Wellbeing Recommender System, a User-Centered Framework for Generating a Recommender System for Healthy Aging

Jordi Torres¹, Meritxell Garcia¹, Garazi Artola¹,², Teresa Garcia-Navarro¹, Isabel Amaya², Nekane Larburu¹,², and Cristina Martin¹,³

¹Vicomtech Foundation, Basque Research and Technology Alliance (BRTA), Mikeletegi 57, 20009 Donostia-San Sebastián, Spain
²Biodonostia Health Research Institute (Bioengineering Area), eHealth Group, 20014 Donostia-San Sebastián, Spain
³Faculty of Engineering, University of Deusto, Av. Universidades, 24, 48008, Bilbao, Spain

Keywords: Healthy Aging, Recommender System, Quality of Life, Synthetic Data Generation.

Abstract: The needs of the currently aging population require new technologies to support them in order to offer them a decent quality of life. Different interventions have been proposed in the last years to face this challenge, where recommender systems are gaining strength. The general objective of these systems is to promote the adoption of healthy habits among the end users, but sometimes they show limitations in the fulfilment of this goal. To overcome these limitations, our approach offers an easy to maintain, interoperable, and personalized recommender system capable of providing recommendations based on individuals’ daily activity data. A methodology is presented for the generation and management of wellbeing recommendations, which are then tested using a synthetically generated dataset that simulates a variety of user categories. With the evaluation of this data, a technical validation is carried on to assess the performance and scalability of our developed system.

1 INTRODUCTION

The current trend of population aging, especially in developed countries, will pose several challenges to our society, comprising from changes in the structure of health and social services, as well as the financial system and labour markets (Ahtonen, 2012). Predictions indicate an increase of the burden of age-related expenditures in state budgets. Increasing life expectancy comes with a variety of changes in the care of elderly people as we know it today, with growing evidence that a shift from targeting individual diseases to postponing physical deterioration and comorbidities (Goldman et al., 2013). Considering that the elderly population is the group that requires the most from health services and with age is affected by more comorbidities, interventions that can lead to an improvement of quality of life that can turn into an increase in the healthy life-expectancy (Södergren, 2013), which is the years that a person lives free of disability (Jagger & Robine, 2011). In return, the burden to the health services is reduced as the onset of disability is delayed (Beltrán-Sánchez et al., 2015; Mehta & Myrskylä, 2017), and reduced in a shorter span of time. It is known that this can be modified by some factors such as health habits, among others (Fried, 2000).

In this sense, different approaches have been carried out in the last years to promote wellbeing and ensure healthy lives for our elderly. Among them, the most studied approach is probably the use of recommender systems, which have proven to be

---


In Proceedings of the 9th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2023), pages 118-125
ISBN: 978-989-758-645-3; ISSN: 2184-4984
Copyright © 2023 by SCITEPRESS – Science and Technology Publications, Lda. Under CC license (CC BY-NC-ND 4.0)
useful for this purpose (Ceron-Rios et al., 2017), being a well-functioning form of offering tailored health interventions (Azmi et al., 2019b; Park et al., 2012; Sezgin & Özkan, 2013). Although some of these systems are already in use, they present several limitations regarding i) the need of relevant validation studies to demonstrate their usability (Azmi et al., 2019a; Martinho et al., 2019), ii) interoperability and the use of standards for communication with other clinical platforms (Hors-Fraile et al., 2018), iii) the need of continuously updating knowledge base (Berner & Lande, 2016), or iv) user-based personalization (Rist et al., 2018). In view of all this, our approach attempts to address these limitations by presenting an interoperable, technically validated, and multidisciplinary tool that provides personalized recommendations to elderly users. The system is technically validated using stochastic, synthetic data, which allows to assess the potential benefits of the presented tool in the target population, the elderly.

This paper presents the methodology used for the development of our Wellbeing Recommender System (WRS) and the assessment of its potential to aid users, describing its components and functionalities, and showing the steps followed for its technical validation with synthetic data. In addition, the results and conclusions of the approach are also presented.

2 MATERIALS AND METHODS

In this section, the methodology followed for the generation of the dataset used for the validation of the WRS can be found (Section 2.1), the creation and management of the recommendations is described (Section 2.2.1), the data evaluation process by a rule engine is detailed (Section 2.2.2), the modelling of the data used by the platform is explained (Section 2.2.3) and lastly the communication protocols and standards of the platform are detailed (Section 2.2.4. All these principal components and the general workflow of the WRS are represented in the next Figure 1.

2.1 Synthetic Data Generation

In this study, we tested the WRS using a stochastic, synthetic database named "Synthetic Database for Recommendation System 2022 (SDRS2022)" which simulates various user types. In subsection 2.1.1 the methodology followed to define the categories of users is explained, and in subsection 2.1.2 the definition of the metrics and values that were generated is described.

