Towards Health (Aware) Recommender Systems

Hanna Schäfer∗
Technical University of Munich

André Calero Valdez
RWTH Aachen University

Tom Ulmer
FHS St. Gallen

Santiago Hors-Fraile
University of Seville

Alan Said
University of Skövde

Raghav Pavan Karumur
University of Minnesota

Helma Torkamaan
University of Duisburg-Essen

Christoph Trattner
MODUL University Vienna

ABSTRACT
People increasingly use the Internet for obtaining information regarding diseases, diagnoses and available treatments. Currently, many online health portals already provide non-personalized health information in the form of articles. However, it can be challenging to find information relevant to one’s condition, interpret this in context, and understand the medical terms and relationships. Recommender Systems (RS) already help these systems perform precise information filtering. In this short paper, we look one step ahead and show the progress made towards RS helping users find personalized, complex medical interventions or support them with preventive healthcare measures. We identify key challenges that need to be addressed for RS to offer the kind of decision support needed in high-risk domains like healthcare.

KEYWORDS
Health Recommender Systems; Health Informatics; Patient Modeling; Disease Modeling; Online Health Interventions;

ACM Reference format:
https://doi.org/http://dx.doi.org/10.1145/3079452.3079499

1 INTRODUCTION
Nowadays we rely on recommender systems to help us in making choices related to our entertainment and e-commerce - finding media to consume or products to purchase. Amazon, Spotify, Trivago, and Netflix – all rely on recommender algorithms to boost their sales. But when it comes to areas concerning our health such as nutrition, exercise, medication, diagnoses, and treatments, recommender systems are still in their infancy concerning trustworthiness and reliability [39]. While there is great potential for development of Health Recommender Systems (HRS) [7], the potential risks of using them are also great.

In this work, we attempt to shape popular discussion topics dealing with HRS. The variety of those topics results from the number and diversity of stakeholders involved in health systems. Taking the patient perspective, HRS need to provide a simple interaction, empowerment through explanations of the proposed recommendations, and safety against harmful recommendations. This will allow the patients to trust the system. For clinicians and experts on the other hand, what matters is a precise and correct representation of their domain knowledge and processes. Finally, health care providers, insurance companies, and clinics are interested in success rates, study results, and financial benefits of the new systems.

Thus, in the following sections we will discuss HRS from the user, system, and evaluation perspective starting with the traditional systems, continuing with the current approaches, and ending with a glimpse into future challenges.

2 TRADITIONAL RECOMMENDER SYSTEMS
In traditional RS, user preferences are derived from ratings and utilized to predict the users rating on new items. Below we shortly discuss those basics and why they cannot completely fulfill the purpose of advanced HRS.

User Preferences. RS typically elicit and determine the user preferences by explicit and implicit user actions (such as ratings, likes, etc.). This works well, since those systems want to find the items that represent the users’ taste.

Advanced HRS would require information on user needs, context, diseases, ethnicity, etc. to find the optimal item, since those characteristics might contrast the users’ preference or each other. For example in food recommendations [9, 10], the user’s preference for ice cream, may contrast his need for recommendations of foods for diabetics.

Rating Prediction. For many years, the primary focus of RS research lay on the task of predicting users’ preference, i.e. what a certain user’s rating for a certain item is going to be [36]. Initiatives like Netlix Prize1, further established rating prediction as the "default" recommendation scenario. This led to rapid advancements in rating prediction, establishing matrix factorization methods like Singular Value Decomposition as the state of the art [22].

Similar approaches have already been used in the health domain, i.e. for smoking cessation [1, 27]. While the question of privacy

1www.netflixprize.com

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

DH'17, July 02-05, 2017, London, United Kingdom
© 2017 Association for Computing Machinery.
ACM ISBN 978-1-4503-5249-9/17/07...$15.00
https://doi.org/http://dx.doi.org/10.1145/3079452.3079499
within the Netflix Prize scenario might be worrisome for health applications, some privacy solutions have already been developed without critical loss of accuracy performance [29].

