Reference Design Model for a Smart e-Coach Recommendation System for Lifestyle Support Based on ICT Technologies

Ayan Chatterjee, Martin W. Gerdes, Andreas Prinz, Santiago G. Martinez, Anine Christine Medin

Faculty of Engineering and Science
Faculty of Health and Sports Sciences
University of Agder
Grimstad, Norway

e-mails: {ayan.chatterjee, martin.gerdes, andreas.prinz, santiago.martinez, anine.medin}@uia.no

Abstract—As acknowledged by the World Health Organization (WHO), the demographic development shows that by 2030, eight out of ten foremost causes of death will be connected to risk conditions of lifestyle diseases, regardless of gender. Chronic illness associated with modifiable lifestyle factors will be accountable for the highest death worldwide. Health behavior change should be given priority to avoid serious damages. An Electronic Coach (e-Coach) system can empower people to achieve a healthy lifestyle with early risk predictions and appropriate tailored lifestyle recommendations. Research in e-Health has the potential to provide methods to improve personal healthcare with Information and Communication Technologies (ICT). An Electronic Health (e-Health) virtual coaching recommendation system can monitor people and convey the appropriate recommendations to improve their lifestyle and prevent against non-communicable or lifestyle diseases. This paper addresses the potential of selected emerging information and communication technologies to make e-Health systems smarter, more collaborative, and more efficient. As a result of the analysis, a reference design is discussed here to develop and validate the performance of a smart e-Coach system utilizing ICT, Internet of Things (IoT), and Artificial Intelligence (AI) to provide individual lifestyle recommendations aiming at a healthier lifestyle to prevent obesity and overweight. The healthcare sector is still looking for collaborative, user-friendly, optimized, cost-effective, reliable systems for health e-Coaching with automatic, meaningful, empirical evidence-based, and personalized lifestyle recommendations. The focus is on this system’s implementation and validation, considering obesity and overweight as a case study.

Keywords- e-Coach; e-Health; AI; IoT; Human-centered design; ICT; Lifestyle diseases; Recommendation; Decision support; Data security and privacy; Ethics; Ontology.

I. INTRODUCTION

Globalization, rapid urbanization, poor dietary habits, and a sedentary lifestyle negatively impact people’s health across the globe, affecting people from different socioeconomic groups. Key risk factors are excessive use of alcohol, inappropriate nutrition, physical inactivity, surplus consumption of salt, intake of saturated fat, smoking, drinking of excess sugar-containing foods, and beverages. All these factors can contribute extra weight gain, elevated blood glucose level, high blood pressure, higher total cholesterol in the blood, followed by different physiological impacts such as obesity, cardiovascular disease, hypertension, and diabetes type II, as illustrated in Figure 1. Non-communicable or lifestyle diseases are the most common cause of death worldwide. It is the leading unconditional probability of dying between ages 30 and 70 years [12]. ‘Obesity and overweight’ is one of the major lifestyle diseases that lead to other health conditions, such as Cardiovascular Diseases (CVDs), Chronic Obstructive Pulmonary Diseases (COPDs), cancer, diabetes type II, hypertension, and depression. It is a nutritional disorder and represents an excess amount of energy storage as mass in the body.

e-Health monitoring has become increasingly popular providing ICT-based remote and timely care support to its patients and healthcare providers. A health e-Coach system is an interactive, secure, monitoring, and caregiving system to produce automatic, meaningful, empirical evidence-based and personalized lifestyle recommendations to attain personal wellness goals. The goals are achieved by maintaining a balance in social life, physical activities (such as sleeping, exercising), eating, and daily stress level. The referred reference e-Coach model will capture personal, physiological, activity (sleep pattern, exercise), nutrition (fruits, vegetables, salads, energy drinks, and alcohol) and contextual data (spatial, temporal and weather) from secure wearable bluetooth enabled devices, manual interactions, feedback, and customized questionnaires over time. The collected time-series data will train machine learning models for behavior analysis, early prediction of wellness trends and risks.