2.1.1 Elderly Population Characterization

The generated dataset simulates profiles of potential users of the WRS. With the purpose of obtaining a real representation of the future users, we based the data generation on the population groupings defined in the SHAPES project, whose aim is to build a large-scale platform EU standardized open platform for long-term active and healthy aging. Using the personas method (Pérez-Montoro, 2017) eight mutually exclusive types of personas were defined:

- **Active**: older adults between 65-75 years old with good health and an active lifestyle. They can be either retired or still working and have an active social life. Smoking, alcohol and caffeine consumption are closely related (Torres-Collado et al., 2018);
- **Chronic**: older adults from 65 on that suffers of multimorbid conditions such as diabetes and oncological disease. Even though their limitations, they try to maintain their autonomy and active life (Arnautovska et al., 2018);
- **Musculooskeletal**: older adults that suffer mobility difficulties. They have high risk of falls and the fear of falling limits their daily life activities. Sleep duration is associated with the musculoskeletal pain (Lavigne et al., 2011). They try to maintain their autonomy but there is a risk of social isolation (Auais et al., 2018);
- **Neurodegenerative**: lonely elders with memory decline, isolated from society and with need of homecare (McCabe et al., 2014). Past high alcohol consumption habits are related with the prevalence of neurodegenerative diseases (Kamal et al., 2020). Sleep patterns become altered due to the neurodegenerative process (Owen & Veasey, 2020);

Figure 1: Data execution workflow using the WRS.
Lonely: older adults socially isolated, with no support that needs homecare assistance. Through physical activity, they can reduce their loneliness feeling (Pels & Kleinert, 2016), which in turn is closely related to their sleep quality (Jia & Yuan, 2020);

Drug dependency: older adults that have high alcohol consumption habits. They have a high percentage of hospitalizations and hospital visits (Choi et al., 2015), and it is considered that they can maintain a controlled autonomy. Abuse of alcohol severely affects sleep quality (Devenney et al., 2019);

Fragility: older adults over 85 years old with high fragility suffering from falls, exhaustion and weight loss (Scheibl et al., 2019). They are highly dependent and need of a professional or informal caregiver;

Deafblind: elders characterized by their difficulties in socialising. They have a minimal dependency condition (Bodsworth et al., 2011). Due to their disability, can develop high alcohol consumption habits (Fellinger et al., 2012).

Not only were these categories defined, but also how belonging to one of them influenced the metrics of the study. For this, a research was done in the literature to identify the variables and values needed to model the categories, as detailed in the following subsection.

2.1.2 Metrics for Synthetic Users

To create stochastic, synthetic data, metrics were used that represent the habits and lifestyle of the elderly population. The criteria for these metrics were based on literature related to physical activity, sleep, and liquid intake. For physical activity, the elderly were grouped into five age groups (65 to over 85 years old) based on the difference in expected daily steps (Tudor-Locke et al., 2013). For sleep data, the literature suggests a decline in sleep duration with aging (Faubel et al., 2009), and a range of less than 7 hours or more than 9 hours was used as a high-risk symptom or alteration in sleep duration, respectively. For liquid intake, the elderly tend to consume less water as they age and in Europe, liquid intake is not uniform due to differences in beverage categorization (EFSA Panel on Dietetic Products, Nutrition, and Allergies (NDA), 2010). The European Food Safety Authority (EFSA) analyzed reports to create reference values for liquid intake. The values were used to make reference values for water, alcohol, and caffeine-based beverages to monitor daily drinking behavior and anticipate potential health problems such as alcoholism and hypertension.

2.2 WRS Components

The WRS design is composed of the following modules: (i) the wellbeing recommendations manager to create and maintain the content of the different recommendations that the system returns, (ii) the wellbeing rule engine to generate the personalized recommendations by evaluating user’s data, (iii) a wellbeing ontology to keep the knowledge representation homogeneous across the platform and (iv) an interoperability module that transforms the input/output data following clinical standard communication protocols such as FHIR HL7 so that the system can be integrated with other platforms. In the next subsections, each component is described more in detail.

2.2.1 Wellbeing Recommendations Manager

The recommendations included in the WRS are not limited to a single domain; due to the possible use of different sources that can be used to gather data, the format used to store the recommendations in a digital format is domain independent.

In this paper, the recommendations included are divided into 3 groups, depending on the target that the recommendations are associated with. These categories are (i) liquid intake; (ii) physical activity and (iii) sleep. For each one, specific recommendations were defined based on agreed criteria published and accepted in the literature (see Section 2.1 for more details). Once identified, the wellbeing recommendations were modelled into several rules, which are the knowledge base used by the rule engine described in subsection 2.2.2.

