Semantically-enabled Personal Medical Information Recommender

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

Institute of Computer Science, FORTH, Heraklion, Crete, Greece
{kondylak, koumakis, psaraki, troullin, hatzimin, elenikaz, kmarias, tsiknaki}@ics.forth.gr

Abstract. Word wide web has become the first choice of patients to inform themselves about their disease, side effects and possible treatments. While patient's knowledge from internet is widely regarded as having a positive influence on the treatment, a lot of criticism exists for the quality and the diversity of the available information. In this paper we demonstrate the Personal Medical Information Recommender (PMIR), a semantically-enabled, intelligent platform that empowers patients to search in a high quality set of web documents for relevant medical knowledge. In addition, the platform automatically provides intelligent and personalized recommendations, according to the individual preferences and medical conditions. To demonstrate the platform example patients will be used to show the functionality of the system. Then we will allow conference participants to *directly* interact with the system to test its capabilities.

1 Introduction

During the last decade, the number of users who look for health and medical information has dramatically increased. However despite the increase in those numbers and the vast amount of information currently available online, one important challenge is the problem of the quality and the amount of information that can be found online [1]. It is widely accepted that it is very hard for a patient to accurately judge the relevance of some information to his own case.

Although similar approaches exist already, on the e-Health (such as WebMD (www.webmd.com) and HONSearch (www.hon.ch/HONsearch/Patients/)) and other domains [2], they are not dynamically adapted according to patient's preferences or his/her medical record. In this paper we demonstrate a Personal Medical Information Recommender (PMIR). PMIR [3-5] 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. The PMIR is integrated into an existing personal health record (PHR) as a set of individual apps. One of those apps is the *Document Registry* app which medical experts are using to register and annotate high-quality web documents. Then, the patient is able to select the *PMIR search* app to look for useful information. In addition, as the patient logs in to his PHR account, automatically, appropriate useful documents are recommended 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 platform that exploits individual patient's profiles to provide recommendations.

2 Architecture

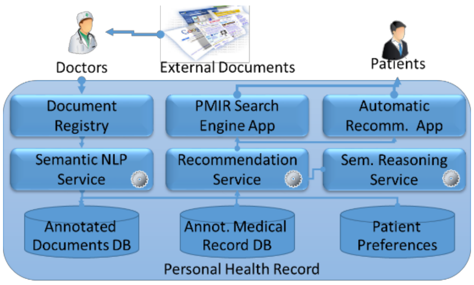

Fig. 1. - PMIR Architecture

The architecture of the system is shown in Fig. 1 and consists of three layers: The *database layer* including three different databases, the *service layer* including three services and the *front-end layer* including three individual apps.

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, domain experts (doctors etc.) use the *document registry* app in order to identify and register appropriate web documents (web pages, pdfs, docs etc.). Currently a possible cooperation with HONSearch is under consideration in order to register web documents already identified from similar multi-stakeholder consensus approaches. The selected unstructured text goes through a semantic lifting and indexing phase using *the Semantic Natural language processing (NLP)* Service. During this process a) the newly inserted web documents are annotated with terms from the SNOMED-CT, the LOINC and the RXTerms ontologies and b) the frequencies of the annotations in the documents and the collection are calculated. In this demonstration we assume that all annotations are correct. The results are then saved in the *Annotated Documents Database (DB)*.

On the other hand, using the PHR system, the patient is able to log and review his medical record. Similarly to the documents, the text entered by the patient is annotated using the aforementioned ontologies and stored in the *Annotated Medical Record DB*. In addition, users are able to search for relevant, high quality information using the *PMIR Search engine* app. User questions are also pre-processed and annotated at runtime using the Semantic NLP service and sent to the *recommendation service*. The results are returned and the user is able to rate the returned results. Each click and rating is recorded in the *Patient Preferences DB*.

As already described the recommendation service gets as input a) the annotated documents DB, b) the annotated medical record, c) the annotated user question and d) the preferences of the patient and returns a ranked list with the results. When the patient logs-into the system some recommendations are also automatically suggested to him using all input described except the annotated user question. The algorithm used for the recommendation service is a variation of the vector space model [6]. Documents are represented as vectors $d_i = (w_{1,d}, w_{2,d}, \ldots, w_{n,d})$ where each $w_{j,i}$ is a weight for the annotation j in document i, and reflects the importance of that term – based on the term frequency of the annotations in the document and the collection. On the other hand user requests are represented as vectors $r = (w_{1,Q},\ldots, w_{t,Q}, w_{1,P},\ldots, w_{m,P}, w_{1,I}, \ldots, w_{n,I},)$ where $w_{j,Q}$ is a weight for the *j*th annotation of the input query Q, w_{iP} is a weight for the *i*th annotation of patient profile *P* and w_{kI} is the weight of the inferred annotations I. Weights $w_{k,Q}$ and w_{mP} are calculated based on the term frequency inverse document frequency of the annotations in the collection and patient profiles respectively. Concerning the inferred annotations, for each annotation on the user query and the profile we call *Semantic Reasoning service* to get its three super-classes and sub-classes that are similarly used for matching annotations in documents.