Once the health risk is predicted, the e-Coach system will produce lifestyle recommendations to target individuals to improve their well-being, health, and physical performance by determining and recommending individual lifestyle changes. Emerging e-Coach systems reveal certain limitations in terms of technologies, quality, and performance. To overcome those limitations, an integrated and collaborative e-Coach service is described, addressing the cooperation of secure cloud-based fitness and well-being services with the public health and the care infrastructure.

e-Health research helps us to solve e-Coach associated challenges, such as activity and wellness status determination, an ontology for e-Health information, the selection of an e-Coach strategy, the generation of automated, meaningful, empirical, evidence-based and personalized lifestyle recommendations and their automatic assessment compared to a set goal. This study is deeply focusing on what to coach and how to coach, including the
design, development, testing, and evaluation of observational evidence-based, context-specific, and individual lifestyle recommendations. Here, we have highlighted a reference design model for e-Coach and recommendation generation for obesity and overweight, and we are currently working on its implementation and the validation phase.

A high-level reference design is presented in this paper with special consideration of upcoming technologies, such as data ontologies from Semantic Sensor Networks (SSN), IoT, AI, Decision Support Systems (DSS), big data, data mining, and Human-Computer Interaction (HCI). The objective of the considered health e-Coach system, as illustrated in Figure 5, can be summarized as follows: a) improve personal healthcare with the help of ICT; b) deliver high quality, evidence-based, secure, cost-effective, timely care to support people for sustaining a healthy lifestyle; c) predict the health risks and generation of risk alerts; d) ensure a rule for personal data security, portability, and remote access; e) normalize the heterogeneous format of personalized data with appropriate e-Health ontology model; f) analyze human psychology, human behavior and; g) follow defined medical guidelines for the generation of lifestyle recommendations.

The remainder of this paper is structured as follows. In Section II, we summarize the related work and highlight the differences between our current and existing work. Section III presents health e-Coach design factors. In Section IV, we describe an e-Coach reference model for obesity and overweight. The paper is concluded in Section V.

II. RELATED WORK

Different projects have been conducted by different study groups on health e-Coach to generate a personalized recommendation plan for a healthy lifestyle. The idea is to give remote care to participants with a suggestion for a healthy diet and fitness plan to prevent obesity and obesity-associated non-communicable diseases. Behavior and health are strongly linked. Healthy eating and regular physical activity can lead to a healthy lifestyle [1]. Having proper e-Coaching recommendation plan may help people to accomplish health goals to maintain healthy behavior. Research projects have addressed the following two types of recommendations approaching obesity:

(a.) food-based recommendations where a Machine Learning (ML) algorithm with the experience sampling method has been used to develop an interactive support system for dietary recommendations. Sensors are used to collect data related to state (emotion, social activity), context (e.g., location, time, and weather), and collected data are sent to a secured storage over the IoT infrastructure. After that, a decision support system is developed to predict human behavior and generate timely feedback for activity. It uses a classification algorithm (decision tree approach) to warn people against unhealthy food, and a clustering algorithm (Hierarchical Agglomerative) for profiling. It generates feedback (passive feedback or cognitive behavioral therapy or adaptive or active) with adaptive messages (intervention) [4], including proper food suggestions, alerts for allergic or harmful substances, recommendations for relaxation and yoga [3]. The solution for feedback is either Short Message Service (SMS) based, web-based, mobile application based, or combined [4].

(b.) physical activity-based recommendations, where a tracker has been maintained for daily step count, metabolic equivalent of tasks, kilocalories, and distance to reduce sedentary behavior. A healthy diet, enough physical activity to ensure energy balance, not smoking, and limited alcohol are four key factors to prevent obesity significantly. Data has been captured over time and studied (data processing and transformation, such as cleaning, formatting, pre-processing) with ML algorithms to give feedback whether the target is achieved or not. Based on the result, the personalized and predictive model recommends changing personal behavior, daily routine, or diet plan [5].

Obesity may lead to social isolation or loneliness, or depression [6], that may intensify the chance of premature death. Demisris et al. [7] proposed a theoretical framework for assessment and visualization of older adults’ wellness. It is an enhanced screening platform for wellness and a telehealth component captures vital signs and a customized questionnaire. The system also incorporates a gait analysis component. Based on the analysis of multiple health and
wellness parameters, the system aims to identify early trends or patterns to help in health risk prediction for overweight people, and recommends a wellness solution with technological tools and techniques. The wellness parameters of the system are physical well-being, social well-being, and spiritual well-being. Gerdes et al. [2] proposed a holistic concept of a coaching system for the Conceptualization of a Personalized e-Coach for Wellness Promotion. The paper describes a Continuous Process of Wellness Management with a gap to complete the cycle for continuous process. The gap is closed with a case study of obesity and overweight, and the reference design model discussed here.

