Practical Considerations for Exploiting the World Wide Web to Create Infobuttons

James J. Cimino[a], Jianhua Li[a], Mureen Allen[a], Leanne M. Currie[b], Mark Graham[a], Viktoria Janetzki[a] Nam-Ju Lee[b] , Suzanne Bakken[a,b] , Vimla L. Patel[a]

[a] *Department of Biomedical Informatics, Columbia University, New York, New York, USA*
[b] *School of Nursing, Columbia University, New York, New York, USA*

***Background**:*

We are studying ways to provide automated, context-specific links (called "infobuttons") between clinical information systems (CIS) and other information resources available on the World Wide Web. As part of this work, we observed the information needs that arose when clinicians used a CIS and we classified those needs into generic questions. We then sought general methods for accessing information resources to answer the questions. ***Methods**: For each generic question, we identified a satisfactory resource and then developed a method for retrieving from it the information relevant to the question. We then studied these methods to characterize them into general approaches.* ***Results**: We identified six general approaches and describe them in detail. These approaches range in complexity from simple, hard-coded links to intelligent agents and calculators.* ***Conclusion**: Web-based information resources can be exploited using a relatively small number of methods, although the specific methods require custom solutions. Standard methods for accessing Web-based resources would simplify the task of linking to CISs.*

Keywords:

Information Needs, Information Retrieval, Decision Support, Internet

Introduction

Informatics researchers have long sought ways to provide links between clinical information systems (CIS) and on-line information resources, with the intention of using these resources to answer information needs that arise during use of the CIS.[1,2] The usual approach to designing such links has been to ask the question, "What information needs can the resource serve?" and then build a "hard-wired" link that uses information in the clinical system to drive the retrieval from the resource.[3] With the advent of the World Wide Web, links have become easier to create and the number of available resources has increased, reducing some of the barriers to link creation. The result has been the creation of a variety of such links, often called "infobuttons",[4-9] from Web-based clinical systems to Web-based resources.

We are approaching this task from a different perspective, by asking the question, "What information needs do clinical information system users have in a given situation?" and then finding resources that will answer these questions.[10] We can readily find resources that address the questions, but the methods by which we link to these resources (and by which we automate the retrieval of information from them) are proving to be heterogeneous with respect to program interfaces, user interfaces, information organization, and terminology. This paper describes our experience with different techniques for creating infobuttons.

Materials and Methods

As part of a formal study of the information needs that arise while using a clinical information system,[11] we identified 154 instances of information needs.[12] Our analysis of these needs shows a number of recurring "generic" questions. For example, we encountered a variety of needs that could be represented with the generic question, "What is the definition of X?", where X is a context-specific concept of interest (e.g., a term found in a patient's medical record report). For each of the questions, we sought to provide a Web-based link to some system that would answer the question (for example, an on-line medical dictionary to provide definitions of terms). Just as a resource might answer any number of specific questions, if they fit the pattern of some generic question, we found that Web-based resources use a variety of "generic" methods by which information can be automatically retrieved. For each of our generic questions, we identified the resource, the generic method by which information could be retrieved, and the terminology used by the system.

Results

Table 1 shows the generic questions we were able to detect in the information needs we observed. For each need, we have identified a resource that addresses the need, developed a method for accessing the resource, and identified the terminology used by the resource, also listed in Table 1. Each of the methods falls into one of six categories, which are described below and depicted in Figure 1.

Simple Link

In some cases, a resource exists that is on the one hand static and on the other hand provides broad coverage of a topic, as with a text document. If a generic question relates to the specific domain covered by the resource, and the resource is not so voluminous that a user might get lost trying to find an answer, then the solution is simply to provide the user with a link to the document.

For example, we have found that clinicians reviewing laboratory test results may have questions about treatment of diseases related to those results. When the test is any of a number of tests related to rheumatoid arthritis, the American College of

Table 1: Information needs found in [11,12] and the methods (described in the text of the paper) used to resolve them

Information Need	Context	Resource	Method	Terminology
What are the institution's transfusion guidelines?	Laboratory Results	Institutional policy manual on the Web[1]	Simple Link	N/A
What specimen do I collect for lab test X?	Laboratory Results	Laboratory manual as a set of Web pages[2]	Concept-Based Link	Laboratory Test Codes
What is the definition of X?	Radiology Reports	Merriam-Webster[3]	Simple Search	N/A
		Medicine on the Net[4]	Simple Search	N/A
Various questions about disease/finding X	Various	UpToDate[5]	Simple Search	N/A
		PubMED[6]	Concept-Based Search	MeSH
		Harrison's Principles of Internal Medicine[7]	Intelligent Agent	N/A
Give me patient instructions for drug X.	Pharmacy	Lexicomp[8]	Intelligent Agent	Drug Name
What are the dose forms for drug X?	Pharmacy	Micromedex[9]	Intelligent Agent	Drug Name
How do I use/taper drug X?	Pharmacy	Micromedex[9]	Intelligent Agent	Drug Name
What drugs interact with drug X?	Pharmacy	Micromedex[9]	Intelligent Agent	Drug Name
What is the patient's anion gap?	Laboratory Results	Local Web-based anion gap calculator[1]	Calculator	Laboratory Test Codes
Does the patient have iron deficiency?	Laboratory Results	Local Web-based iron stores calculator[1]	Calculator	Laboratory Test Codes

