NSF Workshop on Task-based Information Search Systems*

Diane Kelly, Jaime Arguello and Robert Capra (Editors)
School of Information and Library Science
University of North Carolina at Chapel Hill
Chapel Hill, NC USA 27599-3360
{*dianek,jarguell,rcapra*}*@email.unc.edu*
http://ils.unc.edu/taskbasedsearch/

Abstract

This workshop gathered leading researchers in information retrieval, human-computer interaction and information behavior to discuss challenges and opportunities associated with the development of systems and tools to support people involved with complex, multi-search and multi-session search tasks. Such task-based search systems present many challenges and the goal of this workshop was to enumerate, discuss, and document these challenges into a research agenda to guide future research. This article includes a description of the motivations and goals of the workshop and syntheses of major workshop activities including pre- workshop identification of challenges by each participant and discussion breakout groups. Major themes of the workshop included the development of domain-neutral modeling techniques to represent tasks, task properties and task-related search behaviors, interface support tools to assist with a variety of task-related information behaviors and the identification of techniques and tools to evaluate task-based search systems. The most critical need identified was the development of task models; this was viewed as essential for addressing the challenges of tools and evaluation measures.

*Authors and Participants (listed alphabetically): Eugene Agichtein (Emory University), Jae-wook Ahn (Drexel University), Nick Belkin (Rutgers University), Pia Borlund (Royal School of Library and Information Science), Katriina Byström (University of Boras), Ben Carterette (University of Delaware), Fernando Diaz (Microsoft Research), Abdigani Diriye (Carnegie Mellon University), Susan Dumais (Microsoft Research), Luanne Freund (University of British Columbia), Gene Golovchinsky (FxPAL), Jaap Kamps (University of Amsterdam), Bill Kules (Catholic University of America), Birger Larsen (Aalborg University), Christina Lioma (University of Copenhagen), Jingjing Liu (University of South Carolina), Gary Marchionini (University of North Carolina), Catherine Smith (Kent State University), Mark Smucker (University of Waterloo), Simone Stumpf (City University London), Jaime Teevan (Microsoft Research), Elaine Toms (University of Sheffield), Pertti Vakkari (University of Tampere), Arjen de Vries (CWI), Ryen White (Microsoft Research), Barbara Wildemuth (University of North Carolina)

1 Introduction

When people seek information, they typically do so in order to resolve some underlying need or task, such as finding a bus schedule to plan travel, finding a recipe to make a dish for a potluck dinner, or finding the homepage of an author of a recently read book to see what other books she has published. While contemporary search engines are good at helping people resolve these types of look-up tasks, they are not as useful in helping people engaged in more complex tasks whose resolution might require multiple search sessions and multiple search strategies. Instead, 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, internalized understanding of a problem or topic. Contemporary search environments are tailored to support a small set of basic search tasks and provide searchers with few options to search and interact with information, and little to help them synthesize and integrate information across sessions.

This report defines research challenges related to the development of task- based information search systems that were elicited during a NSF-sponsored workshop held at the University of North Carolina at Chapel Hill in March 2013. This workshop gathered leading international researchers in information retrieval, human-computer interaction and information behavior 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- based search systems and the goal of this workshop was to enumerate, discuss, and document these issues into a research agenda that could help guide work in this field. Specifically, this workshop focused 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.

2 Need for Workshop and Related Efforts

A small group of researchers have been working for some time on task-based information seeking and retrieval. One of the first reviews of this approach was written in 2003 [5], and since that time a steady and growing stream of research has been published. While this research has generated several notable task models and documented how task type and task properties can impact search behavior, there has yet to be any concentrated efforts to further develop this area and combine the findings with system design and development. There has been little work developing different indexing, retrieval and/or ranking functions, or developing different interfaces and interaction

techniques for different task types. Furthermore, there have been few efforts to generate evaluation methods and measures, and relevance models that are tailored to different tasks and consider multi-session search.

In 2006, Byström, Sundin and Limberg convened a group of researchers in Sweden to better understand task-based research [3], and in 2009, Anderson *et al.* organized a panel discussion at the annual conference of the American Society for Information Science and Technology (ASIST) exploring conceptual and methodological approaches to studying task in information science [2]. One outcome of both workshops was a call for the development of more task models and the integration of such models into system design. While this meeting and panel generated some research in these areas, the audiences and outreach was small, so the ideas have not widely spread. This meeting and panel only involved people from one community, while the problem requires researchers and perspectives from several communities, including information retrieval, human-computer interaction and information behavior.

