Lectures on Information Retrieval. Agosti, M., Crestani, F., Pasi, G. (eds.) Springer-Verlag, 2000: 157-178. (Lecture Notes in Computer Science; 1980)

Users in Context

Peter Ingwersen

Department of Information Studies, Royal School of Library and Information Science, Birketinget 6, DK 2300 Copenhagen S, Denmark
pi@db.dk

Abstract. Users as actors in interactive information retrieval (IIR) are seen in the contexts of their perceived work tasks and information seeking behaviour. The paper models IIR processes by demonstrating a variety of approaches, ranging from Ingwersen´s cognitive communication model for IR interaction, over Saraceveic´ stratified model which includes a typology of relevance conceptions, to Borlund´s model of work task perception, information need development and relevance assessments. Other associated models and perspectives of IIR are discussed when appropriate to the major focus points of the contribution: information need development and typology; understanding of relevance in IIR; and experimental problems in IIR.

1 Introduction

Mainstream IR research, e.g. the Cranfield and TREC traditions [1], [2], assumes users as an experimental constant, commonly represented in models by sets of queries. Only in interactive IR (IIR), e.g. in the interactive TREC track as well as in all cognitive and user-driven IR research, users are seen as (more) dynamic actors, that is, as variables in the research settings [3].

In a typical TREC-like experiment one operates with a set of approximately 60 queries. Each query is regarded as a *true* representative of a *stable* information need. In non-interactive IR one expects that user requests always are identical to the information need or gap of knowledge expressed by the request. If a request for information is consisting of few terms, like often on the Web, that is what the user asks for – not that the knowledge gap might be vague. Since the non-interactive experiment assumes the query as a constant, the information need or gap *must* be seen as a stable phenomenon throughout a retrieval session. Aside from the problem of keeping the non-interactive experiment under control there exist at least two reasons for keeping requests and information needs constant. First, the goal of non-interactive experiments is to observe the *retrieval performance* of competing systems or algorithms – not really to find out why algorithms function as they do in real-life settings. Secondly, it is paramount that an assessor judges the retrieval outcome for relevance. In the non-interactive TREC experiments the role of the assessor is to generate the query and later to make assessments after one run of the systems. Thus, he acts like a user in batch mode. Obviously, if other users (or assessors) participated in experiments with *several* consecutive iterations or runs, human learning processes might occur and vary, and the one-run assessments would turn out to be inadequate or wrong.

Belkin et al. as well as Iivonen looked into this phenomenon in the early interactive track of TREC [4], [5] during which one assessor per query indeed was used like in common TREC. They found that searcher inconsistency was paramount but that one of the applied search strategies led to the "best" performance. That "best" strategy is of course unpredictable. In one-run non-interactive experiments, where several assessors are used per query, there are required more than 30 different queries in order to make the variation between assessors statistically insignificant. Below this number the ranking of the involved systems may alter in terms of performance measures [6].

The investigative problem in IIR is consequently to allow for the inclusion of the disturbing variable during experimentation, i.e. the user framed by his or her world, and, at the same time, to keep control of the experiment or investigation. Otherwise it becomes difficult to compare what one wants to compare, for instance, two best-match algorithms, two different human query modification methods, two ways of visualisations in interfaces, etc. Below, we discuss central concepts that should be taken into account in IIR research with users in action.

First, the paper discusses the basic IIR models, including the simplistic mainstream model. The models in focus are those by Ingwersen [3] and Saracevic [7], and their precursors and derivatives. The models are viewed in context of information seeking. This is followed by a discussion of information need development and dynamics over time in relation to perceived work tasks or interests, and a section on relevance conceptions and models. Some examples of experimental design conclude the paper.

2. Interactive Information Retrieval Models

Figure 1 outlines the interactive IR model as depicted by Ingwersen within the framework of the cognitive viewpoint [3], [8]. The model has gone through several modifications due to new empirical research results in international co-operation.

Basically the model operates with five central components. From the left the information objects and their representations, including thesaural nets, are in interaction with the IR system setting during retrieval. The interface component would commonly be seen as part of the entire system and functions as query generator, based on some input from the user component, for instance, in the form of a request or by selection of some visual object on the interface. The individual user – or team of individuals – displays a cognitive space that is assumed to consist of a world model developed over time from cultural and social experiences. The world model is represented by different and dynamic cognitive structures. These are assumed responsible for the actual *perception* of a work task or interest framed by the current cognitive state of the individual. According to this state, which deals with both conceptual domain-related knowledge *and* retrieval and seeking knowledge, the individual user may be in a problematic situation and state of uncertainty in the attempt to solve the problem derived from the perceived work task. If not solvable intrinsically by the cognitive state, including tacit knowledge, the individual may recognise a knowledge gap [9] or need for information. This is the intentionality behind engaging into information seeking behaviour and IIR. This view of the cognitive processes involved in IIR on the user side derives from the well-known ASK hypothesis put forward by Belkin et al. in 1982 [10]. The hypothesis operates with the similar concepts of problematic situation, cognitive uncertainty, and information need. In Figure 1 the concept of work task or interest, e.g. also of cultural nature, is introduced to explain why people get into problematic situations. There is hence a strong emphasis on the social interaction between the individual user and the situated context surrounding that individual, also over time. The actual work task may thus originate from the social-organisational environment or be produced by the user himself. The environment may take the form of scientific, professional, or social domains with recognised strategies, goals and preferences as well as tasks to be fulfilled.

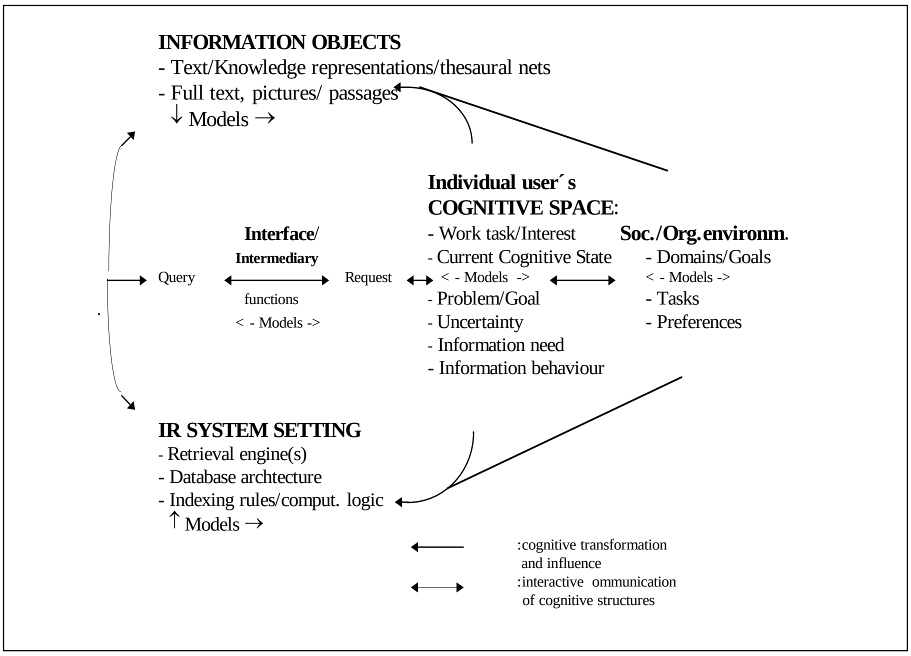

Fig. 1. Cognitive model of IR interaction. Extension of Ingwersen [8, p. 16]

