*Medical Informatics in a United and Healthy Europe*
*K.-P. Adlassnig et al. (Eds.)*
*IOS Press, 2009*


*doi:10.3233/978-1-60750-044-5-903*

# Challenges and Opportunities of Using Recommender Systems for Personalized Health Education

Luis FERNANDEZ-LUQUE [a,1], Randi KARLSEN [a,b], Lars K. VOGNILD [a]
[a] *Northern Research Institute (Norut), Tromso, Norway*
[b] *Computer Science Department, University of Tromso, Norway*

**Abstract.** The use of computers in health education started more than a decade ago, mainly for tailoring health educational resources. Nowadays, many of the computer-tailoring health education systems are using the Internet for delivering different types of health education. Traditionally, these systems are designed for a specific health problem, with a predefined library of educational resources. These systems do not take advantage of the increasing amount of educational resources available on the Internet. One of the reasons is that the high availability of content is making it more difficult to find the relevant one. The problem of information overload has been addressed for many years in the field of recommender systems. This paper is focused on the challenges and opportunities of merging recommender systems with personalized health education. It also discusses the usage of social networks and semantic technologies within this approach.

**Keywords.** recommender systems, health education, personalization, tailored health education

## 1. Introduction

Health Education aims to promote, maintain or restore the health of individuals by informing and persuading. Internet is becoming one of the main sources of Health Education and Information, and there are a wide range of resources from different providers. In fact, recent studies show that most Europeans search in the Internet for health information [1]. Despite the availability of many good educational resources, the task of identifying the relevant ones can be very difficult for the user. Additionally, the users have to deal with the appearance of many health resources of a very low quality, or even with misleading information, such as content promoting anorexia (aka proana). One of the main reasons for the difficulty of finding relevant educational resources is that the health information space on the Internet is huge and it is leading to an information overload problem.

The use of computers in Health Education started in the mid 90s as a tool to assist the health educators for tailoring resources [2–4]. Computer-Tailoring Health Education Systems (CTHES) could be seen as Expert Systems for automating many of the tasks performed by the health educators. CTHES are used to create educational

[1] Corresponding Author: Northern Research Institute, Postboks 6434 Forskningsparken, 9294 Tromsø, Norway; E-mail: luis.luque@norut.no.

resources adapted to the unique persons needs. The most common approach is the creation of motivational messages based on acquired parameters about the health or mental status of the user. Internet is also increasingly being used in CTHES as a channel to deliver different types of educational resources (text, videos, web, etc.). However, Internet is not used normally as a source of additional resources. One of the reasons is the health information overload problem described in the first paragraph. This problem has to be tackled in order to reach the full potential of the Internet for health education. This problem has been studied for many years in the Information Filtering field where one of the most common approach is Recommender Systems (RS) [5, 6]. An RS provides recommendations of items (pieces of information) which are predicted to be interesting for the user. For example, an RS could deliver recommendations of health web sites based on information in the user's health record.

In MyHealthService [7], we are exploring the usage of RS for Health Education. During the SoA we identified many challenges and opportunities but found few examples [8, 9]. This paper provides an overview of RS and CTHES highlighting the challenges and opportunities of using RS for health education. In addition, we discuss the opportunities of using Social Networks and Semantics.

## 2. Recommender Systems and Personalized Health Education

De Vries [2] defined computer-tailoring health education as "the adaptation of health education to one specific person through a largely computerized process". This adaptation implies two processes: 1) the selection of an adequate resource and 2) its adaptation to the patient's needs. Their main components are the user models, the resources library and an algorithm. The educator's expertise is needed to build the algorithm, which has to select and personalize an educational resource for a patient.

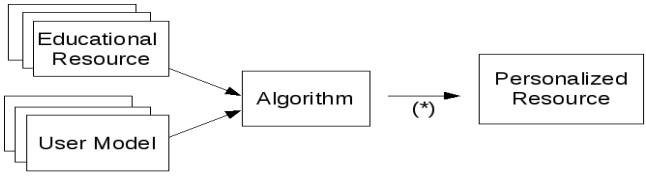


**Figure 1.** Main elements of an RS or CTHES. (*) In RS it is only the selection of the resources, and in CTHES it comprises the selection and the adaptation.

RS emerged in the mid 90s, as a set of techniques to solve the problem of the information overload. Since then, many techniques have been studied by the industry and the research community, especially since the raise of the Internet. Their objective is to select items which satisfy the interest of the users. As in the case of CTHES, also RS comprise mainly 3 elements: a user's model, a collection of resources and a recommender algorithm. CTHES are normally tackling a specific health problem with the supervision of healthcare professionals. Traditionally these systems address a well known set of users and educational resources and the personalization is based on a predefined set of health related parameters. The specialized context and the dependency on the experts, in CTHES, make it difficult to: 1) take advantage of additional web educational health resources and 2) adapt the systems to the unique and changing health of the patients (e.g., with more than one health problem at the same

time). In contrast, RS are designed for more general scenarios with less predefined knowledge of their users and resources and relaying mainly on the knowledge gained through the user's interaction. This approach requires less involvement of experts when including new users and resources, but has several weaknesses that we will address in following sections.

