ABSTRACT

Access to the right information is a significant contributor to success in many endeavors. It is, however, difficult to characterize what constitutes right information. This is an important question for systems development projects, which continue to exhibit a sub-par track record of success. This paper describes patterns of information seeking such as nature of information sought and sources of information consulted in the context of tasks performed during systems development projects. The analysis uses task-oriented information seeking as a theoretical perspective, inferring patterns from longitudinal data collected from multiple student teams engaged in real-world systems development efforts. The results show that the nature of tasks themselves varies for routine versus innovative projects, with implications for the nature of information sought and sources consulted. Some of the counter-intuitive findings include increasing incidence of genuine decision tasks over time; and use of the web for genuine decision tasks versus people for routine tasks. Implications of the findings for practice are discussed.

Keywords

Information Needs, Systems Development, Systems Integration, Information Seeking

INTRODUCTION

In spite of significant research related to tools, methods and modeling approaches, the track record of systems development and systems integration projects continues to be sub-par (Charette 2005). The causes of failure have been described as (a) technological as well as (b) organizational (Yoon et al. 2000). With increasing research in these two directions, we have begun to understand the need for better models and methods (Wand and Weber 2002) as well as more effective change management practices (Robey and Markus 1984). With studies of ERP implementations, there is also a recognition of paradoxes associated with process redesign and implementation (Boudreau and Robey 1996; Robey and Newman 1996).

The research reported in this paper explores a third contributor to the success of systems development and integration efforts: the need for and availability of information for tasks that systems developers must perform. The inquiry is driven by the paucity of research, within the IS discipline, related to information seeking behaviors of systems development professionals. Although such inquiries have been conducted in other contexts (e.g. chief executive’s data needs (see, e.g. Rockart 1979)), similar inquiries for information systems development efforts have been lacking. The research conducted and findings reported in this paper use task-oriented information seeking as the theoretical perspective with the fundamental constructs of task, information, and information source (Hansen 2005). The choice ensures that information seeking behaviors, that is the nature of information sought and the information sources consulted, are mapped against tasks that individuals and teams perform during systems development and integration. The paper, therefore, addresses two specific research questions:

• What is the information sought by systems developers?
• What information sources are consulted by system developers?

The research method uses multiple exploratory case studies. Although the results suggest a possible correlation between information seeking behaviors and eventual success by system developers, establishing this is not the primary intent of this paper. Instead, the key contribution of the paper is an improved understanding of patterns of information seeking behaviors (information sought and sources consulted), and their mapping against the nature of tasks carried out by systems developers and integrators.
PRIOR WORK

This section reviews prior work in two streams: (a) systems development and integration efforts with an emphasis on information needs; and (b) theories of information seeking.

Systems Development and Integration

Systems development and integration, sometimes referred to as enterprise integration (EI) represents the act of designing, implementing and deploying IT-based solutions in organizational settings. Although authoritative definitions can be difficult to pin down, trends indicate that systems development efforts are moving from engineering of stand-alone applications to those requiring linking of applications to support cross-functional processes (Smith et al. 2002). They integrate different system functionalities (Lee et al. 2003) spread across multiple organizational units, in order to align processes and integrate data. Regardless of the set of labels used to describe these (Cummins and Knovel 2002; Lee et al. 2003), they represent efforts that the next-generation of systems development and integration professionals are likely to engage in (Lee et al. 2003; Purao et al. 2008).

These efforts are more complex compared to traditional systems development projects; are wider in scope, require larger outlays in terms of money, time and other resources (Sumner 1999). The complexities in these new breed of projects make them even more difficult to carry out and manage. As a result, the nature of information needed cannot be known a priori because methodologies for carrying out such projects are not readily available nor are they standardized in a manner that will allow a straightforward instantiation for multiple projects (Linthicum 2000; Schmidt 2002; Scott and Vessey 2002; Wing Lam and Shankraman 2004). Accepted wisdom in industry suggests that because of the long term nature of these efforts, they represent a journey (Schmidt 2002) that must pay attention to the interplay between the scale of the integration effort and the dynamism of the organization. Individual efforts that may take place along this journey can, therefore, encompass varying scope, multiple and different stakeholders, and a variety of tasks –, requiring significant ad hoc decisions related to the phases included in a viable plan, the nature of tasks within each phase, and responsibilities assigned to members of a team.