References to Web-based resources. Note that "<>" indicates location into which the concept of interest ("X") is inserted.
1. Not available outside of Columbia's intranet.
2. http://cpmclabinfo.cpmc.columbia.edu/<>#Collection ("<>" uses concept-specific document codes from the data dictionary)
3. http://www2.merriam-webster.com/cgi-bin/mwmednlm?book=Medical&va=<>
4. http://www.medicinenet.com/script/main/srchCont.asp?li=MNI&SRC=<>&op=MM
5. http://www.utdol.com/application/vocab.asp?search=<>&submit=Go
6. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed&cmd=Search&term=<>[MeSH+Terms]
7. Not available outside of Columbia's intranet. See http://harrisons.accessmedicine.com.
8. Not available outside of Columbia's intranet. See http://www.lexi.com.
9. Not available outside of Columbia's intranet. See http://www.micromedex.com/

Rheumatology provides a convenient, authoritative place to find the answer.[13] However, since it is a static document (in this case, a Portable Document Format (PDF) document image), there is no straightforward method to retrieve just the section that is relevant to the clinician's specific question. The simple solution, then, is to provide a Uniform Resource Locator (URL) link to the whole document (*www.rheumatology.org/research/guidelines/raguidelines02.pdf*). If the user clicks on this URL, the user's Web browser will retrieve and display the document.

Concept-Based Link

In some cases, we have found resources that are composed of a set of documents, with unique URLs for specific concepts or sets of concepts. The URLs in such resources typically have one part in common (such as the server name and the directory tree) and only a small part (such as the actual file name) that is unique. In such cases, if we can map the concept of interest to the unique part of the address, we can dynamically create a simple link to the relevant document.

For example, we have found that clinicians reviewing laboratory test results may have questions about the tests themselves (e.g., "What specimen do I collect for lab test X?"). Our clinical laboratory provides a Web-based resource containing such information. Users normally interact with the resource by searching through a long alphabetized list of topics. However, we became aware that each topic is a link to a static document (e.g., the topic "Rheumatoid Factor" links to the URL *cpmclabinfo.cpmc. columbia.edu/chapter/mono/ss001700.htm*, while the topic "Legionella Antigen, Urine" links to the URL *cpmclabinfo. cpmc.columbia.edu/chapter/mono/new1050946339.htm*). We matched each of our laboratory test terms to the relevant document in the resource and stored the document names in our data dictionary. We can now create infobutton links by extracting a concept-specific part from the data dictionary (e.g., *new1050946339.htm*) and combining it with a static part (*cpmclabinfo.cpmc. columbia.edu/chapter/mono/*) to produce a link to the document that is specific for the concept of interest. This acts as a simple link, so that if a user clicks on the link, the user's Web browser will retrieve and display the specified document.

Simple Search

Many information resources provide users with a "Search" feature that allows them to type a word or phrase into a text box and

