Patient Empowerment through Personal Medical Recommendations

Haridimos Kondylakis, Lefteris Koumakis, Eleni Kazantzaki, Maria Chatzimina, Maria Psaraki, Kostas Marias, Manolis Tsiknakis

Computational Medicine Laboratory, FORTH-ICS, Heraklion, Crete, Greece

Abstract

Patients today have ample opportunities to inform themselves in the internet about their disease and possible treatments. While this type of patient empowerment is widely regarded as having a positive influence on the treatment, there exists the problem that the quality of information that can be found on online is very diverse. This paper presents a platform which empowers patients in two ways: First it allows searching in a high quality document repository, and secondly it automatically provides intelligent and personalized recommendations, according to the individual preferences and medical conditions. To the best of our knowledge our platform is the only one which combines a search engine with automatic recommendations exploiting indivudual patient profiles and preferences.

Keywords:

Medical Information; Recommendations; Information Retrieval

Introduction

Medicine is undergoing a revolution that is transforming the nature of healthcare from reactive to preventive. The changes are catalysed by a new systems approach to disease which focuses on integrated diagnosis, treatment and prevention of disease in individuals. This will replace our current mode of medicine over the coming years with a personalized predictive treatment. While the goal is clear, the path is fraught with challenges. One of these challenges is the problem of the quality and the amount of information that can be found on online [1] since health information is one of the most frequently searched topics on the internet.

During the last decade, the number of users who look for health and medical information has dramatically increased. Already from 2002 a percentage of 80% of all adults online in the United States were estimated to have looked for health information whereas the 23% of the Europeans were using the internet to be informed about their health needs [2]. However despite the increase in those numbers and the vast amount of information currently available online, it is very hard for a patient to accurately judge the relevance of some information to his own case.

This paper focuses on current research activities related to the implementation of a Personal Medical Information Recommender (PMIR) within the EURECA EU[1] research project. The project aims to build an advanced, standards-based and scalable semantic integration environment enabling seamless, secure and consistent bi-directional linking of clinical research and clinical care systems which, among others, will empower patients to extract the relevant data out of the overwhelmingly large amounts of heterogeneous data and treatment information.

PMIR [3] is targeted at improving the opportunities that patients have to inform themselves in the internet about their disease and possible treatments, and providing to them personalized information and recommendations. Its goal is threefold: (1) to deliver relevant information to patients, based on their current profile as represented in their personal healthcare record (PHR) data, (2) to ensure the quality of the presented information by giving medical experts the chance to control the information that is given, and (3) to facilitate an easy uptake of the new system by minimizing the necessary manual effort.

The PMIR is integrated into a PHR as a set of individual apps. The patient is able to select the *PMIR search app* to look for useful information and a medical expert can select the *Semantic Annotator app* to register high quality web documents. Moreover, as the patient logs in to his PHR account, automatically, appropriate useful documents are recommend to him by the *Automatic Recommendation app*. Both the search engine and the automatic recommendation mechanism exploit the individual patient profile and patient preferences to provide personalized information. To the best of our knowledge PMIR is the only medical information search engine and recommendation system that exploits the individual profile of the patient to provide personalized empowerment.

The rest of this paper is structured as follows: Section 2 presents the PMIR architecture, describing the available components. Then, Section 3 describes similar approaches and Section 4 demonstrates the design of our evaluation. Finally, Section 5 concludes the paper and discusses future directions.

