ITEA3

Personal Healthcare Empowerment

Deliverable 3.2

USER PROFILE MODELING

WP3 – Coaching Framework
T3.2 User profile (patient) modelling

Document Type	Document
Document Version	V0.7
Access Level	Confidential
Contractual Submission Date	M10
Actual Submission Date	M14
Editors	ISEP
Contributors	ISEP, EXPERIS, ARD

Document Contributors

Partner	Author	Role
ISEP	Diogo Martinho	Contributor
ISEP	Constantino Martins	Contributor
ISEP	Goreti Marreiros	Contributor
ISEP	Ana Almeida	Contributor
EXPERIS	Ester Sancho	Contributor
EXPERIS	Julia Gomez	Contributor
EXPERIS	Gema Maestro	Contributor
ARD GROUP	Bilge BAŞDEMİR YALÇIN	Reviewer
ARD GROUP	Rabia AKTAŞ	Reviewer

Document History

Date	Version	Editors	Status
25/10/2018	0.1	Diogo Martinho, Constantino Martins	Table of Content Initial Draft
31/10/2018	0.2	Diogo Martinho, Constantino Martins, Goreti Marreiros	Table of Content Update, Description of Each Section
16/11/2018	0.3	Diogo Martinho, Constantino Martins, Goreti Marreiros, Ana Almeida	User Modelling Definitions and State-of-the-Art
07/02/2019	0.4	Diogo Martinho, Constantino Martins, Goreti Marreiros, Ester Sancho, Julia Gomez, Gema Maestero	Proposed Model for PHE
10/02/2019	0.5	Diogo Martinho	Merged Version with all Contributions
28/02/2019	0.6	Diogo Martinho	Final Document Updates according to Reviewers Feedback
28702/2019	0.7	Diogo Martinho, Ester Sancho	Released

Table of Content

List of Figures

List of Tables

Glossary

Acronym	Meaning
API	Application Programing Inteface
AQI	Air Quality Index
BMI	Body Mass Index
CORD	Chronic Obstructive Respiratory Disease
CPAP	Use of Continuous Positive Airway Pressure
DDD	Domain Dependent Data
DID	Domain Independent Data
GPS	Global Positioning System
HW	Healthy Workplaces
NR	Nor Related
OSA	Obstructive Sleep Apnea
OWL	Web Ontology Language
PHE	Personal Health Empowerment
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
Y	Yes

EXECUTIVE SUMMARY

Study of the current State-of-the-Art on existing user characteristics (which can be differentiated according to Domain Dependent Data and Domain Independent Data) and techniques to model different user profiles (which can be both statistical and non-statistical), and proposal of a common user model for the PHE.

1 Introduction and Project Background

Rising costs of healthcare due to the ageing population and related increase of non-communicative diseases urges for finding ways to save expenses by diminishing the need for care and making the current care more efficient. At present, healthcare provision is reactive, and process driven, treating patients according to predefined pathways with limited possibilities to consider the individual needs or abilities. Health authorities and care providers are finally noticing the one resource that had remained unused – the person or patient him/herself! By starting with the primary need of the person – to be healthy – and including him/her into the process in an active role, new paradigms for care become possible. Significant cost reductions can be achieved by preventive solutions to help the person adopt a healthy lifestyle – thus reducing the number of patients – and by providing the person with tools to actively participate in the treatment when diseases do arise – thus decreasing the burden on care personnel.

The main goal of this proposal is to empower people to monitor and improve their health using personal data and technology assisted coaching.

Developments in technology have enabled the empowerment of people for self-care more than ever before. Smart phones and tablets and quantified self-style self-monitoring wellness devices are commonplace. Wellness oriented solutions often suffer from short-term use due to quickly diminishing interest from their users and from lack of possibilities to utilise them in conjunction with clinical healthcare treatments. Patients are left alone with their problems in between therapy or treatment, and the possibly collected personal data is left unused.