The identified gaps in current health e-Coach studies are summarized as a lack of a.) context-specific empirical, tailored recommendation; b.) medical guidelines to follow to generate lifestyle recommendations; c.) personal, health, and wellness data protection plan (security and accessibility); d.) consent to recruit participants in the study as a part of fair data economy; e.) human-centric designing (co-design, user experience, safety, cost); f.) credibility analysis; g.) complexity analysis; h.) methodology, guidelines, protocols, i.) analysis of efficacy, sustainability, maturity, j.) proper research (legal mobile apps and their data protection policy, ligal body sensors, or actuators to collect important physiological data and storing them at unsafe storage); k.) post-research participant’s data security plan.

Our study is focused on mitigating the gaps, as described above, and transform distributed, heterogenous health and wellness sensor data into meaningful information to build a machine learning model for health risk prediction and to generate an effective individualised recommendation in case of obesity, as well as to turn it into a behavioural motivation for an effective human-eCoach-interaction. The concept of the Reference System for Telehealth and Telecare services and System Architecture for Future Telehealth and Telecare Services [8] has been extended here to design the health e-Coach system to guide obese and overweight people.

III. HEALTH E-COACH DESIGN FACTORS

In this section, we first describe the essential factors to be considered while designing, developing, testing and validating the performance of a health e-Coach system to achieve the research goal.

A. Perspective

The reference digital health e-Coach system will capture personal, physiological (blood pressure, heart rate, blood cholesterol level, blood glucose, height, weight), contextual (time, location, environment, ocial), and behavioral (physical activity, nutrition, sleep) data over time from medically approved secure wearable sensors, customized questionnaires, manual interaction (interviewing), mobile apps, and feedback forms. Data will be stored in a secure storage and will be utilized to train a machine learning model for behavior analysis and the early prediction of wellness trends and risks, as illustrated in Figure 2. Once wellness trends and health risks are predicted, the e-Coach system will deliver corresponding alerts, suggestions, and lifestyle recommendations to its target individuals to improve their well-being, health, and physical performance by determining and recommending optimal individual lifestyle factors (example - healthy diet, regular exercise). The perspectives of the discussed “health e-Coach” system here are to deliver lifestyle recommendations to prevent obesity and overweight, and, thereby, help participants to lead a healthier lifestyle to reduce the risk of chronic illness.

B. Approach

We are going to follow the well-established design-science-research methodology [9] to system design, development, and evaluation of the e-Coach system. The overall tasks can be summarized as follows: (1.) systematic literature review; (2.) design and development (feasibility study, data collection, development of a recommendation engine, e-Coaching through interaction); (3.) trial run; and (4.) model evaluation with performance parameters (goal achievement). The feasibility study includes (a.) what to measure? [independent variables (e.g., demographics) and dependent variables (e.g., BMI)]; (b.) how to measure? [spending time to explore and find the right instruments and methods]; (c.) type of data to be collected [personal, physiological, contextual and behavioral] and determination of data collection process; (d.) policy for the recruitment of the participants; (e.) preparation of consent form; (f.) data security and privacy - General Data Protection Regulation (GDPR); (g.) usage of a standard framework for e-Health lifestyle recommendation and determination of health e-Couch efficacy.

C. Data Collection

Personal, physiological, contextual, nutrition, and activity related data are most important in this study to determine the wellness status and we are going to capture a selected set of data in Table I from participants with minimal effort. The frequency of data collection (personal, physiological, contextual, activity, nutrition), the specific target group to be monitored, modelling of an SSN ontology, the appropriate list of sensors and assessment tools, the development of a gateway for data collection, storing of clean, normalized data in the storage (as data volume will be high), and ensuring its security are important. Management of data from continuous observations, management of incomplete data is in focus, as illustrated in Figure 3.

