MEDINFO 2017: Precision Healthcare through Informatics
A.V. Gundlapalli et al. (Eds.)

doi:10.3233/978-1-61499-830-3-1004

Accessing Reliable Health Information on the Web: A Review of the HON Approach

Célia Boyer[a], Arnaud Gaudinat[d], Allan Hanbury[e], Ron D. Appel[b, c], Marion J. Ball[b, c], Michel Carpentier[b, c], Jan H. van Bemmel[b, c], Jean-Paul Bergmans[b], Denis Hochstrasser[c], Donald Lindberg[c], Randolph Miller[b], Jean-Claude Peterschmitt[b, c], Charles Safran[b], Michèle Thonnet[b], Antoine Geissbühler[b]

[a] *Executive Director – Health On The Net Foundation, Geneva, Switzerland*
[b] *Foundation Council members – Health On The Net Foundation*
[c] *Co-founders – Health On The Net Foundation*
[d] *BiTeM group, HES-SO/HEG Geneva, Information Science Department, Carouge, Switzerland*
[e] *Institute of Software Technology & Interactive Systems, TU Wien, Vienna, Austria*

Abstract

Accessing online health content of high quality and reliability presents challenges. Laypersons cannot easily differentiate trustworthy content from misinformed or manipulated content. This article describes complementary approaches for members of the general public and health professionals to find trustworthy content with as little bias as possible. These include the Khresmoi health search engine (K4E), the Health On the Net Code of Conduct (HONcode) and health trust indicator Web browser extensions.

Keywords:

Quality Indicators; Internet; Trusted Health Information

Introduction

The Web has become an important and significant health-related resource. (N.B., the terms resource, website and webpage are used interchangeably herein for health content on the Web.) By providing copious medical and healthcare information, the Web has empowered the public. Some 60% of Europeans [1] and 72% of Americans [2] have used the Web to address healthcare questions. Internet use offers many conveniences, including accessibility 24 hours a day, anonymity when conversing with others, and analyses on a wide range of subjects. The Web provides information on how to contact local and national experts, access to sage opinions (e.g., best therapies, effectiveness research), and connectivity to a massive quantity of health information resources. Alarmingly, some 75% of people using the Web for health purposes do not differentiate among the facts they obtain in terms of accura- cy or credibility [3].

A 2013 Pew Research survey indicated that eight out of ten online health inquiries start by using a search engine‖ [2]. Somewhat uniquely, Web-based health information can directly impact a person's health status, providing benefits, but also harm - in a few extreme cases, resulting in death [4]. Naive users do not realize that search engine results mix trustworthy information with unreliable and even purposefully manipulated health information. The most commonly used search engines currently provide no support to differentiate quality assessments of health content. Even worse, general search engine results can lead to biased medical content (deviation from the truth"), and can lead users to make inappropriate healthcare decisions [5]. PageRank was identified early as a promising way for health websites to indicate to consumers that they were providing quality information to consumers [6].

However, mechanisms underlying page rankings, e.g., hyperlinks and browsing history, merely indicate popularity of a webpage is [7]. Many organizations have previously attempted to guide Web users to high-quality health websites, but this remains a challenge. For example, Google investigated possible statistical estimating schemes to judge the correctness of

facts [7]. Nevertheless, in the rapidly evolving health domain, facts are not absolute – the best test or therapy for a condition today may not be so tomorrow. Information correct for one person's situation may not be for others. Thus a crucial need exists to assess the trustworthiness of a given health Web resource. This paper presents recent new quasi-automated methods to filter healthcare web site content to assess trustworthiness and readability. The ultimate goal is to promote laypersons' easy access to quality health information [8, 9].

The problem of trust

The concept of trust is elusive to define because a multitude of factors contribute to it. Grandison and Sloman [8] defined the quality of a page as content targeted at the right audience. Gil and Artz [9] listed 19 different factors affecting how users determine trust in webpage content, including: topic, context and criticality, popularity, authority, recommendation, bias, appearance, honesty and currency of information.

