Modeling the Information Behaviour of Software Engineers Using a Work - Task Framework

Luanne Freund

Faculty of Information Studies, University of Toronto, Ontario, Canada.
freund@fis.utoronto.ca.

Elaine G. Toms

Faculty of Management, Dalhousie University, Halifax, Nova Scotia, Canada.
etoms@dal.ca.

Julie Waterhouse

Centre for Advanced Studies, IBM Canada, Markham, Ontario Canada.
juliew@ca.ibm.com.

Faced with the rapid proliferation of digital information resources within organizations, employees need targeted search systems in order to be effective in their jobs. The goal of this study is to develop a model of information behaviour within a work-task framework that can be used to inform the design of a workplace information search system. In a two-phase process, we conducted a series of studies using multiple methods to identify workplace characteristics and to understand how they influence the needs, search strategies, and information sources used by software engineers working as services consultants in a large high-tech company.

Introduction

Information retrieval has become a key issue for organizations that produce and use large amounts of information in digital format. Organizational information resources exist in a wide range of formats, are often scattered across different types of

repositories, and are not heavily hyperlinked (Fagin *et al*., 2003; Hawking, 2004). The corporate intranets that contain these resources are often behemoths. IBM's intranet, for example, contains 50 million URLs on 7000 unique hosts (Fagin *et al*., 2003). Finding information in these organizational information spaces can be extremely difficult. People who use information in their work, spend about 15-35% of their work time searching for it, but they are successful only about 50% of the time (Feldman, 2004). Increasingly, employees are suffering from information overload, resulting in stress and loss of job satisfaction (D. Allen & Wilson, 2003; Edmunds & Morris, 2000) and organizations are incurring costs related to poor information access (Feldman & Sherman, 2003).

One approach to this problem is to develop search systems capable of retrieving results that are targeted to the work that people do. Our research is exploring ways of using information about the workplace search context to supplement keyword search systems and re-rank results based on their likely usefulness for task completion. In particular, we take the approach that information retrieval in the workplace is embedded in specific work and task frameworks, and that many of the decisions that people make about information are unrelated to semantic content, or topical relevance. Rather they are shaped by situational relevance, which is determined by factors within these contextual work and task frameworks (Schamber, Eisenberg, & Nilan, 1990). However, to date, information retrieval systems have made little use of the relationships between work and task context and search behaviour.

Our goal in this study is to develop an empirical model of how contextual factors shape the information behaviour of a group of software engineers in order to inform the design of a contextual search system for this work domain. We opted to study engineers, as their information seeking behaviour has been studied extensively, and is known to be task-oriented and strongly shaped by their work context (Leckie, Pettigrew, & Sylvain, 1996). The following questions guided this work:

- What are the contextual factors that influence the information seeking and searching behaviour of this group in the workplace?
- What is the nature of the relationship between these contextual factors and the perceived relevance of information sources and documents?

We believe that contextual search systems have the potential to reduce the effects of information overload and the challenges of information access in the workplace, by providing users with smaller sets of more useful documents. This study serves as the first step in developing a proof-of-concept contextual search system grounded in the information behaviour of a particular user community.

Previous Work

Contextual studies of information seeking and work

Context plays a powerful role in shaping how people search for information and in determining the information they select and subsequently use. This has been recognized in key models of information behaviour (Dervin, 1999; Wilson, 2000) and interactive information retrieval (Jarvelin & Ingwersen, 2004; Saracevic, 1996). Taylor (1986) noted two decades ago that information needs consist of two distinct but inter-linked parts: the topic or subject matter of the need, and the situational elements that will impact how the information is used. Contextual approaches are also common in fields such as Human Computer Interaction and Human Factors, which use a range of methods that focus on tasks and situations: contextual design, situated action, task analysis, cognitive work analysis, and activity theory (Button, 2003). In contrast, studies of information retrieval have traditionally approached search acontextually, by focusing on topics or queries divorced from the context of users and tasks. Search systems founded on the topic model are not able to filter results based on utility or value, which leaves the human searcher to manually review large amounts of information in order to meet their needs (Lawrence, 2000). In response to this limitation, contextual search has recently risen to the top of the research agenda for information retrieval (Allan, 2003)

Information behaviour of engineers

There is a large body of research on the information behaviour of engineers, which has

informed this study. Engineers have been characterized as service-oriented professionals with a strong knowledge base in their subject, whose work focuses on completing rather complex tasks and finding solutions to technical problems (Hertzum & Pejtersen, 2000; Leckie *et al*., 1996). Leckie et. al. (1996) have proposed a general model of information seeking of professionals, based on a survey of research on engineers, medical professionals and lawyers. This model emphasizes the impact of the work context, expressed as work roles, tasks, and information sources on the nature of information needs and seeking behaviour. Some additional contextual factors identified in this model are: demographics (career stage, speciality), project stage, features of the problem situation, and characteristics of the information need (frequency, importance, predictability, complexity). A similar contextual approach to engineers' information seeking is expressed in Taylor's (1991) Information Use Environments, and is well expressed by Pinelli (2001): the complexities of the decision processes used by engineers to evaluate knowledge require an understanding of the personal, situational, contextual, and community characteristics in which the engineer works.

Studies of engineers have identified some well-defined patterns of information behaviour. They are known to be highly motivated and heavy consumers of information: estimates suggest that they spend 20%-80% of their work time seeking and manipulating it (Leckie *et al*., 1996). However, they are often constrained by time and budget considerations (Fidel & Green, 2004), such that they typically seek a small amount of "good enough" information (Orr, 1970). Engineers tend to rely heavily upon verbal communication with colleagues, internal reports and personal information stores as sources (T. J. Allen, 1977; Hertzum & Pejtersen, 2000; Taylor, 1968). This behaviour has been interpreted as a function of Zipf's Principle of Least Effort (King, Casto, & Jones, 1994; Pinelli, Bishop, Barclay, & Kennedy, 1993), but recent research suggests that this principle does not adequately explain how engineers select among written information sources (Anderson, Glassman, McAfee, & Pinelli, 2001). Alternate explanations have also been put forward to explain engineers' preference for obtaining information from colleagues: the role of trust (Hertzum, 2002; Hertzum, Anderson, Anderson, & Hansen, 2002), the added value of feedback and contextualization (Hertzum & Pejtersen, 2000), and the challenges associated with finding written information (Hertzum & Pejtersen, 2000; Kwasitsu, 2003).

