

A Framework for Improving the Quality of Health Information on the World-Wide-Web and Bettering Public (E-)Health: The Medcertain Approach

Gunther Eysenbach^a, Christian Köhler^a, Gabriel Yihune^a, Kristian Lampe^b, Phil Cross^c, Dan Brickley^{cd}

^a University of Heidelberg, Heidelberg, Germany

^b Finnish Office for Health Care Technology Assessment, STAKES, Finland

^c Institute for Learning and Research Technology, University of Bristol, UK

^d W3C World Wide Web Consortium, M.I.T, Boston., USA

Abstract

There has been considerable debate about the variable quality of health information on the world-wide-web and its impact on public health. While central authorities to regulate, control, censor, or centrally approve information, information providers or websites are neither realistic nor desirable, public health professionals are interested in making systems available that direct patient streams to the best available information sources. National governments and medical societies have also recognized their responsibility to help users to identify "good quality" information sources. But what constitutes good quality, and how can such a system be implemented in a decentralized and democratic manner? This paper presents a model which combines aspects of consumer education, encouragement of best practices among information providers, self-labeling and external evaluations. The model is currently being implemented and evaluated in the MedCERTAIN project, funded by the European Union under the Action Plan for Safer Use of the Internet. The aim is to develop a technical and organisational infrastructure for a pilot system that allows consumers to access meta-information about websites and health information providers, including disclosure information from health providers and opinions of external evaluators. The paper explains the general conceptual framework of the model and presents preliminary experiences including results from an expert consensus meeting, where the framework was discussed.

Keywords:

Internet; Public Health; Quality of Information; Ethics; Consumer Health Informatics; MedCERTAIN; Publishing

Introduction

Several studies have evaluated the quality of medical information on various venues of the Internet such as the World-Wide-Web¹, newsgroups² and email consultations³⁻⁵. As

the number of health related websites continues to grow, and information technology and consumer health informatics are becoming integral parts of modern public health concepts and national health care policies in developed countries⁶, implications of Internet information for public health are widely discussed^{7,8}.

Health professionals should take the lead in guiding patients to the best available medical information on the web. Medical organisations and governments have begun to develop national gateways (Healthfinder.gov, NHSDirect/NeLH), portal sites and other forms of infomedaries such as seals of approval ("kitemarks") in an effort to help consumers to locate trustworthy information resources. The challenge of developers of such systems is to establish a reliable and evidence-based process for selecting and recommending certain sites to consumers and to develop criteria for including sites into such infomedaries. The universal question in this process is "What constitutes quality and how can it be assessed?". A new scientific discipline is emerging around the question on how to appraise information for consumers, how to guide consumers to the best available information and how to help consumers in appraising information.

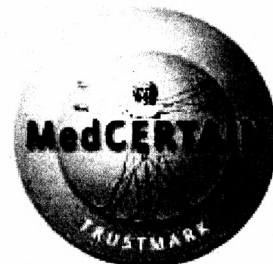


Figure 1- The MedCERTAIN trustmark is a dynamically generated interactive button included on the homepage of certified websites.

This paper explains the approach of MedCERTAIN (MedPICS Certification and Rating of Trustworthy and Assessed Health Information on the Net)⁹. The aim of this EU funded project is to develop software which demonstrates the use of electronic labels (meta-data).

Methods

The MedCERTAIN approach combines aspects of consumer education, helping information providers to implement best practice guidelines, encouraging self-labeling by information providers, external evaluations and monitoring. One main difference to usual “kitemarking” (seal of approval) approaches is that “quality” is not defined as something that can be objectively and universally measured, but that preferences and needs of users can be taken into account and that at the end the user determines whether he/she regards the resource as trustworthy. To obtain the MedCERTAIN trustmark, the information provider has to disclose certain information in a standardised, computer readable format using the “Health Information Disclosure, Description and Evaluation Language” (HIDDEL, formerly known as medPICS) expressed in XML/RDF. The same language can be used by gateways to express opinions of evaluators about other sites.

Information labels - an analogy to food labels

The MedCERTAIN labeling concept – a combination of self-labeling and third-party control - can be best explained by drawing an analogy to food labels.

