Collaborative Information Retrieval

in an Information-intensive Domain

Preben Hansen[*)] **and Kalervo Järvelin**[+)]

** SICS – Swedish Institute of Computer Science, Box 1263, 164 29 Kista, Sweden.*

Tel.: +46 (0)8 633 15 54, Fax: +46 8 751 72 30, E-mail address: preben@sics.se

+ University of Tampere, Center for Advanced Study, FIN-33014 University of Tampere,

Finland, Tel. : +358 3 215 6953, E-mail address: kalervo.jarvelin@uta.fi

Abstract

In this article we investigate the expressions of collaborative activities within information seeking and retrieval processes (IS&R). Generally, information seeking and retrieval is regarded as an individual and isolated process in IR research. We assume that an IS&R situation is not merely an individual effort, but inherently involves various collaborative activities. We present empirical results from a real-life and information-intensive setting within the patent domain, showing that the patent task performance process involves highly collaborative aspects throughout the stages of the information seeking and retrieval process. Furthermore, we show that these activities may be categorized and related to different stages in an information seeking and retrieval process. Therefore, the assumption that information retrieval performance is purely individual needs to be reconsidered. Finally, we also propose a refined IR framework involving collaborative aspects.

Keywords: Collaborative Information Retrieval; Work-tasks; Information Seeking; Task Performance Process

[*] Tel.:: +46 (0)8 633 15 54   Fax: +46 8 751 72 30
E-mail address: preben@sics.se (Preben Hansen)

1 Introduction

Today, it is essential for professional workers to both stay informed and inform their work environments in order to effectively manage knowledge and stay competitive, effective and innovative. The flow of information (e.g. gathering, assimilation, and creation of information) involves increasingly complex means. It is also well recognized that people act in a social and organisational context together in groups when trying to solve seeking problems (Karamuftuoglu, 1998, Munro *et al.*, 1999; Soininen & Suikola, 2000). Therefore, in order to understand aspects of *collaborative activities* from an information seeking and retrieval (IS&R) perspective, we need to investigate the manifestations of collaboration.

However, one prevailing assumption in information retrieval (IR) is that problem understanding, query formulation and retrieval is basically viewed as an *individual activity* and that the searcher performing the task is in a rather *isolated situation*. Decontextualized approaches to information seeking and retrieval activities cannot but yield narrower findings than investigations involving information access in dynamic and real-life work place environments as mentioned above. Therefore collaborative information handling activities should be related to both information seeking (IS) and information retrieval (IR) processes.

We have investigated the problem of collaboration in the context of the handling of patent applications. As to the knowledge of the authors, there are no prior studies regarding collaboration in relation to the information seeking and retrieval process in the patent domain. The approach in this study was to analyse the problem of collaborative activities and information access through the subtasks of the patent application handling process. In this article we empirically investigate collaborative aspects of IS&R processes in the information-intensive work setting of patent domain. We collect both qualitative and quantitative data on users performing real-life IS&R tasks to identify and classify different types of collaborative activities.

So far, there is no widely accepted definition of collaboration, sometimes also referred to as cooperation. In this study, collaboration is specifically related to information retrieval (IR), and thus, we work from the following broad and preliminary definition of *Collaborative Information Retrieval (CIR)*: *CIR is an information access activity related to a specific problem solving activity that, implicitly or explicitly, involves human beings interacting with other human(s) directly and/or through texts (e.g., documents, notes, figures) as information sources in an work task related information seeking and retrieval process either in a specific workplace setting or in a more open community or environment.* This definition is preliminary, and needs refining through empirical observations and investigations. CIR means active and explicit retrieval of information for solving a specific task. On a general level, it can be said that sharing information is usually about sharing *already* acquired information, while CIR deals with searching *for* information. Sometime these do coincide.

The article is organized as follows. In Chapter 2 we discuss the concept of collaboration and its relation to IS&R and CSCW. We also present some characteristics of collaboration in general which are relevant to our study. Chapter 3 describes the research questions that guided our empirical study and the study design. We propose a conceptual framework for collaborative information seeking and retrieval in a professional information intensive domain. We also present a short overview of the patent handling process. In Chapter 4 we present the empirical results from the study and finally, in Chapter 5, we lay out some conclusions and implications for future research.

2 Collaboration in Information Seeking and Retrieval

2.1 Collaboration in prior IS&R Research

Research in Information Seeking & Retrieval (IS&R) views, in professional settings, the IS&R process embedded in a task performance process (e.g. Allen, 1977; Byström & Järvelin, 1995; Hansen, 1999; Byström & Hansen, 2002). While collaboration is understood as an important feature of IS&R, there actually is very little *empirical* knowledge about collaboration in IS&R processes. An early example is Allen (1977) who studied the differences between the information-seeking behaviour between engineers and scientists. The study provided an understanding and proof that engineers and scientists have different information-seeking behaviour. One of the findings was that these two groups used information sources differently. Allen points out important aspects of the information-seeking behaviour relevant for our study: the importance of personal contacts and discussions between engineers and that there are gatekeepers in organisations. Allen also studied patterns of communication within a small research laboratory and found a typical communication network. These networks showed central points around which communication was centred. In the middle of these networks, key persons were identified – technological gatekeepers (Allen, p. 145). In larger organisations, there are networks of such gatekeepers. Allen proposes to “…organize information dissemination around them [the gatekeeper]…”(Allen, p. 180).

Pinelli et al (1993) discuss the information-seeking behaviour of engineers from within a conceptual framework. In this paper, the engineers are viewed as information processing systems throughout the seeking process. A central concept is that information processing is the difference between information possessed and information required and defined as uncertainty, (p. 189). The conceptual framework assumes that, in response to e.g. a task, specific types of data, information and knowledge are needed. The engineers act, according to the authors, along two alternatives: create the information or search existing information. If an engineer decides to search for information, there are two types of information channels available: informal or collegial networks (oral interpersonal communications with colleagues, gatekeepers and personal collections of information) and formal information systems (libraries, librarians, information specialists and information retrieval system).

O’Day & Jeffries, (1993a) discusses sharing information and sharing search results within group situations at four levels: sharing results with other members of a team; self-initiated broadcasting of interesting information; acting as a consultant and handling search requests made by others; and archiving potentially useful information into group repositories. Other researchers have also pointed out that collaborative activities may occur in earlier stages of IS&R processes (Kuhlthau, 1993; Marchionini, 1995), but did not go deeper into CIR issues.

Recent research in IS&R extends our knowledge on how people access, retrieve and judge information. Karamuftuoglu (1998) proposes an approach, called *social informatics*, which seeks to include the relationships between humans within an IR process. Fidel and colleagues (2000) describe a project focusing on collaborative activities of members of a work-team within an organization when performing IS&R tasks.

Sonnenwald and Pierce (2000) studied information behaviour in a dynamic work context of command and control (C2) at the battalion level in the military. The authors highlighted phenomena they call interwoven situational awareness, which is defined as individual, intragroup and intergroup situational awareness (c.f. awareness in Chapter 2.2). They also found that there was a continuing necessity of information exchange during their work operations. Even though the authors do not explicitly talk about information searching in detail, their findings provide valuable insight into intra- and intergroup communication in order to acquire information needed.

