Information search activity: an overview

Draft version of a paper accepted by *European Review of Applied Psychology*

Jérôme DINET, Aline CHEVALIER and André TRICOT

Abstract. This article exposes a brief history of information search models, from general and behavioural models to cognitive ones and presents recent theoretical advances in this research field. Based on these models, information search can be considered as a complex cognitive activity that involves various individual, contextual and social factors. Then, these factors and their impact on this activity are presented. We also suggest research issues to be developed in order to increase our knowledge in this area and we plead for further research on specific concepts depending on future and/or recent areas of interest.

Key-words: information searching, models, cognitive and affective factors.

Résumé. Cet article présente un bref historique des modèles de la recherche d’information ainsi que les avancées récentes dans ce domaine de recherche. Dans un premier temps, les modèles généraux sont présentés pour parvenir aux modèles cognitifs de cette activité complexe qu’est la recherche d’information. Dans un second temps, différents facteurs individuels, contextuels et sociaux influençant cette activité et devant être considérés par les recherches sont également exposés. Pour conclure, des pistes pour les recherches à venir sont présentées en insistant particulièrement sur certains concepts qu’il convient d’approfondir par de nouvelles études.

Mots-clés : recherche d’information, modèles, facteurs cognitifs et conatifs.
I. INTRODUCTION: A BRIEF HISTORY

People seek out and use information constantly as part of their daily life. Information related to work, leisure, health, money, family, and a host of other topics, is sought from a huge range of sources. Increasingly, some of these sources are digital. The information field continues to evolve rapidly as digital technology changes the nature of information and how people interact and learn from/with information issued from the Web. And because the Web assumes greater roles in education and learning, the past 20 years have seen the growth of scientific research devoted to human behaviors and mental processes involved in information search, especially in two scientific fields: human-computer interaction (HCI) and information retrieval (IR). Whether you are looking for your very first job, switching careers, re-entering the job market after an extended absence, choosing between different candidates for a vote, obtaining information about a product you need, or just preparing your upcoming holidays, you have to search for information within numerous documents disseminated through an increasing number of websites; we therefore must develop new information searching skills so as to read them and exchange them, to become information literate, information literacy being the ability to recognize the extent and nature of an information need, then to locate and evaluate the needed information in diverse contexts (Probert, 2009).

Several disciplines are concerned with understanding how people search and make use of information, the channels they employ to gain access to information, and the factors that inhibit or encourage information searching behavior (Wilson, 1997, 1999). “Information searching behavior is the ‘micro-level’ of behavior employed by the searcher in interacting with information systems of all kinds” (Wilson, 2000, p.49).

An increasing number of researchers in information science and in cognitive psychology have recently begun to explore the behaviours and the cognitive processes involved in information search activity. The main purpose of this paper is to present recent theoretical advances in this
Cognitive psychologists study information seeking by probing the thoughts of people conducting information searches, simulating presumed cognitive functions with computer algorithms, and observing information behaviors in specific environments (e.g., offices, schools, libraries).

IR corresponds to the scientific area of study concerned with searching for documents, for information within documents, as well as that of searching into structured databases (such as digital libraries) or the Web. IR is interdisciplinary, mainly based on computer science, information science and cognitive psychology. Initially, one of the goals of IR was to create systems (such as Online Public Access Catalogs or OPAC in the libraries) that could provide access to books, journals and other documents (e.g., CD-Roms, DVD). Today, Web search engines are the most visible IR applications and studies are often interested in database structure, search language, and indexing rules.

In parallel, the fields of computer science and psychology adopted the information-processing model of cognition and developed a new field known as human–computer interaction (HCI), this field includes predictive models of simple human behaviours (Card, Moran & Newell, 1983; Lazar, Feng & Hockheimer, 2010; Sears & Jacko, 2007). HCI is the "discipline concerned with the design, evaluation and implementation of interactive computing systems for human use" (definition from the ACM SIGCHI Curricula for Human-Computer Interaction: http://hcibib.org/; retrieved September 2011). Information scientists, psychologists and human factor specialists, who sought to understand information seeking and develop better information systems, quickly adopted HCI design methods and usability evaluation techniques. As IR became increasingly important in all areas, some researchers included the human as part of the information retrieval process. They shifted the focus directly onto the human–information seeker rather than on the information resource (e.g.,
Bates, 1989; Belkin, Oddy, & Brooks, 1982; Saracevic & Kantor, 1988).

In these two major research disciplines, human-centred information seeking has become an important approach for studying information phenomena. Marchionini (2006) argues that the intersection of these two communities from HCI and IR defines an emerging human–computer information retrieval (HCIR) research domain. While information scientists have been developing new and more general theories of human information seeking and information behaviour, the broader cognitive sciences have undergone radical shifts.

Today, we take a more holistic view of humans as situated actors within an environment which strongly influences their thoughts, processes and behaviors. This environment includes one’s body, the physical, social and digital environments, such as in the situated cognition theories (e.g., Brown, Collins, & Duguid, 1989), distributed cognition (Hutchins, 1995), and, more generally, social aspects of human behaviour. This change concerning human cognition is beginning to change the way researchers think about how people seek and use information. For instance, it is reflected in the interest in collaborative search and social networking systems. These developments in cognitive science have strongly influenced scholars in the information field who have shaped models of human information behaviour (e.g., Fidel & Pejtersen, 2004; Ingwersen & Järvelin, 2005; Kuhlthau, 1988; Marchionini, 1995). The impact that these developments in cognitive science have had in information schools is reflected in the study of collaborative information search and in the relationships among people (Dinet, 2008). Although a computerized information search (e.g., by using the Web and search engines) is typically a lonely activity, it is increasingly clear that information search tasks take place in a global setting of people and resources unlimited by space or time.

