasks and their role in studies of search behaviors. Paper presented at HCIR 2009: Bridging Human-Computer Interaction and Information Retrieval, Washington, DC, October 23, 2009.
    This position paper identifies some of the issues with task characterization and operationalization in interactive IR studies.

## Gene Golovchinsky

I think the biggest obstacle to the deployment of task-aware systems is lack of understanding when such systems may be useful. When it's clear that records of prior interaction can be used to inform subsequent system behavior, this information is already incorporated into systems. There are no significant technical difficulties to start down this road. The biggest challenge is one of perception: because Google doesn't do something, doesn't mean that that something isn't possible or desirable in other contexts.

## Jaap Kamps

To build an information access tool that actively supports a searcher to articulate a whole search task, and to interactively explore the results of every stage of the process. There is a striking difference in how we ask a person for information, giving context and articulating what we want and why, and how we communicate with current search engines. Current search technology requires us to slice-and-dice our problem into several queries and sub-queries, and laboriously combine the answers post hoc to solve our tasks. Combining different sources requires opening multiple windows or tabs, and cutting-and-pasting information between them. Current search engines may have reached a local optimum for answering micro information needs with lighting speed. Supporting the overall task opens up new ways to significantly advance our information access tools, by develop tools that are adapted to our overall tasks rather than have searchers adapt their search tactics to the "things that work."

References

1. Ian Ruthven. Interactive information retrieval. ARIST, 42(1): 43-91 (2008).
   http://dx.doi.org/10.1002/aris.2008.1440420109
   Solid overview of how much we know about the interaction, also immediately highlighting how little we know about the mechanics of interaction during a process of performing a complex task.
2. Arjen P. de Vries, Wouter Alink, Roberto Cornacchia: Search by strategy. *ESAIR* 2010:27-28
   http://doi.acm.org/10.1145/1871962.1871979
   Interesting new approach to formulate a complex query (or search strategy) for tasks of increasing complexity.

Can we make a retrieval system aware of the searcher's stage in the information seeking process, tailor the results to each stage, and guide the searcher through the overall process? A search session for a non-trivial search task consists of stages with different sub-goals (e.g., problem identification) and specific search tactics (e.g., reading introductory texts, familiarizing with terminology). Making a system aware of a searcher's information seeking stage has the potential to significantly improve the search experience. Searchers are stimulated to actively engage with the material, to get a grasp on the information need and articulate effective queries, to critically evaluate retrieved results, and to construct a comprehensive answer. This may be of particularly great help for those searchers having poor information or media literacy. This is of obvious importance in many situations: e.g., education, medical information, and search for topics "that matter." Some special domains, such as patent search and evidence based practices in medicine, have clearly prescribed a particular information seeking process in great detail. Here building a systems to support (and enforce) this process is of obvious value.

References

1. Marcia J. Bates: Where should the person stop and the information search interface start? Inf. Process. Manage. 26(5): 575-591 (1990) http://dx.doi.org/10.1016/0306-4573(90)90103-9
   There is a need for a new discussion on what role the system and user play, and how the interface supports the task progress as well as the information seeking process.
2. Forest Woody Horton: Understanding information literacy: a primer; an easy-to-read, non-

technical overview explaining what information literacy means, designed for busy public policy-makers, business executives, civil society administrators and practicing, UNESCO, 2008 http://unesdoc.unesco.org/images/0015/001570/157020e.pdf
Information/Media literacy research has many relations and basically outlines what type of information seeking behavior should be promoted by the system.

Meta questions on how to foster collaboration between research groups in computer science and information science, and in academia and industry, so that we could work *together* on solving some of these challenges in the near future.

## Bill Kules

Research Problem: Design of exploratory search tasks for search system evaluation

Evaluation is an essential part of developing search tools that are more task aware, particularly for exploratory search, which is a recognized challenge for information seeking systems and an area of active research and development.

For any user study, tasks must be carefully constructed to balance ecological validity with experimental control. For exploratory search, this is a particular challenge, because we are trying to induce search behaviors that are inherently open-ended. Individual searchers have to interpret the task, formulate their own queries and evaluate the results based on their understanding of the information need and their own knowledge and experience. At the same time, we wish to maintain some level of experimental control to permit comparisons between systems and longitudinally.

Borlund (2003) developed the concept of a simulated work task, which forms the basis for many user evaluations of search systems. Many studies have used the simulated work task as the basis for search tasks, but tasks are rarely comparable between or even within studies, limiting our ability to build up a corpus of results in a manner similar to the TREC studies. Recent work has started to formalize attributes of exploratory search tasks and provide suggestions for how to create and validate such tasks (Kules and Capra, 2012; Wildemuth and Freund, 2012). There are a number of open questions to be investigated. Three of them are:

1. What is an appropriate, parsimonious set of attributes to define exploratory search tasks?
2. How can we quantify (can we quantify) measures for these attributes?
3. Given that searchers individually interpret tasks and results, what comparisons does this allow us to make between systems and studies?

References

1. Pia Borlund. The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, 8(2), 2003.This paper developed the concept of a simulated work task. It has formed the basis for much user-focused systems evaluation.
2. Bill Kules and Robert Capra. Influence of training and stage of search on gaze behavior in a library catalog faceted search interface. Journal of the American Society for Information Science and Technology, 63(1):114–138, 2012.This paper extends the simulated work task concept to create exploratory search tasks. It incorporates 7 additional criteria and describes an iterative procedure to create and validate exploratory search scenarios for a semi-controlled laboratory study.
3. Barbara M. Wildemuth and Luanne Freund. Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval*, HCIR '12, pages 4:1–4:10, New York, NY, USA, 2012. ACM.This paper reviews and critiques 84 papers related to exploratory search task design and proposes a set of design recommendations.

## Birger Larsen

One way of progressing for systems to become more task aware is to facilitate research by considering if it is possible and fruitful to extend the Cranfield paradigm to support experiments with task based search. What are the demands on topics and relevance assessment to support task based experiments, and what additional procedures and performance measures are needed? Can the complexity be handled and what could be learned from such experiments?

References

1. Marianne Lykke, Birger Larsen, Haakon Lund, and Peter Ingwersen. Developing a test collection for the evaluation of integrated search. In *Proceedings of the 32nd European Conference on Advances in Information Retrieval*, ECIR'2010, pages 627–630, Berlin, Heidelberg, 2010. Springer-Verlag.This poster paper describes the iSearch test collection, where we put much more emphasis on obtaining through and structured descriptions of the work tasks and information needs. This may be one step towards task based search as it facilitates experiments with extended task descriptions.

## Christina Lioma

One potentially interesting aspect of task-aware search is the ranking model that estimates the relevance of the retrieved results. Traditionally, ranking models are grounded on mathematical estimations, such as metric distance or probabilities, and often include empirically-tuned parameters. It is not uncommon to use the exact same ranking model in different search tasks. However, relevance should not necessarily always be treated uniformly across different tasks. Task-based ranking models could be considered, taking as a starting point advances in dynamic similarity measures, which are partly tuneable at query time manually by the user (Bustos and Skopal 2006), or which accommodate various different task-based similarity functions (Ciaccia and Patella 2009). These papers present the two examples of dynamic similarity measures mentioned above:

References

1. Benjamin Bustos and Tomáš Skopal. Dynamic similarity search in multi-metric spaces. In *Proceedings of the 8th ACM International workshop on Multimedia Information Retrieval*, MIR '06, pages 137–146, New York, NY, USA, 2006. ACM.
2. Paolo Ciaccia and Marco Patella. Principles of information filtering in metric spaces. In *Proceedings of the 2009 Second International Workshop on Similarity Search and Applications*, SISAP '09, pages 99–106, Washington, DC, USA, 2009. IEEE Computer Society.

## Jingjing Liu

For multi-session tasks, how can search systems perform better, at different stages, and for different task types (e.g., tasks with different structures, difficulty/complexity levels, life vs. scholarly tasks, actionable vs. informational tasks, etc.)?

Frequently seen in everyday life, multi-session tasks are usually complex and require multi-sessions to complete. While IR systems do a decent job with simple search tasks, there's much room for them to improve in multi-session tasks. How can systems be better designed to facilitate users' finding and re-finding of information in multi-session tasks? What system features will be supportive and preferred by users? Understanding multi-session task features, user behaviors, and system features are all important to address this question.