Hansen and Järvelin (2000) described a set of data collection methods used in a real life work setting investigating the information seeking and retrieval processes performed by patent engineers. The study design was done from a task-based approach. One of the main preliminary results in this study was that the patent engineers were involved in different collaborative activities such as collaboration related to internal or external activities and collaborative activities related to individual or group related activities.

Herzum and Pejtersen (2000) report on the importance of providing support for people when searching information systems. Two case studies were conducted involving engineers and the authors found that people searched for documents to find people and searched for people to obtain documents. Furthermore, they interacted socially to acquire information without engaging in any explicit search activity. It is necessary to consider that there is a need to consult people with specific competencies and experiences. These findings provide further knowledge on that people do engage in collaborative information seeking and retrieval activities.

Cooperative and collaborative activities within an organisation often involve information sharing. In a recent paper, Talja (2002) described and classified different types of information sharing. Her study was done through empirical observation and shows that collaborative seeking is a common seeking strategy. The following types of information sharing were identified:

- Strategic sharing
- Paradigmatic sharing
- Directive sharing
- Social sharing
- No sharing

The study clearly shows that collaborative seeking is as natural and common a seeking strategy as individual seeking. However, this study does not really go deeper into the retrieval stages of the seeking and retrieval process.

In Information Retrieval, collaborative activities have been studied and observed among search *intermediaries* (e.g., O’Day & Jeffries, 1993a; O’Day & Jeffries, 1993b; Crabtree & al., 1997). The focus of the present paper however is collaboration within work groups of colleagues – i.e., end-users – and therefore we bypass this area. Empirical studies on collaboration in IR among end-users are scarce indeed.

2.2 Collaboration in CSCW

The motivation for the study of CIR is that most of the IS&R work assumes that the information seeker is an individual rather than acting in a group. In this regard, research in Computer Supported Collaborative Work (CSCW) could provide some understandings about concepts and collaborative aspects of IS&R activities. CSCW deals with collaboration in organisations and work groups, and systems supporting collaboration, such as organizational memory, organizational information handling and information sharing. We do not review the CSCW-literature in whole, but focus on findings relevant for collaboration in IS&R.

In a study made by Harper and Sellen (1995) in the professional works setting of the International Monetary Fund (IMF), it is reported that re-use of documents will most often involve paper documents or “paper-like” equivalents. Even though the study mainly deals with sharing of information, the findings are important. One interesting finding reported is that social interaction is not as important to the sharing of *objective* information as it is to the sharing of *interpreted* information. According to the authors this means that there is information that can stand alone and separate from social processes. Harper and Sellen focus on the sharing of information and do not explicitly deal with the retrieval of information.

In HCI and CSCW we find a large body of literature were attempts are made to facilitate finding information through social networks, (e.g. The Answer Garden by Ackerman & Malone, 1990). Another attempt is made by McDonald & Ackerman, (1998) reporting on the Information Lens. They conducted a five-month field study on how people in a medium-sized organisation find the expertise to construct, maintain and support their software systems. The study deals mainly with how people share information. They distinguish two steps: expertise identification and expertise selection. The identification involves identifying the resources and then to obtain them. The selection involves choosing among people with required skills and expertise.

A professional work setting involving collaborative IS&R activities is a complex and varied environment. In their paper, Romano et. al. (1999) describes how user experiences with IR have informed the development of a prototype of a Collaborative Information Retrieval Environment. The prototype is a system that is dedicated to support collaborative information searching. The paper presents a literature study of both the IR and Group Support System (GSS) which resulted in a list of interesting commonalities to be taken into account such as that both IR systems and GSS's are a single user systems; both have a single user perspective; and that many systems fail to support how both individuals and teams work together.

Hertzum (2000) reports on how information seeking is interwoven into co-operative work and investigates the role of people as information sources during the work of a system design task. The author found that software engineers were looking for practical experience rather than hard facts, and they were also looking for commitments rather than information. These are truly of great importance for system design, however, the author do not go deeper into information retrieval aspects of collaborative activities.

Based on the CSCW literature, relevant cooperative activities between humans may for instance be classified as:

- *asynchronous* or *synchronous* activities
- activities based on *traditional human* communication or *computer-mediated*
- *loosely* or *tightly coupled* activities

Moreover, the *awareness* of people, activities, or objects may vary. These are considered in more detail below.

Traditional *human communication may be asynchronous* or *synchronous* - asynchronous through ordinary mail and book/journal reading; and synchronous through human face-to-face real-time communication and ad-hoc social interactions. *Computer-mediated communication* may also be asynchronous through e-mail, searching the Internet and log viewing, and synchronous through video conferencing (e.g. Erlich & Cash, 1994; Haake *et al.*, 1999).

In *loosely coupled activities*, the system will take advantage of recommendations from other people through observations of their information seeking behaviour such as search paths and annotations; recommendations based on usage rates, and explicitly stated recommendations. *Tightly coupled* activities may in the context of IS&R include sharing queries and strategies for their refinement, and feedback and judgement phases with others. It may also include sharing objects such as reviews (Haake *et al.*, 1999).

Awareness is an unexplored issue in cooperation in IR. In this study we are interested in awareness from three aspects: people, activities, and objects. Awareness of *colleagues* can be divided into either awareness of *individuals* or *groups.* Secondly, awareness of *information* can be investigated by focussing on, e.g., topics, types of sources and information types.

Finally, *information sharing* in professional settings may include sharing the same need for information, search strategies, and sharing retrieved objects (Robertson and Hancock-Beaulieu, 1997; Talja, 2002). Technologies supporting this kind of sharing are CSCW-tools, personalized systems, content-based information filtering (Pemberton, Rodden & Proctor, 2000) and social information filtering, such as the systems Information Lens (Malone et al., 1986), and Answer Garden (Ackerman & Malone, 1996).

3. Study design

3.1 Research Questions

The aim of the present study is to develop our understanding of *collaborative* activities within IS&R processes and to identify what kind of collaborative activities that can be observed in an IS&R process. We assume that an IS&R situation is not merely an individual effort, but inherently involves various collaborative activities. Our specific research questions are:

1. *How* do collaborative activities manifest themselves and *how often* to they occur?
2. *When* do collaborative activities take place in a IS&R process?
3. *What* are the characteristics of collaborative activities?

3.2. Study setting, data and methods

This study is part of a larger study performed at the Swedish Patent and Registration Office[1] (PRV), Stockholm, Sweden. Data collection was performed on-site in a real-life work setting involving patent engineers (PEs) performing their real-life work-tasks. The data was collected in order to discover what PE's actually did during their work processes. We performed on-site observations and applied a set of combined qualitative and quantitative data collection methods. The data were collected during 2 months (of which 5 weeks for the on-site observations). A part of our longitudinal data describes how, what and when collaborative activities manifest themselves. Altogether 9 patent engineers from the PRV voluntarily participated in the study. The patent engineers worked either alone or in pairs in each office. However, as will be shown in this paper, the patent engineers were frequently involved in collaborative activities. The low number of participants is explained by the data collection methodology chosen. All 9 participants were observed performing 12 unique patent work-tasks. All participants performed 1 task, except for participants # 2, 3, and 4, who performed 2 tasks each. Patent work tasks were observed at different stages, which enabled us to cover all the main stages in the information handling process of

[1] PRV, Stockholm, Sweden, http://www.prv.se

each patent application and thus, all stages were represented in our study. Due to time constraints (it takes up to 2 years to complete a patent application) we limited our unit of observation to well-established sub-tasks within the patent handling process. The individual tasks usually took between 1-5 days to complete. During the work task, the investigator had to adapt to normal work procedures, and thus, interruptions, other sudden duties and natural events influenced the observations.

