Collaborative Search and Sensemaking of Patents

Suresh K. Bhavnani[1]
School of Information
University of Michigan
Ann Arbor, MI 48109, USA
bhavnani@umich.edu

Matthew Scholl[2]
School of Information
University of Michigan
Ann Arbor, MI 48109, USA
schollm@umich.edu

Gavin Clarkson
School of Information,
School of Law,
and Native American Studies
University of Michigan
Ann Arbor, MI 48109, USA
clarkson@umich.edu

Current addresses
[1]Center for Computational Medicine
and Biology, Medical School
University of Michigan
Ann Arbor MI 48109, USA
bhavnani@umich.edu

[2]Accenture
Chicago, IL 60601, USA
matthew.r.scholl@accenture.com

CHI 2008, April 5 – April 10, 2008, Florence, Italy
ACM 978-1-60558-012-8/08/04.

Abstract
Despite the large number of patent searches conducted by professional patent searchers and inventors, little is known about how such searches are actually performed. Here we describe a qualitative study of experienced patent searchers as they conducted in-context searches at a technology transfer office. Based on studies of expert search and sensemaking in other domains, we expected the professional searchers to (1) use well-formed search strategies that were effective for patent search, and (2) rapidly make sense of the novelty of an invention by constructing new representations to organize existing patents that appear relevant. Instead, we observed the searchers perform simplistic preliminary searches and then exchange their search process and results with inventors and patent lawyers to collaboratively make sense of the patentability and licensability of the invention. Furthermore, their sensemaking consisted of selecting known representations of patents to organize the new information, an approach we call "weak" sensemaking. These results suggest implications for designing systems that support the observed collaborative "weak" sensemaking with the goal of helping the users to more effectively determine the patentability and licensability of an invention.

Keywords
Qualitative study, Interviews, Patent Search.

ACM Classification Keywords
H.3.3 Information storage and retrieval---Search process.

Introduction

Rapid advances in fields such as biotechnology have led to a sharp increase in the number of patent applications by universities and corporations [2], and to a corresponding increase in the number of online patent searches. However, while several studies have analyzed existing patent search systems and databases [e.g., 8], and developed new effective patent search algorithms [e.g., 4], little is known about how patent search itself is conducted by experienced patent searchers. Here we describe a qualitative study of one class of professional patent searchers called licensing representatives ("reps"), who help to determine the patentability and licensability of inventions at technology transfer offices in most large universities.

Prior Research

In contrast to the lack of user studies on patent search, there is a substantial body of literature on other search topics including collaborative information retrieval [e.g., 6], design of systems to support exploratory search [e.g., 3], and differences between expert and novice searchers [e.g., 1]. For example, experts in different domains have acquired domain-specific search strategies that help them perform comprehensive searches by visiting sources in specific sequences [1]. This body of research has resulted in several theories and models such as the Sensemaking theory [7], which states that when users search for information, they also attempt to make sense of that information by continually developing and refining representations (e.g., meaningful categories) to aid in the performance of real-world tasks. For example, when users search for cars, they also attempt to construct categories of cars (e.g., domestic versus foreign cars), to assist in making higher-level decisions (e.g., buy a reliable car).

The above theories and models guided our expectations of how search is conducted by professional patent searchers. We expected the searchers to (1) have well-defined search strategies in the domain of patent search, and (2) construct new representations to organize the found relevant patents with the goal of helping them make sense of the novelty of the invention.

Analysis of Real-World Patent Searching

To understand how professional patent searchers search for patents and determine the patentability of an invention, we conducted a qualitative study of the technology transfer office (TTO) in a large state university. We focused on a TTO because of the growing number of patent applications by universities, and the huge legal costs incurred when patents are incorrectly issued or used. Furthermore, we were motivated to study the TTO because they gave us permission to observe, analyze, and publish real-world patent searches, in return for specific insights to improve their search process.

Participants

The TTO consisted of 26 employees, 7 of whom were licensing representatives (reps), the focus of our study. As described by the TTO website, inventors who need to determine whether their invention is patentable or

licensable are required to disclose their invention through an online form. Once an invention is disclosed, a rep with experience in the domain of the invention is assigned to the disclosure, and proceeds to determine if the invention is patentable or licensable. If an invention is determined to be patentable, then the disclosure is forwarded to an external paid patent lawyer, who performs additional searches, which may lead to a patent application at the United States Patent and Trademark Office (USPTO).

The 7 reps in the office ranged in patent search experience from 2-18 years, and specialized in specific domains such as software and biomedical inventions. Informed by studies on expert searching and sensemaking in other domains, our goal was to understand the domain-specific strategies used by the search experts, how they determined the novelty of an invention by making sense of existing patents, and the difficulties they encounter during the above process.

Method

We used in-context observations of real searches, followed by semi-structured interviews [5] to understand the reps' search and sensemaking behavior. The reps were requested to contact us when they planned to perform a search related to a new invention disclosure. The reps were asked to perform their searches as they normally would, after which we conducted follow-up semi-structured interviews to understand the search behavior. The semi-structured interviews consisted of questions about (1) the nature of the invention, (2) the search goal, (3) the search process, (4) the individual sensemaking of found patents, and (5) the motivation for search termination. During our initial observations, we repeatedly observed the reps perform short inconclusive searches, and describe how they required more information from the inventors and lawyers. We therefore included in the interview two more questions about the collaborative search and sensemaking, and the information exchanged between the reps, inventors, and lawyers.

