Personalization needs extension towards task stages in collaborative research work tasks

Sanna Kumpulainen
University of Tampere
Kanslerinrinne 1
33014 Tampere
Finland
sanna.kumpulainen@uta.fi

Hugo Huurdeman
University of Amsterdam
Turfdraagsterpad 9
1012 XT Amsterdam
The Netherlands
h.c.huurdeman@uva.nl

Heikki Keskustalo
University of Tampere
Kanslerinrinne 1
33014 Tampere
Finland
heikki.keskustalo@uta.fi

ABSTRACT

People have various goals and intentions while interacting with information items. The intentions are derived from larger, motivating work tasks, and the intentions during the stages of task performance may vary. Further, the information activities are often conducted in collaborative work settings. We present preliminary findings on a collaborative research task process indicating that task-stages are also observable in collaborative tasks. We show that trust building and common understanding of the task emerge and discuss the implications to evaluation of information interaction. Based on our understanding, we claim that the personalization based on individuals' past actions is not necessarily enough, but needs extension towards task stages in collaborative research tasks.

CCS CONCEPTS

• Information systems~Users and interactive retrieval

KEYWORDS

Task-based information interaction, Task stages, Collaborative tasks

1 INTRODUCTION

Understanding and learning from user interactions involves a number of different aspects from understanding the task and intent, to developing user models and personalization services. Users' understanding of the overall task and information needs develop as they interact with the information available. A detailed understanding of the relation between data sources, tools, and activities in the research cycle, as well as how this interaction affects the understanding about the task goals is essential.

In the field of personalized information retrieval, most research has generally focused on adapting the result presentations by re-ranking results or revising the initial search results. Personalization works in case of well-defined fact-finding type of tasks. However, it is widely agreed that information intents are derived from a larger, motivating task [1-4] and that it includes various types of information activities [3]. Larger tasks may include information searching that is manifested as information seeking sessions which inherit their goals form the task they are part of [4]. After retrieving useful information, people start working with the relevant items [3]. We present preliminary findings about history researchers' information interactions during a collaborative research effort in a naturalistic setting. By taking a task-based approach, the study of task stages in a collaborative process is enabled. We discuss the implications for personalization towards tasks, task-stages and collaborative work. While naturalistic settings may lack control and generalizability, they provide invaluable data when studying performance during the larger tasks.

1.1 Background

Personalization attempts to ensure that content and services are tailored to individual users' personal preferences, goals and context while at the same time making the reuse of information easier. The information items that are relevant to the users' activities and interests are highlighted and brought to users. Personalization aims to answer to information needs specific to each user's interests (topical), prior experience (search histories) or location (context of action), or suggest recommendations based on other users' behavior (social recommendation). These entail building a model of users' interests by examining their past interactions with the system to identify information that might be pertinent to the users.

It is widely agreed that the information intentions and usefulness assessments are derived from a motivating task. The motivating task behind information behaviors involves *stages* [1,5]. During the stages, the searchers' understanding of the task and useful information are changing. Accordingly, searching behavior is changing. Therefore, we argue that personalization should be evaluated against the task stages.

Collaboration is required in work settings, which are too complex for individuals to deal with [6]. Hansen's [7] study on patent engineers' information seeking tasks uses a task-based approach in a collaborative real world setting. It studies the work process of patent engineers in detail, and analyzes the process as stages. We add to this by analyzing the collaborative activities as task stages and describe how the growing understanding of the task affects the task performance.

Task-based personalization features have been discussed by few researchers [8-11]. The studies conceptualize tasks at the search session level or as successive search sessions, and focus on the aspects of searching. We want to take a step further and address the larger task performance process from the viewpoint of information activities in a *collaborative effort* during an information intensive task. Taking a too narrow window into information behaviors may lead to insufficient representations of the users and their tasks. Therefore, the scope of research needs widening towards task stages in collaborative settings.

1.2 Task-based evaluation

There are few task-based approaches in IR evaluation. Belkin and others [12] proposed an evaluation frame consisting of three levels. It measures performance on a task goal level, information interaction level and information seeking strategy level. The activities are partially analytical, and may occur in varying order during the task performance. Liu and Belkin [13] proposed a task aware personalization based on dwell time. They also found that previous topical knowledge affected the assessments of document usefulness. Järvelin and others [3] propose an evaluation frame consisting of several information activities that take place during a learning task. In their model, information interaction is conceptualized as behavioral and cognitive activities related to task planning, searching information items, selecting, working with them, and, synthesizing and reporting.