A *pilot study* was initially conducted in order to validate our methodology. After the pilot, a *group introduction* was organized in which information about the components of the project was provided to the participants. In the main study, the following data collection methods were used:

- *Semi-structured and open-ended interviews.* Interviews were performed before and after the main data collection period. The pre-interviews were used to collect data about demographics, experience and knowledge levels and descriptions on search procedures etc. The post-interviews were done to follow-up interesting and valuable issues detected and identified during the main data collection period. All 9 participating patent engineers were interviewed. Each interview took between 1 to 2 hours.

- *Electronic "Diary*". An electronic “diary” was designed for the participants so that they could describe their daily activities. The electronic diary sheet was designed in order to be able to collect different types of data and contained a formal outline involving specific steps to be followed when writing the diary. The outline was based on a) an initial interview with 2 patent experts from within PRV; and b) on the experience of the pilot study. The stages in the diary represented the main and common steps in the patent handling process. This general outline acted as a "reminder" for the participants. The participants were asked to freely add whatever they thought was of interest. The diary also contained empty fields so that logging information[2] could be inserted from database searches. In summary, the diaries were designed to capture the following data: full versions of log histories from database searches; descriptions of problems to be solved and how the PE did solve it; personal comments on the search problems and the work process in whole; time stamps for performing sub-tasks; search terms, search strings, classification codes used and usually a discussion related to them; collaboration with colleagues; handling work tasks within their own department/group etc. Each of all the 9 participants was asked to fill out their own diary. The data were collected during 2 months period. The participants were asked to send back the diaries to the investigator at the end of each working day. This ensured the investigator to do a quick check of the content and if there were any problematic issues or peculiarities, the investigator could immediately go back to the PE and ask for a clarification or complementary information. The diaries were combined with the post-interviews made at the end of the data collection period. Specific issues were drawn from the diaries and discussed during the interview. The diary was handled electronically and the diary sheet was sent to each participant. They were asked to copy the original sheet as many times as necessary during the data collection period. Each participant got an ID-key to be used together with the actual date to mark up each diary to be sent back to the investigator. They were given a web page were they could download their diaries individually. The investigator could then monitor the archive for incoming diaries.

[2] Log information was handled in such a way that it did not reveal any unauthorized information

- *Focus observations.* In parallel to the diaries, continuous on-site visits were performed, observing participants in their work. During observation, a list of key questions was asked and notes were taken. Awareness of unexpected situations and activities was emphasized. The subjects were encouraged to "talk aloud". Each single observation could take up to 7-8 hours per day and for each task, 3-6 daily visits were conducted. The observations were constrained by activities such external/internal courses and confidential internal meeting or other duties for which the investigator did not have any access to. Each task was followed up to 5 days in a total (could be spread out on a 2 week period).

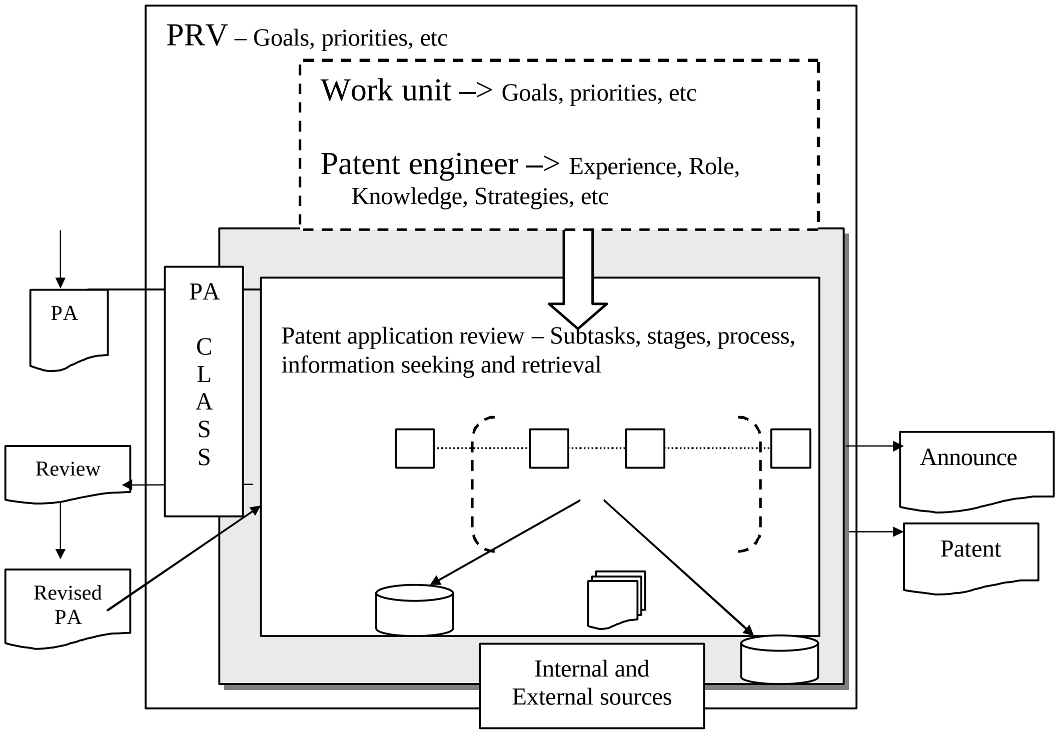
3.3. A short overview of the patent work task performance process

Generally, the task performance process at PRV is formally well structured and involves a certain set of stages (Figure 1). Initially, the patent application (PA) arrives at PRV and is *registered*. The application could be filed as either a national or an international application. The PA then undergoes a first review and a *classification process* and is *assigned* to a department and to a specific patent engineer (PE) with expert knowledge in the area of the PA. The PE then starts *preparations* and *planning* for the task and *reads the application* in order to *identify and define the problem(s)* and claims made. Further information is collected and reviewed and *requirements* for the handling process are decided. The goal is to *identify the information need* in order to solve the patent application task. The *information need formulation* stage involves the process of describing the identified need and specific conditions for further processing. The next stage is to *identify relevant and appropriate sources* to be used.

When an electronic source is selected, an IR task process is initiated. For this purpose, the PE has both internal (in-house created DB: s) and external (commercial DB: s) sources. *Query formulation and reformulation* involve the iterative formulation and reformulation of the identified information need through constructing more or less complex query sequences. The search outcome then undergoes *relevance assessment and judgment.* When a satisfactory set of documents has been retrieved and judged as relevant, the next phase of the process involves the use of these documents. The retrieved documents are used for *writing different reports according to the type of PA.* Depending on the type of PA, the PA may go back to the *applicant for revision.* When the revised version comes back to the patent office, the PA undergoes an *inspection*. Finally, when both the applicant and the patent office have agreed, a *public announcement of patent* is performed and the patent is *filed nationally and internationally*. What is not shown explicitly is that there are large time intervals between the patent handling phases.

The tasks investigated in this study are the IS&R tasks (the shaded box) and start when the patent engineer receives a classified PA and ends with judgements and decisions on whether the retrieved documents should be used for the final report.