The data were recorded using audio, video, and screen-capture tools in addition to hand-written notes. The logistics of coordinating time-critical real-world observations resulted in 16 searches conducted by 5 reps (1 rep did not conduct patent searches, and another was located off-site) over a period of 9 months. Each rep conducted between 1-6 searches. The searches were representative of those typically conducted at the TTO.

Analysis

All searches and interviews were transcribed, and the answers were categorized in terms of the 5 pre-determined questions (nature of the invention, the search goal, the individual search process, individual sensemaking, and motivation for search termination) and the two questions that emerged during the interviews (the collaborative search and sensemaking, and the information exchanged between the reps, inventors, and lawyers). The results of the analysis were refined and verified through follow-up interviews with reps and inventors.

Results

SEARCH PROCESS. A preliminary analysis of the online searches revealed that 14 of the 16 searches used a small number of queries (1-7) composed of 1-3 keywords issued primarily to the USPTO website, and took less than an hour (30-45 minutes) with most of

the time spent on reading patents. In the remaining 2 cases, the reps entered 35 queries (as there were multiple novel elements in the invention requiring different sets of keywords) and 19 queries (which consisted of variations on a small set of terms). Ten searches used simple Boolean expressions with an "and" or "or". Even though the reps were experienced searchers, none of them used the 7-step USPTO recommended structured search strategy. Finally, none of the searches led to a conclusive decision about the patentability or licensability of the invention, and all searches ended in the need for more information from the inventor or lawyer. Our future research will include a detailed analysis of the above searches.

SEARCH GOAL. The reps repeatedly stated that their goal was to conduct preliminary searches (versus comprehensive searches conducted by the patent lawyers), followed by discussions with the inventor and the lawyer. This explains why there was little motivation to conduct comprehensive searches.

> "I'll usually do surface-level patentability searches, which means I'll do a search similar to what I am doing here, put them on paper and send them to the inventor and to the attorney. Then when we do our next conference call we'll discuss those patents."

Furthermore, the reps explained that if an invention was patentable, but there was no potential company that would license that invention, then the university would most probably not proceed with applying for a patent. The reps therefore appeared to play a crucial role in collaboratively determining whether the invention was potentially novel *and* could generate revenues for the university, before additional funds were spent to hire a patent lawyer.

SENSEMAKING. While the search goal was to do "surface-level" patentability searches, the reps appeared to be doing more than just finding relevant patents. During the search, they also attempted to make sense of the kind of *overlap* between the patents found and the invention, with the goal of determining its patentability. The analysis helped to identify 5 different types of overlaps. The following two excerpts describe two different types of complete overlap (shown in bold) between the invention and found patents. The first describes a possible **exact overlap** with an existing patent, and the second describes a **broad overlap** that covers a class of inventions, of which the invention in question is just an instance. In both cases the prior art invalidated the novelty of the invention.

> "…**if they are the same**, this is a smoking gun, and we don't need to go any further."

> "I think what we're going to find here is that most patents are **going to dominate** what the invention here is … it's just a modification on the design."

The following quotes describe 3 more types of overlap: a **large overlap** (which is not too promising), **a small overlap** (which is promising) and **no overlap** between patents found and the invention. The last case could be promising (as the invention is highly novel), or in this case determined not to be promising (as a lack of patent activity indicates a problem area).

> "Even though the art is **a bit congested** in terms of you know everybody's looking at this target … there might be some narrow sort of claims."

> "The ones that they pulled up really are just the heating side, and no mention of the use of a compression device. … There's probably some compression stuff, **but not the combination of the two**."

> "… it **does not appear that anyone is patenting actively** GPS for surgical purposes, which typically means there's a problem with the approach."

Often, the reps assessed the patentability of **multiple novel elements** within an invention, each of which could have one of the above 5 overlaps with existing patents. Figure 1 shows a sketch by one of the reps as he explained the complexity of determining the patentability of such an invention.

> "For the technology we had this morning, there's probably 4 or 5 novel elements within it, so it becomes a little harder to do these patent searches. Any one of those novel elements could be a patent."

The overall sensemaking appears to be a process of mapping the novelty of an invention to a small set of a priori representations of the overlap (e.g., exact overlap, broad overlap). However, unlike the "strong" form of sensemaking typically described in the literature [8] (where new representations are found or constructed), the above suggests a "weak" form of sensemaking, where previously determined representations are selected to organize new information [personal communication, Dan Russell].