Mainly in an educational context, Kuhlthau [5] and Vakkari [1] studied larger task performance processes as stages. In the models, one important aspect is the actor's *perceived task* at each stage of task performance. The focus formulation point is especially important in both models, since it changes the searchers' perceptions about the task and how to proceed. Further, it affects also the outcome of the task. Vakkari's model [1] is adapted from Kuhlthau's model and discussed the actual information searching in more detail. It is based on a longitudinal study with students preparing research proposals individually, and has three stages:

Stage 1. Pre-focus: Prior to focus formulation conceptualization of the task is fragmentary and vague. The searcher is unable to construct the task and unable to express precisely what kind of information is needed. Background information is preferred.

Stage 2. Focus formulation: At this point, searching is more focused and the searchers know what they need. Their conceptual structures are building up. During the focus formulation, searchers also need background information but the domain and topic of task is clearer. More methodological and procedural information is required to understand in which ways they can achieve their goals.

Stage 3. Post-focus: During the final stage, the conceptualization of the task is solid. Searchers are retrieving topic specific information. The means of identifying the documents and information resources is comprehensive. Their searching is directed and more close-ended. Searchers are able to assess the usefulness of the information retrieved, because the task goals are clearer.

When a complex task process starts the first aim of the searcher is to understand what the task is. In situations when the searchers' mental models lack conceptual structures for accurate representations of the task, exploratory searching takes place. According to White and Roth [14] people engaged in exploratory searches are generally unfamiliar with the domain of their goal, unsure about the ways to achieve their goals and unsure about their goals. This is similar to complex task as defined by Byström [15]. Searchers need help in *creating understanding about the task*. In terms of stage models, this means help with formulating a focus.

2 RESEARCH SETTING

We studied information interactions in a history research project aiming at supporting information interactions during a research task. This study focuses on the observations and discussions about the information activities of historians while retrieving, accessing, and analyzing collections of historical documents as primary sources. At the beginning of the research, we interviewed the historians and collected critical incident demonstrations during which the participant chooses an ongoing or recently completed task to present in more detail. After the interviews, we observed a series of collaborative meetings in a participatory setting. The meetings aimed at improving information access to the historical document collections from the Second World War. The collections consisted of personal letters (approximately 7 000 letters), a database with 90 000 entries of detailed records on individuals and over 140 000 photographs with caption texts and some metadata. During the meetings we collected users' intents (task goals). Two history researchers and one with a background in visual journalism, interested in the history of emotions and experiences, collaborated with two information scientists. Two research assistants were also involved.

The data included four interviews, five critical incident demonstrations and 12 collaborative meetings collected during five months period (September 2017 to end of January 2018). Interviews and critical incident demonstrations lasted one hour each, collaborative meetings lasted approximately two hours per meeting. During the collaborative meetings, the researchers discussed about the current state of their work and future directions. The setting was natural work context and during the meetings, the participants reflected their thoughts about their task. During the observations, we collected hand written notes. The interviews were audio recorded and transcribed into text files and analyzed with a qualitative research tool. During the data collection, we focused on behavioral and cognitive activities related to information interactions, in particular task planning, searching information items, selecting, working with them and synthesizing and reporting, as defined in [3]. The selected framework helped the investigators to handle the complex situations during observations.

Naturalistic field studies are rare in information searching research, because they are tedious and time consuming. However, they provide invaluable data when studying performance during the larger tasks. This kind of approach may lack standardized procedures; have only small number of participants; and data are derived in varying situations lacking control. It is a tradeoff between control and reality, but we believe that both – controlled experimental settings and research "in the wild" – are important to examine.

3 RESULTS

The collaborative meeting materials were analyzed with thematic analysis [16]. The analyses focused on what happened during each meeting and the participants' understandings about the task concepts, structure and procedures (i.e., means how to reach the goals) acted as basis for the task stage division. Table 1. shows an overview of the themes.

Table 1. Task stages and emerging themes

Stage	Meeting	Task relevant themes	Collaborative themes
Pre-Focus	1-3	Domain knowledge Source descriptions State of affairs Task goals & overall search goals	
	4-5	Procedural info Possible handles Dimensions of information sources Detailed description of sources	Trust building
Focus formulation	6-8	Focus formulation Procedures and workflow Task goals and outcomes	Timelines Milestones Shared concepts
	9 -10	Procedural evaluation Shared conceptualizations in use Conceptual clarifications Detailed descriptions of the work task Timelines Task goals	Shared focus Reflecting the process
Post-focus	11	Detailed descriptions of possible information tools Detailed agreed procedures	Focused understanding of the goals Focused understanding of the procedure
	Meeting 12	Comparing characteristics of search tools Methods and procedures	Reflections on the existing tools Division of labor

In the pre-focus stage, the participants were discussing the historical materials in detail. In our case, they were domain experts and knew what each information collection included. They aimed at an integrated system, in which all the items could be explored easily. However, during the pre-focus stage, the themes of the meetings turned from the materials into the conceptual aspect. There was a strong aim to create a common understanding what the task goals were. This was not specifically easy, since the participants represented distinct academic cultures. The procedural discussions were vague and some possible methods were discussed. The discussions were careful and tentative, since not all the participants were familiar with each other. Trust building was important, and when it started to build the discussions grew deeper.