Innovations in the project are expected in;

1. analytics on heterogeneous personal health sources to provide insight in the relation between behaviour and health
2. methodologies to develop interactive, dynamic and personalised coaching programmes,
3. modularisation of a scalable coaching framework
4. innovative motivating approaches for long-term adherence
5. innovative business models for preventive.

The Personal Health Empowerment project aims to achieve significant cost reductions for preventive solutions to help the person adopt a healthy lifestyle and providing the person with tools to actively participate in the treatment when diseases do arise by empowering people to monitor and improve their health using personal data and digital coaching. As a result, these will be causing to reduce the number of patients and decrease the burden on care personnel.

The results of the project include:

- Innovative technologies for vital signs, activity and behaviour monitoring
- Personal health analytics and visualisation tools
- Methodology and tools for the development of interactive and dynamic coaching programs (content & functionality)
- A modular reference framework for coaching application development and deployment

- Motivating self-care applications
- Validated pilot with users in the target groups for lifestyle management
- Exploitation plans for partners including go-to-market plans with disruptive business models

The project innovations will have a large impact on healthcare provision in the future, providing both evidence and means to realise people-centric and preventive health care, and allow for cost-saving solutions with increased patient involvement. It will address societal challenges including ageing, rising dependency ratio, lifestyle-related diseases, and healthcare efficiency to provide care in a more personalised and efficient way.

1.1 Deliverable Scope and Objectives

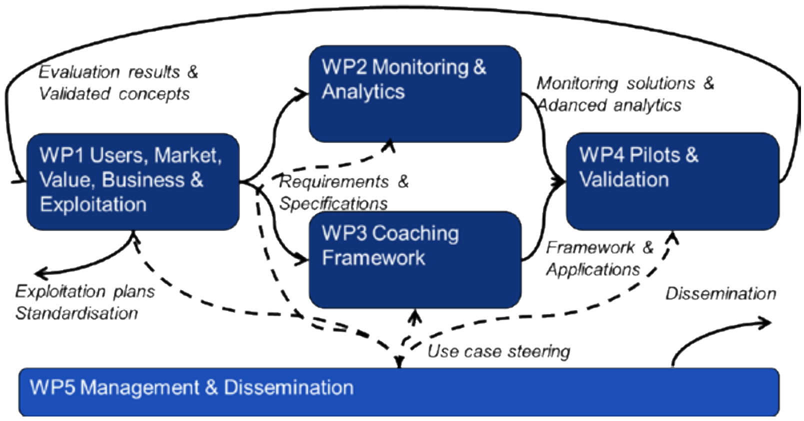

Figure 1 – PHE WPs Interconnections

2 User Modelling Definitions and State-of-the-Art

This section describes different definitions associated to User Modelling which includes main characteristics and techniques existing in the literature. In the end of the section, we look at existing proposals and the work developed over the last years related to User Modelling.

2.1 User Characteristics

A user model is composed by a set of characteristics that adjust the content, presentation and navigation to each user. These characteristics can be domain-dependent and domain-independent and are related with beliefs about the user, which include preferences, knowledge and attributes, or are an explicit representation of properties of individual users and user classes.

2.1.1 Domain Dependent Data (DDD)

Domain dependent data is related with system responses tailored according to the domain knowledge of a user [1]. For this, it is necessary to perceive user current state and knowledge regarding concepts and relations inherent to the domain, predict how the user will interpret system responses, understand the many different goals and plans of each user, predict and respond to different mistakes while the user is using the system and identify the most adequate way to present information to each user. Different methods can be used to measure user knowledge and expertise regarding the domain: Direct Dialogue and Indirect Acquisition.

2.1.1.1 Direct Dialogue

This type of interaction is performed directly with the user in order to assess his/her expertise regarding the domain. For this, the system should incorporate features to allow users to input and share their knowledge (for example, using questionnaires or forms) and mechanisms to process the inserted data to correctly measure user knowledge regarding the domain.

