Salton Award Keynote: Information Interaction in Context

Kalervo Järvelin
University of Tampere
Finland
kalervo.jarvelin@uta.fi

Abstract
This article is a lightly edited text of the author's Salton Award Keynote: Information Interaction in Context, presented at the 41st SIGIR conference in Ann Arbor, July 9th, 2019. It first gives some personal background and then discusses some important areas of information seeking and IR in the author's research work. These include task-based information behavior and interaction, natural language processing to improve document ranking in mono- and cross-language IR, and IR evaluation metrics. Finally, the article proposes a way to organize research on information interaction.

1 Introduction

Dear colleagues and friends, let me begin by saying how pleased and honored I am to receive the Gerard Salton Award from SIGIR. I only had the chance to meet Professor Salton once in person, at SIGIR'91 in Chicago, where he invited me as a visiting scholar to Cornell. I could not do that then, but had been inspired, for long, by his great pioneering work on automatic information retrieval and the SMART system. I also had conveyed his ideas to my students. He was a man of vision and authority. He could have pushed me to a more system-oriented path within IR, than the one I actually followed.

I'm also humbled by the honour of having been chosen to stand among the previous 11 recipients of this award. Their contributions to IR are of greatest value. I have greatly enjoyed their writings, many of which have deeply affected my thinking about IR. Honestly, I never believed that I would be giving this acceptance talk.

As the tradition is, based on several examples by prior Salton Award winners, I'd like to present today a personal reflection on IR – one possible version. I will begin with some personal history in Information Science and IR. Then I'd like to present some recent studies looking at information retrieval and information interaction in context. I will finish by proposing my personal view on what should be included in the study of IR.

2 Some Personal Background

I came to the field of IR through Information Science. I have spent almost my entire career at the University of Tampere, where I started as an MSc student in 1972. My initial aim was to obtain a degree in librarianship. However, my first professor, Sinikka Koskiala, soon introduced me to IR, information behaviour and Information Science through inspiring books by Wilfried Lancaster [24], Gerard Salton [32], and Manfred Kochen [21]. The former two are well-known in this community, and need no introduction. It gives me great pleasure to mention Dr. Kochen, since he worked here, at the University of Michigan, from mid-sixties to late eighties. His 'The Growth of Knowledge' was an inspiring introductiioin to Information Science. Professor Koskiala also advised me to begin studies in Computer Science. This led me to research in database management systems, which I'm not *that* known for, but have nevertheless continued until recently.

Within Information Science, the subfield of information behaviour, or user and information needs studies, was quite popular at that time. It was seen to offer a scientific foundation for the development of libraries. Indeed, in the 1960's, William Paisley [28] and others (e.g. [1, 27]) had proposed quite influential and advanced frameworks, focusing on a *person's* information behaviour. However, the vast majority of the studies were user studies, or institution-centred studies, focusing on the role of users in the life of libraries, and paying mere lip service to the more advanced models.

Later, I was exposed to ideas by Gernot Wersig (Information sociology) [40], Nick Belkin (the ASK hypothesis) [5], Brenda Dervin (sense-making theory) [8], Douglas C. Engelbart (human augmentation research) [9], and Steven Alter (decision support systems) [2], among others. These led me to view Information Science as a discipline studying people as *actors* in their work/life. This means that they are *not to be considered as users* of some information delivery institutions *per se,* nor as information seekers *per se,* but as *subjects* perceiving their own information environment and interacting with it.

Many among us, who design and develop information services and systems, hold the view that information is always beneficial; that all relevant information should be sought for; and consumed. This is a popular rationalistic view on people from systems viewpoint, IR-rationalism. People are here seen as users with *prescribed, normative, behaviour* within the frame of the system. Some management scientists and sociologists tell a different story, aiming to describe real behaviour. For example, Nathan Caplan, Professor Emeritus of Sociology here at University of Michigan, studied information interaction by decision-makers [7]. He described several interaction styles, some of which intentionally ignored relevant information not supporting decision-makers' current perspective.

Such influences gradually shifted the focus of Information Science toward the role of information services in the life of people. They also contributed to the transformation of Information Science *from an institution focus,* or systems focus, to *a person in context focus*.