It is appropriate to characterize this new breed of systems development and integration projects as information-intensive. Not only do they require knowledge such as business processes, organizational data, and the technology (Linthicum 2000), but also about technology alternatives, standards and trends in the technology marketplace, and information related to how different parts of the project must be coordinated (Schmidt 2002). There is little research to understand information needs of professionals in this new era. Exceptions include works by Detlor and Freund (Detlor 2003; Freund et al. 2005) who emphasize information seeking as an activity that an individual engages in within a team context. This observation provides the context and motivation for this study: understanding and characterizing information needs of individuals and teams engaged in systems development and integration projects. Next, we turn to theories of information seeking to understand the alternatives available for conceptualizing information seeking behaviors.

Theories of Information Seeking

Several theories of information seeking have been proposed in the IR (information retrieval) stream. Table 1 provides a selective review to highlight the choices available and underpins the fundamental constructs described next.

<table>
<thead>
<tr>
<th>Theory</th>
<th>Description</th>
<th>Source(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pioneering model</td>
<td>Information need as a trigger for the overall information seeking behavior</td>
<td>(Wilson 1981)</td>
</tr>
<tr>
<td>Sense-making theory</td>
<td>Views information seeking as a sense-making process used by an individual to construct a bridge between a context and a desired situation</td>
<td>(Dervin 1983, 1992)</td>
</tr>
<tr>
<td>Extension model</td>
<td>Extension of the Pioneering model with additional triggers, intervening variables, search behaviors, and feedback based on information use</td>
<td>(Wilson 1997)</td>
</tr>
<tr>
<td>Integrated framework</td>
<td>Outlines five facets – personality, matter, energy, space, and time – as shaping information behavior</td>
<td>(Sonnenwald and Iivonen 1999)</td>
</tr>
<tr>
<td>Process-oriented model</td>
<td>Multi-stage model with stages that include: starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending</td>
<td>(Ellis 1989)</td>
</tr>
<tr>
<td>Identifying Activities</td>
<td>Activities that are part of the information seeking behaviors: initiation, selection, exploration, formulation, collection, and presentation</td>
<td>(Kuhlthau 1991)</td>
</tr>
<tr>
<td>Information retrieval model</td>
<td>Addresses the interaction between users and information retrieval systems that serve to satisfy human information needs</td>
<td>(Ingwersen 1996; Saracevic 1996; Spink 1997)</td>
</tr>
</tbody>
</table>
Table 1. A Selective Review of Theories of Information Seeking

<table>
<thead>
<tr>
<th>Theory Description</th>
<th>Source(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task-oriented information seeking (TaskInfoSeek)</td>
<td>(Hansen 2005)</td>
</tr>
<tr>
<td>Tasks at three levels: the work task triggering the information needs; the</td>
<td></td>
</tr>
<tr>
<td>information seeking tasks embedded in the work task; and the information</td>
<td></td>
</tr>
<tr>
<td>retrieval tasks that are part of the information seeking task</td>
<td></td>
</tr>
</tbody>
</table>

Of these, task-oriented information seeking (TaskInfoSeek) is appropriate for this research because of the primacy it affords the context: the work tasks performed by individuals and teams as part of the systems development and integration effort. Table 2 outlines examples showing domains in which TaskInfoSeek has been applied and highlights constructs used.