In IIR *situated context* plays a central role [11]. For each of the five components there exists context. Context is signified by the notion of "models" associated with each component. The notation signifies that models of activities and solutions, behaviour and possible future situations external but crucial to the function of the component are embedded in that component, e.g. by the designer or generator or by learning. With respect to information objects the system setting, interface, user, and the social environment act as context, situated in a given retrieval or seeking activity. Authors do commonly attempt to envisage that context, for instance by generating their texts or images in such a way that they are acceptable by the prevailing domain paradigm and by (at least some) future readers or viewers. On the Internet authors also think about the system settings to which they load their objects.

The individual user is in an interesting position by being dynamic and self-contained. He or she is *influenced* by the social-organisational environment, domain, and work tasks as kinds of situated context. In fact, he or she may seek the required information to lessen the uncertainty and fill the knowledge gap by social interaction with that environment. For instance, the individual user may inquire a colleague about information. On the other hand, the same individual may look to available retrieval systems and engage in IIR – indeed also preceding or proceeding social interaction activities. One might ague that the IR systems, information objects and interfaces form part of the social-organisational environment. This would typically be the case in enterprises or organisations, but also in scientific communities. The problem for the person is now twofold. He or she must engage the information objects in order to obtain data concerning the *perceived work task* and problem. This engagement requires adequate domain knowledge to form part of the current cognitive state. However, in order to engage a system, that is, to perform *search tasks*, the cognitive state in addition must possess adequate retrieval knowledge.

We may consequently observe that the IIR consists of different types of engagements during acts of interaction. Seen from the user point of view the interaction with the social environment certainly requires conceptual (domain) knowledge but also behavioural and personal communicative skills. In order to assess the relevance of the incoming information adequate cognitive structures (conceptual presuppositions or pre-understanding) are required [12]. When interacting with IR systems the personal communicative skills are replaced by levels of retrieval knowledge, including system knowledge. Simultaneously, the user must possess sufficient conceptual domain knowledge to reach into information space to find some objects. When the system feeds back some data from retrieved information objects, the semantics or original contextual properties of that data are not supported by any behavioural signs or attitudes, as is the case in personal communication. The conceptual demand on the user as the recipient of data to carry out interpretations, assess relevance, to learn something and to fulfil his work task is hence of higher magnitude during IIR. This phenomenon also concerns the search task data fed back from the system through the interface: does the user understand the retrieval and system structures, the logic, commands, icons, etc.?

Figure 1 incorporates the simplistic retrieval model applied to mainstream system-driven retrieval research. If we make a vertical cut on the left-hand side of the model between the notion of “query” and the interface component, we observe a triangular interactive model consisting of the information objects, the system setting, and a set of queries. We then take the set of queries and one of the other two components as constants in an experimental setting typical of the mainstream approach. The remaining component is the only variable to be studied across systems.

From the rather complex scenario, Figure 1, that involves users and their social contexts, cognitive as well as probabilistic uncertainties in objects, we may understand that experimental settings in IIR fundamentally are forced to apply social science methodologies – rather than methods and settings adhering to an early age of Physics [13].

2.1 Complementary IIR models

Quite recently Belkin et al. [14] generated a four-dimensional model of IIR strategies seen as episodes of information seeking. Based on this model and that of Ingwersen, Figure 1 [3], Saracevic made a comprehensive alternative model of IIR incorporating a relevance typology [7].

Belkin et al.´s episodic model of information seeking strategies (ISS) considers the types of search a system must support. It might hence be regarded a model of IR interaction behaviour rather than an information seeking model. The underlying idea is that people commonly engage in multiple searching behaviour, both during single IR sessions and across sessions in a longitudinal sense. The goal of the model is to support retrieval (or seeking) by making design and implementation of IR systems adapt to the changing requirements of the systems. The model consists of 16 types of episodes by means of a four-dimensional classification of IR modes. Each mode contains a binary number of values and each type of behaviour is hence defined by the four-dimensional values. The four modes are *method of searching* (scanning or browsing); *mode of retrieval* (recognition or specification of relevant objects); the *goal of retrieval* (learning about the system and information space or finding relevant information); and the *resource considered* (information objects or meta-information).

The model has been applied to Web searching and navigation studies, for instance by Pharo. Pharo´s test seems to show that the model is not exhaustive enough and that there is a potential for interdependency between the method of searching and mode of retrieval [15, p. 211].

Saracevic´ alternative model of *stratified interaction* displays a three level structure consisting of surface, cognitive, and situational strata [7]. The surface level deals primarily with the computational data processing based on a query. In relation to Figure 1 this level concerns the interaction between information objects and system setting instigated by a query. The cognitive level embraces the process of perceiving information during man-machine interaction in relation to the perceived need for information. Here, the interaction involves the user through an intermediary mechanism. The situational level refers to the *information use* with respect to a perceived work task in context of an environment. As a central point we observe the longitudinal dimension of use or utility stressed in the models. Saracevic stresses that ideally the system components ought to adopt to users and vice versa. The model is clearly associated to that of Ingwersen from 1992 [8, p. 148] and its extended version [3, p. 9) depicted on Figure 1. Its strength is its comprehensiveness. The model incorporates a typology of relevance consisting of five different types of relevance. 1) Algorithmic relevance, which is the relation between the system output matched with the query features; 2) Topical relevance, i.e. the relation between the aboutness of information objects and the query; 3) Pertinence or cognitive relevance, i.e. the association between the perceived information need of the user and system output; 4) Situational relevance, which is seen as the usefulness of the objects to the current interest or work task of the user; 5) Affective/motivational relevance, associated with the goal of the user.

Situational relevance and information utility are not new concepts in information science and IR [16]. However, with the re-incorporation of situational relevance and situated context into contemporary IR research and information seeking models the issues of *informativeness* and *information use* become central research objects. This issue extends the *timeline* commonly observed during IR investigations.

Traditionally IIR research stops the experiment with users assessing the retrieval output after a number of iterations during one session. The assessments can be of situational nature where the user judges the *perceived usefulness* of the retrieved objects by means of interpretation on site of titles, summaries, or full objects. Obviously, the real usefulness or degree of informativeness and the ensuing *actual use* of (parts of) the information objects in relation to a work task is a quite different measure. It can only be taken when the user has digested their contents and associated information with the task in question, often after social interaction. Informativeness and actual use makes it obvious increasingly to view information seeking and retrieval as a whole, as proposed analytically by Vakkari [17], [18]. Longitudinal studies of information behaviour and search processes should hence play a more central role as an empirical foundation for future Information Seeking and Retrieval (ISR) models and research.

