An automated HONcode detection system informs internet users of HONcode compliance

Célia Boyer\textsuperscript{a}, Lijilana Dolamic\textsuperscript{b}, Mayoni Ranasinghe\textsuperscript{c}, Vincent Baujard\textsuperscript{d}

\textsuperscript{a} Health On the Net Foundation, Geneva, Switzerland

Summary

The HONcode provided by Health On the Net (HON) foundation is the most successful third party certification initiative. In nearly 20 years, it has acquired a database of over 8000 trustworthy health websites, has been translated into over 30 languages and become not only a well-respected name amongst health information providers but also an increasingly well-recognised brand amongst health information end-users. The HONcode, having begun in the mid-1990s has managed to stay current and relevant two decades on, because of its ability to change and remain relevant to the present times. The past few years have been no different, and HON has been heavily involved in bringing out certain updates to further its mission of unrestricted trustworthy health information online. In this paper we present the HONcode certification process and the main changes taking place in the certification process to enable its continued success and sustainability.

Key words: HONcode; certification; information quality; internet

Introduction

The internet has become an important source of health-related information for patients. According to a recent study, not less than 60% of Europeans [1] and 72% of Americans [2] use the internet in a quest for health-related information. As 38% of these health information seekers stated that they have managed the medical condition on their own, without asking for professional help [2], the need for the medical information available on the internet to be trustworthy is of the utmost importance.

The most popular approaches to achieving this goal are self-evaluation such as DISCERN [3] or third party certification such as HONcode provided by Health On the Net (HON) foundation. The HONcode was established in 1995 as a pragmatic solution to help end-users to identify transparent health-related information online and differentiate it from information that might prove to be incorrect or even potentially dangerous. The HONcode has experienced a constant growth in number of certifications and has been translated into a large number of languages, thus enabling websites in other languages also to be regulated according to the HONcode certification. The speed of growth of the internet content and the ease of publishing calls for the certification process to become simple and flexible. The easy access to the information is closely related to the advances in the domain of information retrieval, given that 77% of health-related searches start from a general search engine [4]. The possibility to filter health websites with the use of a tool automatically detecting HONcode quality criteria as an add-on to a search engine could be an excellent tool.

HONcode certification process

A study [5] conducted by IPSOS found that 26% of the respondents knew whether the health website they visit is HONcode certified. As a result recommendations such as “the HONcode seal is a reliable indicator of website quality, and we can confidently advise our patients to search for this marker” [6] and “physicians should recommend the HONcode seal to their patients as a reliable indicator of website quality or, better yet, refer patients to sites they have personally reviewed” [7] were made. Studies conducted over the years have demonstrated that websites bearing the HONcode seal obtained higher quality and content scores [8, 9]. Others recommend that more patients should be directed toward HONcode certified websites [10]. Being a non-governmental organisation in official relations with the World Health Organization is in itself a demonstration of the dedication of HON to the promotion of reliable health information. The HONcode certification motivates behavioural changes in the production process of health websites. Websites requesting the HONcode already make changes during the production process. A study carried out in 2010 demonstrated that only 0.6% of health websites not requesting the HONcode certification respect the eight HONcode ethical standards vs 89% of certified websites. The certification process led health websites to respect the ethical and quality standards of HONcode, with full disclosure of the production process of the health website [11]. To initiate the HONcode assessment process (figure 1), a website’s editor must voluntarily request a review via an online form. The review starts with a self-assessment phase. Once a site has requested the HONcode certification a HONcode expert assesses the candidate website using precise guidelines for each principle. When the principle justification is found on a page, the extract and the web address are added to the HONcode file and stored into a database. If a principle is totally or partially not respected, the recommendations are sent to the site editor.
This educational process for the information provider [12] gradually transforms the website into a transparent and trustworthy one. Once the health website respects all the HONcode principles, the unique HONcode identification and seal are issued. The seal can be revoked; the website can lose the certification if it no longer respects the HONcode principles and chooses not to comply with them despite repeated warnings from the HONcode team respecting the quality procedure.

The HONcode has also been instrumental in creating public awareness of the importance of discriminating the health information available online. Given that the internet is growing at an immense rate, we can assume that the amount of online health information is growing accordingly. Not all of this information will be credible. In our experience, we have seen more “bad” health websites than “good”.

Given that the HONcode certification is a manual process for the time being, it is of paramount importance that the process remains unchanged, regardless of the reviewer. A study was carried out to evaluate the robustness of this process. The results demonstrated that the HONcode is, as shown below, robust and replicable.