After the collection of necessary data, training, and cross-validation of the machine learning model with appropriate features for health risk prediction are important. The selection of proper machine learning and deep learning algorithms for time-series analysis of collected health and wellness data and their comparative study concerning accuracy score, classification report, the confusion matrix are all important and need research focus for generating an effective recommendation plan. Once risk prediction is made, it is necessary to deliver alerts, recommendations, and suggestions to the participants so that they can follow them. This requires interactive human-computer interaction, building trust about the system, integrating offline (human) coaching psychology to the e-Coach, selection of an effective recommendation plan, and improving it further.
### TABLE I. OVERVIEW OF TYPE OF DATA TO BE COLLECTED

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal Data</strong></td>
<td>Age, gender, mobile, email, postcode, income, education, social participation, medical history, food preferences, habit, waist to hip ratio, smartphone app usage for behavioral tracking, internet connectivity</td>
</tr>
<tr>
<td><strong>Physiological Data</strong></td>
<td>Blood pressure, heart rate, blood glucose, total blood cholesterol, weight, height</td>
</tr>
<tr>
<td><strong>Contextual Data</strong></td>
<td>Temporal and spatial data, weather data</td>
</tr>
<tr>
<td><strong>Activity Data</strong></td>
<td>Raw acceleration, energy expenditure, metabolic rate, physical activity intensity, total steps, activity bouts, sedentary bouts, sleep latency, total sleep time, sleep efficiency, duration of activity and non-activity</td>
</tr>
<tr>
<td><strong>Nutrition Data</strong></td>
<td>Intake of energy drinks, alcoholic beverages, intake of core foods such as vegetables or fruits, intake of discretionary foods such as unhealthy “junk foods.”</td>
</tr>
</tbody>
</table>

---

D. Secure Wearable Sensors and Data Security

Physiological and activity data are generally collected via secure wearable devices. The wearable activity monitors are connected with a smartphone or tab or computer via Bluetooth nearfield communication technology. Participants will install a developed mobile application that will transfer the collected activity and physiological data to a private secured storage, in compliance with GDPR or similar. Though different wearable health sensor devices are available on the market, we need to choose only legal devices. Wearable sensor devices should have regulated access control policy, approval by the authorized medical board (example, Food & Drug Administration (FDA), US) to collect, process, and store health related personal data.

In our study, collected personal, physiological, activity, nutrition, and contextual data, as described in Table I, will be processed securely during transmission using encryption, and there will be no possibility of information leakage. Transport Layer Security (TLS), a Virtual Private Network (VPN), authentication (form based, active directory, API key), and authorization (access control) will be used in our model to secure data during transmission and during storage. Data will be stored at a secure storage with Fast Healthcare Interoperability Resources (FHIR) standard and data will be accessed with HTTP based REST API following an API key authentication.

---

Figure 2. e-Coach overview of obesity and overweight.

A dedicated platform will be used for collecting, storing, analyzing, and sharing sensitive data in compliance with privacy regulations (GDPR). The API key, access-control list, two-layer authentication with form based authentication (Unique User-id (UUID) or username and password), and active directory security or RSA key are appreciated to protect the system from illegitimate users.

Linking between unique user-id and personal identifiers (such as name/email/phone/postcode) should be stored in a secure environment, and there will be no clue to perform any cross-identification or backtrack legal formulation of any participant. In line with the participant’s consent or request, information about the participants must be anonymized by removing personal identifiers or deleted after the successful deployment of the project.

---

E. Knowledge Presentation

Integrating raw, unstructured, high volume data from heterogeneous sources like wearable sensors, questionnaires, mobile applications, interviewing, and creating a compact, machine-understandable, structured, meaningful information from it, is challenging. But, it is important to extract knowledge from data in order to convey tailored lifestyle recommendations.
Semantic web is a concept and a group of technologies that support the development of machine-interpretable data and allows machines to process, explore, and analyze web documents effectively. It has offered an universal framework for describing data with common data formats, specifying data about data with semantic web languages URI, XML/ XML schema, RDF, RDFS, OWL, RIF/ SWRL rules, SPARQL, and assisting better knowledge representation, advanced access, formal analysis of resources, uphold standardization, increase interoperability among heterogeneous sources and networks, data integration, data discovery, and situation awareness. The semantic sensor network incubator group (SSN-XG) has developed an SSN ontology to model sensor devices, systems, and processes observations and to allow its network, sensors, and data to be organized, installed, managed, queried, controlled through high-level specification. It is a domain-independent ontology [10].

