

# Requirements document for the general public health search

Natalia Pletneva<sup>1</sup>, Alejandro Vargas<sup>1</sup> et Célia Boyer<sup>1</sup>

<sup>1</sup>*Fondation Health On the Net, Genève, Suisse*

*Nataly.Pletneva@healthonnet.org*

## Abstract

This article presents the results of the survey conducted by the Health on the net (HON) Foundation in a framework of the EU project KHRESMOI (FP7 – ICT-2009.4.3). The aim of the survey was to identify the most important user requirements and difficulties in a process of online health information retrieval to be further used for use case elaboration. It was necessary to design a specific questionnaire based on HON logs, previous HON surveys and literature review. The questionnaire was developed in a period from December of 2010 to March 2011 in collaboration with the Society of physicians of Vienna. Further, it was translated from English into French, Spanish and German and put online.

385 answers were collected representing mostly the opinions of highly educated users from healthcare (not physicians) and IT areas of occupation. Overall, the representatives from 42 countries around the world filled in the questionnaire with top contributors from France and Spain.

Currently most of the respondents reported connecting to the Internet via Wi-Fi, and many of them are using mobile devices (laptops and smartphones mostly). The Internet was mentioned to be the second source of the health information after the physician, due to its accessibility and “easiness”. The most researched topics were about general health, chronic diseases and lifestyle. Illness-related information can be prioritized in a following way: treatment/therapy, detailed and general disease description, drugs and scientific articles. The most important characteristics of a search engine are relevance and trustworthiness of the results. This appears to be the main problems too, that current results retrieved using a general search engine do not satisfy these requirements. Additionally, advertisement banners, contradicting information found on the different web sites and difficulties related to very precise queries are among other barriers.

## **Introduction**

The percentage of people going online is constantly growing. According to the Internet World statistics, nearly 2 billion people, or 28.7% of the world's population, are using the Internet [1]. In 2010 national bodies reported that 52,5% of adults were looking for health content in the Internet in Spain [2] and 39% in the UK [3].

Online health surfing can be very beneficial for the novice users in terms of feelings of reassurance, confidence, relief [4], but due to overwhelming quantity and uneven quality of online health information it might be also dangerous and health threatening [5]. Although medical professionals are qualified to evaluate the information quality, they can also be overwhelmed with its quantity and annoyed with associated time consumption.

The quality of online health information is often questionable. The information provided may be often incomplete and in some cases even misleading [6] [7] [8]. The consequences on users' health are unpredictable: so far there are no studies that online health information causes harm. Nevertheless, with experience both users' scepticism and the demand for high quality information are growing. For example, in the USA, among those looking for health and medical information online, the number of people dissatisfied with their search results (from 6% to 9% in the last five years) or with the reliability of information (from 5% to 8% of those who have ever looked for online health information from 2005 to 2010) has been increasing [9]. Despite mostly positive outcomes reported by PEW, 25% of e-patients said they were overwhelmed by the amount of information found online, 22% felt frustrated, 18% confused and 10% frightened by the serious or graphic nature of the information found online [4].

As a consequence of health information search the Internet influences the doctor-patient relationship. Doctors remain the most significant and valuable source of information for patients. In France, in 2010, patients preferred asking doctors rather than the Internet (89% vs. 64%) [10]. On the other hand, an international study of 2008 revealed that 88% turn to their physicians to validate online information, but at the same time, the same number (88%) turn to other sources to validate information from their doctors [11].

Some studies [12] [4] [13] proved that health search starts most commonly from use of a general search engine as Google®, Yahoo®, etc. However, the trustworthiness of the search results is often questionable as it was discussed above. It was also proved that a search engine specially designed for health and medical information retrieval provides more reliable results [14].

## **Methodology**

This article presents the results of the patient questionnaire realized within the EU project KHRESMOI (FP7 – ICT-2009.4.3). KHRESMOI aims to allow members of the general public to obtain reliable and understandable medical information by developing automated and semi-automated approaches to classifying information. To develop the new search engine focused on biomedical information as KHRESMOI it is necessary to understand what is valued and desired by the general population looking for online health information as well as identify the usability barriers for further elaboration of a use case. Despite a variety of studies dedicated to the quality of online health information and influence of the Internet on doctor-patient relationships, there are no studies studying in depth the process of health information retrieval by the members of general public. Thus, the aim of this article is to collect such opinions in order to understand how they are searching for online health information via conduction of an Internet-based survey.

