WRAPIN : a tool for patient empowerment within EHR

Michel Joubert\textsuperscript{a}, Arnaud Gaudinat\textsuperscript{b}, Célia Boyer\textsuperscript{b}, Antoine Geissbuhler\textsuperscript{b}, Marius Fieschi\textsuperscript{a},
HON Foundation Council members\textsuperscript{b}

\textsuperscript{a} LERTIM, Faculté de Médecine, Université de la Méditerranée, Marseille, France
\textsuperscript{b} Health On the Net Foundation (HON), Geneva, Switzerland

Abstract

Legal and technologic trends are making medical records progressively more patient-accessible. In the near future, information technology may make it even easier to provide patients a chance to review their records. One may wonder, however, about the practical use of this technology by patients. Understanding his/her own health record will certainly be one of the main concerns of patients. WRAPIN has been designed to provide patients and citizens with trusted health information. It will help to determine the reliability of documents by checking the ideas contained against established benchmarks, and enable users to determine the relevance of a given document from a page of search results. First, we present what is, in our opinion, the most original and important patient-centred WRAPIN characteristics and functionalities. Then, we compare these characteristics with those of representatives of two main trends in information retrieval: systems based on the popularity of web sites, and on the clustering of web sites. This comparison shows that, even though patients are tempted to use popular search engines, these are not sufficiently specialized in the medical domain to help them understand their own HER.. Finally, we discuss the complexity of medical readings over the Internet and the efforts that are still required in this domain.

Keywords : Abstractive and Indexing; Information Storage and Retrieval; Electronic Health Record; Internet; Quality Control

Introduction

In a short note Dr PH Chew related a visit to one of his patients [1]. The patient said: “Doctor, I would like to inform you that for the past two weeks my corpus spongiosum has not been filling adequately (…) I read on the Internet that Metoprolol\textsuperscript{®} can cause impotence.” Dr Chew wrote: “I realised that the internet has provided this patient with the knowledge that his recently acquired impotence was drug-induced.” It is true that it is now easy to log onto a web site and obtain this kind of information about Metoprolol. We learn from this account that Internet is currently an essential source of information for patients. It can also be used by patients to understand and check diagnoses and treatments.

Legal and technologic trends are making medical records progressively more patient-accessible. In the near future, information technology may make it even easier to provide patients a chance to review their health records. Studies of technology that allows patients to access their records over the Internet have already been conducted [2, 3]. In contrast to paper records, electronic medical records should be perfectly legible. Internet-accessible records can be viewed repeatedly and in the context of rich sources of medical information available on the World-Wide-Web, potentially increasing the potency of the intervention [4].

The Health on the Net Foundation Code of Conduct (HON-code) [5, 6] for medical and health Web sites addresses one of Internet’s main healthcare issues: the reliability and credibility of information. The HONcode defines a set of transparency rules to: 1) hold Web site developers to basic ethical standards in the presentation of information; and 2) help ensure readers always know the source and the purpose of the data they are reading. MedHunt [7] is a dedicated, medical, full-text search engine, using a global database including HONCode accredited and selected Web sites retrieved by MARVIN, the HON’s robot. The global database currently includes two types of documents: 1) the accredited Web sites which are updated daily and reviewed manually by the HON team; and 2) the selected health Web sites which have been automatically retrieved from the Web by the robot MARVIN.

WRAPIN has been designed to provide patients and citizens with trusted health information [8]. WRAPIN provides an entirely new facility allowing comparisons of health/medical documents with an interconnected knowledge base, in order to discover whether the information exists in the published literature and to provide a summary of the ideas contained. It will help to determine the reliability of documents by checking the ideas contained against established benchmarks, and enable users to determine the relevance of a given document from a page of search results. This will allow patients to gather information in a trustworthy environment. Moreover, the resources referenced in the WRAPIN database are indexed with MeSH.
key-words. That allows to provide end-users with trusted definitions of health concepts.

A search engine for trustworthy health web sites

The main innovation of WRAPIN is to make available a tool to determine information reliability by automatically checking a document against matching sources from databases of known quality. Figures 1.a and 1.b respectively illustrate the submission of a text to be analysed and an extract of a related Web site of known quality. WRAPIN is fundamentally different from even the best medical search engines currently available, since an entire document may be proposed as the search query. A conventional series of search terms can also be used. A special effort has been made in the indexing process of health resources, since the extraction of MeSH terms from queries is guided by a component based on the UMLS knowledge sources [9].

Whereas other systems display search results in summary form, with bits of text extracted from matching documents, WRAPIN attempts to show how a document matches a query. To better define the results, MeSH terms (and synonyms) are highlighted in the results summaries and throughout the text of matching documents. Figure 2.a and 2.b respectively illustrate an extract from the answer to the question “ablation of a lung tumor”, and the display of a retrieved page where terms of the question are automatically highlighted.