OWL (defines what to write) processes web information when RDF (defines how to write with triple-store-resource, property and value) is used for describing. RDFS, which is an extension to RDF, provides modeling and structured vocabularies for RDF data. OWL (turtle syntax) with SPARQL API (a query engine for querying and updating RDF models using the SPARQL standards [13]. SPARQL is a query language and a protocol for accessing RDF designed by the W3C RDF data access working group. SPARQL queries RDF graphs., Protégé (an open-source ontology development tool with OWL support), Apache Jena (a free and open source Java framework for building semantic web, linked data applications and extract knowledge from ontology), SSN Ontology, which gives sensor descriptions and observations, and SNOMED-CT (a systemized nomenclature of medicine since 1965 and designed as an ontology and it is an organized list of a wide variety of clinical terminology defined with unique codes namely ICD) have all been used in our research to model the health e-Coach ontology for the case study: obesity and overweight and the same ontology modelling is beyond the scope of this paper.

F. Human-Centric Design

In our research, we will use human-centric approaches to design and develop the desired health e-Coach system and make it more usable, utilizing human factors or ergonomics and usability learning and skills [ISO 9241-210 - Ergonomics of Human-Centered System Interaction]. A solid user-centered design that will be used has been discussed here. It follows a cyclic process of 1. plan 2. analyze 3. design and 4. test and evaluate.

Accomplishing user-centered design is very challenging, and it needs a comprehensive process, beginning with planning, analyzing and closing with testing and improving it further. Such a system can only be accomplished by putting the user experience at the heart of the development process. It is a process that places human needs and limitations at a higher priority compared with other targets during design thinking and implementation. To achieve a successful human-centric design, designers should focus on the following [10]: a.) co-design; b.) define the problem that needs to be solved well; c.) understand user characteristics and needs; d.) perform research on user experience and understand user behavior and impressions; e.) the developed prototypes must be tested and validated by real users to ensure smooth user experience, design consistency, reduction of learning effort, and building trust; f.) a continuous communication between user and the designer/design to gain feedback to ensure further design improvements; g.) the design should be shaped in the user’s cognitive knowledge; h.) empathic system experience and according to system development that users love to use daily.

We envision to integrate elements of the human process of coaching in our research, such as feedback/ rating, preference sharing, user-friendly human-computer interaction, intelligent interactive session, problem solving, timely alert (SMS/e-Mail), wellness vision, motivational suggestions, encouragement, assessing human psychology and sentiment based on physiological and contextual data, educating people to make the human-e-Coach-interaction effective for obesity recommendations. In human centered design, all the human factors, social factors, and technology factors intermingle together under the human activity umbrella. The human centered design process will flow in between three main phases: research, prototype, and implement keeping human at the heart of the design.

G. Requirement Framework

The subsequent technical requirement domains have to be considered in the design, development, and evaluation of the discussed health e-Coach system and its supporting infrastructure components are as follows (Table II):

<table>
<thead>
<tr>
<th>TABLE II. REQUIREMENT FRAMEWORK FOR E-COACH SYSTEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasibility</td>
</tr>
<tr>
<td>Scalability</td>
</tr>
<tr>
<td>Data Governance</td>
</tr>
<tr>
<td>Interoperability</td>
</tr>
<tr>
<td>Usability</td>
</tr>
<tr>
<td>Reliability, Robustness, Availability</td>
</tr>
<tr>
<td>Training</td>
</tr>
</tbody>
</table>

H. Goal

The goals and focus of the discussed health e-Coach system can be summarized as follows – a.) encourage co-design; b.) creation of a compact, intelligible abstraction
from massive, raw, unstructured observations for health and wellness data using e-Health ontology; c.) protection against illegitimate access to the system and personal health records; d.) collection of personal, physiological, activity, and nutrition data to build a machine learning model for health risk prediction and to understand if there is any negative change in behavior pattern; e.) model automatic, meaningful, and empirical evidence-based and personalized lifestyle recommendations to achieve health wellness goals. Continuous monitoring of the willing participants and evaluation of time-series monitoring data will produce detail knowledge about the individual relation between nutrition, activity, and Basal Metabolic Rate (BMI). Based on the obtained knowledge, the lifestyle recommendations, as depicted in Figure 4, with regards to nutrition and activity, will be given to the participants to avoid obesity and overweight.

