Theoretical, Empirical and Practical Approaches to Resolving the Unmet Information Needs of Clinical Information System Users

James J. Cimino, MD; Jianhua Li, MD; Suzanne Bakken, RN, DNSc; and Vimla L. Patel, PhD
Department of Medical Informatics, Columbia University, New York, New York, USA

We hypothesize that when clinicians review clinical data in an electronic medical record, the information needs that arise are predictable, based on a number of situational factors. Because our theory does not say, exactly, what those needs are, we are using an empirical approach (observation) to detecting and categorizing them. For each need, we can construct an "infobutton" that links the clinical data to an on-line information resource. We have constructed an Infobutton Manager to match the data being reviewed by clinicians with context-appropriate infobuttons. This paper describes how the theory, observations, and practical solutions can come together to improve clinician decision making by resolving information needs.

INTRODUCTION

Seventeen years have passed since Covell, Uman, and Manning's landmark paper on studying clinicians' information needs was published.[1] Since then, several studies have further examined this aspect of the clinical decision-making process.[2,3] Although these studies hypothesized about the potential for computers to help with unmet needs, none has examined the specific information needs that arise during clinicians' interactions with clinical information systems (CIS). Although such interactions generally represent only a fraction of clinicians' time, we believe they are disproportionately important for several reasons:

- they occur when clinicians learn new information about patients,
- they occur when clinical decisions may be made (especially with order entry systems),
- the information needs that arise during the interactions may be predictable (based on the CIS function being used),
- the clinician is in a setting (using a computer) in which information needs may be resolved, and
- the retrieval of relevant information may be accomplished automatically by combining the prediction of the need with the clinical information that triggers the need.[4]

We have previously examined the technical feasibility of this approach by creating links within a Web-based CIS to *infobuttons*, context-specific questions that perform automated retrieval.[5] That work showed that the technical aspects of the process are relatively simple, compared to the task of predicting the clinicians' information needs. Furthermore, the method of hardwiring specific infobuttons to specific parts of the CIS not only makes them application-specific, but institution-specific as well.

We have developed a theoretical approach to characterizing and addressing the specific kinds of information needs that arise when using a CIS. We are combining this with an empirical approach to determine the actual needs that arise in actual settings. This paper describes the two approaches, as well as a practical prototype we have developed that implements infobuttons in a way that is dynamic (with respect to selecting information needs), as well as being application- and institution-independent.

THEORETICAL APPROACH

Cognitive theory,[6] supported by relevance research,[7,8] holds that information needs are, in part, *situated* ñ that is, they are based on the context in which they arise. Cognitive factors (memory, knowledge, and strategies) interact with situational and social factors in the decision-making process. This theory can include the interactions between users and CISs. We therefore hypothesize that:

- the information needs arising during clinicians' use of a CIS are predictable based upon specific context-dependent aspects of the interaction,
- these aspects are readily detectable by the CIS,
- common information needs that arise in a particular context can be predicted, and
- automatically recognizing and responding to needs can improve decision making.

Unmet information needs that result in delayed and uninformed decisions are a leading cause of medical errors.[9] We are seeking to explore how our hypothetical approach might be exploited to help clinicians make better-informed (and thus less erroneous) patient care decisions.

For example, suppose that a CIS user is reviewing a laboratory test that shows a positive antibody for Rheumatoid Factor. We could use the knowledge that the user is reviewing laboratory results to select a small list of potentially relevant questions ("What is the differential diagnosis for *test* *result*?", "What are the guidelines for evaluating *test result*?", etc.), customize them based on the specific test result being reviewed by the user ("What is the differential diagnosis for Rheumatoid Factor?", "What are the guidelines for evaluating Rheumatoid Factor?", etc.), and present them to the user. Each question could provide a link to an on-line resource appropriate for that question, and could access the resource in a way that takes the user directly to the relevant information.

Starting with our theory about context-specific information needs of CIS users, the above example illustrates how we can generate testable hypotheses about the feasibility of predicting and responding to such needs: after creating a system to predict the questions and generate the resource links, we can measure the impact of the system on increasing the rate at which information needs are satisfied. Prior to such a study, however, we need a better understanding of the cognitive aspects of the process.