Figure 1. Schematic model of the Patent application handling process at PRV


PRV – Goals, priorities, etc
Work unit –> Goals, priorities, etc
Patent engineer –> Experience, Role, Knowledge, Strategies, etc
PA
PA CLASS
Patent application review – Subtasks, stages, process, information seeking and retrieval
Review
Revised PA
Announce
Patent
Internal and External sources

The patent engineers work in different topic areas, such as "Communication technology" and "Optics", etc. In general, the patent engineers handle three different types of applications: application written by a professional patent bureau, written by an internal patent department within a company, and finally applications written by private persons.

The overall goal is to protect investments that individual and companies have made into new technological innovations and developments and to stimulate the competitiveness in Sweden in a just and fair way. The handling of the patent applications, which is done mainly through classification, searching, retrieving, inspecting and judging relevant information within the patent domain, will ensure that the applied invention is treated in a fair way.

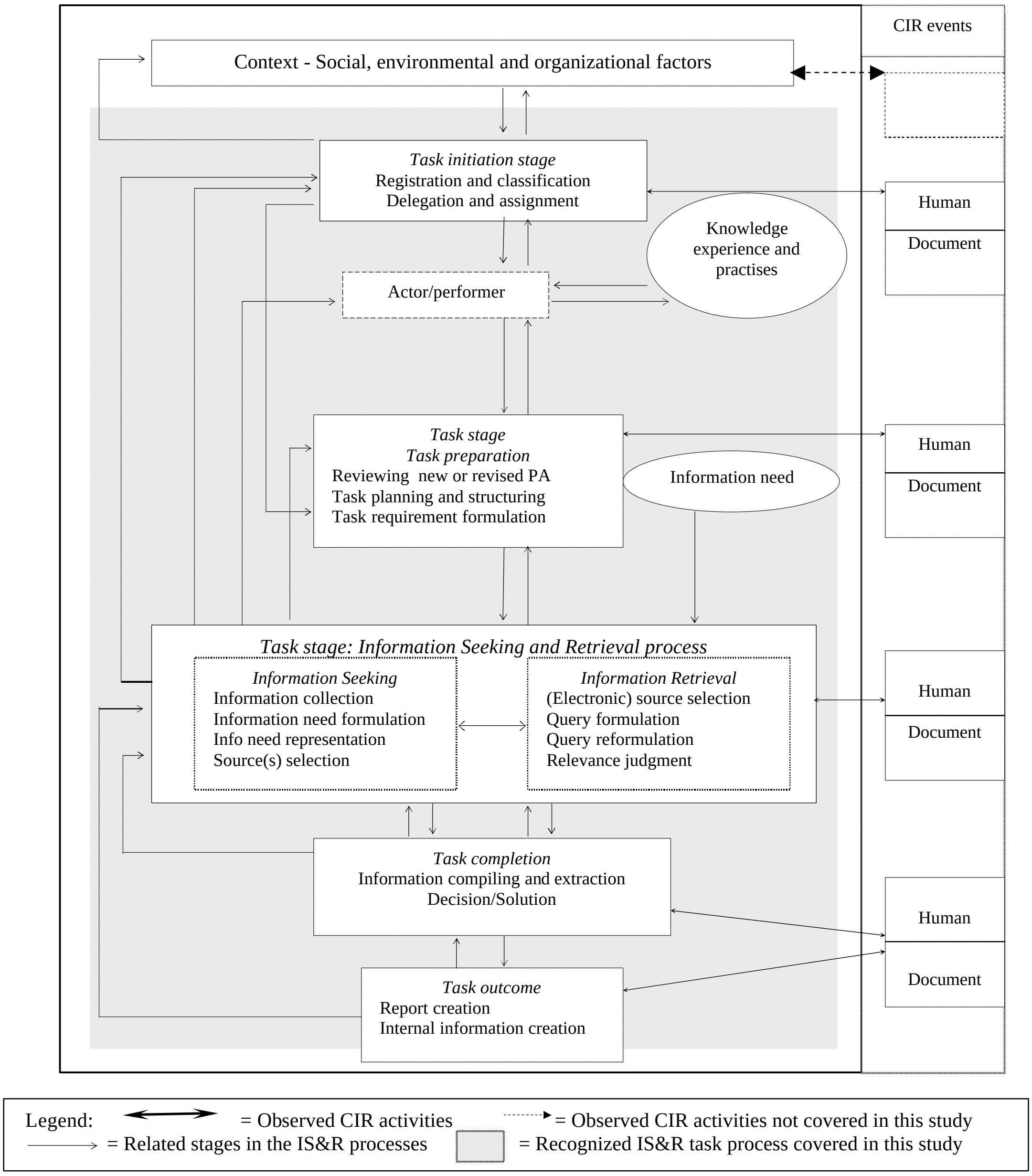
Based on the interview with the patent engineers, the goals of patent work can be described along three different levels: the organisational level, the group level, and the individual level. These three levels cover different aspects. In short, on the organizational level we find issues such as protecting ideas; supporting the development and growth of Swedish industry and further development of patents; disseminating information and knowledge; helping applicants; and providing applicants with high quality patent searches. On the group or team level the following important issues were recorded: providing information, knowledge and protection of applied invention if necessary; processing as many applications as possible; creating and developing praxis and consensus within the work team regarding judgments etc. Finally, the individual level resulted in issues such as to give the application a good qualitative judgment; provide patent searching as a service; to find what is already known and therefore not to accept redundant applications or parts thereof thus identifying patents that really are unique and therefore possible to patent; to give the applicant a good and strong protection for his ideas. Due to the long process time of the patent application, there is a set of constraints that the patent engineers encounter. The most important are (in descending order): Time (too little time to investigate the problem satisfactorily, interruptions (this could be meetings, external phone calls etc), problem solving support (colleagues that need support to solve a specific problem such as a sub-topic outside the engineers own area.); IT connections failure; and finally costs (awareness of searching expensive databases).

The Swedish patent and registration office handles approximately 4000 national patent applications and 5000 international applications annually.

3.4. A Framework for CIR in the patent domain

Our framework for CIR in the patent domain is based on empirical observations and adapted from a general IS&R model. It divides the patent task process into 3 main levels: the Work Task level (WT), which includes stages such as the initiation, preparation and planning of the task as well as the task completion, the Information Seeking Task level (IST) and the Information Retrieval Task level (IRT). Each level has a set of sub-processes as can be seen in Figure 2. The framework views the IS&R as dynamic and interactive processes embedded in a larger work-task environment (Hansen, 1999; Byström and Hansen, 2002) and explicates CIR activities at each IS&R stage.

Figure 2. CIR framework for the patent domain.