### *2.1. Algorithms*

In CTHES, the algorithms for personalization are based on the human expert knowledge, normally defined in a set of rules, grounded on behavior theories, whereas in RS, the algorithms are normally based on the automatic analysis of the interaction of the users. The most common approaches in RS are collaborative-based and content-based. In the first one the recommendations are based on the ratings by similar users and in the second one on the user's previous ratings. Both approaches success dependent on users ratings, and in addition the content-based systems need a description of the features of the resources. There are also more advanced types of RS algorithms based of knowledge on how the characteristic of a user matches a resource. All these different types of algorithms can be combined in hybrid RS [6]. However, in the context of RS for health education the problem of the lack of information about new users could be compensated by accessing their Electronic Health Record (EHR), which is a common approach in CTHES.

**Table 1.** Challenges and opportunities in the algorithms

| CTHES | RS | RS for Health Education |
|---|---|---|
| Grounded on medical expert knowledge for a specific health problem.<br><br>High dependency on the human experts: it makes it difficult to increase number of users or resources. | *Collaborative-based:* based on the ratings of similar users<br>(-) need of critical mass of users<br>(-) lack of knowledge about new items and new users<br>*Content-based:* based on the user's previous ratings of similar items<br>(+) results don't depend on other users<br>(-) need of information about the items and previous user ratings<br>(-) super specialization of the recommendations | RS approaches could be used to gather aspects, which are not taken into account by CTHES approaches (such as personal preferences).<br><br>The RS algorithms could learn from the previous interactions with the system.<br><br>RS could be used to incorporate additional resources from the Internet. |

### *2.2. User Modeling*

In CTHES the user models traditionally contain the parameters predefined by the experts, normally covering health and mental status. These parameters are captured from questionnaires and the patient's health records. In RS the initial knowledge about the user is limited and the use of extensive questionnaires is not common because it could hinder a user's interest in using the system. Also, RS normally do not have access to any kind of database about the users and therefore the most common approach is to gather knowledge from the analysis of user's interaction with system (e.g., ratings, usage patterns, etc). The strength of this approach is the automatic user modeling, but its weakness is the time needed to establish the user model. In CTHES the library with the education resources are created by the professionals. To incorporate a new resource it has to be described by the professionals and it may imply to adjust the tailoring algorithm. This approach guarantees the quality of the resources but makes it difficult to dynamically incorporate additional resources from the Internet. However,

RS are designed for scenarios of continuous increase in the amount of resources and adding new resources are much easier. In content-based RS the new resources have to be described while in collaborative-base RS this is not even necessary. The challenge in collaborative RS is the lack of knowledge about resources that have few user ratings. In the context of Health Education, an RS approach could establish knowledge on a resource based on information provided by the patients. For example, the ratings and user comments on resources could complement the description of the resources created by the professionals. In addition, RS approaches are more suitable for the integration of additional educational resources available on the Internet.

**Table 2.** Challenges and opportunities in user modeling

| CTHES | Recommender Systems | RS for Health Education |
|---|---|---|
| *Questionnaires:*<br>(+) based on well-known medical theories<br>(-) can be a burden to the patients<br>*Health Records (EHR, PHR):*<br>(+): could be a huge source of knowledge<br>(-): integration problems | *Automatic user-modeling:* the recommendations are based on the user's ratings of the items<br>(-) lack of information about new users (cold start problem)<br>(+) the users don't need to input data<br>(+) easy to maintain | 1.- The access to the PHR of the users could reduce the problem of the lack of information about new users<br>2.- The analysis of the user's ratings could enrich the traditional CTHES approach by taking into account the user's preferences |

### *2.3. Opportunities with Social Networks and Semantic Web*

Social Networking platforms are used in both RS and CTHES, not just to deliver content but also to harvest knowledge about the users and the content. For example, in RS the analysis of Social Networks has been used to discover emerging topics of interest for guiding the recommendation algorithm [10, 11]. In CTHES Social Networking platforms have been used to gather information about the users. For example, in the project RiskBot researchers are using automatic analysis of public profiles in MySpace for tailoring public health interventions [11].