Empowering consumers to select healthy nutrition is an important public health aim. The question of which “high-quality food” can be recommended is as complex as the question of what health information to recommend. Both questions are difficult to answer universally, i.e. if the individual consumer and the individual circumstances are not known. In other words, only in extreme cases (rotten food or fraudulent information) it makes sense to put tags on foodstuff or websites discouraging (or recommending) their use.

What is needed instead are elements of consumer education, self-labeling and external quality control.

First, consumers have to be educated about healthy components of a balanced nutrition. Different consumers may be asked to look for different things, e.g. some may want to lose weight and may be taught that they should especially look for low fat products; others may suffer from hypertension and should avoid food rich in sodium.

Secondly, producers of food have to display ingredients on standardised labels, telling consumers for example the amount of fat and sodium contained in their products.

Third, these labels (and the products itself) have to a certain degree to be evaluated and compliance with rules be enforced. The US 1990 Nutrition Labeling and Education Act (NLEA) for example regulates clearly how such nutrition

labels should look like, when language such as “fat-free” may be used, which health claims may be used and how they have to be worded.

Taken together, consumer education, self-labeling and elements of external monitoring empower consumers to make informed choices.

MedCERTAIN: Electronic labels for websites

The MedCERTAIN approach plays a similar role on the World-Wide-Web as food labels in the real world. The concept is also based on the pillars of consumer education, encouragement of self-labeling and third-party evaluation.^{9,10}

First, consumers are to be educated and told, what to look for if using a health website. We may for example teach them about privacy issues or about best practices in e-commerce, which internal quality management mechanisms in the production of information one may expect, and so on. The DISCERN group¹¹, Health on the Net Foundation¹², American Medical Association¹³, Internet Healthcare Coalition¹⁴ and the Hi-Ethics group have all contributed to this process and produced codes of ethics or similar documents that aim to educate health information providers and users alike. However, work still needs to be done, as unclear language and lack of consensus among these documents has recently been criticized.¹⁵

Secondly, information providers should display clear labels containing relevant information which allow users to assess the quality and trustworthiness of an information provider and to make informed choices. The first step in the MedCERTAIN certification process is therefore to let the information provider answer a questionnaire requiring disclosure of certain information, such as who is behind the information service, who are significant sponsors, what are internal quality processes, who is the target audience, what is the aim of the website, how it has been evaluated, and so on. This meta-information will be presented to the consumer in a standardised and accessible format (Fig. 2).

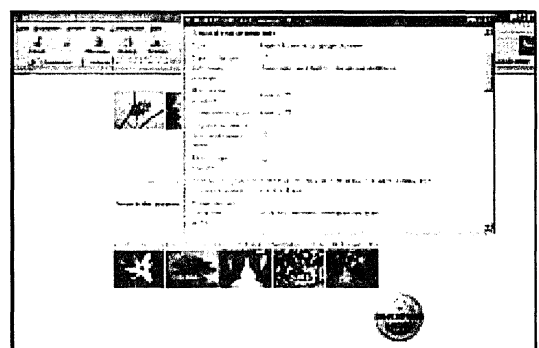


Figure 2- A click on the trustmark connects to the MedCERTAIN server, which will open a window displaying disclosure information and further meta-information entered by the information provider and other sources

By making this meta-information also accessible as com-

puter-readable metadata, users may more specifically search for sites meeting their own needs, or may get advice and alerts if a site does not comply to individual pre-set preferences. This is the concept of “downstream-filtering”¹⁶.

On this level (called level 1 certification), trust develops through active disclosure by the information provider in a standardised form, informing users about his processes on quality management, qualification and training of staff, etc. MedCERTAIN provides a stimulus and a technical framework for implementing web ethics, by providing a common access structure and terminology for disclosure information.

However, as with food labels, a certain degree of external control, monitoring and evaluation is necessary, otherwise everybody could exploit these labels for marketing purposes. It is a familiar phenomenon that webmasters are abusing meta-data to increase traffic to their sites by using deceptive meta-information (e.g. by using the keyword “sex” as description for their websites). Pure self-certification systems may lead to more harm than good. Therefore, MedCERTAIN will, as much as possible, make sure that the meta-information provided by the information provider is correct, adequate and complete, and that formal ethical requirements of the site are complied with. On what is called a level 2 evaluation, a member of the MedCERTAIN team will visit the website and check the resource for compliance with formal criteria and verify the information given by the health information provider. This evaluation can be done by a non domain-expert (Table 1).