CIR events
Context - Social, environmental and organizational factors
Task initiation stage
Registration and classification
Delegation and assignment
Knowledge experience and practises
Actor/performer
Human
Document
Task stage
Task preparation
Reviewing new or revised PA
Task planning and structuring
Task requirement formulation
Information need
Human
Document
Task stage: Information Seeking and Retrieval process
Information Seeking
Information collection
Information need formulation
Info need representation
Source(s) selection
Information Retrieval
(Electronic) source selection
Query formulation
Query reformulation
Relevance judgment
Human
Document
Task completion
Information compiling and extraction
Decision/Solution
Human
Document
Task outcome
Report creation
Internal information creation
Legend:
= Observed CIR activities
= Observed CIR activities not covered in this study
= Related stages in the IS&R processes
= Recognized IS&R task process covered in this study

The framework presents CIR activities that have been identified at each stage. Task performance may involve other processes as well, but in this study we will mainly discuss the collaborative activities in relation to the IS&R process (the area with grey shading) and collaboration within it. The framework could be interpreted as describing 2 processes in parallel: a first and *primary process* is one that describes the general information seeking and retrieval process and the relationships between the different components. Additionally, on the right side of the framework, we describe a *secondary process* of collaborative activities involving both document-related and direct human-related CIR events. The processes are interrelated and may describe a general IS&R model from a collaborative perspective. The CIR process may be further developed into a more fine-grained level. With document-related we mean collaborative activities that are based on documents and textual information. These may be written notes that are used by others or search logs that are processed and (re)-used for similar tasks. Document-based collaborative manifestations could be made explicitly (the activity is done with the goal to be shared with others) or implicitly (the activity may eventually lead to a collaborative activity, but is not primarily intended to do so). The direct human CIR activities are of a collegial character, e.g., discussing and adjusting to a consensus within a department on how to proceed in specific situations. The direct human CIR activities are mainly synchronous and the document-related activities are to a great extent of asynchronous. Finally, this IS&R framework describes processes within the Patent domain. In this framework, it is the sub-tasks within the three main stages that are of importance. It may not be the case that CIR always takes place at all the three stages (WT, IST, IRT). The framework aims at describing the patent handling processes and also were one can expect to find CIR activities.

4. Results

4.1 Type of collaboration

The *first research question* concerned *how* collaboration activities occurred during information seeking and retrieval processes. In order to detect collaborative actions, we analysed transcriptions of the on-site observations and the diaries using content analysis. A set of collaborative activities was identified. In the second phase of analysis we found the following two *main categories* of collaborative activities: document- and human-related.

4.1.1 Document-related collaborative activities

This group of collaborative activities involves creating or using documents (electronic or paper-based), such as "working notes" that may contain information about search strategy, query terms and classification codes (see Task 111 below) etc. Such working notes are written down by the engineer and stored for later perusal by the engineer himself and her colleagues within the own work group. Other relevant document-related activities include the use of reports, reference documents, and notes accompanying the main PA, which the applicant has judged valuable for the patent engineer. We found 100 document-related CIR activities (Table 1) of which two are transcribed below:

> "In the incoming patent application (PA), I looked for referred patent documents made by the applicant [in this case made by a patent bureau]. I

> will now collect these documents and read them.” (These documents were then filed together with the patent application[3]; task 52 - diary notes).

> “I also take notes on classes, search terms and document numbers in order to use them later or for others to use them for similar tasks.” (Task 111 – diary notes).

In the sample of task 52, the references in the application is made to serve several purposes: a) they are used implicitly as evidence to show the patent office that the applicant have the necessary underlying knowledge regarding the area in general and the topic of the application specifically; b) they are used as pointers to the patent engineer examining the application in some specific direction or recommending these references as means for further understanding and discussions. The references are a product of earlier searches made by the applicant intended to support the formal application and the patent engineers in her/his problem solving. If the engineer finds it interesting, the documents are retrieved. The collaborative aspect of the sample Task 111 is that the notes are written down, stored in paper or electronic form by the patent engineer. These notes are in some cases stored together with notes made by other patent engineers in the same work group. These notes can later be reused by other within/outside the work group that created them. The retrieval aspect of the notes is that they are classified and ordered according to subject area and furthermore, colleagues could use them for query formulation. They are both retrieval using either search facilities on the electronically stored notes or manually using the small paper archives managed by individual PE or work groups.

[3] Authors remark.

4.1.2 Human-related collaborative activities

This category comprises of activities that directly use knowledge possessed by other humans in the patent handling process. Examples are asking colleagues internally and externally for advice and expert judgements. It also includes asking individuals or groups within or outside one's own department and subject area. We found 55 human-related CIR activities (Table 1). The following are transcribed examples:

> "X gets a visit from a colleague. They both discuss what applications X will work on in the coming period. The applications should be within subject area. Different strategies for searching and sources to be used were discussed." (Task 15 – observation note).

> "A colleague from my group came today and asked me for advice on how to proceed regarding the classification of a specific patent application." (Task 86 – "talk aloud" during observation)

In both cases, the patent engineers synchronously collaborated involving, among other things, specifying relevant information sources, search strategies and classification codes.

4.1.3 Summary

The findings definitely show that IS&R tasks are not performed in isolation. In fact, another result from the analysis is the development of the two categories of CIR activities. Extensive numbers of collaborative events and activities have been observed (Table 1). Furthermore, they do not only refer to direct activities between humans, but are also related to document-based activities.

Table 1. Distribution of unique CIR activities by type across individual PA tasks

	Type of CIR activity		
Task	Document-related	Human-related	Total
1	7	2	9
2	9	7	16
3	2	3	5
4	11	2	13
5	15	3	18
6	15	4	19
7	6	5	11
8	12	5	17
9	7	8	15
10	10	10	20
11	4	6	10
12	2	0	2
Total	100	55	155
Mean	8,33	4,58	12,92
%	65	35	100

Table 1 shows that the both document-related and human-related CIR activities occurred in all tasks except for one task. A total mean of nearly 13 collaborative activities was observed and thus, collaborative activities are an important and essential aspect of the IS&R processes at PRV. The results also show that both document-related (mean 8,3) activities per task and human-related (4,6) activities per task are common. These are a fairly high number of events. In tasks 2, 5, 6, 8, 9, and 10 altogether 15-20 collaborative events were observed. This points to a dynamic and interactive information handling process. Tasks 1, 4, 5, 6 and 8 show 50% or greater share of document-based collaborations than human-related. In task 2, 3, 9 and 11, we observe a minimal but notable bias toward human-related CIR events.

4.2 Collaboration in ISR processes

Our *second research question* investigates *when* collaborative activities do occur. 12 observed patent handling tasks related to roughly 3 main task phases in the whole patent handling process and showed the following distribution: 10 were in the initial phase, 6 in the middle phase, and 5 at the completion phase, which makes a total of

21. This distribution is explained by the fact that some of the 12 unique tasks were observed during more than one phase.

We decomposed the IS&R activities into stages employing our patent CIR process model (Figure 2). We then identified collaborative activities and mapped them to the model. Basically, the process is described as containing the following levels from a work task perspective:

- Work task level
    - TI - Task Initiation Stage;
    - TP - Task Preparation and Planning Stage;
    - TC - Task Completion Stage–Usage and creation stage.
- IST - Information Seeking Task level;
- IRT - Information Retrieval Task level, and

At the work task level, the Task Initiation stage (TI) involves classification and task assignment; the Task Preparation and Planning stage (TP) involves reviewing new or revised PA, structuring the task and formulating requirements for the task in order to be soled. At the Information Seeking level (IST) the following sub-processes are involved: information need formulation; source(s); relevance judgement. The next level is the Information Retrieval level (IRT) that includes source selection, query formulation and relevance judgement. The final main stage is the Task completion stage (TC) and belongs to the Work Task level, involving activities such as information use and creation. The framework in Figure 2 depicts collaborations at different stages. The single activities within the whole process were counted and categorized and then mapped to the corresponding stage in the IS&R process.