There have been several workshops focused on task, including Larsen, Lioma and de Vries's Task-based and Aggregated Search Workshop [4] and the Second Strategic Workshop on Information Retrieval in Lorne (SWIRL) [1]. Task-based search was discussed within the context of several larger themes identified by SWIRL participants, and was also presented briefly as a mini-theme. A recent NII Shonan Meeting, which was held in Japan in October 2012, focused on whole-session evaluation of interactive information retrieval systems. While not specifically on the topic of task-based search, the ideas generated during this workshop are relevant to task-based evaluation, since the types of tasks with which we are concerned are those that take place across many sessions. A number of our invitees participated in one or many of these past events, which allowed us to build-on the findings from these previous meetings.

These recent meetings and discussions demonstrate that this was an opportune time to host a workshop focused on task-based search, and that this topic is recognized by many leading researchers in information retrieval as significant.

3 Structure of Workshop and Report

Prior to attending the workshop, participants were asked to submit short statements identifying significant challenges and important research papers, with annotations describing the significance of the papers. At the start of the workshop, three presentations by leading researchers were given on core area: task-based information seeking behavior presented by Pertti Vakkari; search engines and task presented by Susan Dumais; and interfaces and task presented by Gene Golovchinsky. Following these presentations, two additional shorter presentations were made Nick Belkin and Ben Carterette about recent efforts to address evaluation of search sessions. Belkin described a recent workshop at the Shonan Village in Japan about whole-session evaluation of interactive information retrieval, while Carterette reported on the National Institute of Standards and Technology's Text Retrieval Conference (TREC) Session Track.

During the second-half of Day 1 of the workshop, participants selected some of the most challenging issues related to task-based search and discussed and analyzed these issues in small groups of 4-6 people. On Day 2, these groups sub-divided into smaller groups to discuss more focused aspects of the broader issues discussed the prior day. Participants were asked to develop a research agenda, including the specification of a research study that might initiate the agenda.

Group leaders were elected on each day and asked to take notes, share summaries of the discussion during common plenaries, and generate reports describing the discussion.

This SIGIR Forum article contains syntheses of the homework responses, the core presentations and the breakout groups. Additional materials related to this workshop can be found at the workshop website: http://ils.unc.edu/taskbasedsearch/, including all the homework responses, the full text of the breakout group reports and a bibliography of research relevant to task-based search.

4 Core Area Presentations

To start the workshop, three speakers were invited to give overviews of three "core" areas related to task-based search. The purpose of these talks was to give attendees grounding for the workshop topic and help people understand task-based research efforts from different communities. The first core area presentation was given by Pertti Vakkari and was titled, "Task-Based Information Seeking." In this presentation, Vakkari reviewed prior research efforts to define work tasks and search tasks, described the difficulties involved in defining tasks, and outlined important dimensions of task complexity. He also discussed issues and challenges involved in evaluating task-based search. Next, Sue Dumais gave a presentation titled, "Task-Based Search: A Search Engine Perspective." Dumais motivated the importance of this work by noting that long search sessions are very common and that tasks often extend over devices and long periods of time. She showed examples of common tasks and discussed methods for automatically detecting tasks and sub-tasks. Dumais also outlined a number of ideas about how to support users' task-based searches, including query histories, richer snippets, integration of verticals, inline answers, customization, and support for richer sensemaking. Finally, Gene Golovchinsky presented on "Interactivity and Feedback." In his talk, Golovchinsky described how two types of feedback could be considered; from person to system, and from system to person. He outlined ideas about how to increase the use of relevance feedback, and discussed ways that systems can provide hints about potential actions to users such as showing which documents are new, which terms are effective, and ways to reformulate a query. Golovchinsky also presented examples of how interfaces can support users in interacting with the past (e.g., previous actions/results) or the future (e.g., reformulations, suggestions). He described how interface components could help users in query formulation through previews and other persuasive features that nudge people to take positive actions.

5 Challenges Identified by Participants

Prior to the workshop, attendees were asked to identify one or two outstanding research questions that need to be addressed in order for search systems to become more task-aware. In the following paragraphs, we highlight some of the key challenges identified. The responses from our attendees fell under four broad categories: (1) Task Modeling, (2) System Predictions, (3) Interactions and Auxiliary Tools, and (4) Evaluation.