Table 2. Examples of Applications of the Task-oriented Theory of Information Seeking

<table>
<thead>
<tr>
<th>Study of</th>
<th>Description</th>
<th>Constructs</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Analysts</td>
<td>Institutional resources affect choice of information sources and hence the outcome of their activities</td>
<td>Information Sources</td>
<td>(Baldwin and Rice 1997)</td>
</tr>
<tr>
<td>Pastoral clergy</td>
<td>Information seeking affected by combinations of the organization which decides the resources available</td>
<td>Information Sources</td>
<td>(Wicks 1999)</td>
</tr>
<tr>
<td>Engineers</td>
<td>Use of internal and external sources of information in the research and development projects</td>
<td>Internal and External Sources</td>
<td>(Ellis and Haugan 1997)</td>
</tr>
<tr>
<td>Dentists</td>
<td>Different roles assumed by dentists and effect of tasks on their choice of information sources</td>
<td>Task, Information Source</td>
<td>(Landry 2006)</td>
</tr>
<tr>
<td>Planning architects</td>
<td>Tasks core to the function, type of information used and information channels (sources) consulted</td>
<td>Task, Information Type, Information source</td>
<td>(Serola 2006)</td>
</tr>
</tbody>
</table>

Fundamental constructs from TaskInfoSeek include Task, Information Types, and Information Sources. The first, Task, is conceptualized differently from the notion of task analysis (Kirwan and Ainsworth 1992) or generic task (WFMC 2008). Instead, tasks are conceptualized to include work tasks, information seeking tasks, and information retrieval tasks. The second construct, Information, is an equally all-encompassing construct that requires a conceptualization different from that prevalent in the IS discipline. A useful characterization is suggested in the IR stream by Buckland (Buckland 1991): information-as-process, information-as-knowledge and information-as-thing. The first refers to the act of informing or being informed; the second refers to meta-information, that is, it reflects what the communicated-information is about, which is also adopted in this study; the third, information-as-thing denotes the “object” that contains and communicates the information (e.g. a text document or person). For TaskInfoSeek, it is information as knowledge about something that represents the most appropriate choice. The final construct, Information Sources, roughly maps to Buckland’s (1991) conceptualization of the carrier of information. The analysis by Nilakanta and Scamell (1990) about information sources (e.g. books, periodicals, people, electronic, and others) and communication channels provides one possible classification. Byström (2002) suggests another based on resource type, e.g. people, documentary sources, and visits. The theory choice, TaskInfoSeek, and the fundamental constructs provide appropriate starting points for development of the conceptual framework.

CONCEPTUAL FRAMEWORK

The conceptual framework for the study represents an adoption and instantiation of TaskInfoSeek (Hansen 2005). It conceptualizes the information seeking task as embedded within the work task. Although TaskInfoSeek identifies the information retrieval task as a separate task, for the purpose of this work, the distinction is not emphasized. The core activity investigated, therefore, is the information-seeking task, defined as “…information seeking tasks focus on the satisfaction of an entire information need (consisting of different types of information, subject topics, etc.) through…consultations of channels and sources…” (Hansen 2005). The definition accommodates the constructs Information Type and Information Sources. The construct Work Task represents the actual activities performed by individuals and teams as part of information systems development and integration efforts. Figure 1 shows the adaptation of the TaskInfoSeek model for the study.
Figure 1. Conceptual framework for the study following TaskInfoSeek

Work Tasks and Information Seeking Tasks

A taxonomy of work tasks is adopted from related empirical studies (Bystrom and Jarvelin 1995; Kim and Soergel 2005) to ensure they represent the attributes relevant for TaskInfoSeek. Although these are similar to Campbell’s (Campbell 1988) characterization that includes decision tasks, judgment tasks, problem tasks and fuzzy tasks, the Bystrom and Jarvelin (1995) taxonomy is used because of its conceptual proximity to TaskInfoSeek. Table 3 outlines these task types.