In the same way, owing to the fact that the populations of developed countries count more and more older people while computer use is increasingly affecting all aspects of life
(Billip, 2001; Czaja et al., 2001; DeOllos & Morris, 2000; Ellis & Allaire, 1999; Mellor, Firth & Moore, 2008) and because various cognitive and metacognitive abilities related to information search processes decline with aging (e.g., attention, working memory, information processing speed; Czaja, Sharit, Nair & Rubert, 1998; Hawthorn, 2000; Selwyn, Gorard, Furlong & Madden, 2003), it is increasingly important to understand strategies used by elderly end-users in their search for information through computers and to understand the impact of decline on their strategies.

Be as it may, we jumped progressively from a HCI point of view to a Human-Information interaction point of view. As Knight and Spink (2008) noted, the historical context of the major information behaviour model developments is closely aligned with two on-line technology revolutions. The first involves the creation of early online IR systems, used by “information professionals” (i.e., experts) who usually search on behalf of the person who will ultimately use the information found. The second major development has been the advent of Web search engines, which have made available to any Web user an immeasurable amount of information, with its own unique set of information characteristics. Research into IR, interactive IR and the resulting development of information behaviour models has reflected this dramatic shift in both the end-user/searcher and the information environment.

Because the main purpose of this article is to present recent theoretical advances in the information search field, we expose several models of information search (from information science oriented models to cognitive models; Section II) before discussing various factors involved in information searching (Section III) and, in conclusion, we present further research issues to be developed.

II. FROM INFORMATION SCIENCE ORIENTED MODELS TO COGNITIVE MODELS
For several years now, many theoretical models have tried to describe and predict the human behaviors involved in the information search tasks, and specifically when this activity is
performed in computer systems. Some of these models elaborated by psychologists have attempted to infer the cognitive processes involved. Some of them have become a very popular means of representing complex cognitive processes by trying to capture their essential elements and the relationships between them, usually graphically (Johnson, 2009). In general, theorizing in information behaviour has been at the modelling stage (Case, 2007) with few associated empirical and experimental tests, although models are often inductively derived from grounded theoretic work (Wilson, 1999). These models have limitations because they may oversimplify cognitive processes and they may lead people to over generalize. According to some authors (e.g., Case, 2007; Dinet & Tricot, 2008; Johnson, 2009), most of these models describe information search as sequential actions rather than psychological processes. It is for this reason that in the next section, we distinguish the information science oriented models from the cognitive models: the first models investigate information search behaviour from a generalist view while the second investigate more specifically cognitive processes involved during information search activity. Among all the models, we chose only those that had significantly contributed through the study of specific approaches or dimensions.

II.1. Information Science oriented models

Historically, studies conducted by Guthrie and colleagues (Guthrie, 1988; Guthrie & Dreher, 1990; Guthrie & Mosenthal, 1987) are often considered as the first ones to investigate human information behavior, in “traditional” environments (e.g., textbooks, notices). These studies incorporated the following four components: (a) goal formation: identifying the goal of the search task; (b) text selection: selection of an appropriate text; (c) information extraction and integration: integrating the extracted information with prior knowledge; and (d) evaluation: recycling with monitoring and evaluation of progress towards the goal. The individual repeats this cycle until s/he reaches the information targeted.

II.1.1. The Marchionini’s model (1995)
Marchionini (1995) describes information search modes as directed, semi-directed and undirected (i.e., browsing). In directed searching, the searcher has a specific topic in mind and searching behaviour follows a predetermined path. In semi-directed searching, the searcher has an approximate idea about the topic and searches for information to gather information that will enable him/her to become more focused. In undirected searching (i.e., browsing), the searcher has some interest in the topic and searches the resources to get a better view or to find another topic.

Marchionini’s information searching model follows the eight following steps (Figure 1): (1) Recognition and acceptance of a problem; (2) Definition and understanding of the problem; (3) Choice of a search system; (4) Formulation of the query; (5) Execution of the search; (6) Examination of results; (7) Extraction of useful information; and (8) Reflection and iteration as necessary or conclusion of the search.

![Figure 1. The eight search steps of the search process according to Marchionini (1995)](image)

According to Marchionini’s model, the paradigm of information searching is a natural and necessary mechanism of human existence (Marchionini 1995). Marchionini defines
information searching fundamentally as an interactive process within an information environment (Knight & Spink, 2008). Indeed, understanding the information environment is as important as understanding the searchers psychological processes, as it is the interaction between the two that establishes and reveals the actual information seeking strategies of the user. The key difference between Marchionini’s information seeking context and the previous information seeking contexts is that Marchionini investigates specifically the human behaviours involved in electronic/digital environments.

II.1.2. The Information Search Process (ISP) model

ISP model was first published in 1985 but has since been greatly enriched (Kuhlthau, 1988, 1999a 1999b; Shannon, 2002). Initially, Kuhlthau (1988, 1999a, 1999b) developed a holistic model of the information searching process that takes into account the emotion, cognition and physical experience at different stages of the information search (Figure 2).