These rules are defined in a domain independent format, so that the same representation format can be re-used in different domains, not only the three specified ones. Furthermore, the recommendations can be returned in different languages, thus broadening the potential user’s population that could benefit from this solution. The generated recommendations can be displayed to the end-users in different devices (i.e., users’ phone, computer, etc), as the communications to/from the WRS are managed via a REST API. As a result, if an organization plans to integrate a recommender system for its users, employees, etc, it is possible by sending the data in a HTTP request, as detailed in subsection 2.2.4, and then visualize the answer as needed, either in a web, or as notifications of a mobile application for
instance. Lastly, since recommendations may become obsolete, or new recommendations may want to be introduced into the system, a web-based rule authoring tool (AT) was integrated to manage the formalized content in a simple way by any user. This tool was developed in (Torres et al., 2020), and was used to first introduce the rules, and later edit them when needed. This is a need that is critical for the adoption of the system as the contents need to be updated with new evidence otherwise they will become outdated (Sim et al., 2001).

All the introduction of data was done using the aforementioned tool, which in conjunction with the ontology presented in subsection 2.2.3, eased the introduction of the knowledge into the system, while reducing the possibility of input errors of human origin. An example of the use of the AT for the formalization of rules is shown in Figure 2, where the definition of the conditions that conform are rule is done using the interface of the AT. Individual conditions or groups of conditions can be defined, as well as the relation between them. In a similar way, recommendations are introduced using the same interfaces.

2.2.2 Wellbeing Rule Engine

The number of rules formalized in the WRS can grow substantially over time as new content is added into its knowledge base. As a result of it, a correct management of the rules and its execution is crucial to ensure the scalability of the system. As mentioned before, the base of the WRS is the work presented in (Torres et al., 2020), where a business rule engine (Drools) is used to perform the mass execution of the rules that conform the knowledge base.

Once users’ raw data is received, it is transformed from the input FHIR message to internal JSON instances. For liquid intake data, measurements an aggregation by drink (i.e., water, beer, tea, juice, etc) and by the type of the drink (alcoholic beverages, carbonated drinks, caffeine beverages, etc). The use of business rule engines allows the platform to perform well, independently of the size of the knowledge base, guaranteeing the scalability of our system.

2.2.3 Data Model

The WRS system uses an ontology to keep its knowledge base homogenous and easy to access for new rule introduction. The ontology helps reduce errors in representing the knowledge of the protocols in rules and the information for both the conditions and recommendations of the rules is modeled in it. The ontology can be used by other solutions that need the same data, as it can be accessed through a REST API. Additionally, the information can be coded in different languages to better fit the user's sociolinguistic profile.

2.2.4 Interoperability

Interventions aiming to improve the habits of the users can employ one or more component/platform to support the patient during the intervention process. As a result of this, the need to interchange data between different systems becomes a necessary feature that enhances the interoperability of the developed platforms. In the present work, the issue of interoperability was addressed adopting the FHIR standard developed by HL7 for exchanging clinical data between systems. All the interchanges of data, both input and output requests, are done via FHIR resources. These resources are processed by a parser component that serializes/deserializes the data between the FHIR and the internal formats used to manage the users’ data.

Figure 2: Introduction of conditions using the integrated Authoring Tool.
Although the WRS is presented as a standalone solution, it is designed to be capable of working in conjunction with other components, as was the case in the ecosystem of digital solutions developed under the SHAPES project, where the interchange of clinical data between the different components of the ecosystem is done via, among others, FHIR resources.

## 3 RESULTS AND DISCUSSION

With the profiles defined with the persona methodology and the identified metrics described in Section 2.1, a dataset of 100 subjects between 65 and 90 years of age has been created, with the following characteristics (see Table 1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean and SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>78.24 ± 8.4</td>
</tr>
<tr>
<td>Gender (%)</td>
<td>53 (Female) 47 (Male)</td>
</tr>
<tr>
<td>Physical condition</td>
<td></td>
</tr>
<tr>
<td>14 (Active)</td>
<td></td>
</tr>
<tr>
<td>10 (Deafblind)</td>
<td></td>
</tr>
<tr>
<td>17 (Drug dependency)</td>
<td></td>
</tr>
<tr>
<td>12 (Musculoskeletal)</td>
<td></td>
</tr>
<tr>
<td>10 (Lonely)</td>
<td></td>
</tr>
<tr>
<td>14 (Fragility)</td>
<td></td>
</tr>
<tr>
<td>16 (Chronic)</td>
<td></td>
</tr>
<tr>
<td>7 (Neurodegenerative)</td>
<td></td>
</tr>
<tr>
<td>Steps</td>
<td>4689.14 ± 8434.1</td>
</tr>
<tr>
<td>Sleep (hours)</td>
<td>6.08 ± 2.2</td>
</tr>
<tr>
<td>Water intake (ml)</td>
<td>1205.07 ± 1012.3</td>
</tr>
<tr>
<td>Alcohol intake (ml)</td>
<td>386.58 ± 195.3</td>
</tr>
<tr>
<td>Caffeine intake (ml)</td>
<td>933.48 ± 358.5</td>
</tr>
</tbody>
</table>

The process of validation consisted of the evaluation of the data by the WRS system for each patient, the storage of the generated recommendations, and lastly, an analysis of the correlation between the different profiles of the generated users and the type of the wellbeing recommendations obtained.