Much of the literature on engineers has focused on the identification of criteria for information selection. A range of studies has shown that accessibility is the key criteria (Leckie *et al*., 1996; Pinelli *et al*., 1993); however, recent work has brought this into question as well. In a questionnaire of 872 engineers and scientists, Anderson, Glassman, McAfee, & Pinelli (2001) tested a number of hypotheses relating to the selection of information carriers. With respect to the selection of written sources, the factors: accessibility, technical quality and prior experience with a source were not correlated, whereas importance to my work was. Although importance is not a well-defined concept in this study, the authors suggest that it may reflect a relationship between the source and its likely contribution to task completion. A similar finding was reported by Hertzum (2002), who found that the key criteria for selection for software engineers was appropriateness to task. Another recent study based on analysis of interviews with 31 engineers from a range of fields found that the concept of accessibility is multidimensional, and engineers rely on a range of associated factors in selecting information: familiarity with source, efficient to use, physically close, in the right format, and the right level of detail (Fidel & Green, 2004). Studies that focus specifically on software engineers have found that the currency of information is particularly important (Hertzum, 2002) as is the organization and format of the documentation (Forward & Lethbridge, 2002.

Information use among engineers seems to be closely tied to work and task context, and to issues of information access. This points to the importance of developing more effective, context-sensitive search systems for engineers. However, while the literature provides a clear picture of general patterns of behaviour and factors which shape this behaviour, there is very little research that identifies causal relationships between the factors and the behaviours.

Task-based information retrieval

There is a growing interest in the role of tasks in information seeking and retrieval (Bystrom & Hansen, 2005; Vakkari, 2003). A task can be described as "a piece of activity to be done in order to achieve a goal" (Vakkari, 2003), however in terms of

information retrieval, it is useful to view at least three levels of tasks: work tasks, which motivate information seeking tasks, which in turn motivate information search tasks (Bystrom & Hansen, 2005). Studies have sought to classify and characterize tasks in order to determine the nature of the relationship between tasks and information behaviour. Algon (1999) used multiple methods to derive a classification of task types for project related work and found that different tasks types were correlated with different sub-steps and information behaviours. Bystrom and Jarvelin (1995) studied the relationship between task complexity and information behaviour in a work environment. They classified tasks in 5 groups, according to a priori determinability, ranging from the least complex (automatic information processing) to the most complex (genuine decision tasks). Some key findings are that as task complexity increases, the complexity of the information need increases, the need for domain and problem-solving information increases, more internal sources are used, and the number of sources increases. Vakkari (2001) has focused on a process model of task completion. This work extends Kuhlthau's (1993) Information Searching Process model to identify changes in search terms and tactics at different task stages. Some of the findings are that as work on an information search task proceeds, searchers use more and varied search terms, which are more expressive of their topic; they use a greater number of search tactics; and they judge fewer documents relevant as their domain knowledge and understanding of the topic increase. This research has not yet resulted in information retrieval systems that incorporate these models; however there is strong evidence to suggest that there is value in taking a task-based approach workplace searching. However, tasks, just like information needs, are strongly shaped by situational and contextual factors (Bystrom & Hansen, 2005). For this reason, we have taken a broad approach to studying the work context and information behaviour of our study population.

Phase I: Understanding the Information Users and their Use Environment

In Phase 1 of our project, we used multiple methods to explore many parameters of the work environment and the information behaviour of our target group, software services consultants in a large high-tech firm. Findings from this initial phase shaped the

research design used in Phase 2.

Population

Our population was a group of software engineers working as software services consultants for a particular product line in a large hi-tech firm. The software services group, which includes over 1,000 individuals world-wide, works with industrial and commercial clients who have purchased their software products to provide expert assistance and guidance in implementing systems and solving problems.

Methods

We used four techniques to explore the work and information use of software consultants:
Focus group: Using electronic meeting software, five consultants identified and ranked the importance of the kinds of tasks they do, the problems they encounter, and the information sources they use.
Semi-structured interviews: Five consultants elaborated on their work practices and information use.
Job shadowing: Three work days were spent with one consultant at a customer site to observe workplace information practices directly.
Resource audit: Documents and sources of digital information were assessed to understand the types of resources available to this group.

Results

Findings from Phase 1 allowed us to construct a general model of the work and information practices of software consultants, and to understand the needs of these users and their work environment.

Results

Nature of the Work: Software services consultants work with clients to provide a range of services with respect to specific software products. They are, in essence, technical product experts, and handle a wide range of consulting projects (engagements) that vary from a single day to over a year; they may be working on the customer site or remotely from their own office. Some of the key tasks they do, as identified in the focus group, are: troubleshooting, performance analysis, consulting and advising, software design, and migrations. They often serve on engagements as a single consultant, but interface with other software engineers who work for the customer. This is a field in which technology and related information are highly dynamic. These consultants rely heavily upon information to do their work. They are commonly involved in on-site engagements working directly with clients who are very knowledgeable, so they are expected to be experts among experts. For this reason, consultants are constantly either trying rapidly to "get up to speed" on various topics, or are looking up specific information on-the-fly, in order to solve problems or complete tasks. Overall, they spend a considerable amount of time trying to obtain the information they need.

Information Sources: The documents used by this group of software consultants are widely dispersed on intranet sites, internal databases, and also in external web sites. We identified over 30 web sites and databases relevant to the work of this group, containing well over 50,000 documents. Documents range from brief e-mail messages to book-length manuals. These documents exist in a broad range of genres and file formats.