The questionnaire was developed by the HON team from December 2010 to March 2011. Initially the questionnaire had 17 questions and finally it has been release with 53 questions. The grounded theory approach has been used by analysing previous HON Foundation surveys, HON Foundation's logs and conducting the literature review. The development of the questionnaire was discussed with the project partners the Society of physicians of Vienna (GAW). The questionnaire was further translated into French, Spanish and German and uploaded on the HON Foundation web site<sup>1</sup>. The questionnaire was open from end of March – end of April.

### ***Structure of the questionnaire***

Part 1. Use of the internet (9 questions)

- User experience description
- User frequency health seeking

Part 2. Current Search of health information (8 questions)

- Current search of health information
- User preferences and requirement identification related to health information search
- What do users look for

Part 3. Use of search engine for health topics (14 questions)

- User preferences and requirement identification, related to use of a search engine
- User behavior while typing a query and analyzing search results

Part 4. Difficulties and barriers (3 questions)

- Technical and non-technical obstacles
- Advertisement influence

Part 5. Let's dream of citizen-centered health search engine (7 questions)

- Search results display
- Identification of helpful tools
- Funding

Part 6. About you (12 questions)

- Socio-demographic characteristics
- Availability of resources in a mother tongue

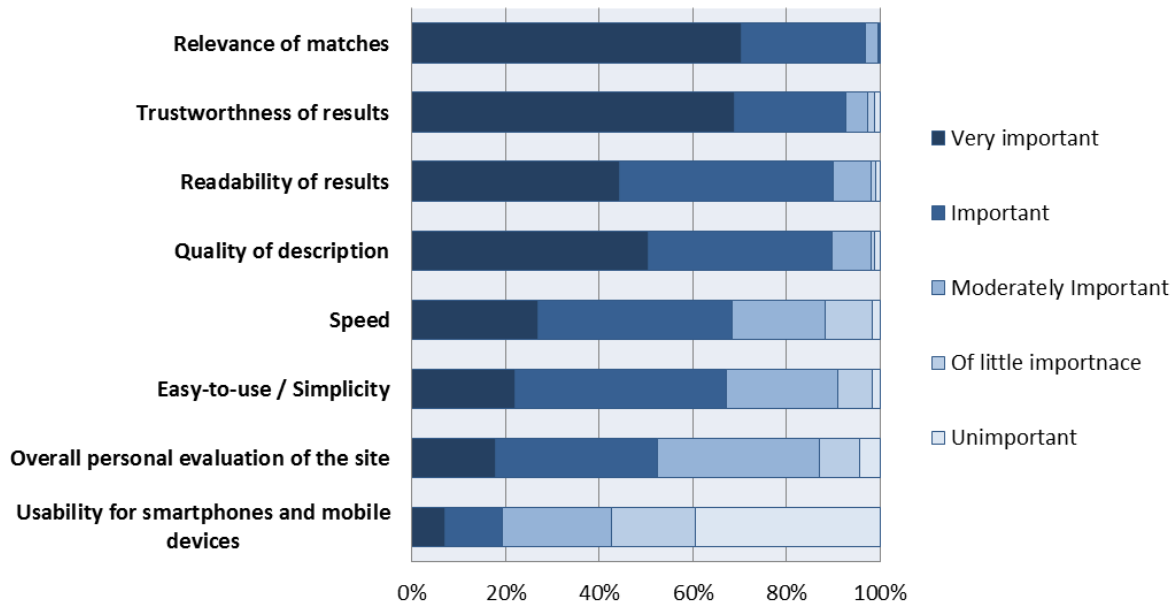
### **Results of the survey and discussion**

385 members of "general public" participated in the survey, slightly more females (53%) than males (43%), what corresponds with a general tendency that women are more active health information seeker, especially in using the Internet for these purposes [15] [16].

The most important characteristics of a search engine for the users are relevance of matches (97% consider it as important or very important), trustworthiness (93%) and readability (90%) of results and quality of description (90%) (see Figure 1). Easy-to-use or simplicity is still important (67%) despite the fact that most of the participants self-assess themselves as good or professional Internet users.

