Procedural Knowledge Search by Intelligence Analysts

Bogeum Choi
University of North Carolina at Chapel Hill
North Carolina, USA
choiboge@live.unc.edu

Sarah Casteel
University of North Carolina at Chapel Hill
North Carolina, USA
scasteel@live.unc.edu

Robert Capra
University of North Carolina at Chapel Hill
North Carolina, USA
rcapra@unc.edu

Jaime Arguello
University of North Carolina at Chapel Hill
North Carolina, USA
jarguello@unc.edu

ABSTRACT

Prior studies have explored the information-seeking practices of specific professional communities, including lawyers, physicians, engineers, recruiters, and government workers. In this research, we investigate the information-seeking practices of *intelligence analysts* (IAs) employed by a U.S. government agency. Specifically, we focus on the needs, practices, and challenges related to IAs searching for procedural knowledge using an internal system called the Tradecraft Hub (TC Hub). The TC Hub is a searchable repository of procedural knowledge documents written by agency employees. Procedural knowledge (as opposed to factual and conceptual knowledge) includes knowledge about step-by-step procedures, techniques, methods, tools, technologies, and skills, and is inherently task-oriented. We report on a survey study involving 22 IAs who routinely use the TC Hub. Our survey was designed to address four research questions. In RQ1, we investigate the types of work-related objectives that motivate IAs to search the TC Hub. In RQ2, we investigate the types of information IAs seek when they search the TC Hub. In RQ3, we investigate important relevance criteria used by IAs when judging the usefulness of information. Finally, in RQ4, we investigate the challenges faced by IAs when searching the TC Hub. Based on our findings, we discuss implications for improving and extending searchable knowledge base systems such as the TC Hub that exist in many organizations.

CCS CONCEPTS

• **Information systems → Users and interactive retrieval**.

KEYWORDS

procedural knowledge, procedural search, intelligence analysts, survey study, qualitative analysis

ACM Reference Format:
Bogeum Choi, Sarah Casteel, Robert Capra, and Jaime Arguello. 2022. Procedural Knowledge Search by Intelligence Analysts. In *Proceedings of the 2022 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '22), March 14–18, 2022, Regensburg, Germany.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3498366.3505810

CHIIR '22, March 14–18, 2022, Regensburg, Germany

ACM ISBN 978-1-4503-9186-3/22/03…$15.00
https://doi.org/10.1145/3498366.3505810

1 INTRODUCTION

People use search systems to complete a wide range of tasks, including work-related tasks performed in a professional setting. In this respect, much research in interactive information retrieval (IIR) has aimed at understanding the information-seeking practices of specific professional communities, such as patent lawyers [14, 17, 19], healthcare workers [7, 8], software engineers [10], government workers [4, 25], and recruiters [23]. When studying a specific professional community, important research questions include: (1) What higher-level work tasks are searchers trying to accomplish when gathering information? (2) What are contextual factors that may influence their needs and search strategies? (3) What types of information do people seek and why? (4) What are important criteria used to judge the usefulness of information? (5) How do people leverage existing features of the search system? (6) What challenges do searchers face and what novel features can we introduce to alleviate those challenges? Our research in this paper aims to understand the information needs and search practices of *intelligence analysts* (IAs) employed by a U.S. government agency. Specifically, we investigate the needs, practices, and challenges of IAs when searching for *procedural knowledge*.

To explain our research objectives, we must first define *procedural knowledge* and *procedural search*. In general, procedural knowledge is knowledge about how to perform a specific task or a type of task. Research in education distinguishes procedural knowledge from factual and conceptual knowledge [2]. Procedural knowledge includes knowledge about step-by-step procedures and knowledge about the tools, techniques, technologies, methods, heuristics, and skills related to a procedure [2]. It also includes knowledge about *when* to use a procedure and *why* (i.e., case-based reasoning). Procedural search relates to information-seeking tasks in which the primary objective is to acquire procedural knowledge. Naturally, a procedural search task may involve gathering information about how to execute a process in a given scenario (e.g., under certain constraints). However, procedural search tasks go beyond tasks with a "how-to" intent (e.g., "How do I do XYZ?"). Other objectives of procedural search tasks may include: (1) gathering background information about a specific technology or process; (2) comparing the features of different tools; (3) evaluating the pros and cons of a specific technology; and (4) exploring ways to improve a process.

In these examples, there may not be an immediate task to be performed. However, the goal is still to acquire procedural knowledge.

IIR researchers have studied procedural search tasks from different perspectives. Prior studies have estimated the extent to which people use search engines (especially web search engines) to gain procedural knowledge [3, 6, 29, 30]. For example, Bailey and Jiang [3] developed a taxonomy of common web search tasks and reported that searching for "how-to" information was among the top categories. Other studies have observed how people search for procedural knowledge in a laboratory setting [9, 21, 28]. For example, Urgo et al. [28] found that participants perceived procedural search tasks to involve more *creativity* than factual and conceptual search tasks. Finally, systems-oriented research has developed different tools to support procedural search tasks, including algorithms for recommending alternative procedures with the same objective [1, 22, 26], providing tips and advice [30], and suggesting queries related to subtasks of the task at hand [31].

In this paper, we report on a survey study involving 22 IAs employed by the U.S. National Security Agency (NSA). As explained in Section 3.1, NSA IAs routinely perform a wide range of analytic tasks that require gathering, analyzing, and evaluating complex information. Specifically, our survey focused on how IAs use a specific search system called the Tradecraft Hub (TC Hub) to search for procedural knowledge. That is, all participants were self-identified as users of the TC Hub. The TC Hub is an internal search system that provides access to *tradecraft documents*. In the intelligence community, tradecraft documents are known for describing "specific techniques and methods used to perform intelligence analysis" [12, p. 333]. Tradecraft documents are written and uploaded to the TC Hub by agency employees. In this respect, the system is organic and grassroots in nature. The TC Hub was developed to serve as a repository of procedural knowledge. It enables agency employees to share procedural knowledge with other employees with similar work-related needs. To illustrate, tradecraft documents may provide background information about a specific data source, tool, technology, or process; information on how to execute a procedure (e.g., a specific type of analysis); and tips and advice based on personal experience and organizational "best practices". To summarize, the TC Hub was designed to help IAs learn from each other and avoid "reinventing the wheel".

Our survey was designed to investigate four research questions:

- **Work Task Objectives (RQ1):** What types of work-related objectives are IAs trying to accomplish when searching for information in the TC Hub?
- **Information Types (RQ2):** When searching the TC Hub, what types of information are IAs trying to gather and why?
- **Relevance Criteria (RQ3):** What are important relevance criteria used by IAs when deciding whether a specific tradecraft document is useful (or not useful) in achieving their objective?
- **Challenges (RQ4):** What challenges do IAs face when searching the TC Hub for information? What are the factors that may influence these challenges?