Information access: Documentary information sources are almost always accessed in digital format. Consultants access internal information sources in the company's intranet and internal databases, as well as on the web and in their own personal document collections, which they commonly store in directory structures on the hard drive of their portable computers. They use a standard battery of tools to find documents: search engines, directories, known URLs, bookmarks, and hypertext links. In keeping with other studies of engineers, this group also relies heavily upon written or verbal information exchange with colleagues. They use a wide range of tools to

communicate with colleagues, including internal and external forums, virtual teamrooms, e-mail, instant messaging, chat, telephone, and face-to-face communication.

Discussion

This phase of the study revealed an extremely complex information behaviour scenario. In particular, the large number of scattered information resources and access mechanisms and the great breadth of knowledge required to do this job contribute to this complexity. Although we were able to draw some general conclusions with respect to the information behaviour of this group, we found that their information needs, strategies, and sources seemed to be contingent upon the different engagements and tasks in which they were involved. As one consultant noted: "Well, there are different types of engagements. From short-term to long-term, from support and problem-resolution to advice and architectural review. Those are completely different games. So it very much depends...on the engagement, it's a different thing that you do...different types of information" (P07). It was on this basis that we undertook the second study.

Phase II: Linking Work Tasks with Information Behaviour

The goal of this study was to develop a model of how the work-task context influences strategies for finding and selecting information. In particular, we were aiming to: a) identify the key contextual factors, and b) determine the nature of the relationships between these factors, the consultants' information behaviour, and the information itself.

Methods

In designing the methods for this phase, we were constrained by some aspects of the work practices of software consultants (Singer & Lethbridge, 1996). The work of

software consultants is specialized and varies greatly from one engagement to the next. They are often working out of the office at customer sites, and under severe time and cost constraints. As we discovered in Phase I, much of their work is done via computer, so there is relatively little social interaction involved that can be easily observed by a third party. Finally, much of the time they are working with confidential information. Our decision to collect data through interviews rather than other methods such as direct observation or diaries was based on these constraints.

Participants

Fourteen software services consultants from two product groups in one company participated. All participants are software consultants who provide a range of services for customers. They have worked as consultants for 2-10 years. All have university degrees in a range of disciplines: engineering, computer science, commerce, systems analysis, and public administration; four have graduate degrees. They varied from generalist consultants at the intermediate to senior level, to those with distinct roles as project managers, information technology architects, or specialists.

Interviews

We conducted detailed, semi-structured interviews of 1-2 hours in length with the 14 participants. Half the interviews were conducted by telephone because of distance and time constraints. We adapted a contextual model of information seeking to serve as a framework for our interviews (Kari & Savolainen, 2003). This model identifies a large number of contextual layers that are likely to influence how people look for information. We focused our study on the layers that facilitate understanding information behaviour within this particular work context rather than across organizations and cultures. Thus, the interview focused on understanding consultants' information behaviour in the following ways:

- By engagement: Participants described several specific consulting engagements

in which they were recently involved, and identified the main information problems and resources used in each.

- By task or role: Participants discussed the specific tasks involved in their work by relating to a seed list of eight tasks derived from the pilot study (design, programming, migration, security, testing, performance tuning, installation, and configuration).
- By information goal: Participants discussed their information strategies with respect to a list of six different information goals derived from the pilot study (learning about, finding out how to do something, finding advice, looking up facts, finding a solution, finding a tool).
- By problem: Participants discussed how they would go about finding information with respect to a set of eight problems drawn from technical newsgroups.

Data Analysis

The interviews were digitally audio-recorded and transcribed for qualitative analysis using Qualrus software. Coding focused on the identification of themes, indigenous categories of tasks, information problems and sources, and correlations between work practices and information behaviour. In the analysis, we identified about 240 segments of text that discussed some aspect of information behaviour in the context of work practices.

Results

Information Access Issues

The interviews revealed that access to information is extremely important to this group: "finding the information quickly and finding the right information is absolutely critical in our position" (P06). Participants indicated that on average they spend about 20-30% of their time looking for and consulting information sources. One participant noted that this figure would have been higher if not for time constraints: "it is 30% not because we don't need more information, but because that is what we can spare. We also need to

work....You can never have enough. Its just that you spend 30% of your time finding just about enough then you have to go" (P07).

Indeed, participants confirmed the need for improved search systems, commenting on:

- **The overwhelming volume of information and information sources: "the amount of data resources ... are just immense. I mean, I bet you I use probably 1% of them" (P12).**
- **The lack of a central repository or search portal starting point: "we have sites all over the world, and you really don't know where to go. You might have your favourite sites, but sometimes it takes hours..." (P10).**
- **The instability of these sources, which change locations, change names, and become obsolete very quickly;**
- **Difficulties using existing search engines: "A lot of times it's also: How do I phrase this so I don't get 2,000 responses? If I look for certain words, I know I'm going to get thousands of responses and they don't mean anything when I sift through them" (P03).**
- **The poor quality of some of the information: "So there are lots of bad articles out there. How to say which ones are good is very difficult...there are lots and lots of people out there with incorrect information" (P06).**