Hence, a patient staying in a given foreign hospital could query WRAPIN and obtain information from data of his/her patient record expressed in his/her natural language. This is particularly significant in Europe where there exist numerous different natural languages, and where people move from one country to another, sometimes everyday, for work reasons.

Another way to adapt the information retrieval process to patients’ queries is to propose to rephrase queries. In the case of WRAPIN, this is done by suggesting a list of terms or relevant qualifiers, known as facets, aimed at making a query more precise. This is illustrated by Figure 3.a in which the initial query is “glioblastoma”. The list of terms used to rephrase a query is created dynamically with regards to the query content and is not predefined. These facets are based on the MeSH thesaurus. WRAPIN exploits other facets based on the medical domain. This kind of facet provides a useful way of helping citizens to better qualify their queries with pre-defined categories linked
to a style of query. For instance, a query related to a disease activates categories such as "overview", "causes and risk factors", "screening and diagnosis", "complications", "treatment", and "prevention", as illustrated by Figure 3.b.

A last feature that WRAPIN exploits is the categorization of sources according to the complexity of the pages they publish. For instance, Medline is considered to be a complex source publishing technical information which is difficult to read by the general public.

A tool for patient empowerment within EHR

Figure 4 shows the general architecture of WRAPIN involving EHR. Extracts of an EHR may serve as a query to the system. Text is analysed, whatever its language (presently five European ones), by a multilingual sub-system represented in a grey rectangle. The result of this treatment is sent to the search engine which exploits an index of medical resources. Results of the query are sent to a module which displays them and suggests rewordings to render the query more accurate. If rewording is performed, a complete cycle is then processed until such time as the user is satisfied with the answers received.

On account of the above functionalities, WRAPIN constitutes an original Web search engine for the health domain. It will allow patients to involve themselves in their own care process, and to ask about and verify information related to their health care. If, for instance, a patient reads in his/her personal French EHR « … exérèse d’un gliome révélé par l’examen anatomopathologique associée à une radiothérapie … » (...ablation of a glioma, diagnosed by a pathology exam, associated with radiotherapy...), and if he/she wants more details in relation with this text, a copy-paste of it allows him/her to query WRAPIN and to obtain a list of trustworthy, related Web sites, as illustrated by Figure 5.

Facing the ever greater sharing of personal EHR, one may wonder what practical use patients can make of this system. Understanding one’s own medical record will certainly be a major concern for patients. In Table1, we compare the WRAPIN features with those of representatives of two main trends in information retrieval: systems based on the popularity of Web sites (e.g., Google), and clustering of Web sites (e.g., Vivisimo).

Discussion

The World-Wide-Web is today probably the world’s largest source of information. Among other functions, it enables patients and others to gather information about their health condition, to verify information they received and perhaps did not entirely take in, and to check the validity of their treatment. In short, it allows them to understand. WRAPIN is a tool they can use to query the health Web. One of the main features of WRAPIN is that it registers only certified quality sites. This allows it to convey reliable health information to patients and other users. Its wide-ranging features and its ease-of-use make it an unrivalled tool empowering the patient in his/her own health care process. Even if patients are tempted to use popular search engines, these are not sufficiently specialized in the medical domain to help them understand their own EHR. The comparison we made between features offered by other search engines shows that WRAPIN, as an instrument dedicated to the health domain, provides a better guarantee in terms of information quality and reliability of Web sources. Thus, thanks to its functionalities, it provides an easy-to-use search engine coupled together with EHR.

The WRAPIN search engine has been operational since 2005, having evolved from a prototype which was tested and improved by evaluation outcomes [8]. The goal of this evaluation was to assess if WRAPIN really met the user needs and par-
Figure 4. EHR as an input to the WRAPIN search process.

Table 1. Comparison of WRAPIN features with those of Google and Vivisimo.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Google</th>
<th>Vivisimo</th>
<th>WRAPIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Querying facilities</td>
<td>Limited in length</td>
<td>Limited in length</td>
<td>Copy and paste a part of a patient record</td>
</tr>
<tr>
<td>Quality of information</td>
<td>Related to page ranking only</td>
<td>Related to page ranking only</td>
<td>Related to trustworthy sources</td>
</tr>
<tr>
<td>Coverage/overabundance*</td>
<td>901,000,000</td>
<td>821,000,000</td>
<td>112,000</td>
</tr>
<tr>
<td>Cross language query**</td>
<td>No</td>
<td>No</td>
<td>Limited to medical terms***</td>
</tr>
<tr>
<td>Display of the results</td>
<td>Colorized keywords from cache or toolbar</td>
<td>Colorize from the toolbar</td>
<td>Colorize keywords on fly</td>
</tr>
<tr>
<td>Refinement/rephrasing</td>
<td>Limited to English</td>
<td>Corpus-oriented</td>
<td>Corpus-oriented plus a multilingual specific access to diseases</td>
</tr>
<tr>
<td>Health literacy****</td>
<td>English facets only for diseases</td>
<td>No</td>
<td>Health information is categorized by its source Direct access to health definitions Health spelling suggest</td>
</tr>
<tr>
<td>Transparency</td>
<td>Commercial</td>
<td>Commercial</td>
<td>Non-governmental</td>
</tr>
<tr>
<td>Conflicts of interests</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend: *number of answers to the "health" query, Nov 26, 2006; ** e.g., the ability of a patient to query his/her health record from a foreign country; *** English, Spanish, French, Italian, Portuguese, German; **** access to easy-to-read information