Based on our findings, we discuss implications for designing novel search tools to support procedural search tasks within an organization. Our study focused on the use of a specific IR system (i.e., the TC Hub) by members of a specific professional community (i.e., NSA IAs). However, systems such as the TC Hub exist in *other* professional environments [11, 13, 18, 20]. Generally speaking, the TC Hub is a type of enterprise search system – intended to help a specific community of workers make decisions and complete work-related tasks based on their personal and organizational objectives [15]. Systems such as the TC Hub are an important resource in many organizations. Therefore, it is important to understand how they are being used and how they can be improved and extended to better support workers.

2 BACKGROUND AND RELATED WORK

Professional Search. Our goal in this research is to learn about the information-seeking practices of intelligence analysts. In this respect, we build on prior work aimed at understanding how people search in a professional setting to support their work-related activities. Tait [27] argued that *professional search* is unique in two respects. First, tasks are often *assigned* versus *self-generated* based on personal interest. For this reason, searchers may lack the prerequisite knowledge to search effectively and judge relevance. For example, searchers may be unfamiliar with the task domain or may not fully understand the organizational goals supported by the task outcomes. Second, professional searchers may have work-related constraints. For example, they may need to document the search process to prove "due diligence".

Prior work has sought to understand professional search in domains such as legal [14, 17, 19], human resources [23], healthcare [7, 8], government [4, 25], and engineering [10]. Several studies have investigated professional patent searchers [14, 17, 19]. Results from these studies provide insights about: (1) the types of work tasks involving patent search (e.g., trend analysis) [19], (2) the effects of the task type (e.g., novelty verification) on specific behaviors (e.g., completion time) [14], and (3) desired system features (e.g., navigation support) [17]. Russell-Rose and Chamberlain [23] investigated the information-seeking strategies of recruitment professionals. Results found a strong preference towards complex Boolean queries that evolve as the recruiter forms a "mental model of the ideal [job candidate]" [23, p. 673]. Ely et al. [7] investigated the obstacles faced by physicians when answering questions related to patient care. Results found that physicians did not pursue answers to 45% of their questions, mostly because they doubted the existence of relevant information. Ely et al. [8] reported on the types of questions physicians struggle with: (1) diagnosing patients with rare symptoms, (2) answering simple questions under *complex* constraints, and (3) determining relations between elements.

Prior work has also investigated the information-seeking practices of government workers [4, 25]. These studies analyzed work-related tasks through the lens of *apriori* determinability, a measure of task complexity rooted in uncertainty about the requirements, processes, and outcomes of the task. Byström and Järvelin [4] found that complex tasks required more information about the task domain and alternative approaches to the task. Additionally, complex tasks were more likely to involve humans as "information sources". Saastamoinen et al. [25] found that complex tasks required more information synthesized from different sources.

Freund et al. [10] studied the information-seeking practices of software engineers. Results found several trends. First, participants frequently engaged in search tasks involving procedural knowledge (e.g., troubleshooting a problem). Second, participants reported experiencing challenges related to information overload and inaccurate/obsolete information. Finally, for complex tasks, participants preferred information from people with firsthand experience.

Each of the studies above focused on professionals in a specific domain. Russell-Rose et al. [24] conducted a survey of professionals across four *different* domains in order to uncover common and uncommon trends. In terms of similarities, results found a strong preference towards Boolean search due to its transparency, reproducibility, and portability across search systems. In terms of differences, healthcare workers valued system features that improve recall (e.g., wildcard operators) and recruiters valued features that improve precision (e.g., recency-based sorting).

Defining Procedural Knowledge. Understanding how intelligence analysts search for procedural knowledge begs the question: What is procedural knowledge? To answer this question, we leverage the Anderson and Krathwohl (A&K) taxonomy [2]. The A&K taxonomy was developed to help educators *precisely* define learning objectives for students. The A&K taxonomy situates learning objectives at the intersection of two orthogonal dimensions: knowledge type and cognitive process. The taxonomy defines four knowledge types: factual, conceptual, procedural, and metacognitive knowledge. A&K define procedural knowledge as "how-to" knowledge about performing a task. In this respect, procedural knowledge involves knowledge about step-by-step procedures, algorithms, techniques, methods, heuristics, and skills. Additionally, it involves knowledge about when to use a procedure to solve a problem.

The cognitive process dimension describes the types of mental activities associated with the learning objective. The taxonomy defines six cognitive process (from simple to complex). A *remember* objective involves being able to recall information verbatim. An *understand* objective involves being able to summarize or explain. An *apply* objective involves between able to execute a process. An *analyze* objective involves being able to describe relations between elements. An *evaluate* objective involves being able to critique elements. Finally, a *create* objective involves being able to generate a novel solution or knowledge representation.

Importantly, the A&K taxonomy does not only distinguish procedural knowledge from other types of knowledge. It also provides a useful framework for understanding how procedural knowledge tasks can vary by complexity. For example, the following objectives involve procedural knowledge but vary from simple to complex: (1) memorize a procedure, (2) summarize a procedure, (3) execute a procedure, (4) identify the similarities and differences between multiple procedures, (5) evaluate multiple procedures and select the best one, and (6) generate a novel procedure. In Section 4.1, we leverage the A&K taxonomy to characterize the work task objectives participants mentioned in their survey responses.

Understanding Procedural Knowledge Search. Understanding how people search for procedural knowledge is important because people *already* use search systems to support tasks involving procedural knowledge. Völske et al. [29] analyzed one billion natural language queries (NLQs) issued to the Yandex search engine. NLQs accounted for 4% of all query traffic, and a substantial portion were queries of the form "how to [verb]". Interestingly, many "how to [verb]" queries ended with terms associated with user-specific constraints (e.g., 'do-it-yourself', 'at home') and the type of media being sought (e.g., 'images', 'videos'). Eickhoff et al. [6] analyzed queries issued to Bing over a one-month period. The authors estimated that 3% of all search sessions had *knowledge acquisition intent* involving either declarative or procedural knowledge. For procedural knowledge queries, the most characteristic n-grams included "how do", "how to", and "can I", which suggests that procedural searches often involve uncertainty about task *feasibility*. Bailey and Jiang [3] analyzed Bing search sessions and found that procedural search sessions were the 3rd longest (13 queries on average).

Prior studies have also observed how people search for procedural knowledge [9, 21, 28]. Urgo et al. [28] had participants complete learning-oriented tasks focusing on either factual, conceptual, or procedural knowledge. Participants perceived procedural knowledge tasks to involve more *creativity*. The authors noted that procedural search tasks required participants to modify procedures to fit their unique preferences and constraints. Ertl [9] investigated the effects of prior knowledge and collaboration on learning outcomes associated with procedural knowledge. Participants had better learning outcomes when they had more prior knowledge and collaborated with others. Pardi et al. [21] investigated search behaviors during web searches for procedural knowledge. Results found a strong preference for visual content.