Table 1 – MedCERTAIN levels of evaluation

Trust-mark Level	Evaluation by whom	Evaluation Target	Example
Level 1	None (self-certification)	-	Disclosure information
Level 2	3 rd party non-domain expert	Actual structure and process	Checking formal criteria, e.g. provision of authorship
Level 3	3 rd party domain expert	Product (information)	Accuracy; any glaring biases, harmful or deceptive information?
Level 4	Researcher	Outcome	Knowledge change, Validity of health assessment tools

In what we call a level 3 evaluation, medical experts (or other users) will look at the actual product, i.e. evaluate the content of information sources, much as a gourmet tester may taste and review food for a consumer magazine. These higher level aspects of quality (accuracy, comprehensiveness, balanced information) are difficult or even impossible to be rated in a “reliable” and objective fashion. What appears to be “accurate” for one expert may be “inaccurate” for other experts. In cases of disagreement it is more important for the consumer to be able to tell which organisation or individual experts says what about a given resource. When

the user clicks on a MedCERTAIN level 3 trustmark, he will therefore be able to access the meta-information on which organisation says what about a given site. What is difficult to achieve in the real world (if I buy a book it takes considerable efforts to find out what reviewers say what or which organisations recommend it), is – in principle – possible on the world-wide-web on a click of a button, if people use a common vocabulary to describe resources, and if these opinions are harvested and aggregated. This vocabulary is the HIDDEL (MedPICS) vocabulary being developed as one deliverable of the MedCERTAIN project.

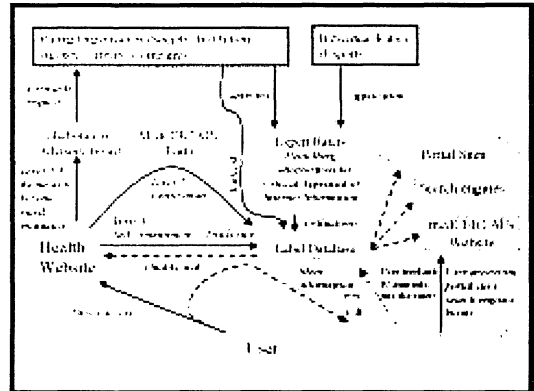


Figure 3- Overview of the MedCERTAIN certification framework

Results and Discussion

Focus group evaluations show a high degree of consumer interest. An invitational expert consensus workshop, held in September of 2000 in Heidelberg brought together 80 medical researchers, health professionals, consumer representatives, representatives of government agencies, inter-governmental and non-governmental medical associations, librarians and computer scientists from 20 different countries to discuss quality issues and the MedCERTAIN model. The concept of a trustmark was defined in a consensus statement¹⁷ as follows:

“Any certificate, symbol, sign, based on information or meta-information (information about information) provided by a third-party that is aiming

- to enhance peoples trust into a certain product, service, relationship, information provider or piece of information perceived as being trustworthy by the organisation issuing the trustmark; and/or
- to enhance the ability of people to evaluate the trustworthiness of information, services or products.”

Table 2 – MedCERTAIN deliverables

	Deliverable	Date Due
D1.1	Setup of project website (http://www.medcertain.org)	8/ 2000
D1.2.	Web-based database of a trusted community of raters (>100 raters)	9/2000
D2.1.	Report on applicability (in the European context), reliability and validity of medPICS	10/2000
D2.2.	Final version of the rating vocabulary (MedPICS 1.0, now called HIDEDEL)	11/2000
D3.1.	Provision of simple mechanisms and tools for raters to rate/annotate websites using bookmarks, what's related menu, bookmark uploads and remote bookmark storage and technology to aggregate these data	1/2001
D3.2.	Provision of more advanced data acquisition tools such as HTML- and JAVA interfaces for the raters and implementing digital signatures	6/2001
D4.1	Develop ratings query API, search/browse interfaces, what's-related service, HTTP proxy add-on. A proposal to W3C for PICS 2.0 services in XML/RDF.	2/2001
D4.2	Provide for user feedback mechanisms and automatic detection of suspicious sites	4/2001
D5.1	Report on pilot Finnish testbed	8/2001
D5.2	Final report on overall rating results	11/2001
D6.1	Report on technical functionality of the rating concept and deployment in other frameworks	8/2001
D6.2	Report on usability and impact of rating concept and service	9/2001
D7.1	Setting up further demonstration projects e.g. in hospitals, libraries and schools in Finland, Germany and elsewhere	9/2001
D7.2	Production of a project brochure for consumers	11/2001