In the focus formulation stage, the participants started increasingly to engage into the process, and there were clearly expressed notions about the increased understanding of the goals and procedures. The clarified situation was boosting the collaboration, and exact timelines and milestones were set. The concepts were shared, so that the participants from different disciplinary backgrounds could use specific terms learned from others. The awareness of the state of the task had increased, and clear statements about it were made. There were some reflections and assessments about how the process had advanced.

In the post-focus stage, the level of details increased. The process took a distinctly new turn as goal-directedness increased. Participants were able to compare the existing search tools and even discuss the qualities of the possible future search tools in detail.

4 DISCUSSION

In this paper, we applied Vakkari's model [1], previously only used in learning tasks by single users, within a context of collaborative research work tasks. Our findings indicate that task stages exist and that a similar focus formulation process as in [1] can be found in collaborative tasks. During the pre-focus stage, the vague understanding of the topic, structure and goals of the task and lack of procedural knowledge was obvious. During this stage, support for conceptual structuring is needed. Typical search box tools are often used in pre-focus stages, but if one does not know any suitable search keys and cannot identify useful information, the results may turn out irrelevant. If the perceived task is muddled at the beginning of the process – in terms of unclear knowledge about the domain, procedures or required input – it seems more complex than in the post-focus stages, in which the knowledge about the topic has been created. At the pre-focus stage, people with scarce domain knowledge need support to expand and differentiate their conceptual model of the topic, cf. [17]. The task process is moving from a vague conception towards a clearer understanding of the task, with a coherent conceptual structure. The searchers typically engage in exploratory searching before coming up with any specific direction. When creating a direction, the searchers may benefit from trails across the information environment [14,18] and in case of historical documents, following a pre-built historical path across useful information items [19]. Faceted search features, as well as suggestions for additional query words and formulations can provide handles for exploratory searching at this stage [11].

During the post-focus stage, the most appealing methods seem to be re-retrieval aids such as previous queries and search histories [1,11]. In post focus stages, the perceived task complexity is lower, due to learning during the task performance process. After focus formulation, the actors are able to assess the relevance of the information items and are able to reflect their choices against the task goals. For the tasks of lower complexity, such methods as explicit relevance feedback are suitable, as suggested by [20]. The searchers require more control over their work in this phase, due to the more expert like conceptual map of their task.

One information activity that was prevalent during the task process was annotating the historical materials cf. [21]. Ontologies or dictionaries that help to make explicit the semantic content of the

historical documents could be useful here. Boot [22] suggested that annotations function as a basis of recommendations since they may show important information about objects and locations. Further, annotation patterns help to create profiles of persons and research interests.

This study focused on the historical domain, so the findings are depending on some domain-specific characteristics. User needs and intents differ according to domain specific features. However, domain expertise is an important aspect of information behavior. Vakkari's [1] and Kuhlthau's [5] models have mostly been constructed in the context of searchers with low domain knowledge, making it unsure if these could be applied to our context. As it turned out, our participants were experts in their own domain, but they simultaneously had low domain knowledge in the domains of the collaborating experts. For example, history researchers were not familiar with the principles of the NLP systems, and the information scientists initially did not understand the intents of the historians. Thus, there were novice-type focus formulation needs for each expert role.

This research has some limitations, and the analyses are preliminary. We did not study the information activities in detail during the task process, only the aspect how the task stages evolved during a collaborative task performance. Moreover, we did not aim at generalized findings, but were seeking for human aspects of information interactions in a collaborative research task. However, the research is still ongoing, and we will later evaluate the future tools within a task-based setting.

There is a body of research on several aspects of personalization in information retrieval, e.g., query reformulation, result re-ranking, user modeling and social aspects. However, when deciding which aspects of information interactions are important, typical technologies have two weaknesses. First, lack of task awareness: needs evolve during the task performance process, which should to be taken into account, specifically in terms of task stages. Second, lack of collaborative viewpoint: during collaborative tasks, trust creation among the participants and a shared understanding of the task are important. Thus, support mechanisms, which can contribute to trust and task understanding, are essential. To build trust, it is important to foster and capture "deep dialogues" between group members [23]. For shared understanding, personalization should be extended towards information interactions of the collaborative group, instead of the individual.