3 Task Complexity and Information Behavior

Some key factors in context formation are the tasks people are aiming to perform while interacting with information. Theoretical reflections led me to think that task type and complexity may affect an actor's information behaviour. In particular, tasks of different complexity level pose different difficulties – such as insufficient or imprecise data, or lack of knowhow. They may require different types of knowledge – such as general domain knowledge, specific knowledge on the problem at hand, or problem-solving knowledge. And finally, this knowledge may be constructed from data or information available in different types of sources, which are to different degrees within reach to the person.

These ideas are quite self-evident once stated, but not that simple to study empirically. In fact, I gave up the empirical part and continued my formal research within databases. It took an innovative master's student, later Dr. Katriina Byström, supervised by Prof. Pertti Vakkari, to figure out how to gain empirical evidence on the effects of task complexity on information behaviour at a detailed process level, as opposed to questionnaire-based opinions.

Byström and myself defined work task complexity in terms of perceived a priori determinability of information inputs, processes, and outputs [6]. We distinguished five levels of complexity from Automatic information processing tasks, which may be completely described at the outset, to Genuine decision tasks, which cause embarrassment in all aspects of inputs, processes, and outputs. We found evidence on differing information needs, differing sources and channels between the task complexity classes.

Self-evident ideas are however not always accepted. The joint article manuscript "Task complexity affects information seeking and use" was rejected at one leading Information Science journal as being "woolly" but later accepted by the IP&M journal with minor knitting.

After a few years incubation the idea got quite popular in interactive IR and information seeking studies. Others have extended them theoretically Pertti Vakkari [36, 37, 39] or proposed classifications of work tasks and search tasks (e.g., Li & Belkin [26], Wildemuth & Freund & Toms [41], and Kelly & al. [18]).

For example, Li & Belkin [26], proposed a faceted classification of tasks, which differentiated *objective* task complexity from *subjective* task complexity and offered other task-related facets like interdependence, salience, urgency and difficulty.

Task analysis and classification obviously enhance one's capability of examining the relationships between work tasks and various aspects of information interaction.

4 Ranking in Focus

Following the rise of test-collection based IR research, we established an IR Lab and the research group FIRE in the beginning of the 1990's at the University of Tampere. The IR community focused then on developing ranked retrieval systems that scaled up to handling large collections. That research was without question very important and successful. The results were remarkable. Still we thought that a broader person-oriented scope would be beneficial to the field. To advance this effectively, we thought, one needs credibility earned through becoming a full member of the community. This would be achieved through carrying out successful research with similar goals – that is, test-collection based research with focus on ranking. In addition to making use of the TReC test collections, we decided to build a Finnish language test-collection – as part of Eero Sormunen's doctoral project [34]. This test-collection brought two important benefits: one in the development of robust NLP methods for IR, and the other in the come-back of graded relevance assessments in IR evaluation. The former was inspired by the morphological complexity of the Finnish language. And in fact, Finnish has been used as the litmus test for the effectiveness of many morphologically oriented IR methods. We were quite successful in research into:

- the application of NLP methods in IR across languages. We used
  - reductive methods, incl. lemmatization and stemming, for conflating word forms.
  - We also studied generative methods, including inflectional stem generation and full

inflectional form generation, producing possible word forms for an un-stemmed index

  - We worked with many languages, both European and Indian

- We also developed cross-language IR:
  - employing all main approaches: dictionary-based, machine translation, comparable corpus and combined methods
  - and finishing with approximate string matching methods for searching on words not translatable using other approaches.
  - the Pirkola method of performing dictionary-based translation, developed in Ari Pirkola's thesis, was top-scoring despite its simplicity for years around 2000 [29].

We also examined ranking in the context of ontologies and query expansion [19], and structured queries and documents [3].

5 User-oriented Metrics?

The second consequence of our test-collection approach was the revival of graded relevance assessments in test-collection-based IR evaluation. The seminal Cranfield test-collections had graded assessments, but in TReC, the assessments were binary at that time. We were inspired by the seemingly loose relevance criteria employed in the TReC test collections: a document is topically relevant if it has at least a single sentence pertaining to the topic. Since some documents contain much more, we wondered, if effectiveness evaluation by such highly relevant documents would matter, in fact, give a different picture on IR method effectiveness.