3 METHODS

3.1 Background and Recruitment

To investigate RQ1-RQ4, we conducted a survey of intelligence analysts (IAs) employed by the U.S. National Security Agency (NSA) who are experienced users of the Tradecraft Hub (Section 3.2). To provide some background, the NSA is responsible for "global monitoring, collection, and processing of information and data for foreign intelligence and counterintelligence purposes."[1] Specifically, the NSA specializes in *signal intelligence*, defined as "intelligence derived from electronic signals and systems used by foreign targets".[2] IAs at NSA perform a wide range of analytic tasks, including: (1) identifying relevant information sources on foreign intelligence targets; (2) assessing the validity and relevance of foreign intelligence; (3) analyzing foreign target intelligence; (4) monitoring target intelligence for changes and anomalies; and (5) producing intelligence reports to support policy making.

The survey was administered with the help of research partners at NSA and the Laboratory of Analytic Sciences at North Carolina State University, a research lab funded by the U.S. Department of Defense. Our research partners assisted with recruitment and ensuring that survey responses did not contain classified information. To recruit participants, the survey was advertised on several internal mailing lists and discussion forums. The recruitment materials included a video explaining the purpose of the survey. As described Section 3.3, our survey asked participants to describe specific work-related tasks that required searching the TC Hub for information. We knew that participants were unable share classified information. Therefore, the video also instructed participants to describe their work tasks by using *analogies* from an unclassified domain (e.g., journalism) or by using generic language (Section 3.4). To further

[1] https://en.wikipedia.org/wiki/National_Security_Agency
[2] https://www.nsa.gov/about/faqs/sigint-faqs/

prevent the disclosure of classified information, responses were reviewed independently by two NSA Classification Advisory Officers (CAOs) before they were sent to us. The CAO reviewers did not modify responses. However, in a few cases, statements were redacted. Ultimately, we obtained responses from 22 IAs. Participation in the study was voluntary and participants did not receive any monetary compensation. The study was approved by NSA's Human Research Protection Program.

3.2 The Tradecraft Hub

Our survey focused on understanding procedural knowledge tasks supported by an NSA-developed search system called the Tradecraft Hub (or "TC Hub"). In this section, we describe the TC Hub.

The TC Hub is a search system for so-called "tradecraft documents". In the case of the TC Hub, tradecraft documents are written and added to the system by agency employees. Tradecraft documents can be added by any authorized user using an entry form provided by the system. Users can also edit and delete previously authored documents. Importantly, tradecraft documents are intended to support other analysts with work-related tasks. In this respect, the TC Hub is an internal repository of procedural knowledge. The TC Hub has several important features.

First, when a document is added to the TC Hub, it is categorized by the author into one of six categories. A *background* document contains background knowledge about a specific subject. For example, it may provide an overview of an existing tool or technology (e.g., historical origins, applications, and strengths/weaknesses). Similarly, a *definition* document is intended to define a concept, method, or technology/tool. It is intended to be more concise than a *background* document. A *how-to* document is intended to describe how to perform a specific task or process. It should provide step-by-step instructions on how to complete the task and may also discuss inputs, outputs, and ways to interpret the outputs. A *lessons learned* document is intended to describe experiential (i.e., first-hand) knowledge about a specific technique, tool, or resource (e.g., successes, failures, and recommendations). A *critical review* document is a commentary on another tradecraft document. It should describe ways in which an existing document could be improved or extended. Finally, an *operational document* is intended to focus on procedural knowledge that is highly specific to an individual's organizational context. In contrast to a *how-to* document, an *operational document* may not be as generalizable. Importantly, when searchers query the system, they can filter search results based on these categories. However, there is no central authority overseeing the categorization process. Therefore, searchers may disagree with how a document is categorized.

Second, the system enables users to tag documents using any keywords of their choice. This social-tagging feature has potential benefits and shortcomings. On one hand, social tags have the potential to make documents more discoverable. On the other hand, the taxonomy is not centrally controlled. Therefore, social tags can be ambiguous and redundant with other tags. Third, searchers can "like" documents and some pre-authorized users can "endorse" documents on behalf of their internal organization.

Finally, from a search perspective, users can search the TC Hub in several ways. Searchers can issue Boolean and unstructured queries. Additionally, searchers can filter results along different dimensions, including: (1) publication date, (2) document type (e.g., background, definition, how-to, etc.), (3) author, and (4) social tags. Also, the system includes a "more like this" feature that allows searchers to find similar tradecraft documents.

3.3 Survey Design

Our survey asked three "general questions" about the TC Hub.

- **Q1:** What do you like about the TC Hub? Why?
- **Q2:** What do you dislike about the TC Hub? Why?
- **Q3:** What challenges do you encounter when using the TC Hub?

Additionally, the survey asked participants to describe two instances in which they used to TC Hub to find information. Participants were asked to recall and answer questions about one *positive* and one *negative* experience. In response to each experience, participants were asked to answer the same ten questions.

- **Q4:** What were you looking for?
- **Q5:** Why were you looking for this information?
- **Q6:** What did you already know about the topic?
- **Q7:** What knowledge did you use to support your search process?
- **Q8:** Did you find what you were looking for? Please describe.
- **Q9:** How much did you already know about the TC Hub?
- **Q10:** What features of the TC Hub did you use?
- **Q11:** What steps did you take?
- **Q12:** Did you encounter any difficulties? If so, please describe.
- **Q13:** What affected your search (e.g., helped, hindered)?

3.4 Using Analogies and Generic Descriptions

Our survey asked participants to describe two work-related tasks. For each task, participants were asked about the task itself (Q5), their prior knowledge (Q6), their approach to finding information (Q10), and any challenges encountered (Q13). Participants had to address these questions without disclosing classified information. This posed an interesting challenge—How does one learn about the practices and needs of a community that cannot share precise details about their tasks? To address this challenge, we instructed participants to use analogies and/or generic descriptions. Both strategies were pilot-tested with "live" interviews with IAs. Participants were instructed on how to use analogies and/or generic descriptions in a video they were asked to watch before responding to our survey. Ultimately, both strategies were successful and are therefore a methodological contribution.

To explain the use of analogies, the video instructed participants to *"imagine a scenario and information need similar to yours from an unclassified domain such as investigative journalism, genealogy, finance, home repair, DIY projects, or cooking."* The video included an example from journalism: *"I had an audio recording that incriminates a popular CEO, but I didn't know if it was legitimate. I had other recordings of the same person to compare it against. I wanted to find different methods for authenticating the recording, understand their pros and cons, and choose the most reliable one."* The video explained that this analogy would enable us to know that the task involved: (1) finding alternative solutions to a problem, (2) comparing the alternatives, and (3) selecting the best one based on specific criteria.

Overall, participants were able to use analogies to describe their tasks and scenarios in insightful ways. For example, P3 stated:

"Let's say my problem is fixing a leaky faucet. I don't want to see all products that aim to fix leaky faucets. I'd like to see reviews by homeowners who haven't redone a bathroom themselves but want [to] fix the problem without having a lot of plumbing knowledge." In this case, the participant wanted to evaluate different ways to fix a problem based on specific criteria (i.e., appropriateness for a novice). Similarly, P23 stated: *"[I wanted to] learn how to select a ripe banana at the grocery store. However, searching on 'ripe banana' gets you articles about apples [...] the expression 'ripe banana' does not appear but plenty articles have 'ripe apple".* In this case, the participant wanted to know how to determine a latent state (i.e., ripeness) based on observable evidence (i.e., the outside of a banana). Additionally, the response illustrates the challenge of using query-terms that frequently appear in a non-relevant context.

