

Conference on ENTERprise Information Systems / International Conference on Project
MANagement / Conference on Health and Social Care Information Systems and Technologies,
CENTERIS / ProjMAN / HCist 2015 October 7-9, 2015

The influence of documents, users and tasks on the relevance and comprehension of health web documents

Melinda Oroszlányová^{a,*}, Cristina Ribeiro^{a,b}, Sérgio Nunes^{a,b}, Carla Teixeira Lopes^{a,b}

^aDEI, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

^bINESC TEC, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

Abstract

Search engines typically estimate relevance using features of the documents. We believe that several features from the user and task can also contribute to this process. In the health domain there are specific characteristics of web documents that can also add value to this estimation. In the present work, using a dataset composed by set of annotated web pages and their assessment by a set of users regarding their relevance and comprehension, we analyse what characteristics affect documents' relevance and what characteristics influence how well users comprehend them. We have conducted a bivariate analysis using characteristics of the above data collection. The strongest relations we have found are linked to the task features, suggesting a direct association between tasks' clarity and easiness and both the relevance and the comprehension of the content. The language of the document, its medical certification, the update status, the content in pathology definitions, the content in prevention, prognosis and treatment information, are other characteristics valued by consumers in terms of relevance. Users' previous experience on health searches and, particularly, on the topic being searched, their gender, the language and terminology of their queries were shown to be related to their success in the search tasks. We have also found that lay terminology, knowledge about the medico-scientific terms and the language of the documents are good indicators of comprehension. Documents containing links and testimonies, and the ones recently updated were observed to be better understood by users, as well as blog posts and comments.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of SciKA - Association for Promotion and Dissemination of Scientific Knowledge

Keywords: health information retrieval, web, relevance, comprehension;

* Corresponding author. Tel.: +351 22 508 14 00.

E-mail address: melinda@fe.up.pt

1. Introduction

According to the International Telecommunications Union's estimations, 40.4% (ca. 3 billion) of the population was using the Internet worldwide in 2014⁸. Due to the immense amount of information that has become available to users in the past decades, people give preference to the Internet as a source when seeking information^{14,9}. Consequently, the number of people being affected by such information has also increased, what gave rise to studies^{6,4,5,3} about people's usage and search habits on the Web, in the health domain. National surveys about how Internet users in the U.S.A. search for health information on the Web have reported that, on a daily basis, millions of American adults have been using online resources for their health concerns since 2002^{6,4,5}. In 2002, 80% of Internet users have looked online for health topics⁵, 61% of American health seekers said the Internet has improved the way they take care of their health⁵, and 68% of the users said that health searches had some impact on their decisions related to health care⁵. However, 73% of health seekers have turned away from a health website during a health search, due to the commercial interests of the site, the lack of information sources, absence of update, or other information quality indicators which were not displayed on website⁵. A similar survey from 2006⁴ reports that, on a typical day, 7% of American adults (> 18 years) search for health information. Users' access to the Internet and their interest in health related information were both found to be influencing aspects in health search surveys⁵. In 2010, among all adults in the U.S., 74% went online, and 59% looked online for health information⁵. These proportions were shown to be influenced by the health status of the users⁵, and depending on the discovered information, users' findings may as well have an impact on their decision making.

Our goal is to study how the features of users, tasks and some specific characteristics of web documents are related to the documents' level of relevance and how well users comprehend them, aiming that our findings can be used by search engines to provide more useful documents to the users. To do so, using a previously annotated dataset composed by a set of annotated web pages, relevance judgments and users, we conduct a bivariate analysis regarding these two variables.

2. Literature Review

At the end of the 90's, a guide to the Internet introduced by Pallen¹² became an important resource for the healthcare providers of that time. It urged them to share information on health topics with the public. Later on, researchers started to focus on the information about specific medical topics available on the Web. They saw that, for the users who search for health related information, it might be difficult to determine the reliability of web pages. Furthermore, online health seekers started to look for information not only for themselves, but also often for someone else^{6,4,5,3}. This way, health search has an impact not only on the users' own but also on others' health care routine^{6,4,5,3}. Every search engine has to estimate the relevance of the information accessed via web pages, referred to as the relevance of a document to a user¹³, and knowing how relevance depends on the user and document's characteristics can bring insightful orientations to search engine developers. Therefore, the study of relevance is of interest in health information retrieval (IR). Users evaluate web pages for information seeking and decide about the utility of the given web site for different types of information-seeking tasks, based on certain criteria. These features, which make web documents relevant for information seeking, include textual, structural and qualitative aspects, non-textual items and physical properties of web documents¹⁶. Key relevance criteria identified by health information users during web searches in a previous study² included research, topic, scope, data, influence, affiliation, web characteristics, and authority, indicating the complexity of web users' relevance judgements. Other studies of user-defined relevance criteria reported specificity, topicality, familiarity and variety as most frequently used in relevance judgements¹⁴. In information retrieval processes, the role of the user is central in relevance assessment, which is influenced by user, task, query and document characteristics (e.g. age, gender, health search experience, medical specialty, task clarity)¹⁷.