Architecture

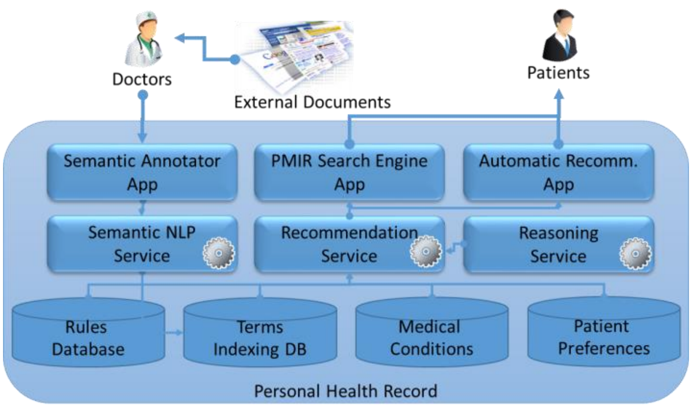

Figure 1 - PMIR Architecture

[1] http://eurecaproject.eu/

The architecture of the system is shown in Figure 1 and consists of three layers: The database layer including four different databases, the service layer including 3 services and the front-end layer including three individual apps. All modules are implemented within a PHR system through which the whole functionality of the PMIR is offered. Bellow we analyze in detail each one of the aforementioned modules.

Personal Health Record

Personal health record systems are a rapidly growing area of health information technology which provides the ability for an individual to own and manage a complete and secure digital copy of his health and wellness information. An electronic personal health record helps patients to gather and manage their medical information in one accessible and secure location and to take a more active role of their own health management.

Within EURECA, the adopted PHR system used is IndivoX [4], a free and open-source health platform which integrates health information across sites of care and over time. IndivoX was selected by performing an extensive evaluation on the existing PHR systems currently available [5]. IndivoX contains user's personal information like demographics and allows storing and editing different types of information like medications, laboratory results, problems, procedures and allergies. All information is stored in the *Medical Conditions Database* shown in Figure 1. Data entries are annotated using three different ontologies to enable interoperability, namely the LOINC[2], the SNOMED-CT[3] and the RxTerms[4]. LOINC is a universal code system for measurements, tests and observations and is used for annotating laboratory information, whereas for problems, procedures and allergies the SNOMED-CT ontology is used. The SNOMED-CT provides terms, synonyms and definitions used extensively in clinical documentation. Finally for information about medications the RxTerms ontology is used, derived from RxNorm for capturing medication history recording or prescription writing.

IndivoX, has been extended within EURECA to include the components of the PMIR and to enable patient empowerment using novel techniques such as *adaptive presentation*, *adaptive navigation* and *intelligent alerts* [6-7]. Adaptive presentation deals with the content presentation in a manner that best suits individual users' needs, adaptive navigation comprises with the ways to alter visible links to support hyperspace navigation and intelligent alerts notify the patient about possible drug interactions etc. In this paper we will focus only to the components of the PMIR.

The Semantic Annotator App & Service

The first step in providing useful information to patients is the identification of useful, reliable, high-quality online health information and its appropriate and efficient use. To cope with the unprecedented volume of healthcare information available on the net, PMIR uses domain experts (doctors etc.) in order to identify and register appropriate web documents. Currently a possible cooperation with HON[5] is under consideration in order to register web documents already identified from similar multi-stakeholder consensus approaches.

Using the *Semantic Annotator app*, an expert is able to register external documents that contain useful information to be further elaborated. Those web documents are high quality web resources (web pages, pdfs, docs etc.) selected carefully by the appropriate experts targeting patients. The interface of the app is shown in Figure 2.