EMPIRICAL APPROACH

In the above example, the context-sensitive information required for matching need to resource is simply the part of the CIS being reviewed and the clinical data on the screen. However, other information might be required as well. Is the user a nurse, physician, patient,[10] or patient's caregiver? Is the patient male or female? Is the patient newborn, child, adult or geriatric? Does it matter that the result is from a test of serum or cerebral spinal fluid? Each of these aspects could theoretically influence the kind of question asked, as well as the type of and manner by which the resource is accessed. We can speculate about the importance of different aspects, but our theory does not help us make predictions about the actual relative importance of these factors.

Our previous work with infobuttons focused on the technical aspects of linking predetermined questions to predetermined resources. Those aspects, it turns out, are the easy part. Knowing what questions are likely to be asked, and what factors influence the choice of questions, is the hard part, as earlier work with a predecessor of infobuttons has shown.[4] In the absence of a more complete theory of human cognition, we must resort to an empirical approach to solving this part of the problem.

Data Collection

Several studies of physicians and nurses in in-patient settings, using focus groups[11] and direct observation,[12] demonstrated a wide variety of information needs, only some of which were related to information found in patients' electronic medical records. In order to further characterize these needs, we are undertaking a direct observation of actual usage of WebCIS, the CIS at New York Presbyterian Hospital.[13] We are using a portable usability laboratory (PUL) to capture video and audio of clinicians using WebCIS. The PUL uses a video converter connected between the computer and monitor of a personal computer. The converter captures all activity on the computer screen and converts it to standard videotape format (NTSC). This signal, together with an audio signal captured by microphone, is conveyed a distance of up to 75 feet to a videocassette recorder. A research assistant observes the recording on a television monitor and, using proprietary software on a laptop computer, can pan and zoom the converter to enlarge areas of interest (such as locations on the screen where the user is entering data). Additional technical details are available in our study of clinician interactions with an outpatient medical records system.[14]

The PUL is attached to a clinical workstation at a nursing station on a general medical ward. Consenting physicians and nurses using the workstation are asked by a second research assistant to "think aloud" about what they are doing and what they wish they could do while using the system. The researcher will make note of any non-computer-based information resources the clinician makes use of (including books, journals, colleagues, etc.).

Data Analysis

The videotape will be transcribed, annotated, and coded in the manner previously described.[14] We will identify each occurrence of explicit information needs (where the subject expresses an information need) and implicit information needs (where the subject accesses some information resource). We will classify contextual aspects of each occurrence by: subject type, patient age and gender, CIS task, and clinical data being reviewed. Finally, we will identify the specific need, expressed as a question.

Once the contexts and questions have been extracted from the raw data, we will generalize the questions into forms that are less specific with regard to the actual clinical data in the context. For example, if several subjects have information needs about the significance of laboratory test results, we will create a generic form of the question that is related to the set of all laboratory results and has a placeholder for the specific result (e.g., "What is the significance of *laboratory result*?"). On the other hand, if subjects articulate information needs about medication doses when they are reviewing laboratory tests of drug levels, we will create a generic question that is related to the set of tests related to drug levels (e.g., "What is the dose of *drug*?"). As we map