As a result of the evaluation of the SDRS2022 dataset, a total of 523 recommendations were obtained. This showed at first that the system was capable of dealing with high amounts of data in a reduced amount of time, but it was not easy to check whether the recommendations obtained were related with the different patient profiles. An aggregation of the recommendations based on the profile of the user that generated them was performed, grouping first the recommendations belonging to each of the profiles described in Section 2.1, followed by a categorization of the rules depending on the nature of its recommendation. The categorization was different for each of the wellbeing recommendation categories. In the case of the sleep recommendations, they were grouped depending on the number of hours slept. If the patient slept the recommended hours, it was considered as a sleep notification, if it was not the case, it was considered a warning, differentiating between excess (sleep excess warning), and lack of slept hours (sleep lack warning). A similar categorization was done with the activity recommendations, differentiating between recommendations when the end-user reached activity_notification, and did not reach its daily step goal activity_warning. Lastly, in the liquid intake case, the recommendations were grouped depending on the class of drink, that is, recommendations including alcoholic beverages were grouped under alcohol warnings, and recommendations of caffeine/theine-containing drinks such as coffee, tea or carbonated drinks were grouped as caffeine warnings. The obtained results regarding the recommendation type and patient profile can be seen in Figure 3.

The recommendations of the system reflected the main characteristics of the user profiles, as the highest amount of alcohol-related recommendations took place in the Drug dependency category which also presented over sleep disturbances and low physical activity as described in Section 2.1. Similarly, the Frailty users were not able to meet the daily step target, which can be correlated with their frail physical condition.
Other observed behaviours match Section 2.1. Sleep disturbances are seen in the Neurodegenerative group, high physical activity disturbances in Chronic, and poor sleep quality in Lonely. Active, Neurodegenerative, and Lonely groups had high alcohol consumption. Chronic’s low average recommendations could be due to heterogeneity, and similar recommendations for Active, Neurodegenerative, and Lonely groups suggest a need to improve profile modelling to prevent overlap.

Although results show that the recommendations are in line with the habits described in Section 2.1, a more real categorization of the patients can be done by not limiting them to just one category, resulting in a more accurate approximation of the WRS value.

The technical performance of the system was evaluated by measuring the time it took to process the SDRS2022 dataset (parse the FHIR observations, evaluate data, and return the recommendations in FHIR format). The evaluation was done using a PC with 16GB of RAM and an Intel i5-9400F processor and the results for 100 users are presented in Table 2.

<table>
<thead>
<tr>
<th>Table 2: Technical characteristics of the evaluation process.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
</tr>
<tr>
<td>Total observations</td>
</tr>
<tr>
<td>Number of generated recommendations</td>
</tr>
<tr>
<td>Recommendations/user</td>
</tr>
<tr>
<td>Evaluation time (ms)</td>
</tr>
</tbody>
</table>

4 CONCLUSIONS

In this paper we present a rule-based recommender system (WRS) capable of providing multidomain wellbeing recommendations for the elderly population. The aim of this system is to help users to follow healthy lifestyle habits that will help them improve their quality of life.

The WRS is integrated with a rule authoring tool (AT) that allows for updating its knowledge base with the latest evidence on wellbeing recommendations. The system is designed to maintain consistency across its knowledge base by incorporating an ontology that defines the variables necessary for defining wellbeing rules.

The potential impact of the formalized recommendations was evaluated by analyzing the triggered recommendations for each user profile. The analysis showed that the formalized rules are useful.
in providing information about users’ habits. However, the dataset used can be improved by not limiting users to a single category and by improving the methods of data generation to simulate more realistic users.

Future research will develop a methodology to analyze behavioral changes over time. The WRS conforms to clinical standards, and new features can be added easily to improve the impact on the end-user’s life. The generation of synthetic data will also be enhanced to simulate users’ habits and enable multi-category inclusion. This will allow to test the methodology for detecting changes in behavior.

ACKNOWLEDGEMENTS

This project has received funding and clinical professionals’ advice by gewi-Institut für Gesundheitswirtschaft e.V. under the European Union’s Horizon 2020 research and innovation programme under Grant Agreement No 857159. The funding sources had no involvement in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication. The study complies with the current laws of Spain and Europe. All authors declare that they have no competing interests.

REFERENCES