Long-term investigations of cognitive behaviour and interactive retrieval over time by Wang come here to mind [19]. Wang studied empirically from a cognitive approach the alterations of the perceived information needs over a research project, represented by the distribution of articulated unique and novel versus overlapping search terms. Wang & White goes further by investigating the actual application of documents at the *reading stage*, in particular in relation to the decisions of giving citations to the works used.

Hence, we observe how bibliometric citation studies and mapping of scientific communication patterns, using citations as representations of use, may walk hand in hand with IIR, for instance in terms of presentation of information spaces [3], [21]. The original idea of applying citation indexing as an alternative form of representation in IR derives from Garfield in 1968 [22]. Figure 1 illustrates the spiral of user inquiry, IIR, and use in context by the arrows re-connecting the user/environment to the information objects. In scientific communication the latter contains the results of author interpretations and selections of earlier contributions in the form of citations provided on the reference lists in their journal articles.

The close association between information seeking and IIR has very recently been addressed in a number of general models of information behaviour. Wilson [23] outlines 1999 the central earlier models of information seeking and other aspects of information behaviour. In the form of a nested model he places retrieval in IR systems within the sphere of information seeking processes that again are under the general umbrella of information behaviour. The discussion concerns, for instance, the sense-making framework by Dervin & Nilan [9], Kuhlthau´s phenomenological stage process model [24], Belkin et al.´s episodic model [14], Ingwersen´s cognitive communication model for interactive IR [3], and Saracevic´ stratified IR model [7]. Wilson suggests that the discussed models are *complementary* rather than conflicting, and that his proposed problem-solving model [23, p. 266] provides a basis for relating the models in relevant research strategies.

3 The Work Task - Information Need Relationship

In a pioneering effort Vakkari [17] provides a detailed analysis of theory growth in information seeking, in particular the growth of a theoretical research program on the relation between work task complexity and information seeking. This analytical work refers back to the empirical studies made by Byström & Järvelin [25] on work task complexity and information needs from 1995. The general trend observed is that the more complex the work task the less the user know about or can define his or her information need, i.e. what is unknown at present. In a cognitive sense that implies that one may assume that basically only something is known about the task or problem at hand, and perhaps nothing about what is required of information to fulfil it. In less complex situations, e.g. in the case of routine tasks or problems to be solved, the information need seems more articulated. To IIR this means that one should not always ask what the user wants, but rather about *why* he or she wants it. This goes very well together with the holistic cognitive IR theory proposed 1996 by Ingwersen [3] in which poly-representation (or multi-evidence) of the information space *as well as* of the user´s cognitive space is suggested as a way to avoid dead end retrieval situations. The theory is associated with the aforementioned ASK hypothesis on the user side [10], and the plausible inference technique by Turtle & Croft [26] as well as the Dempster-Shafer uncertainty logic and multiple-evidence principles proposed by van Rijsbergen & Lalmas [27].

Table 1. Matrix of four distinct cases of human intrinsic information needs, given a perceived work task situation, and the corresponding seeking and interactive information retrieval behaviour - simplified version of Ingwersen [3, p. 15].

Intrinsic information need vriables – given a perceived work task	**Well-defined**	**Ill-defined**
Stable	***Verificative*** ***Conscious topical*** *Querying* Filtering behaviour	***Muddled task & info.need*** Search loops
Variable	***Conscious topical*** *Query-Navigation* Dynamic interaction	***Defined work task*** ***Muddled info.need*** *Browsing* Try-&-error behaviour

Ingwersen [3] proposes to view the *perceived work task* as a rather stable cognitive state during retrieval session time, but not in a longitudinal sense. The corresponding information need, however, may be dynamic and develop simultaneously due to interpretation and learning processes during IIR. According to Ingwersen one may view the basic cases of information needs in the form of a matrix, defined by two dimensions. One is the degree of perception of the need, i.e. how well is the information need defined in the mind of the user at a given point in time. The other dimension corresponds to the degree of variability of the need over time, i.e. the motivation and ability for change. In line with the holistic cognitive view each individual user will act differently to the same given work task due to differences in the perception influenced by the socio-cultural history of the person and his or her perception of the current context. Notwithstanding, each user should have a degree of understanding of the work task. Otherwise there would not exist any *reason* or cause for engaging in information seeking or IIR. Depending on the current cognitive state, the user may belong to one of the four cases of human information needs at the initiation of the IIR session, Table 1. During retrieval he or she may move to other cases and hence require different kinds of retrieval support.

On Table 1 the transition between the four cases is continuous. The matrix operates with three kinds of intrinsic information needs [28]. *Verificative* needs signify that the user wishes to verify information objects with known non-topical (structured) data, such as author names, client address, cited authors, journal name, etc. This type is assumed to be stable during a session period until objects have (not) been retrieved. *Conscious topical* needs for information imply that the user wants to clarify, review or pursue information in known subject matter and domain. Known subject matter signifies topical (unstructured) data on contents, such as terms, concepts, image representation, etc. This type is assumed to be either of intrinsically stable nature, like the verificative one, or variable over session time. The third kind of information need is called *muddled* or ill-defined. The user is engaged in the exploration of new concepts and relations outside known subject matter or domain, or the known data are incomplete and cognitively vague. In reality one may observe needs that are mixed of verificative and conscious topical ones. We will functionally regard such blends as belonging to the conscious topical kind of information need. We can see that OPACs commonly deal with verificative needs, having some difficulty with the topical ones. Bibliographic and full-object databases are traditionally suited to both verificative and topical information needs.

If users constantly acted rationally, that is, if they expressed *everything they know* about their perceived need and work task, IIR could possibly handle the well-defined information needs quite properly. Besides, the system-driven approach to IR, that information needs are stable and queries (requests) exactly mirror the underlying needs, would be much more in line with reality. However, the problems in IIR are not confined to muddled need situations alone, whether stable or variable. People tend to act at random, to be *uncertain*, and *not to express everything* they know. Instead they express what they assume is enough and/or suitable to the intermediary and/or IR system. They compromise their statements under influence of the current *context and situation.* This context involves the perceived work task or domain *as well as* the perception of the search task, i.e. the understanding of the system and knowledge sources to engage with.

This phenomenon is called the *Label Effect.* The effect was predicted by Taylor in 1968 [29] and empirically verified and discussed by Ingwersen 1982 [30], [8].

The Label Effect means that users, even with well-defined knowledge of their information gap, tend to label their initial request for information verbally by means of very few terms or single concepts. It implies two obstacles to successful IR: First, intermediary mechanisms have difficulty in reaching out into the proper directions in information space where data are located relevant to that particular user. Due to the lack of *context* in the request a multitude of directions are indeed possible. This is what we observed in the scientific online age 10-20 years ago, and the same phenomenon is dominant today in web searching. Secondly, intermediary mechanisms may not be capable of distinguishing between users with detailed, some or no knowledge about their information requirements, that is, whether the user is intrinsically well or ill-defined concerning the ASK. It becomes hence difficult for the system to support adequately the user in his or her retrieval endeavour.