Participants were also able to use generic descriptions to provide insightful answers. For example, P23 stated: *"A colleague asked me how to do a particular procedure [using] a tool. I hadn't done it in a while, so I couldn't remember the exact steps."* In this case, the participant wanted to find out how to execute a procedure using a specific tool. Additionally, the participant had prior knowledge that the task was feasible. Another participant (P19) stated: *"I needed to know specifically how [a] piece of technology interacted with the telecommunication system."* In this case, the participant wanted to know how using a specific tool might affect other factors.

In all of the above examples, responses lack precise details. However, they provide insights about: (1) the cognitive processes associated with the task; (2) the different components of the task (e.g., data sources, tools, technologies, processes, etc.); (3) important bits of prior knowledge (e.g., task feasibility); (4) relevance criteria (e.g., simplicity); and (5) challenges.

3.5 Data analysis

To investigate RQ1-RQ4, we conducted a qualitative analysis of our survey data. First, all authors reviewed a subset of the data to agree on the dimensions we wanted to code. We determined four dimensions associated with our RQs: (1) work task objectives, (2) information types, (3) relevance criteria, and (4) challenges. Next, two of the authors each coded 50% of the data and focused on extracting descriptions of the work-task objectives, types of information sought, and relevance criteria that could be used as dimensions for faceted filtering. The analysis for RQ4 (i.e., challenges) required more conceptual abstraction. Thus, the two authors met several times to review each other's codes and consolidate them into a set of themes. Finally, all authors met to discuss the most salient themes and their definitions.

4 RESULTS

4.1 RQ1: Work Task Objectives

Our survey asked participants to describe two work-related tasks that motivated them to search the TC Hub. To better understand the higher-level goals of using the TC Hub at work, we classified work-task objectives along two dimensions: (1) the cognitive process and (2) the artifact involved in the work task. Essentially, we analyzed work tasks by considering the primary "verb" and "noun" used to described the task. In Figure 1, we present the different types of work task objectives observed in our data.

The "verb" was used to classify work tasks into a specific cognitive process from the A&K taxonomy (Section 2). The cognitive process of a task is related to the primary mental activity. Work-task objectives were classified into the cognitive processes of *understand*, *apply*, *analyze*, *evaluate*, and *create*. Since our participants focused on tasks that involved acquiring knowledge they could act upon, we did not observe any *remember*-level tasks (i.e., memorize information).

The "noun" was used to characterize tasks based on the primary *artifact* associated with the task. Our "noun" categories were derived from the data itself. Our participants described work tasks associated with four types of artifacts: data, tools, technology, and process/method. Some work tasks were associated with a specific *data* source used in an analytic task. Other tasks involved a specific tool or technology. A *tool* is a specific piece software (e.g., the R stats package) and a *technology* is a tool category (e.g., machine learning). Tasks also involved *processes and methods*, which may include one or more data sources, tools, and/or technologies. A process/method might involve a specific type of analysis or work-related procedure (e.g., accessing a data source). Finally, some work tasks did not involve a specific data source, tool, technology, or process, and were more exploratory in nature (e.g., "I wanted to find articles to create learning materials for new employees."). Therefore, we also included a *general* category. To illustrate our task classification process, consider a task that involved "comparing the features of two software tools". This task would be classified as '*analyze/tool*' because it involves analyzing the similarities/differences between tools. Of course, tasks may involve more than one cognitive process. For example, *analyzing* the similarities/differences between tools may also require *understanding* each tool in isolation. In such cases, tasks were assigned to the most complex process applicable, related to the ultimate objective of the task.

Figure 1 shows the different types of work task objectives described by participants. Each task is described in generic terms and is situated at the intersection of a cognitive process and artifact. The values in parentheses indicate the number of tasks of each type. In total, we analyzed 44 work tasks. However, some work tasks involved multiple goals. Therefore, the numbers in parentheses sum to 52. The results in Figure 1 show three important trends.

First and foremost, we observed a *wide* range of work task objectives. In terms of cognitive process, we expected most tasks to be associated with the cognitive process of *apply*—using procedural knowledge to execute a procedure (e.g., perform an analysis). However, tasks also had objectives that primarily involved understanding, analyzing, evaluating, and creating. In terms of artifacts, most tasks involved a process/method. However, we also observed tasks primarily associated with a specific data source, tool, or technology, as well as tasks that were exploratory in nature (i.e., general).

Second, most work tasks were related to the cognitive processes of *understand* and *evaluate*. In terms of *understand*, some tasks focused on understanding the general purpose of a data source, tool, or technology. Other tasks focused on understanding the functionality of a tool and the inputs/outputs of a process. In terms of *evaluate*, our analysis points to important criteria used by IAs when evaluating alternatives. As expected, participants evaluated alternatives based on their effectiveness, reliability, applicability, and suitability for a specific audience (e.g., domain novices). However,

	Data (7)	Tool (10)	Technology (6)	Process/Method (21)	General (8)
Understand (17)	Learn about a dataset. (2)	Understand the purpose of a tool. (3) Understand the methods implemented by a tool. (1) Learn how to use a tool. (1) Learn about new tools being used for a specific purpose. (1)	Understand the purpose of technology. (1)	Learn about novel ways to execute a process. (1) Understand the inputs, outputs, and variables of a process. (1)	Find useful articles for a specific audience (e.g., new employees). (3) Refresh understanding of a topic. (3)
Apply (12)		Determine how to use a specific tool to solve a problem. (1)		Determine how to execute a process. (9) Determine how to execute a process and interpret the outcomes. (1) Determine how to access a dataset. (1)	
Analyze (6)	Analyze the relations between data sources. (1)	Compare two tools. (2)	Find similar technologies that serve the same purpose. (1) Understand the effects of using a technology in a specific context. (1)	Determine whether one method can replace another. (1)	
Evaluate (13)	Evaluate the credibility of a data source. (1) Evaluate the usefulness and pitfalls of a data source. (1) Determine the best uses of a data source. (1)	Evaluate the suitability of tools for a novice. (1)	Determine the popularity of a technology. (1) Evaluate the continued use of a technology. (1) Evaluate the suitability of a technology to analyze a dataset. (1)	Determine the root cause of a problem. (3) Evaluate the best way to execute a process. (2) Evaluate whether a process represents "standard practice". (1)	
Create (4)	Create an analysis workflow using a dataset. (1)			Create an educational "use case" for a specific technique. (1)	Create learning materials for a topic. (2)

Figure 1: Work task objectives organized by cognitive process and artifact.

other participants mentioned more nuanced evaluation criteria, such as the extent to which an alternative is a "popular choice" or is "standard practice". Finally, some tasks had evaluation criteria that were open-ended. These included tasks to learn about the benefits, drawbacks, and best uses of a data source, tool, technology, or process, as well as tasks to determine whether one alternative is a suitable replacement for another.