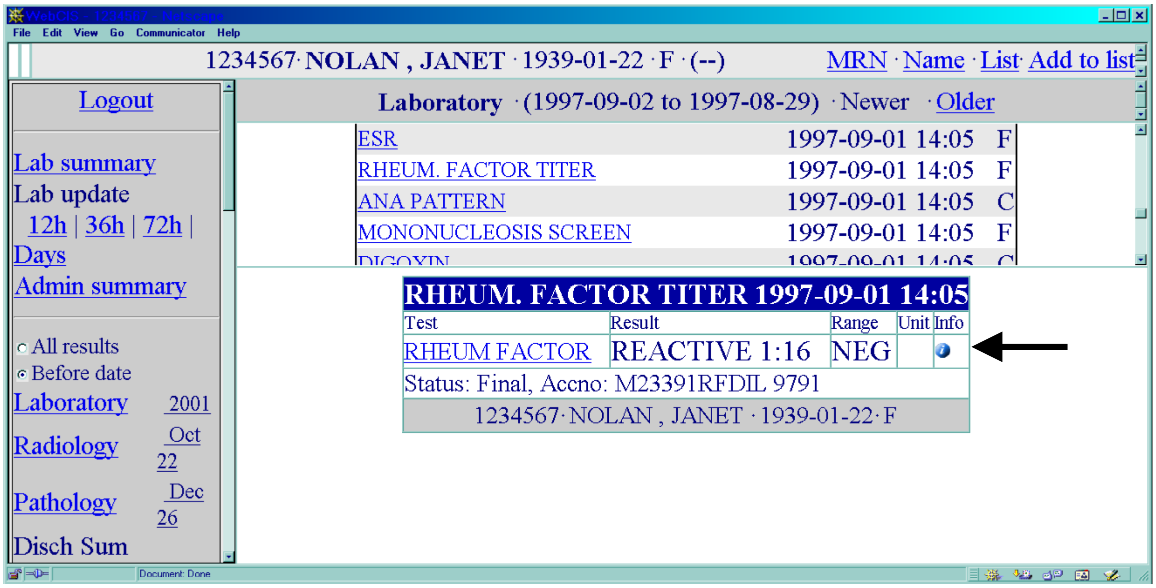

Figure 1: WebCIS screen showing laboratory results and link to the Infobutton Manager (arrow)

specific information needs to generic questions, we aggregate the context-specific information from each occurrence, such that we will know which questions were asked only by nurses, only about neonates, only about drug levels, etc.

PRACTICAL APPROACH

Several researchers have developed ways to link CISs to information resources.[15,16,17] Although they did not carry out studies of context-dependent information needs, each developed a database of likely questions, based on the answers that could be obtained from then-available resources. Resource matching was done using the semantic types of the data being reviewed, with some use of the National Library of Medicine's Information Sources Map.[18]

Efforts prior to the ascendancy of the World Wide Web were hampered by difficulties with access to disparate types of resources, as well as with integration with their respective CISs. As work with infobuttons has shown, these barriers have become significantly more surmountable.[5] In particular, the ubiquity of Web-based resources greatly simplifies access to knowledge. Web-based CISs ease the integration tasks; however, they are still subject to the rules and limitations of software development.

The Infobutton Manager

We have taken a two-step approach to developing practical ways to integrate links (i.e., infobuttons) to resources in our CIS. The first step uses an *Infobutton Manager* (IM) to keep track of all questions, contexts, and information resources. The IM is a CGI (Common Gateway Interface) program that can be accessed from any Web-based CIS through a simple hypertext reference (HREF) that "posts" (a method of passing data to a CGI) context information.

For example, in Figure 1, the white ìiî in the (blue) circle is a link to the IM. The link passes the following: Institution=NYPH, UserType=MD, AgeGroup=E, Gender=F, Context=LabDetail, and ClinicalConcept =36144. The first two variables tell the IM about the user: a physician from New York Presbyterian Hospital. The next two variables tell the IM about the patient: an elderly female (we do not pass date of birth for confidentiality reasons). The last two variables tell the IM about the clinical data being reviewed: details of a laboratory test, where the specific test is represented by code 36144. Because the Institution is "NYPH", the IM can look up this code in the Medical Entities Dictionary (MED).[19]

The Context Table

The IM matches the context information against rows in a Context Table. The Context Table has columns corresponding to each element of the context information. Each row in the Context Table corresponds to one "context"; each column has one or more values corresponding to context variables.

For most of these columns, the number of options is relatively small (for example, we currently only match on the user types "RN" and "MD"). However, for Clinical Context the number of values is potentially huge. In order to make the Context Table more manageable, the IM takes advantage of the MED hierarchy: only the class of concepts of interest needs to be included in the table.