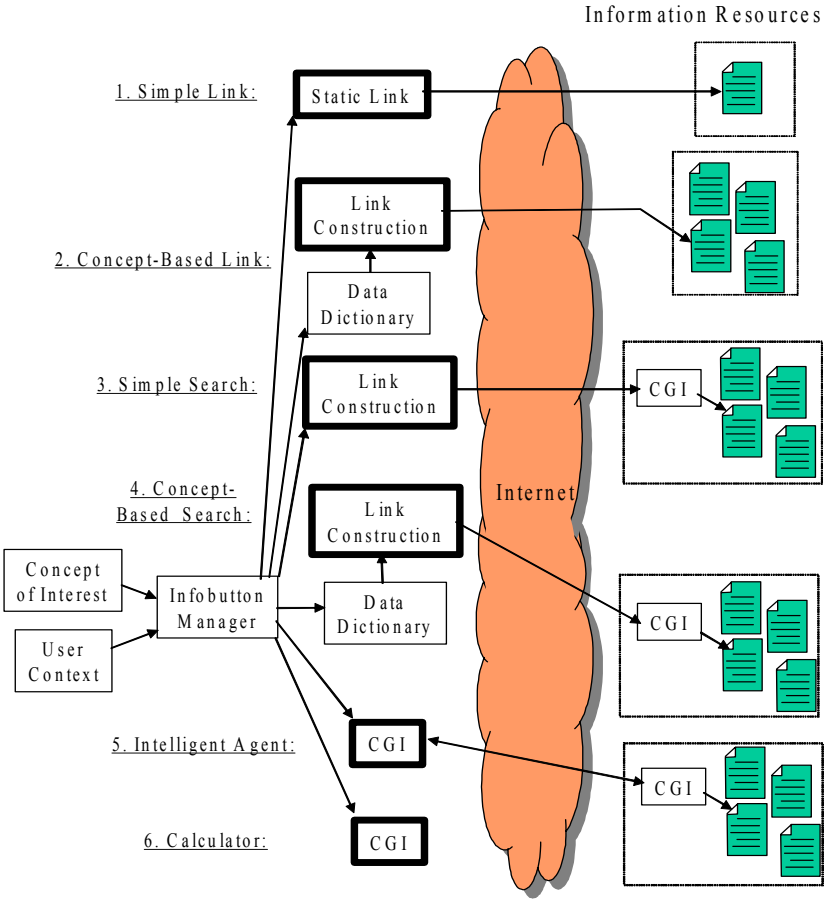

Figure 1 - Six different infobutton methods (heavy boxes). When a user clicks on an infobutton link in the clinical information system, the Concept of Interest and the User Context are passed to the Infobutton Manager, Which then selects relevant resources. The details of each method are described in the text. CGI stands for Common Gateway Interface, the standard for executing programs via the World Wide Web

then passes the text to a Common Gateway Interface (CGI) program. The CGI can often be evoked directly (by what is referred to as the "get" method), using a URL composed of the CGI's name and address and any required parameters, including the user's term of interest.

For example, suppose that a clinician encounters the word "malacoplakia" in a text document and wants to know its meaning. To obtain an answer from the Merriam-Webster on-line medical dictionary, one can simply insert "malacoplakia" into a generic URL (see Table 1, Note 3) to form the specific URL *www2.merriam-webster.com/cgi-bin/mwmednlm?book=Medical&va=malacoplakia*. If a user clicks on a link containing this URL, the CGI (*mwmednlm*) will perform a search and return a page to the user's Web browser for display.

Concept-Based Search

In some cases, the search features provided by information resources will perform better with (or may even require) a term from a controlled terminology, rather than text input. In these cases, we can use our data dictionary to map from the concept of interest to a form recognized by the resource.[4]

For example, we have found that a variety of clinician information needs related to diseases and findings can be addressed with PubMed searches. We have included in our data dictionary many MeSH translations obtained from the US National Library of Medicine's Unified Medical Language System (UMLS).[14] Our infobutton manager[10] can convert the concept of interest to a MeSH term in the same manner used to obtain a concept-specific link. It can then insert the translated term into the PubMED URL, in the manner used with the Simple Search method, above.

Intelligent Agent

Unfortunately, not all resources can be accessed through simple links or CGI calls containing terms of interest. In some cases, the resources require the user to navigate through the resource to reach specific information of interest. Intelligent agents are programs that can automatically carry out this navigation by simulating the actions (mouse clicks and data entry) of a human user.

For example, we found Micromedex (see Table 1, Note 9) to be very useful and appropriate for resolving several different types of information needs related to drug products. However, navigation through the user interface to access the desired informa-

tion can take four or more steps. It also requires using explicit brand or generic names, rather than free text. In order to exploit Micromedex for our users' information needs, we constructed an intelligent agent program that translates a concept of interest to a generic drug name, executes the Micromedex search CGI, captures the document returned by Micromedex, extracts the useful link from the document, executes this link, obtains a second document back from Micromedex, and provides the topic-specific section of the second document back to the user.

Calculator

Finally, we have found some situations where no suitable resource exists on the Web, either locally at our institution or elsewhere. Interestingly, this situation has not yet arisen for text-based information. However, we have encountered several instances where the clinicians requested help with a calculation based on laboratory values. In each case, we were unable to find suitable sites that could be readily exploited. We therefore created our own programs to accept laboratory results and perform the appropriate calculations.

Discussion

Our experience demonstrates that a variety of Web-based resources can be exploited to service clinician information needs in the context of using a CIS. However, several challenges remain before we can reach the ultimate goal of being able to call the Infobutton Manager from any clinical system and have it be able to provide links to any appropriate resource. As with many integration tasks, the challenges can be characterized as those related to interaction method (format) and those related to representation (terminology).