---

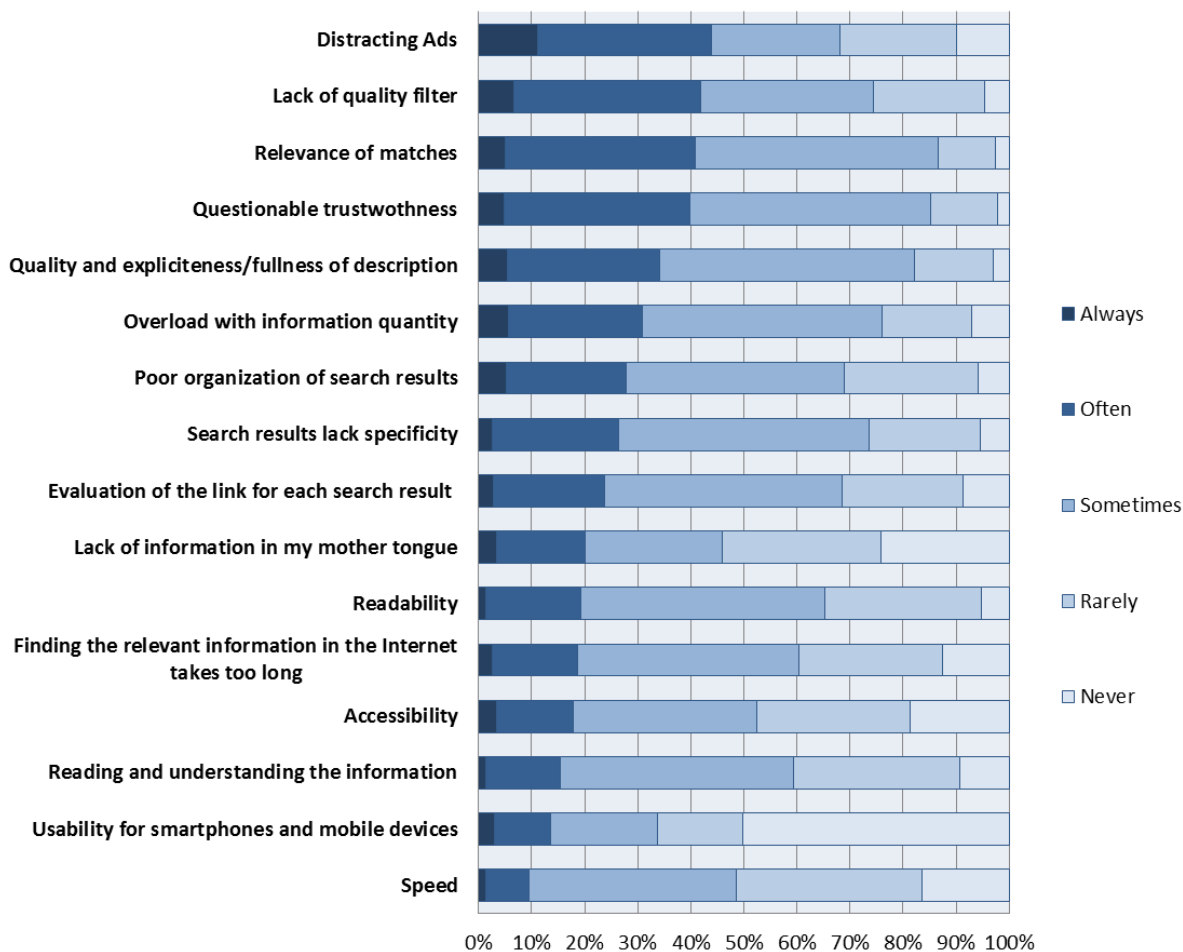
<sup>1</sup><http://www.healthonnet.org/kpat>



**Figure 1. What are the most important characteristics of search engine for you?**

Most of the respondents usually type two-three words (60%) in a search bar, 29% type up to five words. The terms being introduced into a search bar are medical terms as diagnosis or drug names in 82% of cases. 14% prefer to type a query in a question format in a same way they would have asked their physician. Some mentioned that they can type a physician's name. The most widely used advanced options are language (70%), date range (57%) and country (50%). Some respondents indicated that they include a web site name, restrict a domain and exclude certain terms, use a function of geolocation, search for video and images and use notifying alerts for subscribed tags.