2.1.1.2 Indirect Acquisition

Indirect acquisition method allows the system to assess user knowledge indirectly according to how the user performs different actions. Depending on this assessment the user knowledge regarding the domain is classified in different levels which in turn are updated over time as the user works with the system.

2.1.2 Domain Independent Data (DID)

Domain independent data is not related with user expertise regarding the domain but to his/her cognitive abilities which indicates how the user perceives, thinks, remembers, behaves and solves different problems[1]. In other words, domain-independent knowledge corresponds to the phycological characteristics of the user. There are many different psychological models and tests that can be used to assess user personality such as the Myer-Briggs Type Indicator, the Eysenck's Pen Model and the Big Five Model.

2.1.2.1 Myer-Briggs Type Indicator

Myer-Brigg Type Indicator model [2] is a model used to identify personal characteristics and preferences. This model considers four different areas of personality based on the Carl Jung's Psychological Types [3] and which are perception, judgment, extraversion and orientation. These four areas combined result in sixteen different types and the scores on each dimension represent the strength of each dimension.

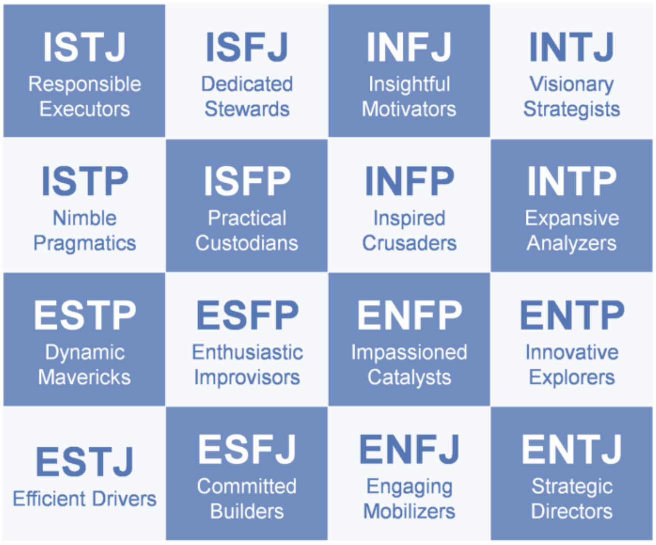

Figure 2 – Myer-Briggs Type Indicator

2.1.2.2 Eysenck's Pen Model

In 1950 [4], Eysenck proposed the PEN model using three dimensions to describe different personalities. These dimensions are: extraversion-introversion; Neuroticism versus Emotional

Stability; and psychoticism versus impulse control. According to Eysenck, individuals with high levels of extraversion are more social, talkative and outgoing, while individuals with high levels of introversion are more quiet, shy and less social. Individuals with high levels of neuroticism experience more stress and anxiety, while individuals with low levels of neuroticism experience more stable emotional levels. Individuals with high levels of psychoticism are more likely to show impulsive, irresponsible and miscalculated behaviour while individuals with low levels of psychoticism tend to be more controlled and organized.

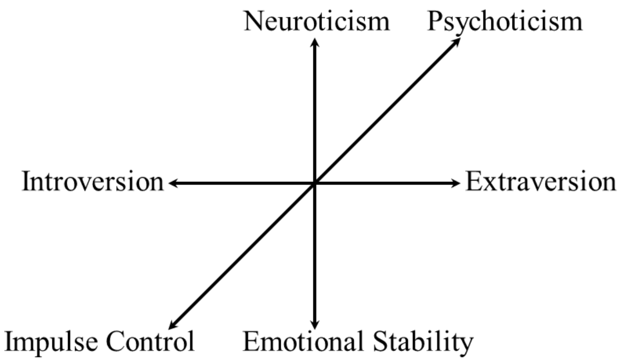

Figure 3 – Eysenck's PEN Model

2.1.2.3 The Big Five Model

The Big Five Model, also known as the OCEAN model has been proposed and developed over the last century by different researchers such as [5-8] and considers the existence of five main traits of personality which are extraversion, agreeableness, openness, conscientiousness, and neuroticism.