A closer observation of the matrix, Table 1, suggests the following issues of concern to IIR research. Mainstream IR research is fundamentally concerned with the investigation of the well-defined and stable case of the matrix. Indeed, we have such kinds of needs, for instance, in connection with patent retrieval and filtering, i.e. selective dissemination of information (SDI) as it is called in information science. In this case IR may support users by means of querying and/or confined navigation. Users will be expected to be less uncertain and be capable of query modification as well as assessing topical relevance as well as pertinence due to the rich cognitive state and situational relevance due to work task perception.

In the case concerned with well-defined but variable information needs people are assumed to be willing (or forced) to learn and shift focus after initial engagement. We may expect exploratory navigation and *stages of uncertainty* throughout the IR session, in line with the “berry-picking” exploratory behaviour suggested and modelled by Bates [31]. The cognitive uncertainty is empirically found to increase during the initial stages of IR (and seeking) processes due to interpretative problems of the retrieved data [32] and the quality of the cognitive state. Situational relevance assessments are possible due to known work task, but query modification as well as topicality and pertinence assessments may be unreliable at initial stages of engagement with the system.

The muddled and variable kind of information needs seem to require means of browsing rather than querying due to the inherent Label Effect. Cognitive uncertainty will be expected to be high and we may observe try and error behaviour during searching, since new adequate search features may be hard to recall from memory or non-existent. However, the motivation and curiosity of the user may make the session progress. With the exception of situational relevance assessments judgements of topicality or pertinence are assumed very cumbersome during many stages of the IIR process, as is query modification. In fact, the only cognitive structures assumed to be present are those associated with the perceived work task or interest.

The final case of ill-defined but stable information needs also assumes uncertainty as to the work task definition. It might be highly complex [25] but the work task may also be vaguely represented in a cognitive sense for another reason. In the case of human mediators (librarians), Ingwersen found [30] that they rarely possess a complete picture of the work task or problem of the end-user. What they often only know is “something” – a few terms or concepts – extracted from the user during personal communication. The Label Effect clearly appears in such cases and the mediator runs into search loops. In order to break the dead end the mediator´s cognitive state must rapidly absorb knowledge about the user situation. In generalist circumstances, for instance, in public libraries, this "getting to know" the underlying situation is often hampered by lack of domain knowledge on the mediator side. On the other hand, in specialised information services in organisations the mediators often know of the current tasks of their end-users, due to collaboration, and the muddled case can be solved or moved to another case in the matrix. However, when being in the fourth case the searcher have severe difficulty with respect to all kinds of relevance assessments as well as query modification activity. The reason that public librarians after all often succeed is grounded in their extensive retrieval knowledge that may guide them to probably proper locations in information space.

The matrix, Table 1, demonstrates that only in one-two cases can we expect users to act according to plans in rational ways, i.e. in the well-defined cases. This difference also lies in the notions of navigation versus browsing. Navigation is seen as purposeful moves by links or similar activity in networks of information objects. The user seeks to fulfil a goal, either by navigating in a confined space or by a more exploratory behaviour – but constantly with the work task or end goal in mind. Browsing signifies an activity of randomness in searching. The searcher is open to novel paths and serendipity effects may occur. Recently, Hong [33] has published a model of intentions and shifts that take place during IIR, also applying the episodic IIR model by Belkin et al. [14]. Hong´s model primarily concerns the well-defined cases discussed above.

Due to the randomness, vagueness, and Label Effects of users, at least during the initial stages of IIR, retrieval research might profit from also concentrating on the perceived work tasks and their function in IIR. Essentially, the concern is to make possible for the system to obtain several *simultaneous evidences* or representations of the user work task and underlying situation. Such representations do not exclude non-topical data types, such as meta-data. The system is hence better equipped to support the user in his or her information retrieval activities.

4. Relevance Issues in IIR

Relevance has become a major area of study in information science. A wide variety of subject fields have tried to deal with this concept. Theoretical frameworks abound, and yet, relevance is also a concept that is intuitively understood, but very difficult to define. Nevertheless, since information science was first seen as a distinct discipline in the 1940s, relevance has been identified as one of its fundamental and central concept [34].

In the past, studies have concentrated on either a systems-centred or user-centred or cognitive approaches to information retrieval. In systems approaches to IR, relevance is considered to be a property of the system, whereas in user-oriented and cognitive traditiona, relevance is *directly associated* to the cognitive processes of the users and their changing knowledge and needs regarding information, stimulated by the context. Furthermore, there are many kinds of relevance, not one only, as discussed by Mizzaro [35]. It is clear that the concept of relevance covers a very wide area of knowledge, and it is perhaps owing to this diversity that the latest studies concentrate on the *interaction* between the user and the system in trying to establish what relevance really is. It is during this interaction that an important dimension must be added, namely that of *time.* As the cognitive state may change over session time or across sessions both the information need and relevance may change for the same user. This time dimension can be measured and plotted in terms of *information-seeking* stages and successive searches, as empirically shown by Wang [19] and Spink et al. [36].

Saracevic´ stratified model of IIR also offers an integrated framework to incorporate a system of relevance, and states that "[T]he effectiveness of IR depends on the effectiveness of the interplay and adaptation of various relevance manifestations, organised in a system of relevances. Thus the major direction of R&D in information science should be toward increasing the effectiveness of relevance inter-plays and interactions. This should be the whole point of relevance research in information science" [7, p. 216]. Following the framework, discussed above, relevance manifests itself on different levels or strata. Relevance inferences may differ at various levels, but the inferences are always interdependent, and IR evaluation is all about comparing relevance inferences from different levels. Relevance can be typified at different levels of manifestation, and we can study its behaviour and effects within and between strata. As briefly defined above, Saracevic´ relevance system contains the following relevance manifestations: algorithmic; topical; cognitive relevance or pertinence; situational; and motivational & affective relevance.

Saracevic´ underlying assumption is that relevance is rooted in human cognition. This is summarised by the following words: "As a cognitive notion relevance involves an interactive, dynamic establishment of a relation by inference, with intentions toward a context" [7, p. 206]. On the surface it would mean that relevance is a subjective phenomenon. However, also an objective manifestation can be found, namely that made available by the IR system itself directly through its algorithms. Since the algorithms are cognitive representations the assumption holds. In the early days of IR experimentation [1], and during the following two decades, relevance was seen as either objective, i.e. system relevance that could be measured, e.g., by recall and precision, or a subjective relevance in the form of utility. With the article by Schamber et al. in 1990 [34] this dual notion of relevance was scattered by the introduction of situational relevance manifestations, and more recent analytical as well as empirical contributions on the matter all treat relevance as a multi-dimensional phenomenon. Simultaneously, relevance is not anymore consisting of a binary scale used during assessment activities in all IR experiments. The empirical studies by Borlund & Ingwersen [37] as well as Greisdorf & Spink [36], [39] are examples of a non-binary approach to relevance *and* scaling.