In IR studies, the concept of relevance is central, and it is also the main concern of any IR system¹¹. There are several types of user-based relevance, which depends on the context and on the user (Saracevic, 1996). Here we consider situational relevance (i.e., utility), expressed by the relevance of the documents to the users' tasks¹⁷, and motivational relevance, which refers to the relationship between the users' intents and goals, and the information

object¹³ (i.e., the web documents), and want to predict the relevance of a document to a user, using their characteristics.

Several studies have been conducted not only considering accessibility of relevant content, but also its comprehension, when evaluating health-related information available on the Web. They conclude that users need a high reading level in order to comprehend web-based health-related information¹. How well the users understand the health content, has been shown^{1,10,7} to be possibly related to their further decision making. Users may trust and prefer online information seeking when the content of web documents they find is understandable, useful and credible for them^{14,9}. Given this context, it is important to analyse what influences users judgments pertaining the relevance and comprehension of web documents. In this work, with the help of the available data collections¹⁷, we show that other than document features, user characteristics are also good descriptors and possible predictors of relevance and comprehension.

3. Description of the data set and statistical analysis

This study is based on an existing data set composed by an annotated sample of health web documents. This set of documents was initially collected for an user study¹⁷, and was later automatically and manually annotated. The data set is composed by the 30 Google top-ranked documents for 8 information situations using 4 different queries with different language and medical terminology (lay or medico-scientific). The documents were assessed by a researcher and 10% of them were also assessed by an external judge (health professional). The manual annotation was considered well defined by the judge. In the user study, 40 users assessed the relevance and comprehension of web documents. Data about the users, their judgements, tasks and the documents of the cited study are used in this work to explore the relationship between them. The main characteristics of the users include their gender, age, proficiency in English language, experience, frequency and success in web and health search, as well as their health status and literacy, the language and terminology they use. The characteristics related to the users' search tasks concern the queries (language and terminology), and the assessments of the tasks' clarity, complexity and familiarity, previous search, information idea about the tasks and technical terms used in them. They also included users' judgements of their task completion status, comprehension and of the relevance of the content. The main characteristics of the documents were grouped into four categories related to their content, web document, responsible entity and website.

In Section 4 and 5 we describe the relationship between pairs of variables, when one of them is relevance and comprehension. We use inferential statistics, depending on the nature of the variable (continuous, nominal or ordinal), in order to determine how a variable compares to another. We use a one-tailed Spearman's rank-order correlation for ordinal versus continuous and ordered variables, where the ρ statistic measures the strength of association between the two variables. The significant results suggest positive association between the variables with positive value of ρ statistic, and negative association, when ρ is less than zero. The threshold for the variables included in the further analysis was set to $\rho > |0.1|$. For ordinal and nominal variables, the Chi-squared test of independence is used. It determines whether there is a significant association between the variables, with the null hypothesis that the two tested variables are independent. The null hypotheses that there is no association in the underlying bivariate population are rejected when the probability associated with the test statistic (p -value), that is of observing a sample statistic as extreme as the test statistic, is less than the significance level ($\alpha = 0.05$). In this case we can conclude that there is a relationship between the two variables. We calculate the standardized residuals (SR) in order to show the dependence for the significant results ($\alpha = 0.05$) from the Chi-squared test. They indicate the importance of a cell (in the cross-table) to the ultimate Chi-square value. For visualizing bivariate relationships, we use mosaic plots with residual-based shading. They are composed of tiles with areas proportional to the observed frequencies in the corresponding cells of the contingency table. Positive values of SR are interpreted as cells having greater observed frequency than would be found under independence assumption, and negative values indicate cells which occur less often than under independence. The $SR > |2|$ are considered significant at $\alpha = 0.05$ (denoted with *), and $SR > |4|$ at $\alpha = 0.0001$ (denoted with **). With $SR < |2|$ we are unable to say whether they are significant.

