

Association for Information Systems AIS Electronic Library (AISeL)

ECIS 2008 Proceedings

European Conference on Information Systems
(ECIS)

2008

Knowledge Workers' Use of Electronic Information Sources

Taline Jadaan

Viktor Institute, Gothenburg, Sweden, taline@viktor.se

Dick Stenmark

University of Gothenburg, dist@ituniv.se

Follow this and additional works at: <http://aisel.aisnet.org/ecis2008>

Recommended Citation

Jadaan, Taline and Stenmark, Dick, "Knowledge Workers' Use of Electronic Information Sources" (2008). *ECIS 2008 Proceedings*. 149. <http://aisel.aisnet.org/ecis2008/149>

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2008 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

KNOWLEDGE WORKERS' USE OF ELECTRONIC INFORMATION SOURCES

Taline Jadaan, Viktoria institute, Hörselgången 2, 2tr, S-41756 Gothenburg, Sweden,
email: taline@viktoria.se

Dick Stenmark, IT University of Gothenburg, Department of Applied IT, S-41296
Gothenburg, Sweden, email: dist@ituniv.se

Abstract

Knowledge workers are those who interact knowledgeable with information by creating, reading, analyzing and acting upon it. Being able to find relevant information is thus an important task for any knowledge worker, but how is this achieved? By interviewing knowledge workers about their information seeking activities, we have produced novel findings. Firstly, we suggest that the knowledge worker moves between and within three different information environments – the local, the organizational and the global – and are thus forced to switch between tools to satisfy an information need. Therefore we suggest that future tools need to be designed to allow seamless interaction across all environments and tools. Secondly, the knowledge worker does not use the intranet search engine but finds intranet information via URLs received from colleagues. Thirdly, the knowledge worker seems to appreciate the judgment of fellow employees and to trust human filtering more than computer algorithms. Fourthly, surprisingly often the knowledge worker searches manually in the local and organizational environment, despite the existence of search tools. In contrast, when the public web is searched, search engines are often used heavily. We discuss how these findings are useful insights for the design of future information seeking tools.

Keywords: Information seeking, knowledge workers, electronic information environments.

1 INTRODUCTION

Information has increasingly become the cornerstone of all organizational work and organizational members spend much of their days producing, seeking, sharing and consuming information. Numbers from the early nineties suggest that engineers spend 40-66% of their time sharing information (King et al., 1994) or 20-80% of their time seeking and manipulating it (Leckie et al., 1996). Engineers were amongst the first users of information to be studied (Fidel and Green, 2004), and the research concerning their information-seeking behavior is now in its fourth decade (Kwasitsu, 2003). Since being able to find information often is a prerequisite for being able to perform once duties, the interest in studies of information seeking behavior amongst ordinary business people has increased over the last decades. Many aspects of information seeking behavior have been covered over the years, but the topic is far from exhausted, we argue. In particular two circumstances speaks in favor of continued inquiry into this matter; firstly, the fact that advancements in information technology has changed dramatically both the workplace and the very nature of work (Ware and Grantham, 2003) in the last decade makes previous findings dated and calls for a revision of these findings, and, secondly, most studies have been conducted by information scientists who has primarily been interested in the users' behavior and less so in understanding how to design the technology.

The purpose of the paper is thus twofold; to further our understanding of information seeking behavior and to sketch out implications for the design of future information systems. By addressing both these two important issues we help span the gap that currently exists between the information science and the information systems disciplines (Ingwersen and Järvelin, 2005).

Engineers have by some been called "knowledge workers" since they need to collect, use, evaluate and generate a huge amount of information in order to complete their work (Cheuk, 1998). According to other definitions, a knowledge worker is someone who interacts knowledgeable with information and sees information not only as something derived from knowledge but as something that changes knowledge (Schultze, 2000; Sellen et al., 2002). Being able to find relevant information is thus an important task for any knowledge worker, including engineers. Information seeking is a process that encompasses many activities that are not necessarily related to information systems (IS). However, being IS researchers, we have in this study focused particularly on the knowledge workers' use of IS, the outcomes of this use, and what design implications for new information seeking tools. To successfully enhance practitioners' information seeking, argue Fidel and Green (2004), one needs to examine the specific factors that cause them to prefer one source over another. This study is this situated in a real organizational context and our research question is: *How do knowledge workers go about to satisfy their information needs*; i.e. what kind of information sources and tools do they use?