<table>
<thead>
<tr>
<th>Model of the Information Search Process</th>
<th>Initiation</th>
<th>Selection</th>
<th>Exploration</th>
<th>Formulation</th>
<th>Collection</th>
<th>Presentation</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feelings (Affective)</td>
<td>Uncertainty</td>
<td>Optimism</td>
<td>Confusion</td>
<td>Clarity</td>
<td>Sense of direction / Confidence</td>
<td>Satisfaction or Disappointment</td>
<td>Sense of accomplishment</td>
</tr>
<tr>
<td>Thoughts (Cognitive)</td>
<td>vague</td>
<td>focused</td>
<td></td>
<td></td>
<td>increased interest</td>
<td>increased self-awareness</td>
<td></td>
</tr>
<tr>
<td>Actions (Physical)</td>
<td>seeking</td>
<td>relevant</td>
<td>information</td>
<td>seeking</td>
<td>pertinent information</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. The six stages model of ISP according to Kuhlthau (2004)

Her model was motivated by an interest in investigating why students behaved the way they did. Kuhlthau (1999a) wanted to address a recognized gap between information systems and users’ natural process of information seeking to improve information system design and guide the encountering process made by intermediaries (librarians and professionals). Her model has been tested in two longitudinal studies and has further been verified in two large-scale studies (Kuhlthau, 1999a, 1999b). Accordingly, the ISP can be
divided into six (Kuhlthau, 1988) and later on into seven stages (Kuhlthau, 1999a), which
differentiate information searched for, ways of searching and relevance assessments, while
moving the seeker from the initial state of recognized information needed to the goal state of
solving. According to Kuhlthau, information search consists of seven stages, each involving
components from one or more of three domains (emotions, cognition, actions): (1) Initiation
(recognize information need); (2) Selection (identify general topic); (3) Exploration
(investigate information in general; (4) Formulation (formulate focus); (5) Collection (gather
information pertaining to focus); and (6) Presentation (answer). Lastly, (7) the individual can
check or revise information search activity.

One of the main features of the Kuhlthau’s model (1988) is the integration of the
cognitive, emotional and physical activity related to information search. Over the last two
decades (Kuhlthau, 1999b), strategies used by the end-user to carry out each of the seven
stages have been introduced in the model and added in other recent theoretical models (e.g.,
Ellis, 1989; Hyldegard, 2009). Kuhlthau’s primary focus was the affective and cognitive
processes experienced by information seekers. For instance, through their actions, people
search for information relevant to a general topic in the beginning stages of the search process
and then pertinent to the focused topic toward closure. Formulation of a focus or a personal
perspective of the topic is a pivotal point in the search process (Kuhlthau, 2004). At that
point, feelings shift from uncertain to confident, thoughts change from vague to more clear
and interest increases.

II.2. Cognitive models

If models issued from information science give a large amount of interesting information
about information search behaviours, cognitive models investigate in particular mental
processes involved to explain difficulties and to predict performances.

II.2.1. The precursors
At the beginning of the 1990’s, Ingwersen proposed that information behaviour results from a series of dynamic interactive processes, which occur at multiple levels within the “cognitive space” of the user and the “information space” of digital systems. By using these multiple representations for information behaviour, Ingwersen (1992, 1996) was able to model an interactive process, not only between a user and the electronic system, but also between the user and the information objects within the system (Figure 3). The model is primarily focused on understanding the actual information system being used, and the interactive cognitive processes that occur between the user and the system when s/he retrieves and finally uses information (Knight & Spink, 2008).

Figure 3. Ingwersen’s model of the information search process (1996)
At about the same time, Saracevic (1996) elaborated the “stratified interactive model” of information behaviour, which is part of a more general model of information use: acquisition – cognition – application. His interactive model is based on the assumption that users interact with computer systems in order to use information i.e. to apply the information acquired through a cognitive process. The interactive model proposed by Saracevic (1996) is simplified to three levels (i.e., strata): surface, or the level of interaction between the end-user and the system interface; cognition, or the level of interaction with the content of the texts
and/or their representation; and the situation, or the context that provides the initial problem at hand. His model shows a strong resemblance to that of Ingwersen (1996) with three strata of interaction (Figure 4): at the surface level, a user interacts with a system through an interface by issuing commands or queries; at the cognitive level, the user interacts with the output of the system (e.g., texts, images); and at the situational level, “users interact with the given situation or problem-at-hand which produced the information need and resulting question” (Saracevic, 1996, p. 203).

Figure 4. Saracevic’s model of stratified interaction (1996)

II.2.2. The Information Seek Cycle model

The Information Seek Cycle (ISC) model was developed by David, Song, Hayes and Fredin (2007). Their model is cyclical and consists of three stages (Figure 5): (1) Preparation: the preparation stage starts when the end-user has to make choices from a menu of links; (2) Exploration: in this stage, the end-user navigates and explores the results of the choices and processes the information displayed; (3) Consolidation: after exploring and processing information, the end-user consolidates by evaluating the results against the goals set during the preparation stage. The outcomes of the evaluation have a direct impact on the recalibration/adaptation of goals that are carried into the preparation phase of the next cycle (Figure 5).
Figure 5. The cyclic nature of information seeking according to David et al. (2007)

Previous studies have shown that the users’ performances could be, at least partially, predicted by the difficulty level of the goals to be achieved (Locke & Latham, 2002). According to the ISC model, other psychological processes are involved, such as the intrinsic motivation or self-efficacy. For instance, Compeau and Higgins (1995) showed that computer self-efficacy is particularly relevant within the ISC because it has both a “looking-back component” derived from experiences from past cycles and a “looking ahead” component that influences the quality of effort and task persistence in future cycles. In information searching on the Web, the paper written by Rodon and Meyer (2011) in this special issue concerns specifically these impacts of Web self-efficacy on the performances of end-users.