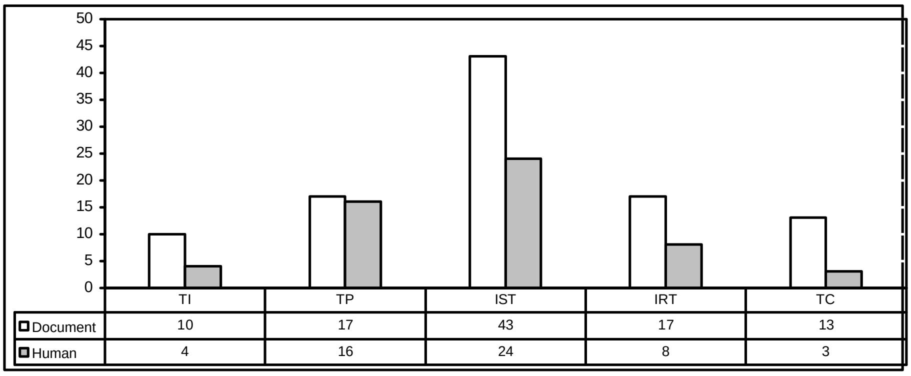
Table 2. Activities by collaborative *categories* along IS&R process stages

The IS&R process stage	*Collaboration categories*					
	Document-related		Human-related		Total	
	#	%	#	%	#	%
TI	10	10	4	7	14	9
TP	17	17	16	29	33	21
IST	43	43	24	44	67	43
IRT	17	17	8	15	25	16
TC	13	13	3	5	16	10
Total N=155	100	100	55	100	155	100

Table 2 shows how different categories of CIR activities were distributed across IS&R process stages. We see that collaborative activities were observed at all stages in the ISR process. Altogether 155 collaborative activities related to the IS&R process. In all, 65% of the activities were document-related which implies that the patent engineers experience a need to (re) use document-based information created by themselves or by others. This suggests that the patent system need to be designed carefully to facilitate reuse of document-related information. Not surprisingly, the task preparation (TP) stage shows a rather high score of collaborative activities (17% document-related events and 29% human-related events). Furthermore, the information seeking stage (IST) had the highest score of 43% document-related events and 44% for human-related collaborative events.

Altogether 55 of CIR activities (or 35%) were human-related. In this category, the stages of planning the task (TP) and the information seeking stage (IST) show a high score of CIR activities (29% respectively 44%). However, very interesting indications emerge in the information retrieval (IRT) stage in which a total of 25 (16%) CIR events were observed for both the categories. This stage is normally believed to involve individual activities of information processing, but as can be seen, collaboration is frequent. Furthermore, a rather low number of human-related CIR activities occur in the task completion stage (5%). In order to check the reliability of the categorisation of the data, we performed an intra-categorizer reliability check. Reliability was found to be 95% after a re-categorization of the data 3 months later.

Figure 3. Distribution of collaborative activities by IS&R process stage (N=155)



	TI	TP	IST	IRT	TC
□ Document	10	17	43	17	13
▨ Human	4	16	24	8	3

Legend: TI=Task Initiation; TP=Task Planning; IST= Information Seeking Task;

IRT= Information Retrieval Task; TC=Task Completion.

In Figure 3 we can see the distribution of all 155 CIR occurrences over the IS&R process stages. The results show that there are *document-related* CIR activities in *all* sub-stages in the IS&R process. Clusters with the highest degrees of document-related activities were found in the information-seeking task (43 events). We found a relative high degree of document-related events at the task planning stage, information retrieval task and information completion (TC) stage (17, 17 respectively 13). There is a high level of *human* CIR activities at information-seeking task (24) involving e.g. formulating information needs.

Table 3: The distribution of CIR activities over IS&R stages

	WT	***IST***	***IRT***
	Ti,tp,tc	ist	irt
	#	#	#
Doc	40	43	17
Hum	23	24	8
	63	67	25
%	41	43	16
n=155			

Table 3 shows the distribution of IS&R stages along the three major task levels of patent handling. In all 41% of the CIR activities occurred at the work-task level (ti,tp,tc) and 43%, respectively 16%, at the seeking and retrieval levels. If we decompose the activities further into document- or human-related events we observe an interesting pattern where 63% of all activities within the work task level were document-related events. The same pattern can be observed at the information seeking level (64%). Roughly, one third of the activities at all stages were human-related.

Finally, we want to decompose the data further according to three categories earlier proposed and described (Järvelin & Repo, 1983; Byström & Järvelin 1995; and Järvelin & Wilson, 2003): *problem solving knowledge* (PSK - deals with how problems should be treated and what knowledge is needed to solve problems); *problem knowledge* (PK - describes the structure and properties of the problem at hand); and finally, *domain knowledge* (DK - which deals with known facts and theories in the domain of the problem). Because several types of knowledge can be applied during one single collaborative activity, the numbers in Table 4 sum up to 222 activities.

Table 4: Distribution of document- and human-related activities across knowledge types and main work-task stages

	Work Task (WT)		*Information Seeking Task (IST)*		*Information Retrieval Task (IRT)*	
	DOC	HUM	DOC	HUM	DOC	HUM
PSK	22	21	22	12	14	4
	10%	10%	10%	5%	6%	2%
DK	22	13	39	21	10	3
	10%	6%	18%	10%	5%	1%
PK	2	0	5	3	9	1
	1%	0%	2%	1%	4%	0,5
Total	46	34	66	35	33	8
%	21%	15%	30%	16%	15%	4%
Total (doc,hum)		80		101		41
Total % (doc,hum)		36%		46%		18%

A total of 222 identified knowledge activities were recorded in the data. This means that in some of the 155 individual CIR-activities, two or more types of knowledge were exchanged. In order to check the reliability of the classification of the data into knowledge types, we performed an intra-classifier reliability check. Reliability was found to be 78-82% after a reclassification of the data 2 1/2 weeks later.

The analysis of the data clearly shows that on the work task-level, patent engineers acquire both document-based and human-based problem solving knowledge (PSK), while domain knowledge (DK) is mainly acquired through documents. Problem knowledge is not needed at this stage. Regarding the information seeking stage, domain knowledge is dominant, both in document and human-related activities (39 respectively 21 activities) followed by document-related PSK. At the retrieval stage, collaborative activities are mainly document-based. These findings do not record the length or duration of the exchanges.

4.3 Collaborative characteristics

Finally, the *third main research question* focuses on the characteristics of the two CIR categories: document-related and human-related activities. For this purpose, we analysed all 155 occurrences distributed over the 12 patent handling tasks in detail. The following is a selected list of characteristics:

Document-related activities. Collaborative activities observed related to documents can be classified as follows:

- *Sharing information objects/documents* such as articles, patent applications, working notes and reports. Working notes created by colleagues may be informative and reflect on previous processes involving a specific document and its connections and relationship to other documents. These documents may be part of other information retrieval processes. (26 instances[4]).
- *Sharing* different types of *contextual relationships* between individual, and sets of, information objects. Means for describing the relationships between documents may be through:
    - *Annotations* are content-based comments assigned to a document or specific sections of a document.
    - *References* are assigned to refer to topical or content-based relationship in a broader sense.
    - *Citations* are made to point out relationships to more specific parts of other documents and sections of documents.

[4] Each CIR activity may result in more than one instance, hence the larger number of instances.

These relationships are recorded on working notes or the actual patent applications that are filed. These can be stored and used as independent documents. (23 instances).