The rest of the paper is structured as follows. The next section presents related work on knowledge workers' information seeking behavior. Thereafter we describe the site and the method used in this study. In section 4, we account for our empirical findings, and discuss and analyze these in section 5. Finally, we answer our research questions and conclude the paper in section 6.

2 RELATED WORK

Engineers represent a diverse group of professionals. However, Byström and Järvelin (1995) note that most engineers can be characterized as subject specialists who perform rather complex tasks. As task complexity increases, so does the complexity of the information needed by the engineers, while the number of useful information sources decreases. Further, Hertzum and Pejtersen (2000) characterize engineers as having a certain degree of freedom in choosing the way they want to accomplish their work, and being expected to make informed decisions in a number of situations where many possible solutions are available. Apart from the specifics of the task at hand, the choices made by engineers

depend to a large extent on their understanding of the context of the task and, consequently, on their success in obtaining information about this context.

Having studied how people organize their office desks, Malone (1983) found that people organize their personal information environment in order to be able to (re-)find things at a later stage but also to remind them of things to do. Whittaker and Sidner (1996) came to the same conclusions having studied personal information management of e-mail; people organize their e-mail in order to be able to (re-)find things but also to serve as a reminder of tasks to do. There are thus obvious analogies between how we organize physical information in our offices and digital information in our computers. However, finding a piece of paper in an office is a manual and often laborious task, not easily supported by IT. Finding digital information should be easier, but unfortunately we often struggle in this area too.

Surprisingly little research is found on knowledge workers' information-seeking habits and practices. Cheuk (1998) referred to engineers as knowledge workers since their work includes collecting, using, evaluating and generating huge amounts of information, but as IT continues to transform the workplace, this is now true also for many other professions.

Martin and Metcalfe (2001) studied knowledge workers from information awareness and information dissemination points of view, focusing on how knowledge workers kept themselves informed in times of information saturation and information overload. Based on a literature review they argued that even though Internet and web environments enable knowledge workers to access information directly, librarians will continue to have an important role filtering, customizing and providing awareness of information relevant the knowledge workers' needs. This study did not present any empirical data.

Jones et al. (2001; 2002) coined the phrase *Keep Found Things Found* (KFTF) while examining how managers, information professionals and researchers managed to (re-)find the information they needed. Jones and colleagues focused in particular on web-based information and found that the two features explicitly supported by the web browsers (i.e., bookmarks and history lists) were relatively underused. Instead, their users engaged in a large variety of methods, including send email to themselves, pasting URLs into documents or printing web pages. Jones et al. did, however, not present any concrete design implications for better IS.

Sellen et al. (2002) studied how knowledge workers used the web. In their work, they took a rather broad approach to web use, including activities such as *transacting*, e.g., paying bills, *communicating*, e.g., chatting, *housekeeping*, e.g., checking whether a site was up to date, *finding*, e.g., searching for something specific and well defined, *information gathering*, e.g., providing background information on a topic, and finally *browsing*, e.g., non goal-driven information searching. Sellen et al. found that information seeking, i.e., finding, information gathering and browsing, accounted for 86% of the observed activities.

More recently, Géczy et al. (2007) looked at knowledge workers browsing the intranet and found that the knowledge workers had focused interests and explored only a narrow subset of the available resources. The knowledge workers examined in their study achieved their goals in few steps through repeatedly applied browsing patterns. The study did however not reveal what sort of information needs the knowledge workers had or what other information sources beside the intranet that was used.

What is evident from our literature review is that research has often focused in on only one information context at the time, such as email (Whittaker and Sidner, 1996), the public web (Jones et al., 2001; 2002; Sellen et al., 2002) or the intranet (Géczy et al., 2007). Not many look at how multiple sources are being used in combination. We argue that a more holistic approach to organizational information seeking is needed.