It is particularly to know if it can help patients and individuals to judge information available on the Web. The analysis of the results made clear the important impact of WRAPIN to meeting needs expressed by great number of users. There were three types of questionnaires: technical, satisfaction and ergonomic. The evaluation was online and compared WRAPIN with another Web search engine. There were four types of queries: free and pre-formatted questions, and free and pre-formatted URL for more than thirty medical specialties. The use of WRAPIN has, on the whole, been perceived as informative, reliable and trustworthy. Yet in a number of cases, the replies given by the system were irrelevant, but it was also true for the other engine. The rate of irrelevance/impertinence should be scientifically reckoned, and finally reduced. It seems the upfront lexical analysis of the queries, as well as the ergonomic of the reformulation leave room for improvement. In the view of the testers, WRAPIN is over-performing other engines when the replies are pertinent. By providing a synthetic and reliable reply to a query, or an assessment of a document further to the submission of an URL, WRAPIN goes beyond the other engines.
and brings a valuable tool to users seeking medical and health care information. It has been felt helpful that humans reviewed Web sites for appropriateness and correctness.

Usually, when citizens look for health information they look for it on the Internet with search engines like Google. It is obvious that, in order to have a more complete information related to various unclear issues of their medical health record, citizens will rush to Google again and try to clarify them. In this case, unfortunately, their search will not worry the reliability of found information. In this paper, we have presented several criteria which show that WRAPIN can be used by citizens in this context and represents a good and interesting alternative to existing non-specialized search engines. However, a more detailed evaluation should be performed in order to validate the relevance of WRAPIN for such a task and to identify more precisely appropriate functionalities needed within an EHR-oriented perspective.

To facilitate understanding of personal medical information by health consumers, the use of accessible, non-technical language is essential. Information complexity is often quantified on the basis of the complexity of words and sentence formulation, for instance, the number of words in a sentence or syllables in a word [10]. In a domain as specific as health, lexical knowledge is of paramount concern. Today, WRAPIN informs users about the a priori complexity of information sources. For instance, Medline is categorized as a technical source compared to Medhunt, since the presence of specific medical terms increases the level of complexity for the general public. Hence, a complexity measure based on the use of a lexicon seems promising.

Moreover, a major effort is still required if patients and non-patients are to have efficient access to health-related information. As noted by McCray in a thorough study into reading capability and the understanding of health-related documents [11], patients need to interact in a variety of health care settings, including doctors’ offices, clinics, and hospitals. They also need to interact with a broad range of health-related information, including treatment instructions, patient education materials, prescriptions, bills, and insurance forms. In addition, they are being asked to take increasingly greater responsibility for their own health-care and disease management. The role of the Internet in the health-care system is just beginning to be understood. Still needed are better tools to assess and modify the comprehensibility of health materials as well as methods for improving access to information. McCray’s conclusion, to which we totally subscribe, is that “health informaticians, developers of health information, and health care providers all need to work together to ensure that everyone has an equal opportunity to access, understand, and use health information”.

Acknowledgments

The WRAPIN project, IST-2001-33260, has been supported by the European Commission. The authors have a special thought for Pr Jean-Raouïl Scherrer who was the originator of WRAPIN and passed away during the project. The authors wish to acknowledge all the members of the WRAPIN consortium, especially Michèle Thonnet, Anne Strauss (MISS, Mission d’Informatisation des Systèmes de Santé), Robert Baud (HUG, Hôpitaux Universitaires de Genève), Patrick Weber (NICE Computing), Pierre Uziel (XR Partner Information Systems), and Juan Bonal (Thales Information Systems). They thank Mr George Morgan who revised their English.

References


Address for correspondence

Michel JOUBERT
LERTIM
Faculté de Médecine
27, boulevard Jean Moulin
13005 Marseille
FRANCE
e-mail: mjoubert@ap-hm.fr
URL: http://cybertim.timone.univ-mrs.fr/