5.1 Task Modeling

The most basic unanswered research question is: What is a task? While some tasks (such as buying a car) are clearly defined, other tasks (such as learning a new skill) are more difficult to identify because they evolve or are embedded within a larger goal hierarchy. As a first step, it seems necessary to more clearly define what a task is and to determine whether search systems should model tasks as flat structures, with clearly defined start and end points, or as hierarchical structures.

Tasks differ across a number of attributes or characteristics. Task characteristics can be a function of the task structure, the user, and/or the user's context. Example characteristics include the task complexity, salience, urgency, and difficulty. Several attendees proposed the need to develop a more comprehensive typology of tasks. Responses also highlighted the need to understand which task characteristics influence search behavior and which characteristics have little effect across users. Along the same lines, some responses called for further research on how different task types influence relevance, search strategy, search interactions, and search outcomes. A deeper understanding of these relationships would help determine which task characteristics have actual design implications for the search system (for a search engine's ranking algorithm and/or the interface).

As mentioned above, certain tasks have clearly defined start and end points and can be defined as a sequence of sub-tasks. Future research should also investigate how task stage influences search behavior. Do users employ different relevance criteria at different stages in the task? Do they desire different types of functionality from the system?

Finally, if we view tasks as being made-up of smaller sub-tasks or components, future work should also consider whether certain components generalize across tasks. To explain this point, an analogy was made to the "cut and paste" sub-task, which is widely used in desktop applications. Knowing which sub-tasks generalize across tasks could inform the development of auxiliary tools that could help users across a wide range of tasks.

5.2 System Predictions

Several responses focused on predictions a system could make about the task associated with a search. These include predicting task type and task stage. If task type and stage influence search behavior, then it seems possible to predict task type and/or stage from user interactions with the search engine.

Many tasks require multiple search sessions. In order to make predictions about the task type and/or stage, a system would need to keep a record of all the searches associated with a particular task. This would require the system to detect when a user is embarking in a new task and when a user has completed a task. Likewise, the system would need to maintain an inventory of a user's "open" tasks and match the current search session with an open task. All of these challenges would be exacerbated if we considered searches across multiple devices. For example, search sessions associated with the same task may look very different from a desktop computer vs. a mobile device.

Beyond predicting task type and stage, we could also imagine a system that tracks every user's search trail for each task and then retrieves search trails in response to a user's new task. The general idea would be to have the system return search trails that are relevant to the task.

In addition to citing different types of predictions a system could make, several responses focused on sources of evidence a system could use to inform its predictions. One response called for a tighter integration between desktop and search applications. Users often use desktop applications to complete work tasks. Desktop applications could potentially accumulate information about a user's current task, and, if the task results in information-seeking, the application could convey this information to the search engine. Another alternative would be to capitalize on the fact that users within the same information-use environment accomplish similar types of tasks. Thus, the search engine could potentially share evidence across users in same environment.

Finally, additional research is needed to determine what kinds of information about a user's task a search engine should try to predict and what kinds of information the search engine should elicit from the user directly. What are the appropriate mechanisms for eliciting task information and what are the appropriate times to elicit information?

5.3 Interactions and Auxiliary Tools

Ultimately, the goal of making predictions about a user's task is to customize the search experience. Several attendees proposed research in developing specialized interactions and auxiliary tools. Above, we mentioned the possibility that task type and/or task stage may affect how users judge the relevance or utility of search results. One avenue to explore would to develop and integrate ranking functions that focus on the document attributes that are most important for a particular task type or task stage. Another alternative would be to dynamically adjust the presentation of results to highlight the most important document attributes.

Users use a wide range of tools to accomplish their tasks. Future work should also consider developing auxiliary tools that help users integrate and make better use of the information found during their searches. As an example, for tasks that require comparing between different alternatives, the search engine could surface a spreadsheet application.

5.4 Evaluation

Evaluation is a critical component of IR. It is necessary for both tuning system parameters and for comparing between alternative solutions and interfaces. In terms of evaluation, the main research question is: How to do we evaluate systems in a more "realistic" way? Ideally, the evaluation should consider the search engine's ability to help the user accomplish a task from start to finish. Thus, we require evaluation methods and metrics that operate across multiple queries and search sessions.

Reponses from our participants made reference to two common methods for evaluating search systems: user studies and test-collection-based evaluation. Within the context of task-based search, user studies typically use simulated tasks in order to study search behavior and/or system performance for different task types. One of the main challenges, however, is that different studies use different simulated tasks. Future research should focus on developing simulated tasks that can be shared across research groups. The simulated tasks should have clearly defined characteristics and should be empirically validated. That is, they should be tested to ensure that they elicit similar behavior across similar users. A re-usable set of simulated tasks would have two important benefits: they would help make user studies reproducible and they would allow the simulated

tasks to be validated across different populations.