II.2.3. The IPS and IPS-I Models of Brand-Gruwel

Based on the Big6™ model (Eisenberg & Berkowitz, 1990; Eisenberg & Johnson, 2002), Brand-Gruwel and colleagues distinguished six stages in information searching (Brand-Gruwel, Wopereis, & Vermetten, 2005; Walraven, Brand-Gruwel, & Boshuizen, 2008; Figure 6): (1) task definition, (2) information-seeking strategies, (3) location and access, (4) use of information, (5) synthesis, and (6) evaluation. Based on various concerns (see Boekhorst, 2000; MacKenzie, 1994), this model has been adjusted and transformed into a descriptive model. The most important modification in the model is the addition of a regulation category instead of that of evaluation.
Figure 6. The Big Six Skills and their relationships with information search activities according to Brand-Gruwel and colleagues (2005, 2008)

Based on experimental studies involving undergraduate and Ph.D students while they searched for information to write a short text on a specific topic (Brand-Gruwel et al., 2005; Walraven et al., 2008), the cited authors distinguished five dimensions in their model: (1) The information-problem definition is always executed at the beginning of the information searching process; (2) The information seeking: the main dimension is the selection of relevant sources; (3) The information consideration consists in identifying and assessing the information about its quality and relevance in order to determine if it is linked with the information-problem or not; (4) The information processing refers to a deep process. The reading, content construction and evaluation of information relevance are involved. The analysis, selection and information structuring are very important in the content construction; (5) The organization and layout of information refers to the main task to be performed.

Throughout these five dimensions, regulation activities are constantly involved. The individuals have to control their execution, to guide the task to be performed, to manage the time, to test the content quality as well as to evaluate the product and the process. This model
is interesting because (1) it involves cognitive and meta-cognitive processes, and (2) it describes in precise terms skills involved in information seeking.

Recently, Brand-Gruwel, Wopereis & Walraven (2009) proposed the IPS-I Model (Figure 7), for Information Problem Solving on the Internet. This model combines five constituents (i.e., five steps: definition of a problem, information searching, information scanning, information processing and organization) and four regulation activities carried out during the entire IPS-process: orientation, monitoring, steering, and evaluating. From a cognitive point of view, these four regulation activities correspond to metacognitive processes (to plan, to modify strategy if needed, to monitor his/her own activity). Moreover, during the performance of an IPS-task and the search for information on the Web, three skills are determining factors: reading skills (e.g., to read hypertext successful), evaluating skills (e.g., to assess credibility of a source), and computer skills (e.g., typing and navigation skills).

![Figure 7. The IPS-I Model from Brand-Gruwel, Wopereis & Walraven (2009)](image)

III.2.4. The Comprehension-based Linked model of Deliberate Search model

Kitajima, Blackmon and Polson (2000) have developed a model of cognitive processes involved in the information search activity: The “Comprehension-based Linked model of Deliberate Search” or CoLiDes model. CoLiDes has been improved many times by the authors (e.g., Blackmon, Kitajima & Polson, 2005; Blackmon, Kitajima, Polson & Lewis, 2002).
In this model, the actions that an individual can achieve from a web page concern two types of processes in parallel (Figure 8): attention processes and selection processes of action. Each of these two processes occurs at several periods of activity: (1) when formulating the goal: on basis of internal or external question, the individual develops a mental representation of the goal; (2) when analyzing the Web page: back to the web page, the individual carries out a visual analysis of the structure and content of that page, distinguishing the categories of information according to their location, typography, etc.; (3) when segmenting in areas of content: based on domain and procedural knowledge acquired through explicit learning or experience, the end-user segments the page into significant areas on the basis that information in the same area is related semantically; (4) when analyzing the content areas: once an area is identified, the user/reader will focus her/his attention on its content and decide to activate knowledge from memory in order to understand; (5) upon the understanding of the content on the basis of information contained in the considered area, the user/reader constructs a new mental model. Then s/he compares the mental representation of these contents with her/his initial representation of the goal: if the semantic relatedness between these two mental representations (i.e., mental representation of the goal and mental representation of the webpage content) is high, then the user/reader considers that the information contained in this area will enable her/him to achieve her/his objective; (6) on basis of the understanding of the content related to the link, the user decides to select the one that seems best suited to achieve his/her goal; (7) in the selection and implementation of the action: this phase corresponds to the effective implementation of behaviour, i.e., click on the link.
As with other models described above, CoLiDes assumes that the information search process involves top-down and bottom-up processes. The user/reader uses the perceptual characteristics extracted from the web page (e.g., title, cues, typography, etc.) as well as domain and procedural knowledge stored in long-term memory about these perceptual (visual) characteristics and the Web in general to navigate. More recently, Van Oostendorp et al. (2011), in this special issue, developed a new cognitive model of web-navigation taking into account the semantic information from graphical elements present on a web page to compute the information scent value of the hyperlinks, on the basis of CoLiDes.

II.2.5. Searching for information with a search engine tool

In addition to the general cognitive models of information search systems that we have previously presented, other authors have developed new models specifically related to search engines (e.g., Pak & Price, 2008; Stronge, Rogers, & Fisk, 2006). For instance, influenced primarily by work carried out by Sutcliffe and Ennis (1998) and Marchionini (1995), as well as by several other researchers including Borgman (1986; 1999) and Kulthau (1999a, 2004), Sharit, Hernandez, Czaja and Pirolli (2008) have elaborated a new model to describe the influence of various domains of Internet-related knowledge and cognitive abilities on Internet
information-seeking performance. This model seems particularly relevant to understand and explain age-related differences (Figure 9).