Openness – Trait associated to characteristics such as imagination and insight. People who have high openness tend to have a broad range of different interests about the world and other people and are willing to learn new things and enjoy new experiences.

Conscientiousness – Trait associated to characteristics such as thoughtfulness, good impulse control, and goal-directed behaviour. People who have high conscientiousness tend to be organized and mindful of details.

Extraversion – Trait associated to characteristics such as excitability, sociability, talkativeness, assertiveness, and emotional expressiveness. People who have high extraversion tend to be outgoing and value social interactions.

Agreeableness – Trait associated to characteristics such as trust, altruism, kindness, affection, and other prosocial behaviours. People who have high agreeableness tend to value cooperation.

Neuroticism – Trait associated to characteristics such as sadness, moodiness, and emotional instability. People who have high neuroticism tend to experience mood swings, anxiety, irritability, and sadness.

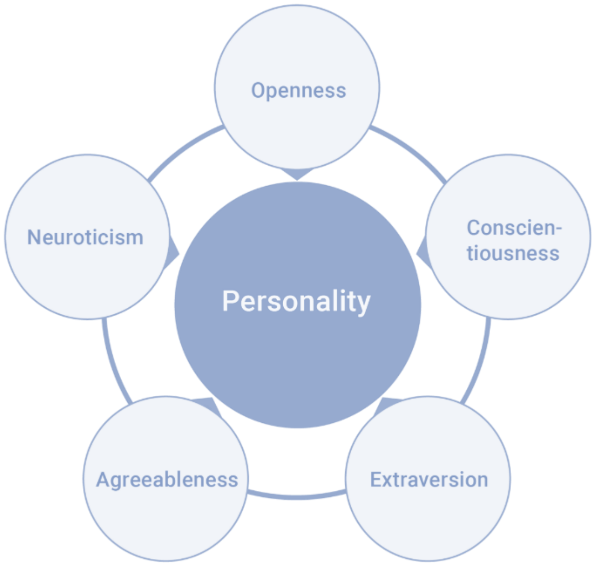

Figure 4 – Five Traits of Personality

2.1.3 Common Characteristics

In Table 1, it is presented common characteristics for User Modelling considering the definitions presented in previous sections.

Table 1 – Common Characteristics in User Modelling

Model	Profile	Characteristics	Descriptions/Examples
Domain Independent Data	**Generic Profile**	Personal Information	Name, email, password, etc.
		Demographic Data	Age, Gender, etc.
		Patient Background	Smoker, Pregnant, etc.
		Allergies	Allergies which the patient may have
		Deficiencies: visual or others	Sees well, uses eyeglasses, etc.
		Domain of application	Localization of the user, etc.
		Inheritance of the characteristics	Creation of stereotypes that allow to classify the user
		Knowledge (Background Knowledge)	A collection of knowledge translated in concepts. Possibility of a qualitative, quantitative or probabilistic indication of concepts and knowledge acquired for the user
	Psychological Profile	Cognitive Capacities	
		Traits of Personality	Psychological profile (introvert, extrovert, etc.)
		Personal Preferences	Likes and Dislikes
		Inheritance of characteristics	Creation of stereotypes that allow to classify the user
Domain Dependent Data		Objectives	Questionnaires to determine user objectives
		Complete description of the navigation	Kept register of each page accessed
		Knowledge acquired	A collection of knowledge translated in concepts.
		Medication Intake	Data related to patient intake of medication
		Context model	Data related with the environment of the user
		Aptitude	Definition of the capacity to use the system
		Task Preferences	Definition of the individual preferences with the objectives to adapt the navigation and contents

2.2 Techniques for User Modelling

After identifying the data related to each user characteristics, it is then possible to define the algorithms that will process this data and in turn affect the computational environment. These algorithms are mainly defined using statistical and non-statistical techniques.