For example, the context information above would match against the Context Table rows:

NYPH | MD,RN | F,M | Y,M,E | LabDetail | 2248
NYPH | MD,RN | F,M | Y,M,E | LabDetail | 32945
NYPH | MD,RN | F,M | Y,M,E | LabDetail | 41480

For the most part, these matches are straightforward. The last field represents the class code for relevant concepts. In the first match, the code 2248 is the MED Code for "Single-Result Laboratory Test". In the second case, 32945 is the MED Code for "Autoantibody Tests". In the third match, 41480 is the MED Code for "Rheumatoid Factor Tests". Each of these concepts is a superclass of the laboratory test shown in Figure 1.

The Infobutton Table

In addition to the context matching information, each row in the Context Table contains an Infobutton ID. This ID is a unique key in a second table, the Infobutton Table. This table contains the actual links to the information resources, along with natural language questions for display to the user.

The IM uses the MED to translate the test name into a measurable substance (in this case, "Rheumatoid Factor"). It can then generate the HTML code shown in Figure 2. When displayed by a Web browser, the HTML code shown in Figure 2 would look like this:

What is the differential diagnosis of Rheumatoid Factor?

What are guidelines for evaluation of Rheumatoid Factor?

How is rheumatoid arthritis managed?

Figure 2: Examples of hypertext links generated by the Infobutton Manager

What is the differential diagnosis of Rheumatoid Factor?
What are guidelines for evaluation of Rheumatoid Factor?
How is rheumatoid arthritis managed?

The links yield,* respectively: 78 citations from PubMed, 23 guidelines from the National Guideline Clearinghouse, and the American College of Rheumatology's "Guidelines for the Management of Rheumatoid Arthritis".[20]

DISCUSSION

For any given information need, if an answer exists, it is likely to exist on the Internet. If we can identify these needs, we can resolve them. The IM is a potentially powerful tool for implementing such solutions. Because it is a CGI, it is readily integrated into Web-based CISs. Because it is table-driven, it is easy to add new information needs and contexts. Because it provides HTML links, it is able to use the World Wide Web. Finally, because of its institution-independent architecture, it can be used as easily around the world as it is at NYPH. However, for all the promise of the IM, it presents several challenges.

The first version of the IM is functioning and has been integrated into a test version of WebCIS. As we gain experience with its use, it is likely that we will encounter limitations of the current system. For example, we may find that the context items we have identified are insufficient for expressing all the necessary nuances. As new attributes are needed, we will expand the Context Table to accommodate them.

Although many Web sites permit direct queries by HTML references such as those shown in Figure 2, we may find that some information needs can best be resolved by accessing resources with more complex interfaces. In such cases, we will need to create agents, or "bots", that can interact with the resources to obtain the necessary information. The Infobutton Table will then contain links to agent intermediaries instead of the outside resources themselves.

The institution-independent nature of the IM relies on the institution-independent nature of the context attributes. We believe that, to a large extent, this is achievable: Web-based CISs typically provide their users with functions like laboratory results review, and they can be expected to know the age and gender of their patients, as well as the professional roles of their users. However, standardizing the clinical concepts across institutions remains challenging. In some cases, we may be able to find a standard terminology for representing the concepts (such as LOINC[21] for laboratory results). In other

* as of July 5, 2002.

cases, we may need to provide modules for the IM to allow it to translate from institution-specific codes to something it recognizes (the UMLS may be a possibility). In the worst case, we can simply add multiple rows in the Context Table for each institution, while leaving the Infobutton Table free of such specificity.

Of course, the major issue to resolve will be the identification of the actual information needs in order to represent them as contexts and queries. The empirical approach that we have used to date - direct observation using the PUL at a clinical workstation in the nursing station - has demonstrated utility in capturing actual information needs. Experience to date leads us to believe that the IM will provide sufficient representational power to handle the common needs. The IM will then need to be evaluated from technical (does it work?), ergonomic (is it usable?), and cognitive (is it useful and, if so, how useful?) perspectives. However, if it succeeds in reducing the rate at which clinicians' information needs go unmet, the real challenge will be to show that this translates into a reduction in errors and improvement in the quality of health care.

CONCLUSIONS

We believe a combination of theoretical, empirical, and practical approaches are required to improve resolution of clinician information needs. The Infobutton Manager provides an architecture for realizing this combination that appears to adequately address the practical requirements. Addressing the cognitive challenges as they interact with social and contextual variables is the next step.