Figure 4. Recommendation plan for obesity and overweight.

IV. PROPOSED MODEL DESCRIPTION AND RESULT

The holistic obesity and overweight e-Coach model is based on the co-design initiative, and the participant is located at its core. As illustrated in Figure 5, the complete model is divided into seven components – 1.) formulation – This includes: systematic literature review to gain complete knowledge about obesity and overweight, its causing factors and associated results, feasibility study (technical and financial), defining what to coach and how to coach, identification of needed data to be collected, defining frequency of collection, and e-Health ontology design; 2.) selection of participants – This includes: advertisement with defined inclusion and exclusion criteria, consent preparation, consent signing, training, participant’s account creation, handover of wearable sensor devices; 3.) infrastructure – This includes: the needed cloud infrastructure for data collection (network elements, firewall, protocol, active directory authentication, VPN) and secure database storage; 4.) data collection – This includes the collection of physiological data, environmental contextual data, personal data, activity data, and nutrition data from participants via secure wearable sensors, manual interaction, interviews, smartphone apps, feedback, and customized questionnaires overtime at defined frequencies; 5.) decision support system – This is needed to train a machine learning model for behavior analysis, early prediction of wellness trends and risks; 6.) e-Coach and recommendation plan – The e-Coach system will help to model lifestyle recommendations to achieve health wellness goals and will advise about physical activity, nutrition, and other relevant factors for healthier lifestyle coaching. The e-Coach system integrates a proper recommendation plan that follows clinical guidelines; 7.) human e-Coach interaction – This means incorporating elements of the human process of coaching, such as feedback/ rating, preference sharing, user-friendly human-computer interaction, intelligent chat service, problem solving, timely alert (SMS/ e-Mail/ Voice), wellness vision (physical/ social/ emotional/ spiritual), motivational suggestions, encouragement, assessing human psychology and sentiment based on physiological and contextual data, educating people to make the human-e-Coach-interaction effective for lifestyle recommendations.

Health e-Coaching is a continuous process that collects health and wellness data from participants continuously via wearable clinically approved sensors, smartphone apps, questionnaire over time and process them with AI algorithms to predict health risk, behavioral pattern, change in habits, random variation in lifestyle over a week/ month/ day-by-day, and assessment of behavioral data to generate evidence-based, real-time, context-based lifestyle recommendations, suggestions, alerts and followed by, the evaluation of individual goal. The role of ICT experts (researchers/ developers/ technology experts) is to design, develop, test, and evaluate the performance of an e-Coach system in active collaboration with health experts (nutritionists/ nurses/ physicians/ psychologists). The health experts help to design and validate the rules for needed data collection, lifestyle recommendations, health risk prediction, and determination of vital physiological signs.

V. CONCLUSION

We are working on the assumed hypothesis that ‘an effective e-Coaching mechanism can prevent obesity and overweight with automatic generation of personalized lifestyle recommendation.’ The data collection and lifestyle recommendation process should neither create any side effects nor any major discomfort to the participants. Hence, researchers need to collect the maximum amount of personal, physiological, activity, nutrition data with a minimum amount of collection, and measurement overhead. The entire measurement process should not be burdensome and, hence, we are going to collect a listed set of data in Table I with the participant’s consent. Integrating offline human coaching psychology to an e-Coach, selection of an effective lifestyle recommendation plan, and improving it further to reach the
convergence criteria of overweight to normal weight, is a computationally hard problem. This reference design model can be extended to e-Coach people against other lifestyle diseases.

REFERENCES


[8] M. Gerdes, F. Reichert, J. P. Nytun, and R. Fensli, "Reference design for smart collaborative telehealth and telecare services based on iot technologies", International Conference on

Figure 5. Holistic obesity and overweight e-Coach model.