2.2.1 Statistical Techniques

2.2.1.1 Linear Modelling

Linear Modelling is a technique which takes the weighted sum of known values and predicts the value of an unknown quantity [9]. These models are usually very inexpensive and easy to learn and understand. Furthermore, these models can be also extended and generalized without much effort. Two examples could be using a linear model to predict user's ratings of different activities suggested by the system or using linear model to assess the association between total cholesterol and body mass index.

2.2.1.2 Beta Distribution

The Beta Distribution is a predictive model which considers the number of correct predictions and the number of incorrect predictions and then generates both and estimate and a confidence level [10]. It is easy and cheap to calculate since it only requires two numbers (the number of hits and misses) to measure both estimate and confidence level. An example could be using a Beta Distribution model to track users' preferences by the number of likes and dislikes they provide to system for any suggested activity.

2.2.1.3 Markov Model

A Markov Model follows a structure very similar to a Linear Model and consists of a set of states, a set of probabilities which determine the likelihood of transition between these states and, for each state, a set of observation/probability pairs [9]. For example, a Markov Model could be used to predict user most frequent actions while using the system by looking at his past performed actions.

2.2.1.4 Bayesian Networks

A Bayesian Network is a directed acyclic graph where nodes denote variables and the arcs connecting nodes represent causal links from parent nodes to child nodes [9]. Each node is associated with a conditional probability distribution which assigns a probability to each possible value of this node for each combination of values of its parent nodes. These models are usually very flexible as they can provide a compact representation of any probability distribution, they can explicitly represent causal relations and they allow predictions regarding more than one variable (unlike many other statistical models which only considers a single variable). Examples of Bayesian Network models could be to predict the most adequate type of suggestions for a user according to the type of action being performed, or to predict error rates while the user is using the application.

2.2.1.5 Rule Induction Model

Rule Induction Model consists of learning sets of rules that predict the class of an observation from its attributes [9]. These models can represent rules directly or represent rules as decision trees or in terms of conditional probabilities. A rule itself is not considered a model and therefore, this type of models always considers a set of rules which collectively define a prediction model, or the knowledge base.

2.2.2 Non-Statistical Techniques

2.2.2.1 Overlay Model

An overlay model assumes that the user's knowledge is a subset of the domain knowledge. An overlay user model can thus be thought of as a template that is "laid over" the domain knowledge base. Domain concepts can then be marked as "known" or "not known" (or with some other method, such as an evidential scheme), reflecting beliefs inferred about the user. Overlay modelling is a very attractive technique because it is easy to implement and can be very effective. Unfortunately, the underlying assumption of an overlay model, that the user's knowledge is a subset of the domain knowledge of the system, is quite wrong. An overlay model cannot account for users who organize their knowledge of the domain in a structure different from that used in the domain model, nor can it account for misconceptions users may hold about knowledge in the knowledge base [11].

The overlay model consists of (a subset of) the concepts from the underlying domain model. For each concept, the overlay model contains data that represents (an estimation of) the individual user's knowledge about or interest in this concept (or some other relationship with this concept)[1].

In this method, the user knowledge is related, layer to layer, to the Domain Model, producing the user knowledge model (Figure 5) [12]. The expression of the knowledge level of each concept is dependent on the Domain Model itself: this value can be binary (knows or ignores), qualitative (good, average, weak, etc.) or quantitative (the probability of knowing or not, a real value between 0 and 1, etc.).