RS Evaluation. Traditional RS were evaluated primarily based on how well a certain recommendation approach can reproduce recording of earlier interactions between users and items [12], i.e. either predicting whether a user has liked/viewed an item, or predicting the rating a user has already given to an item. However, recently, rating prediction has come under more scrutiny, mainly because ratings given by users do not necessarily reflect the actual intent of the users [11]. Instead RS are starting to look into different utility functions such as the ranking’s quality in terms of diversity, novelty, and general utility [18].

For HRS this utility function could be extended by medical utility functions such as treatment duration or pain relief. Those measures may even be affected by their health and state of mind [47]. Besides these single utility functions, HRS also need to consider potential harms, contradicting utilities as well as ethical issues. This highlights the importance of user studies in the evaluation of HRS compared to numerical measures.

3 STANDING AT THE BRINK

Many attempts on implementing HRS have been made in the last five years. Table 1 shows the top concepts mentioned in the titles of papers found on Google scholar published during these years. Clearly, not only the number of papers on HRS is increasing but also their variety in topics. Some of those trends such as personalization, lifestyle, and the distinction between patient and expert systems will be discussed in the following section.

User Profiling and Personalization. The biggest trends in HRS (table 1) considering the user perspective are medical user profiling and medical personalization of recommendations.

The challenges regarding user profiling already start with the retrieval and selection of user health data. Given that data is available in a standardized format, it needs to be summarized into a patient profile [45]. The closest process to traditional RS is to recommend health artifacts (e.g. articles, doctors, therapies, fitness plans, etc.) based on their popularity with similar patients. However, as the field of HRS offers a great variety of user needs, the profile needs to fit all intended goals. Lopez-Nores et al.’s HARE, take this challenge by implementing ‘property-based’ collaborative filtering that combines traditional content based information from the electronic health record with collaborative methods [24].

Personalization takes these approaches one step further and has been shown to be an important factor affecting patient satisfaction. Personalized recommendations increase the details a recommendation can provide and improve the users’ understanding of their medical condition. Luo et al. propose an application that can recommend relevant home medical products to patients based on their electronic health record [25]. Besides the quality and popularity of these products they can be selected based on diseases of the user (e.g. Alzheimer’s). Others recommend artifacts based on stage of progression in their specific disease [28]. Diving deeper into models for specific diseases, Hu et al. [16] use Patient Similarity Analytics to derive optimal long time prognosis of treatment plans and to visualize them. To do this, they combine data from heterogeneous sources such as demographic information, diagnosis, medication, and lab tests. Then they optimize them for physical health, mental health, quality of life, and risks.

All these systems create first personalized user models based on digital health records. The challenges include availability and quality of personal health data, the variety of user needs that can result from very similar health records and even the adaption of recommendations to the knowledge level of the user/patient.

Persuasion, Empowerment, and Trust. An important change when moving from traditional RS to HRS is the relevance of how recommendations are presented and how the user can interact with them e.g. in a mobile system (table 1). These concepts include persuading the user in case of inconvenient recommendations, empowering the user by interactively guiding his decision, and creating trust, using expert input or explanatory interfaces.

Existing HRS can be more effective when combined with persuasive elements that not only inform but also steer the users. For instance, Radha et al. [34] tried to improve recommendations in lifestyle adaptation for hypertension control and prevention by using a strategy that maximizes engagement and motivation. Alongside persuasion, the nudging theory [43] can be used to lead people towards healthier habits. For instance, to accomplish behavior change, Reimer et al. [35] introduced a self-learning framework as a comprehensive approach for nudging in health apps. One important part of it is personalization, which is based on an individual goal achievement graph, tailoring the different nudge types to the user behavior [19] and triggering situation-specific nudges.

Although nudging has been proven effective in many cases, it raises some ethical concerns. If the user is aware of being manipulated into changing his behavior, it might have less effect. On the other hand, if he is being persuaded subconsciously, the recommender should take more responsibility in terms of safety. The opposite approach would be to empower users by showing them the internal logic of the recommender. For example, Hammer et al. [13] integrated rule-based domain knowledge into HRS to empower an elderly population of patients in their lifestyle decisions. In contrast to subconsciously nudging people, explanation nudges