Figure 9. A model of search engine information-searching behaviour (Sharit et al., 2008)

These models suggest that information searching should be considered as a cyclical activity (like models presented in previous sections) within cognitive processes and that different abilities (such as working memory, spatial ability, reasoning, flexibility) are involved at different stages. Among these models, the model of Sharit et al. (2008) seems to be the most precise. These authors consider information searching to be related to problem-solving and decision-making activities whereby the problem-solver’s knowledge and other mental representations are manipulated in order to achieve a goal. In their model, the problem-solving process is divided into the three following sub-processes:

- The representation of the problem to be solved, in which the statement of the problem is internalized through the construction of a mental representation of the facts to be searched;
- The planning consists of generating a method in order to come up with a solution. It often requires dividing the problem into sub-goals;
The execution involves carrying out the operations that were developed during the planning process.

These processes are iterative as planning may generate further insights into the problem and thus promote modified problem representations. Moreover, in their model, Sharit et al. present the possible impacts of domain and technical system knowledge as well as the possible role of cognitive abilities during navigation and evaluation activities related to the information search.

This model highlights the role of cognitive abilities and knowledge, but other factors (not considered in the current model) are also involved, such as social and emotional ones. We develop these points in the next section.

III. DIMENSIONS AND FACTORS INVOLVED IN INFORMATION SEARCHING

In accordance with Lazonder and Rouet (2005), three sets of variables can affect an individual’s behaviour during an information search task: contextual variables (e.g., time, individual vs. collaborative, material and equipment), resource variables (e.g., interface design, number of websites) and individual variables (e.g., level of domain knowledge, procedural knowledge, language skills). In addition to this study, we propose research issues that should be developed in order to increase our knowledge in this area and stress, in particular, the need for further research on five specific concepts depending on future and/or recent areas of interest.

III.1. The social dimension

As previously shown, information searching is an important and integrated part of problem-solving because it affects the quality of the solution to a specific task. While various models of information seeking have been suggested (Wilson, 1999; for a review, see Dinet & Tricot, 2008), the majority implicitly assumes that the information searcher is an individual. For example, the cognitive point of view in information science has commonly focused on
attributes of the individual in order to understand the cognitive and emotional motivations for information behaviour. But several researchers in information science and in cognitive psychology have recently begun to explore the social and contextual dimensions of information behaviour and information seeking, thereby challenging the assumption of the individual information seeker (e.g., Hertzum, 2002; Hyldegard, 2006; Karamuftuoglu, 1998; Lazonder, 2005). From our point of view, if collaborative information search activity is inevitably performed in a social context (and numerous studies have investigated this issue), the social dimension is not necessarily collaborative. For instance, even if teachers continually encourage the pupils to search for information by using technologies (ICT) in classrooms, “to search” is often stressed and an individualized way of working dominates (Jedeskog & Nissen, 2004).

As Hyldegard (2006) suggested, previous works on Collaborative Information Behaviour (CIB) is characterized by the importance of social factors in acquiring, retrieving, seeking, managing, sharing and generating information. Even though participants are always students or adults, some interesting results have been obtained in relation to the following question: Did group members behave differently from individuals? For instance, in a recent study conducted among students, Lazonder (2005) investigated the impacts of collaboration on web search behaviour and search outcomes by comparing pairs of students and individuals. The results mainly showed that pairs used a wider range of searching strategies than individuals did, were more proficient in monitoring and evaluating their searching behaviours, were faster, and produced a greater number of correct responses to the searching tasks.

Nevertheless, even though it is implicitly assumed — in most cognitive models of information search — that the information seeker is an individual, it is commonly acknowledged that individuals (at schools, at work, etc.) often work in groups or teams that may affect their behaviours as well as their thoughts and feelings (Hyldegard, 2006).
In the same way, in examining the behaviour of scientists in acquiring information relevant to research and development, Sheen (1992, p. 28) notes that: “...some technologists effectively manage to draw a boundary round their expertise in order to protect their position and status within a firm: external information sources are utilised but then internalised and used to develop a personal power base.” In this current special issue, the paper written by Bétrancourt et al. (2011) develops this social dimension in the information search activity.

III.2. The emotional/affective factors

Even if the role of affective factors on information search performances was stressed by some authors towards the middle of the 1990s (Nahl & Tenopir, 1996), two main components were proposed for information search process models (e.g., Sutcliffe & Ennis, 1998): activities as high level components that describe the information searching task in general terms (physical actions and/or cognitive actions) and strategies considered as plans or methods to deal with specific aspects of the information searching problem. In other words, because one of the major contributions of computer science and cognitive science to psychology has been the information processing model of human thinking (with the metaphor of brain-as-computer), emotions were absent. But emotions, such as anxiety, frustration, uncertainty, play a crucial role in the context of online searching. For instance, some studies have investigated the relationship between anxiety and learners’ performances in information search tasks (e.g., Kracker & Wang, 2002), between consumers’ satisfaction and their behaviors on different websites (Erevelles, Srinivasan & Rangel, 2003; Morris, Woo, Geason & Kim, 2002) and between patients’ fears and their behaviors on websites related to health (Matthews, Sellergren, Manfredi & Williams, 2002). In other words, since the early 2000s, numerous studies have investigated causes and/or effects of various emotions during the Web search (e.g., Arapakis, Konstas & Jose, 2009; Arapakis, Jose & Gray, 2008; Bilal & Kirby, 2002;

More recently, evidence from neuro-physiological studies have highlighted the importance of emotions in Human – Computer Interaction (Picard, 2001, 2003). Yet, the disciplines that study Human – Computer Interaction (HCI) and Information Retrieval (IR) have only recently started to investigate this phenomenon and gain understanding of its causes and effects by collecting objective data (Ethier, Hadaya, Talbot & Cadieux, 2008; Julien, 2007; Julien, McKechnie & Hart, 2005; Picard, 2003; Picard & Klein, 2002; Nahl & Bilal, 2007; Lopatovska & Arapakis, 2010).