Another promising approach in both RS and CTHES is the use of Semantic Technologies to enhance the integration of the different components of the system and for the creation of semantic-based algorithms [12–15]. Semantic technologies could enhance the adaptability, for example adding a new data source could be done by changing the data models and not having to redesign the system. This can be of great interest for the reason that the health of the patients is unique but also changing. The application of semantic technologies for health education can be catalyzed by the appearance of semantic standards for health information (e.g., OpenEHR, HL7 v3, etc.) and semantic repositories for health information sites [16].

We believe that the user generated content on the Internet is a promising source of information for powering RS based health education. In [9] tags created by users are used for delivering health information within a patient community. We presented some early results on users publishing about their health on Internet [17], and we have detected personal health information disclosed in comments on YouTube videos [18].

## 3. Conclusions

The application of Recommender Technologies for Health Education could become one solution to tackle the health information overload on the Internet. The use of

Recommender System for health education is in an early stage and few examples can be found in the literature [8, 9]. This lack of previous studies on the applicability of RS for Health Education was the main rationale of this paper, where we provided an overview of the two different fields: RS and CTHES, and discussed the promises of RS for health education. We also discuss the emerging approached in both fields towards the usage of Semantic technologies and Social Networks for enhancing their performance, and report our own promising results about patient generated content.

**Acknowledgments.** I would like to thank the rest of the MyHealthService's team. This research belongs to the Tromsø Telemedicine Laboratory, co-funded by the Research Council of Norway.

## References

[1] Andreassen, H.K., Bujnowska-Fedak, M.M., Chronaki, C.E. et al. (2007) European citizens' use of e-health services: A study of seven countries. *BMC Public Health* 7:53.
[2] de Vries, H., Brug, J. (1999) Computer-tailored interventions motivating people to adopt health promoting behaviors: Introduction to a new approach. *Patient Education and Counseling* 36(2):99–105.
[3] Cawsey, A. et al. (2007) Adaptive information for consumers of healthcare. In Brusilovsky, P. et al. (Eds.) *The Adaptive Web: Methods and Strategies of Web Personalization*. Springer-Verlag, Heidelberg, LNCS 4321:465–484.
[4] Lustria, M. et al. (2009) Computer-tailored health interventions delivered over the web: Review and analysis of key components. *Patient Education and Counseling*, 74(2):156–173.
[5] Adomavicius, G., Tuzhilin, A. (2005) Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering* 17(6):734–749.
[6] Burke, R. (2002) Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction* 12(4):331–370.
[7] Burkow, T.M. et al. (2008) An easy to use and affordable home-based personal eHealth system for chronic disease management based on free open source software. In *Proceedings of MIE 2008*, *Studies in Health Technology and Informatics* 136:83–88.
[8] Witteman, H., Chignell, M., Krahn, M. et al. (2008) A recommender system for prostate cancer websites. *AMIA Annual Symposium Proceedings 2008,* 1177.
[9] Khan, S.A. (2008) A Tag based Recommendation Engine to Suggest Information Resources in an Online Community for Health Promotion. *AMIA Annual Symposium Proceedings 2008*, 1002.
[10] Groh, G., Ehmig, C. (2007) Recommendations in taste related domains: Collaborative filtering vs. social filtering. In *Proceedings of the 2007 ACM Conference on Supporting Group Work*, ACM, New York, 127–136.
[11] Bonander, J. et al. (2008) Personally tailored health information: a health 2.0 approach. In Eysenbach, G. (Ed.) *Medicine 2.0 Proceedings, Social Networking, Collaboration, Participation, Apomediation, and Openness. Journal of Medical Internet Research* 10(3):e22.
[12] Sheth, A., Nagarajan, M. (2009) Semantics-empowered social computing. *IEEE Internet Computing* 13(1):76–80.
[13] Middleton, S.E. et al. (2004) Ontological user profiling in recommender systems. *ACM Transactions on Information Systems* 22(1):54–88.
[14] Erriquez, E., Grasso, F. (2008) Generation of Personalised Advisory Messages: An Ontology Based Approach. In *Proceedings of CBMS 2008*, IEEE, 437–442.
[15] Cantador, I. et al. (2008) Ontology-based personalised and context-aware recommendations of news items. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, Washington, DC, USA, 562–565.
[16] Stamatakis, K. et al. (2007) *Aqua, a System Assisting Labelling Experts Assess Health Web Resources*. In *Proceedings of the 12th International Symposium on Health Information Management Research (iSHIMR 2007)*, 75–84.
[17] Fernandez-Luque, L. et al. (2008) Study of the epatient as a generator of content. In Eysenbach, G. (Ed.) *Medicine 2.0 Proceedings, Social Networking, Collaboration, Participation, Apomediation, and Openness. Journal of Medical Internet Research* 10(3):e22.
[18] Fernandez-Luque, L. et al. (2009) An analysis of personal medical information disclosed in YouTube videos created by patients with multiple sclerosis*. In this volume*.