This wondering or embarrassment inspired the development of IR evaluation metrics based on graded relevance assessments – the generalized recall-precision metric and the Cumulated Gain metrics family, foremost the nDCG metric. This was joint work with Dr. Jaana Kekäläinen [14 - 16], and later also Dr. Heikki Keskustalo [20]. The former metric generalized *recall* as the share of the total relevance mass retrieved at a given rank, and *precision* as the relevant mass proportion within the documents retrieved at a given rank. Also the calculation of the *generalized MAP* was defined. Others have presented similar and extended ideas. For example Robertson, Kanoulas & Yilmaz proposed the GAP, graded average precision metric [30].

We proposed the normalized discounted cumulative gain, the nDCG, as a user-oriented metric. This was a reaction to IR evaluation often giving credit to marginal documents ranked late in search results. This offers *statistically* significant differences in evaluations based on effectiveness differences that do not matter *in the searcher's practise*, we thought. The nDCG was to be user-oriented, because it was based on *graded* relevance, allowed *weighing* of relevance levels, and *discounted* the gain incurring from documents found late on the result list. Roughly, the argument was that the metric rewards search engines that are able to retrieve highly relevant documents early. The steepness of the reward can however be adjusted to accommodate various search scenarios. There are quite a few notable proposals extending the basic metric in various ways regarding user models and multiple query sessions, for example.

Nevertheless, basically (n)(D)CG was intended, and has been used, as a *metric for single ranking evaluation* in experimental IR. Such evaluations are most typically carried out *without searchers* – except for indirectly through relevance assessments – with the primary goal of improving ranking.

Here the nDCG has served IR evaluation well.

However, when human searchers have been incorporated in evaluation studies, the results have often suggested that better ranking does not necessarily translate to better search task, or work *task, performance*. For example, in Turpin & Scholer's [35] experiment, searchers employed search engines with controlled mean average precision (MAP) of between 55% and 95%. The study results indicate that there was no significant relationship between system effectiveness measured by MAP and a simple precision-based search task. Smith & Kantor [33] found that searchers using intentionally degraded systems were as successful in searching as those using a standard baseline system. In achieving this success, they altered their behavior in compensation. Vakkari & Huuskonen [38] examined how students' search effort for an assigned learning task was associated with retrieval effectiveness, and how this was associated with the quality of learning task outcome. They found that effort in the search process degraded precision but improved task outcome. In these examples, human effort compensated bad retrieval results on the way to good task outcome.

Therefore it is not clear, which IR system or interface affordances should be developed to improve the effectiveness of human task performance. A broader context is needed because optimal solutions may be missed if the R&D work on IR systems/services is not informed by their use context.

This view was developed further in the joint book 'The Turn', with Peter Ingwersen [12]. The book aimed at integrating IR and Information Seeking research theoretically. To learn about that integration empirically, one needs to study actors in their own information environments.

Simultaneously there was taking place a transformation within the search engine industry, focusing on the behaviour of their users as the key source of signal for improved ranking and recommendation, and placement of ads. This meant extensive server-side logging of user behaviour in the single search engine context. Analysis of the logs has led to significant improvements in user-system interaction (and serious privacy problems that have surfaced recently). Still, this approach has a hazy take on people interacting with information in their work and everyday life contexts, which offer multiple information sources, services and systems.

Consequently, some of us moved from system focus to actor focus along Diane Kelly's [17] Continuum of types of IR research.

6 Life at Workplaces

In approaching the human context I was inspired by the book Laboratory Life (by Bruno Latour and Steve Woolgar [25]). It describes the adventures of Bruno Latour, a famous French sociologist and philosopher, in finding out what is going on in a Californian biological research laboratory, using ethnographic methods. The authors studied how the social world of the laboratory produced papers and other texts. One has to do ethnographical *observation* because, if one *asks* the scientists, their statements systematically conceal the nature of the activity, which typically produces their research reports. By generalization, asking *doers in other domains* about their work, produces a rationalized account not matching reality. We inferred that, if one wants to learn about real information interaction, one must go observe how people perform their activities. Consequently, we went to work places in several doctoral projects:

- in the patent domain (Preben Hansen: Task-based Information Seeking and Retrieval in the Patent Domain) [10],

- in molecular medicine (Sanna Kumpulainen: Task-based information access in molecular medicine: task performance, barriers, and searching within a heterogeneous information environment) [22],
- in social work (Saila Huuskonen: Recording and Use of Information in a Client Information System in Child Protection Work) [11], and
- in public and business administration (Miamaria Saastamoinen: Information Searching in Authentic Work Tasks: A field study on the effects of task type and complexity) [31].