4. Relevance bivariate analysis

In the following, the variables referring to the situational and motivational relevance of health related information in the documents are analyzed regarding their possible relationship with the characteristics of web documents and users, according to the statistical strategy described in Section 3.

Situational relevance is given by the relevance judgements of the users¹³, having 3 levels (0-non-relevant, 1-partially relevant, 2-totally relevant). Motivational relevance refers to the relationship between the user's goals and intents, and the information objects¹³, which are in this study the retrieved web documents of health information. It is expressed by the user's accomplishment, feeling of success and satisfaction, described by the variable task completion status in the present work, evaluated by the users on a scale with five levels (1-extremely unsuccessful; 5-extremely successful).

4.1. Situational Relevance

User

Female users tend to find the health-content more useful ($SR = 2.05^*$) than male users ($SR = 2.57^*$). We cannot judge how relevant is a document for users who have previously searched about the health topic, due to the standardized residuals being less than |2|. The habit of searching in English is weakly associated with documents' relevance ($p = 0.11^*$). Moreover, users who feel more successful in web search have a higher tendency to judge documents with higher relevance scores ($p = 0.13^*$).

Task

We observed that users completing the search tasks in their mother tongue (Portuguese, $SR = 4.80^{**}$) do not find the content relevant, but we cannot state that with English queries ($SR = 1.63^*$) they would find more relevant documents. Nor can we conclude whether the terminology they use in the query or their previous search experience about the topic would indicate or not relevant content, because the standardized residuals are less than |2|. Users who were not familiar with the technical terminology used in the search tasks ($SR = 2.20^*$), found the documents more relevant than the users who knew the technical terms ($SR = 3.46^*$). How well users comprehend the documents is weakly associated with documents' relevance ($p = 0.24^*$), and how clear is the search task is also weakly associated with documents' relevance ($p = 0.20^*$).

Content of the document

We found the combination of HON code categorization (i.e., medical certification of documents) of the documents for consumers and not relevant content rare ($SR = -2.41^*$) what suggests that HON categorization is a good indicator of pages valued by consumers. According to our analysis, Portuguese documents ($SR = -2.80^*$) are less relevant than English documents ($SR = 2.09^*$). Health consumers value testimonies ($SR = 2.82^*$), and they have some awareness of the quality of the information. That is, HON certified documents tend to be more relevant ($SR = 2.82^*$). Recent updates (less than 1 year ago) of the documents seem to be important for their relevance ($SR = 3.22^*$). Commercial intent ($SR = -2.40^*$), advertisements ($SR = -3.14^*$), videos ($SR = -2.05^*$), and pathology definition ($SR = 2.74^*$) are suggested to be good indicators of pages valued by consumers. We observed that health consumers, in general, do not value scientific publications ($SR = -3.89^*$). The standardized residuals for the indication of the place of treatment in the document and relevance (1.67) are too low to suggest any statistically significant relationship between these two characteristics. However, the indication of prevention ($SR = 2.70^*$), prognosis ($SR = 2.54^*$) and treatment ($SR = -2.58^*$) seem to be important to the relevance of a document, what might depend on the search task being done by the users. Using health medico-scientific terminology is weakly associated with documents' relevance ($p = 0.15^*$). We observed a weak negative association, with p statistic less than 0.1, between the rank of the document and relevance ($p = -0.24^*$).

Web document

Comments in blogs ($SR = 2.16^*$) seem to be valued by the users, but articles in journals ($SR = -3.41^*$) and academic work ($SR = -2.33^*$) do not seem to be valued by the users.

Entity responsible for the document

Authors' contacts displayed on the documents are shown to be important ($SR = 2.49^*$) for the users' relevance judgement. We cannot show a statistically significant relationship between documents with scientific character and their relevance, due to the low value of the corresponding standardized residuals (1.66) between these two characteristics.

Website

The domain 'uk' of web documents shows good scores ($SR = 2.27^*$), what is aligned with English documents having higher relevance scores than Portuguese documents ($SR = 2.14^*$). Websites of collaborative type, such as blogs ($SR = 2.03^*$) and Wikipedia ($SR = 3.27^*$), seem to be valued by the users, which is in line with the results of the comments in blogs. The indication of privacy policy of the website ($SR = -2.11^*$) seem to be important to the relevance of a document, but due to small values of standardized residuals we are not able to conclude statistically significant relationship between the team ($SR = 1.28^*$) and process of revision ($SR = 1.47^*$) and relevance.

