ELSEVIER

A multi-level model of information seeking in the clinical domain

Peter W. Hung *, Stephen B. Johnson, David R. Kaufman, Eneida A. Mendonça

Department of Biomedical Informatics, Columbia University, Vanderbilt Clinic, 5th Floor, 622 West 168th Street, New York, NY 10032, USA

Abstract

Objective: Clinicians often have difficulty translating information needs into effective search strategies to find appropriate answers. Information retrieval systems employing an intelligent search agent that generates adaptive search strategies based on human search expertise could be helpful in meeting clinician information needs. A prerequisite for creating such systems is an information seeking model that facilitates the representation of human search expertise. The purpose of developing such a model is to provide guidance to information seeking system development and to shape an empirical research program.

Design: The information seeking process was modeled as a complex problem-solving activity. After considering how similarly complex activities had been modeled in other domains, we determined that modeling context-initiated information seeking across multiple problem spaces allows the abstraction of search knowledge into functionally consistent layers. The knowledge layers were identified in the information science literature and validated through our observations of searches performed by health science librarians.

Results: A hierarchical multi-level model of context-initiated information seeking is proposed. Each level represents (1) a problem space that is traversed during the online search process, and (2) a distinct layer of knowledge that is required to execute a successful search. Grand strategy determines what information resources will be searched, for what purpose, and in what order. The strategy level represents an overall approach for searching a single resource. Tactics are individual moves made to further a strategy. Operations are mappings of abstract intentions to information resource-specific concrete input. Assessment is the basis of interaction within the strategic hierarchy, influencing the direction of the search.

Conclusion: The described multi-level model provides a framework for future research and the foundation for development of an automated information retrieval system that uses an intelligent search agent to bridge clinician information needs and human search expertise.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Information retrieval; Information seeking; Online searching; Search strategies; Problem solving; User expertise; Cognitive model; Clinician information needs; Intelligent agent

1. Introduction

Multiple information needs arise in the patient care setting during the normal course of work [1–4]. Gaps in knowledge can adversely affect the ability of a physician to make the best care decisions. Unfortunately, these needs often remain unmet for a variety of reasons, despite the presence of an ever-growing variety of online resources.

Some common obstacles that prevent physicians from pursuing or finding answers to their information needs include: doubt that relevant information exists, lack of time to initiate a search, uncertainty about where to find the information, and the absence of relevant topics in the resources searched [5]. Additionally, the process of searching an information resource for an answer may be disruptive to natural work flow. A recent survey of more than three thousand physicians showed that while the Internet was perceived as an important source of information, barriers to Internet information use were also high, with 57% of physicians reporting significant navigation and searching

* Corresponding author. Fax: +1 212 342 1647.
E-mail address: peter.hung@dbmi.columbia.edu (P.W. Hung).

difficulties, 49% reporting too much information to scan, and 45% reporting the belief that specific information to answer their question was not available [6]. These findings underscore a need to develop information retrieval systems that can help bridge these barriers.

One challenge is that while physicians have expert knowledge in the domain of patient care, they often do not possess the expertise necessary to translate their information needs into search strategies that produce the desired answers [7,8]. Beyond the initial barrier of articulating one's information need, many physicians do not know how to choose and search online resources effectively in order to find correct answers to their questions [9,10]. Furthermore, online information resources differ significantly in terms of their features and search interfaces, such that an effective search strategy for one resource may not work for another. Because of this, a clinician may rely on a few "favorite" resources while lacking the expertise to effectively search other more appropriate resources for a particular information need.

On the other hand, health science librarians represent a wealth of search knowledge within the biomedical domain. As a result of formal training and practical experience, they are adept at matching questions to information resources that are likely to yield appropriate answers. They also routinely employ systematic techniques for conducting a productive search for information, and often know how to navigate specific information resources in ways that casual searchers do not [11,12].

Often overlooked is the fact that the searcher's choice of strategies and techniques plays a central role in the effective retrieval of information, and that "online information retrieval is a problem-solving activity of a high order, requiring knowledge and understanding for consistently good results" [13]. Creating a system capable of such problem solving—one that employs an intelligent agent to make decisions about how to look for an answer by generating flexible search strategies based on human expert search knowledge—could be a particularly useful approach to meeting the information needs of clinicians. Loosely related efforts in biomedicine have focused on constructing hand-coded, pre-determined queries for narrow question types and particular resources (e.g., diagnosis questions in PubMed) [14–17]. In other domains, researchers have built systems that incorporate particular algorithmic search strategies based on those used by reference librarians [18], but little has been done to explore and develop technologies that are able to accept a defined information need and autonomously generate complex, adaptive search strategies that will conduct a high precision search to retrieve an appropriate answer, changing course as necessary during the process. A system that employed an intelligent search agent with such functionality would undoubtedly be a valuable tool for satisfying specific information needs that arise within the context of clinicians' work, particularly if supported by a user interface that accepts questions and their context in a form that clinicians are familiar with, i.e., natural language.

Before one can feasibly build a system that leverages human expert search knowledge, the nature, structure, and process of searching for information, as well as where and how search experts apply their knowledge, must be understood. Therefore, it is important to elucidate an appropriate theoretical model of information seeking that will guide and facilitate the formal representation of search expertise. Towards this purpose, the present paper describes the development of a theoretical model of online information seeking within the biomedical domain.

Before moving further, we more precisely describe the target of our modeling efforts. Consider an information retrieval system that entails the participation of two parties: an information *requester* (e.g., a clinician) and an information *searcher* (e.g., a human librarian or intelligent agent). The question that we are trying to answer is: given a well-defined information need that has been both adequately expressed by the requester and understood by the searcher, what types of search knowledge would the searcher require in order to solve the problem of retrieving an appropriate answer that meets the information need? The focus of the model to be presented is therefore on the generation and execution of a search strategy in the context of a single search session and in response to a single information request (as opposed to a more open-ended information seeking process involving multiple search sessions over time, as described elsewhere in the literature). Additionally, our interest lies particularly in the use of electronic resources and the strategies utilized to identify and seek information from these resources.

From the outset, we emphasize the limited scope and pragmatic nature of our model, as it is ultimately meant to serve as the architectural basis of a strategy-generating module to be used by an intelligent agent within an automated search system. As such, the model that we present does not aspire to be exhaustively descriptive nor explanatory of information seeking (e.g., how information needs emerge, why people conduct a search in the first place, what they do with the answers). The paper does not explicitly present a cognitive model of information seeking, but rather an explanatory vocabulary that can be used to articulate such a model. The goal of the current model is to present the dimensions of search knowledge needed to execute a context-initiated search process.

2. Background

In presenting a nested view of information seeking behavior, Wilson differentiates between information behavior and information seeking. *Information behavior* is defined as "those activities a person may engage in when identifying his or her own needs for information, searching for such information in any way, and using or transferring that information." More finely grained is *information seeking*, which concerns "the variety of methods people employ to discover, and gain access to information resources" [19]. The model developed in this paper can best be described as