The documents are first ranked by the cosine of the angle between the document vectors and the request vector, and these weights are multiplied by preference weights that change the order of the presented results. The preference weights are calculated based on personalized preferences (user clicks and ratings). Finally, documents with low similarity (below a threshold) are discarded, so as to exclude irrelevant document. Null queries can be submitted as well, that correspond to retrieval of documents pertaining to the medical record and the patient's preferences. This is used when calling the *Automatic Recommendation App*. Finally when no preferences are available (the cold start problem) they are not considered for recommendations.

3 Demonstration & Conclusion

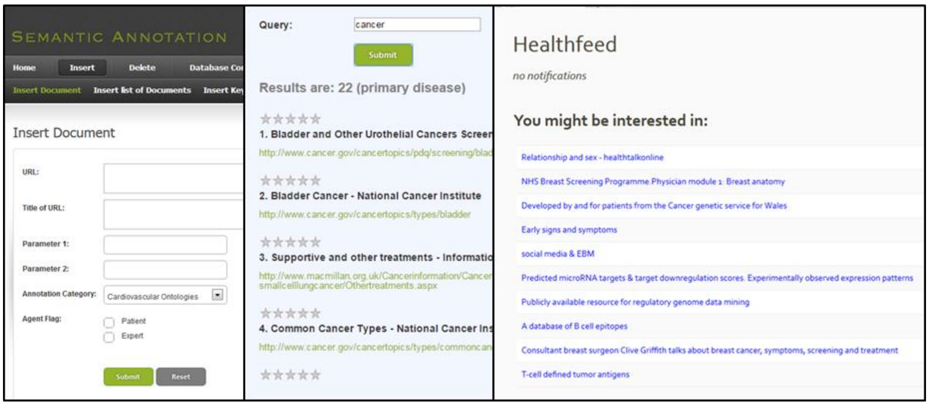

Fig. 2. The three main demonstration items, the semantic annotator (left), the semantic search engine (center) and the intelligent recommendations apps (right).

To demonstrate the functionalities of the PMIR system, dummy patients will be used. A first prototype of the Demo is online (http://139.91.210.42/ - login: peter, password: peter123 and select the *Recommender* App). A richer and more stable version of the platform is about to be released. The demonstration will proceed in five phases, the main components of which are shown in Fig. 2: (1) *Semantic Annotator*: The demonstration will start by presenting the app that the medical experts are using to register, delete and modify external documents that contain useful information to a targeted set of patients. The annotations performed automatically will be explained and some examples will be presented. (2) *Patient Medical Record:* Then, the demonstration will proceed by presenting the information a dummy patient has already stored in his account and how this information is also annotated using the semantic annotator service. (3) *Semantic Search Engine:* In this phase, we will search for useful information using the semantic search engine. Modifications will be performed to the patient profile and then the adaptation of these results according to the modified patient profile will be demonstrated. In addition some results will be rated and clicked and the change in the order of the results will also be shown. (4) *Intelligent Recommendations:* Besides allowing the patient to search information by himself we will also present some recommendations proposed to him automatically. Again, we will demonstrate how these recommendations are adapted dynamically by changing the profile of the user. (5) *Hands-on" phase:* In this phase conference participants will be invited to directly interact with the system, either guided or by accessing the online PHR system using their laptops.

To conclude, in this demonstration, we present a new platform that 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 such as additional comorbidities, patient preferences etc. We demonstrate our platform and allow the direct participation of the conference participants to test its capabilities. The extensive evaluation performed shows excellent result that will be published in a follow-up paper. To the best of our knowledge no other system, providing medical information, is able to be dynamically adapted in such a diverse environment.

Acknowledgments.

This work was partially supported by the iManageCancer (H2020-643529), the p-Medicine (FP7-270089) and the EURECA (FP7-288048) EU projects

4 References

1. Berg, G.M., Hervey, A.M., Atterbury, D., Cook, R., Mosley, M., Grundmeyer, R., Acuna, D.: Evaluating the quality of online information about concussions. JAAPA, 27(2), pp. 1-8 (2014)
2. Fossati, M., Giuliano, C., Tummarello, G.: Semantic Network-driven News Recommender Systems: a Celebrity Gossip Use Case. SeRSy, pp. 25-36 (2012)
3. Kondylakis, H., Koumakis, L., Rüping, S., Kazantzaki, E., Marias, K., Tsiknakis, M.: PMIR: A Personal Medical Information Recommender, MIE, pp. 1193 (2014)
4. Kondylakis, H., Koumakis, L., Kazantzaki, E., Chatzimina, M., Psaraki, M., Marias, K., Tsiknakis, M.: Patient Empowerment through Personal Medical Recommendations, MEDINFO, 216, pp.1117 (2015)
5. 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., Recommendation Services, 25, pp. 39-61 (2013)
6. Salton, G., Wong, A., Yang. C.: A vector space model for automatic indexing. Communications of the ACM, 18, no. 11, pp. 613-620 (1975)