7 CONCLUSIONS

Personalization aims at providing the most useful information items to the users. Useful information is retrieved to serve the purposes of larger motivating tasks that progress in stages. We conducted a small-scale study on a few participants in a natural work setting. Regardless of the uncontrolled research setting, we found that collaborative tasks evolve similarly to individual tasks.

We performed a collaborative task with focus on supporting the information interactions of historians. We observed that the needs of the participants changed significantly, as the task stage proceeded (pre-focus to focus to post-focus). We propose that the

task stages should be studied further in collaborative settings for two reasons. First, understanding the needs at different stages, and how the needs evolve (e.g., to build trust, to share concepts) facilitate including collaboration in system design. Secondly, recognizing the shifts between the stages is required in task-aware personalized support, since the needs are changing at those points. Hence, we believe that personalization would benefit of extension toward task stages in collaborative research work tasks.

ACKNOWLEDGMENTS

REFERENCES

[1]   Vakkari P. A theory of the task-based information retrieval process: A summary and generalisation of a longitudinal study. J Doc 2001; 57: 44-60.

[2]   Li Y, Belkin NJ. A faceted approach to conceptualizing tasks in information seeking. Inform Process Manag 2008; 44: 1822-1837.

[3]   Järvelin K, Vakkari P, Arvola P, Baskaya F, Järvelin A, Kekäläinen J, Keskustalo H, Kumpulainen S, Saastamoinen M, and Savolainen R. Task-based information interaction evaluation: The viewpoint of program theory. ACM Transactions on Information Systems (TOIS) 2015; 33: 3.

[4]   Byström K, Hansen P. Conceptual framework for tasks in information studies. J Am Soc Inf Sci Tec 2005; 56: 1050-1061.

[5]   Kuhlthau CC. Inside the search process: Information seeking from the user's perspective. J Am Soc Inform Sci 1991; 42: 361.

[6]   Shah C. Collaborative information seeking. J Am Soc Inf Sci Tec 2014; 65: 215-236.

[7]   Hansen P, Järvelin K. Collaborative information retrieval in an information-intensive domain. Inform Process Manag 2005; 41: 1101-1119.

[8]   Belkin NJ, Hienert D, Mayr-Schlegel P, and Shah C. Data requirements for evaluation of personalization of information Retrieval–A position paper. Proceedings of Working Notes of the CLEF 2017; .

[9]   White RW, Jose JM, and Ruthven I. A task-oriented study on the influencing effects of query-biased summarisation in web searching. Inform Process Manag 2003; 39: 707-733.

[10]   Kotov A, Bennett PN, White RW, Dumais ST, and Teevan J. Modeling and analysis of cross-session search tasks. Proceedings SIGIR 2011; 5-14.

[11]   Huurdeman HC, Wilson ML, and Kamps J. Active and passive utility of search interface features in different information seeking task stages. Proceedings CHIIR 2016; 3-12.

[12]   Belkin NJ, Cole M, and Liu J. A model for evaluation of interactive information retrieval. Proceedings of the SIGIR 2009 Workshop on the Future of IR Evaluation 2009; 7-8.

[13]   Liu J, Belkin NJ. Personalizing information retrieval for multi-session tasks: Examining the roles of task stage, task type, and topic knowledge on the interpretation of dwell time as an indicator of document usefulness. J Am Soc Inf Sci Tec 2015; 66: 58-81.

[14]   White RW, Roth RA. Exploratory search: Beyond the query-response paradigm. Synthesis lectures on information concepts, retrieval, and services 2009; 1: 1-98.

[15]   Byström K. Task complexity, information types and information sources : Examination of relationships. Tampere University Press 1999.

[16]   Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006; 3: 77-101.

[17]   Vakkari P. Exploratory searching as conceptual exploration. Proceedings of HCIR 2010; 24-27.

[18]   Kumpulainen S. Trails across the heterogeneous information environment: Manual integration patterns of search systems in molecular medicine. J Doc 2014; 70: 856-877.

[19]   Clough P, Ford N, and Stevenson M. Personalizing access to cultural heritage collections using pathways. Proceedings PATCH 2011; 12-19.

[20]   White RW, Ruthven I, and Jose JM. A study of factors affecting the utility of implicit relevance feedback. Proceedings SIGIR 2005; 35-42.

[21]   Melgar L, Koolen M, Huurdeman H, and Blom J. A process model of scholarly media annotation. Proceedings CHIIR 2017; 305-308.

[22]   Boot P. Mesotext: digitised emblems, modelled annotations and humanities scholarship. Amsterdam University Press. 2009.

[23]   Holton, J. A. (2001). Building trust and collaboration in a virtual team. *Team Performance Management: An International Journal*, 7(3/4), 36–47.