It is important to understand the users' difficulties in searching and evaluating online health content. It appeared that distracting ads (44% say it is the case always or often), lack of quality filters (42%), relevance of matches (41%) and questionable trustworthiness (40%) are the most important ones (see Figure 2). Thus, not surprisingly users want to have relevant and trustworthy results. Technical issues such as accessibility in terms of customization of letter size and font colours (18%), usability for smartphones and other mobile devices (14%), speed (9%) are the least important.



**Figure 2. What types of difficulties do you face when searching for health information?**

Another problem that the respondents do not consider important is reading and understanding the information, or health information literacy (15%) (see Figure 2). Additional questions were asked regarding the way the respondents perceive the level of understanding of medical information. 88% state they often or always understand health/medical information found online. This result is difficult to be interpreted. Are the respondents answering positively because at the end they were able to answer their initial question, no matter how long it took or is it reflecting the fact that their health knowledge and capacity of adaptation are higher? High self-esteem can be explained by the high education level of the participants, thus they self-report being able to research medical and health information in the Internet. However, very little evidence exists justifying positive correlation between higher education level and higher health literacy [17]. Moreover, as for the general public and those who are online the education level stratification differs from the respondents of the survey, the question cannot be left aside. The problem of health literacy was also recognized by international organizations such as the WHO [18]. One of the approaches to ensure the access to “readable” and understandable information for the users was elaborated in the PIPS project (2004-2008)<sup>2</sup> when automatically estimating the readability level of medical websites was developed and tested. A user thus can choose a readability level of “easy”, “medium” and “difficult” depending on his/her knowledge of the particular domain [19].

## Conclusions

<sup>2</sup><http://www.pips.eu.org/>

The results presented in the report reflect the views of educated people, a vast majority of participants graduated from a university, many of them have Master and PhD. Apparently, this group of people cannot represent whole e-patients community, neither Internet community. Nevertheless, having highly educated respondents allows collecting different ideas which can be more elaborated and conscious than from the general public. It is important to contrast their opinions to the ones of people who have no such Internet experience, but still are using the Internet to obtain health information.

Summarizing the results of the survey we can identify the following milestones in creating a user-centred search engine.

First of all, most of the respondents are connected to the Internet via Wi-Fi, trend of using mobile devices will keep growing too. Thus, if by now the problem of web site accessibility from a mobile device is not among the top, it will evolve.

Secondly, the Internet is the second source of the information after the physician, and it is more likely to be used due to its accessibility and easiness. The most researched topics are about general health, chronic diseases and lifestyle. Illness-related information can be prioritized in a following way: treatment/therapy, detailed and general disease description, drugs and scientific articles.

Thirdly, majority of e-patients are using a general search engine to find the answers to their health queries. The most important characteristics are relevance and trustworthiness of the results. This appears to be the main problems too, that current results do not satisfy these requirements. Additionally, advertisement banners, contradicting information found on the different web sites and difficulties related to very precise queries among other barriers.

Fourthly, the ideal representation of the information will be categorization of the results into different groups with the possibility to do it automatically (the same way the doctor will say) or manually. Helpful tools which are highly appreciated by the respondents are advanced search, medical dictionary/thesaurus, suggested relevant topics, search of images and risk factor tools. However, taking into account low internet literacy of general population we will need to develop the tutorials to explain them how to use all these options. Most users will tolerate ads in order to keep the search engine free, but would like to be ensured that the ads have undergone a rigorous quality control. Some of the respondents proposed to have a possibility to create a profile and keep the search history for the follow up investigation of the topic. Some respondents have also mentioned they would like to have a free access to medical literature. Quite a lot of users want to see the search results appearing in social networking web sites, but it is important to identify clearly which source is original (for example, article), and which is a discussion or opinion about it.