4.1 Dimensions of Relevance

Saracevic operates with the following attributes that makes it possible to distinguish between the above mentioned five manifestations of relevance, Table 2.

Table 2. Attributes of relevance according to Saracevic [7].

Attributes of Relevance	
Relation	Relevance always implies a relation, often in communication or exchange.
Intention	The relation in expression of relevance involves intentions such as objectives, roles, expectations (motivation).
Context	Intention always comes from a context, and is always directed toward that context.
Inference	Assessment (often graduated) of the effectiveness of a given relation.
Interaction	Inference is accomplished as a dynamic process of interaction, and interpretations of the other attributes change as cognition changes.

Looking at the attributes of relevance as listed above, Table 2, it is clear that relevance always indicates a relation. Different manifestations of relevance indicate different relations. It would therefore seem that the trend moves toward viewing relevance in IR not as a single definition of relevance, but as a system of relevances (note the plural). Consequently no single relevance in the system can be viewed in isolation. Relevance exists as an interacting system of manifestations on different levels.

According to Cosijn & Ingwersen [39] who analysed Saracevic´ framework [7] and plotted his manifestations against the attributes, it is interesting to note that the relevance manifestations are moving from a systems approach to a user- and socially oriented approach. Thus the whole spectrum is included in relation to the IIR components, Figure 1. More importantly their analysis demonstrates that the attributes function in *different dimensions* for the different manifestations of relevance, Table 3.

This raises two issues associated with the original framework: First, should the defined manifestation of *affective or motivational relevance* be regarded as part of a linear scale of moving from objective to subjective relevance? One may argue that motivational relevance is the same as the *intent attribute.* Cosijn & Ingwersen suggest to replace it by a *socio-cognitive relevance* as the ultimate manifestation of relevance on a linear scale, as proposed by Ørom under the label of contextual relevance [40], and corresponding to domain-related relevance proposed by Hjørland [41]. Secondly, one may regard affective relevance as a dimension of relevance influencing *all* the preceding subjective relevance types. Cosijn & Ingwersen argue [39] that affective relevance is *not* a discrete category or part of a linear scale. It should rather be viewed as part of, and influencing the subjective types of relevance (topical, pertinence, situational and socio-cognitive relevance).

Table 3. Attributes and manifestations of relevance. Revision of Saracevic´ framework [7] by Cosijn & Ingwersen [39]. Socio-cognitive relevance has replaced affective-motivational relevance in right-most column.

	Manifestations of Relevance				
Attrib-tes of Rele-vance		⇔	**Affective**	**Relevance**	⇔
	Algorithmic	**Topical**	**Cognitive / Pertinence**	**Situational / Utility**	**Socio-Cognitive**
Rela-tion	Query ⇒ Information objects (feature-based)	Subject/topic expressed in query ⇒ information objects	State of knowledge/cognitive information need ⇒ Information objects	Situation, work task or problem at hand as perceived ⇒ Information objects	Situation, task or problem at hand as perceived in socio-cultural context ⇒ Information objects
Inten-tion	(a) System dependent (b) Intent/ motivation behind algorithm	(a) User /assessor expectations (b) Intent/ motivation behind query	Highly personal and subjective, related to information need, intentions and motivations	Highly personal and subjective or even emotional. Related to goals, intentions and motivations	Personal, subjective / org. strategy. Related to user's experience, traditions, scientific paradigms
Con-text	Tuning search engine performance (e.g. TREC)	All types of subjective relevance are, by definition, dependent (user´s/assessor´s context-context)			
Infe-rence	Weighting and ranking functions	Interpretation of aboutness and subject matter at semantic level	Subjective and individualised process of cognitive/pragmatic interpretation, selection and filtering	User's ability to utilise information objects in a way meaningful to user	Users' (or group's) ability to utilise information objects, meaningful to environment
Interac-tion	Automatic relevance feedback or query modification	Relevance judgements are content dependent	Relevance judgements are content, feature, form & presentation dependent	Including interaction *with* environment	Including interaction *within* environment
		Increasing	**Time**	**Dependence**	⇒

Table 3 displays in tabular form the final analysis result with the socio-cognitive relevance manifestation replacing affective or motivational relevance. The table illustrates the difference between topicality observed by a person external to the user-system interaction process, for instance, an assessor, and topicality assessed by a user having a real information need. The relation and intention attributes work differently depending on whether the judge is internal or external to an investigation. External observers can only assess the aboutness of the *expressed* need, in the form of a query or request, in relation to retrieved information objects. Due to the Label Effect such an assessment may be rather different from that made by the user, in particular if he or she possesses an intrinsically well-defined information need. The Borlund & Ingwersen model, Figure 2, originating in a simplified form from [37], takes this distinction into account, naming the user´s topical assessment as intellectual topicality.

External assessors may judge algorithmic relevance by means of (intellectual) topicality, as done in non-interactive TREC after a one-run session. However, the *time* issue plays a distinctive role, as discussed earlier on. The more IIR iterations by users the more likely the inconsistency in between users and between users and assessors. *Pertinence* as well as situational relevance belongs clearly to the sphere of the individual user. Pertinence is a result of pragmatic interpretations of objects, for instance, their novelty, also over session time, their way of being presented, their credibility, and/or in relation to features different from subject matter, such as a publishing journal. Affective issues come into play, like in connection to the situational and socio-cognitive relevance manifestations. *Situational relevance* is influenced by the context, e.g. the domain, the perceived work task, organisational preferences, etc., Figure 1, but the individual user is the deciding factor in relation to usefulness of objects on site. At a later stage in the process of fulfilling the work task the user may be further influenced by social interaction. Hence, we move into the region of *socio-cognitive* relevance. Examples of this manifestation are, for instance, reviewer meetings in editorial boards or conference programme committees assessing contributed papers, co-author negotiations of which articles to use and cite, or workshop and research seminar discussions of papers and ideas. Schamber provides a comprehensive review of the historical and contemporary issues of relevance up towards 1994 [42].

4.2. Experimental Issues on Relevance and Work Tasks in IIR

Saracevic´ typological relevance framework as well as the work task - information retrieval issue have inspired several empirical research efforts and further modelling of relevance phenomena recently. For instance, Borlund & Ingwersen [37] explored the possibility of applying situational relevance and topicality assessments in interactive IR experiments using non-binary relevance assessments and by modelling Saracevic´ framework [7] in relation to real as well as simulated work tasks. That model can be modified to emphasise the relation between a *perceived* work task, the dynamic information need developing over time, and relevance categories for test persons as well as assessors in IR experiments - Figure 2. The same approach also evaluates performance measures that relate types of relevance and suggests novel relevance ranking measures, also suitable for non-interactive IR [43]. Järvelin & Kekälainen [44] are also concerned with relevance ranking in non-binary mode by proposing and testing the cumulative gain the user obtains by examining the retrieval results up to a given ranked position. Spink et al. [36], [38] empirically investigate what they call regions of relevance, including the application of non-binary assessments in large-scale interactive studies. They develop the relevance framework further. In an empirical research environment Vakkari [45] and Vakkari & Hakala [46] investigate the development of relevance, relevance criteria and contributing information types of searched documents in task performance over an academic term period. Similarly, Wang & White investigate and provide a cognitive model of the actual application of documents at the reading stage in particular in relation to the decisions of giving citations to the works used [20].