References

1. Jaime Arguello, Wan-Ching Wu, Diane Kelly, and Ashlee Edwards. Task complexity, vertical display and user interaction in aggregated search. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 435–444, New York, NY, USA, 2012. ACM.Arguello et al. (2012) addressed task features (complexity) and the system interface feature (vertical display) as well as their interaction with users in aggregated search. This could be very relevant to and beneficial in dealing with multi-session tasks.
2. Alexander Kotov, Paul N. Bennett, Ryen W. White, Susan T. Dumais, and Jaime Teevan. Modeling and analysis of cross-session search tasks. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '11, pages 5–14, New York, NY, USA, 2011. ACM.Kotov et al. (2011) showed that it is possible to effectively model and analyze users' cross-session search behaviors. Two problems they dealt with were: 1) identifying related queries to a current one from previous sessions, and 2) given a multi-query task, predicting if the user will return to the task in the future. This research is helpful for search systems to determine task context and suggest queries for multi-session tasks.

What task features make a search difficult? And how can systems better support "difficult" tasks according to the reasons why they are difficult?

Byström, K. & Järvelin (1995) and Byström, K. (2002) explored the effect of task complexity (defined as the a prior determinability of information inputs, processing, and outputs) on people's information seeking and use in a work task environment. These studies found that with the increase of task complexity, increased the complexity of information needed, the needs for domain information and problem solving information, and the number of sources, but decreased the success. There is a strong link between information types acquired and sources used, and that task complexity has a direct relationship to source use.

Although it is not the same concept as task difficulty, according to Li & Belkin (2008), both represent the information seeker's perception that the information seeking is not easy. More qualitative studies like these are needed to understand what task features make IR system users feel "difficulty". These will help design systems can better support "difficult" tasks according to the reasons why they are difficult.

References

1. Katriina Byström and Kalervo Järvelin. Task complexity affects information seeking and use. *Information Processing and Management*, 31(2):191– 213, March 1995.
2. Katriina Byström. Information and information sources in tasks of varying complexity. *Journal of the American Society for Information Science*, 53(7):581–591, 2002.
3. Yuelin Li and Nicholas J. Belkin. A faceted approach to conceptualizing tasks in information seeking. *Information Processing and Management*, 44(6):1822–1837, 2008.

## Gary Marchionini

An overarching problem is two-fold: user context elicitation and use. By this I mean determining what and how information seekers learn over sessions and correspondingly how systems might assist this process.

A second, more specific problem is how to represent search history to users.

References

1. Robert Capra, Gary Marchionini, Javier Velasco-Martin, and Katrina Muller. Tools-at-hand and learning in multi-session, collaborative search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 951–960, New York, NY, USA, 2010. ACM.

## Catherine Smith

Research Problem 1: the need to study transitions between task-specific applications and search sub-tasks.

As Belkin (2009) stated, "… we might say that an ultimate goal of, and challenge for IR research is to arrange things such that a person never has to engage with a separate IR system at all (although I am quite willing to agree that there are certainly circumstances in which such engagement might be indeed desirable.)." In this view, the burden of acquiring useful task descriptions (useful to the retrieval system) might be handled by applications that support "parent-task" goals (with a parent-task defined as any task that invokes an information search sub-task). In order to exploit task-related data available from such an application, we need to study transitions between search sub-tasks and parent-tasks.

References

1. Nicholas J. Belkin. Really supporting information seeking: A position paper. Information Seeking Support Systems. Technical Report. 2009.

Following from above, as an example of transitions, one can imagine search sub-tasks interleaved with active reading, where reading is the parent-task. An application like the one described by Hinkley, Bi, Pahud, & Bixton (2012) might collect implicit and/or explicit task-related data, which it could pass to a search utility when search sub-tasks are invoked. We need to describe transitions, and investigate how transitions may be improved for the user. Toms, Villa, & McCay-Peet (2013) is an example of an experimental study along these lines. The authors state their objective as, "… to explore the boundaries of the work task and search process to examine how users integrate search with the larger task" (p. 16). The study used an active reading interface which was developed by the researchers, and was an integral component of a larger experimental retrieval system. Work on this problem would be further advanced by collaborations with HCI researchers designing task-specific applications. This is particularly important if we are to consider an architecture that enables coupling of task applications and a search utility.

References

1. Elaine G. Toms, Robert Villa, and Lori McCay-Peet. How is a search system used in work task completion? *Journal of Information Science*, 39(1):15–25, 2013.
2. Ken Hinckley, Xiaojun Bi, Michel Pahud, and Bill Buxton. Informal information gathering techniques for active reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1893–1896, New York, NY, USA, 2012. ACM.

## Mark Smucker

In our work on time-biased gain (TBG), Charlie Clarke and I have written about how TBG has no presupposed notion of gain, or of user interfaces, or even of retrieval systems. What matters to time-biased gain is that we have some way of estimating the gain achieved by the user over time.

This workshop concerns itself with search tasks where the user wants to "cultivate a deeper understanding of a problem or topic" and where the task requires "sustained interaction and engagement with information". The notion that lengthy interactions with information are central to task-based search implies that not only will gain likely be spread out over a long time period, but that gain may not be simply accumulated on acquisition of relevant material. I wonder to what extent our notions of gain in search must change. Today we think of gain as finding relevant documents, but will that be the correct model of gain for task-based search?

Cooper (1973) discusses the notion that each document encountered in a search session should have some positive or negative utility. In Cooper's formulation of the problem, the retrieval system's job is to deliver documents and the user can report to us the utility of each document. If we see our IR systems as becoming more than tools for retrieval of documents, we may need new measures of gain. For example, if our IR systems became designed for supporting creative work, we might need a measure of gain similar to the creativity support index of Carrol and Latulipe (2009). Or, perhaps we need to start measuring and modeling negative utility along the lines of searcher frustration as done by Feild, Allan, and Jones (2010). Once we know how to measure gain for users, we will then be faced with the task of how to incorporate these notions of gain into our Cranfield-style evaluations of task-based search.

References

1. Erin A. Carroll and Celine Latulipe. The creativity support index. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, pages 4009–4014, New York, NY, USA, 2009. ACM.
2. William S. Cooper. On selecting a measure of retrieval effectiveness. *Journal of the American Society for Information Science*, 24(2):87–100, 1973.
3. Henry Allen Feild and James Allan. Task-aware search assistant. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 1015–1015, New York, NY, USA, 2012. ACM.
4. Mark D. Smucker and Charles L.A. Clarke. Time-based calibration of effectiveness measures. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 95–104, New York, NY, USA, 2012. ACM.

## Simone Stumpf

There are two areas for task-based search systems that I think are interesting to explore:

1. Understanding task-based search for non-text items.    Not all search is for text-based items; users' search tasks also include images, music and videos. Research into searching for these items is limited and fragmented. Previously, there has been some work to understand how users search for images (Westman 2009), however there is a growing realization that more information is needed that take the context and background of the user into account to support them in their task-based search. More recently, there has been increasing interest in searching for and in videos (Smeaton 2007).

2. Providing better cues and "scent" in task-based search.    Search engines results pages on the web have moved on from being just a collection of ranked items and they now provide subtle cues for the user to get to the information that they want via snippets, visual previews, etc. However, there are two issues surrounding this. Firstly, this functionality is usually not available to users on their personal storage systems and they may rely on cues of association (Chau et al. 2008). Secondly, there is a lack of understanding of the role these cues play in users' task-based search (Woodruff et al. 2001).

References

1. Allison Woodruff, Andrew Faulring, Ruth Rosenholtz, Julie Morrsion, and Peter Pirolli. Using thumbnails to search the web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 198–205, New York, NY, USA, 2001. ACM.
2. Alan F. Smeaton. Techniques used and open challenges to the analysis, indexing and retrieval of digital video. *Information Systems*, 32(4):545– 559, 2007.
3. Stina Westman. Image Users' Needs and Searching Behaviour, pages 63–83. John Wiley & Sons, Ltd, 2009.
4. Allison Woodruff, Andrew Faulring, Ruth Rosenholtz, Julie Morrsion, and Peter Pirolli. Using thumbnails to search the web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 198–205, New York, NY, USA, 2001. ACM.

Fundamentally - what is a "task"? There are so many different understandings of this term and it really matters, as any systems that are developed rest on a basic assumption of what is meant by "task".