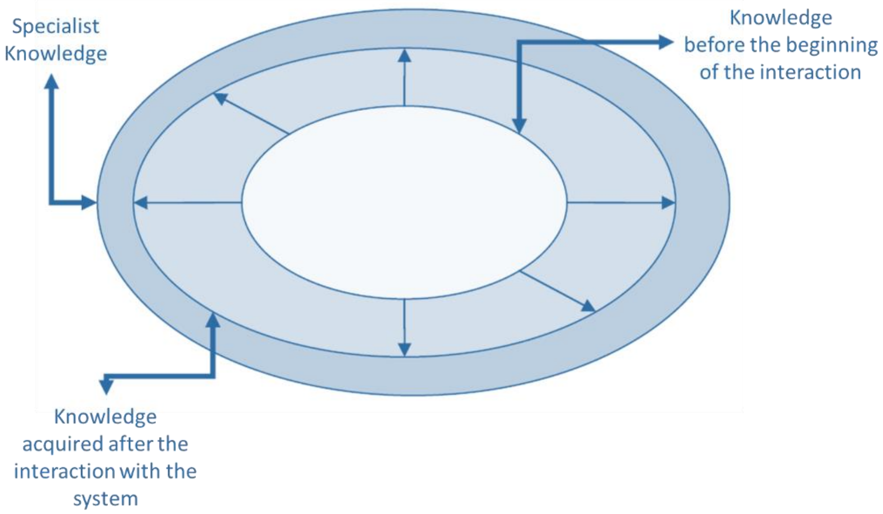

Figure 5 – Representation of the Overlay Model, adapted from [12]

[1] https://www.eelcoherder.com/images/teaching/usermodeling/03_user_modeling_techniques.pdf

2.2.2.2 Perturbation Model

The perturbation model can represent user beliefs that the overlay model cannot handle. A perturbation user model assumes that the beliefs held by the user are similar to the knowledge the system has, although the user may hold beliefs that differ from the system's in some areas. These differences in the user model can be viewed as perturbations of the knowledge in the domain knowledge base. Thus, the perturbation user model is still built with respect to the domain model but allows for some deviation in the structure of that knowledge [11].

Perturbation model represents learners as the subset of expert's knowledge plus their mal-knowledge [13].

This method considers that the knowledge and the student aptitudes are a perturbation of the specialist knowledge, and not a subset of his knowledge (as in the previous model) (Figure 6) [12]. This method can be used to represent knowledge that is beyond the Domain Model defined by the specialist.

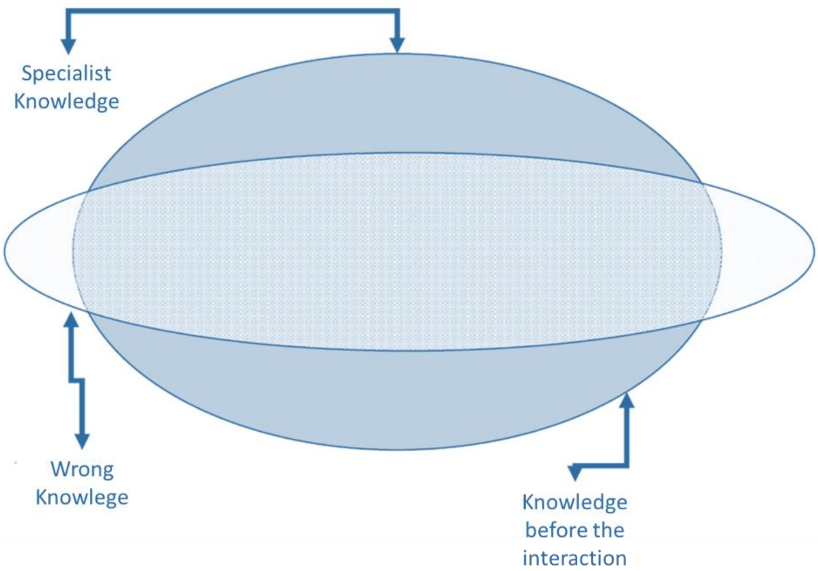

Figure 6 – Representation of the Perturbation Model, adapted from [12]

2.2.2.3 Knowledge Modelling

Process of creating a computer interpretable model of knowledge or standard specifications about a kind of process and/or about a kind of facility or product. The resulting knowledge model can only be computer interpretable when it is expressed in some knowledge representation language or data structure that enables the knowledge to be interpreted by software and to be stored in a database or data exchange file.