3 RESEARCH METHOD AND CONTEXT

The study took place in a large Swedish IT consultancy firm. The company has approximately 6.800 employees in Europe, North-America, South-America, Asia and Australia, and some 1.000 of these works in Sweden. The company has used email since the early 1980s and implemented a corporate intranet in the mid-1990s. The recent intranet that was implemented is highly standardized where all content is published by specific web masters with the help of Microsoft's Content Management Systems (CMS) tool. In Addition to the CMS-based intranet, the company also uses Teamplace, a collaboration environment for projects and small working groups to share information, document management and collaboration. The Teamplace environment can be searched using the search engine built into MS Sharepoint, while the remaining intranet is indexed by a commercial intranet search engine. The company has also several department hard drives to facilitate information storing and sharing. Microsoft Outlook is used as the e-mail application and the company uses MSN Messenger for instant messaging,

The data was gathered via thirteen semi-structured interview with various knowledge workers whose job roles included project coordinators, mainframe technicians, procurement managers and system developers. The respondents were randomly selected from the company's phone book, and contacted via email. Individual in-depth interviews were scheduled and subsequently carried out at the respondents' work place. The interviews, which were all recorded and transcribed, varied in length between 46 and 88 minutes and focused on, what information the respondents needed to carry out their work task, what sort of information sources the used to obtained this information, and how they went about finding the information.

During the data analysis phase we started with an open coding phase where the data was repeatedly read and grouped into concepts that were suggested by the data itself rather than being informed by theory (cf. Orlikowski, 1993). A first read-through of our data resulted in 215 instances where information was needed. Examining these in detail, removing doubtful cases and concatenating episodes where multiple sources were used to fulfill the same need, resulted in a reduction to 88 completed search episodes. As relationships between sources and tools begun to emerge, we re-read and re-interpreted previous transcripts paying attention to these patterns. This iterative process, where the respondents' statements were contrasted and compared with one another and to our tentative understanding, continued until no further modifications to our conceptual model could be inferred.

One interesting aspect that emerged out of our analysis when trying to match information seeking to particular sources was the observation that respondents often were forced to switch between different information sources to accomplish his/her work task. When doing so they also needed to switch tools. For example, a respondent could search the local hard drive by manually examining a folder structure, then turn to the intranet and try to find information by navigating the menus, and thereafter use a search engine to search the public web – all during the same search episode. We therefore chose to structure our empirical finding based on the three *information environments* that our data suggested, in order to attain a better conceptualization. We named the three information context *the local*, *the organizational*, and *the global* environments, respectively. The *local environment* comprises the knowledge worker's own hard drive and his or her own e-mail inbox. The *organizational environment* includes the email system, shared file servers, group disks, document management systems and the corporate intranet, and the *global environment* consists of the Internet and external databases.

4 EMPIRICAL FINDINGS

The employees used a variety of different information sources, search tools, and strategies that often blended into one another. When reporting the empirical findings, we will use the three information environments described above – *the local*, *the organizational*, and *the global* – to organize and report our empirical data.

4.1 The local information environment

The local hard drive is a source of personal information. All respondents who keep documents on their individual hard drive tried to order it according to their own folder structure. The folders names were typically based on different projects or work task. To find documents by name, ten out of the thirteen respondents used the built-in search application in Windows operating system. However, most of the respondents claimed not to save documents on the local hard drive due to the fact that the documents they used were continuously updated and replaced, and that keeping a local copy would thus not be useful.

“I don’t save documents on my local hard drive because I’m always in need of the latest version. All documents that belong to a certain project have a specific project e-room where one can store and download the needed documents. This makes easy for the other members of the project to access them.” Chris

Others did not trust themselves being able to (re-)find information if they stored it on a file server or in the intranet. They preferred to keep a local copy just in case. Some also argued that keeping a copy on your own machine speeded things up.

“Personally, I think the intranet is poorly structured. I spend much time searching our intranet for some trivial stuff I know should be out there. Like meeting minutes. This makes you insecure, and then you start to save a lot of information locally and use desktop search to find it. Stuff that I am personally responsible for, I always keep copies locally – it’s much quicker.” Joan

Storing documents on the local hard drive meant that Windows built-in search function could be used to locate the files if the name or part of the name was known. However, searching document by name was often not sufficient and many respondents needed to be able to search for content. The organization did not provide any desktop search application but three respondents had on their own accord downloaded Google’s desktop search to help them re-find documents.