References

1. Victor M. Gonzalez and Gloria Mark. Constant, constant, multi-tasking craziness: managing multiple working spheres. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 113–120, New York, NY, USA, 2004. ACM.This paper is one take on what a task is but not necessarily for searching.

## Jaime Teevan

Defining task boundaries. While some tasks (like buying a car or planning a trip) are clearly defined, others (like planning summer activities or doing research) are much harder to identify because they evolve, change, are part of larger tasks, and consist of sub-tasks. It can be very hard for a person -- let alone a computer -- to clearly identify task boundaries, but clear task definition may be important for tools that want to support task-based search.

References

1. Rosie Jones and Kristina Lisa Klinkner. Beyond the session timeout: automatic hierarchical segmentation of search topics in query logs. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, CIKM '08, pages 699–708, New York, NY, USA, 2008. ACM.Jones and Klinkner [CIKM 2008] discuss the challenges with identifying tasks (which they call "search missions") versus sessions, and present one way to do so automatically.

## Elaine Toms

In this discipline, we seem trapped in the user-centred paradigm; not everything information-oriented is about the user and search behavior. A task may exist in isolation from the people who accomplish it. Tasks emerge out of an organizational environment and their resolution supports some organizational outcome. Tasks have clear objectives, but may have multiple outcomes, and multiple ways of reaching that outcome. One could conceive of the user as a convenient slave/robot/handmaiden to get the task done. The challenge is twofold:

1. Understanding the process: think Henry Ford and project the automobile assembly line a century later when the task has information components that have to be mixed, scrapped, stirred, and moulded although not in quite the same physical way. We do not know how similar typical "knowledge work" tasks are to, for example, the tasks that occur on an automobile assembly line, although there have been clues as demonstrated by the work on how people write papers and proposals.
2. Understanding which "sledge hammer, drill or screwdriver" the "slave" needs to get the job done; perhaps less like the automobile assembly line, much of knowledge work requires human intervention in the form or decision making that requires intense cognitive activity. What tools does the slave need to assist with the job?

In the context of knowledge work, what are those generic tasks that are shared by many contexts, that is, which ones are comparable to, for example, the "cut and paste" tasks of the desktop application work? which ones are context specific, for example comparable to the produce a slide show in a presentation software? Which ones require finding data and/or information? Which ones require using information? Which ones rely on the talents of the slave because the technology is still not sophisticated enough to do the task from beginning to end, and how do we assist the slave with more useful tools?

Does the approach used by Bartlett in decomposing a bioinformatics task fit with other types of "knowledge work" tasks? The Kulthau and Vakkari work on writing proposals and papers goes a long way toward decomposing task in an educational context (although they may not see it that way). What "cognitive protheses" do we need to develop to support task completion? An interesting and short note that defines this concept:
http://www.lpi.usra.edu/publications/reportsCB-1089/ford.pdf

Why have we never done a formal requirements analysis for any of our information solutions? Take even the digital library. Its design is based on past practices.

References

1. Joan C. Bartlett and Elaine G. Toms. Developing a protocol for bioinformatics analysis: An integrated information behavior and task analysis approach: Research articles. *Journal of the American Society for Information Science*, 56(5):469–482, 2005.

## Pertti Vakkari

Understanding more in detail how larger tasks are related to search tasks and searching. By tasks I mean information intensive work tasks, which generate several search sessions. Empirical results hint that various aspects of search process like term selection, querying, relevance judgment and the information utilized vary between search sessions when task performance proceeds. In order to understand the role of various activities (stages) in the search process within and between sessions it is necessary to understand the whole search process and how it is associated with task performance. This is important 1) theoretically for understanding the phenomenon we are interested in, 2) for system design to better match the tools with human activities from the angle of both search tasks and work tasks, and 3) for creating evaluation procedures and metrics for task-based search.

References

1. Jingjing Liu and Nicholas J. Belkin. Searching vs. writing: Factors affecting information use task performance. *Proceedings of the American Society for Information Science and Technology*, 49(1):1–10, 2012.

   In an experimental longitudinal setting Liu & Belkin studied the associations between newspaper article writing tasks and information searching and use in three points in time during the preparation of the article. The results in this and other articles from the same experiment are important because they have extended our understanding of how some features of tasks and task performance are related to various aspects of searching and utility assessments.

2. Pertti Vakkari and Saila Huuskonen. Search effort degrades search output but improves task outcome. *Journal of the American Society for Information Science and Technology*, 63(4):657–670, 2012.

   In a field study Vakkari & Huuskonen examined how medical students' search effort for an assigned essay writing task was associated to precision and relative recall, and how this was associated to the quality of the essay. They found out that effort in the search process degraded precision, but improved task outcome. The poorer the precision, the better the quality of the essay. The findings concerning the whole process are important, because they suggest that traditional effectiveness measures in information retrieval are not sufficient for task-based searching. They should be complemented with evaluation measures for search process and task outcome.

## Arjen P. de Vries

How do we design and evaluate search systems (and their retrieval models) given that we know that relevance is not just topical?

References

1. William S. Cooper. Expected search length: A single measure of retrieval effectiveness based on the weak ordering action of retrieval systems. *American Documentation*, 19(1):30–41, 1968.The paper is useful because it takes an effort oriented view on the evaluation of systems (actually, developing also a measure that quantifies a system's success in terms of the reduction in effort that can be expected). I have been surprised about the fact that the measures proposed in this work are hardly ever used. (Last year's best SIGIR paper may bring it back into the picture, who knows.)
2. Kevyn Collins-Thompson, Paul N. Bennett, Ryen W. White, Sebastian de la Chica, and David Sontag. Personalizing web search results by reading level. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, CIKM '11, pages 403–412, New York, NY, USA, 2011. ACM.  This paper considers readability in IR, demonstrating how it is a relevance criterion that matters.

How can we tailor the retrieval model to the task? What part can we automate in this tailoring process, and what part will remain the designer's task?

References

1. Stefano Ceri, Alessandro Bozzon, and Marco Brambilla. The anatomy of a multi-domain search infrastructure. In *Web Engineering, volume 6757 of Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 2011.The Search Computing project developed a powerful search environment, that allows to define searches that integrate multiple sources. The more traditional database approach that motivates this project does emphasize that we could perhaps benefit from more abstraction in defining how search will take place, instead of the usual focus on a specific instance.
2. Norbert Fuhr. Logical and conceptual models for the integration of information retrieval and database systems. In *East/West Database Workshop*, Klaenfurt, pages 206–218. Springer Verlag, 1994.Describes IR as generalization of the DB approach, and argues for a view that would into both the DB, IR and HCI aspects of retrieval in a unified manner.

## Ryen White

1. Characterizing and supporting cross-session and/or cross-device search tasks, including "slow search" support that capitalizes on time between search episodes . Motivation: Complex tasks persist over time. People are using multiple devices more frequently. Need ways to support transitions between devices that capitalizes on the time that search engines may have – both in predicting whether a searcher will resume the task, deciding what action to take to help them (e.g., finding more/better results while the searcher is away from the search engine), and helping them restore their task state.

2. Leveraging on-task behavior of the current user (personalization) and similar users (those in related cohorts). Motivation: On-task behavior is most relevant for personalization. Need ways to automatically identify search tasks and use this task-relevant information to adapt the search experience (results and UX) within the current session and beyond. Also potential benefit from using other searchers' on-task search behavior, especially for addressing the "cold start" problem associated with new users.

3. Understanding and modeling the impact of task and user characteristics on information search behavior. Motivation: Attributes of the user (e.g., their domain knowledge), the search task (e.g., it's complexity), or their relationship (e.g., user familiarity with tasks of this type) affect search behavior. Better understanding these effects and developing user/task models that consider these effects can help design better systems and methodologies (including user simulations learned from sources such as logs) to evaluate these systems.

4. Automatically identifying components of search tasks and guiding users through those stages. Motivation: Complex search tasks have multiple aspects. Automatically identifying those parts can help systems guide users through the stages in a useful sequence. Tours or trails could be shown to searchers as an alternative/complement to existing result lists. These tours can be manually created or determined algorithmically from sources such as search log data.