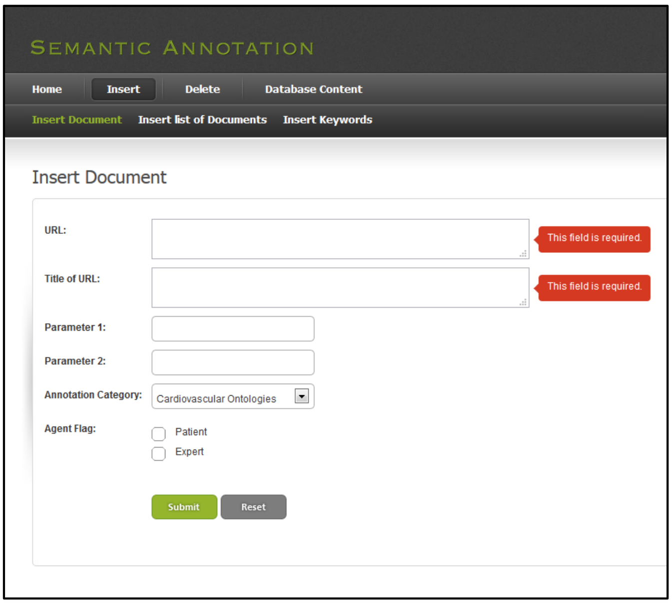

Figure 2 - The interface of the semantic annotator

In addition, to effectively exploit those textual resources simple string matching with patient search terms is not enough. To augment the quality of the search, the unstructured text documents go through a semantic lifting and indexing phase. During this phase the *Semantic Annotator Service* is called periodically, once each week, to annotate the newly inserted web documents with terms from the SNOMED-CT, the LOINC and the RXTerms ontologies. The whole process is asynchronous since the time required to annotate each document is significant. Each document is parsed using Natural Language Processing (NLP) algorithms for tokenization, lemmatization and part of speech [8]. Then, the Semantic NLP service extracts ontology terms, concepts and semantic types of the aforementioned ontologies and all terms and documents are stored in the *Terms Indexing DB.* This database is later used to match documents with patient profiles and patient queries.

The PMIR Search Engine App

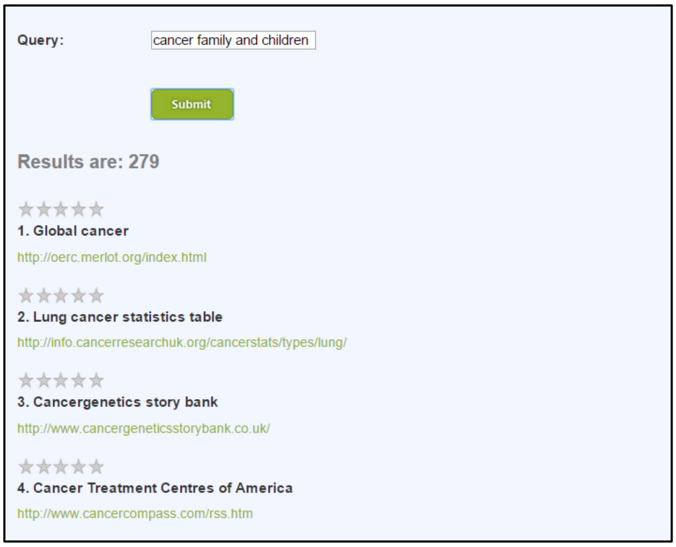

Figure 3 - The interface of the PMIR Search Engine App

Using the PHR system, the patient, besides logging and reviewing his medical information, is able to search for relevant, high quality information using an intelligent search

[2] http://loinc.org/
[3] http://ihtsdo.org/snomed-ct/

[4]http://wwwcf.nlm.nih.gov/umlslicense/rxter-mApp/rxTerm.cfm
[5] https://www.healthonnet.org/HONcode/Conduct.html

engine provided as an individual PHR app. The interface of the search engine is shown in Figure 3. The Search engine app sends the user query to the recommendation service and then the app visualizes the returned results.

The patient besides searching for relevant results is able to rate a result according to his opinion. The clicks and the rankings of each user are stored in the *Patient Preferences* database.

Automatic Recommendations App

Besides allowing the patients to search for useful information, interesting web documents are also automatically recommended to the patients. The information is presented to them at the first page they see when they log-in to the PHR. These documents are also returned by another execution of the recommendation service. The interface of the app is shown in Figure 4.