In 1996, the Health On the Net (HON) Foundation determined that the International Committee of Medical Journal Editors (ICMJE – www.icmje.org/) conventions and recommendations for printed medical journals could be applied to online information. Thus, HON developed the HONcode, a set of ethical, honesty, transparency and quality standards related to health website content production. Note that HONcode certification provides a metric that determines if the processes underlying a website's construction and maintenance conform to standards of excellence; it does not evaluate the veracity of the site's content per se. The HONcode certification process implies that the health website editors are both motivated and committed as they need to invest time to meet HONcode criteria into the future. The website editors receive no direct financial return or incentive for such efforts from HON. Moreover, in 2014 the annual assessment of certified websites was changed into a contribution-based program enabling HON to continue offering certification. Prices range from 50 euros for not-for-profit websites to 325 euros for high-ranked commercial websites. Certified health websites agree to display the HONcode seal on the website (Figure 1), to be continuously scrutinized and to implement the recommendations made by HON.

Figure 1 – HONcode seal displayed on certified websites

Based on more than two decades of research and pragmatic experience accrediting websites using HONcode criteria [10], the authors now believe that those criteria capably capture the trustworthiness of a health Web resource. Studies have demonstrated that websites that conform to HONcode quality standards contain more reliable health information than randomly selected health-related websites [18]. The presence of the HONcode symbol on a website informs the user that the site respects a quality standard. Thus the user, when faced with multiple, contradictory and sometimes questionable information, can trust those sites that are HONcode certified. Unfortunately, HON lacks the resources to review and revisit all health-related websites continuously. Also, HONcode certification is a voluntary process whereby health website editors must request that HON reviews their sites, which requires awareness of the HONcode initiative. So, the problem at hand is how to estimate the trust level of an arbitrary health website by quasi-automated means.

Previous relevant work on health website trustworthiness

Other groups have explored natural language processing (NLP) approaches to facilitate access to quality health information [11, 12]. The authors and colleagues have also examined automated identification of information trustworthiness using the HONcode criteria combined with an experimental multilingual automated detection system [13, 14]. The latter NLP-based approach also included readability-level scoring, which rates how easy it is for the average user to understand the web page content. In addition to developing methods that analyze content that users have found, HON has also devel- oped services that help Web users to directly access trustwor- thy health information. HON has actively participated in the development of the Khresmoi for Everyone (K4E – http://everyone.khresmoi.eu) health search engine, which pro- vides access to trustworthy health websites. The JAMA benchmark criteria, the DICERN score (http://www.discern.org.uk) and Medplineplus.gov and the HONcode have been widely covered and compared in the literature for assessing health and medical-related websites [15, 16, 17]. The Table 1 below summarizes the specificities of the four instruments.

Why automate HONcode certification?

Previous approaches to the HON certification process have been described elsewhere [10, 13]. Because HONcode certification has been until now carried out manually, the number of sites that can be assessed by HON reviewers on a daily basis is limited. This article focuses on quasi-automated detection of compliance with the HONcode criteria in a manner that complements the human task. Yet, ample room exists for greater awareness of online health information quality amongst the general public. Additionally, it is not easy to determine which websites are certified when using a general search engine. This is why, in addition to the HONcode Web browser extension, the automated detection of HONcode principles has been studied and developed.

Table 1 – Comparison of main initiatives to identify trustworthy online content

Initiatives	Types of initiative	Main differences
DICERN	16 questions	No implementation
JAMA benchmark criteria	4 main criteria	No implementation
HONcode certification	8 criteria Certification process conducted by trained HON health professionals	See Table 2 - Voluntary approach - Motivate health website editors to improve the production process of their sites - Search engines with access to certified websites - Manually assessed and curated according to published guidelines
MedlinePlus	Manually selected webpages	- Manually curated list according to 15 criteria - From the U.S.A. government

Methods

HONcode quality criteria

The proposed approach is based on the HONcode. Table 2 lists the HONcode principles.