1 HONcode guidelines: http://www.hon.ch/HONcode/Guidelines/guidelines.html

Results on the HONcode inter-reviewer agreement level

Thirty-six websites were evaluated for each HONcode criterion by three HONcode senior reviewers. The inter-reviewer agreement was evaluated using both percent agreement and Fleiss’ kappa [13].

In table 1 the values of Fleiss’ kappa are rather small when compared with percent agreement. For the criteria “complementarity” the kappa value of −0.113 indicates disagreement, contrary to the percent agreement of 79.63 for this criterion. Two effects have been documented that might cause the misrepresentation of the inter-rater reliability by kappa [14]. The prevalence problem appears when one observation is coded more often than others, resulting in the kappa estimation being very low, which is the case for the “complementarity” criterion in our study. Taking into account the particularity of the data for this criterion, the kappa would not be the correct statistic to be used. These results show that the probability of the expert giving an incorrect evaluation is quite low, especially for “easy” criteria such as “contact details”. However, this probability is somewhat higher for more “complicated” criteria such as the “justifiability” criterion. This brief study identifying the level of agreement between expert reviewers gives the baseline with which to evaluate the similarity of results between the automated and the manual systems.

HONcode reassessment contribution system

One of the reasons for the success of the HONcode certification has been its ability to remain dynamic and to adapt to the changing requirements of the internet, such as the incorporation of guidelines relevant to social media and Web 2.0 or interactive websites. HON has always been ready to take risks and make the necessary changes to ensure the continued improvement/enhancement of the HONcode certification.

Since its inception, thanks to government and public support, HON has provided all its services completely free of charge, including the HONcode annual assessment. Currently, under less economically stable times, our governmental sponsors have had to make fiscal compromises. It is for that reason that sponsors have encouraged us to become financially independent. A change in this direction was carried out in 2014, when the HONcode reassessment review was made a contribution-based service, where websites make a payment for the annual review to take place.

Table 1
Inter-reviewer agreement, percent vs Fleiss’ kappa.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Percent agreement (%)</th>
<th>Fleiss’ kappa</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authority</td>
<td>92.59</td>
<td>0.745</td>
<td>Substantial agreement</td>
</tr>
<tr>
<td>Complementarity</td>
<td>79.63</td>
<td>−0.113</td>
<td>Poor agreement</td>
</tr>
<tr>
<td>Privacy</td>
<td>85.19</td>
<td>0.614</td>
<td>Substantial agreement</td>
</tr>
<tr>
<td>Reference (att.)</td>
<td>88.89</td>
<td>0.756</td>
<td>Substantial agreement</td>
</tr>
<tr>
<td>Justifiability</td>
<td>74.07</td>
<td>0.463</td>
<td>Moderate agreement</td>
</tr>
<tr>
<td>Contact details</td>
<td>95.37</td>
<td>0.471</td>
<td>Moderate agreement</td>
</tr>
<tr>
<td>Financial disclosure</td>
<td>87.04</td>
<td>0.716</td>
<td>Substantial agreement</td>
</tr>
<tr>
<td>Advertising policy</td>
<td>85.19</td>
<td>0.691</td>
<td>Substantial agreement</td>
</tr>
<tr>
<td>Date (attribution)</td>
<td>79.63</td>
<td>0.492</td>
<td>Moderate agreement</td>
</tr>
</tbody>
</table>
This was done to ensure sustainability of the HONcode as well as to maintain the independence of the HONcode certification. So far, the response has been quite optimistic and we are hopeful that the uptake of the contribution-based service will continue to rise. As the introduction of fees to this service can potentially interfere with the “access for all” mission of HONcode, HON is currently in the process of creating a “HONcode for All” programme of sponsorship for websites who are unable to finance their own HONcode reassessment. Our mission is, and always has been, the provision of trustworthy health information to all with no exclusions, and in creating this programme it is our goal to ensure this, despite any changes to the certification process.

**Limitation of the certification**

Three main concerns arise when we speak of the HONcode certification. The certification is voluntary and this limits its coverage. Although there is a significant number of websites requesting the HONcode certification, there is also a large number not doing so, simply because they are not aware of such a standardisation code. The second concern comes from continuous evaluation of the websites. The HONcode seal is attributed to the website at a certain moment. It guarantees respect of principles at that given moment; however, continuous checking is difficult because of the volume of data to be monitored. And, finally, another concern is the scalability of the HONcode manual certification. If all health websites requested the certification how could HON respond to the demand while also carrying out other review activities such as the annual reevaluation? One answer to all three limitations would be automation of the review process. HON is currently in the process of developing an automated system and has already released an automatic tool with a few functionalities which may one day be the stepping stone to a fully automatic HONcode review system.