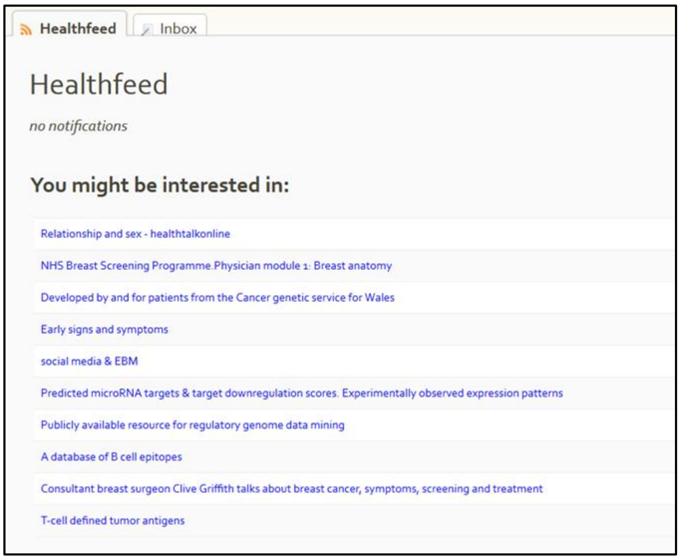

Figure 4 – Automatically recommending useful documents to patients

The Reasoning Service

The recommendation service takes advantage of the *Rules Database*. The development of these specific rules aims to allow generalizing the queries that the patient enters though reasoning. For example, assume that the terms "*corboplatin*" and "*cancer"* exist in a user query. We can easily identify through reasoning that the first term belongs to the *Drug* category and the second term belongs to the *Disease* category. Then we could search for the more general query "*Drug for a Disease"* in addition to the user input. Those rules are stored in the *Rules DB* and are exploited each time a user issues a query in order to increase the recall of the user search.

Table 1: The basic semantic types.

Disease	Drug	Medical Procedure	Tissue
Biomedical	Cell	Organism Function	Finding
Body Part	Gene	Clinical Attribute	Patient
Diagnosis	Age	Molecular Sequence	Device
Symptom	Virus	Injury or Poisoning	Vitamin
Laboratory	Food	Temporal Concept	

To create these rules we selected terms from several clinical, medical, biomedical and pharmaceutical ontologies. Every term in the ontology is mapped to a semantic category as an abstraction level of knowledge. The selected semantic categories for the PMIR search engine are shown in Table 1. Based on the combination of these primitive semantic categories the experts can create quite easily generalization rules. Some of those rules are shown in Table 2 and have a special meaning in the clinical domain; for example, the rule "*Drug for Disease*" is generated to describe a treatment for a specific disease.

Table 2: An example list of rules generated by the combination of prime rules

Drug for Disease
Patient took Drug for Disease
Drug for Disease in Body Part
Drug for Symptom

Such a reasoning scheme provides to the search engine additional knowledge for more effective search.

The Recommedation Service

The results presented to the patient, by the PMIR Search Engine and by the Automatic Recommendations apps, are provided by the *recommendation service.* The service considers the following databases to make the results of the query as personalized as possible:

a. *Patient Preferences Database*: This database contains user preferences that are acquired as the patient browses the results presented to him. His selections are logged and he is also able to rank by himself a result as relevant or irrelevant to augment similar future searches.
b. *Medical Conditions Database*: The medical conditions of the patient as they have been logged by the patient himself. The entries in the database are annotated using SNOMED-CT, LOINC and RXTerms terms while the patients does data entry.
c. *Terms Indexing DB*: The database with the annotations of the external documents that the medical doctors entered for indexing.
d. *Rules Database*: This database includes rules that are used for generalizing patient queries as described in a previous section.

The recommendation service is the bridge between the patient's information, preferences, the rules and the indexed documents. The content that is identified as relevant is scored according to its dynamic relevance to the patient. In addition to the objective relevance, the system takes into account subjective information such as the content which the patient has already seen, or the type of content the patient prefers (articles for the general public or specialist information, text or image, etc).