<table>
<thead>
<tr>
<th>Task Type</th>
<th>Characteristics</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated information</td>
<td>Completely determinable outcome; automatable in principle</td>
<td>Net Salary computation</td>
</tr>
<tr>
<td>processing tasks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal information-</td>
<td>Case-based or rule-based problem-solving, Accompanies information for task</td>
<td>Tax Calculation, straightforward, but needs</td>
</tr>
<tr>
<td>processing tasks</td>
<td>completion</td>
<td>case-based analysis</td>
</tr>
<tr>
<td>Normal decision tasks</td>
<td>Routine and structured, but case-based arbitration has a major role</td>
<td>Hiring an employee</td>
</tr>
<tr>
<td>Known, genuine</td>
<td>Some sense of structure but no fixed procedure; Time required to complete task</td>
<td>Deciding about the location for a new</td>
</tr>
<tr>
<td>decision tasks</td>
<td>is not easily determinable</td>
<td>factory</td>
</tr>
<tr>
<td>Genuine decision tasks</td>
<td>Unexpected and/or unstructured tasks; accompanies minimal information; No clear</td>
<td>Collapse of a business due to external</td>
</tr>
<tr>
<td></td>
<td>idea of outcome or procedure to be followed</td>
<td>factors requiring an action</td>
</tr>
</tbody>
</table>

Table 3. Classification of Information Seeking Tasks (Bystrom and Jarvelin 1995)

Two other constructs within TOIS are elaborated next, identifying multiple alternatives for each.

Information Types

For the construct ‘Information Types,’ a possible taxonomy is suggested by Butler’s (Butler 1993) three kinds of information, namely task information (about task requirements and strategies), objective information (about the relation between performance and task demands), and normative information (about performance levels for some reference group). Bystrom and Jarvelin’s (1995) view categorizes it into problem information, domain information, and problem-solving information suggesting another alternative. Vakkari’s (2000) classification as background information, faceted background information, and specific information provides a third possibility. Each provides a possible starting point to be extended based on the data gathered via the exploratory case studies. Because task-oriented information seeking is the key focus of this study, the categories of information types adopted are tied to the definition of task information.
Information Source Types

For the construct ‘Information Source Types,’ prior work (Bystrom 2002; Bystrom and Jarvelin 1995; Nilakanta and Scamell 1990) suggests categories such as internet, intranet, people, documents, and email. The taxonomy provides an initial conceptualization subject to extension and refinement given the exploratory case-study method followed. It is described next.

RESEARCH METHODOLOGY AND DATA COLLECTION

The research follows an exploratory case study methodology (Mason 2002) with techniques that include within-case and cross-case analyses to shape interpretations (Eisenhardt 1989). The method is well-suited to the study due to its exploratory nature (Pare 2004) and well-established focus, in that the study context and phenomenon of interest are clearly laid out.

Setting

The data for the study was obtained from teams of individuals engaged in industry-sponsored systems development and integration projects. The individuals, who were students at the time, had completed one or more industry internships (i.e., they possessed at least some of the qualities of the population of professional developers such as teamwork). Data was collected from 19 such teams, each consisting of 4 to 5 team members. The data collection involved self-reported weekly status reports over the project duration. They included information about: (a) tasks performed for the project; (b) information needed for the tasks, and (c) sources consulted. The longitudinal data collection took place over a period of 10 weeks. More than 1,000 task descriptions were gathered from 19 teams; with each team contributing 55 task descriptions on an average.

Coding

Codes for classifying the information seeking tasks, information sought and information sources consulted emerged from the data, informed by prior research. Each status report was analyzed to identify individual task completion episodes. The episodes were coded (LeCompte and Schensul 1999) with task type, types of information sought and types of information sources consulted. Of the more than 1,000 episodes, 928 task episodes could be classified for task type and information source type, and 782 could be further classified for information types. The analysis used these 782 episodes. The taxonomy of information seeking tasks (see Table 3 earlier) was used to classify the tasks. The taxonomy of information types from prior research was extended, based on the data gathered, with categories such as technology, enterprise, domain, project, process, and client-team. The taxonomy of information sources from prior research was further refined, based on the data, to include sub-categories such as documents from clients, team, past experience, and external sources. Other sub-categories of the category people included self, team, client, and internal.

Analysis

For the purpose of analysis, the episodes were aggregated to the team level because of the choice of Team as the level of analysis. The analysis was conducted as a hermeneutic process, marked by two stages. First, a number of descriptive analyses were performed. This stage explored descriptions of each independent construct including longitudinal data analysis and composite sub-category analysis. Several displays were created such as longitudinal trends, composite distribution and others. These allowed the researchers to understand the data with different lenses. They also allowed frequent checks on different interpretations the researchers attempted. The second stage involved interpretations of relationship between constructs, and correlations across constructs. This analysis followed a sequential strategy – a within-case analysis strategy was followed by an across-case strategy. Each set of findings was also checked against prior literature.