4.2. Motivational Relevance

User

Female users tend to complete the search tasks more successfully ($SR = 4.21^{**}$) than male users ($SR = 4.88^{**}$). Users with previous health-search experience tend to be very successful in completing the search tasks ($SR = 4.88^*$), but we also observed that users without health-search experience are the most successful ($SR = 6.29^{**}$) in the completion of search tasks. Users' success with web search ($p = 0.26^*$), their proficiency in English language ($p = 0.16^*$), their health status ($p = 0.12^*$), and their health search in chats ($p = 0.11^*$) are weakly associated with motivational relevance. There is a weak negative association between users' health search in newsletters ($p = -0.19^*$), RSS ($p = -0.18^*$), forums ($p = -0.18^*$), and social networks ($p = -0.11^*$), respectively, and motivational relevance.

Task

We found that users feel more successful in completing search tasks in their mother tongue (Portuguese, $SR = 5.44^{**}$) than in English language ($SR = 2.82^*$). Users who previously searched about the topic tend to be moderately successful with the search tasks ($SR = 4.87^{**}$). We observed that users who knew the technical terminology felt very successful in completing the search tasks ($SR = 2.51^*$), but those who did not know technical terms tended to be the most successful ($SR = 3.27^*$). Users complete the search task easier with lay terminology ($SR = 4.22^{**}$) than with medico-scientific terms in the query ($SR = -4.35^{**}$). However, they still feel extremely unsuccessful with the completion. Among the task characteristics which satisfy the threshold criteria ($p > |0.1|$), the level of clarity of the search task ($p = 0.65^*$), its easiness ($p = 0.64^*$), and whether the users have any idea about the information ($p = 0.58^*$), show the strongest positive association with motivational relevance. The task familiarity ($p = 0.25^*$) is weakly associated with motivational relevance, as well as the correct task answer ($p = 0.17^*$), situational relevance ($p = 0.17^*$), and comprehension ($p = 0.17^*$). The strongest relationship among the task characteristics (i.e., task clarity) and motivational relevance ($p = 0.65^*$) is displayed on Fig. 1.

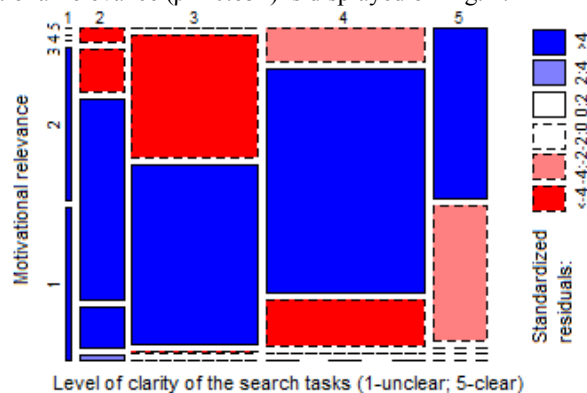


Fig. 1. Relationship between the task's clarity and motivational relevance (where 1 - extremely unsuccessful; 5 - extremely successful).

5. Comprehension bivariate analysis

In this section we analyze the users' understanding of the documents (or comprehension) regarding its possible relationship with the characteristics of web documents and users, according to the statistical strategy described in Section 3. The variable comprehension has 3 assessment levels (0-did not understand, 1-partially understood, 2-understood).

5.1. User

Female users tend to completely understand the documents' content ($SR = 2.42^*$), while male users claim that they do not understand ($SR = 6.46^{**}$) the content. Among the variables that satisfy the threshold criteria, the success of the users with web search ($p = 0.21^*$) and their health literacy ($p = 0.21^*$) are the most strongly associated with comprehension. However, this value is still very low, indicating a weak association. The users' proficiency in English language ($p = 0.19^*$), and health search in English language ($p = 0.16^*$) are also weakly associated with comprehension, as well as the frequency of web search ($p = 0.16^*$), and the level of satisfaction of the users' health information need on chat ($p = 0.14^*$).

5.2. Task

We found that users conducting search tasks with lay terminology ($SR = 3.15^*$) tend to completely understand the documents, and those, who use medico-scientific terms in the query ($SR = 3.47^*$), tend to understand the content only partially. Users who knew the technical terminology totally understood the content ($SR = 2.40^*$), but those, who did not know technical terms, showed a tendency to not being able to understand the document ($SR = 3.42^*$). We observed that users who do the search tasks in their mother tongue (Portuguese, $SR = 2.20^*$), totally understand the content of the retrieved web documents, and those, who search in English language ($SR = 3.08^*$), tend to understand the content only partially. We cannot judge the level of comprehension of the users according to their previous search about the topics of the tasks, due to the standardized residuals being less than $|2|$.