Therefore, understanding affective reactions is essential to improve Human-Computer Interaction and the design of systems (Dufresne, Prom Tep, Sénécal & Courtemanche, 2010).

Today, three kinds of methods exist to investigate emotions during Human – Computer Information Retrieval (HCIR) activities (Lopatovska & Arapakis, 2010): (i) observation methods (i.e., the observation of facial, vocal and gesture cues as emotional stimuli), (ii) self-report methods (i.e., asking participants to describe their own emotions), and (iii) physiological signal processing methods. From a methodological point of view, physiological data provide interesting information about emotions during HCIR. For instance, Galvanic Skin Response (GSR), the heart rate and blood volume pressure (BVP) are relevant to identify end-users’ emotional state such as frustration or stress (e.g., Dufresne, Courtemanche, & Prom Tep, 2010; Picard, 1997).

In the same way, because one of the main differences between collaborative and individual information search is the role of affective factors, several empirical studies (e.g., Hmelo-Silver, 2003; Hmelo-Silver & Bromme, 2007; Howes, 2008; Issroff & del Soldato, 1996; Jones & Issroff, 2005; Jones & Pellegrini, 1996; Kumpulainen, 1996; Vass, 2002; Wenger, 1998) have investigated the affective factors involved in side-by-side collaborative
settings between partners. Nevertheless, recent literature shows that little is yet known about the quality and processes of pair interaction in collaborative information searching contexts (for exception, see Hyun, 2005; Large, Beheshti and Rahman, 2002). Thus, information about the emotions during users’ interactions in a context of collaborative search for information is, hence, necessary in order to widen our knowledge of users’ behaviours in the collaborative context where emotions are crucial (Maldonado, Lee, Brave, Nass, Nakajima, Yamada, Iwamura, et al., 2005).

III.3. The expertise effects

The effects of expertise in information search have been studied mainly in regard to domain expertise and the information search itself.

The domain expertise refers to the knowledge about the topic of research (for instance, a psychologist who searches for information linked to psychology). Experts plan and prepare their activities, they are more effective and are faster than novices (e.g., Marchionini et al. 1995; Brand-Gruwell et al. 2005; Navarro-Prieto et al. 1999). Wildemuth (2004) describes experts who gave a conceptual specification by gradually adding in additional concepts during their information search. Vibert and colleagues (2009) raise the question of transfer of expertise from one domain to another. Their results showed that when researchers in biology were able to search effectively Medline references in neuroscience (compared to researchers in neuroscience), these biologists, non-domain experts, were as effective as researchers in neuroscience, domain specialists. But, interestingly, the two groups did not proceed in the same way: non-experts spent more time reading the record (i.e., bibliographic references) and consulted more abstracts while experts spent more time reading documents.

Expertise in information search can be exercised in all fields of knowledge. Borgman (1986, 1999) emphasizes the conceptual aspects of this expertise that is not to be confused with semantic and lexical knowledge useful for formulating query. Moreover, technical skills
related to computers are not necessary for their implementation. She uses the notion of mental model to understand the contours of this expertise that she defined as the ability to construct a mental model in an information space, a key factor in predicting the success of a search.

Fields et al. (2004) examined the strategies of research librarians in the digital library of their society and those of their "clients", scientists and engineers, domain experts. The documents are organized and multiple accesses (by keywords, or access full-text, categories and links) are available. Librarians have to perform several activities:

- They reformulate their queries several times before getting satisfactory results that show a mastery of syntax of the query language and resources contained in the data bank. Beyond this mastery, what characterizes their activity is that they seek to develop their knowledge of key words and terms which seems to have a "discriminating power". To each stage of research, librarians pay attention to the size of the result set, enlarging or reducing it. They showed how they could assess their lack of domain knowledge or their complaints and address the quality of their queries.

- They organize their activities and this ability seems essential. Other than their ease in the formulation of requests, how they organized these queries "strategically" qualified expertise. In fact, librarians do not seem to plan or anticipate their queries, they adapt their query formulation with regard to the evolution of the search information context. Their formulation strategies depend not only on each displayed result but also on the entire process formulation, the history of interaction. The approach is entirely exploratory, the librarians did not seek to get optimal response from the first interaction. By comparison, novices formulate fewer queries, often with too many words, their restatements are based on simple strategies such as the use of synonyms, and they leave web pages quickly.

- They have a more precise perception of the size of the results. By comparison, novices have often difficulty to estimate the size (i.e., the number) of potential relevant documents in the
database and on the Internet.

Novices give an impression of impatience, by running many Websites and many search engines, clicking more often and spend little time planning (Tabatabai & Shore, 2005; Chevalier & Kicka, 2006). According to some authors, novices perform more movements in information retrieval (reiteration of requests) and have difficulties in providing the appropriate requests (Wildemuth, 2004) while for others the opposite is true (Ihadjadène, Chaudiron & Martins, 2003). Novices in the field decide earlier-on in the process on the usefulness of documents compared to experts (Brand-Gruwel et al. 2005).