Finally, our analysis suggests that work task objectives follow Zipf's law. That is, a few tasks were fairly common and most tasks were rare (observed only once). This is perhaps not surprising. It is common knowledge that search tasks and queries follow Zipf's law [16]. However, this points to an important challenge for systems like the TC Hub—they need to support a wide range of uncommon tasks and scenarios. The most common task type involved "determining how to execute a process", classified as *apply*/*process*. However, we also observed unexpected and more nuanced task objectives. For instance, we observed tasks that involve *creating* as the primary cognitive process (e.g., designing workflows, use cases, and educational materials).

4.2 RQ2: Information Types

As described above, our survey asked participants to recall two work-related tasks that motivated them to search the TC Hub for information. For each work-related task, they were also asked to describe the type of information they were looking for and why they sought this information (i.e., for what purpose). We identified five main types of information sought by participants: (1) background information, (2) term definitions, (3) procedure applicability, (4) detailed step-by-step information, and (5) advice.

Background information. Many participants (n=16) mentioned that they needed background information. Background information refers to information that enables someone to gain a high-level understanding of a topic (e.g., the historical context of an analytic process). Some participants sought background information because they wanted to stay knowledgeable about different aspects of a topic. For example, P10 said: *"General background information about the topic would be helpful. To use the making breakfast analogy, a history or background on why humans generally prefer a meal first thing in the morning."* Using this analogy, P10 searched for tools, techniques, and analytic procedures related to "making breakfast".

Other participants sought a different type of background information, namely *specialized* background information. Participants described work-related tasks that had highly specific requirements and constraints (e.g., they needed to use a specific tool, technology, or data source). For this reason, many participants sought background information about those specific elements of the task in the context of their specific task scenario or situation. For instance, P21 said: *"Initially I was searching for background information about the specific Agency capability/dataset. I was not searching for broad background on the general topic as a whole. Going back to my police* *officer/license plate example, I was searching the Hub for how to access and search through license plate database information, which is specifically only available to the Agency."* Similarly, P12 was surveying use cases for a specific analysis and stated: *"I was looking for background information about this analytic. I did not know how it worked, what the output looked like, and it was important to me to know exactly what data was used and how."* In these cases, participants still sought background information to enhance their high-level understanding of a topic. However, they also had specific situational constraints that needed to be accounted for in their search.

Term Definitions. A few participants (n=6) mentioned that they wanted definitions for key concepts or variables involved in a specific analytic task. This type of information seemed to be more important for domain novices. For instance, P12 said: *"I did not know anything about the inputs or outputs of this analytic prior to researching on the hub... I wanted definitions of any non-obvious key terms that would be found within the output."* The need for term definitions was closely related to three important trends: (1) the heavy use of technical jargon in TC Hub articles, (2) the lack of articles on foundational topics, and (3) the fact that different organizational entities use different terminology. For example, P1 said: *"There is a lot of very technical content, with very specific technical terminology."* Similarly, P19 noted an overall lack of "quick" definitions in tradecraft documents: *"There are no definitions, and I've heard that a lot from people I train."* Additionally, P14 pointed out that different groups use different terminology, which compounds the issue: *"Changing terms/acronyms or where two different groups have come up with [different] names for the same thing."*

Procedure Applicability. Several participants (n=11) described needing information that describes the *contexts* in which a procedure is applied. This type of information can be useful when someone is trying to evaluate the fit between a given procedural solution and their unique circumstances: When/where can I apply procedure X? How does procedure X function in a specific context Y? For instance, P1 looked for information about *"how the analytic methods fit into an overall analysis workflow"*. They wanted to understand *"the why of using the tool"* in the context of a specific mission. Additionally, P9 wanted information about *"what tasks could be accomplished with the tool"*. Lastly, P19 needed to know *"how a specific piece of technology interacted with the telecommunication system"*. In all these cases, participants wanted information about the applicability (e.g., effectiveness or unintended outcomes) of a procedural solution in a specific context.

Detailed steps. Participants (n=12) also mentioned needing detailed steps. Not surprisingly, participants looked for how-to instructions to help them perform a given task. For example, P12 said *"I wanted details about how to go through steps to accomplish the task. I was looking for an explanation of all the data that is incorporated in the analysis...I needed this information to test this specific analytic to create a use case for using this as part of a larger effort [agency-level mission]."* More interestingly, some participants mentioned wanting to know the *rationale* behind specific steps. In other words, participants wanted to know not only how to execute the steps of a procedure but also *why* each step is important. For example, P4 said: *"[I needed] general background information accompanied by the logic of which steps to take and why, rather than just step-by-step instructions [about] particular tools."* Participants wanted step-by-step instructions on how to execute a procedure and interpret the outcomes. They also wanted to understand the logic behind the steps, which may enable someone to *modify* the procedure.

Advice. Several participants (n=7) mentioned wanting advice from other people familiar to a topic. For example, P5 wanted to diagnose a problem when they only had some hypotheses about the root cause. Specifically, P5 stated: *"I was looking for any experts who have experience with my hypothesis and symptoms."* In this case, the participant essentially wanted to consult someone who has already dealt with the problem and had "similar symptoms". Some participants also expressed their appreciation for the grassroots nature of the TC Hub, which enables learning from other people's experiences and expertise. P10 said: *"[I needed] advice from others often, because a key feature of Tradecraft Hub... [is] learning from others."*

4.3 RQ3: Relevance Criteria

In RQ3, we explore the criteria our participants used to determine document relevance. Based on their responses, we identified five main categories of relevance criteria: (1) intended audience, (2) level of details, (3) specificity vs. generalizability, (4) task constraints, and (5) authorship information.

Intended audience. Participants often encountered difficulties in their searches due to assumptions about the intended audience of an article. For example, P1 said: *"[the authors] used specialized terminology and made a lot of assumptions about the reader's existing knowledge."* P19 echoed this feeling about prior knowledge: *"the [articles] were written by people in the know, for people in the know, and not for people learning the topic."*.

Participants also understood the challenges involved in writing articles that can be understood by a wide audience. For example, P15 empathized: *"[People] who write [articles] do genuinely try to pass on good information. Just not all of them are 'teachers', or fully consider the variety/range of their consumers (trainees)."*

These results suggest that search systems should provide mechanisms to filter results based on the intended audience and prior knowledge needed. One approach could encourage authors to include metadata about the intended audience when they submit an article. Another approach would be for the system to provide faceted filters based on reading level, type/amount of specialized terminology, or estimates of the prior knowledge needed.