Among the variables that satisfy the threshold criteria, the level of easiness of the search tasks and the relevance judgements show the strongest association with comprehension with $p = 0.24^*$. The level of clarity ($p = 0.18^*$) and the completion status of the search tasks ($p = 0.17^*$) are weakly associated with comprehension, as well as the users' idea about the information in the search tasks ($p = 0.11^*$). The relationship between task easiness and comprehension with the highest positive value of 0.24^* is displayed on Fig. 2.

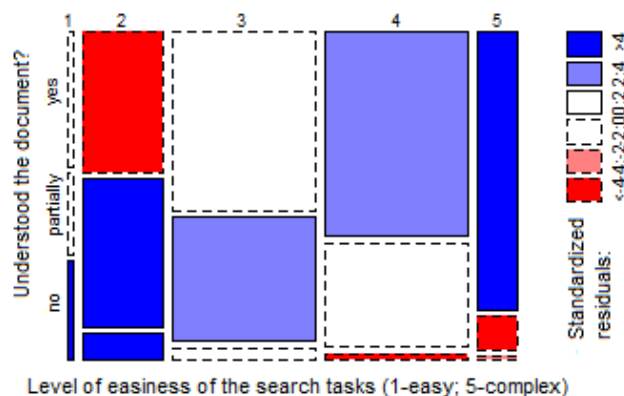


Fig. 2. Relationship between the easiness of the tasks and users' understanding of the documents' content.

5.3. Content of the document

We found that documents with HON code categorization for professionals were only partially understood by the users (SR = 3.93*). Upon the HON code categorization of documents for consumers, we cannot make judgements of the level of comprehension by the users, because the standardized residuals were less than |2|. Users totally understood the documents which contained links (SR = 2.93*). We observed that health consumers, in general, did not understand scientific publications (SR = -7.54**). According to our analysis, documents in the mother tongue of the users (Portuguese, SR = 4.71**) were more understandable for them than English documents (SR = 3.94*). Documents with the most recent updates (less than 1 year ago) seem to be understood the best (SR = 2.29*) by the users. We found that users tend to understand the documents with epidemiologic data only partially (SR = 4.11**), as well as the documents with the indication of clinical cases (SR = 3.23*). Users tend to completely understand documents containing testimonies (SR = 2.09*), and HON code certified documents (SR = 2.06*). They do not understand the content without images (SR = 2.93*), and partially understand the content with video (SR = 2.87*). We observed that documents without the indication of pathology definition (SR = 3.69*), diagnosis (SR = 2.66*), and treatment (SR = 2.62*), were not understood by the users. Users tend to understand less the documents without advertisements (SR = 2.80*), and it is rare that they do not understand the documents with commercial intent (SR = -3.11*). Upon the indication of prognosis and prevention, we cannot make judgements of the users' level of comprehension of the documents, due to the standardized residuals being less than |2|.

The usage of specific medical vocabulary in the content ($p = 0.16^*$) is weakly associated with comprehension, and the rank of the document ($p = -0.11^*$) exhibits a weak negative association with comprehension.

5.4. Web document

Comments in blogs (SR = 2.54*) seem to be totally understood by the users, but articles in journals (SR = 5.83**) only partially, and academic work (SR = 4.26**) is not understood.

5.5. Entity responsible for the document

We found that documents with scientific character are partially understood by the users (SR = 3.42*). When the name of the responsible entity was not indicated, users did not understand the content (SR = 2.32*). The standardized residuals for the indication of the name of the author were less than |2|, which does not let us make judgements of the users' level of understanding the documents.

5.6. Website

We found it rare that the documents from websites with domain 'br' (SR = -2.32*) and 'es' (SR = -2.33*) were totally understood by the users. Websites of collaborative type, such as blogs (SR = 2.51*) seem to be completely understood by the users. Content from websites that allow document sharing (SR = 3.30*) tends to be partially understood, as well as documents from digital journals (SR = 2.67*), digital magazines (SR = 2.24*) and digital libraries (SR = 3.18*) related to institutes of scientific profile. Documents from websites which did not indicate their objective were found to be partially understood very rarely (SR = -2.80*). The standardized residuals for the indication of team of revision were less than |2|, what does not let us make judgements of the users' level of understanding. It is rare that the documents with the indication of the process of revision (SR = -2.6*) are not understood by the users. The indication of the privacy policies on the website seems to be important for documents comprehension (SR = 2.19*).