III.4. The impact of aging

An increasing number of elderly people use computers and the Web to search for information, especially health/medical-related information, but also to communicate (Fox & Fallow, 2003; Karavidas et al., 2004; Meischke, Eisenberg, Rowe, & Cagle, 2005). The elderly often retrieve health/medical information from the Web, which used to be available to experts exclusively (Bromme, Jucks, & Runde, 2005; Williamson, 1995, 1997). Today, seniors can have immediate access to a vast amount of health-related Websites. The information they retrieve can help them to understand specific diseases and to increase their therapeutic adherence (Morahan-Martin, 2004). These new end-users have specific needs related to the aging-related declines in cognitive processes.

Over the last decade, many studies have explored computer use and health-related information needs among the elderly adults. Three main types of research can be identified: (a) studies involving interviews and/or focus groups of seniors (e.g., Karavidas, Lim, & Katsikas, 2005), (b) surveys given to organizations that offer information and/or services to senior users (e.g., Mühlhauser & Oser, 2008), and (c) experimental studies in which seniors are observed using computers to find information (e.g., Becker, 2004; Czaja et al., 2001; DeOllos & Morris, 2000; Ellis & Allaire, 1999; Selwyn, Gorard, Furlong, & Madden, 2003).
Despite the fact that these three types of studies provide very interesting information on the effective information needs of the older adult population, two different methodological points have to be considered: the first one concerns the variation in age range. This variation may be mirrored by the ages used in the scientific literature: some studies chose 50, 60 or 65 years as the beginning age. The second one concerns the origins of the small samples conducted that prevent us from generalizing results.

There are considerable social and economic reasons linked to a better understanding of elderly people’s needs (Hawthorn, 2000; Schneider et al., 2008; Zajicek, 2004). For instance, the number of seniors is growing more quickly than that of all the other segments of the population. This will impact on the cost of social care unless technological solutions can be found to enable people to stay longer in their homes. Designers of information systems must take into account the special needs of such a significant population who often find current products difficult and complicated to use (Schneider et al., 2008), but older users often fail to consider and anticipate their navigation difficulties, and the special features of their needs (Newell et al., 2006). Making interfaces that meet the needs of older users is a challenge for information systems designers. This challenge actually requires designers 1) to acquire an in-depth knowledge of older users’ needs, 2) to involve older people in their design process and 3) to adapt usability testing. In response to the first requirement, some researchers have recently studied the strategies and the difficulties that older adults experience while searching for information (see Aula, 2005; Stronge et al., 2006; Sharit et al., 2008; Dommes, Chevalier, & Lia, 2011; Westerman, Davies, Glendon, Stammers & Matthews, 1995), and attempted to synthesise the physical and cognitive factors involved in computer use (see e.g. Hawthorn, 2000; Chevalier et al., 2008; Taveira & Choi, 2009; Westerman & Davies, 2000).

Indeed, aging is often associated with executive functions declines (Salthouse, Atkinson, & Berish, 2003) and more precisely with a decrease of working memory span.
(Zacks, Hasher, & Li, 2000), inhibition failure (Lustig, Hasher, & Tonev, 2001) and decrease of cognitive flexibility (Kray & Eppinger, 2006). Consequently, older adults experience more difficulties than their younger counterparts to search for and find information from information systems even if the performance of older adults can be comparable to that of younger adults if given sufficient additional practice (Westerman & Davies, 2000). For instance, in a recent study comparing young and older web users, cognitive flexibility was shown as strongly involved in the reformulation of unsuccessful requests and its decline with aging affected information search performance (Dommes et al., 2011). More precisely, older adults obtained poorer search performance with fewer correct answers. They interacted little with the system, and showed specific difficulties in getting out of impasses and reformulating unfruitful requests. Regression analyses suggested that age differences in the number of requests were moderated by differences in cognitive flexibility and that the older participants did not leverage their increased vocabulary in producing new keywords.

Education level may also acts as barriers to information access thereby frustrating an information-seeker, specifically for older adult users because the education level has a positive impact on the anxiety: the individuals with a higher education level have a lower anxiety level (Ellis & Allaire, 1999; Karavidas, Lim, & Katsikas, 2005). In addition the stereotype concerning the fact that adults are unable, unwilling, or afraid to use technology still remains (Adler, 2006; Howze & Redman, 1992) even if some studies have shown that older adults have a positive attitude towards computers and technologies (Mitzner et al., 2010). Indeed, consistent with these stereotypes, several large scales usage studies have found that older adults do not use certain technologies to the extent that younger adults do.

Data are scarce, and further studies are still required for a better understanding of the age-related differences observed in searching for web-based information. In this present
special issue, Etcheverry et al. (2011) develop the respective contribution of inhibition and recollection processes to age-related difficulties in the Web search activity.