Contextual Parameters of Information Behaviour

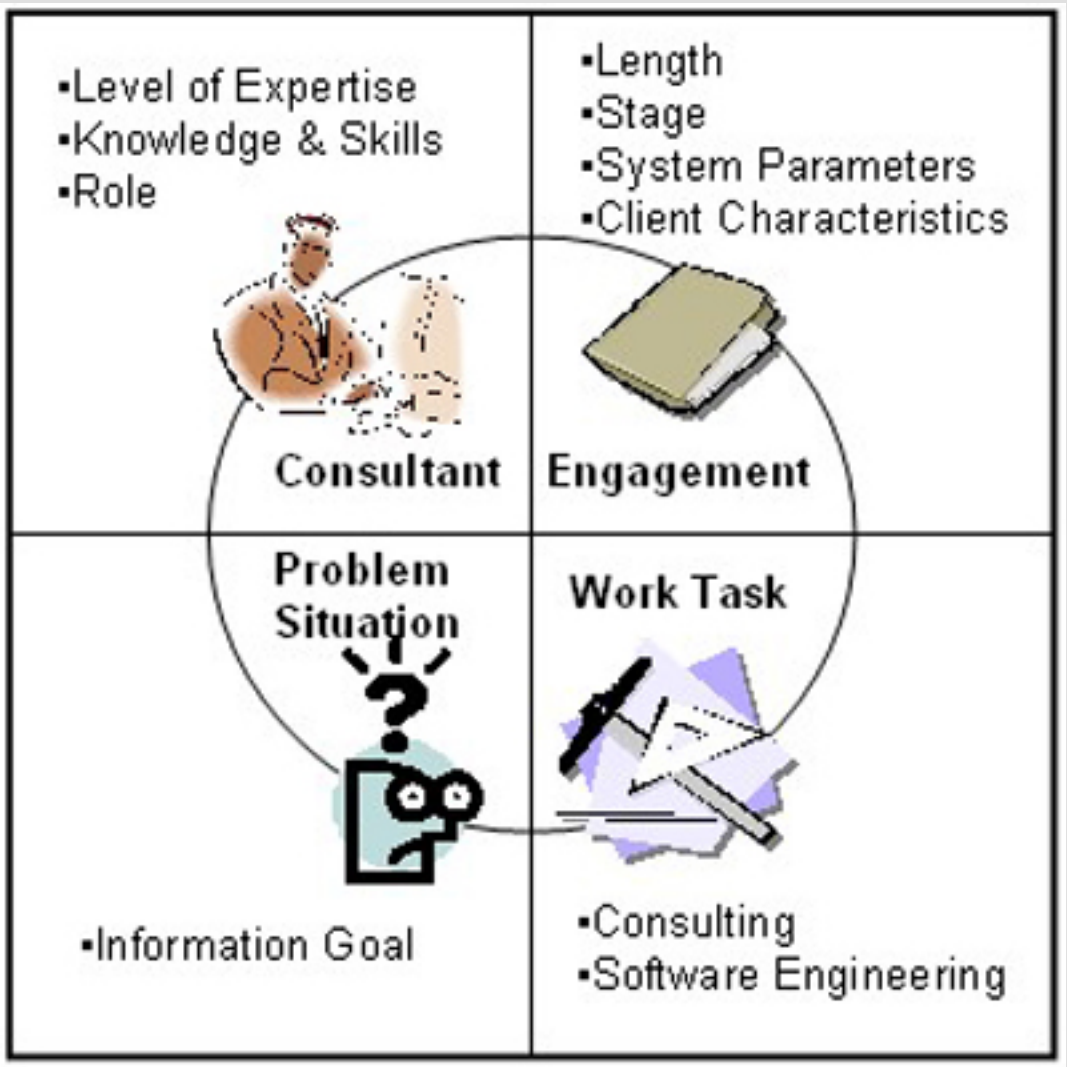


Figure 1. Spheres of Context

The main contextual spheres influencing information behaviour that emerged from the data are illustrated in Figure 1. These relate to the consultant as user who has a particular engagement (project) that involves particular work tasks, which give rise to problem situations. A more detailed discussion of each of these spheres is included below. Participants described conditions within each sphere that would influence the type of information they were looking for, and the strategies they used to obtain this type of information. They expressed the relationships between these conditions and their information strategies in a number of different ways:

- **Information characteristics: topic, level of detail, length, etc.**
- **Communication channels: people or written documents**
- **Sources of documents: Internet or intranet web sites, or repositories**

- **Document genres: Product documentation, technical articles, etc.**

Sphere 1: The Consultant

Expertise: **The level of expertise of a consultant influences the nature of the work they are assigned, the level of detail of the information that they seek, and their intended use of the information. Novice consultants tend to be assigned to work with stable technologies that are well-documented, and to do more hands-on tasks, which require procedural information. Novices "go for the product detail right away...They'll go and try to figure out the how-to "; whereas, more senior consultants "tend to understand the why" (P08). Senior consultants tend to work with products that are under development, and with tasks such as architecture, planning, and design, which require more high-level, conceptual information. They are more likely to use internal information that is not publicly available and is more difficult to access.**

Knowledge & Skills: **Both novice and expert consultants frequently encounter tasks, products, and technologies with which they are not familiar, because of the pace of change in the software domain. When this happens, they will spend more time searching for information, and will make use of overview documents: "for example if I'm not familiar with performance testing, then if I know the best practices then I can pretty much start right away" (P04). On the other hand, when consultants are on familiar ground, they are more likely to have documents in their personal collections to which they refer, and will tend to be searching for more specific and detailed information.**

Role: **The roles that consultants take on are strongly related to the underlying tasks. Typical roles are: project manager, architect, and technical specialist. Project managers will be more likely to use process-oriented documents and to documents as templates. Architects will focus on high-level conceptual and planning documents while technology specialists will use more product documentation that contains detailed technical and procedural materials, as well as specifications.**

Sphere 2: The Consulting Engagement

Engagements are consulting projects that provide services to clients. Each engagement defines a set of core contextual elements for the work of consultants: the client, the technologies, the products, and the timeframe or length of the project. In addition to these, an engagement is usually defined by the nature of the tasks involved. Tasks are dealt with separately in the next section.

Client.

A number of client parameters seem to influence information behaviour, but to a minor degree. For instance, the level of in-house IT skills of the client may influence the level and amount of information a consultant uses on an engagement. Also, consultants working with more important clients tend to have easier access to experts as information sources and to informal sources of information within the company.

Length.