FINDINGS

The findings are reported in three categories following the constructs of interest: tasks, information, and sources.

Varying Patterns of Task Distribution

The information seeking tasks were analyzed, based on the categories, across the weeks to understand their distribution over time. The data shows that over 95 percent of the tasks are accounted for by three types – normal information processing tasks, normal decision tasks, and known genuine decision tasks. Figure 2 shows the fraction of tasks for each phase (aggregated across teams) classified in these three dominant categories. The trends show a counter-intuitive outcome. As the project progressed, the number of known genuine decision tasks increased, while the number in the other two categories, normal information processing and normal decision, fell.
A possible interpretation for this result may be the nature of integration efforts. As the effort progresses, new information and linkages with existing systems may become apparent, presenting genuine decision situations. This outcome is commensurate with the idea of integration as one that requires a journey that is different from conventional, stand-alone software development effort (Castro et al. 2002; Chung et al. 1991). A fine-grain analysis of these trajectories was done by separating the phases into specific weeks.

Figure 3 shows the results as the fraction of tasks for each week (aggregated across teams) classified in the three dominant categories. With the weeks on the longitudinal scale, the graph shows that tasks in the two extreme categories (normal information processing and known genuine decisions) increased as the middle category, normal decision tasks, reduced. One possible interpretation of this finding is that as the effort progressed, student teams were able to move some semi-structured tasks requiring decisions to normal information processing tasks. Together, the distribution of tasks, across phases and across weeks, points to a progression that is different from traditional information systems development projects. Instead of starting with a set of requirements that describes an organizational or business problem (Chung et al. 1991), they point to the need for iteration (Castro et al. 2002; Wand and Weber 2002) that goes beyond just requirements gathering. They suggest that requirements for systems integration projects may not be as evident to the analyst as the project begins. Instead, they may surface as the project progresses leading to more genuine decision tasks later in the project lifecycle (Linthicum 2000).

The individual projects were also analyzed following the information seeking task categories. Four models emerged from this analysis. These were:

- **Uphill model**: displaying a rise in task count for increasing task complexity
- **Downhill model**: displaying a drop in task count for increasing task complexity
- **Hockey-stick model**: displaying a U-shape in the curve across increasing task complexity
- **Bumpy model**: displaying a bump in the curve across increasing task complexity

To further understand patterns of information seeking tasks highlighted by the four models, the projects themselves were classified as routine versus innovative (Remenyi and Heafield 1996). The routine projects were described as ones where the team knew what to do and how to do it; whereas innovative projects were described as the team attempting something that was not done before. Fifteen projects were classified by two independent raters (four were not due to lack of some details). Of these, the raters agreed on 12. The remaining 3 were jointly assessed following a discussion. Each project is described following a project type (routine or innovative) and mapped against the four models in Table 4 below.
Project Type (Characterized as Routine vs. Innovative by two raters) and Project Descriptions

<table>
<thead>
<tr>
<th>Routine</th>
<th>Innovative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Evaluation of project management software</td>
<td>2 Interface to integrate information across different websites</td>
</tr>
<tr>
<td>7 Researching business value of web service-based integration</td>
<td>3 Evaluating software and developing a small module</td>
</tr>
<tr>
<td>8 Prototype for textbook exchange website</td>
<td>11 Investigation of API for ER-Win and design a web service</td>
</tr>
<tr>
<td>14 Improving an internal application for business growth</td>
<td>13 Design project portfolio management software for a university</td>
</tr>
</tbody>
</table>

Task Distribution Model (Based on data)

<table>
<thead>
<tr>
<th>Routine</th>
<th>Innovative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downhill</td>
<td>Uphill</td>
</tr>
</tbody>
</table>

Table 4. Mapping Project Types against Models of Information Seeking

A few observations can be made from the table. First, all four routine projects follow a Downhill Model, supporting the interpretation that routine projects deal with work similar to what has been accomplished in the past and therefore, do not require many novel problem-solving tasks. Second, innovative projects accommodate different models with two dominant ones: Uphill Model and Hockey-stick Model. Both suggest a rise in genuine decision tasks as the project progresses, pointing to the need for novel problem-solving and as a result, complex information seeking tasks. Further analyses are not reported here due to space constraints.