6. Conclusion and Future Work

In this research, we analyzed the influence of web document features, user and tasks characteristics, on the relevance, users feeling of success, and understanding of the documents.

The statistically significant relationships discovered in the bivariate analysis, suggest association between the specificity of the vocabulary used in the documents, the position of the document in the ranking and situational relevance. As expected¹⁷, relevance decreases with the rank position of the document, which can be explained by the definition of ranking which is an automatic ordering of the documents by relevance¹³. We found that HON certification, documents classified by HON as being “for consumers”, the language of the document, and recent updates are good indicators of relevance. Commercial intent, advertisements, and pathology definition, as well as the indication of prevention, prognosis, and treatment were also suggested to be good indicators of pages valued by the consumers. Blogs and comments in them, Wikipedia, the domain ‘uk’, authors' contacts, and privacy policy of the website were shown to be important for the users' relevance judgements, too.

Our findings also suggest significant positive association between the clarity and easiness of the search tasks, and users' familiarity with them, their information idea and success in web search and motivational relevance. Previous searches on the health domain and, particularly, on the search task topic, the gender of the users and the language of their query were shown to affect their success, as well as the terminology used in the search.

We observed significant positive association between the users' success in web search, their health literacy, the easiness of the search tasks, and the relevance of the information in the documents and comprehension. Female users showed higher tendency for better understanding of the content. Users understand documents better when they use lay terminology. However, users familiar with the medico-scientific terms, and those who searched in their mother tongue, also commonly assessed the documents as totally understood. Documents containing links and testimonies, written in the users' mother tongue, and with recent updates, were found to be better understood by the users, as well as blogs' posts and comments. Although we reported significant associations, all the correlation values are low. As future work we plan to conduct a multivariate analysis and build two prediction models, one for relevance and another for comprehension, and find an answer to whether we can infer the relevance and comprehension of a document using its characteristics and the user's.

Acknowledgements

This work was partially supported by Project SIBILA - Towards Smart Interacting Blocks that Improve Learned Advice, reference NORTE--07--0124--FEDER--000059, funded by the North Portugal Regional Operational Programme (ON.2--O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF), and by national funds, through the Portuguese funding agency, Fundação para a Ciência e a Tecnologia (FCT).

References

1. Berland, G. K. (2001). Health Information on the Internet: Accessibility, Quality, and Readability in English and Spanish. *JAMA*, 285(20), 2612-2621.
2. Crystal, A. G. (2006). Relevance Criteria Identified by Health Information Users During Web Searches. *JASIST*, 57(10), 1368-1382.
3. Espanha, R. V. (2008). *Health and the Internet: Autonomy of the User*, LINI WORKING PAPERS N°6.
4. Fox, S. (2006). *Online health search*, Pew Research Center's Internet & American Life Project.
5. Fox, S. (2011). *The Social Life of Health Information*, Pew Research Center's Internet & American Life Project.
6. Fox, S. R. (2002). *Vital decisions: A Pew Internet Health Report*, Pew Research Center's Internet & American Life Project.
7. Hersh, W. (2008). *Information Retrieval: A Health and Biomedical Perspective (Health Informatics)* (3rd ed.). Springer.
8. International Telecommunications Union. (2014). *World Telecommunication/ICT Indicators database 2014 (18th Edition)*.
9. Kim, J. (2009). Describing and Predicting Information-Seeking Behavior on the Web. *JASIST*, 60(4), 679-693.
10. Lalmas, M. M. (2006). Readability Applied to Information Retrieval. *Advances in Information Retrieval*. Springer.
11. Manning, Ch. D., Raghavan, P., Schütze, H. (2009). *Introduction to Information Retrieval*. Cambridge University Press.
12. Pallen, M. (1995). Guide to the Internet. The world wide web. *BMJ*, 311, 1552-6.
13. Saracevic, T. (1996). Relevance reconsidered. *Proceedings of the Second Conference on Conceptions of Library and Information Science (CoLIS 2)*, (pp. 201-218).
14. Savolainen, R. (2008). Source preferences in the context of seeking problem-specific information. *Information Processing and Management*, 44, 274-293.
15. Savolainen, R. K. (2006). User-defined relevance criteria in web searching. *JDOC*, 62(6), 685-707.
16. Tombros, A. R. (2005). How Users Assess Web Pages for Information Seeking. *JASIST*, 56(4), 327-344.
17. Lopes, C. T., Ribeiro, C. (2010). Context effect on query formulation and subjective relevance in health searches. *IIIX '10 Proceedings of the third symposium on Information interaction in context*, pp. 205-214.