References

1.      Eugene Agichtein, Ryen W. White, Susan T. Dumais, and Paul N. Bennet. Search, interrupted: understanding and predicting search task continuation. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 315–324, New York, NY, USA, 2012. ACM.
2.      Alexander Kotov, Paul N. Bennett, Ryen W. White, Susan T. Dumais, and Jaime Teevan. Modeling and analysis of cross-session search tasks. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '11, pages 5–14, New York, NY, USA, 2011. ACM.
3.      Yu Wang, Xiao Huang, and Ryen W. White. Characterizing and supporting cross-device search tasks. In *Proceedings of the sixth ACM International Conference on Web Search and Data Mining*, WSDM '13, pages 707–716, New York, NY, USA, 2013. ACM.
4.      James Pitkow, Hinrich Schütze, Todd Cass, Rob Cooley, Don Turnbull, Andy Edmonds, Eytan Adar, and Thomas Breuel. Personalized search. *Communications of the ACM*, 45(9):50–55, September 2002.
5.      Bin Tan, Xuehua Shen, and ChengXiang Zhai. Mining long-term search history to improve search accuracy. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '06, pages 718–723, New York, NY, USA,

2006. ACM.

6.  Zhen Liao, Yang Song, Li-wei He, and Yalou Huang. Evaluating the effectiveness of search task trails. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, pages 489–498, New York, NY, USA, 2012. ACM.

7.  Ryen W. White, Wei Chu, Ahmed Hassan, Xiaodong He, Yang Song, and Hongning Wang. Enhancing personalized search by mining and modeling task behavior. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, New York, NY, USA, To Appear. ACM.

8.  Diane Kelly and Colleen Cool. The effects of topic familiarity on information search behavior. In *Proceedings of the 2nd ACM/IEEE-CS joint Conference on Digital libraries*, JCDL '02, pages 74–75, New York, NY, USA, 2002. ACM.

9.  Katriina Byström and Kalervo Järvelin. Task complexity affects information seeking and use. *Information Processing and Management*, 31(2):191– 213, March 1995.

10.  Yuelin Li and Nicholas J. Belkin. An exploration of the relationships between work task and interactive information search behavior. *Journal of the American Society for Information Science and Technology*, 61(9):1771– 1789, 2010.

11.  Ryen W. White, Susan T. Dumais, and Jaime Teevan. Characterizing the influence of domain expertise on web search behavior. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, WSDM '09, pages 132–141, New York, NY, USA, 2009. ACM.

12.  Ryen W. White, Ian Ruthven, Joemon M. Jose, and C. J. Van Rijsbergen. Evaluating implicit feedback models using searcher simulations. *ACM Transactions of Information Systems*, 23(3):325–361, 2005.

13.  Catherine Guinan and Alan F. Smeaton. Information retrieval from hypertext using dynamically planned guided tours. In *Proceedings of the ACM Conference on Hypertext*, ECHT '92, pages 122–130, New York, NY, USA, 1992. ACM.

14.  Adish Singla, Ryen White, and Jeff Huang. Studying trail finding algorithms for enhanced web search. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '10, pages 443–450, New York, NY, USA, 2010. ACM.

15.  Ahmed Hassan and Ryen W. White. Task tours: helping users tackle complex search tasks. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12, pages 1885–1889, New York, NY, USA, 2012. ACM.

16.  Randall H. Trigg. Guided tours and tabletops: tools for communicating in a hypertext environment. *ACM Transactions of Information Systems*, 6(4):398–414, 1988.

17.  Alan Wexelblat and Pattie Maes. Footprints: history-rich tools for information foraging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 270–277, New York, NY, USA, 1999. ACM.

## Barbara Wildemuth

Many studies have used one or more attributes of search tasks as an independent variable and examined various search behaviors (e.g., search terms selected, search strategy formulation and re-formulation, or browsing behavior) as the dependent variable. Many of these have found some type of effect, but not all of them have. Which task attributes are most worthwhile to incorporate in future studies of this type? Are there any that consistently show no effect on search behaviors or outcomes?

References

1. Yuelin Li and Nicholas J. Belkin. An exploration of the relationships between work task and interactive information search behavior. *Journal of the American Society for Information Science and Technology*, 61(9):1771– 1789, 2010.Li's taxonomy of task attributes contributed to this empirical study; it's useful because it demonstrates a strong (though not perfect, I think) conceptual foundation for an empirical study.
2. Elaine G. Toms, Luanne Freund, Richard Kopak, and Joan C. Bartlett. The effect of task domain on search. In *Proceedings of the 2003 Conference of the Centre for Advanced Studies on Collaborative research*, CASCON '03. IBM Press, 2003.  This is an older study, but it makes me think that we should try to pick some of the low-hanging fruit first. It may be relatively easy to detect the domain in which a person is searching; can we then tune the search engine to better support the searcher?  Many interactive IR studies use (search) task complexity or difficulty as an independent variable. Yet these concepts are rarely defined clearly and have been operationalized in a variety of ways. So that the results of future studies can be compared with each other, we need to come to some agreement on the definitions of search task difficulty and search task complexity. In addition, in many studies, it's not clear whether the focus is on the search task or the work task, so we may also need to come to some agreement on definitions of work task difficulty and work task complexity.

Many interactive IR studies use (search) task complexity or difficulty as an independent variable. Yet these concepts are rarely defined clearly and have been operationalized in a variety of ways. So that the results of future studies can be compared with each other, we need to come to some agreement on the definitions of search task difficulty and search task complexity. In addition, in many studies, it's not clear whether the focus is on the search task or the work task, so we may also need to come to some agreement on definitions of work task difficulty and work task complexity.

References

1. Donald J. Campbell. Task complexity: A review and analysis. *The Academy of Management Review*, 13(1):40–52, 1988.This is a classic review, but outside our field. Campbell treats task complexity as 1) primarily a psychological experience, 2) an interaction between task and person characteristics, and 3) a function of objective task characteristics. Applied to our work today, this framework may provide a basis for our own definitions of task complexity.
2. Katriina Byström and Kalervo Järvelin. Task complexity affects information seeking and use. *Information Processing and Management*, 31(2): 191– 213, March 1995.

3. Jingjing Liu, Chang Liu, Michael Cole, Nicholas J. Belkin, and Xiangmin Zhang. 2012. Exploring and predicting search task difficulty. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12. ACM, New York, NY, USA, 1313-1322.The first of this pair focuses on work tasks and the second focuses on search tasks. While we tend to focus on search tasks (as interactive IR researchers), it may be equally important (or more important) to attend to the complexity or difficulty of work tasks.

**Other research topics**

The role of simulated work task situations/scenarios: Do we all agree that use of simulated work tasks is "best practice" in developing search tasks for experimental studies? Across all types of tasks?

## Biographical Sketches

We are pleased to have the following confirmed attendees for the NSF Task-Based Search Systems Workshop in Chapel Hill, NC:

**Eugene Agichtein**
Emory University, USA

Eugene Agichtein is an Associate Professor of Computer Science at Emory University, where he founded and leads the Emory Intelligent Information Access Laboratory (IR Lab). The active projects in IR Lab include mining searcher behavior and interactions data, modeling social content creation and sharing, and applications to medical informatics. Dr. Agichtein obtained a Ph.D. in Computer Science from Columbia University, and did a Postdoc at Microsoft Research. He has published extensively on web search, information retrieval, and web and data mining. Dr. Agichtein's work has been supported by DARPA, NIH, NSF, Yahoo!, Microsoft, and Google, and has been recently recognized with the A.P. Sloan Research Fellowship and the 'Best Paper' award at the SIGIR 2011 conference.

**Jae-wook Ahn**
Drexel University, USA

Jae-wook Ahn is working as a data scientist at Drexel University since September 2012. He worked as a post-doctoral CRA-NSF Computing Innovation Fellow at the HCIL (Human-Computer Interaction Lab), University of Maryland at College Park. He got his Ph.D. in Information Science from University of Pittsburgh in 2010 and was a member of PAWS (Personalized Adaptive Web Systems) and TALER (Teaching and Learning Research) Labs. He has worked for various projects on exploratory user interfaces for personalized information retrieval, visualization, and recommendation systems. He also worked on dynamic network visualization of participatory social media for his postdoc project. He is currently working for the Digging into Data and Meaningful Concept Display projects, on analyzing metadata from heterogeneous digital libraries and visualizing concepts in museum search environments.