Table 2 – The eight HONcode principles

Principle	Detail
HC1 - Authority	Indicates the qualifications of the authors
HC2 - Complementarity	Information supports, does not replace the doctor-patient relationship
HC3 - Privacy policy	Respects the privacy and confidentiality of personal data submitted to the site by the visitor
HC4 - Attribution of reference criteria date	Cites the source(s) of published information Dates medical and health pages
HC5 - Justifiability	Backs up claims relating to benefits and performance
HC6 - Transparency	Presentation is accessible; contact information is present
HC7 - Financial disclosure	Identifies funding sources
HC8 - Advertising policy	Clearly distinguishes advertising from editorial content

Automated detection of HONcode principles

HON has conducted extended research on benchmarking and assessing natural language processing (NLP) methods for multilingual automated detection of HONcode principles [13, 14, 19]. The first step of the manual and any automated certification process is to determine where on the candidate web site information relevant to each HONcode criterion appears. This requires around 25% of the time of manual HON assessors. The automated machine learning algorithm training data set (ground truth) comprises the notes that previous expert manual HON reviewers created during their reviews – and include extracts of texts justifying that each principle was met [13, 14] in real life settings. For each HONcode principle, a specific classifier has been created, except for the Attribution principle (HC4), which has been divided into one classifier focusing on references (HC4-Reference) and one on date (HC4-Date), giving nine distinct classifiers (Figure 2). The resulting panel of NLP algorithms has been tested, compared, evaluated, fine-tuned and applied in order to develop classifiers for each of the HONcode principles. This resulted in the selection of the Naïve Bayes (NB) supervised machine learning algorithm [13,14]. The classifiers are used to identify the presence of the HONcode criteria into health webpages in English and French.

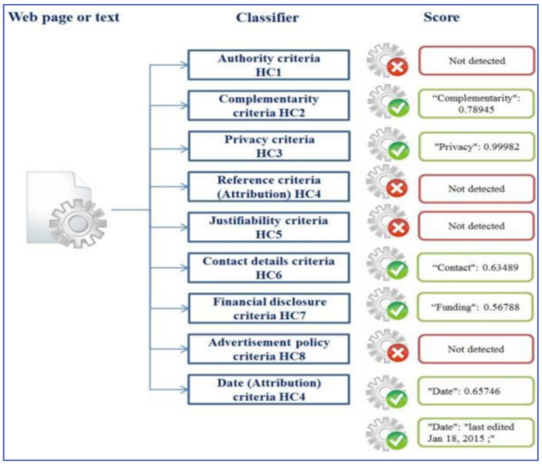

Figure 2– 9 classifiers for automated detection of HONcode criteria

All of the classifiers are applied to Web page as a page can mention several principles. All the principles detected within a given domain name are associated to that domain name. A page is considered classified as per a given principle once it has reached a certain threshold set by the benchmarks [13, 14]. The classification is done using the key terms specific to each HONcode principle as determined by machine learning. The text in colored boxes in Figure 3 illustrates how the classifier score is elaborated based on a collection of key terms identified. The HONcode reviewer extracts the full text corresponding to the principle, as shown in yellow in Figure 3. However, the automated detection of the presence of HONcode criteria only provides an overall estimate of a health website's level of credibility.

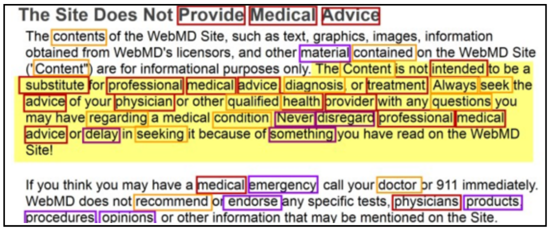

Figure 3– Terms detected by the automated system vs. those detected manually by HON assesor for HC2-Complementarity criterion

An evaluation of how the classifiers work on more complex and real health websites, as opposed to simply extracts, as was done for the systematic evaluation of the classifiers, was conducted [13]. This evaluation comprised 27 highly complex English-language health websites evaluated both by the expert HONcode reviewers and the automated HONcode system. Prior to the use of the classifiers on the health websites, various processes to extract the content of the websites were run. These websites were then parsed by the automated system. In parallel, the websites were assessed manually by the HONcode experts. The average manual HONcode evaluation time for a health website is 80 minutes. Human reproducibility of the results of a HONcode certification ranges from 80% to 95% depending on the principles [13]. Lastly, the results are compared with the HONcode experts verifying the compliance of HONcode criteria (the automated system only detects the presence of words characterizing the HONcode criteria). Globally the automated detection of HONcode criteria performs well for most of the criteria. This evaluation has led to improvements in the criteria related to HC2-Complementarity, HC4-Date and HC8-Advertising policy. Further investigation shows that for the HC4-Date criterion, the named entity recognition (NER) technique was better adapted than machine learning and a sliding window as a classification unit was necessary for the HC2-Complementarity criterion in order to detect text with other criteria included in the same page. Upon further evaluation, these improvements demonstrate their efficiency [19] and thus have been adopted and used within the automated detection of HONcode criteria (Figure 5).