Level of details. Another relevance criteria participants discussed was the level of details in a document. This manifested along three dimensions: (1) the amount of content in a document, (2) the scope of a document, and (3) whether a document had all the information needed for a task versus being just "one piece of the puzzle". A common distinction mentioned was whether the document broadly described a process (e.g., gave an overview of it), or whether it gave specific details about how to execute the procedure (e.g., step-by-step instructions).

Another distinction was whether the document was self-contained or not. Participants wanted to know if the document provided enough information about a procedure that they would not need to do additional searches. P12 described this in terms of completeness: *"[Hub articles] are not always complete... The information was in the*

Hub, just not... condensed together... I had to do a lot of additional searching."

These results suggest that systems should provide features that allow users to filter the search results based on the scope of a document and level of details provided.

Specificity vs. Generalizability. Documents in the TC Hub exist along a continuum from specific to general in focus. For example, a highly specific document might describe how to connect a Microsoft Surface Pro 3 computer to a Dell U3219Q 4K monitor. Conversely, a more generalizable document might give an overview of how to connect any laptop to any monitor (e.g., connection types, converter cables, software settings, etc.).

Our participants described wanting to be able to find/filter documents based on their level of specificity vs. generalizability. For example, P2 described needing more generalizable documents: *"The best article I could find... looked at the data format in a more narrow context than I was hoping... Many of the articles are for much more specific tasks than what I was looking for in this case"*. Similarly, P4 described: *"I was looking for an overview article... and found only articles about how to apply these techniques [using] particular tools"*.

These results suggest an opportunity for search systems to provide filters that allow participants to indicate the level of specificity or generalizability of documents they wish to retrieve. Future work could consider how to train classifiers for this facet.

Task requirements/constraints. Participants also described that their tasks often had specific constraints that were important relevance criteria during the search process. For example, a common constraint was that they only had certain tools available. These constraints impacted participants' search interactions with both SERPs and landing pages (i.e. the tradecraft documents). On SERPs, users described situations where it would be helpful to be able to include their constraints as part of their search. On landing pages, participants noted that they often had to skim through large amounts of text to determine if the procedure met their unique constraints. For example, P2 described: *"...most articles don't list constraints on what is needed/assumed for a technique to work."*.

Search systems could assist users by algorithmically extracting prerequisites, inputs, tools, and techniques that are described in the articles (e.g., this recipe involves braising, finely chopping, and reducing). These extracted constraints could then be provided on the SERP as faceted filters or could be highlighted on landing pages (e.g., in a sidebar) so that users can more quickly determine if the document is relevant based on their unique constraints.

Author information. Participants also described using author information to help determine document relevance. For example, P3 said: *"...there are mental shortcuts to using the TCH. For example, [name] has a great reputation, I'll read her [articles] first."* Similarly, P23 described: *"I knew what office was officially responsible for this software, so I knew which authors to look for as the authority."*

4.4 RQ4: Challenges and Desired Features

RQ4 examines the challenges that participants faced and desired features they described. We identified four main challenges: (1) having to "wade through" a lot of information, (2) vocabulary problems, (3) information quality and redundancy, and (4) gaps and category mismatch. We also identified two desired features commonly requested by participants: (5) identifying similar/related concepts and (6) providing explanations for specific search results.

Wading through lots of information. One of the main challenges reported by participants was having to "wade through" many documents (some irrelevant and of poor quality) to identify a document relevant to their specific needs. Participants mentioned wanting to have better ways to review the results on SERPs to make this process more efficient. For example, P22 stated: *"When you do get lots of results, I would like to have them better categorized."*.

Participants also described difficulties parsing information-dense documents to locate the specific information relevant to their needs. P23 gave the following example: *"Say you want to learn how to select a ripe banana... Buried in a 12-page article on produce procurement, there's a section on bananas, with a subsection on selecting them at the correct stage of ripeness."* P22 attributed these issues to a lack of standard document structure: *"there's no standard way of writing an article, which can result in lots of time wading through [information]."*

Since TC Hub documents often contain dense and lengthy text, it would be beneficial to provide ways to help users scan and navigate within a document. For example, the system could highlight relevant sections of a document with respect to the query, or provide an automatically generated table of contents. To help standardize document structure, the system could also provide templates for authors to use when writing a specific type of TC Hub article.

Vocabulary problems. Another challenge involved vocabulary problems including: (1) not knowing what search terms to use, (2) not comprehending technical jargon, and (3) variations in vocabulary usage. First, participants described situations where they did not know how to articulate their needs. For example, P6 described: *"It's difficult to get answers for the 'unknown unknowns'... If I'm searching for [auto] maintenance articles do I search for 'rotors' or 'discs'? Is it a 'serpentine belt' or an 'auxiliary belt'"'?*

Second, participants described challenges understanding documents that contained too much jargon or specialized vocabulary. P13 even suggested that the system should discourage the overuse of jargon at the document creation stage: *"It would be nice if the Hub prompted writers to simplify their jargon to make articles understandable to new employees."*

Third, participants discussed challenges related to polysemy and synonymy. For example, P15 described: *"Depending on job role... the same words don't mean the same thing, but... there aren't better words"* P15 also noted that: *"I use different terminology than the ones used in the article."*

Out-of-date information and redundancy. Participants also noted challenges in dealing with information that was out-of-date or redundant. For example, P21 noted: *"For every good article there's at least five articles that contain inaccurate (or obsolete) information, and even endorsed articles can sometimes be misleading."* As a possible solution, P25 suggested that the system should allow users to provide feedback (i.e., flagging): *"I'd love to see more ways for readers to give specific quantifiable feedback... to flag questions or out-of-date material. For example, 'XXX software has been replaced by YYY software'."* Similarly, Freund et al. [10] found that software engineers faced challenges with filtering obsolete information.

Participants also mentioned issues related to redundancy in documents, especially ones written on common topics. For example, P25 noted: *"When you seek to generate an article... [there should be] some smart code in the background to show you articles similar, [so that] we have fewer versions of pretty much the same information."*

These results highlight users' needs to determine not just *if* information exists in the TC Hub, but if *current, up-to-date* information exists. This is a common issue with procedural knowledge in rapidly changing domains such as information technology. Search systems could help users filter results based on date modified and flag articles that contain outdated or redundant information.

Gaps and categorization mismatch. Another challenge participants noted were gaps and inconsistent categorization of articles in the TC Hub. These highlight several challenges in supporting search in small- to medium-sized knowledge bases. First, the knowledge base may be missing articles on important topics. For example, P2 observed: *"There are some important foundational topics that no one has written an article about yet that really should exist. Think of it like having a cookbook for making pies that doesn't include any recipes for making a pie crust."* The search system could help identify missing information by tracking common queries that result in few results, no clicks, or quick reformulations. The system could also recommend people who might have knowledge in a missing topic based on their previous searches or articles they authored.

Second, when users are unsure if particular information exists in the knowledge base, search becomes a more difficult process. Users may not be sure if a lack of relevant results is due to missing content or problems with their query formulation. For example, P5 described: *"I was not confident in my initial research as I knew it was a newer technology and would likely not have been published yet. I was more confident in searching for similar technologies."*

Third, although every document in the TC Hub has a document category assigned by its author (e.g., background, how-to), participants noted inconsistencies in how documents are categorized. For example, P19 noted: *"If you're searching for background and I really need background, but they put it in how-to, you'll never find it."*.