Figure 2 illustrates the manifestations of relevance set into an experimental IR model.

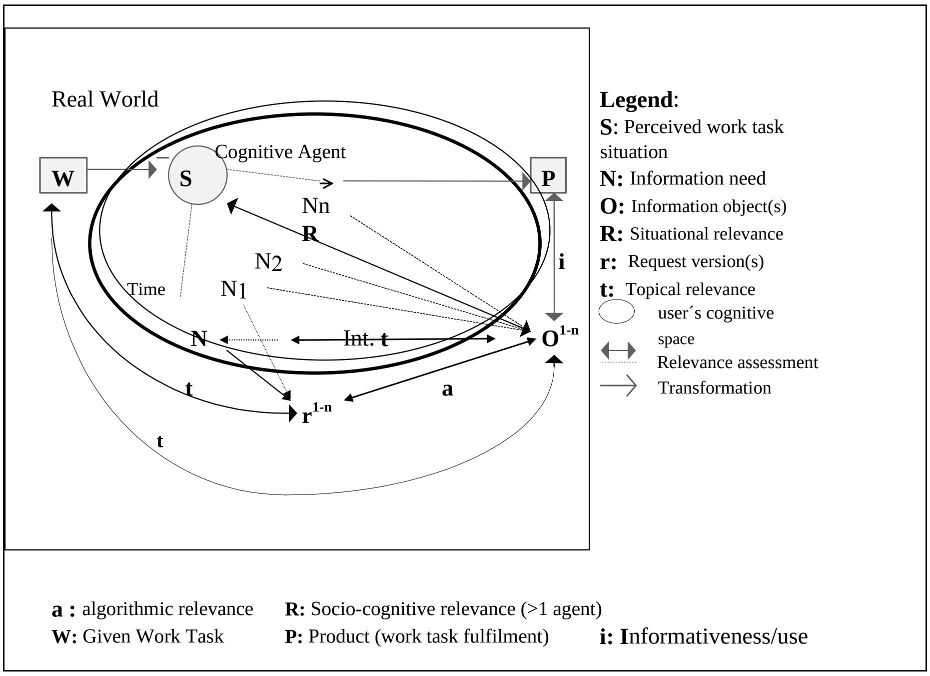

Fig. 2. Manifestations of relevance in interactive information retrieval with a given work task, cognitive spaces of (two) cognitive agents (ellipses), information objects retrieved from a system (not shown), and the final information product.

On the figure two cognitive agents are symbolised by ellipses. A cognitive agent may in this setting be a user or an assessor. The traditional and most commonly used type of relevance is the algorithmic relevance (a) which expresses the degree of match between the request/query version (r^{1-n}) and the retrieved information objects (O^{1-n}) resulting in a ranked output. The topical founded, though intellectually influenced type of relevance, intellectual topicality (Int. t), is signified by the topical nearness (aboutness) between the retrieved objects (O^{1-n}) and the topic of the information need version (N^{1-n}). Situational relevance (R) signifies the relationship between the retrieved objects (O^{1-n}) and the perceived cognitive work task situation (S) originated from (W), for instance in the form of usefulness. Because intellectual topicality may continue into pertinence assessments, i.e. the relation between objects and current information need, e.g. in the form of novelty or applicable features of objects, a dotted line continues the intellectual topicality line of assessment towards the need (N). In short, the two relevance manifestations might be difficult to distinguish during experiments. If two users interact as a team, socio-cognitive relevance assessments may be observed as negotiated decisions on the usefulness of objects, under influence of the understanding by the agents of the work task. The fulfilment of (part of) a work task commonly results in a product which can be observed in the real world. The relationship between previously observed relevance assessments, e.g. situational, intellectual topical, or algorithmic, and the actual use of objects in the product provides an observable measure of informativeness (i), as done by Wang & White [20].

It is important to note that the *time dimension* is incorporated into the model, Figure 2, signified by the sequential versions of the information need ($^{N1\text{-}n}$) and the corresponding request versions ($^{r1\text{-}n}$). The perceived work task (S) is regarded rather stable during session time, but not across sessions. Since the request versions/queries form part of the real world they are observable. One may hence assess the relationship between requests and the given work task in order to observe manifestations of *interpretations* of the original work task (t), also across test persons.

This sort of assessment is *ideally* of topical nature. Similarly, an observer might focus on the nearness of the given work task and the retrieved objects, also those that have been judged highly relevant by the body of users (or systems). Indeed, this is what TREC assessors commonly do. They create the task and observe the (topical) relevance of the output, pooled from the algorithmic results of the involved IR systems. In this respect one may refer back to Saracevic´ notion that relevance inference are interdependent and IR evaluation is all about comparing relevance inferences from different levels or strata [7]. However, it is easy to see that such ideally objective observations and measurements in fact are *subjective*. *The observer turns into a cognitive agent*, as depicted in the model, Figure 2, in line with a user. Consequently, the observer/assessor suffers from the *same problems* of perception and interpretation of the work task, as do users. The objective “topical” assessments by the assessor turn out to be of intellectual nature or perhaps in the form of pertinence or even situational relevance. In particular, if the assessor has generated the task, he or she should have an idea about in which situated context the task is supposed to function. The assessor is then like a user with a real information need. How do we then know which relevance manifestation the assessor actually applies when assessing? How much influence has the context on him? Some of these facets of relevance assessments were already discussed in 1973 by Cooper [47].

If the assessor has not generated the work task he or she becomes similar to a user during either a brief one-run session (in non-interactive TREC) or a real IIR experiment with multiple runs with given simulated topics or information needs to fulfil. In the latter case the *presentation and form* of the given work task becomes crucial for the outcome. Common TREC “topics” are rarely situated or expressing a context of use. This is the reason for explicitly to apply cover stories or *simulated work task situations* during IIR experiments [37], [43], [48], preferably in a classic social science setting (placebo experiments) if comparing two systems or features of interfaces [13].

The model, Figure 2, demonstrates the influence of a real or a given simulated work task. Simulated work tasks or cover stories provide a context or *reason* for that people should look for information. Each test person is then free to make his or her own perception and interpretation of that context, forming his own perceived information need. The more limited the context the more open semantically the possibilities for interpretation. Borlund currently evaluates this research methodology, by comparing real needs for information with simulated ones, and by investigating characteristics of well-functioning simulated work tasks [48], including their domain-dependency. Other proposals for applying search and work tasks as instrumental in IIR investigations are currently under development [49]. The advantage of applying such simulations is that users are free to interpret the situations which, when given to a number of test persons, makes the experimental setting controllable and manageable.