**Jaime Arguello (organizer)**
University of North Carolina, USA

Jaime Arguello is an Assistant Professor at the School of Information and Library Science at the University of North Carolina at Chapel Hill. His research interests are in information retrieval, with a focus on aggregated search—the task of automatically integrating multiple search engines in a single interface. His research has been published in conferences including SIGIR, CIKM, WWW, ECIR, IIiX, HLT-NAACL, CHI, ICWSM, TREC, and DG.O. Jaime is a recipient of the SIGIR 2009 Best Paper Award and the 2011 ECIR Best Student Paper Award. He teaches courses in

Information Retrieval and Text Data Mining. He received his Ph.D. and M.S. in Computer Science from Carnegie Mellon University and his B.S. in Electrical Engineering from Washington University in St. Louis.

**Nick Belkin**
Rutgers University, USA

Nick Belkin is Professor II of Information Science at the School of Communication & Information at Rutgers University. His recent research has focused on personalization of interaction with information, with specific reference to the influence of the task which motivates information seeking on interpretation of searcher behaviors with respect to predicting document usefulness.

**Pia Borlund**
Royal School of Library and Information Science, Denmark

Pia Borlund is Professor at the Royal School of Library and Information Science, Denmark, and docent at the School of Information Sciences, University of Tampere, Finland. Her research focuses on methods for evaluation of systems that support interactive information retrieval (IIR). Her interest in IIR systems evaluation, design and usage brings together three broad areas: interactive information retrieval, human-computer interaction, and information seeking (behaviour). Pia Borlund has conducted research on frameworks and guidelines for performance evaluation of interactive information retrieval systems centered on the context instrument 'simulated work task situation' by involvement of users. Her current research focuses on methodological issues, test design and recommendations for evaluation of user search interaction.

**Katriina Byström**
University of Boras, Sweden

Katriina Byström of the Swedish School of Library and Information Science, University of Borås does research on task-based information seeking and retrieval in workplaces and on information architecture. She is an active member of the academic Library and Information Science community and the co-founder and associate editor of the international Journal of Information Architecture. Her latest research projects are: Better Search Engine focusing on work task based search support and Better Web, the development of digital information and communication milieus on the web. Professor Byström chairs the European Network for Workplace Information (ENWI).

**Rob Capra (organizer)**
University of North Carolina, USA

Robert Capra is an Assistant Professor in the School of Information and Library Science at the University of North Carolina at Chapel Hill. His research interests

include human-computer interaction, personal information management, and digital information seeking behaviors, tools, and interfaces. He holds a Ph.D. in Computer Science from Virginia Tech and M.S. and B.S. degrees in Computer Science from Washington University in St. Louis. At Virginia Tech, he was part of the Center for Human-Computer Interaction where he investigated multi-platform interfaces, information re-finding, and interfaces for digital libraries. Prior to Virginia Tech, he worked in corporate research and development, spending five years in the Speech and Language Technologies group at SBC Communications (now merged with AT&T Labs) where he focused on voice user interfaces, speech recognition, and natural language processing.

**Ben Carterette**
University of Delaware, USA

Ben Carterette is an Assistant Professor of Computer & Information Sciences at the University of Delaware. His research interests include all aspects of information retrieval experimentation, from experimental design to test collection construction to evaluation measures to statistical analysis of experiments, and especially how the needs of users can be better modeled at each stage in a batch-style evaluation. In addition to publishing in venues such as ACM TOIS, SIGIR, CIKM, ECIR, and ICTIR, he has co-coordinated five TREC tracks and co-organized five workshops on IR evaluation.

**Arjen de Vries**
Centrum Wiskunde & Informatica, Delft University of Technology, The Netherlands

Arjen P. de Vries is a tenured researcher at CWI leading the Information Access research group, and a full professor (0.2 fte) in the area of multimedia data management at the Technical University of Delft. De Vries studies the intersection of information retrieval and databases. In November 2009, De Vries co-founded Spinque, a CWI spin-off that provides integrated access to any type of data, customized for information specialist or end user, to produce effective and transparent search results.

**Fernando Diaz**
Microsoft Research, USA

Fernando Diaz is a researcher at Microsoft Research New York. His primary research interest is formal information retrieval models. Fernando's research experience includes distributed information retrieval approaches to web search, interactive and faceted retrieval, mining of temporal patterns from news and query logs, cross-lingual information retrieval, graph-based retrieval methods, and synthesizing information from multiple corpora. Fernando received his PhD from the University of Massachusetts Amherst in 2008. His work on federation won the best paper awards at the SIGIR 2010 and WSDM 2010 conferences. He is a co-organizer of the Temporal Summarization track and Web track at TREC 2013.

**Abdigani Diriye**
Carnegie Mellon University, USA

Abdigani Diriye is a postdoctoral researcher at Carnegie Mellon University's Human-Computer Interaction Institute. He received his PhD from University College London. His PhD studies examined the role search interfaces play during information-seeking, and how we can build more useful and usable search systems. Previously, he has worked in the area of collaborative, social and multimedia search whilst interning at FX Palo Alto Labs, Microsoft Research and the Knowledge Media Institute. Currently, his research focuses on ways we can leverage human- and machine-generated data to support people when searching and sensemaking on the Web.

**Susan Dumais**
Microsoft Research, USA

Susan Dumais is a Principal Researcher and manager of the Context, Learning and User Experience for Search (CLUES) Group at Microsoft Research. Prior to joining Microsoft Research, she was at Bellcore and Bell Labs, where she worked on Latent Semantic Indexing (a statistical method for concept-based retrieval), interfaces for combining search and navigation, and organizational impacts of new technology. Her current research focuses on user modeling and personalization, context and information retrieval, temporal dynamics of information, interactive retrieval, and novel evaluation methods. She has worked closely with several Microsoft groups (Bing, Windows Desktop Search, SharePoint Portal Server, and Office Online Help) on search-related innovations. Susan has published more than 200 articles in the fields of information science, human-computer interaction, and cognitive science, and holds several patents on novel retrieval algorithms and interfaces. Susan is also an adjunct professor in the Information School at the University of Washington. She is Past-Chair of ACM's Special Interest Group in Information Retrieval (SIGIR), and serves on several editorial boards, technical program committees, and government panels. She was elected to the CHI Academy in 2005, an ACM Fellow in 2006, received the SIGIR Gerard Salton Award for Lifetime Achievement in 2009, and was elected to the National Academy of Engineering (NAE) in 2011.

**Luanne Freund**
University of British Columbia, Canada

Luanne Freund is an Assistant Professor at the iSchool at the University of British Columbia, Canada. She received her PhD from the University of Toronto in 2008. Luanne's research is focused on interactive information retrieval, human information behaviour, and the effects of task and document genre on search. Current projects include the Systematic Review of Imposed Search Tasks (http://ils.unc.edu/searchtasks), which investigates the use of assigned search tasks in experimental studies; E-Informing the Public (http://diigubc.ca/research/egovernment), which is focused on the design of task

and genre enhanced search systems to support public access to e-government information; Next Generation Information Access - NGAIA (http://diigubc.ca/ngaia), which is focused on the problem of domain-specific information retrieval, and Access to News Media, which seeks to support information seeking in the online news domain. Her research is funded by the Canadian Social Sciences and Humanities Research Council and the Graphics, Animation and New Media Network of Centres of Excellence, Canada.

**Gene Golovchinsky**
FxPAL, USA

Gene Golovchinsky is a Sr. Research Scientist at FX Palo Alto Laboratory, Inc. His research interests include the design of interfaces for interactive information seeking, collaborative search, HCIR, dynamic hypertext, the role of visualization in information retrieval, and pen-based computing.

**Jaap Kamps**
University of Amsterdam, The Netherlands

Jaap Kamps is an Associate Professor of information retrieval at the University of Amsterdam. His research interests span all facets of information storage and retrieval, a common element is the combination of textual information with additional structure, such as document structure, Web-link structure, and/or contextual information, such as metadata, anchors, tags, or clicks. He leads research projects on retrieval with structured information, and projects with cultural heritage institutions (museums, archives, libraries).