Finally, a number of responses expressed the need to develop test- collection based evaluation methods for task-based search. Test-collection evaluation has many benefits: it allows us to measure small improvements in performance and to reproduce results. The main challenge is that it requires modeling a user's sustained interaction with the search engine. To this end, one avenue to explore is how to simulate users. This brings us back to one of our initial points. Further studies are needed to understand how different task characteristics affect search behavior. Results from these studies would allow us to conduct more realistic user simulations.

6 Breakout Group Discussions

During the workshop, breakout groups were formed to discuss sub-topics. Each breakout group prepared a summary report based on their discussions. Group reports are clustered according to the major topic of focus: (1) modeling tasks and behaviors; (2) creation of tools and support for task- based search; and (3) evaluation. In addition, one small group focused on identifying ways to move the general research agenda forward.

6.1 Modeling Tasks And Behaviors

The first topic that guided breakout groups was modeling tasks and behaviors. This topic focused on the identification, elicitation, modeling and tracking of tasks, processes and states, including the identification of frameworks for conceptualizing task and relevance models. Group members included Katriina Byström, Luanne Freund, Jingjing Liu, Gary Marchionini, Pertti Vakkari and Barbara Wildemuth. This group identified as their main focus the question of how to represent and model a broader conception of task-based searching that extends beyond discrete, transactional searches, with a specific view on the transition points from one kind of activity to another, and from one goal or task to another. The group noted that most current models of search tend to be low level and overly simplistic, and only offer evidence of transition probabilities within a single search session (e.g., between querying and viewing results, not between different tasks or systems). The group thus decided to focus on developing a framework that connects models of information seeking tasks and information search tasks through the transitions between these tasks, including the probabilities of transitions and transition triggers.

In guiding the development of their framework, the group considered existing models and frameworks and established three key steps to guide their work: (1) identification of model elements; (2) identification of model structure; and (3) identification of task-based factors that are likely to influence the model, such as task type, task stage, and prior knowledge and expertise of the searcher. They proposed that the main benefit of the model would be to provide support for searchers' moves and decisions (e.g., by recommending specific tools). The group went on to identify challenges associated with developing and empirically testing such a model, including complications related to domain-specific information behavior and limitations associated with data collection tools. The group divided on the following day and proposed two research studies. The first was a cross-cultural, ethnographic study of information seeking behavior situated in the context of team-based patient care in the medical domain. The second sub-group proposed a series

of studies, also in the health care domain, which focused on individuals and the health acceptance model.

In addition to these groups, two other smaller groups, which were initially part of the larger tools group, proposed research agendas that addressed the first topic area of the workshop. The first group (Fernando Diaz, Catherine Smith, Simone Stumpf and Elaine Toms) focused on the identification of task primitives. The group motivated their proposed research by observing that task models are needed to support task-based search, but little is known about the fundamental characteristics and dependencies between task activities and searching, which they term task- dependencies, and the extent to which these dependencies vary across domain. To investigate task dependencies across domains, the group suggested a standardized framework for decomposing task structure, which they term task-primitives, which would facilitate the discovery and documentation of universal task dependencies. In order to arrive at these task-primitives, the group proposed to record, decompose and analyze in situ task activities. The research agenda proposed by another group (Jae- wook Ahn, Gene Golovchinsky and Birger Larsen) sought to map and understand which tools/components/widgets are most useful for which moves/activities/actions. As a starting point, the group proposed to identify general abstractions and patterns that underlie information seeking behavior which can be used to suggest tools and compare behaviors across systems.

6.2 Tools and Support

Three breakout groups focused on tools to support task-based search. The first breakout group included Fernando Diaz, Sue Dumais, Jaap Kamps, Cathy Smith, Simone Stumpf, Elaine Toms, and Arjen de Vries and focused on the topic of "tools to support workflow." They formed two sub-groups to address the topic–one that considered "bringing task into search," and another that discussed "bringing search into task." The group focused on support for complex tasks that require synthesizing data from multiple sources across multiple search sessions. They also focused on tasks that have a specific output goal, commonly in the form of an aggregated set of information (family tree, written report). Progress toward such outputs can be measured and there was consensus in the group that such tasks are not well-supported by current search systems, tools, or apps. In considering how to embed search into the work task, the group advocates for considering a broad view of information access methods and considering how such information access is embedded across all aspects of the work task. The group also considered how the work environment can impact search. Here, they suggest exploiting the structure and constraints of the task output to help support both the task-specific product needs and the information access process. Extracting structure and sub-tasks out of the work environment was identified as a major challenge. The sub-structure and dependencies among sub-tasks are important to account for, but users must also have flexibility to fluidly move among or skip components, and to backtrack. Collaborative work adds another level of complexity. The group felt that today's web/cloud-based computing is blurring the boundaries between search and work tasks and that the time is right to integrate search into work and daily life tasks.