5. Concluding Remarks

In contrast to non-interactive IR experiments IIR investigations rely on a relatively high number of test persons and work tasks/information need situations (search jobs) that make the results statistically reliable, whether the goal is measuring performance, functionality or behavioural aspects of retrieval interaction. The rule of thumb is that behavioural or cognitively related investigations of quantitative nature require, as a minimum, 40-50 participants and 2-3 work tasks of simulated and/or real nature per person in order to be reliable. In the case of qualitative empirical studies the number of persons may be less. Since we clearly are talking about sampling the numbers will depend on the target population. The sample population as well as the chosen test situations and systems under study should always belong to a defined knowledge or work domain. Naturally, if the entire population, say in an organisation, is analysed smaller studies are still valid since no sampling takes place. But one may not be able to generalise beyond the organisation or domain. In performance comparisons between systems or interface features one may decrease the number of test persons to minimum 25-30 but should apply an increased number of test situations or search jobs per person. In particular care should be taken of the cross tabulation of test situations applied between two groups of test persons where one acts as the control group of the other. Similarly, the test situations can be made to function as sets of control. The search jobs should be permuted in order to normalise learning effects using the systems under testing and to avoid that particular job sequences define the searching behaviour [37].

Another dimension of IIR and information seeking studies and investigations is the question of who to use as test persons. We observe often in information research that information or computer science students are used as participants, since they are easy to approach. This is not a good idea if the purpose of the study is to observe how a chemical information system and interface function. Another facet of this dimension is the problem that face interface investigations. Novel interfaces or interface functionality are always new to everybody. Many consecutive sessions are thus required in order for the participants to accustom to the interface prior to the serious investigation of the interaction is carried out. Too many interface solutions have prematurely been disregarded due to studies made over a few sessions by users new to the system/interface.

Users in IIR evaluations cannot be asked to perform an unlimited number of work tasks. On average, a test search may take 20-30 minutes per person. Four work tasks per test session are consequently seen as the maximum, in particular if the researcher also applies post search interviewing. Since the investigation often requires more than four tasks to be performed by each participant, several test sessions are necessary and (unwanted) learning processes may take place between sessions. Careful planning, the application of several data collection methods (triangulation) and normalisation of the test situations can solve some of these problems [13].

By applying several simulated work task situations it becomes possible to induce controlled parameters in the search jobs, such as features unfamiliar to the user population, in order to investigate IR system functionality and performance during interaction with knowledge weak populations. In general, IIR experiments should make more use of *differentiated* work task situations in order to observe, for instance, which kind of situations or information needs are best fulfilled by which kinds of interface functions or retrieval algorithms. As we have seen there exists a variety of work task situations, of corresponding information need types as well as of relevance types. We also know about several algorithmic means to retrieve information objects at the theoretical level or as black box experiments. Such means include methods of visualisation, query modification and relevance feedback. From of these features studies involving users we experience often that their theoretical attractiveness in practical interaction raises new issues and problems. For instance, Beaulieu has found that users have severe problems in understanding what relevance feedback and human query modification entail [50]. Further, the common probabilistic weighting methods often produce the same information objects again and again via relevance feedback algorithms on the top of the ranked output, only changing the ranking order of relevance from iteration to iteration. This is because all relevant objects chosen over the session have equal weights. Campbell [51] has thus investigated to alter the algorithms so that the present visualised objects that are chosen to be of interest have highest weights. One should note that not only the direct interactions between man and system are researchable in IIR. In fact, the entire system of interactions and transformations, demonstrated on the model, Figure 1, are contributing to our understanding of IIR. However, our comprehension of how the information and system spaces interact with the cognitive and social ones is far from complete [52]. The grey areas on the IIR research map are *legio*.

References

1. Cleverdon, C.W., Keen, E.M. Factors Determining the Performance of Indexing Systems. Vol. 1: Design. Vol. 2: Results. Aslib Cranfield Research Project, Cranfield, UK (1966).
2. Harman, D.K. The TREC Conferences. In: Sparck Jones, K. and Willitt, P. (eds.): Readings in Information Retrieval. Morgan Kaufmann, San Francisco (1997) 247-256.
3. Ingwersen, P. Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. Journal of Documentation **52**, 1 (1996) 3-50.
4. Belkin, N.J., Kantor, P., Fox, E.A., Shaw, J.A. Combining the Evidence of Multiple Query Representation for Information Retrieval. Information Processing and Management **31** (1995) 431-448.
5. Iivonen, M. Consistency in the Selection of Search Concepts and Search Terms. Information Processing and Management **31**, 2 (1995) 173-190.
6. Voorhees, E.M. Variations in Relevance Judgements and the Measurement of Retrieval Effectiveness. In: Croft, B.W., Moffat, A., van Rijsbergen, C.J., Wilkinson, R. and Zobel, H. (eds): Proceedings of the 21st ACM SIGIR Conference on research and development of Information retrieval. ACM Press/ York Press, Melbourne, Australia (1998) 315 – 323.
7. Saracevic, T. Relevance Reconsidered '96. In: Ingwersen, P., Pors, N.O. (eds): Information Science: Integration in Perspective. Royal School of Library and Information Science, Copenhagen, Denmark (1996) 201-218.
8. Ingwersen, P. Information Retrieval Interaction. Taylor Graham, London (1992).
9. Dervin, Brenda, Nilan, M. Information Needs and Uses. In: Williams, M.E. (ed.): Annual Review of Information Science and Technology, Vol. 21. Learned Information for the Am. Soc. for Information Science, Medford, NJ (1986) 3-33.
10. Belkin, N.J., Oddy, R., Brooks, H. ASK for Information Retrieval. Journal of Documentation **38**,3 (1982) 61-71.
11. Wilson, T.D. On User Studies and Information Needs. Journal of Documentation. **37**,1 (1981) 3-15.
12. Winograd, T., Flores, C.F. Understanding Computers and Cognition: A New Foundation for Design. Addison-Wesley, Norwood, NJ (1986).
13. Pors, N.O. Information Retrieval, Experimental Models and Statistical Analysis. Journal of Documentation. **56**,1(2000) 55-70.
14. Belkin, N.J., Cool, C., Stein, A., Thiel, U. Case, Scripts, and Information Seeking Strategies: On the Design of Interactive Information Retrieval Systems. Expert Systems with Applications **9** (1995) 379-395.
15. Pharo, N. Web Information Search Strategies: A Model for Classifying Web Interaction. In: Aparac, T., Saracevic, T., Ingwersen, P., Vakkari, P. (eds.): Proceedings from the 3rd International Conference on Conceptions of Library and Information Science, CoLIS 3: Digital Libraries: Interdisciplinary Concepts, Challenges and Opportunities. University Of Zagreb, Zagreb, Croatia (1999) 207-218.
16. Wilson, P. Situational Relevance. Information Storage and Retrieval **9** (1973) 457-471.
17. Vakkari, P. Growth of Theories on Information Seeking: An Analysis of Growth of a Theoretical Research Program on the Relation between Task Complexity and Information Seeking. Information Processing & Management **34** (1998) 361-382.