Together, these challenges illustrate the importance of including features to help users more easily determine when the information they seek does not exist in the knowledge base. Additionally, the system could warn authors when their classification of a new document seems unusual. A simple approach might leverage machine learning—training classifiers to predict whether a new document has been categorized appropriately (i.e., consistent with a gold-standard dataset of preclassified articles).

Pointing out similar/related concepts. The TC Hub includes features to create links between articles and to recommend "Related articles". While participants appreciated and used these features, they also expressed a desire for more ways to connect with related material. For example, P1 observed: *"It would be useful if searches turned up conceptually similar content so I didn't have to make guesses about what terms would turn up the content I'm looking for"*.

Explanations of why results are shown. Participants also described wanting explanations about how and why related items were included in the search results. For example, P6 noted: *"I had just seen this 'new' data in my results and I curious to learn more about it. Why was is showing up in my results? Can I use it in conjunction with other data?"*.

These results suggest possibilities for search systems to help users gain a better contextual overview of how various concepts, technologies, and tools relate to the topic of their search. For example, visualizations to show process maps and concept maps could help users understand the broader context of their search.

5 DISCUSSION AND DESIGN IMPLICATIONS

In this section, we discuss how our results have implications for designing systems to support searches for procedural knowledge.

SERP filtering by different types of complexity. Document complexity was an important relevance criterion for our participants. We identified three important dimensions of procedural knowledge complexity: (1) target audience, (2) level of detail, and (3) generalizability. *Target audience* refers to the intended audience for which a document was written (e.g. novices vs. experts). *Level of detail* refers to the extent to which the information in a document is self-contained and can be used to solve end-to-end problems. *Generalizability* refers to the extent to which the information in a document can be generalized across different tasks. These dimensions play important roles in helping users determine document relevance during procedural search tasks. Future search systems should explore ways to allow users to filter results using these facets. To classify documents along these dimensions, a system could leverage user-provided annotations or machine-learned classifiers based on lexical features.

Document-level facets to highlight information within a document. Our participants discussed having difficulties "wading through" long articles to determine if they had the specific type of procedural information they needed. We envision *document-level facets* that could be displayed in a sidebar and show key concepts, terms, and constraints mentioned in the article. Clicking on one could highlight those relevant passages in the document. This type of interface could help searchers quickly determine the *applicability* of a procedural document based on their unique requirements and constraints (e.g., data sources and tools available).

Result overviews and common terms. Participants also described difficulties in getting an overview of new, unfamiliar domains. Participants reported taking explicit search steps to discover important concepts, terms, and vocabulary. Procedural knowledge search interfaces could aid users by providing displays of concepts and terms related to the current query at the top of the SERP (i.e. going beyond simple query suggestions). The display would provide novice searchers an important overview of the domain and vocabulary. The concepts/terms could also be clickable to allow participants to explore them in more depth (e.g. show background articles about that concept/term).

Procedural similarity. Our participants described needs to identify information with procedural similarities to their current task. For example, when learning to bake pastries, it might be helpful to see related documents about baking cakes. Participants also expressed needing to find procedural knowledge about tools and techniques similar to ones that they were already familiar with. For example, if I know how to make meringue, what other things can I do that use a similar technique (or that build on this technique). These scenarios highlight needs for search systems that can incorporate aspects of *procedural similarity.* Whereas traditional document similarity measures are often grounded in lexical and semantic similarity, procedural similarity measures are needed to help retrieve procedural documents that involve similar inputs, requirements, steps, techniques, processes, and skills. Prior research has developed algorithms to predict procedural similarity and may provide a starting point [1, 22, 26].

Features to help assess if information actually exists in the repository. Participants also reported sometimes having difficulties determining if the information they sought *existed* in the TC Hub. Search systems could address this issue through several approaches. First, prior work on algorithms for predicting missing content (e.g., [5]) has shown promising results and could be used to: (1) help users determine if what they are looking for does not exist and (2) proactively provide feedback to the knowledge base administrator(s) about what information is missing and who might be qualified to write articles to fill the gaps. Second, search systems could provide visualizations to help users understand what parts (or percentage) of the collection they have already seen to help them determine the completeness of their search strategies.

Connecting people. In addition to searching for knowledge stored in documents, participants used the TC Hub to connect with people. This is an important role that procedural knowledge bases are likely to play in an organization, and search systems should support these uses. The most obvious use case involves connecting searchers with experts (e.g., people who have authored popular articles on the subject). Additionally, based on our survey responses, we can imagine two other use cases. A second use case could involve connecting searchers with peers with a similar background who have searched for similar information. Searchers might benefit from connecting with peers who have had similar goals and have overcome similar challenges. A third use case could involve connecting experts with novices. Based on our survey responses, experts often write tradecraft documents intended for novices (e.g., new employees). Therefore, when a tradecraft document is being uploaded, a system could suggest novice users who have looked for related information. Novice users might be able to provide feedback on the understandability and usefulness of a new article.

Standardized content labelling. User-generated content labelling and folksonomies have known limitations (e.g., divergent tag vocabularies and sparse labelling if the user base is small). Procedural knowledge systems could help address these issues. Systems could provide labelling suggestions to users based on a standardized vocabulary (e.g., this document has: overview, step-by-step instructions, involves knowing specific prior knowledge). In addition, at document creation, the system could predict and verify the classification of documents (e.g. as how-to or background documents).

6 CONCLUSIONS

In this paper, we present results of a survey study that examined the needs, practices, and challenges that intelligence analysts (IAs) faced in using an internal procedural knowledge base called the Tradecraft Hub. We classified the work task objectives that led participants to search for information in the TC Hub along two dimensions (the cognitive process and the artifact involved) and found a long-tailed distribution of task types. Additionally, we identified specific types of information sought by IAs during procedural knowledge tasks and important relevance criteria used to determine the usefulness of information found, which included intended audience, level of details, specificity, and task constraints. Finally, we described challenges that users face when searching for procedural knowledge and proposed specific system features to aid users during these types of tasks. Our results extend prior work to understand how users search in professional contexts and users' needs when searching for procedural knowledge. Our results have implications for the design of systems to support searching for procedural knowledge, especially in organizational contexts.

ACKNOWLEDGMENTS

This research was funded by the Laboratory for Analytic Sciences at North Carolina State University.