We mapped information environments and the actors' behaviour in them. We used a mixture of data collection methods including interviewing people, and observing, photographing and client-side logging their information interaction.

We obtained a rich picture of, for example:

- how information seeking and retrieval is coupled with larger task processes;
- how people collaborate on both IS and IR;
- integrated use of multiple information sources and systems which was more intense in complex tasks;
- how barriers made system limitations visible and caused workarounds;
- what kind of queries and sessions were utilized for solving information problems in various contexts; and
- how much time the actors spent in each information system type in tasks of various types.

I will now discuss two sample findings. The first one is from Sanna Kumpulainen's thesis [22]. Figure 1 represents some real life information interaction for a single work task in research in molecular medicine. Multiple channels and sources are harnessed for information interaction in this task.

This figure presents the step-by-step performance of a complex work task consisting of 23 steps over two hours. The top line indicates the eight queries employed. The bottom line approximates task duration in chunks of 15 minutes. Note that the chunks are, graphically, of varying lengths because some time-slots involve many events while others may involve just examining a single found resource. The shaded column is a one-hour break.

The middle area represents the structure of the information environment of the researchers and is divided into broad horizontal channels: from top down they are search engines, web sites, literature databases, biological databases, the PC tools, and other. Each track within a channel represents the same unique resource throughout the session while different tracks indicate different resources of the same type. The solid arrows represent the workflow, the dashed arrows represent data flows, and the dotted arrows represent transitions by links.

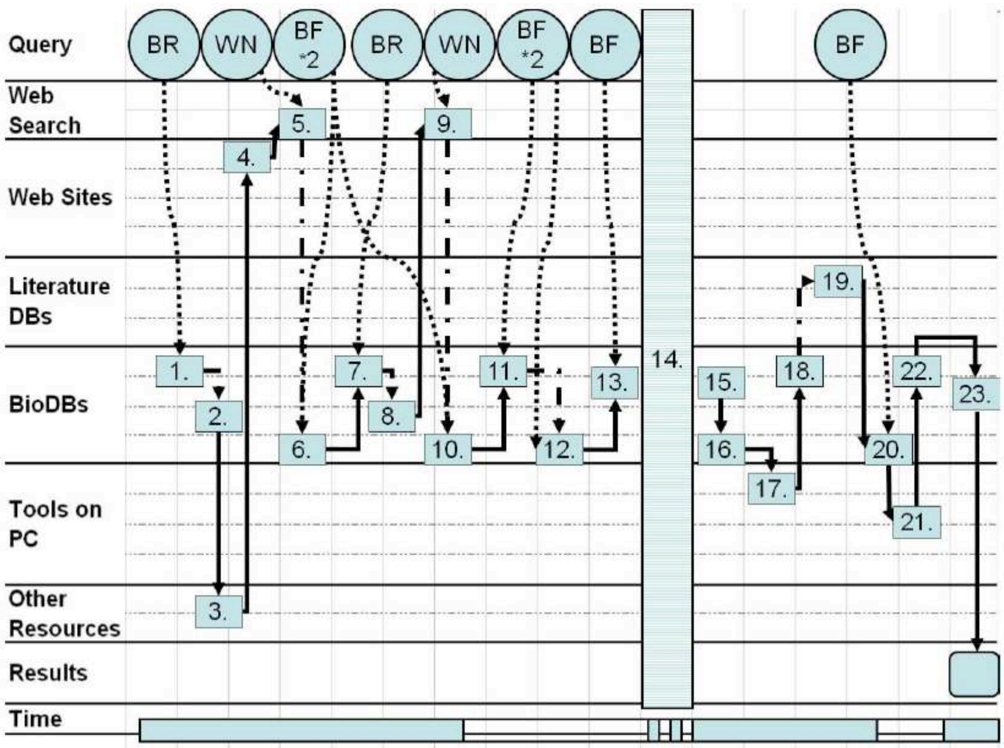

Figure 1. Sample work task process in molecular medicine [22, p. 45]. QUERY LEGEND: BF – bio factual; BR – bio resource; WN – Web navigational

While the majority of action takes place between three bio resources in this task, this session spreads across six channels. Overall, we note that the task progresses through several tools in several channels and may return to a given tool after some visits to other tools. The task performer harnesses and integrates the tools for the task as is most natural and convenient for her. This depends on her personal – perceived – information environment. If we focus on the eight queries alone, especially through server-side logging showing only a subset, we obtain a very limited view on the task performer's information interaction.