**Diane Kelly (organizer)**
University of North Carolina, USA

Diane Kelly is an Associate Professor at the School of Information and Library Science at the University of North Carolina at Chapel Hill. Her research interests are in interactive information search and retrieval, information search behavior and evaluation methods and metrics. Her research has been published in several conferences and journals including ACM SIGIR, ACM CHI, CIKM, IIiX, JCDL, Transactions on Information Systems, Information Processing and Management, JASIST, IEEE Computer and CACM. She teaches undergraduate and graduate courses on research design, interactive information retrieval and foundations of information science. She has served on the UNC Behavioral Institutional Review Board (IRB) since 2005. She received a Ph.D. in Information Science and a Graduate Certificate in Cognitive Science from Rutgers University and an undergraduate degree in Psychology from the University of Alabama.

**Bill Kules**
Catholic University of America, USA

Bill Kules is Associate Professor and Associate Dean for Research and Administration in the School of Library and Information Science (SLIS) at The Catholic University of America. His research interests include faceted search tools to support complex information seeking tasks like exploratory search, and online teaching/learning environments to support engagement and enjoyment. Dr. Kules was a co-organizer of the workshop on Human-Computer Interaction and Information Retrieval (HCIR), an annual meeting of researchers and practitioners whose work spans the fields of human-computer interaction and information retrieval. He was also an organizer of the 2005 Workshop on Exploratory Search Interfaces, a guest editor for the April 2006 CACM Special Section on Supporting Exploratory Search, and is a guest co-editor for a special issue (in press) on HCIR for the journal Information Processing & Management. Before joining SLIS Dr. Kules spent 20 years designing and implementing information systems for a variety of applications, including wireless telephony, customer service, banking, and a multimedia web sites. He earned his Ph.D. in Computer Science at the University of Maryland.

**Birger Larsen**
Royal School of Library and Information Science, Denmark

Birger Larsen (born 1974) is Associate Professor at the Royal School of Library and Information Science in Copenhagen, Denmark since 2006. From 2010 he is leader of the 'Information Systems and Interaction Design' research group. He has a passion for research that involves the activities, processes and experiences arising in the meeting between users, information, and information systems in a given context - with the goal of optimising these to empower users in their task and problem solving. His main research interests include Information Retrieval (IR), structured documents in IR, XML IR and user interaction, domain specific search, understanding user intents and exploiting context in IR, as well as Informetrics/Bibliometrics, citation analysis and quantitative research evaluation.

**Christina Lioma**
University of Copenhagen, Denmark

Christina Lioma is an Assistant Professor and Freja research fellow at the Department of Computer Science, University of Copenhagen, Denmark. Her research focuses on the computational processing of language, mainly in the areas of information retrieval and computational linguistics.

**Jingjing Liu**
University of South Carolina, USA

Jingjing Liu is an Assistant Professor in the School of Library and Information Science at University of South Carolina. She received her Ph.D. in Information Science from Rutgers University in 2010. Her research focuses on the design and evaluation of information systems that support information seeking and use, and

the accomplishment of work tasks and search tasks of various types. Her recent projects deal with personalization of information retrieval, examining people's search behaviors in various types of tasks and predicting document usefulness based on user behaviors, understanding and predicting search task difficulty, understanding users' knowledge change in the search process, and exploring the factors that affect people's information task performance. In her research, Jingjing has been gaining and using rich experience and expertise in task type control, design, and analysis of the effects of task types on various aspects of information retrieval. She has published in journals such as Information Processing and Management and Journal of Documentation, as well as conference proceedings such as ACM SIGIR and CIKM.

**Gary Marchionini**
University of North Carolina, USA

Gary Marchionini is Dean and Cary C. Boshamer Professor in the School of Information and Library Science at the University of North Carolina. He teaches courses in human-information interaction, interface design and testing, and digital libraries. He founded the Interaction Design Laboratory at SILS. He has led projects (NSF funded) on user interfaces for video retrieval, statistical tables, and multi-session and collaborative search. He has more than 200 publications over his career.

**Catherine Smith**
Kent State University, USA

Catherine L. Smith is an assistant professor at the Kent State University School of Library and Information Science. She studies how people use interactive search systems, using theories from cognitive psychology in exploring and explaining search behavior. Her current research activities focus on search expertise and how expertise is gained in formal instruction. In a second research area, she studies the effects of semantic priming in search interaction. Cathy's work is motivated by the idea that advanced systems should help people learn how to search when information needs are unfamiliar, uncommon, or complex.

**Mark Smucker**
University of Waterloo, Canada

Mark Smucker is an assistant professor in the Department of Management Sciences at the University of Waterloo, and is cross appointed with the David R. Cheriton School of Computer Science. Mark's research interests include the design, analysis, and evaluation of interactive information retrieval systems. Mark's recent work has focused on making information retrieval evaluation more predictive of actual human search performance. Mark has been a co-organizer of two TREC tracks, and he co-organized a SIGIR 2010 workshop on the evaluation of retrieval systems via the simulation of interaction. He is a recipient of the SIGIR Best Paper award (2012) as well a recipient of the University of Waterloo, Faculty of Engineering's Teaching

Excellence Award (2012). Mark earned his computer science PhD from the University of Massachusetts Amherst in 2008. Prior to his PhD, Mark worked for seven years in industry on projects ranging from recommendation systems to the study of internet advertising behavior.

**Simone Stumpf**
City University London, UK

Simone Stumpf received a PhD in Computer Science in 2001 and a BSc in Computer Science with Cognitive Science in 1996, both from University College London. She joined City University London in 2009 and currently holds a Senior Lecturer position. Previously, she conducted research at Oregon State University and University College London. Her research centres on end-user interactions with intelligent systems and personal information management systems. She is a member of the End Users Shaping Effective Software (EUSES) consortium, an international collaboration to develop and investigate technologies that support end-users to directly influence software behaviour. Dr Stumpf also has professional experience as a User Experience Architect.

**Jaime Teevan**
Microsoft Research, USA

Jaime Teevan is a Senior Researcher in the Context, Learning, and User Experience for Search (CLUES) Group at Microsoft Research, and an Affiliate Assistant Professor in the Information School at the University of Washington. She studies how people use digital information, particularly as related to their social and temporal context, and builds tools to help better support these information interactions. Jaime was named a Technology Review (TR35) 2009 Young Innovator for her research on personalized search. She co-authored the first book on collaborative Web search, and was Chair of the Web Search and Data Mining (WSDM) 2012 conference. Jaime also edited a book on Personal Information Management (PIM), edited a special issue of Communications of the ACM on the topic, and organized workshops on PIM and query log analysis. She has published numerous technical papers, including several best papers, and received a Ph.D. and S.M. from MIT and a B.S. in Computer Science from Yale University.

**Elaine Toms**
Sheffield University, UK

Elaine Toms is Professor of Information Science, and Head of the Information Retrieval Research Group, Information School, University of Sheffield, UK. Her research focuses on how people interact with information systems, including interacting with the content and with the tools to support use; and how to evaluate such systems in a user-centred way. Her current interest lies in the topic of this workshop: how to create more task-specific search tools and widgets to aid the process, and the integration of search into systems that support tasks.

**Pertti Vakkari**
University of Tampere, Finland

Pertti Vakkari is professor of Information Studies at the School of Information Sciences, University of Tampere, Finland. He has been involved in collaborative European research projects. Vakkari has been a member in several program committees of conferences in information studies like ISIC, IIiX and ECIR. He is a member of the editorial board of journals "Journal of Documentation" and "Information Processing & Management". Vakkari's research interests include task-based information searching and use, the use of digital libraries, the evaluation of information search systems, fiction retrieval, and perceived outcomes of public libraries. His publications include several monographs and readers and over 100 articles. He has received ASIS&T SIGUSE Award for Outstanding Contributions to Information Behavior.

**Ryen White**
Microsoft Research, USA

Ryen White is a Senior Researcher at Microsoft Research. His research interests lie in understanding search interaction and in developing tools to help people search more effectively. He received his Ph.D. in Interactive Information Retrieval from the Department of Computing Science, University of Glasgow, United Kingdom, in 2004. Ryen has published over 150 conference papers and journal articles in Web search, log analysis, and user studies of search systems. He has received six best-paper awards, including two at the ACM SIGIR conference (2007, 2010), one at the ACM SIGCHI conference (2011), and one in JASIST (2010). His doctoral research received the British Computer Society's Distinguished Dissertation Award for the best Computer Science Ph.D. dissertation in the United Kingdom in 2004/2005. Ryen has co-organized numerous workshops on information seeking, in particular exploratory search, including an NSF-sponsored invitational workshop, and has guest co-edited special issues in these areas for a variety of outlets, including Communications of the ACM and IEEE Computer. Since 2008, he has co-organized the annual HCIR Symposium. Ryen has served as area chair for many top conferences such as SIGIR, WSDM, WWW, and CIKM, and currently serves on the editorial board of ACM TOIS, ACM TWEB, and the Information Retrieval Journal. In addition to academic impact, his research has been shipped in many Microsoft products, including Bing, Xbox, Internet Explorer, and Lync.