Consultants tend to distinguish between very short "critical situation" type engagements and the long-term engagements that last for more than a month. Very short engagements are often associated with troubleshooting tasks. The time constraints mean that they do not use lengthy product manuals or training materials, and have no time to try out a hands-on solution. In this case, they are likely to use make use of search engines to find specific detailed solutions and published problem reports, or to call up a colleagues or request help through product support channels. For the middle-range projects, there is more time to consult with colleagues through forums and use a broader range of information materials. For long-term projects, consultants are more likely to invest time reading lengthy manuals, using training materials, and taking online courses.

Stage.

Some consultants also noted that the project stage is associated with different patterns of information use. The initial period involves more information searching, and is also the time when most background reading or learning is done. The next stage focuses more on detailed and procedural information, and the final stage is concerned more with project management tasks and associated information.

System.

The system parameters provide the framework for the technical work involved in an engagement, and as such, they are critical in determining the relevance of information. The main system parameters are: the products and versions in use, the platform, and the operating system. There are two broad characteristics of products that influence information behaviour: ownership and level of maturity. When seeking information about the company's own products, consultants will tend to rely upon product documentation produced internally. When there are third-party products involved, their main source of information is external web sites and product pages. The level of maturity of a product also influences the amount and kinds of information available. Established products are likely to be fully documented in formal public documents that are easy to find. Engagements involving new releases, "1.0 engagements", tend to be very challenging from an information perspective. In these cases, consultants tend to rely primarily upon colleagues and hands-on problem solving.

Sphere 3: Work Tasks

The interviews were designed with technical software engineering tasks in mind, but it soon became clear that tasks in this domain are two-dimensional. The consulting dimension of tasks relates to the role of the consultant vis-a vis the client and the provision of services. The software engineering dimension relates to the technical services provided in an engagement. In a given engagement, work tasks are almost

always two-sided. For example, workers may be engaged in a deployment review, a migration training session, or a hands-on product installation. Both dimensions of work tasks influence information behaviour.

Consulting Task Dimension.

The main consulting tasks that we identified in this domain are: project management, training, mentoring, reviewing, proof-of-concept, technical support, and hands-on work. In general, the consulting task dimension is associated with process-oriented documents, such as roadmaps and templates that provide the consultants with guidelines for performing this type of service. There are differences between tasks that are hands-on and thus primarily technical, and overview tasks, such as reviewing or mentoring, that involve more high-level and generic information. Training engagements tend to use the same types of materials that consultants use to self-educate: course materials, tutorials, best practices, etc. Technical support services are associated with troubleshooting and tend to use more specific and detailed information.

Software Engineering Task Dimension.

The technical dimension of work tasks seems to have a major impact on information behaviour in this domain. It represents the core professional work of software engineers, and there seems to be considerable variation in the way information is used to support the different tasks. We identified a core set of ten software engineering tasks that are common in this domain: Architecture, Design, Implementation, Deployment, Installation, Configuration, Integration, Migration, Performance tuning, and Troubleshooting. In terms of information behaviour, these tasks can be grouped into a few broad categories:

- ***High-level tasks* such as architecture and design tend to require information that is at a fairly high level of abstraction. This type of information is found in formal documents such as technical manuals and white papers. These will often include overviews, design patterns, and best practices.**

- ***Low-level technical tasks*** **such as implementation, deployment, and installation are considered to be routine - "straightforward, day to day" (P05) tasks. They are more detail-oriented and procedural and tend to be associated with specifications, product documentation, and information in the form of step-by-step instructions and guides. These tasks are well supported by internal documents; consultants tend to have a clear idea where to find this type of information.**
- ***Product scenario type tasks*** **such as upgrading, migrating, or integrating existing systems are complex in the sense that they involve multiple systems and/or components that may inter-relate in unknown or unexpected ways. These tasks tend to require detailed information on specific products with respect to particular system configurations. For these types of context-specific tasks, consultants tend to seek out information that is grounded in experience, either through colleagues or through knowledge-sharing channels. The wide range of product and technical information required for these types of tasks makes Internet search engines a common source of information.**
- ***Problem-solving tasks*** **such as troubleshooting, debugging, and performance tuning involve a high degree of uncertainty and generally involve a broader approach to searching for information. These tasks involve system-generated information such as traces and log files, and include the use of tools to test the system. Associated documents tend to be low-level technical information about very specific similar cases. Troubleshooting involves many unknowns, so it requires the kind of information you "cannot read from a book" (P04). When troubleshooting, consultants make heavy use of internal and external forums to find similar cases and to find people with experience who can provide some clues, strategies, or insight.**

Sphere 4: Problem Situation

For any given task, a consultant may need to deal with a number of problems or questions that prompt information seeking and use. Problems typically involve a specific topic, and may fall into general subject categories as well. Consultants tend to

distinguish between the following broad categories of problems: product, technology, consulting, and marketing.

Information Goals.

In addition to a subject or topic, most problems are also associated with an information goal, or intended use. In Phase 1, we identified a set of information goals that are typical in this domain, and the interviews provided more insight into how these different goals influence information behaviour.

Learn about.

When the goal is to learn about something, consultants generally turn to more high-level, conceptual information in formal genres: manuals, courses materials, and technical articles. They are often looking for brief overviews to get a quick introduction to a topic. They will often ask other more knowledgeable consultants to recommend reading materials.

Collect advice.

When consultants need to make decisions, they are most likely to turn to colleagues either directly or through forums. This is most common when the problems are related to specific scenarios and the solutions are not documented in more generic information sources. Best practices documents can be useful for this type of information goal.

Find instructions.

When the goal is to complete some defined task or process, consultants seek out low-level procedural information. They look for detailed step-by-step instructions in product

documentation and technical articles.