I'm convinced that we can and will find similar features in many other task domains, not just in molecular medicine and not just in research contexts.

As another example, consider some of Miamaria Saastamoinen's findings [31]. Among others, she looked at how work task complexity is connected to IR features. She found out that:

- Up to 42 % of work tasks did not include any searching. There was other information interaction, though, with the necessary information available. The remaining tasks had on average 2.5 search tasks per work task. Complex tasks included more search tasks than simple ones.
- In addition, most search tasks were really simple
    - typically containing one query of length one keyword of type "proper name"
    - again, complex tasks had more queries, longer queries, and more varied queries
- Some 50% to 60 % of task duration was spent on sources on the PC or communication media.

Recall that typical IIR studies use simulated work tasks or specified search tasks that require much searching, long sessions, in a single system / source. According to this study the vast majority of real life searching is very simple – while sufficiently effective for task performance. Yet it uses multiple systems and a range different sources.

Moreover, in the flow of authentic work, people seldom seemed to consciously recognize separate 'search tasks' and reflect on their topical content. Search tasks in IR experiments may be an academic construct seen by IR academics and not *that* consciously seen by the people performing tasks "out there in the wild".

7 Task-based Information Interaction and Interface Capabilities

Partly concurrently with these studies, we proposed an evaluation framework for the study of Task-Based Information Interaction (TBII) [13].

Evaluation, in general, is the systematic determination of merit of some object using criteria based on some goals. It aims at analyzing to what extent the object of evaluation attains these goals. To perform evaluation one needs to define (a) the object of evaluation, (b) its goal(s), and (c) indicators of goal attainment. When defining the goals, it is important to understand the difference between the immediate outputs of the system and its longer-term final outcomes. It is also essential to model the mechanisms connecting the object system features and use to the goals. Otherwise we do not understand the system and its development becomes like kayaking in the night.

We define task-based information interaction here as behavioral and cognitive activities related to task planning and monitoring, searching information items, selecting between them, working with them, and synthesizing and reporting the outcome. In carrying out such activities, people use a range of tools, including but not limited to, IR systems. Other systems include systems for calendar and project management; communication; databases and spreadsheets; reference and contact management; document and word processing; image and presentation management; and domain specific processing. Many of these systems offer search functionalities within their own data space. Nevertheless, their main affordances often are other than search. By accepting the broad definition of information interaction, such systems become objects of development and evaluation in our field.

Currently, we do not have enough systematic knowledge on real-life information interaction within these activity types. From the evaluation viewpoint, we have weak understanding about the mechanisms that produce work task results and how various information interactions contribute therein. In particular, we have poor understanding of the contribution of the search activity in task performance.

As a way forward, we should reconsider what Marcia Bates asked in her article "where should the person stop and the information search interface start" almost thirty years ago [4]. Bates analyzed the search activity along two dimensions. The first one was the *level of search activity* like moves, tactics, stratagems, and strategies. The other dimension was the *level of system involvement* in carrying out the activity such as no support, listing possible actions, execution on demand, monitoring and suggestion, and automatic execution. She also proposed where the R&D efforts should be directed. To cut this short, she was advocating research that enhances the person's control of interaction and the systems' support to this, rather than automatic execution -- especially at higher search activity levels.

Bates' approach can be extended to all activity types of information interaction. One may therefore rephrase her question as "where should the person stop and the information *interaction* interface start". This invites us to ponder whether we know, what kind of support exists, is used, desirable, or possible in each activity. Some examples of support by activity type:

- for Task planning and monitoring, process models / templates-to-fill may be offered; this is at the involvement level listing, and at the activity level of strategies,
- for Searching, we have query suggestion (involvement level: listing, activity level: move)
- for Selecting, there is result clustering (involvement level: on demand, activity level: move)
- Working with items: document scrolling and key search (involvement level: on demand, activity level: move)
- Synthesis and reporting: templates to fill (involvement level: listing, activity level: tactic); copy-paste-edit (involvement level: on demand, activity level: move)

Many more examples could be given on existing or possible support features.