<table>
<thead>
<tr>
<th>Year</th>
<th>#Papers</th>
<th>Top-Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>1</td>
<td>HRS, Personalization, Semantic Network</td>
</tr>
<tr>
<td>2013</td>
<td>2</td>
<td>HRS, Context-Aware, Health Data</td>
</tr>
<tr>
<td>2014</td>
<td>5</td>
<td>Trust-Based, Personalization, Diet, Mobile, User Profiling</td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>HRS, Nutrition, Social Network, Mobile, Semantic-Based, Health Data, Patient-Centered</td>
</tr>
<tr>
<td>2016</td>
<td>7</td>
<td>HRS, Tailoring, Smoking, Harm, Interactive, Expert-In-The-Loop, Cloud, Sensor-Network, Mobile, Patient-Centered, Social Network, Patient Satisfaction</td>
</tr>
</tbody>
</table>

Table 1: Using the search term "health recommender systems" on Google Scholar, the authors manually reviewed the relevant HRS publications of the last 5 years and selected the major concepts from their titles.
try to empower both patients and physicians by presenting the
recommenders reasoning in an understandable way. When provid-
ing such explanatory content, nudging can fulfill both persuasion
and empowerment and reduce trust issues by demonstrating the
thoroughness of the evaluated parameters [33].

For the patient, this is currently best solved by involving a hu-
man expert such as a physician in the loop [15]. To generate this
kind of knowledge transfer from experts to patients in an efficient
way, generated explanations need to reach deeper into domain
knowledge. At the same time, given similar content quality, gener-
atated explanations will in many cases meet less trust than the expert
in the loop solution [15]. Thus such a system would also need to
simulate natural human interaction patterns.

Medical Evaluation, Lifestyle Interventions, and Patient Satis-
faction. Since HRS focus on other outcome measures than traditional
RS, new evaluation methods are required. Those include measures
such as medical health improvement or behavior change measures.

One way to evaluate HRS is to view them as a medical interven-
tion and to measure objective medical outcomes, e.g. improving
weight and blood values due to dietary changes[42]. Such interven-
tions are usually evaluated online in A/B tests or randomized con-
trolled trials. Since such trials require a lot of resources, they should be
preceeded by extensive offline evaluation on the recommender
system. In preparing an intervention study the responsibility of
medical effects has to be defined. Since the applications cannot fully
guarantee medical correctness, these studies need to be conducted
with a human medical support system. Moving on, HRS will need to
reach a safety level/ liability comparable to human experts to be
able to fully use medical evaluation and intervention techniques.

For medical interventions regarding lifestyle we often cannot
evaluate objective measures but only the reported behavioral change
[31]. A meta-analysis by Lustria et al. [26] indicates that physical
activities (42%), nutrition (25%), smoking (18%), and drinking (9%)
are the most common scopes for such web-delivered behavioral
interventions. It is well established that messages contribute posi-
tively to smoking cessation [38]. Thus the Smokefree Brain Project2
evaluates a HRS for smoking cessation using a combination of of-
line measures and a post-trial survey. Nutrition is another area
where self-reporting is more common than ground truth measures.
Here HRS can supply feedback messages on nutrition behavior
and transfer knowledge by creating transparent recommendation
explanations. For example, [42] describes a behavioral intervention
study on nutrition recommendations with a matrix factorial design.
Besides blood values, they use dropout and self-reported dietary
intake protocols as outcome measures.

4 VISION OF THE FUTURE

To fully grow their potential, HRS will need to address a lot of
key challenges, such as ethics, domain modelling, user interaction
and privacy. At the same time Cappella et al. [4] note that the
application of RS in the health domain is not free of the current RS
challenges such as sparsity, and a limited number of items. Below
we discuss a selection of prominent future research questions.

User Perspective - Forming a Patient Profile. Although the avail-
able patient data (electronic health records, fitness and lifestyle
trackers) is growing extensively, it is lacking quality, consistency
and compatibility in many places (e.g., charts, records, journals,
diaries, or files). Obtaining periodic information about the patients’
habits, lifestyle changes, socio-economic status, feelings, experi-
ences, perceptions, and behaviors could increase the relevance,
diversity and accuracy of recommendations [17]. Additionally, lat-
est health information from the patients’ family members could be
obtained to make recommendations of diagnoses, treatments or
medications based on potential risks the patient might carry.