Information Sought for Different Tasks

The second set of analyses focused on information sought, mapped against the task categories. Figure 4 shows the results. For each task category, it shows kinds of information sought. The results show some interesting correlations.

First, they show that information related to client-team interactions was sought largely during normal information processing tasks. As tasks became more demanding, the distractions of client interactions were minimized. Next, more domain information instead of project-related information was sought during known genuine decision tasks. Other information types did not clearly show a correlation with certain kinds of information seeking activities although it is interesting to note that domain-related and technology-related information were the most sought categories for genuine decision tasks. Finally, for normal information processing tasks, project-related and technology-related information was sought most often although this category largely included team coordination concerns. Treating technology-related information as problem-solving information (O’Brien and Buckley 2005), this is consistent with Vakkari’s (1999) finding that complex tasks demand more problem-solving information (Serola 2006).

Information Sources Consulted for Different Tasks
The third set of analyses focused on use of information sources. Figure 5 shows the frequency with which three dominant sources were consulted for the information seeking task types. The figure combines People and Email as one source (treating email as consultations with people at a distance); and Intranet and Documents as another source (treating intranet as a repository for electronic documents). The most interesting finding from this set of results is that known genuine decision tasks were supported by information sources on the Internet (instead of people), and normal information-processing tasks required consulting people and email. This finding, in conjunction with an earlier one (showing that innovative projects contain more known genuine decision tasks and routine projects contain more normal information-processing tasks) allows us to assert that routine projects are likely to use people and email as information sources, whereas innovative projects are likely to use Internet as a key information source.

The results extend interpretations found in prior work (Hertzum and Pejtersen 2000) that shows that people are important information sources. They, however, contradict Bystrom’s (2002) finding that People are important information sources as task complexity increases. A possible explanation is the context of the study: enterprise integration projects, which requires drawing on several individuals who contain a slice of information needed, which may be available in compiled form on the Internet.

CONCLUDING REMARKS

The results we have reported, based on the multi-case exploratory study, are a first effort aimed at understanding and characterizing information needs of teams engaged in systems development and integration. Although space restrictions prevent us from showing several details, the results highlighted show the understanding that can be gained from such studies. In spite of caveats related to self-reported data, two attributes contribute to the validity of the results. First, the study aggregates across individuals to the team level ensuring that peculiar influences from individuals are minimized. Second, the longitudinal data collection and use of theoretical perspective borrowed from the IR stream of research provides significant foundation. The outcomes include some counter-intuitive findings such as the relevance of internet as a source of information instead of users and clients for genuine decision tasks. The results have potential to not only shed light on current practice but also inform future research related to knowledge management in these settings. A second critical implication is that for education. In the absence of a clear understanding of information needs, contemporary approaches such as problem-based and project-based learning can be difficult to operationalize. The findings of this study can help to better structure such pedagogical alternatives. The contributions of this research are, therefore, two-fold. First, it shows how a longitudinal study may be conducted with self-reported data that can provide a greater understanding of information seeking behaviors in systems development teams; similar to the investigation by Rockart about chief executives’ data needs (Rockart 1979). Second, the results have the potential for direct application to practice by way of tools and techniques for enhanced knowledge management as well as better pedagogical practice. We hope that our work has provided a first contribution towards this dialog.

ACKNOWLEDGEMENTS

Acknowledgements: The work reported has been funded by the National Science Foundation under award numbers 722112 and 722141. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).
REFERENCES


51. WFMC. 2008. WFMC.org Homepage.
52. Wicks, D.A. 1999. The information-seeking behavior of pastoral clergy: A study of the interaction of their work worlds and work roles. Library and Information Science Research 21(2) 205-226.