“I use Google desktop search constantly, I double click on Ctrl [to bring it up] all the time. If I don’t have the document stored on my local hard drive, then I can switch and search for it on the Internet from within the same application.” Carl

Much of the communication was carried out with the help of email. Email was among other things used to contact customers regarding changes in the specification of requirements and suppliers for latest update other information concerning systems. Emails were regularly saved for their information content but also in order to keep track of chronological details, e.g., when tasks were ordered or when questions were asked.

I save all the conversations which is good in case something turns up and people start to ask ‘Why did you do it like that?’, ‘Did I say that?’. You can then easily show them the email and say, this is what we agreed upon at that time.” Anne

The email system was also used as a storage facility and our respondents reported having requested colleagues to email information they had already received through other channels.

“Even when you speak to someone on the phone I ask them to send an email so I don’t have to take notes, because those notes tend to disappear so easily. If I get it in the computer, it doesn’t go away. But in order to find it again, I have to maintain a folder structure where you can keep things in separate folders. So I keep all the important stuff on my hard disk or in a local file server.” Mike

This behavior added to the flow of emails and to facilitate finding, many respondents used folders to categorize their information. Some information though is not easy to categorize or is too short-term to be worth categorizing. Those are often just kept in the inbox. Due to the high usage, the email-boxes were often overfilled, which resulted in warning messages telling the employees to decrease the amount data in their email-boxes.

“The message ‘Your mailbox is full’ turns up sometimes, telling me that it is time to clean up the mailbox. What I do then is to throw away everything I have in my inbox, what I have there is often not important because I save all the important mails in different folders when I receive it... Anne

In contrast to the local hard drive, most of the respondents never used the search feature in the mail application. Instead, they browsed through the folder where they figured this sort of information would reside.

“No, I read through the emails or look at the subject lines. No, I don’t use search, I read it and I have my folders.” Linn

Email is also frequently used to send and receive links for different web pages regarding project documentation, department meetings, finding information about fellow employees, such as email address or phone number, and other administrative tasks.

4.2 The organizational information environment

As noticed above, the respondents often used email to search for information about fellow employees. However, email was mostly used in cases when the respondents *did not* know in which department the sought-for employees worked, or what email address or telephone number they had. When the organizational belonging of an employee was known, the intranet was the primary source to search for information such as what they looked like or what competence they had. In addition, the intranet was often used to find a certain competence. The respondents very often knew approximately where in the organization a certain competence could be found and they used the intranet to navigate to that department’s web page. Once there, they could read about the employees, find a person with the required competence and then contact him or her via email.

“I usually know in which department the employee works. Let us say that I search for chassis purchaser. The thing I do then is that I click my way down with the help of the menus and it results in many clicks. I find the search engine very difficult to use – it doesn’t work me.” Mike

The respondents seldom used the corporate internal search engine to find intranet information. They preferred to have links sent to them by other colleagues. According to the respondents, the search engine would often return old and irrelevant information, and despite the many returned documents none of them were useful. Instead, they rely on colleagues to know where information is on the intranet and ask them for links. There are many project e-rooms on the intranet where project specific information such as meeting protocol, project direction and conditions are kept, and rather than to search (via the search engine or by browsing), people ask their colleagues to point them right.

“Often you know who works with this so you send them an email ‘Do you have the link to that and that?’ rather than to search on the intranet. You also get a lot of links to intranet sites sent to you, even by management, for corporate information. It’s not like they would write that ‘This information is available on the intranet’ but they write that it can be found following this link and then they add the URL. This shows that they don’t expect us to be able to find it (laughs).” Chris

MSN is mostly used for social purposes and indicates if the employer is at work, or in front of the computer. Even if MSN mostly is used for social purposes, sometimes short-term questions are being asked. The problem with MSN is according to most of the respondents based on that it’s not possible to save the conversation. Instead the employer often asks his colleague to email the information that he asked for. The respondents that worked as programmers often use MSN to send short code parts and ask colleagues for help.

“Most of the time the questions are of a social character, like when are we going for lunch, but sometimes I ask a colleague if he can send a piece of code. In those cases, though, I ask him to email it, because that makes it easier to store it and find it.” Ken

Shared file servers used to be a big source of information before the intranet was implemented. Today, it has lost much of its importance and only a few respondents were dependent on file servers for information seeking. The few file servers still in use typically contained rather static information such as manuals for mainframe and mini computers, operational handbooks, and such types of material. The respondents said that the file servers provided easy access to these resources, but at the same time confessed that most of the time when an error message arise, they would use the Internet to search for a solution to the problem instead of searching through the stored manuals.