However, obtaining such information is still challenging. It re-
 mains important to create an accurate user profile by consolidating
from various sources and picking relevant features for the recom-
manations. This requires standardization of data formats, authen-
ticity of data sources and automated update intervals. Furthermore,
letting the patients modify their own profiles (as in a traditional RS)
may have unintended consequences such as patients (1) reporting
erroneous information, (2) misperceiving or misinterpreting their
conditions (e.g. having a ’’depression’’) [48], (3) manipulating their
situations to get fake diagnoses or treatments suggested to them for
exploiting their employers or other ulterior purposes [17].

User Perspective - Identifying Patient Needs. One aspect of HRS are
the treated illnesses and the different needs that come with them.
Several chronic or life-threatening diseases usually have stages
through which the patients progress. Palliative care for example, is
aimed improving the quality of life by helping patients make
informed decisions about their care [8, 40]. Recent reports show
that patients in such situations are often not educated about their
disease and treatment options until those later stages [6]. In such
contexts, HRS could make an early diagnosis and use techniques
such as disease progression modeling [44] to make appropriate
recommendations of health artifacts to thus satisfy the patients
need to accommodate their therapy to changing requirements [20].

Going further, people with multiple health conditions may have
specific challenges, disabilities and comorbidities. The onset of
a challenge, or comorbidity could reveal an underlying medical
condition. HRS could diagnose such underlying health conditions
early on and thus begin providing early care through recommenda-
tions, which is otherwise not possible. All these aspects of patient
diversity and needs must be regarded in future HRS.

Recommender Perspective - Personalization and Models for Behav-
ioral Change. Overloading the system with all kinds of information
either by the patients themselves, or by the algorithms that inte-
brate electronic health records from multiple heterogeneous sources
invariably result in thousands of features as part of a patient profile.
Yet only a subset may turn out to be useful for a given medical con-
dition, context, or time instance. At the same time it is difficult to
detect, when a recommendation I brought into action, which makes
it harder for the recommender to explicitly learn about positive
reactions and thus about relevant features.

Another approach beyond providing recommendations, is to
serve as a decision-support tools affecting behavior and showing
long-term effects on their users. While some health aspects address
larger target groups, most patients prefer to receive personalized
recommendations, e.g. using individual messages. HRS can serve

2www.smokefreebrain.org

Towards Health (Aware) Recommender Systems

DH’17, July 02-05, 2017, London, United Kingdom
these needs and present individual recommendations in a decision interaction similar the physician-patient-interaction. For instance, Wiesner et al. [46] investigated giving laymen-friendly information to the users based on their individual health records. Lathia et al. [23] discussed the potential of using interactions (e.g., ratings) from RS in providing personalized health-related feedback. HRS could even adapted personalized recommendations to behavioral change phases, such as Transtheoretical Behavioral Change Model [32] or the Health-Belief model [2]. For those theories in which the user experiences different phases (i.e. the Transtheoretical Behavioral Change Model), it is possible to address the phase the user is in, with switching HRS depending on the specific needs in that phase.

Recommender Perspective - Simple and Concise Explanations. Herlocker et al. [14], in their work on explaining recommendations, name multiple reasons why collaborative filtering approaches may not be trusted for high-risk domains as they can occasionally lead to wrong recommendations. The most important of these is that such approaches are based on very sparse data. Therefore, there is a high need for explanations in a high-risk domain such as HRS to help the patients decide how much to trust the recommendations. However, factors such as age, pain, stress, anxiety and memory may limit the amount of information a patient can process [47]. Also, certain age groups may be less familiar with terminology, algorithms, and approaches of RS or less computer literate in general (e.g. older adults). Therefore, any additional information should be as easy to understand as possible, recognizing that the overall goal of RS is to reduce information overload. Sadasivam et al. [37] describe that besides tailoring to a collection of relevant topics the messages should also be written target-oriented. This includes adapting messages with specific terminology and linguistics in the users’ health records and considering the confidentiality of these records [46].