**Automatic HONcode principles detection applied to all health websites**

One challenge that the HONcode certification faces is the continuous surveillance of already certified websites. An automated detection system would be a solution to this problem. Within the KHRESMOI project, HON developed an automated system to identify HONcode conformity of health websites [15] and the initial results have been very encouraging. Further development of this system is ongoing in the scope of the Kconnect project. Illustrated in the figure 2, this tool relies on the application of machine-learning algorithms [16, 17]; it is trained on the HONcode database (figs 1 and 2) of accredited websites. It is currently applied to medical web sites and shows up to 78% precision and 73% recall. The contingency computed between precision and recall indicates that some of the principles of the HONcode, such as “privacy” with precision/recall of 92%/90%, are rather easy to detect. “Justifiability” with the precision/recall of only 45%/33% proves to be the most difficult. Moreover, it appears to be confused with the “complementarity” principle.

Figure 3 gives a sample page conforming to the “complementarity” criterion. In this page the information that the expert is looking for in the process of manual evaluation is marked in blue. Additionally, the terms that the automated system identifies as important for this criterion are surrounded by different colours depending on their level of importance (red – most important, green – least important). In summary, the future of identifying quality, trustworthy health information on the internet will depend on development of advanced search engines with fine-tuned criterion-matching abilities that can guide users to reliable health information websites. An automated HONcode detection system would be the solution to all three limitations of the manual certification process. Automated assistance in conducting HONcode reviews can expedite the current time-consuming tasks of HONcode certification and ongoing surveillance as well as eliminate the chance of human error.

**Automated system for HONcode detection in action**

“KHRESMOI for Everyone” (K4E) is a multilingual, multimodal search and access system for biomedical information and documents. K4E provides access to HONcode-certified websites or websites that have been manually se-
lected, enabling readers to access this information directly without having to wade through multiple pages of dubious material to get there. Besides the search results list containing the webpages most relevant to the user’s query, the K4E provides additional information such as other pages from the same website, related images or disease definitions [18].

The automated system for HONcode detection is integrated into the K4E search engine, to illustrate one possible usage of such a system. For each webpage returned within the search engine results list, the system estimates the level of HONcode compliance. Based on the whole body of the pages trawled from the same website as the result page, the percentage of detected principles is displayed in the form of blue progress bars (fig. 4), hovering over them gives additional information concerning missing principles. This enables users to judge for themselves whether to access the webpage or not.

2 http://everyone.khresmoi.eu

Conclusion

The objective of this article was to show the evolution of implementation and usage of the HONcode. In the 20 years of its existence HON Foundation has promoted the quality of health-related information on the Web. One of the greatest features of the HONcode is its adaptability. Adapting its principal requirements to meet the needs of ever-growing collaborative platforms, and supporting the certification process by the automated system for reliable, faster and permanent evaluation of websites, HON has managed to retain its place as the most important third party certification initiative.

Since its inception, HON has given great weight to the importance of research and analysis with the mission to continuously improve the quality of online health information, while at the same time promoting and creating awareness about the importance of trustworthy online health information.

HON has been involved in multiple research projects to create measures in furthering this mission, one such project being the automatic detection within the Khresmoi and Kconnect projects. HON is extremely proud to be associated with the Khresmoi4Everyone project as it is indeed an excellent tool and will serve as an invaluable resource of trustworthy health information for the general public.

However, despite all this, the HONcode would never have survived if it were not for its users – the health information providers or webmasters, and the health information consumers. Many of the websites within our database go back over a decade and some to the very beginning of HON, a testament to the requirement for such a service.

Acknowledgements:

The automated HONcode detection research and development continues in the scope of European project Kconnect (2015–2018, project No. 644753).

References


Figure 4

K4E trustworthiness information display.
Correspondence:
Célia Boyer
Health On the Net Foundation
Chemin du Petit-Bel-Air 2
CH-1225 Chêne-Bourg
Celia.boyer[at]healthonnet.org
Figure 1
HONcode certification process.
Figure 2
Automated system for HONcode detection.

Figure 3
Terms that the expert and automated system detect for the “complementarity” principle.
Figure 4
K4E trustworthiness information display.