Readability level of health content

The goal of introducing a readability level is to determine how difficult it is to understand medical or health webpages. The idea is to provide users with access to documents targeted at their level of understanding, a level that evolves over time. How difficult is it to understand the health information on a given webpage? Gil and Artz [9] define access to information adapted to the audience as an indicator of quality. Readability levels have been widely investigated from linguistic and stylistic points of view [20], but little has been done at the level of the medical and health domain where complex terminologies are often used, rendering information difficult to understand.

Further investigations have been conducted using machine learning algorithms to categorize the complexity of health information. The readability level is calculated taking into consideration the length of a sentence and the vocabulary and syntax within the medical context [14]. A readability level in one of three categories—easy, moderate or difficult—is then assigned to each health webpage, text or document analyzed.

Results

HON proposes two ways for the public at large to access quality health information: (1) through a dedicated health search engine (such as K4E) with a selection of trustworthy and adapted resources available for the readers; (2) through a browser functionality, including the filtering and highlighting of certified HONcode websites, the automated detection of HONcode criteria and a readability indicator. As general search engines are most often the gateway that laypeople use to access health information online [21], the latter has been an important strategy. In addition, the automated system to detect HONcode criteria aims to assist the HON assessors within the evaluation of a health website's HONcode conformity. Therefore, the automated assistance in conducting HONcode reviews may help in accelerating the current time-consuming tasks of HONcode certification and ongoing surveillance.

Dedicated health search engine

The K4E search engine offers a curated list of online health resources that have been manually checked for quality. K4E includes various functionalities related to quality, such as query formulation, readability and trust indicators. K4E is an alternative to general search engines as it reduces content biases in the search results as the index is curated, the results are based only on the relevance according to search terms, and the organizations behind the search engine are impartial (no advertising). The automated HONcode detection system and the readability level have been integrated into K4E as a demonstrator (Figure 4): a) the trust level indicates to what extent the user can rely on the information provided on the site as determined by the automated HONcode detection system; and b) the readability level indicates how hard it is to understand the health information. The results in a) provide the overall percentage of the HONcode criteria automatically detected and lists those not found for a given health domain.

Usability testing was conducted in a —real-life‖ setting with members of the general public seeking health information online. The evaluation included two types of feedback: the informal feedback given during the session and recorded by the usability testing software Morae (TechSmith's usability testing software, 2014, version 3.3.3.) and the answers to the standard system usability scale questionnaire (SUS), which presented 10 standard SUS statements used to measure usability perceptions and 17 specific statements related to the K4E search engine. The usability test confirmed that the level of —readability‖ and trust indicator score are important user requirements, with a score of 4.34 out of 5 in the Likert scale: strongly disagree (1) to strongly agree (5) [22].

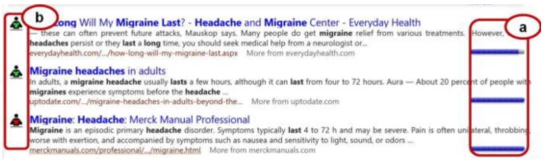

Figure 4 – Trustworthiness and readability filtering for "migraine" with K4E

Quality indicators in a Web browser extension

Using HONcode certified websites as its base, HON developed a Web browser plugin that enriches general search engines (Yahoo, Bing and Google) results with the HONcode seal when health websites are certified. This limits the search only to certified websites and thus reduces its use.