REFERENCES

[1] Eyob N. Alemu and Jianbin Huang. 2020. HealthAid: Extracting domain targeted high precision procedural knowledge from on-line communities. *Information Processing & Management* 57, 6 (2020). https://doi.org/10.1016/j.ipm.2020.102299
[2] Lorin W Anderson, David R Krathwohl, Peter W Airasian, Kathleen A Cruikshank, Richard E Mayer, Paul R Pintrich, James Raths, and Merlin C Wittrock. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, complete edition.*
[3] Peter Bailey and Li Jiang. 2012. User task understanding: a web search engine perspective. https://www.microsoft.com/en-us/research/publication/user-task-understanding-a-web-search-engine-perspective/ Presentation delivered at the NII Shonan: Whole-Session Evaluation of Interactive Information Retrieval Systems workshop. 8-11 October 2012, Shonan, Japan..
[4] Katriina Byström and Kalervo Järvelin. 1995. Task complexity affects information seeking and use. *Information Processing & Management* 31, 2 (1995), 191–213.
[5] David Carmel, Elad Yom-Tov, and Haggai Roitman. 2008. Enhancing digital libraries using missing content analysis *(JCDL '08)*. ACM, New York, NY, USA, 1–10.
[6] Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan Dumais. 2014. Lessons from the Journey: A Query Log Analysis of within-Session Learning. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining (WSDM '14)*. ACM, New York, NY, USA, 223–232. https://doi.org/10.1145/2556195.2556217
[7] John Ely, Jerome Osheroff, M Chambliss, Mark Ebell, and Marcy Rosenbaum. 2005. Answering Physicians' Clinical Questions: Obstacles and Potential Solutions. *Journal of the American Medical Informatics Association* 12, 2 (2005), 217–24.
[8] John W Ely, Jerome A Osheroff, Saverio M Maviglia, and Marcy E Rosenbaum. 2007. Patient-care questions that physicians are unable to answer. *Journal of the American Medical Informatics Association* 14, 4 (2007), 407–414.
[9] Bernhard Ertl. 2009. Conceptual and Procedural Knowledge Construction in Computer Supported Collaborative Learning. In *Proceedings of the 9th International Conference on Computer Supported Collaborative Learning (CSCL'09)*. International Society of the Learning Sciences, 137–141.
[10] Luanne Freund, Elaine G. Toms, and Julie Waterhouse. 2005. Modeling the information behaviour of software engineers using a work - task framework. *Proceedings of the American Society for Information Science and Technology* (2005).
[11] Jonathan Grudin. 2006. Enterprise knowledge management and emerging technologies. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, Vol. 3. IEEE, 57a–57a.
[12] Jessica Jameson, Beverly Tyler, Kathleen Vogel, and Sharon Joines. 2020. *Facilitating Interdisciplinary Collaboration Among the Intelligence Community, Academy, and Industry.*
[13] Steffen Soulejman Janus. 2016. *Becoming a knowledge-sharing organization: A handbook for scaling up solutions through knowledge capturing and sharing.* World Bank Publications.
[14] Hideo Joho, Leif A. Azzopardi, and Wim Vanderbauwhede. 2010. A Survey of Patent Users: An Analysis of Tasks, Behavior, Search Functionality and System Requirements. In *Proceedings of the Third Symposium on Information Interaction in Context (IIiX '10)*. ACM, New York, NY, USA, 13–24.
[15] Udo Kruschwitz and Charlie Hull. 2017. Searching the Enterprise. *Foundations and Trends in Information Retrieval* 11, 1 (2017), 1–142.
[16] Fabrizio Lillo and Salvatore Ruggieri. 2019. Estimating the Total Volume of Queries to Google. In *The World Wide Web Conference (WWW '19)*. Association for Computing Machinery, New York, NY, USA, 1051–1060.
[17] Lisa McDonald-Maier. 2009. esp@cenet@: Survey reveals new information about users. *World Patent Information* 31, 2 (2009), 142–143.
[18] Timothy Morris and Laura Empson. 1998. Organisation and expertise: An exploration of knowledge bases and the management of accounting and consulting firms. *Accounting, Organizations and Society* 23, 5-6 (1998), 609–624.

[19] David Newton. 2000. A survey of users of the new British Library Patent Information Centre. *World Patent Information* 22, 4 (2000), 317–323.

[20] Daniel E O'Leary. 1998. Enterprise knowledge management. *Computer* 31, 3 (1998), 54–61.

[21] Georg Pardi, Yvonne Kammerer, and Peter Gerjets. 2019. Search and Justification Behavior During Multimedia Web Search for Procedural Knowledge. In *Companion Publication of the 10th ACM Conference on Web Science (WebSci '19)*. ACM, New York, NY, USA, 17–20.

[22] Suppanut Pothirattanachaikul, Takehiro Yamamoto, Sumio Fujita, Akira Tajima, and Katsumi Tanaka. 2017. Mining Alternative Actions from Community Q&A Corpus for Task-Oriented Web Search. In *Proceedings of the International Conference on Web Intelligence (WI '17)*. ACM, New York, NY, USA, 607–614. https://doi.org/10.1145/3106426.3106461

[23] Tony Russell-Rose and Jon Chamberlain. 2016. Real-World Expertise Retrieval: The Information Seeking Behaviour of Recruitment Professionals. In *Advances in Information Retrieval*. Springer International Publishing, 669–674.

[24] Tony Russell-Rose, Jon Chamberlain, and Leif Azzopardi. 2018. Information retrieval in the workplace: A comparison of professional search practices. *Information Processing & Management* 54, 6 (2018), 1042–1057.

[25] Miamaria Saastamoinen, Sanna Kumpulainen, Pertti Vakkari, and Kalervo Järvelin. 2013. Task complexity affects information use: a questionnaire study in city administration. *Information Research* 19, 4 (2013).

[26] Pol Schumacher, Mirjam Minor, Kirstin Walter, and Ralph Bergmann. 2012. Extraction of Procedural Knowledge from the Web: A Comparison of Two Workflow Extraction Approaches. In *Proceedings of the 21st International Conference on World Wide Web (WWW '12 Companion)*. ACM, New York, NY, USA, 739–747. https://doi.org/10.1145/2187980.2188194

[27] John I. Tait. 2014. *An Introduction to Professional Search*. Springer International Publishing, 1–5.

[28] Kelsey Urgo, Jaime Arguello, and Robert Capra. 2020. The Effects of Learning Objectives on Searchers' Perceptions and Behaviors *(ICTIR '20)*. ACM, New York, NY, USA, 77–84.

[29] Michael Völske, Pavel Braslavski, Matthias Hagen, Galina Lezina, and Benno Stein. 2015. What Users Ask a Search Engine: Analyzing One Billion Russian Question Queries. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM '15)*. ACM, New York, NY, USA, 1571–1580. https://doi.org/10.1145/2806416.2806457

[30] Ingmar Weber, Antti Ukkonen, and Aris Gionis. 2012. Answers, Not Links: Extracting Tips from Yahoo! Answers to Address How-to Web Queries. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining (WSDM '12)*. ACM, New York, NY, USA, 613–622. https://doi.org/10.1145/2124295.2124369

[31] Zi Yang and Eric Nyberg. 2015. Leveraging Procedural Knowledge for Task-Oriented Search *(SIGIR '15)*. ACM, New York, NY, USA, 513–522. https://doi.org/10.1145/2766462.2767744