Acknowledgments

This work is supported in part by National Library of Medicine grant 1R01LM07593.

References

1. Covell DG, Uman GC, Manning PR. Information needs in office practice: are they being met? *Ann Intern Med* 1985; 103(4):596-9.
2. Cimino C, Barnett GO: Analysis of physician questions in an ambulatory care setting. In Clayton PD, ed.: *SCAMC*; 1991:995-999.
3. Osheroff JA, Forsythe DE, Buchanan BG, Bankowitz RA, Blumenfeld BH, Miller RA: Physicians' information needs: analysis of questions posed during clinical teaching. *Annals Internal Medicine*, 1991; 114(7):576-581.
4. Cimino JJ, Johnson SB, Aguirre A, Roderer N, Clayton PD. The Medline Button. In: Frisse ME, (ed.). *SCAMC*. 1992, 1992:81-85.
5. Cimino JJ, Elhanan G, Zeng Q. Supporting Infobuttons with Terminological Knowledge. *JAMIA*; 1997;4 (Suppl):528-532.
6. Patel VL, Arocha J F. The nature of constraints on collaborative decision making in health care settings. In G. Klein (ed.): *Expertise and naturalistic decision making*. Lawrence Erlbaum Associates, Mahwah, NJ. 2000:78-91.
7. Barry L. User-defined relevance criteria: an exploratory study. *JASIS.,* 1994;45 (3): 149-159.
8. Narayanan S, Bailey W, Tendulkar J, et al. Design of model-based interfaces for a real world information system. *IEEE Trans on Systems, Man and Cybernetics*. 2002;32:11-24.
9. Kohn, KT, Corrigan, JM, Donaldson, MS (eds.). *To Err is Human: Building a Safer Health System*. Inst. of Medicine. Nat Acad Press, 1999.
10. Cimino JJ, Patel VL, Kushniruk AW. What do patients do with their medical records? In: Patel VL, Rogers R, Haux R (eds.). *Medinfo 2001*, 1440-1444.
11. McKnight LK, Stetson PD, Bakken S, Curran C, Cimino JJ. Perceived information needs and communication difficulties of inpatient physicians and nurses. *JAMIA*. 2001;8 (suppl.):453-457.
12. Kubose TT, Cimino JJ, Patel VL. Assessment of information needs for informed, coordinated activities in the clinical environment. *JAMIA*. 2001;8 (suppl.):948.
13. Hripcsak G, Cimino JJ, Sengupta S. WebCIS: large scale deployment of a Web-based clinical information system. *JAMIA*. 1999;6(supl.):804-8.
14. Cimino JJ, Patel VL, Kushniruk AW. Studying the human-computer-terminology interface. *JAMIA*; 2001;8(2):163-173.
15. Miller RA, Gieszczykiewicz FM, Vries JK, Cooper GF. CHARTLINE: providing bibliographic references relevant to patient charts using the UMLS Metathesaurus knowledge sources. In: Frisse ME (ed). *SCAMC*, 1992:86-90.
16. Cimino C, Barnett GO, Hassan L, Blewett DR, Piggins JL. Interactive Query Workstation: standardizing access to computer-based medical resources. *Computer Methods and Programs in Biomedicine*. 1991; 35(4):293-299.
17. Clyman JI, Powsner SM, Paton JA, Miller PL. Using a network menu and the UMLS Information Sources Map to facilitate access to online reference materials. *Bulletin of the Medical Library Association*. 1993; 81(2):207-16.
18. Masys DR. An evaluation of the source selection elements of the prototype UMLS Information Sources Map. In: Frisse ME (ed). *SCAMC*, 1992:295-8.
19. Cimino JJ. From data to knowledge through concept-oriented terminologies: experience with the Medical Entities Dictionary. *JAMIA*; 2000;7(3):288-297.
20. American College of Rheumatology Ad Hoc Committee On Clinical Guidelines. Arthritis & Rheumatism: Guidelines For RA Management ACR Clinical Guidelines Committee. 1996;39(5):713-722.
21. Huff SM, Rocha RA, McDonald CJ, et al. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. *JAMIA*. 1998;5(3):276-292.