4.3 The global information environment

All of the respondents used the Internet in their daily work and it was the typical starting point when wanting to get an overview of things or when looking for new ideas. In these situations, the respondents often had very vaguely expressed information needs.

"If I have an idea I may go to Google and enter the word to see what happens, and often I get new ideas or I get redirected to some vendor site where I can be inspired and see different markets, prices, and such..." Udi

The public web was also the primary source of information when gathering general information about a certain technologies or certain programming languages, or to get code examples that would help a programmer to encode and decode programs

"I search a lot on the internet. I search for code examples, definition for class and how encryption works. There is a site on the web I use it is called Microsoft online code project. You can see others' uploaded code examples and they also explain how they coded them. It's very useful. I normally start by using Google and see if I can find anything of interest. If not, then I'll go to Microsoft online code project." Fred

When searching for information on the public web, the respondents preferred sources they know they could trust. Many of these pages are then being bookmarked so they can easily be re-found.

"I bookmark web pages if I suspect I will need to go there again, if the information was useful or so. This way I don't have to download them to my computer – I can just go there again. But you get a lot of bookmarks and they are difficult to categorize." Teri

As indicated above, the respondents sometimes went from a broad, general search of the entire web to a narrower, more specific search of a particular site. Many of these sites were vendor driven sites where the respondents had to have user accounts in order to access the information. The respondents who regularly visited such sites found these pages to be very useful, in particular to find solution to problems or to understand systems-specific error messages.

"We use different operating system and each one of them has a web page where we can access information and also report errors that have occurred in our operating systems. I also download updates and useful programs and there is also a large database where you can read about reported errors and the solution for them. We usually report the problems via these web sites and later on someone contacts you through email and the communication continues until the problem is solved." Lisa

These vendor-driven sites were often trusted to be of high quality but the searchability of the sites varied. Some were easily to navigate among and the respondents could easily find the needed information or download updates, while others were not.

The respondents were all technology-savvy and experienced web users but there were big differences in how they used search engines. Some searched iteratively, using multiple query terms in different combinations.

"Yes, I used Google and you quickly learn to use more and more search terms. Sometimes you get zero results and that is good, because then you can start to remove search terms one by one until

maybe there are only four left. If you don't use multiple search terms you always get like 200,000 hits and it's difficult to find anything then." Teri

Others displayed a more hap hazardous approach, where they "tossed" in a few words and hoped for the best. According to the respondents, this often worked well enough.

"I usually start with Google and throw in a few search words I think will do it, and then I check the first result page to see if I got anything." Fred.

"You toss in some words and something comes back. You can be clever and use quotation marks and stuff but you usually find it relatively easy anyway." Mike

The general opinion amongst the respondents was that it was easier to find information using the public search engines than it was to find it with the corporate intranet search engine. As mentioned earlier in the organizational personal environment section, respondent often preferred to send an email and ask for a specific link then use the intranet search engine.

5 DISCUSSION

We will, in the discussion section, analyze our empirical findings based the three information environments we have identified and answer our research question by explaining how the knowledge workers go about to satisfy their information needs and what kind of information sources and tools they use. Based on the discussion we will then sketch out design implications for future search technologies.

5.1 Local hard drive, personal inboxes: the local information environment

The local hard disk was not often used as an information source due to the fact that the respondents seldom saved documents locally. The reason for this was that they often needed the latest version of a document and that these documents often were project-specific and frequently changed. Very few of the respondents mentioned any particular strategies for searching amongst their local files. File folders were routinely used to (re-)find documents but the respondents admitted that as soon as the number of folders passed a certain level, it became virtually impossible to find stuff anyway. This echoes what e.g. Whittaker and Sidner (1996) have reported. Still, neither Windows build-in search function nor desktop search was much used. There were exceptions, obviously as discussed at the end of this section.

E-mail was a very frequently used means for communication amongst our respondents, but it also had other functions. The personal inbox was often used as a storage device, helping people to remember and re-finding things, and therefore the email system became an important source of information. When searching the email system, the built-in search function was not often used. Instead, the users resorted to more manual routines. They typically tried to remember who sent the email or when the email was sent or what the email was called and then sort the emails accordingly. Some respondents used folders to categorize information and claimed that this worked for them. Still, most respondents admitted that the amount of email received was overwhelming and that they every now and then had to clean up their email system and throw stuff away.