When this idea of interface support is coupled with task complexity, one may ask *where* more pervasive automation is possible, effective, profitable and even desirable. Probably simple and frequent tasks belong to these. And indeed, the media are already utilizing automation in, e.g., automated routine sports and weather news synthesis and reporting. And at NTCIR, there is a track developing automated systems for university entrance exams. However, in complex unstructured tasks the process needs to be under the performer's control.

I'm repeating: currently, we do not have enough systematic knowledge on real-life information interaction within these activity types. One way to organize knowledge in this area is to study people's information interaction through the four concepts: information interaction activity type, interface level of activity, interface involvement level, and task complexity, see Figure 2.

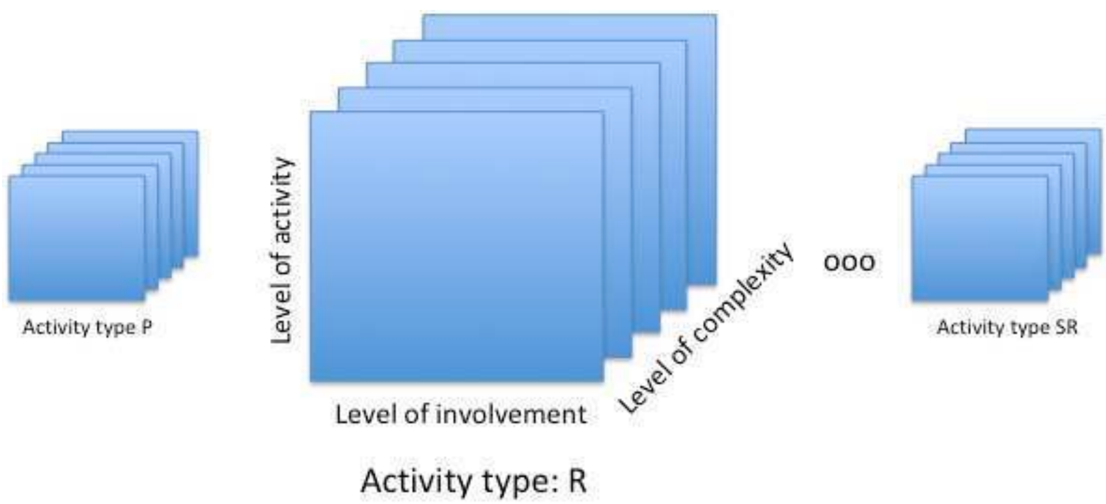

Figure 2. Information Interaction Study Cubes with focus on retrieval (**R**). **P** = planning and monitoring; **SR** = synthesis and reporting; the three dots represent the cubes for **S** = selecting and **W** = working with items.

The idea is that each existing or planned support tool, and findings on its effectiveness, can be placed in the appropriate activity type cube(s), and the cell(s) representing its level of support activity, level of involvement, and the complexity of the task supported. Empty cells indicate areas we might consider for development or assess as futile. I see a lot of research work to do.

8 In Conclusion

Technologies change and transform people's work, everyday life and their concrete information interaction practises. As information systems change, the available information and its sources may change, task processes may change – and ultimately the task goals may change. Studying these changes and what they make available, possible, effective, profitable or desirable help us best design technology that supports people as experts of their own work or life. Just like in the development of Information Science, we need to strengthen in IR the viewpoint of people's information interaction in context– not so much as users than as actors.

Several past recipients of the Salton Award have stressed the importance of people in the study of IR. The number of papers in IR analysing user data has expanded greatly. These developments demonstrate the entire information retrieval community's recognition of the importance of studying people and their information interaction in context. In my view, this is very healthy for the field. Thank you all for this community-wide recognition, for listening to me today – and for the Salton Award.

9 References

[1] Allen, T.J. (1969). Information needs and uses. In: Cuadra, C.A. (Ed.), Annual Review of Information Science and Technology: Vol. 4 (ARIST 4). Chicago, IL: William Benton, pp. 1-29.

[2] Alter, S.L. (1980). Decision support systems: Current practice and continuing challenges. Reading, MA: Addison-Wesley, 1980, 316 pp.

[3] Arvola, P., Junkkari, M., & Kekälälinen, J. (2005). Generalized contextualization method for XML information retrieval. In: Herzog, O. & al. (Eds.), Proceedings of the 14th ACM international conference on Information and knowledge management (CIKM 2005). New York, NY: ACM, pp. 20-27.