2.2.2.4 Behaviour-Based Model

A very common approach to gather requirements for developing a system is to interview and observe the behaviours of users from the intended user population. System design requirements typically characterize the user as one entity with a single set of behaviours, namely expert, novice, or a composite of all the users [14]. The goal of this type of models is to develop a system that can accommodates the great diversity of the user population and improve the users' performance. For this, system users can be categorized into different groups, and then it should be described and modelled each group's behaviours, and finally, this information should be included in both design and operational processes. Users can be categorized based upon similar behavioural characteristics that are important to system interface design and use. User modelling should then describe how users within a specific user group behave in certain situations or perform certain functions.

2.2.2.5 Rule-Based Model

Rule-Based Models can be automatically defined using learning algorithms to identify useful rules (also known as Rule-based Machine Learning Modelling) or can depend on expert-crafted knowledge bases to make inferences about users (traditional Rule-Based Modelling). Examples of this type of models could be using a Rule-Based Model to model user's current abilities, or to predict actions and errors performed by the user. Other examples include using a Rule-Based Model to identify irregular monitoring values captured by the application regarding current user health condition and alert the healthcare professional.

2.2.2.6 Stereotypes

One of the easiest and most common techniques for building models of other people is the evocation of stereotypes. Stereotypes were first introduced in the literature related to User Modelling by Elaine Rich in 1979 [15], and it was brought with the necessity to define a "useful mechanism for building models of individual users on the basis of a small amount of information about them". According to the author, in order to correctly define and use stereotypes it is necessary to collect and use two kinds of information. The first required information is related to the stereotypes themselves which includes the information of different collections of clusters of characteristics or facets. These facets depend on the domain and purpose of the system but may also include information related to the level of expertise while using the system or specific concepts and tasks dealt with by the system. These different facets will result and describe different groups of users. The second kind of information is related to the use of triggers which correspond to the occurrence of different events and that in turn will activate appropriate stereotypes. For example, if a user performs an advanced task while using the system, an "expert user" trigger could be activated. Table 2 shows an example on how to build different stereotypes related to practicing exercise and eating habits.

Table 2 – Stereotypes

	Trigger	**Cluster/Stereotype**	**People who like**	**Will like**
Direct Domain	$exercise$ $food$	ExercisePractitioner FoodEater FoodForExercise	$exercise_1$ $food_1$ $exercise_1$	$exercise_2$ $food_2$ $food_1$
Domain Attributes	$food$	FoodNutrient	$attribute_1$ in $food$	$food_2$
User Attributes	$user$	UserExerciser	**People who have**	**Will like**
			$attribute_2$	$exercise_2$

2.2.3 Ontologies

Nowadays, there is a great necessity to develop systems which can reuse and share knowledge and information for all sort of areas and applications including healthcare. To support such kind of systems, new tools are being developed, also known as Ontologies. On of the most common definitions comes from Gruber which refer to ontologies as "an explicit specification of a conceptualization" [16]. Although it seems a very simple definition, it is widely accepted in the Artificial Intelligence domain. To sum up, an ontology describes a data model, represents concepts and relationships existing in a certain domain. These relationships should allow inferring about all different instances related to the domain. The information represented by an ontology should include individuals (or instances), classes (concepts or types of instances), attributes (concepts' properties which can be mandatory or nor) and relationships (how concepts are related with each other). Some of the most used languages to define and instantiate ontologies are the RDF and RDFS [17] and OWL [18], with the last one being recognized as a standard by the W3C Consortium[2]. There are several advantages associated to the use of ontologies which are:

- Possibility to reuse existing ontologies, considering possible adaptations or extensions of knowledge bass which can promote a significant gain in terms of efforts and investments. Furthermore, this type of structure offers a great availability and possibility to be extended and complemented with concepts of different specific domains and to create an hierarchy/taxonomy.
- Easy access to ontological information, capacity to store thousands of examples, classes, attributes, relationships serving as an efficient search tool and preserve the integrity and share of knowledge between different communities while providing a uniform vocabulary.
- Use Linked Data practices, establishing a global association network between data and different domains.