The second breakout that considered "Tools and Support" included Eugene Agichtein, Jaime Arguello, Christina Lioma, and Ryen White. This breakout group considered how smartphones could recognize the task that a user was trying to accomplish and make recommendations about

apps that could help. To do this, the smartphone would use both implicit behavioral information and other contextual signals (such as location and time of day). The group compared task-based app recommendation to mobile search, and considered several example tasks with multiple steps such as preparing dinner for guests, or planning a day of vacation activities. As with many of the tasks considered for task-based search, these tasks required multiple steps, synthesizing information across sources, and the use of different applications. The group outlined a system architecture with a state-space controller and predictor model that would use features from the user's context and recommendation training data based on explicit and implicit feedback. Several challenges unique to mobile app recommendation were identified. First, at what stage in the task-completion process should an app be recommended? Early in the process users might not recognize the relevance of the recommendation, but delaying too long could be too late to be useful. A second challenge concerned how the system could assist the user by transferring the task-state among the apps being used to complete the task. To evaluate the system, the group proposed research questions that have broader applicability to task-based search systems. First, at what stage are users most likely to accept a recommendation? And second, does knowledge of the sequence of steps influence likeliness to accept a recommendation? They hypothesized that users would be more likely to accept recommendations at the early stages of task completion, and in situations where the user is aware of the task complexity a priori.

The third "Tools and Support" breakout group included Abdigani Diriye, Rob Capra, and Jaime Teevan. This group focused on ways to produce and present session-level and task-based summaries of search results. They outlined ways that search results could be summarized by processing the result set and providing overviews, surfacing common themes and topics, and noting documents that are similar or different from previous searches. Moving up a level, they posited that session-level summaries could help searchers to gain a deeper sense of what content was encountered and help them understand similarities and differences across the information found in the session. At the task-level, they hypothesized that summarization could help users complete complex tasks more quickly, they could simplify and encourage task resumption, and they could help accelerate knowledge acquisition. To generate task-level summaries, they discussed how queries, documents viewed, and browser-level actions could be used as input signals. The group identified several research questions that are important to the development of effective summaries for result sets, sessions, and tasks: 1) What features and attributes make a good summary?, 2) How useful are summaries across different tasks?, and 3) What techniques can be used to construct effective summaries? To evaluate summaries, they propose that measures should focus on quality, utility and usefulness.

6.3 Evaluation

The final topic that guided the breakout group discussions was the development of methods and measures for studying user behavior and evaluating task-based search systems. Participants who investigated this topic were: Nick Belkin, Pia Borlund, Ben Carterette, Diane Kelly, Bill Kules and Mark Smucker. The group focused on two main issues: (1) the limitations of the traditional IR evaluation framework when applied to task- based searching and (2) the establishment of a framework to guide reporting practices to better facilitate cross-study comparisons and the sharing of research infrastructure such as search tasks and questionnaires.

The group began by discussing the standard construal of information retrieval as a tool to help people find documents, which has led to evaluation measures focused on the (topical) relevance of documents, returned in response to a specific query, at a specific point in time. Documents are judged as relevant or not relevant, and a user's interaction with a search system is reduced to being the rate and amount of relevant documents consumed. In task-based search situations, such a construal is inadequate because searching might take place over an extended period of time, the user might issue a number of queries and sometimes the same query might be entered more than once. Furthermore, as person learns through the IR interaction what constitutes relevant information is likely to change. It was proposed that the concept of relevance be replaced by usefulness to indicate that an information object might help a person resolve their underlying tasks even if it is not relevant in a strict sense and also to indicate that human judgments during information seeking are likely to be dynamic. In addition to evaluating documents, it was proposed that the information seeking and retrieval process at the stage-level could be the focus of evaluation, as well as the end result (as opposed to the lists of documents produced by the system along with way). It was also proposed that the scope of the evaluation might be considered since it can range from the entire session, which might span multiple episodes across a number of days, to the usefulness of a particular system feature meant to support a specific aspect of the interaction at a specific point in time.