**Barbara Wildemuth**
University of North Carolina, USA

Barbara Wildemuth is Professor and Associate Dean in the School of Information and Library Science at the University of North Carolina at Chapel Hill. Her research focuses on people's use of information and information technologies. In particular, her studies have included investigations of medical students' searching of online databases, law students' use of a Web-based database of legal resources, and the

effects of different interface designs on the effectiveness of database use. Her most recent work includes a methodological study of the search tasks assigned in interactive information retrieval experiments, being conducted in collaboration with Luanne Freund (UBC) and Elaine Toms (Sheffield). Her recent book, "Applications of Social Research Methods to Questions in Information and Library Science," has been adopted as a text in a number of ILS schools in the United States and abroad. Her teaching responsibilities include courses in various aspects of research methods, human information interactions, and information ethics.

# Bibliography

Eugene Agichtein, Ryen W. White, Susan T. Dumais, and Paul N. Bennet. Search, interrupted: understanding and predicting search task continuation. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 315–324, New York, NY, USA, 2012. ACM.

Jae-wook Ahn and Peter Brusilovsky. Adaptive Visualization for Exploratory Information Retrieval. To appear in *Information Processing and Management*.

Jae-wook Ahn, Peter Brusilovsky, Jonathan Grady, Daqing He, and Sue Yeon Syn. Open user profiles for adaptive news systems: help or harm? In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 11–20, New York, NY, USA, 2007. ACM.

Jaime Arguello, Wan-Ching Wu, Diane Kelly, and Ashlee Edwards. Task complexity, vertical display and user interaction in aggregated search. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 435–444, New York, NY, USA, 2012. ACM.

Joan C. Bartlett and Elaine G. Toms. Developing a protocol for bioinformatics analysis: An integrated information behavior and task analysis approach: Research articles. *Journal of the American Society for Information Science*, 56(5), 469–482, 2005.

Marcia J. Bates: Where should the person stop and the information search interface start? Inf. Process. Manage. 26(5): 575-591 (1990) http://dx.doi.org/10.1016/0306-4573(90)90103-9

Nicholas J. Belkin. Really supporting information seeking: A position paper. Information Seeking Support Systems. Technical Report. 2009.

Nicholas J. Belkin. On the Evaluation of Interactive Information Retrieval Systems. In: Larsen, B., Schneider, J.W., & Åström (Eds.). *The Janus Faced Scholar: A Festschrift in Honour of Peter Ingwersen*. Copenhagen: Royal School of Library and Information Science. 13–21, 2010.

Nicholas J. Belkin. Some(what) grand challenges for information retrieval. *SIGIR Forum*, 42(1), 47–54, 2008.

Pia Borlund. The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, 8(2), 2003.

Pia Borlund, Sabine Dreier, and Katriina Byström. What does time spent on searching indicate? In *Proceedings of the 4th Information Interaction in Context Symposium*, IIIX '12, pages 184–193, New York, NY, USA, 2012. ACM.

Oliver Brdiczka, Norman Makoto Su, and James Bo Begole. Temporal task footprinting: identifying routine tasks by their temporal patterns. In *Proceedings of the 15th International Conference on Intelligent user interfaces*, IUI '10, pages 281–284, New York, NY, USA, 2010. ACM.

Benjamin Bustos and Tomáš Skopal. Dynamic similarity search in multi-metric spaces. In *Proceedings of the 8th ACM International workshop on Multimedia Information Retrieval*, MIR '06, pages 137–146, New York, NY, USA, 2006. ACM.

Katriina Byström. Information and information sources in tasks of varying complexity. *Journal of the American Society for Information Science*, 53(7), 581–591, 2002.

Katriina Byström and Kalervo Järvelin. Task complexity affects information seeking and use. *Information Processing and Management*, 31(2), 191– 213, March 1995.

Katriina Byström and Annemaree Lloyd. Practice theory and work task performance: How are they related and how can they contribute to a study of information practices. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–5, 2012.

Donald J. Campbell. Task complexity: A review and analysis. *The Academy of Management Review*, 13(1), 40–52, 1988.

Robert Capra, Gary Marchionini, Jung Sun Oh, Fred Stutzman, and Yan Zhang. 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, pages 442–451, New York, NY, USA, 2007. ACM.

Robert Capra, Gary Marchionini, Javier Velasco-Martin, and Katrina Muller. Tools-at-hand and learning in multi-session, collaborative search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 951–960, New York, NY, USA, 2010. ACM.

Erin A. Carroll and Celine Latulipe. The creativity support index. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, pages 4009–4014, New York, NY, USA, 2009. ACM.

Ben Carterette, Paul N. Bennett, David Maxwell Chickering, and Susan T. Dumais. Here or there: preference judgments for relevance. In *Proceedings of the IR research, 30th European Conference on Advances in Information Retrieval*, ECIR'08, pages 16–27, Berlin, Heidelberg, 2008. Springer-Verlag.

Stefano Ceri, Alessandro Bozzon, and Marco Brambilla. The anatomy of a multi-domain search infrastructure. In *Web Engineering, volume 6757 of Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 2011.

Praveen Chandar and Ben Carterette. Using preference judgments for novel document retrieval. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 861–870, New York, NY, USA, 2012. ACM.

Duen Horng Chau, Brad Myers, and Andrew Faulring. What to do when search fails: finding information by association. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 999–1008, New York, NY, USA, 2008. ACM.

Lydia B. Chilton and Jaime Teevan. Addressing people's information needs directly in a web search result page. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, pages 27–36, New York, NY, USA, 2011. ACM.

Paolo Ciaccia and Marco Patella. Principles of information filtering in metric spaces. In *Proceedings of the 2009 Second International Workshop on Similarity Search and Applications*, SISAP '09, pages 99–106, Washington, DC, USA, 2009. IEEE Computer Society.

Kevyn Collins-Thompson, Paul N. Bennett, Ryen W. White, Sebastian de la Chica, and David Sontag. Personalizing web search results by reading level. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, CIKM '11, pages 403–412, New York, NY, USA, 2011. ACM.

William S. Cooper. Expected search length: A single measure of retrieval effectiveness based on the weak ordering action of retrieval systems. *American Documentation*, 19(1), 30–41, 1968.

William S. Cooper. On selecting a measure of retrieval effectiveness. *Journal of the American Society for Information Science*, 24(2), 87–100, 1973.

Arjen P. de Vries, Wouter Alink, Roberto Cornacchia: Search by strategy. *ESAIR* 2010:27-28 http://doi.acm.org/10.1145/1871962.1871979

Debora Donato, Francesco Bonchi, Tom Chi, and Yoelle Maarek. Do you want to take notes?: identifying research missions in yahoo! search pad. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 321–330, New York, NY, USA, 2010. ACM.

Henry A. Feild, James Allan, and Rosie Jones. Predicting searcher frustration. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '10, pages 34–41, New York, NY, USA, 2010. ACM.

Henry Allen Feild and James Allan. Task-aware search assistant. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 1015–1015, New York, NY, USA, 2012. ACM.

Norbert Fuhr. Logical and conceptual models for the integration of information retrieval and database systems. In *East/West Database Workshop*, Klaenfurt, pages 206–218. Springer Verlag, 1994.

Norbert Fuhr. Salton award lecture information retrieval as engineering science. *SIGIR Forum*, 46(2), 19–28, 2012.

Gene Golovchinsky, Abdigani Diriye, and Tony Dunnigan. The future is in the past: designing for exploratory search. In *Proceedings of the 4th Information Interaction in Context Symposium*, IIIX '12, pages 52–61, New York, NY, USA, 2012. ACM.

Victor M. Gonzalez and Gloria Mark. Constant, constant, multi-tasking craziness: managing multiple working spheres. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 113–120, New York, NY, USA, 2004. ACM.