18. Vakkari, P. Task Complexity, Problem Structure and Information Actions: Integrating Studies on Information Seeking and Retrieval. Information Processing & Management **35,**6 (1999) 819-837.
19. Wang, P. Users' Information Needs at Different Stages of a Research Project: A Cognitive View. In: Vakkari, P. (eds): Information Seeking in Context. Taylor Graham, London (1997) 307 – 318.
20. Wang, P., White, M.D. A Cognitive Model of Document Use during a Research Project. Study II. Decisions at the Reading and Citing Stages. Journal of the American Society for Information Science **50**,2 (1999) 98-114.
21. Ding, Y., Chowdhury, G.. Foo, F. Mapping the Intellectual Structure of Information Retrieval Studies: An Author Co-Citation Analysis, 1987-1997. Journal of Information Science: Principles and Practice **25,**1 (1999) 67-78.
22. Garfield, E. From Citation Indexes to Informetrics: Is the Tail Wagging the Dog? Libri: International Library Review **48** (1998) 67-80.
23. Wilson, T.D. Models in Information Behaviour Research. Journal of Documentation. **55**,3 (1999) 249-270.
24. Kuhlthau, C.C. Inside the Search Process: Information Seeking from the User´s Perspective. Journal of the American Society for Information Science **42** (1991) 361-371.
25. Byström, K., Järvelin, K. Task Complexity Affects Information Seeking and Use. Information Processing & Management **31**,2 (1995) 191-214.
26. Turtle, H., Croft, W.B. Inference Methods For Document Retrieval. ACM-Sigir Forum **June** (1990) 1-24.
27. van Rijsbergen, C.J., Lalmas, M. An Information Calculus for Information Retrieval. Journal of the American Society for Information Science **47** (1996) 385-398.
28. Ingwersen, P. Cognitive Analysis and the Role of the Intermediary in Information Retrieval. In: Davies, R., (ed.): Intelligent Information Systems. Horwood, Chichester, West Sussex, England (1986) 206-237.
29. Taylor, R.S. Question Negotiation and Information Seeking in Libraries. College and Research Libraries **29** (1968) 178-194.
30. Ingwersen, P. Search Procedures in the Library Analysed from the Cognitive Point of View. Journal of Documentation **38,**3 (1982) 165-191.
31. Bates, M.J. The Design of Browsing and Berry-Picking Techniques for the Online Search Interface. Online and CD-ROM Review **13,**5 (1989) 407-424.
32. Kuhlthau, C.C. The Role of Experience in the Information Search Process of an early Career Information Worker: Perceptions of Uncertainty, Complexity, Construction, and Sources. Journal of the American Society for Information Science **50**,5 (1999) 399-412.
33. Xie, H. Shifts of Interactive Intentions and Information-Seeking Strategies in Interactive Information Retrieval. Journal of the American Society for Information Science **51** (2000) 841-857.
34. Schamber, L., Eisenberg, M., Nilan, M. A Re-Examination of Relevance: Toward a Dynamic, Situational Definition. Information Processing & Management **26**,6 (1990) 755-776.

35. Mizzaro, S. Relevance: The Whole History. Journal of the American Society for Information Science **48**,9 (1997) 810 – 832.
36. Spink, A., Greisdorf, H., Bateman, J. From Highly Relevant to not Relevant: Examining Different Regions of Relevance. Information Processing & Management **34**,5 (1998) 599-621.
37. Borlund, P., Ingwersen, P. The Development of a Method for the Evaluation of Interactive Information Retrieval Systems. Journal of Documentation **53**,3 (1997) 225-250.
38. Spink, A., Greisdorf, H. Regions of Relevance Study, Part 1: Measuring and Mapping Users´ Relevance Judgements. Journal of the American Society for Information Science **51** (2000). In press.
39. Cosijn, E., Ingwersen, P. Dimensions of Relevance. Information Processing & Management **36** (2000) 533-550.
40. Ørom, A. 2000. Information Science, Historical Changes and Social Aspects: A Nordic Look. Journal of Documentation **56** (2000) 12-26.
41. Hjørland, B. Information Seeking and Subject Representation: An Activity-Theoretical Approach to Information Science. Greenwood Press, London (1997)
42. Schamber, L. Relevance and Information Behavior. In: Williams, M. (ed.): Annual Review of Information Science and Technology **29** (1994) 3-48.
43. Borlund, P., Ingwersen, P. Measures of Relative Relevance and Ranked Half-Life: Performance Indicators for Interactive IR. In: Wilkinson, R., Croft, B.W., van Rijsbergen, C.J. (eds.): Proceedings of the 21st ACM-Sigir Conference on Research and Development of Information Retrieval. York Press, Melbourne, Australia (1998) 324-331.
44. Järvelin, K., Kekälainen, J. IR Evaluation Methods for Retrieving highly Relevant Documents. In: Belkin, N.J., Ingwersen, P., Mun-Kew Leong (eds.): Proceedings of the 23rd ACM-Sigir Conference on Research and Development of Information Retrieval. Sigir Forum **34** (2000) 41-48.
45. Vakkari, P. Relevance and Contributing Information Types of Searched Documents in Task Performance. In: Belkin, N.J., Ingwersen, P., Mun-Kew Leong (eds.): Proceedings of the 23rd ACM-Sigir Conference on Research and Development of Information Retrieval. Sigir Forum **34** (2000) 2-9.
46. Vakkari, P., Hakkala, N. Changes in Relevance Criteria and Problem Stages in Task Performance. Journal of Documentation. **56**(5) (2000) 540-562.
47. Cooper, W.S. A Definition of Relevance for Information Retrieval. Information Storage and Retrieval **7,**1, (1971) 19-37.
48. Borlund, P. Experimental Components for the Evaluation of Interactive Inormation Retrieval Systems. Journal of Documentation **56**(1) (2000) 71-90.
49. Reid, J. A New, Task-Orientated Paradigm for Information Retrieval: Implications for Evaluation of Information Retrieval Systems. In: Aparac, T., Saracevic, T., Ingwersen, P., Vakkari, P. (eds.): Proceedings from the Third International Conference on Conceptions of Library and Information Science, CoLIS 3: Digital Libraries: Interdisciplinary Concepts, Challenges and Opportunities. University of Zagreb, Croatia (1999) 97-108.
50. Beaulieu, M. Experiments with Interfaces to Support Query Expansion. Journal of Documentation **53** (1997) 8-19.

51. Campbell, I. Interactive Evaluation of the Ostensive Model Using a New Yest Collection of Images with Multiple Relevance Assessments. Information Retrieval **2** (2000) 87-114.
52. Ingwersen, P. Cognitive Information Retrieval. In: Williams, M. (ed.): Annual Review of Information Science and Technology **34** (2000) In press.