Recommender Perspective - Privacy. When RS are applied to the health domain, a holistic picture including privacy-concerns must be drawn [3]. Users might see utility in data disclosure and might overlook possible de-anonymization risks. The problem lies in the sparsity of data in RS. Narayanan et al. presented a robust de-anonymization approach for sparse datasets [30]. Some approaches to integrate privacy-preserving algorithms into RS already exist [29]. Yet, explaining possible privacy risks to users might in fact reduce trust in the system, even if the system honestly, in a simple and understandable manner, communicates the actual effects of entering personal information [21]. Still, personal privacy needs might differ between users and thus approaches of personalized privacy preservation help maximize individual utility [5] for HRS.

Recommender Perspective - Drug Usage Approval and Complementary Medicine. Regulations, laws and standards for approving usage of drugs, food or other health-products vary strongly across nations. HRS should therefore only consider recommending products that are approved for use locally. Existing medical information portals (e.g., Mayo Clinic) provide unbiased information about alternative systems of medicine (e.g. acupuncture, homeopathy, etc.) as non-personalized recommendations. However, there are various opinions among researchers around the globe on whether such forms of treatment are to be accepted or not. Recommending them to patients in specific locations around the world can therefore have similar challenges which need to be dealt with appropriately.

Evaluation Perspective - User Satisfaction. Since utility and preference might go in different directions for HRS, measuring the overall user satisfaction can be challenging. Should the recommended artifacts be familiar (so that they gain trust and have knowledge of how the treatment is going to work), or novel (help them discover new medicines, therapies, etc.), or serendipitous (for lightening their mood by recommending non-medical artifacts)? How should their implicit data be perceived (as lack of interest in pursuing, or fear of using)? Do patients feel in control of their HRS or do they have difficulties understanding the suggestions? These are just some of the open questions that future HRS need to consider when defining their domain specific user satisfaction measures.

Evaluation Perspective - In Situ Evaluation. In order to prove their worth, HRS will have to be evaluated in real life non-laboratory settings. When risks are real and every recommendation matters, new ways of evaluation will have to emerge. HRS should certainly be evaluated in terms of their ability to improve the quality of care (through accuracy, relevance and early diagnosis) and reduce the costs of that care. They must also be evaluated for factors such as their robustness to false information, consider potential health risk factors based on age, culture, ethnicity, etc. Additionally, the long-term behavioral effects must also be investigated in-situ to address the complexity of health and health behavior.

Evaluation Perspective - Consequences of HRS. The negative consequences of a false recommendation in HRS can be disastrous for health and life of patients. Errors anywhere along the process might have serious ramifications on the patient’s trust and emotions. Explanations (containing details of their health profiles, specific matching features, and relevance scores) may help to a certain extent [14]. On the other hand, recommendations could also “threaten the control, autonomy, and authority of providers based on traditional provider-patient roles” [41]. Thus, it is critical to create recommendations tailored for each person and have the health provider in the loop, depending upon the level of risk associated with the artifact being recommended. For example, patients from poorer socio-economic backgrounds may not be able to afford expensive items, patients with certain diseases may be more prone to anxiety and over-interpretation, and people with specific genetic signatures may be more susceptible to certain medical conditions. Any of these harmful side effects need to be preconsidered in HRS.

5 CONCLUSION - STEPPING STONES

Moving towards HRS is a transition that challenges researchers with new questions both in traditional recommender problems and in domain specific problems. The key concepts in HRS that were tackled within the last 5 years are the personalization of recommender systems to individual patients and their current health context, the balance between persuasion and empowerment as well as the effect of both methods on the users trust and the evaluation methods (e.g. intervention studies) and measures (e.g. user satisfaction) in HRS. Building on this, the identified key challenges can be subdivided into patient/user challenges, recommender challenges and evaluation challenges. On the patient perspective, HRS
Towards Health (Aware) Recommender Systems

will need to collect data from a wide variety of sources, such as electronic health records or lifestyle trackers, interconnect and standardize these data entries, and assure their quality regarding missing or false information. From there the specific user needs and requirements need to be filtered from the available datasets using intelligent user models. On the recommender perspective, those models should then be used to personalize the given recommendation to the user’s health context, history and goals. Once those recommendations are achieved, HRS need to guide the user properly solved for HRS, they can evolve towards more sophisticated systems such as digital health assistants or medical advisors.

6 COMPETING INTERESTS

The authors have declared that no competing interests exist.

REFERENCES