The algorithm used for semantic search is an extension of the vector space model [9] extensively used in information retrieval. Briefly, according to that model, both documents d and queries q are represented as vectors

$$d_i = (w_{1,d}, w_{2,d}, \ldots, w_{n,d})$$

$$q = (w_{1,q}, w_{2,q}, \ldots, w_{t,q})$$

Each $w_{j,i}$ is a weight for term j in document i, and reflects the importance of that term. Those weights are computed on the basis of the frequency of the terms in the document, the query or the collection. At retrieval time, the documents are ranked by the cosine of the angle between the document vectors and the query vector.

The aforementioned algorithm was implemented and extend in our case in the following ways:

a) The vector of the documents is computed based on the following metric $w_{t,d}=idf_t * tf_{t,d}$ where $tf_{t,d}$ is the frequency of the term in the specific document d and idf_t is the inverse document frequency of the term t in all documents in the collection [9]. All those weights are pre-computed and stored in the indexing/terms database to avoid the costly recalculation at runtime.

b) The vector of each query is complemented with a vector from the user profiles. So in our case

$$q = (w_{1,q}, w_{2,q}, \ldots, w_{t,q}, w_{1,p}, w_{2,p}, \ldots, w_{m,p})$$

where $w_{i,q}$ represents a weight for term i of query q and $w_{i,p}$ represents a weight for term j of the profile p. The $w_{i,q}$ is computed similarly to the weights of the document terms. In addition the $w_{i,p}$ is computed on the basis of the frequency of the terms in all user profiles reduced by a factor to give priority to the weights of the query terms.

c) In the case of automatic recommendations, the results are based only on the weights coming out of the profile of the patient. In this case the weights of the profile terms are not reduced to give priority to user query since now we don't have user query terms.

d) In order to take into account the feedback/preferences from the users (ratings and clicks) the ranking of the results is altered based on the ratings and the clicks of the users. More specifically, each individual user can vote a page up or down (in a five-point rating scale) to demonstrate their satisfaction or dislike. All user ratings, as well as user clicks, are correlated with specific weights that contribute to the recalculation of the initial weight of the documents based on the specific user profile.

The extended PHR supplies feedback about the displayed information from the patient, either implicitly or explicitly, such that the relevance of the displayed information can be adopted over time. The process of adapting content to specific user needs can be thought of as two main sub processes. The first sub process involves understanding what content can be most relevant to the current user's interests, and how this content should be organized. The second sub process involves decision on how to effectively present the selected content to the user.

Related Work

There are already several approaches trying to provide to the patients search engines with high quality medical information such as WebMD[6] , MayoClinic Patient Care[7], Medicine Plus[8] etc. However, these engines provide a rather limited set of information and they are not dynamically adapted according to patient's preferences or medical history, which is true in our case.

One of the most well-known search engine is HONSearch [10] which aims to improve the quality of the information intended to both patients and medical professionals by facilitating quick access to the most relevant and up-to-date medical discoveries. Each indexed web document has to fulfill specific criteria in order to be included in the search engine and uses a multi-stakeholder approach to include the relevant web documents. Although the approach seems to be promising there has been also some concerns on the quality [11] on the indexed web sites.

To this direction, but targeting medical experts AskHermes [12] is an online question answering system, trying to answer specialized clinical questions.

On the other hand, different modules already exploit the profiles stored in the PHR systems to automatically present useful information to the patients. For example, the ADE (adverse drug effects) [13] is an alerting system to inform patients about potential risks and the adverse effects of the medications they receive. Interesting approaches to our direction are STEPPS [14] and the discontinued MyDailyApple [13]. STEPPS tries to personalize the automatic retrieval of health information using profiling information from an electronic patient record whereas MyDailyApple provided simple personalized recommendations based on patients Google Health Profiles. However on both cases the lack of semantics leads to poor results and the lack of a search engine to allow patients interaction limits the patient's options.

To the best of our knowledge PMIR is the only system exploiting patient's profiles to provide both automatic and non-automatic high quality information to patients employing semantics, reasoning and exploiting also user preferences.