Catherine Guinan and Alan F. Smeaton. Information retrieval from hypertext using dynamically planned guided tours. In *Proceedings of the ACM Conference on Hypertext*, ECHT '92, pages 122–130, New York, NY, USA, 1992. ACM.

Ahmed Hassan and Ryen W. White. Task tours: helping users tackle complex search tasks. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12, pages 1885–1889, New York, NY, USA, 2012. ACM.

Ken Hinckley, Xiaojun Bi, Michel Pahud, and Bill Buxton. Informal information gathering techniques for active reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1893–1896, New York, NY, USA, 2012. ACM.

Forest Woody Horton: Understanding information literacy: a primer; an easy-to-read, non-technical overview explaining what information literacy means, designed for busy public policy-makers, business executives, civil society administrators and practicing, UNESCO, 2008

Rosie Jones and Kristina Lisa Klinkner. Beyond the session timeout: automatic hierarchical segmentation of search topics in query logs. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, CIKM '08, pages 699–708, New York, NY, USA, 2008. ACM.

Evangelos Kanoulas, Ben Carterette, Mark Hall, Paul Clough, and Mark Sanderson. Overview of the 2011 Session Track. In *Proceedings of 20th Text Retrieval Conference*, TREC '11. 2011.

Evangelos Kanoulas, Ben Carterette, Mark Hall, Paul Clough, and Mark Sanderson. Overview of the 2012 Session Track. In *Proceedings of 21st Text Retrieval Conference*, TREC '12. 2012.

Diane Kelly and Colleen Cool. The effects of topic familiarity on information search behavior. In *Proceedings of the 2nd ACM/IEEE-CS joint Conference on Digital libraries*, JCDL '02, pages 74–75, New York, NY, USA, 2002. ACM.

Alexander Kotov, Paul N. Bennett, Ryen W. White, Susan T. Dumais, and Jaime Teevan. Modeling and analysis of cross-session search tasks. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '11, pages 5–14, New York, NY, USA, 2011. ACM.

Bill Kules and Robert Capra. Influence of training and stage of search on gaze behavior in a library catalog faceted search interface. Journal of the American Society for Information Science and Technology, 63(1), 114–138, 2012.

Yuelin Li and Nicholas J. Belkin. A faceted approach to conceptualizing tasks in information seeking. *Information Processing and Management*, 44(6), 1822–1837, 2008.

Yuelin Li and Nicholas J. Belkin. An exploration of the relationships between work task and interactive information search behavior. *Journal of the American Society for Information Science and Technology*, 61(9), 1771– 1789, 2010.

Zhen Liao, Yang Song, Li-wei He, and Yalou Huang. Evaluating the effectiveness of search task trails. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, pages 489–498, New York, NY, USA, 2012. ACM.

Jingjing Liu and Nicholas J. Belkin. Searching vs. writing: Factors affecting information use task performance. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–10, 2012.

Jingjing Liu, Chang Liu, Michael Cole, Nicholas J. Belkin, and Xiangmin Zhang. Exploring and predicting search task difficulty. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12. ACM, New York, NY, USA, 1313-1322. 2012.

Chang Liu, Michael Cole, Eun Baik, and Nicholas J. Belkin. Rutgers at the TREC 2012 Session Track. In *Proceedings of 21st Text Retrieval Conference*, TREC '12. 2012.

Marianne Lykke, Birger Larsen, Haakon Lund, and Peter Ingwersen. Developing a test collection for the evaluation of integrated search. In *Proceedings of the 32nd European Conference on Advances in Information Retrieval*, ECIR'2010, pages 627–630, Berlin, Heidelberg, 2010. Springer-Verlag.

Stefano Mizzaro. Relevance: the whole history. Journal of the American Society for Information Science, 48(9), 810–832, 1997.

Dan Morris, Meredith Ringel Morris, and Gina Venolia. Searchbar: a search-centric web history for task resumption and information refinding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1207–1216, New York, NY, USA, 2008. ACM.

James Pitkow, Hinrich Schütze, Todd Cass, Rob Cooley, Don Turnbull, Andy Edmonds, Eytan Adar, and Thomas Breuel. Personalized search. *Communications of the ACM*, 45(9), 50–55, September 2002.

Mark E. Rorvig. The simple scalability of documents. *Journal of the American Society for Information Science*, 41(8), 590–598, 1990.

Ian Ruthven. Interactive information retrieval. ARIST, 42(1): 43-91 (2008). http://dx.doi.org/10.1002/aris.2008.1440420109

Adish Singla, Ryen White, and Jeff Huang. Studying trail finding algorithms for enhanced web search. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '10, pages 443–450, New York, NY, USA, 2010. ACM.

Alan F. Smeaton. Techniques used and open challenges to the analysis, indexing and retrieval of digital video. *Information Systems*, 32(4), 545– 559, 2007.

Mark D. Smucker and Charles L.A. Clarke. Time-based calibration of effectiveness measures. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 95–104, New York, NY, USA, 2012. ACM.

Amanda Spink. Multitasking information behavior and information task switching: An exploratory study. *Journal of Documentation*, 60(4), 336–351. 2004. Emerald.

Bin Tan, Xuehua Shen, and ChengXiang Zhai. Mining long-term search history to improve search accuracy. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '06, pages 718–723, New York, NY, USA, 2006. ACM.

Robert S. Taylor. Information use environments. In B. Dervin (Ed.) *Progress in Communication Sciences*, 10, 217–225,1991

Nava Tintarev and Judith Masthoff. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction*, 22, 399–439, 2012.

Elaine G. Toms, Luanne Freund, Richard Kopak, and Joan C. Bartlett. The effect of task domain on search. In *Proceedings of the 2003 Conference of the Centre for Advanced Studies on Collaborative research*, CASCON '03. IBM Press, 2003.

Elaine G. Toms, Robert Villa, and Lori McCay-Peet. How is a search system used in work task completion? *Journal of Information Science*, 39(1), 15–25, 2013.

Randall H. Trigg. Guided tours and tabletops: tools for communicating in a hypertext environment. *ACM Transactions of Information Systems*, 6(4), 398–414, 1988.

Pertti Vakkari and Saila Huuskonen. Search effort degrades search output but improves task outcome. *Journal of the American Society for Information Science and Technology*, 63(4), 657–670, 2012.

Annika Waern. User involvement in automatic filtering: An experimental study. *User Modeling and User-Adapted Interaction*, 14(2-3), 201–237, 2004.

Yu Wang, Xiao Huang, and Ryen W. White. Characterizing and supporting cross-device search tasks. In *Proceedings of the sixth ACM International Conference on Web Search and Data Mining*, WSDM '13, pages 707–716, New York, NY, USA, 2013. ACM.

Stina Westman. Image Users' Needs and Searching Behaviour, pages 63–83. John Wiley & Sons, Ltd, 2009.

Alan Wexelblat and Pattie Maes. Footprints: history-rich tools for information foraging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 270–277, New York, NY, USA, 1999. ACM.

Ryen W. White, Wei Chu, Ahmed Hassan, Xiaodong He, Yang Song, and Hongning Wang. Enhancing personalized search by mining and modeling task behavior. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, New York, NY, USA, To Appear. ACM.

Ryen W. White, Susan T. Dumais, and Jaime Teevan. Characterizing the influence of domain expertise on web search behavior. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, WSDM '09, pages 132–141, New York, NY, USA, 2009. ACM.

Ryen W. White, Ian Ruthven, Joemon M. Jose, and C. J. Van Rijsbergen. Evaluating implicit feedback models using searcher simulations. *ACM Transactions of Information Systems*, 23(3), 325–361, 2005.

Wildemuth, B.M., & Freund, L. (2009). Search tasks and their role in studies of search behaviors. Paper presented at HCIR 2009: Bridging Human-Computer Interaction and Information Retrieval, Washington, DC, October 23, 2009.

Barbara M. Wildemuth and Luanne Freund. Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval*, HCIR '12, pages 1–10, New York, NY, USA, 2012. ACM.

Allison Woodruff, Andrew Faulring, Ruth Rosenholtz, Julie Morrsion, and Peter Pirolli. Using thumbnails to search the web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 198–205, New York, NY, USA, 2001. ACM.

Kaidi Zhao, Bing Liu, Thomas M. Tirpak, and Weimin Xiao. Opportunity map: a visualization framework for fast identification of actionable knowledge. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, CIKM '05, pages 60–67, New York, NY, USA, 2005. ACM.