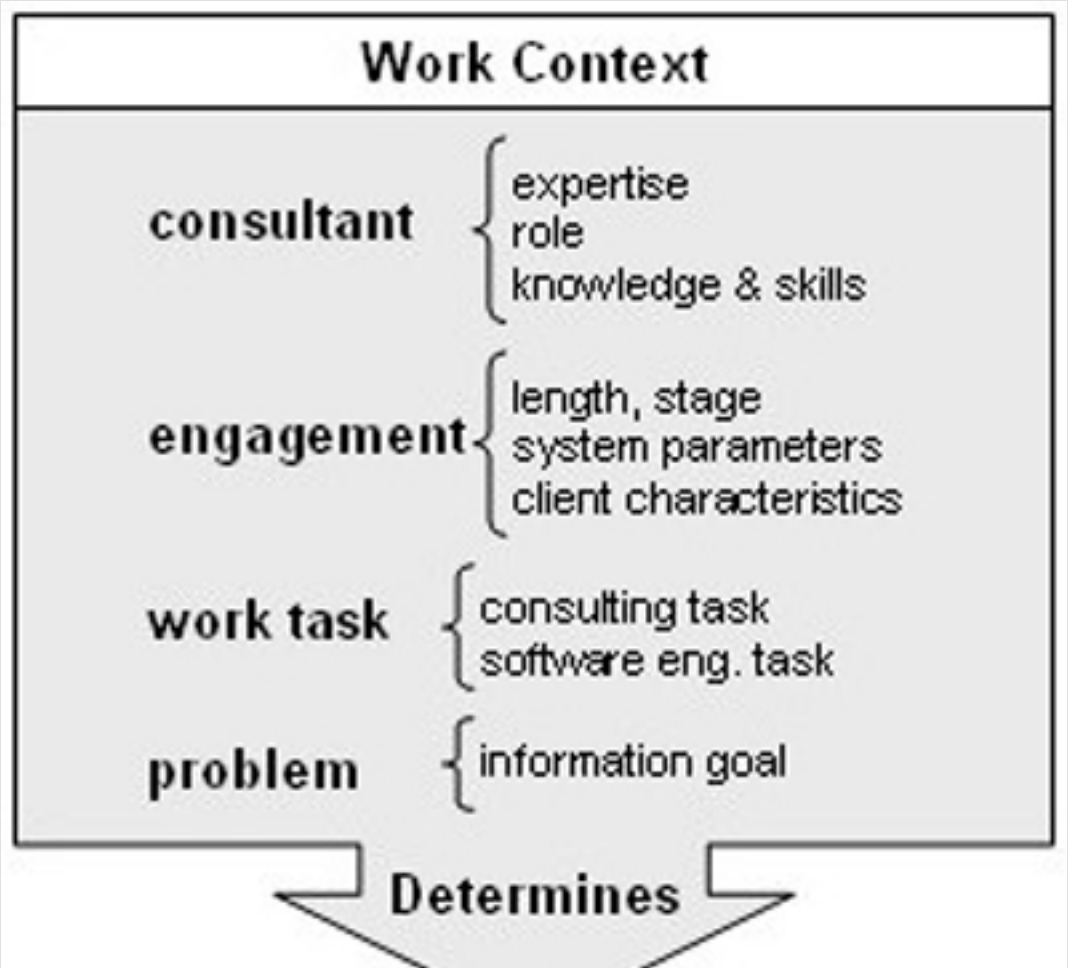
Find facts.

When the goal is to find factual information, such as technical specifications or product settings and features, consultants tend to either go to a known document or to look for an authoritative source through product documentation, internal product web sites, or in external third-party web sites.

Find examples to reuse.

When the goal is to find information to reuse, either as content or for a template, consultants often turn to internal databases that contain materials from past projects: reports, schedules, source code, etc.