Evaluation Design

In this section we will present the design of our initial evaluation in order to test the PMIR system.

To evaluate our system we already registered 500 web documents offering high quality information to patients in the cancer domain. Those documents were selected by contact persons in the cancer domain out of the Information Office at Cancer Charity Tenovus[9], the European Association of Cancer Research[10], the eCancer[11] and the Institute of Cancer policy[12].

The system will be installed locally and tested at the Oxford NHS Foundation Trust[13]. The ethical board of the Trust already accepted to conduct the aforementioned evaluation. In the proposed evaluation, 15 patients will test the system after signing the necessary consent form and they will fill in the PHR system initially with their personal information, medications etc.

Those patients will be separated randomly in 3 groups. The first group will be used as a control group and the results directly out of the Google search will be presented to them. At the second group the recommendation service will be used but without exploiting the knowledge of the patient preferences and profile. Finally for the third group the all available information available will be exploited for each patient. Each patient will answer a usability questionnaire and the results will be compared between each other to identify the usability and the usefulness of our search engine. Although the number of the patients is limited we expect to give us some directions and hints for the future development of our engine.

Conclusion & Discussion

The changing nature of information distribution due to the evolution of the web has important implications for health care. Given the wide use of the web in providing medical information, feeding patients with appropriate content might further enhance the patient's education and experience. The

[6] http://www.webmd.com/

[7] http://www.mayoclinic.org/patient-care-and-health-information

[8] http://www.nlm.nih.gov/medlineplus/

[9] http://www.canceractive.com/cancer-active-page-link.aspx?n=2772

[10] http://www.eacr.org/

[11] http://ecancer.org/

[12] http://www.kcl.ac.uk/lsm/icc/research/centres/icp.aspx

[13] http://www.oxfordhealth.nhs.uk/

validity and the quality of the available healthcare information on the internet is an area of major concern mainly because these have not been well documented [15]. Although healthcare professionals should continue to strive to be the main source of information for patients, we should also be aware that most will continue to use the internet to gather information [16]. A recent publication from Berg et al [17] concluded that the optimal solution for patients is to be guided by healthcare providers to more optimal resources over the web. Delivering accurate sources to patient increases his knowledge and changes the way of thinking which is usually referred as patient empowerment. As a result, the patient’s dependency for information from the doctor is reduced. Moreover, patients feel autonomous and more confident about the management of their disease [18].

PMIR platform focuses on making the available information timelier and more relevant with respect to dynamic influences in the individual patient's treatment. The idea is that even if two patients suffer from the same disease and they are in the same phase of the treatment, their interests on available information may differ based on various factors. For example, patients might have different medical background that is not directly related to the treatment, they might receive additional drugs due to other, independent treatments, or they might be affected by other external factors such as the weather in case of allergies. More subjective factors include for example patient preferences for more simple or more complex information. To the best of our knowledge no other system, providing medical information, is able to be dynamically adapted in such a diverse environment.

For future work we consider to include machine learning approaches for recommending to patients automatically useful information and to extend the reasoning engine to more complex rules. In addition we would like to be able to present a semantically enhanced summary of the indexed document to the patient to augment his understanding. Collaborative filtering is another interesting idea that will be also investigated. In addition we expect that the evaluation of the whole platform will lead us to useful observations that might require the modification of our initial algorithms. Then a second more thorough evaluation will follow. It becomes obvious that delivering accurate, personalized and high quality information to patients is an important topic and several challenging issues remain to be investigated in near future.

Acknowledgements

This work has been supported by the EU projects p-Medicine, EURECA and MyHealthAvatar and has been partially funded by the European Commission under contracts FP7-270089, FP7-288048 and FP7-600929.

References

[1] Berg GM, Hervey AM, Atterbury D, Cook R, Mosley M, Grundmeyer R, Acuna D. Evaluating the quality of online information about concussions. JAAPA. 2014;27(2):1-8.