The emails sometimes contained valuable information themselves, such as evidences of decisions made earlier in the process or information confirming something received through a telephone conversation, but very often the main feature of the email was the URL, pointing to the real information residing on the intranet. Such emails were received both as the result of an explicit request and as unsolicited email from managers. When needing corporate information, an employee would typically email a colleague and ask him/her to send an URL pointing to the information, rather than trying to use the corporate search engine. The respondents explicitly said that they avoided the intranet search engine since it seldom provided them with what they were looking for.

It was very clear that email and intranet were used in combination, where the email system acted as a storage facility for URLs. One would perhaps suspect that Bookmarks (in Netscape or Firefox) or Favorites (in Internet Explorer) would be less frequently used when URLs were kept in emails, but it seems our respondents saved their own URLs as bookmarks and URLs received from others as saved emails. As with email and file folders, bookmarks were difficult to keep track of as the set grows, as already noted in literature (e.g. Jones et al., 2001; 2002; Sellen et al., 2002). Further, no searching or sorting facilities are provided for URLs, which makes this information more difficult to find.

We found it rather surprising that so few used desktop search when those who did use it obviously appreciated it greatly. Joan could not imagine a world without her desktop search since she saved and kept every email she received and documented her experiences in locally stored documents. The main reason for not using desktop search, according to our analysis, is that the organization did not provide it. The three respondents that had installed the desktop search had done so in violation of corporate policy and not all employees were willing or able to do this. The fact that some chose to break with official policy indicates that there is a real need for local searching.

5.2 The email system, intranet, MSN: the organizational information environment

As mentioned earlier, the corporate intranet was often used as an information source but it was seldom searched via the search engine. The search engine produced too much result, the respondents complained, and too many duplicates. In addition, the resulting links were often perceived to be irrelevant. Instead the respondents used other strategies to find information on the intranet

One approach was to guess – based on organizational knowledge – where the information might reside and browse to this location via the menus. This strategy worked rather well for those who had worked for the company for many years and knew their way around. However, even such veterans complained the intranet was poorly structured and their browsing strategy failed when the proper whereabouts of a certain piece of information could not be accurately predicted.

Another more frequently used strategy to find information on the intranet was, as we saw above, to follow links received from colleagues via email. The email could have been received earlier and already reside in the email system, in which case the in-box and other folders had to be searched. If no email containing the URL could be found or if they knew they had no such email, they would email a knowledgeable colleague and ask him or her to send them the URL. This passing, trading and storing of URLs in the email system was very obvious, and regularly used as an information finding means. Jones et al. (2001; 2002) reported that their users often emailed themselves, but we have seen no indications of this in our study.

When comparing searching for URLs via e-mail and searching for the information directly using a search engine, it was obvious that the respondents preferred the former. Where the search engine presented hundreds of links where only one or two were correct, a knowledgeable colleague might have the precise URL that leads directly to the desired information. The reason the search engine is not used, we suggest, is that the respondents prefer the added quality that human filtering provides. The search engines' ranking algorithms keep improving but human judgment still seems to be valued higher. Today we see a growing interest in social software, such as e.g. deli.cio.us, where humans add value to information by tagging pages they found useful. The added value of peer filtering was something that we found in our study as well. Therefore, it seems plausible that traditional search engines would benefit from involving human judgment. This may explain why Google, whose page rank algorithm exploits human linking activities (Brin and Page, 1998), was the preferred search engine amongst the respondents. This will be further discussed in the next subsection.

MSN was not used much for information seeking in the “retrieval” sense. Due to its non-persistent nature, information would be lost once the session was terminated, and thus MSN could not be “searched”. Instead, MSN was used to complement email by indicating who was online and available and thereby able to give an immediate response to an information request. The non-persistent nature

forced users to apply specific strategies. For example, when used to ask for code snippets, those would typically be pasted directly into the code editor in order to be preserved. Likewise, when the respondents requested links to web resources, they asked for the URLs to be sent in emails to be searchable.