Problems of rating systems identified on the consensus workshop and possible solutions include the following:

Reliability of rating instruments. Jadad¹⁸ was among the first pointing out that few Internet rating systems have evaluated their instruments in respect to reliability and validity. This problem can be generalized to the challenge of selecting, evaluating and endorsing *any* kind of printed or electronic information. One could for example argue that the selection process of medical manuscripts for publication in medical journals has similar problems of low inter-observer reliability. No medical journals has ever published figures regarding reliability and validity of their selection process or instruments used by reviewers. Despite these methodological difficulties, users (producers of the information as well as readers) appreciate the input of external evaluators and the input of an editorial board to improve the manuscripts. The scientific community has accepted peer-

reviewed journals as a useful mechanism for users to access filtered information. And it is interesting to note that also in the scholarly publishing world, the processes of evaluating the quality (peer-reviewing) and making the document physically available can be reversed in a sense that papers may also be published first on pre-print servers and later be tagged with meta-information and post-publication reviews¹⁹. In a way, kitemarking, trustmarking and “gateway-building” on the web is a meta-publishing process, having a similar role as medical journals in selecting information worth to be recommended to users. Such “selecting and publishing” processes will inherently always have a limited degree of “reliability”. It remains to be explored whether rating instruments for health information resources on the Internet can be constructed that have a high degree of reliability. The alternative approach is to inform users about different views from different experts, even if there is considerable inter-observer variability.

Dynamics of content. Another argument frequently brought forward to dispute the feasibility of rating or evaluating websites is that the content of websites changes too frequently. However, this argument is only valid if the unit of evaluation is the *information* on the server, not if the trustmark refers to the *information system* and/or *information provider*. To draw an analogy to the printed world: It is well possible to recommend a printed source such as the BMJ to colleagues, even though the actual content of the journal changes weekly. This is because the recommendation of such sources is based on knowing the process (and perhaps people) behind the information production and having evaluated some samples of the product. If an evaluator has assessed both the process and a sample of the product, he may well decide to trust the information source (and thus the information it offers) for a given period of time. Thus, the function of a trustmark would be “not to guarantee for information correctness or usefulness, but to enhance trust into the information provider” [Heidelberg Consensus Recommendations on Trustmarks]¹⁷. Any trustmark approach can and should only indicate that the information producer has taken a number of steps to increase transparency and ensure quality, and/or that process and samples of the product have been evaluated, but cannot and does not want to guarantee the “accuracy” of information.

Costs. Certainly, third-party evaluation comes not at no costs. The number of health related websites worldwide has been estimated as being around 100,000²⁰ and complete coverage by third party evaluation is impossible. However, medical journals such as the BMJ receive and review several thousands of manuscripts per year by using an international network of evaluators who work for free. The Cochrane Collaboration assesses and reviews thousands of clinical trials, also based on an international collaboration of volunteers. A similar collaborative approach may work on the web. The “Heidelberg Collaboration for Critical Appraisal of Health Information” is a proposed group intending to “help people, patients and professionals to identify health information useful to them”²¹, e.g. by ensuring interoperability of rating services, identifying common standards

for recommending websites and exploring possibilities for decentralised, distributed rating systems by creating a network of evaluators, taking into account the power of the Internet.