## References

1. World Internet Usage Statistics News and World Population Stats. *Internet World Stats - Usage and Population Statistics*. [Online] 30 June 2010. [Cited: 30 03 2011.] <http://www.internetworldstats.com/stats.htm>.
2. **6. Instituto Nacional de Estadística.** *Encuesta sobre Equipamiento y Uso de Tecnologías de la Información y Comunicación en los hogares 2010*. 2010.
3. **7. Office for National Statistics.** Statistical bulletin: Internet Access 2010 . *UK national statistics*. [Online] 27 Aug 2010. [Cited: 10 Mar 2011.] <http://www.statistics.gov.uk/pdfdir/iahi0810.pdf>.
4. **11. S, Fox.** Online Health Search 2006. *Pew Internet & American Life Project*. [Online] 29 Oct 2006. [Cited: 20 Dec 2010.] [http://www.pewinternet.org/~media/Files/Reports/2006/PIP\\_Online\\_Health\\_2006.pdf](http://www.pewinternet.org/~media/Files/Reports/2006/PIP_Online_Health_2006.pdf).

5. 12. **R White, E Horvitz.** *Cyberchondria: Studies of the Escalation of Medical Concerns in Web Search.* s.l. : Microsoft Research, 2008.
6. 13. *The Quality of Internetsourced Information for Patients with Chronic Pain Is Poor* . **TB Corcoran, F. Haigh, A. Seabrook, SA Schug.** 2009, Clin J Pain, pp. 617-623 .
7. 14. *Is the Internet a reliable source for dietary recommendations for stone formers?* **Traver MA., Passman CM., LeRoy T., Passmore L. et al.** 4, s.l. : J Endourol., 2009, Vol. 23.
8. 15. *A readability assessment of online Parkinson's disease information.* **Fitzsimmons PR, Michael BD, Hulley JL, Scott GO.** 4, s.l. : J R Coll Physicians Edinb., 2010, Vol. 40.
9. 4. **H., Taylor.** HI-Harris-Poll-Cyberchondriacs. *Harris Interactive.* [Online] 04 August 2010. [Cited: 01 December 2010.] <http://www.harrisinteractive.com/vault/HI-Harris-Poll-Cyberchondriacs-2010-08-04.pdf>.
10. 16. *Vers une meilleure intégration d'Internet à la relation médecins-patients.* *Conseil National de l'Orde des Medcins.* [Online] 06 May 2010. [Cited: 23 November 2010.] <http://www.conseil-national.medecin.fr/article/vers-une-meilleure-integration-d%E2%80%99internet-la-relation-medecins-patients-982>.
11. 17. *Health Engagement Barometer.* *Edelman.* [Online] 2009. [Cited: 22 November 2010.] [http://static.edelman.com/www.edelman.com/healthengagement/docs/Edel\\_HealthBarometer\\_R13c.pdf](http://static.edelman.com/www.edelman.com/healthengagement/docs/Edel_HealthBarometer_R13c.pdf).
12. 18. *How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews .* **Eysenbach G, Köhler C.** 7337, s.l. : BMJ, 2002, Vol. 324.
13. 19. *The Internet as a diagnostic aid: the patients' perspective.* **Schembri G, Schober P.** s.l. : Int J STD AIDS, 2009, International Journal of STD & AIDS, Vol. 20, pp. 231-233.
14. 20. *Comparison of PubMed and Google Scholar Literature Searches.* **Anders ME, Evans DP.** 5, s.l. : Respir Care, 2010, Respiratory care, Vol. 55, pp. 578-583.
15. 3. **Fox, S.** *Health Topics.* *Pew Internet & American Life Project.* [Online] 1 February 2011. [Cited: 16 April 2011.] [http://www.pewinternet.org/~media/Files/Reports/2011/PIP\\_HealthTopics.pdf](http://www.pewinternet.org/~media/Files/Reports/2011/PIP_HealthTopics.pdf).
16. 21. *Health information seeking: a review of measures and methods.* **Anker AE, Reinhart AM, Feeley TH.** 3, March 2011, Patient Educ Couns, Vol. 82, pp. 346-54.
17. 28. **Canadian Council on Learning (CCL).** *A healthy understanding: What have we learned about health literacy in Canada?* Ottawa : s.n., 2008.
18. 29. **WHO.** *Regional Preparatory Meeting on Promoting Health Literacy .* Beijing : ECOSOC, 2009.
19. 30. *Lexically-based Distinction of Readability Levels of Health Documents .* **Borst A, Gaudinat A, Grabar N, Boyer C.** 2, s.l. : AIM, 2008, Vol. 16.