III.5. Information searching and work/professional context

An increasing number of studies, including those discussed in Section III.1 about the social context, describe the tasks of information seeking in work context. Work context refers to the conditions under which the work is performed and the demands such jobs impose on the worker. They attempt to describe the task taking into account the factors which have contributed and given rise to the need for information and those that will allow the use of information found (e.g., Cool & Belkin, 2002; Ingwersen 2005, Johnson 2003). Li and Belkin (2008) proposed a faceted approach to conceptualizing information search tasks. This work aims to identify, describe and classify a range of information seeking strategies in a group of knowledge-intensive workers. The explicit aim of these authors was to inform the work on the subject of information retrieval by explicating the phases of the information interaction tasks. Their approach is based on the scheme of four different dimensions related to information search: mode (recognition - specification), method (scanning - searching), goal (learning - selecting) and resource (information - meta-information). The faceted approach of Li and Belkin (2008) is based on a fairly exhaustive study of literature and show that existing models have tended to be interested only in the partial aspects of tasks and a faceted approach may provide a more holistic model based, in particular, on the source of the task, user, time, product, process, purpose, characteristics of the task and the perception of the user’s task. Finally, the Li and Belkin model (2008) comprises five major facets (communication behaviour; information behaviour; objects interacted with; common dimension of interaction; interaction criteria) and several sub-facets for each of these facets (e.g., for “information behaviour”: create, make decisions, disseminate, organise and preserve).

We give here an example of information task analysis within a work context. We
choose this example because it illustrates that information task analysis and work task
analysis can be done closely. The context is aviation maintenance, where technicians have to
seek information for each task, even if this task is very well known and even if the technician
is an expert in achieving the given task. Maintenance tasks can be classified into two
categories: preparation and completion. The preparation task takes place before completion of
the task itself and aims at the provision of all that is necessary for the implementation of that
(parts, tools, lubricants, etc.). The secondary tasks of information retrieval can be analyzed
from the categorization of Tricot (2004): obtaining new knowledge; fill a lack of knowledge;
complement existing knowledge; be in conformity with the situation, i.e. search for
information even if you don’t personally need it but because you’re obliged to; have detected
a marker of relevance in the situation (as a warning). Barnard et al. (2007) analyzed the
purposes of use of documentation in aircraft maintenance: follow strictly the prescribed steps
in the procedure, find information on a specific topic, check if anything important has been
forgotten, acquire knowledge about a task or system. These four operator requirements can be
mapped to the tasks of information retrieval of Tricot (2004): the need for information about
how complete is the need for new knowledge or need to comply with situation, the need of
specific information to address a lack of knowledge to run a task is the need to supplement
existing knowledge and the necessary information to verify compliance of the task performed
is the need to be compliance with the situation, the need for the information system or
learning a new task is the need for new knowledge or need to supplement existing knowledge.
IV. CONCLUSION

As developed in this article, information search is a complex cognitive activity, which
concerns various scientific fields in cognitive science. Consequently, many models of
information search have been developed and related to individual differences and attributes of
the system device used. However, further studies have to be done in order to expand on the
results obtained and to specify the role of certain variables such as:

(1) Individual differences, and especially the effect of age on information search activity need additional studies concerning not only the elderly, but children and teenagers as well, who search for information either alone or in groups. It is particularly obvious in our opinion that it is time for researchers to consider children and teenagers as more than novices, because today, they are not (Boubée & Tricot, 2010). Expertise in information searching also needs further studies, because if many authors agree with the fact that it combines different factors and dimensions, how well they interact is unclear, as is that which has the greatest impact. Studies show that technical expertise, while often neglected, is fundamental when one wants to understand the global expertise in information retrieval. Authors like Jenkins, Corritore & Wiedenbeck (2003), Marchionini et al. (1995), Ihadjadène et al. (2003) or Kim (2008) argue in very divergent ways on these issues. It therefore seems necessary to emphasize the need to conduct research in the area to better define the different levels of expertise (from novice to expert), the different types of expertise (that of the regular user, it’s professional IR) and different knowledge (in the field, research information, technical). It will then better describe the relationships between these dimensions expertise, including addressing seriously the compensatory hypothesis, that expertise in a dimension may compensate for the lack of expertise in another, as results of Vibert et al. (2009) seem to show.

(2) Information searching in a specific context: the second key domain which needs further research is the context of information searching and how this context involves this activity. We have briefly presented a few studies touching on the social context such as collaborative information search and work context. These two fields are now recognized as important and must be developed. Moreover, we also posit that the research community needs to bring together two poles: the first one concerns theoretical elaboration at a general level (information retrieval as a cognitive and social activity, linking the concepts of information
needs, the others’ representation of information goals, sharing and regulating collective goals, etc.); the second one concerns a much more precise and specific approach of information searching in context (consider, for example studies on information retrieval in the medical field, in the economic field, in the field of leisure, etc.).

(3) The emotional and unconscious aspects of information searching: we have shown above that the emotional aspects have just started to be taken into account in studies on information searching. It also seems that some non-conscious aspects of this activity deserve more attention. For example, we defined at the beginning of this contribution the information search activity that an individual who lacks knowledge in a given situation implements. But how is this individual aware that he/she lacks knowledge? Before one becomes aware of an information need, what mechanisms are implemented?

In this paper, the framework for human information search behaviour research is based on the premise that human information behaviour and strategy cannot be fully described, or understood, without considering the cognitive, affective and social dimensions. Because human information search behaviour has a multi-dimensions nature, research in human information search behaviour should consider these multiple dimensions from a theoretical and methodological point of views. Therefore, the use of multiple data collection and analysis methods that are complementary (e.g. eye tracking, physiological measures, surveys) may reveal more about human information search behaviour than the use of any single method can reveal. In other words, various methods should not be seen as competitors but as methods that can contribute to the understanding of various aspects of human information search behaviour. Finally, the framework may be useful as a tool in helping researchers formulate research questions and provide some ideas about theories and methods that could be used to collect and to analyse data related to human information search behaviour.
References