In order to complement the functionalities offered, HON has developed the Health Trust Indicator Web browser extension that includes quality indicators such as the readability level and the results of the automated detection of the HONcode principles on health websites (Figure 5). The readability score is for a given health webpage while the automated HONcode detection indicator is for the whole health website.

This Web browser plugin allows anyone using these general search engines to know for each result the level of trust and of the site providing the information (Figure 5). Then, the user can select trustworthy health pages and avoid information from websites with a low trust score.

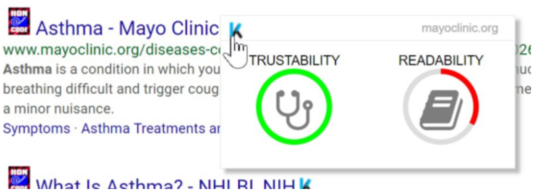

Figure 5 – Google results using the Health Trust indicator Web browser extensions

Discussion

The HONcode and Health Trust Indicator Web extensions currently only have a limited impact as people need to be aware of their availability and the quality issue in order to choose to install such services. A solution would be to have certified websites highlighted directly in major search engines without using a Web extension. This tricky issue was explored from 2006 to 2009 using Google Co-Op Topics. HON tagged and labeled HONcode certified webpages that had been manually categorized according to health subject. These webpages were searchable by end users via the beta version of Google Co-Op. However, the results of this service were inconclusive because users did not use it and it was not financially rewarding. It was eventually dropped.

The K4E vertical search engine specializes in health websites, applying specific domain knowledge in the collection of content and in indexing and query formulation. However, only a limited number of people use health search engines as most people favor the convenience of general search engines.

In addition, K4E does not include a HONcode page rank, which could highlight all links from certified websites and offer a popularity score within the limited circle of certified websites. This idea will be studied for implementation within KConnect's further development.

HON also attempted to mobilize the crowd (crowdcrafting.org/project/healthwebsiteannotationtest/) to assess websites according to the eight HONcode principles with no success, showing that such a task is too large and complex. However, specific and binary tasks can be achievable, such as to tick yes or no if the date was correctly retrived by the automated detection system. Thus the next step will be to propose a prototype with a crowdsourcing function associated with the automated HONcode principle detection system. An assessment and evaluation of the task to be performed by the crowd will be conducted as the task should be simplified as much as possible.

HON contributes to webmaster education but seems to have less direct impact on final users. One part of the solution for health and for information in general found on the Internet has to be education and the development of awareness and critical thinking. Information literacy should be taught rapidly and continuously as soon as a child is able to navigate on the Internet [23].

Conclusions

The HONcode is the most used model for the identification of health sites that are transparent and respect quality criteria [24]. However, more effort in terms of access to trustworthy health information is necessary. Twenty years after the inception of HON, there is no solution that is able to address at a large scale the problem of trust on the Internet, particularly for health information. For the past 20 years, HON and its partners have however made a valiant attempt to investigate solutions to the issue of quality health information online. Access to information is mainly a matter of available content, and a matter of search engine algorithms with strong biases as shown in [5]. The so-called bubble filter phenomenon accentuates the problem; another problem is the profiling of the user in search engine results. These drawbacks indicate that substantial room for improvements in currently available common search engines. An alternative approach would create vertical search engines dedicated to health.

An approach combining quasi-automated trust level indicator categorization with manual HONcode certification offers the possibility to cover health websites that have not been evaluated manually. Such reviews can complement the human work involved in evaluation and monitoring of certified health websites.

The HONcode quality assessment tool, which works through a Web browser extension, is straightforward and allows the user to identify if a site is reliable and respects the HONcode principles. The additional automated quality indicators available via the KConnect Web browser extension have shown through usability testing that users appreciate and favor quality online health information and tools when they are aware of them.

Acknowledgements

The research and evaluation activities performed and presented in this article have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 644753 (KConnect).

References

[1] European Commission, Europeans Becoming Enthusiastic Users of Online Health Information, in: *Digital Agenda for Europe*, European Commission, 2014.

[2] S. Fox and M. Duggan, *Health Online 2013*, Pew Research Center’s Internet & American Life Project, 2013.

[3] S. Fox, *Online Health Search 2006*, Pew Research Center: Internet, Science & Tech, 29 October 2006.