Summary Model



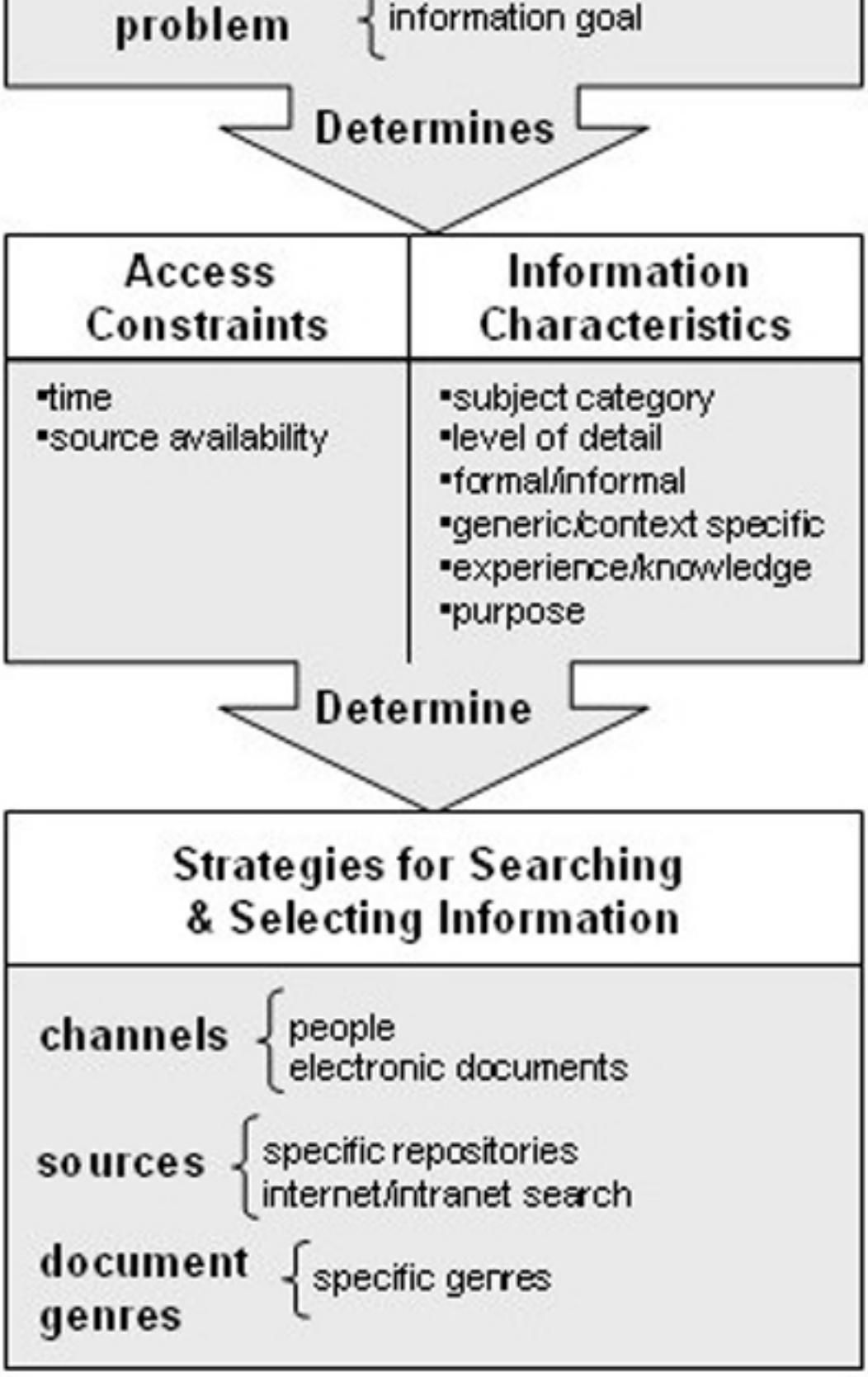


Figure 2: The Influence of Contextual Factors on Information Behaviour in the Software Services Domain

The proposed model derived from this analysis (Figure 2) indicates that conditions within the four contextual spheres determine some information access constraints as well as shaping the general characteristics of the information that is sought. The main constraints are the amount of time available for searching and using the information, and the source availability or accessibility. The information characteristics that emerged as important from this analysis are as follows:

- Broad subject groupings: product, technology, consulting;
- Level of detail: abstract or concrete
- Style: formal or informal;
- Applicability: generic or context-specific;
- Origin: from experience or knowledge;
- Purpose: educational, procedural, documentation re-use, recommending.

These constraints and characteristics, in turn, determine the strategies that consultants use to find information. In particular, they influence choices with respect to channels, sources, and genres.

Discussion and Conclusions

The information use environment within this work domain is large and complex, and participants expressed the need for information search systems that would simplify and streamline the process of finding information. Participants articulated a large number of factors related to their work that shape the decisions they make with respect to information channels, sources and types of documents. The fact that these software consultants have such a rich store of information strategies that are unrelated to the subject matter of the documents, attests to the potential value of incorporating contextual factors into information retrieval systems.

The model presented here is generally consistent with earlier findings on the information behaviour of engineers. Like others, we found clear evidence that individual abilities, work roles, tasks and problem situations drive their selection of channels and sources (Leckie *et al*., 1996; Taylor, 1991). We also found that time was a major

constraint, as was the availability of information sources. Rather than expressing availability as a criterion for selection, our participants seemed to view it primarily as a constraint - perhaps because in their networked environment, electronic access from their laptop is a precondition rather than a mere consideration for information use. There was no evidence that participants even used less accessible sources, such as libraries or print documents.

Rather than use an existing task classification, we opted to develop a set of tasks based on participants' descriptions of their work. For this reason, the task classification is content-based and specific to software engineering consultants. However, the information behaviour associated with these tasks seems to suggest that there are natural task groupings based on known task characteristics such as complexity, uncertainty, task structure, etc. (Algon, 1999). Furthermore, some of the findings support Bystom and Jarvelin's (1995) model of task complexity, in that the more complex tasks do involve use of a broader range of sources, and more reliance upon people as sources. While there was some indication from participants that information behaviour was influenced by project stage, the study did not reveal a strong process model at the task level (Vakkari, 2001). However, this is likely a function of our research design, which was not process-oriented. More work is needed to characterize the task groupings identified in this study in the framework of previous research on tasks.

The main contribution of this model is in the identification of information characteristics that provide the link between work context and information behaviour in this domain. Whereas previous research identified generic criteria that engineers use to evaluate and select information, this model identifies situated selection criteria, i.e. information characteristics that are determined by the work and task context. Arguably, these characteristics come close to expressing Hertzum's (2002) concept of task appropriateness in this domain. Although this set of characteristics was derived through the study of software engineers, the characteristics are sufficiently generic to have potential application in other areas of engineering, and perhaps other task-oriented work domains as well. An interesting potential application of these information characteristics is to identify relationships between work tasks, information goals and document genres. Genres are typically described as a function of form (language and

structure), content (semantic) and purpose (intended use) (Orlikowski & Yates, 1994), features which are closely related to our set of information characteristics. By classifying genres according to these parameters, we may be able to identify which genres best fulfill the information requirements of different tasks and goals. We are following up with further study into these relationships, as they seem to have the most potential to contribute to contextual search systems in the workplace (Freund, Toms, & Clarke, 2005).

We have described a model of the work and task factors that shape software consultants' strategies for finding information. The model outlines four spheres of work context and related conditions that influence the type of information sought and the strategies used to obtain it. The large number of relationships identified in this study suggests that robust contextual models of work domains do have the potential to contribute to improved information retrieval systems. Such systems could provide better support for employees trying to make use of an ever-growing store of digital information to accomplish work tasks.

This study is ongoing, and the next phase is already under way. We are currently conducting a questionnaire with a broader group of software services consultants, in order to validate this model. Future work will include the implementation and testing of this model in a workplace search system.

ACKNOWLEDGEMENTS

The authors thank the IBM Centre for Advanced Studies (CAS) in Toronto, which is supporting this research through a CAS Doctoral Fellowship to the first author, and an SSHRC grant to Toms, which is partially supporting the work. The authors also thank Gordon Lee of IBM for his assistance, the consultants who contributed their valuable time and expertise to this project, and the anonymous reviewers who provided valuable input.