- *Sharing representations of information need*. The representations of the current information need may be stored and reused by other colleagues. Representations may be of two kinds: through classification codes, synonyms, query terms and query structures; and through a narrative description of the problem. Occasionally, these statements and descriptions are saved as a working note and reused in information seeking and retrieval activities. (11 instances).
- *Shared information seeking and retrieval strategies*. Different search strategies can be shared in two ways: a search history can be written down by the searcher to be used in later work-tasks, and log-statistics can be saved, processed and inspected for future use. In this category we may find sources used, statistics on time and number of sessions, documents inspected and documents printed out etc. (21 instances).
- *Sharing decisions, judgments* and assessments made on previous problem solving tasks that are of interest in the current work-task can be recorded on working notes or on copies of patent applications that are filed for archival purposes. "… if this assessment have been done for this problem solving task, parts of it can be used for a very similar task…" (task 111 – diary notes). (21 instances).
- *Communicating and sharing* of *personal and subjective opinions* in written form that, for example, reflect an immediate relationship between the document and its "neighbourhood". (8 instances).

- *Sharing* the *history of an information object*. The history of an information object/ document may be of three types. (36 instances). It may be in the form of a
    - *Document* history in which the document may belong to one or two subject areas assigned by the PE to the document (on paper and/or electronically) in order to be identified and re-used. On paper documents there may be dates and stamps from previous investigations. Comments, decisions may set a history for the document. For the electronic document, annotations in electronic form or in paper form pointing to the "source"-document may be added. The document then has some sort of history that the current user can use for specific purposes;
    - *Log* history, which may involve that all databases and sources, search terms, concepts, query term structure etc, may be stored and reused implicitly or explicitly by other colleagues. They could serve as recommendations or as precise pointers to a problem solving activity*;*
    - *Link* history, which may refer to that the document of interest has links to other documents and has links to itself from other documents. The links made by the PE may be related to activities before the relevance judgment of that document is recorded and could be used for strategy and context-building activities. Also, the actual judgment activity may be recorded as well as information on how the document is used for an end product.

We note that several of the observed and identified collaborative activities correspond to general stages in the IR process (Figure 2). The patent work may not involve identical information needs twice since they are unique to a great extent. However, partial overlap is rather common and the patent engineers may share information representations were parts of a related information need could be reused. Furthermore, the data show that the activities in this section could be both explicitly or implicitly stated for sharing.

Human- related CIR activities. Examples of human CIR activities are the following:

- *Task cooperation*. Sometimes there is a need to share a patent application task due to various reasons. This could be done *sequentially* or in *parallel*. (5 instances).
- Sharing *division* of PA tasks. Colleagues verbally discuss and decide how to divide the incoming patent applications among each other in the subject group or if it is necessary to assign the PA to another group within the organisation. (6 instances).
- Sharing *search strategies*. Search process/strategy was verbally shared and used in a collaborative way if target documents are closely related. In this classification we also find sharing search terms and classification codes. (21 instances).
- Sharing, or asking for, *external and internal domain expertise*. Patent engineers use both internal as well as external expertise to help with problem solving. Colleagues might internally be asked for domain specific knowledge as well as for information retrieval specificities, while external advice might concern clarification, law etc. (20 instances).
- *End product creation*. In the final phase of a task, humans may collaborate to finalize the end product of the problem-solving task. In the case of the patent domain this is often a report covering the outcome of the search and its applicability to the stated claims in the PA. (4 instances).

- Communicating and sharing of *personal and subjective opinions* in verbal form that for example reflects an immediate relationship between the document and its “neighbourhood”. (6 instances)
- Sharing *internal experience*. Ask colleague regarding earlier experience with similar type of applications. This category involves also issues such as procedural, legal and strategic issues. (12 instances).

Human collaborative activities show a pattern that comprises of asking colleagues both internally and externally regarding experiences, and search strategies. One collaborative activity that was not present but expected in the data was the sharing of knowledge about source selection. One obvious reason for the lack of this is that the PE´s has rather high knowledge about available resources and sources and thus they do know what is there to be used. O´Day (1993a) described four levels of information sharing in group situations: sharing results with other members of a team; self-initiated broadcasting of interesting information; acting as a consultant, handling search requests made by others; and archiving potentially useful information into group repositories. Our data suggests at least two additional levels:

- *Case-building activity*: At this level, one collects and adds knowledge such as internal and external documents, people that have been contacted in this specific case, how the case is filed and in what context.
- *History building of an information object*: At this level one collects all information and *traces* made during the task performance process. It involves gathering search terms, classification codes, documents viewed, decisions made, relevance judgments made and how the relevant documents are being used in final reports and the report creation itself.

In summary, we have identified a large set of document and human-related collaborative characteristics involved in the IS&R process. Notably, these collaborative activities do not only belong to the information seeking stage but also to the information retrieval stage.

5. Conclusions

IS&R is more than just the interaction between a single user and a system. In this article we have investigated the manifestations of collaboration in IR. The methods used in the study contributed to a rich and varied set of data. Although the study only involves PA 12 processes, the in-depth analysis of our observations of CIR activities within the patent domain reveals interesting results. We found that collaboration in IS&R is frequent. More frequent than expected. This was evident in *all 12 cases*. The results suggest that collaborative activities are an *important characteristic* of IS&R tasks in professional task-based IR.

We have, using a pragmatic and holistic methodology, pointed out that collaboration is an aspect to be considered, and our empirical findings resulted in the development of a conceptual framework for the description of CIR activities in a professional domain of the patent handling process (Figure 2). The framework points to important aspects related to *when, what* and *how* collaborative activities manifest themselves in work-task performance. We have analysed and discussed two important categories of CIR activities: document-related and human-related. Furthermore, we extracted a set of characteristics describing each of these classes.

Based on the results, we have identified a set of issues that are important regarding collaborative aspects of IS&R.

- *Planning tasks.* Even if there is a formal procedure of structuring the work-task, a patent application may contain aspects that lead to the departure from that procedure. This may also involve issues like how to approach the process or how to consult information sources never used before.
- *Problem definition.* Due to novel areas of invention and the complexity and variation of the sub-problems in a patent application, support is needed to find the right focus and core problem of the patent application. In complex cases the core problem is hidden or divided into several parts.
- *Search topic selection* may also be a problematic area that results in collaborative activities. Once the problem is defined, there may be a need for support regarding *query formulation*. The ability to choose the right query keys might be supported through reusing earlier query formulations. A subject-based "query history" tool might support this need.
- Regarding patent application belonging to a rather well defined area, there are reasons to reuse and share initial, baseline *search paths* and *query construction sequences*.
- When 2 or more subject areas were involved, *relevance assessments* had to be made by several domain experts or senior experts. Due to the specific domain (patent), there was a need to discuss the *final outcome* with colleagues as well as to check with other information previously handled by the patent office before *task completion*.

We find it necessary to develop analytic tools to collect and analyse data collected from real-life setting involving collaborative IS&R processes. We also need frameworks and models in order to develop systems and tools which carry us away from the “single-user interaction” metaphor and viewpoint, and which may support both synchronous and asynchronous communication, communication with colleagues, sharing information such as search strategies and search results.

Future research should focus on what affects CIR processes. Possible research questions could deal with task variation, task complexity or type of task. Furthermore, we need to develop information (seeking and) retrieval frameworks and theories that include methodologies for evaluation of collaborative IR systems. Consequently, this knowledge will also have an impact on the evaluation of IR systems. The issue of relevance judgements will be affected as well by thoroughly understanding CIR.