[4] N. Khomami, Woman dies after taking 'diet pills' bought over internet, in: *The Guardian* April 2015.

[5] R. W. White, and A. Hassan,. *Content bias in online health search.* ACM Transactions on the Web, 2014, 8(4), 25.

[6] K.M. Griffiths, H. Christensen, Website Quality Indicators for Consumers, in: *J Med Internet Res* (2005);7(5):e55.

[7] Xin Luna Dong et al., Knowledge-based trust: Estimating the trustworthiness of web sources, in: *Proceedings of the VLDB Endowment 8.9,(2015)*: 938-949.

[8] T. Grandison. and M. Sloman, *A Survey of Trust in Inter- net Applications*, Institute of Electrical and Electronics Engineers (IEEE) Communications Surveys, (2000).

[9] Y. Gil and D. Artz, Towards content trust of web re- sources, in: *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, (2007) volume 5, is- sue 4, 227-239.

[10] C. Boyer, A. Gaudinat, V. Baujard, A. Geissbuhler. HON Foundation: assessing the quality of health web pages all over the world. In *Medinfo 2007 proceedings Building Sustainable Health Systems* (2007) (p. 1017). IOS Press

[11] Abbasi, A., F. M. Zahedi, and S. Kaza. 2012. Detecting Fake Medical Websites using Recursive Trust Labeling. *ACM Transactions on Information Systems* 2012.

[12] P. Sondhi, V. V. Vydiswaran, & C. Zhai, Reliability pre- diction of webpages in the medical domain. (2012). In *Euro- pean Conference on Information Retrieval* 219-231. Springer Berlin Heidelberg.

[13] C. Boyer and L. Dolamic, Automated Detection of HON- code Website Conformity Compared to Manual Detection: An Evaluation, in: *J Med Internet Res*, (2015); 17(6):e135.

[14] A. Hanbury, C. Boyer, L. Dolamic and J. Palotti, *Report on automatic document categorization, trustability and readability*, KHRESMOI Deliverable D.1.6, 2013

[15] W. M. Silberg, G. D. Lundberg, & R. A. Musacchio. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor–Let the reader and viewer beware. *JAMA* (1997) *277*, 1244– 1245.

[16] I. Bizzi, P. Ghezzi, & P. Paudyal, Health information quality of websites on periodontology. *Journal of Clinical Periodontology*, (2017) *44*(3), 308-314.

[17] AZ Nghiem, Y Mahmoud, R. Som. Evaluating the quality of internet information for breast cancer. Breast. (2016); 25:34-37.

[18] A. U. Joury, M. Alshathri, M. Alkhunaizi, N. Jaleesah, and J. M. Pines, Internet Websites For Chest Pain Symp- toms Demonstrate Highly Variable Content and Quality, in: *Academic Emergency Medicine.* (2016)

[19] C. Boyer, L. Dolamic, P. Ruch and G. Falquet, Effect of the Named Entity Recognition and Sliding Window on HONcode automated detection of HONcode criteria for mass online health content, in: *Proceedings of the 9th In- ternational conference on Biomedical Engineering Systems and Technologies (HEALTHINFO)*, 2016, vol. 5: 151-158.

[20] R. Flesch, A new readability yardstick, in: *Journal of applied psychology*, 32(3), (1948), 221.

[21] M. De Choudhury, M. Ringel Morris and R. W. White, *Seeking and Sharing Health Information Online: Compar- ing Search Engines and Social Media*, 1365–76, ACM Press, (2014).

[22] C. Boyer, J. Hajic, A. Hanbury, M. Kritz, N. Pletneva, P. Schneller, V. Stefanov, Z. Ureskova. 2014 *Report on the extensive tests with final search system KHRESMOI 2014*. Technical Public deliverable, August 2014.

[23] Eisenberg, Michael B., Carrie A. Lowe, and Kathleen L. Spitzer. Information literacy: Essential skills for the infor- mation age. *Greenwood Publishing Group,* 2004.

[24] P. Wilson, How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health in- formation on the internet, in: *BMJ : British Medical Jour- nal* 324, no. 7337 (9 March 2002): 598–602.