References

Algon, J. (1999). *The effect of task on the information-related behaviors of individuals in a work-group enviroment*. Unpublished doctoral dissertation, Rutgers University, Rutgers,N.J.

Allan, J. e. (2003). Challenges in information retrieval and language modeling. *SIGIR Forum, 37*(1).

Allen, D., & Wilson, T. D. (2003). Information overload: context and causes. *The New Review of Information Behaviour Research, 4*, 31-44.

Allen, T. J. (1977). *Managing the Flow of Technology*. Cambridge, Mass.: MIT Press.

Anderson, C. J., Glassman, M., McAfee, R. B., & Pinelli, T. (2001). An investigation of factors afffecting how engineers and scientists seek information. *Journal of Engineering and Technology Management, 18*(2), 131-155.

Button, G. (2003). Studies of work in human-computer interaction. In J. Carroll (Ed.), *HCI Models, Theories and Frameworks: Toward a Multidisciplinary Science* (pp. 357-380). San Francisco: Morgan Kaufmann.

Bystrom, K., & Hansen, P. (2005). Conceptual framework for tasks in information studies. *Journal of the American Society for Information Science and Technology*, 56(?), ?

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

Dervin, B. (1999). On studying information seeking methodologically: the implications of connecting metatheory to method. *Information Processing & Management, 35*, 727-750.

Edmunds, A., & Morris, A. (2000). The problem of information overload in business organizations: a review of the literature. *International Journal of Information Management, 20*, 17-28.

Fagin, R., Kumar, R., McCurley, K. S., Novak, J., Sivakumar, D., Tomlin, J. A., *et al*. (2003). *Searching the workplace web*. Paper presented at the 2003 International World Web Conference (WWW '03).

Feldman, S. (2004). The high cost of not finding information. *KM World, 13*.

Feldman, S., & Sherman, C. (2003). *The high cost of not finding information (No. 29127)*: IDC.

Fidel, R., & Green, M. (2004). The many faces of accessibility: engineers' perception of information sources. *Information Processing & Management, 40*, 563-581.

Forward, A., & Lethbridge, T. C. (2002). *The relevance of software documentation, tools and technologies: a survey*. Paper presented at the ACM Symposium on Document Engineering.

Freund, L., Toms, E. G., & Clarke, C. L. A. (in press). Modeling task-genre relationships for IR in the workplace. *On SIGIR 2005*.

Hawking, D. (2004). *Challenges in enterprise search*. Paper presented at the Proceedings of the Australasian Database Conference, Dunedin, New Zealand.

Hertzum, M. (2002). The importance of trust in software engineers' assessment and choice of information sources. *Information and Organization, 12*(1), 1-18.

Hertzum, M., Anderson, H. H. K., Anderson, V., & Hansen, C. B. (2002). Trust in information sources: seeking information from people, documents, and virtual agents. *Interacting with Computers, 14*, 575-599.

Hertzum, 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.

Jarvelin, K., & Ingwersen, P. (2004). Information seeking research needs extensions towards tasks and technology. *Information Research, 10*(1).

Kari, J., & Savolainen, R. (2003). Towards a contextual model of information seeking on the Web. *The New Review of Information Behaviour Research, 4*.

King, D. W., Casto, J., & Jones, H. (1994). *Communication by engineers: a literature review of engineers' information needs, seeking processes, and use*. Washington, DC: Council on Library Resources.

Kuhlthau, C. C. (1993). A principle of uncertainty for information seeking. *Journal of Documentation, 49*(4), 339-355.

Kwasitsu, L. (2003). Information-seeking behaviour of design, process, and manufacturing engineers. *Library and Informaiton Science Research, 25*(4), 459-476.

Lawrence, S. (2000). Context in search. *IEEE Data Engineering Bulletin, 23*(3), 25-32.

Leckie, G., Pettigrew, K. E., & 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.

Orlikowski, W. J., & Yates, J. (1994). Genre repertoire: the structuring of communicative practices in organizations. *Administrative Science Quarterly, 39*, 541-574.

Orr, R. H. (1970). The scientist as an information processor: a conceptual model illustrated with data on variables related to library utilization. In C. E. Nelson & D. K. Pollock (Eds.), *Communication Among Scientists and Engineers* (pp. 143-189). Lexington, MA: Lexington Books.

Pinelli, T. E. (2001). Distinguishing engineers from scientists - the case for an engineering knowledge community. *Science & Technology Libraries, 21*(3-4), 131-163.

Pinelli, T. E., Bishop, A. P., Barclay, R. O., & Kennedy, J. M. (1993). The information-seeking behavior of engineers. *Encyclopedia of Library and Information Science, 52*(15), 167-201.

Saracevic, T. (1996). *Relevance reconsidered*. Paper presented at the Second Conference on Conceptions of Library and Information Science (CoLIS 2), Copenhagen, Denmark.

Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: toward a dynamic, situational definition. *Information Processing & Management, 26*(6), 755-776.

Singer, J., & Lethbridge, T. C. (1996). *Methods for studying maintenance activities*. Paper presented at the Workshop on Empirical Studies of Software Maintenance, Monterey.

Taylor, R. S. (1968). Question negotiation and information seeking in libraries. *College and Research Libraries, 29*, 178-194.

Taylor, R. S. (1986). *Value Added Processes in Information Systems: Greenwood*.

Taylor, R. S. (1991). Information use environments. *Progress in Communication Sciences, 10*, 217-255.

Vakkari, P. (2001). A theory of the task-based information retrieval process: a summary and generalization of a longitudinal study. *Journal of Documentation, 57*(1), 44-60.

Vakkari, P. (2003). Task-based information searching. *Annual Review of Information Science and Technology, 37*, 413-463.

Wilson, T. D. (2000). Human information behaviour. *Informing Science, 3*(2).