[4] Bates, M.J. (1990). Where should the person stop and the information search interface start? Information Processing & Management, 26(5): 575-591.

[5] Belkin, N.J., Oddy, R.N. & Brooks, H.M., (1982). Ask for information retrieval: Part 1. Journal of Documentation, 38(2): 61-71.

[6] Byström, K. & Järvelin K. (1995). Task complexity affects information seeking and use. Information Processing & Management, 31(2): 191-213.

[7] Caplan, N., Morrison, A. & Stambaugh, R.J. (1975). The Use of Social Science Knowledge in Policy Decisions at the National Level: A Report to the Respondents. Ann Arbor, MI: University of Michigan, Institute for Social Research.

[8] Dervin, B. & Nilan, M. (1986). Information needs and uses. In: Williams, M.E. (Ed.) Annual Review of Information Science and Technology, vol. 21 (ARIST 21). White Plains, NY: Knowledge Industry Publications, pp. 3-33.

[9] Engelbart, D. (1962). Augmenting Human Intellect: A conceptual Framework. Menlo Park, CA: Stanford Research Institute, 134 p.

[10] Hansen, P. (2011). Task-based Information Seeking and Retrieval in the Patent Domain: Processes and Relationships. Tampere: Tampere University Press, Acta Electronica

Universitatis Tamperensis, vol. 1093, 222 p. http://tampub.uta.fi/handle/10024/66773
[11] Huuskonen, S. (2014). Recording and Use of Information in a Client Information System in Child Protection Work. Tampere: Tampere University Press, Acta Electronica Universitatis Tamperensis, vol. 1387, 92+69 p. http://tampub.uta.fi/handle/10024/94893
[12] Ingwersen, P. & Järvelin, K. (2005). The Turn: Integration of Information Seeking and Retrieval in Context. Heidelberg: Springer, 448 p.
[13] Järvelin, K. & al. (2015). Task-Based Information Interaction Evaluation: The Viewpoint of Program Theory. ACM Transaction of Information Systems (ACM TOIS) 33(1), Article 3, Special Issue on Contextual Search and Recommendation, 30 p. DOI: http://dx.doi.org/10.1145/2699660
[14] Järvelin, K. & Kekäläinen, J. (2000). IR evaluation methods for highly relevant documents. In: Belkin, N.J., Ingwersen, P. & Leong, M.-K. (Eds.) SIGIR '2000: Proceedings of the 23rd Annual Conference on Research and Development in Information Retrieval (ACM SIGIR 23). New York, NY: ACM Press, pp. 41-48.
[15] Järvelin, K. & Kekäläinen, J. (2002). Cumulated gain-based evaluation of IR techniques. ACM Transactions on Information Systems (ACM TOIS), 20(4): 422-446.
[16] Kekäläinen, J. & Järvelin, K. (2002). Using graded relevance assessments in IR evaluation. Journal of the American Society for Information Science and Technology, 53(13): 1120-1129.
[17] Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. Foundations and Trends in Information Retrieval, 3(1-2): 1-224.
[18] Kelly, D., Arguello, J., Edwards, A. & Wu, W-C. (2015). Development and evaluation of search tasks for IIR experiments using a cognitive complexity framework. In: Allan, J: & al. (Eds.), Proceedings of the 2015 international conference on the theory of information retrieval. New York, NY: ACM, pp. 101-110.
[19] Kekäläinen, J. (1999). The effects of query complexity, expansion and structure on retrieval performance in probabilistic text retrieval. Tampere, Finland: University of Tampere, Acta Universitatis Tamperensis, vol. 678, 170 p.
[20] Keskustalo, H., Järvelin, K., Pirkola, A. & Kekäläinen, J. (2008). Intuition-Supporting Visualization of User's Performance Based on Explicit Negative Higher-Order Relevance. In: Sung Hyon Myaeng & al. (Eds.), Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (ACM SIGIR '08). New York, NY: ACM Press, pp. 675-682.
[21] Kochen, M. (1967). The growth of knowledge: readings on organization and retrieval of information. New York, NY: Wiley, 394 pp. ISBN-13: 9780471496953
[22] Kumpulainen, S. (2013). Task-based information access in molecular medicine: task performance, barriers, and searching within a heterogeneous information environment. Tampere: Tampere University Press, Acta Electronica Universitatis Tamperensis, vol. 1360. 81 + 80 p. http://tampub.uta.fi/handle/10024/94595
[23] Kumpulainen, S. & Järvelin, K. (2010). Information Interaction in Molecular Medicine: Analyzing Task Processes and Search Logs. In: Belkin, N.J. & Kelly, D. (Eds.), Proceedings of t f the third symposium on Information interaction in context (IIiX 2010). New York, NY: ACM, pp. 95-104.
[24] Lancaster, F.W. (1968). Information Retrieval Systems: Characteristics, Testing, and Evaluation. New York, NY: Wiley.
[25] Latour, B. & Woolgar, S. (1986). Laboratory Life: The Construction of Scientific Facts, 2nd Ed. Princeton, NJ: Princeton University Press, 294 pp. ISBN-13: 978-0691028323
[26] Li Y. & Belkin, N. (2008). A faceted approach to conceptualizing tasks in information seeking. Information Processing and Management, 44(6): 1822-1837.