Legal and liability issues. A false-negative “blacklisting” could negatively impact the reputation and economic health of the medical site evaluated and lead to law suits against the trustmark authority. A false positive “whitelisting” (recommendation of a “bad” site) could harm the economic or physical well being of a consumer or patient and also lead to a law suit.²² However, it remains to be seen how these possibilities play out in practice. For example, few journal editors have been sued over the false-negative rejection of a medical manuscript later shown to be an important contribution to science.

Acknowledgments

MedCERTAIN is funded by the European Commission, Action Plan on Promoting Safer Use of the Internet.

References

- [1] Impicciatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home. *BMJ* 1997;**314**:1875-1879.
- [2] Culver JD, Gerr F, Frumkin H. Medical information on the Internet: a study of an electronic bulletin board [see comments]. *J Gen Intern Med* 1997;**12**:466-470.
- [3] Eysenbach G, Diepgen TL. Evaluation of cyberdocs. *Lancet* 1998;**352**:1526
- [4] Eysenbach G, Diepgen TL. Responses to unsolicited patient e-mail requests for medical advice on the World Wide Web. *JAMA* 1998;**280**:1333-1335.
- [5] Sandvik H. Health information and interaction on the internet: a survey of female urinary incontinence. *BMJ* 1999;**319**:29-32.
- [6] Eysenbach G. Consumer health informatics. *BMJ* 2000;**320**:1713-1716.
- [7] Coiera E. Information epidemics, economics, and immunity on the internet. We still know so little about the effect of information on public health [editorial; comment]. *BMJ* 1998;**317**:1469-1470.
- [8] McLeod SD. The quality of medical information on the Internet. A new public health concern [editorial]. *Arch Ophthalmol* 1998;**116**:1663-1665.
- [9] Eysenbach G, Yihune G, Lampe K, Cross P, Brickley D. MedCERTAIN: Quality Management, Certification and Rating of Health Information on the Net. *Proc AMIA Symp* 2000;230-234.
- [10] Eysenbach G. Towards ethical guidelines for e-health: JMIR Theme Issue on eHealth Ethics. *J Med Internet Res* 2000;**2**:e7. URL: <http://www.jmir.org/2000/1/e7/>
- [11] Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999;**53**:105-111.
- [12] Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health Websites. *Comput Biol Med* 1998;**28**:603-610.
- [13] Winker MA, Flanagan A, Chi-Lum B, et al. Guidelines for Medical and Health Information Sites on the Internet. *JAMA* 2000;**283**:1600-1606.
- [14] e-Health Ethics Initiative. e-Health Code of Ethics. *J Med Internet Res* 2000;**2**:e9 URL: <http://www.jmir.org/2000/2/e9/>
- [15] Baur C, Deering MJ. Proposed Frameworks to Improve the Quality of Health Web Sites: Review. *MedGenMed* 2000
- [16] Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information [see comments]. *BMJ* 1998;**317**:1496-1500.
- [17] Heidelberg Consensus Recommendations on Trustmarks. *J Med Internet Res* 2000;**2**:e12 URL: <http://www.jmir.org/2000/3/suppl2/e12/>
- [18] Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? [see comments]. *JAMA* 1998;**279**:611-614.
- [19] Eysenbach G. The impact of preprint servers and electronic publishing on biomedical research. *Curr Op Immunol* 2000;**12**:499-503.
- [20] Eysenbach G, Sa ER, Diepgen TL. Shopping around the internet today and tomorrow: towards the millennium of cybermedicine. *BMJ* 1999;**319**:1294 URL: <http://www.bmj.com/cgi/content/full/319/7220/1294>
- [21] Draft Heidelberg Collaboration Statement of Purpose. *J Med Internet Res* 2000;**2**:e11 URL: <http://www.jmir.org/2000/3/suppl2/e11/>
- [22] Terry NP. Rating the "Raters": Legal Exposure of Trustmark Authorities in the Context of Consumer Health Informatics. *J Med Internet Res* 2000;**2**:e18 URL: <http://www.jmir.org/2000/3/e18>

Address for correspondence

Dr. Gunther Eysenbach, Dept. of Clinical Social Medicine, University of Heidelberg, Bergheimerstr. 58, 69115 Heidelberg, Email: ey@yi.com URL: <http://yi.com/ey> , MedCERTAIN site: <http://www.medcertain.org>