Acknowledgement

This project was made in cooperation with the Swedish Patent- and Registration Office (PRV), and we would like to thank all the participants in the study at the Swedish Patent Office and especially the project-leaders from PRV, C-O Gustavsson and Stig Edhborg.

References

Ackerman, M. & Malone, T. (1990). Answer Garden: A tool for Growing Organizational Memory. *Proceedings of ACM Conference on Office Information Systems 1990* (pp. 31-39). Cambridge, Massachusetts, United States: ACM

Allen, T. J. (1977). *Managing the flow of technology: Technology transfer and the dissemination of technological information within the R&D organization*. Cambridge, MA: MIT Press.

Byström, K. & Hansen, P. (2002). Work tasks as unit for analysis in information seeking and retrieval studies. The Fourth International Conference on Conceptions of Library and Information Science: Emerging Frameworks and Methods. CoLIS4, (pp. 239-252). Seattle, WA, USA, July 21-25, 2002.

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

Crabtree, A., Twidale, M. B., O'Brien, J. & Nichols, D. M. (1997). Talking in the library: implications for the design of digital libraries. In: Allen, R., & Rasmussen, E. (eds). *Proceedings of the second ACM international conference on Digital libraries.* (pp. 221 – 228). Philadelphia, Pennsylvania, US. July 23 - 26, 1997.

Ehrlich, K & Cash, D. (1994). Turning Information into Knowledge: Information Finding as a Collaborative Activity. Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries, June 19-21, 1994 (pp. 119-125). College Station, Texas, USA.

Fidel, R., Bruce, H., Pejtersen, A. M., Dumais, S. Grudin, J., & Poltrock, S. (2000). Collaborative Information Retrieval (CIR). The New Review of Information Behaviour Research, 2002, pp. 235-247.

Haake, J. M., Wiil, U. K. & Nürnberg, P. J. (1999). Openness in shared hypermedia workspaces: The case for collaborative open hypermedia systems. SIGWEB Newsletter, 8(3), 1999, pp 33-45.

Hansen, P. User Interface Design for IR Interaction. (1999). A Task-Oriented Approach. In: Aparac, T., Saracevic, T., Ingwersen, P. & Vakkari, P. (Eds). *Proceedings of the Third International Conference on Conceptions of Library and Information Science: Digital Libraries: Interdisciplinary concepts, challenges and opportunities* (CoLIS3) (pp. 191-205). Dubrovnik, Croatia. Zagreb: Zavod za informacijske studije Odsjeka za informacijske znanosti: Filozofski fakultet; Lovke: Naklada Benja.

Hansen, P., & Järvelin, K. (2000). The information seeking and retrieval process at the Swedish Patent- and Registration Office. Moving from lab-based to real life work-task environment. In: Kando, N. and Mun-Kew Leong (eds.) *Proceedings of the SIGIR 2000 Workshop on Patent Retrieval* (pp. 43-53). Athens, Greece: ACM SIGIR.

Harper, R. and Sellen, A. (1995): Collaborative tools and the practicalities of professional work at the International Monetary Fund. In: Katz, I., Mack, R. & Marks, L. (Eds.). *Proceedings of the ACM CHI'95 Conference on Human Factors in Computing Systems*. (pp. 122-129). ACM Press, New York.

Hertzum, M. (2000). People as Carriers of Experience and Sources of Commitment: Information Seeking in a Software Design Project. *New Review of Information Behaviour Research*, 1, 135-149.

Herzum, M., & Pejtersen, A. M. (2000). The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing & Management, 36(5),* 761-778.

Järvelin, K. & Repo, A. (1983). On the impacts of modern information technology on information needs and seeking: A framework. In H.J. Dietschmann (Ed.), Representation and exchange of knowledge as a basis of information processes (pp. 207-230). Amsterdam, NL: North-Holland.

Järvelin, K. & Wilson, T.D. (2003). On conceptual models for information seeking and retrieval research. Information Research, 9(1) paper 163 [Available at http://InformationR.net/ir/9-1/paper163.html] accessed 24 October 203.

Karamuftuoglu, M. (1998). Collaborative Information Retrieval: Towards a Social Informatics View of IR Interaction. JASIS, 49(12), 1070-1080, 1998.

Kuhlthau, C. (1993). Seeking meaning. A process approach to library and information services. New York: Ablex publ. 1993.

Malone, T., Grant, K., and Turbak, F. (1986). The Information Lens: An intelligent System for Information Sharing in Organizations. In M. Mantei and P. Orbeton (eds.), *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1986, (pp. 1-8), New York: ACM Press.

Marchionini, G. (1995). Information seeking in electronic environments. Cambridge Univ. press, Cambridge. 1995.

McDonald, D. W., & Ackerman, M. S. (1998). Just talk to me: A field study of expertise location. In: Poltrock, S & Grudin, J. (Eds). Proceedings of the ACM CSCW’98 Conference on Computer Supported Cooperative Work (pp. 315-324). New York: ACM Press.

Munro, A. J., Höök, K. & Benyon, D. (1999). (Eds). Social navigation of information Space. Springer, London. 1999.

O’Day, V. & Jeffries, R. (1993a). Information artisans: Patterns of Result Sharing by Information Searchers. Proceedings of the ACM Conference on Organizational Computing Systems, COOCS'93, (pp. 98-107), Milpitas, CA, New York: ACM Press. 1 – 4 Nov. 1993.

O’Day, V. & Jeffries, R. (1993b). Orienteering in an information landscape: how information seekers get from here to there. Proceedings of the conference on Human factors in computing systems. INTERCHI '93, (pp. 438-445), ACM, 1993, Amsterdam, The Netherlands, New York: ACM Press

Pemberton, D., Rodden, T. & Proctor, R. (2000). GroupMArk: A WWW Recommender System Combining Collaborative and Information Filtering. In: ERCIM Workshop on “User Interfaces for All”, 25-26 October, Florence, Italy.

Pinelli, T., Bishop, A., Barclay, R. & Kennedy, J. (1993): ‘The information-seeking behaviour of engineers’, In: Kent, A., and C. Hall, M. (eds.): Encyclopedia of Library and Information Science, vol. 52, supplement 15, (pp. 167-201), Marcel Dekker, New York.

Roberson, S & Hancock-Beaulieu, M. (1992). On the evaluation of IR systems. Information Processing and Management, 28, 457-466. 1992.

Romano, N. C. Roussinov, D., Nunamaker, J. F. & Chen, H. (1999). Collaborative Information Retrieval Environment: Integration of Information Retrieval with Group Support Systems. Proceedings of the 32nd Hawaii International Conference on System Science (HICSS-32), 1999.

Soininen, K. & Suikola, E. (2000). Information Seeking is Social. Proceedings of the First Nordic Conference on Computer-Human Interaction, NordCHI 2000, (pp. 1-9), 23-25 October 2000, Stockholm, Sweden: Royal Institute of Technology.

Sonnenwald, D. H. & Pierce, L.G. (2000). Information behaviour in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control. Information Processing & Management 36(3): 461-479.

Talja, S. (2002): Information sharing in academic communities: Types and levels of collaboration in information seeking and use. New Review of Information Behavior Research 3, 143-159.