The group also discussed difficulties associated with creating measures that could be used within and across studies of users and tasks. Given a specific type of task, it is highly desirable to have a recommended set of measures that have been validated and calibrated so that they can be used for cross-study comparisons. Measures are likely to vary according to task type, but as of yet, there is no mapping of measures to tasks, and more pressing, there are few measures that are suitable for complex information seeking tasks that take place across multiple sessions. Thus, new measures are needed at both the micro- and macro-levels.

One specific approach taken by this group was to focus on stage-based approaches to information search and consider evaluation measures at this level of analysis. Questions asked included whether there are common stages that users experience while working on tasks; specific intentions of users at different stages; and how users view and describe the usefulness of their experiences at different stages, as well as overall. One important thing that emerged from this discussion was that more needs to be discovered about how people experience and view task-based search. This suggests more exploratory types of approaches to research as starting points for new evaluation measures and approaches. The group suggested a naturalistic, longitudinal study of people using instrumented laptops that would not only log their interactions, but also periodically elicit information from them about their search tasks, goals and experiences. The group further proposed periodic meetings with these participants to review their search histories and interactions in order to better understand successful and unsuccessful moments, the natural history of their search tasks and the types of measures that might be useful to evaluate different types of tasks and different stages of information seeking.

The group further elaborated on the need for a meta-framework for task- based information search studies in order to increase cross-study comparison and interoperability. It was noted that the current variety of research and reporting practices make integration across studies difficult and prevents long-term historical analysis, including meta-analysis, of studies. It further limits replication and reuse of instruments, tasks and measures. The proposed meta-framework suggests

the following aspects of each study be clearly reported: tasks (and methods of task construction), study design, measures, and methods of analysis. The meta-framework would also provide guidelines about reporting practices (e.g., measures of effect size should be reported) and a matrix of measures/tasks for best practice.

7 Conclusions

Users engage in information-seeking in order to accomplish a higher-level task. The grand vision behind the NSF-sponsored workshop on Task-based Information Search Systems is that search systems should be more task- aware. Search systems should be designed and evaluated based on their ability to assist users in accomplishing their higher-level task. Workshop attendees consisted of leading international researchers from information retrieval, human-computer interaction and information behavior. The consensus that emerged from the Workshop was that making search systems more task-aware requires work in different directions. From the user side, it requires further understanding how task characteristics and task stage influence search behavior. From the system side, it requires modeling and tracking a user's task over multiple queries, search sessions, and devices, and designing interactions that guide the user towards task completion. It requires developing evaluation methodologies that more directly measure a system's ability to help users complete the task at hand. Finally, to facilitate research from all perspectives, it requires techniques for modeling tasks, task properties and task-specific search behaviors.

The Workshop provided a stimulating environment for researchers from different backgrounds to share their views about the outstanding research questions in the area of task-based search. Our hope is that this summary inspires researchers and practitioners to work towards building systems that go beyond the query/result-set paradigm and into the task-aware paradigm.

8 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. IIS-1301958. Any options, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The workshop organizers greatly appreciate: Maria Zemankova, NSF Program Manager; Anita Crescenzi, SILS-UNC graduate student and local workshop manager; Kathy Brennan and Wan-Ching Wu, SILS-UNC graduate students and workshop volunteers and the School of Information and Library Science at UNC.

References

[1] J. Allan, W. B. Croft, A. Moffat, and M. Sanderson. Frontiers, challenges, and opportunities for information retrieval: Report from swirl 2012 the second strategic workshop on information retrieval in lorne. *SIGIR Forum*, 46(1):2–32, 2012.

[2] T. Anderson, E. Meyers, K. Byström, L. Freund, and E. Toms. Diverse approaches to "tasks" in information science: Conceptual and methodological insights. *Proceedings of the American Society for Information Science and Technology*, 46(1):1–6, 2009.

[3] K. Bystrm, O. Sundin, and L. Limberg. Task-based research in work and educational settings. In *Proceedings of Nordic Research School in Library and Information Science (NORSLIS)*, 2006.

[4] B. Larsen, C. Lioma, and A. P. de Vries. Report on TBAS 2012: Workshop on task-based and aggregated search. *SIGIR Forum*, 46(1):71–77, 2012.

[5] P. Vakkari. Task-based information searching. *Annual Review of Information Science and Technology*, 37(1):413–464, 2003.