[27] Menzel, H. (1966). Information needs and uses. In: Cuadra, C.A. (Ed.) Annual Review of Information Science and Technology, vol. 4 (ARIST 4). Chicago, IL: William Benton, pp. 1-29.
[28] Paisley, W. (1968). Information needs and uses. In: Cuadra, C.A. (Ed.) Annual Review of Information Science and Technology, vol. 3(ARIST 3). Chicago, IL: William Benton, pp. 1-30.
[29] Pirkola, A. (1999). Studies on Linguistic Problems and Methods in Text Retrieval. Tampere, Finland: University of Tampere, Acta Universitatis Tamperensis, vol. 672. 99 + 84 p. http://tampub.uta.fi/handle/10024/67651
[30] Robertson, S. E., Kanoulas, E. & Yilmaz, E. (2010). Extending average precision to graded relevance judgments. In: Crestani, F. & al. (Eds.) Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval (SIGIR '10). New York, NY, ACM, pp. 603-610. DOI: https://doi.org/10.1145/1835449.1835550
[31] Saastamoinen, M. (2017). Information searching in authentic work tasks: A field study on the effects of task type and complexity. Tampere: Tampere University Press, Acta Electronica Universitatis Tamperensis, vol. 1744, http://tampub.uta.fi/handle/10024/100447
[32] Salton, G. (1968). Automatic Information Organization and Retrieval. New York, NY: McGraw-Hill. 480 p.
[33] Smith, C.L & Kantor, P. B. (2008). User adaptation: good results from poor systems. In: Sung Hyon Myaeng & al. (Eds.) Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '08). New York, NY: ACM, pp. 147-154. DOI=http://dx.doi.org/10.1145/1390334.1390362
[34] Sormunen, E. (2000). A Method for Measuring Wide Range Performance of Boolean Queries in Full-Text Databases. Tampere: Tampere University Press, Acta Electronica Universitatis Tamperensis, vol. 34, http://tampub.uta.fi/handle/10024/67002
[35] Turpin, A. & Scholer, F. (2006). User performance versus precision measures for simple search tasks. In: Efthimiadis, E. & al.(Eds.), Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '06). ACM, New York, NY, USA, pp. 11-18. DOI=http://dx.doi.org/10.1145/1148170.1148176
[36] Vakkari, P. (2001a). A theory of the task-based information retrieval process: a summary and generalization of a longitudinal study. Journal of Documentation*,* 57(1): 44-60.
[37] Vakkari, P. (2001b). Changes in search tactics and relevance judgments in preparing a research proposal: A summary of findings of a longitudinal study. Information Retrieval, 4(3/4): 295-310.
[38] Vakkari, P. & Huuskonen, S. (2012), Search effort degrades search output but improves task outcome. JASIST, 63 (4): 657-670. doi:10.1002/asi.21683
[39] Vakkari, P. & Kuokkanen, M. (1997). Theory growth in information science: applications of the theory of science to a theory of information seeking. Journal of Documentation*, 53*(5): 497-519.
[40] Wersig, G. (1973). Informationssoziologie: Hinweise zu einem informationswissenschaftlichen Teilbereich. Frankfurt, Germany: Athenäum Fischer. 193 p.
[41] Wildemuth, B., Freund, L. & Toms, E. (2014). Untangling search task complexity and difficulty in the context of interactive information retrieval studies. Journal of Documentation, 70(6): 1118-1140.