[2] McMullan M. Patients using the Internet to obtain health information: how this affects the patient–health professional relationship. Patient education and counseling, 2006;63(1): 24-28.

[3] Kondylakis H, Koumakis L, Rüping S, Kazantzaki E, Marias K, Tsiknakis M. PMIR: A Personal Medical Information Recommender, Medical Informatics Europe (MIE), 2014. pp. 1193.

[4] Adida B, Sanyal A, Zabak S, Kohane IS, Mandl KD. Indivo-X: developing a fully substitutable personally controlled health record platform. AMIA Annual Symposium Proceedings. 2010;13:6-10.

[5] Genitsaridi I, Kondylakis H, Koumakis L, Marias K, Tsiknakis M. Towards Intelligent Personal Health Record Systems: Review, Criteria and Extensions, Procedia Computer Science, 2013;21:327-334.

[6] Kondylakis H, Kazantzaki E, Koumakis L, Genitsaridi I, Marias K, Gorini A, Mazzocco K, Pravettoni G, Burke D, McVie G, Tsiknakis M. Development of Interactive Empowerment services in support of personalized medicine, eCancer Medical Science Journal 2014;8(400).

[7] Kondylakis H, Koumakis L, Tsiknakis M, Marias K, Genitsaridi I, Pravettoni G, Gorini A, Mazzocco K. Smart recommendation services in support of patient empowerment and personalized medicine., Multimedia Services in Intelligent Environments – Recommendation Services, 2013;25:39-61, Springer-Verlag.

[8] Jurafsky D, Martin JH, Kehler A, Vander LK, Ward N. Speech and language processing: An introduction to natural language processing, Computational Linguistics, and Speech Recognition, 2000.

[9] Melucci, M.: Vector-Space Model. In: Liu, L., ÖZsu, M.T. (eds.) Encyclopedia of Database Systems, 2009, pp. 3259-3263. Springer.

[10] Boyer C, Selby M, Scherrer JR, Appel RD. The Health On the Net Code of Conduct for medical and health Websites. Computers in Biology and Medicine, 1998;28:603-610.

[11] Khazaal Y, Chatton A, Zullino D, Khan R. HON Label and DISCERN as Content Quality Indicators of Health-Related Websites. Psychiatr, 2012;83:15-27.

[12] Cao Y, Liu F, Simpson P, Antieau L, Bennett A, Cimino J, Ely J, Yu H. AskHERMES: An online question answering system for complex clinical questions. Journal of Biomedical Informatics (JBI), 2011;44(2):277-288.

[13] Roitman H, Mesika Y, Tsimerman Y, Maman Y. Increasing patient safety using explanation-driven personalized content recommendation. IHI, 2010, pp 430-434.

[14] Doupi D, Wijk M, Wyk J, Lei J. Personalizing Patient Education Using Internet Health Resources and EPR Data: Pilot Evaluation of the STEPPS Prototype System. CBMS 2008, pp. 449-454.

[15] Agliardi A, Jadad AR. Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination. BMJ. 2002;324:569–73.

[16] Scullard P, Peacock C, Davies P. Googling children's health: reliability of medical advice on the internet. Archives of disease in childhood. 2010;95(8):580-582.

[17] Berg GM, Hervey AM, Atterbury D, Cook R, Mosley M, Grundmeyer R, Acuna D. Evaluating the quality of online information about concussions. Journal of the American Academy of Physician Assistants. 2014;27(2):1-8.

[18] Wiesner M, Pfeifer D. Adapting recommender systems to the requirements of personal health record systems. In Proceedings of the 1st ACM International Health Informatics Symposium. 2010. pp. 410-414.

Address for correspondence

Haridimos Kondylakis

N. Plastrira 100, Vasilika Vouton,

Heraklion, Crete, Greece, GR-700 13

kondylak@ics.forth.gr