5.3 Internet, the public web, external databases: the global information environment

In our study the public web was used to interact with vendors' web sites to find information such as downloadable problem fixes or the meaning of error codes. Often, such sites require the searcher to be known and registered and they had to log in using a userid and a password. These sites are known to the users and their URLs were either bookmarked or memorized. For users who are not familiar with the systems (and thus do not have user accounts) it is practically impossible to learn what is inside since these sites are seldom indexed by the public search engines. In addition, we note that when user logged in to vendors' sites and thus left the public web, they showed the same behavior as when searching the intranet. They received URLs pointing directly to "their" case or problem, and when visiting the site to search for general information they used menus rather than the search engine.

However, we saw many examples of how our respondents, instead of starting at the vendors' site, first used a public search engine to search for e.g. code snippets and – if unsuccessful – thereafter turn to a vendor's site. Very often the searcher would just "through in" a few query terms and hope for the best, but we also saw more advanced usage where, the user would, for example enter an error message from the mainframe, or use multiple query terms in an iterative process. As mentioned earlier, Google which page rank algorithm exploits human linking was used by the majority of the respondents and they were pleased with it. In contrast, none of our respondents was pleased with the intranet search engine and avoided using it. Instead they relied on URL received from fellow colleagues. This, we argue, suggests that search engines would benefit from human judgment and peer filtering.

5.4 Summing up: Design implications and future work

Almost all respondents explicitly expressed that they felt they needed to save the address (URL) to a particular piece of information. It was obvious, even to some of the respondents, that they were not sure they would be able to (re-)find the information at a later stage unless they saved the URL. This is a known behavior and the browsers have built in support for this in form of bookmarks or favorites, but it is also known to be problematic (Jones et al., 2001; 2002). It is a strategy that does not scale well, there is no search support for bookmarks, they are not easily shared, and they present too little context to help the seeker find the relevant one when searched manually. This is another area where more research is needed, since there is a mismatch between preferences and support. We suggest future search tools redesign their bookmark feature to mimic the situation represented by a saved email, i.e., a fuller context that allows the user to associate to when, where and why a particular URL was saved.

Many respondents kept the bulk of their saved email in the inbox. Likewise, many respondents did not bother to categorize their bookmarks. This may not be problematic. Part of organizing is also the option to leave things (seemingly) *unorganized*. It is true that sometimes items *cannot* easily be categorized and filed, but perhaps more often the user *may not want* to categorize and file an item. In the former case, it is the cognitive load of, on the one hand, coming up with useful category names, and, on the other hand, understanding to what category a particular information item belongs, that keeps the user from filing. Successful filing depends on the ability to be able to imagine future retrieval requirements (Whittaker & Sidner, 1996). In the latter case, the user may not want to assign a piece of information to any particular category (since it may have multiple meanings or application areas) or may not want the information to disappear out of sight. Email folders, file folders and bookmark folders make information disappear from our radar. Malone (1986) suggests that the *spatial location* of information is perhaps more important than its logical classification. Being able to see it –

if only peripherally – enables both being able to find and being reminded. If the information becomes categorized and filed, it disappears out of sight and thus out of mind, and this may explain why people keep piles of documents on their desks and loads of e-mails in their inboxes. Therefore, we suggest future information technologies to more actively support the remembering aspect of a piece of information.

As is evident from our analysis email constitutes both a local and an organizational information environment. An email saved in the personal inbox represents a local aspect, i.e., a discrete piece of personal information that can be filed and sorted according to personal preferences and needs. At the same time, the email system represents an organizational aspect where employees are linked together through a corporate directory. Email users can search for and locate fellow employees, learn their organizational positions, find out their room or telephone number, and with a single mouse click send information requests to any one of them. We suggest that one reason why email is used so frequently in organizational information seeking is that it partly provides a bridge between different information contexts – users are able to search in (two) different environments from within the same tool. To develop this further would be to add awareness (e.g., from MSN) and formal retrieval functions (e.g., from the public Web).

We found that sometimes an information need could not be satisfied within any one particular information environment and the respondents therefore needed to move between three different environments, *the local*, *the organizational*, and *the global*, to satisfy their information needs. In addition, they also often had to switch tools, e.g., from the email client to Google or from the intranet to email, but also when searching within the same environment. This suggests that we need to design our information seeking tools to support context-switching rather than to force the seekers to resort to tool-switching. An in-depth analysis of how users move between tools and information environments has been undertaken and will be presented in a forthcoming paper.