Many of the resources we encountered are fairly open to external requests for their contents, typically through search engines. This openness supports simple requests, such as "Give me a page of text that talks about topic X", but it does not support more specific questions, such as "What is the most common cause of finding X in patients for gender Y and age Z?" We, and others,[8] have achieved more sophisticated queries with intelligent agents. However, their construction is not without cost and they are exquisitely sensitive to changes in format, structure, presentation and content of their target resources. Development of standard methods for interacting with these resources (such as the Z39.50 standard for interacting with bibliographic databases[15] or any of a variety of emerging standards for Web services) would simplify the querying task.

Standards for terminology will also facilitate the development of infobuttons. In particular, we face significant challenges as we attempt to integrate the Infobutton Manager with multiple clinical information systems, each with its own terminology. The implications are even greater for attempts to share infobuttons internationally. In addition, the resources themselves rarely use standard terminologies. Thus, mapping concepts of interest from multiple clinical systems to multiple information resources remains difficult.

Ultimately, we look forward to the availability of resources that not only provide standardized access and use standardized terminologies, but also can make use of additional information such as user type and context, to facilitate the ubiquitous diffusion of infobuttons, for use by a wide variety of healthcare stakeholders.

Conclusion

Despite the challenges of heterogeneous interaction methods and terminologies, we have been successful in creating working infobuttons that address clinicians' information needs as they arise while using a clinical information system. Creative solutions have sometimes been required but, in the end, the Web has proven to be rich in high-quality resources. Our ability to develop infobuttons for a diversity of user questions needs satisfies a major technical challenge to automated methods for resolving the unmet information needs of clinicians as they use clinical information systems.

Acknowledgments

This work is supported in part by NLM grant 1R01LM07593 and Training Grant LM07079-11.

References

[1] Cimino JJ, Johnson SB, A. Aguirre, N. Roderer and P.D. Clayton. The Medline Button. *Proc 16th Ann SCAMC*; 1992:81-85.

[2] Miller RA, F.M. Gieszczykiewicz, J.K. Vries and G.F. Cooper. CHARTLINE: providing bibliographic references relevant to patient charts using the UMLS Metathesaurus Knowledge Sources. *Proc 16th Ann SCAMC*; 1992:86-90.

[3] Cimino JJ. Linking patient information systems to bibliographic resources. *Meth Inform Med*; 1996;35(2):122-126.

[4] Cimino JJ, Elhanan G, Zeng Q. Supporting infobuttons with terminological knowledge. *Proc AMIA Annu Fall Symp*. 1997;:528-32.

[5] Ruan W, Burkle T, Dudeck J. An object-oriented design for automated navigation of semantic networks inside a medical data dictionary. *Artif Intell Med*. 2000 Jan;18(1):83-103.

[6] Ruan W, Burkle T, Dudeck J. A dictionary server for supplying context sensitive medical knowledge. *Proc AMIA Symp*. 2000;719-23.

[7] Baorto DM, Cimino JJ. An "infobutton" for enabling patients to interpret on-line Pap smear reports. *Proc AMIA Symp*. 2000;47-50.

[8] Reichert JC, Glasgow M, Narus SP, Clayton PD. Using LOINC to link an EMR to the pertinent paragraph in a structured reference knowledge base. *Proc AMIA Symp*. 2002;652-6.

[9] Eduljee A, Rocha RA. Practical evaluation of clinical guideline document models. *Proc AMIA Symp*. 2003:836.

[10]Cimino JJ, Li J, Bakken S, Patel VL. Theoretical, empirical and practical approaches to resolving the unmet information needs of clinical information system users. *Proc AMIA Symp*. 2002;170-4.

[11]Currie LM, Graham M, Allen M, Bakken S, Patel VL, Cimino JJ. Clinical information needs in context: an observational study of clinicians while using a clinical information system. *Proc AMIA Symp*. 2003:190-194.

[12]Allen M, Currie LM, Graham M, Bakken S, Patel VL, Cimino JJ. The classification of clinicians' information needs while using a clinical information system. *Proc AMIA Symp*. 2003:26-30.

[13]American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis: 2002 update. *Arthritis Rheum*. 2002 Feb;46(2):328-46.

[14]Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Meth Inf Med*. 1993 Aug;32(4):281-91.

[15]*Information Retrieval (Z39.50): Application Service Definition and Protocol Specification*. ANSI/NISO Z39.50-1995 (http://lcweb.loc.gov/z3950/agency/)

Address for correspondence

James J. Cimino, M.D.
Department of Biomedical Informatics
Columbia University
622 West 168th Street, VC-5
New York, New York 10032 USA
jjc7@columbia.edu