The follow-up interviews also revealed how the reps engaged in collaborative sensemaking to conclusively

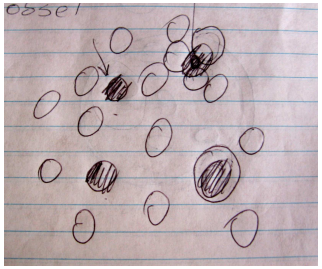

Figure 1. Sketch by a rep to explain how multiple novel elements (black circles) within a single invention can have different types of overlap with existing patents (white circles). The sketch shows examples of **small overlap** (top right), **broad overlap** also called "a dominated patent" (bottom right), and **no overlap** (top and bottom left).

determine the overlap between the invention and patents.

> "I'll ask the inventor to take a look at that and he'll say 'what we're doing when we wrap it is very different'. Or probably he'll say it's much simpler and therefore wouldn't involve all of those steps and therefore would not infringe."

SEARCH TERMINATION. The reps ended their searches when they needed four types of information: **feedback on patents found**, **clarification on the novel elements of the invention**, **potential licensing companies**, and **feedback on the keywords** used. For example, in the following quote, the rep is not sure if the patents he found are relevant to the invention.

> "I would send the few that I found … and say, 'Hey, did you see these? Take a look at these'. Get a little bit of his comments. If he says, 'Oh, these are kind of related', then I just send them to the patent attorney…"

MODE OF INFORMATION EXCHANGE. The reps used different media for saving and transmitting their search process and results, such as post-it notes on print outs as future reminders, and in Word documents. This largely ad hoc approach for recording and exchanging patent search and results tended to be prone to information loss as discussed by the following two reps.

> "I went through the two keywords that I used before. There's a third keyword but I'll dig it up later. I don't remember what it is and I didn't see it on the document. Those were [the inventor's] keywords…"

Design Implications

The results suggest that the reps (1) conduct preliminary searches to rapidly make sense of the patentability and licensability of an invention, (2) terminate their searches when they need specific kinds of information from inventors and lawyers, and (3) exchange their search results with inventors and lawyers to collaboratively make sense of the pantentability and licensability of an invention. Furthermore, there was evidence that the sensemaking

involved a process of determining the novelty of an invention by selecting from a pre-determined set of 5 types of overlap. We refer to this behavior as a "weak" form of sensemaking, because it differs from the current understanding of sensemaking which states that users mainly construct on-the-fly representations.

Given the frequency of the information exchange, we believe there is an opportunity to explicitly support their collaborative sensemaking process. Because the reps seemed comfortable using Venn Diagrams (as shown in Figure 1), the results suggest that they could benefit from a system where they can select from a visual vocabulary of 5 different types of overlap (exact, broad, small, large, and no overlap), denoted as Venn Diagrams. In each diagram, a black circle would represent the novel part of an invention, and white circles could represent a found patent with the kind of overlap that the rep believes is true. These diagrams can be further annotated with questions related to the feedback typically sought about patents, the invention, licensing companies, and the keywords used. The above representations could be exchanged between the inventors and lawyers, each of who could change the degree of overlap if they disagree with the conclusion reached by the rep, and provide feedback on the requested information. The above designs were well-received by the participants in our follow-up interviews.

Conclusions and Future Research

The key contributions of this study are: (1) the analysis of search and sensemaking in the domain of patent search which is currently not well-understood, (2) evidence for a "weak" form of sensemaking where pre-determined representations are selected to characterize the novelty of an invention. Our future research includes a detailed analysis of the observed searches, and methods to support "weak" sensemaking. We are also analyzing searches at another site where similar collaborative and sensemaking behaviors related to patent search have been observed, suggesting that our results might generalize beyond the current site.

Acknowledgements

We thank the study participants and Blomberg, J., Cohen M., Ganesan, A., Kieliszewski, C., Kreulen, J., Markey, K., McQuaid, M., and Saxman, P. This research was funded by the National Science Foundation, Award# IIS 0534903.

Rererences

[1] Bhavnani, S.K., Bichakjian, C.K., Johnson, T.M., Little, R.J., Peck, F.A., Schwartz, J.L., & Strecher, V.J. Strategy Hubs: Next-Generation Domain Portals with Search Procedures. *Proceedings of CHI'03* (2003), 393-400.

[2] Clarkson, G., & DeKorte, D. The Problem of Patent Thickets in Convergent Technologies*, in Progress in Convergence: Technologies for Human Wellbeing*, Annals of the New York Academy of Science, 1093, 2006, 180-200.

[3] Hearst, M. Clustering versus Faceted Categories for Information Exploration. *Communications of the ACM* 49(4), 2006, 59-61.

[4] Kang, I.S., Na, S.H., Kim, J., & Lee, J.H. Cluster-based patent retrieval. *Information Processing & Management*, 43, 5, (2007) 1173-1182.

[5] Kuniavsky, M. *Observing the User Experience: A Practitioner's Guide to User Research*. 2003. Morgan Kaufmann Publishers, San Francisco, CA.

[6] Poltrock, S., Grudin, J., Dumais, S., Fidel, R., Bruce, H. & Pejtersen A.M. Information Seeking and Sharing in Design Teams. *Proceedings of GROUP'2003*, 239-247.

[7] Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. The cost structure of sensemaking. *Proceedings of CHI'93* (1993), 269-276.

[8] Trippe, A.J. Patinformatics: Tasks to tools. *World Patent Information, 25* (2003), 211–221.