6 CONCLUSIONS

By interviewing knowledge workers about their information seeking activities, we have tried to understand how knowledge workers go about to satisfy their information needs. Some interesting and novel observations have been made.

Firstly, we suggest that the knowledge worker moves between and within three different information environments; *the local*, *the organizational* and *the global*. Unfortunately, most of today's information seeking tools are designed for and constrained to one of these environments. Knowledge workers thus have to switch tools when moving to other environments but sometimes also while in the same environment. We suggest that future tools are designed to allow for seamless interaction across all environments and sources.

Secondly, surprisingly many of the knowledge workers searched manually. In email, they tried to remember from whom or on what day an email was received or whom to contact to receive an URL to the required information. This manual behavior applies also to the intranet (where menus were preferred to the search engine) and to the local hard drive (where the users relied on folders rather than desktop search).

Thirdly, the knowledge workers used bookmarks to remember the whereabouts of corporate external information but save URLs in emails for intranet information. Thus, the email system and the intranet were therefore often used in tandem; they search the email system to find intranet information. This is attributed to the social dimensions of email, which leads us to the fourth conclusion.

Fourthly, knowledge workers seem to appreciate the judgment of fellow employees and to trust human filtering more than pure computer algorithms. It therefore seems like a fruitful idea to let future search tools exploit and explore the knowledgeable actions performed by fellow employees, as is currently done in many social media or web 2.0 applications.

These findings suggest that current search tools and their features are not yet perfected and that more research has to be devoted to information seeking and finding in corporate environments.

References

- Brin, S. and Page, L. (1998) The anatomy of a large-scale hypertextual Web search engine, *Computer Networks and ISDN Systems*, 30, 1-7, 107-117.
- Cheuk, B. W. (1998). Modelling the Information Seeking and Use Process in the Workplace. *Information Research*, 4 (2), October 1998.
- Fidel, R. and Green, M. (2004). The many faces of accessibility: engineers' perception of information sources. *Information Processing & Management*, 40 (3), 563-581.
- Géczy, P., Izumi, N., Akaho, S. and Hasida, K. (2007). Long Tails and Analysis of Knowledge Worker Intranet Browsing Behavior. In *Proceedings of BIS 2007*, Poznan, Poland, April 25-27, 2007, pp. 584-597.
- Hertzum, M. and Pejtersen, A.M. (2000). The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing and Management*, (36), 761-778
- Ingwersen, P. and Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context*, Springer, Dordrecht, The Netherlands.
- Jones, W., Bruce, H. and Dumais, S. (2001). Keeping found things found on the web. In *Proceedings of CIKM '01*, Atlanta, GA., November 5-10, pp. 119-126.
- Jones, W., Dumais, S. and Bruce, H. (2002). Once found, what then?: A study of "keeping" behaviors in the personal use of web information. In *Proceedings of ASIST 2002*, Philadelphia, PA., November 18-21, pp. 391-402.
- Kwasitsu, L. (2003). Information-seeking behavior of design, process, and manufacturing engineers. *Library & Information Science Research*, 25 (4), 459-476.
- Leckie, G.J., Pettigrew, K.E., and Sylvain, C. (1996). Modeling the information seeking of professionals: A general model derived from research on engineers, health care professionals, and lawyers. *Library Quarterly*, 66 (2), 161-193.
- Malone, T. W. (1983). How do people organize their desks?: Implications for the design of office information systems, *ACM Transactions on Information Systems*, 1(1), pp. 99-112.
- Schultze, U. (2000). A Confessional Account of an Ethnography about Knowledge Work. *MIS Quarterly*, 24(1), pp. 1-39.
- Sellen, A. J., Murphy, R. and Shaw, K. (2002). How knowledge workers use the Web. In *Proceedings of CHI '02*, pp. 227-234.
- Ware, J. and Grantham, C. (2003). The future of work: Changing patterns of workforce management and their impact on the workplace. *Journal of Facilities Management*, 2 (2), 142-159.
- Whittaker, S. and Sidner, C. 1996. Email overload: exploring personal information management of email. In *Proceedings of CHI '96*, Vancouver, Canada, April 13-18, pp. 276-283.