2.2.3.1 User Ontologies

A user ontology classifies all the relevant characteristics and associated partitions of users into classes with corresponding associated information. In other words, a user ontology includes all the characteristics that can describe the user as a person [19]. Using sharable data structures containing user's features and preferences will enable personalized interactions with different devices for the benefit of the users [20]. A user ontology can be defined using OWL description language which contains the following elements: C – a set of concepts (entities and instances in user ontology); R – the relationship between classes or instances in the user ontology; I – a set of instances and A – a set of rules and restrictions [21]. Several works have been proposed in the literature regarding the definition and use of user ontologies. For example, in [22] it is proposed a Person Profile Ontology model which is responsible for modelling the profile of the user using five main classes: Person (can be either the assisted person, doctor, relative, etc,), Habit (daily activities performed by the assisted person), Impairment (visual, mobility, speech and other impairments associated to the assisted person), Contact Profile (email, phone number and other mechanisms to contact the assisted person) and Preference (preferences of the assisted person such as device preferences).

[2] The World Wide Web Consortium: https://www.w3.org/

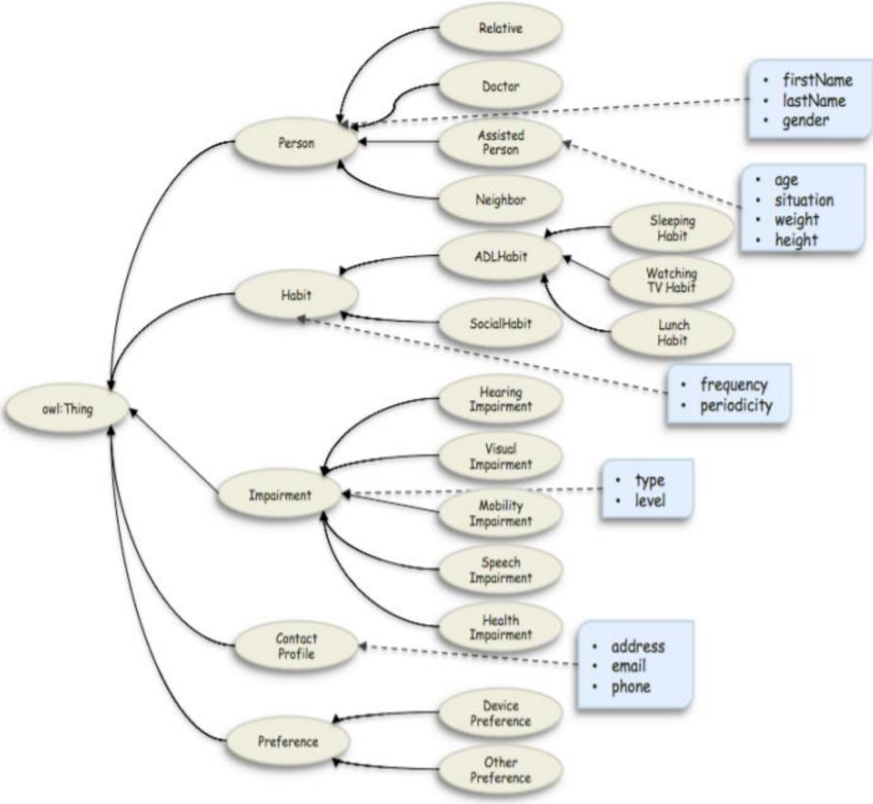

Figure 7 – Person Profile Ontology, adapted from [22]

In [23], it is proposed a user ontology to model information of users using smart home applications. They divided the user ontology in two main components, one component related to static information of the user (such as name and age) and the role of the user (whether the user is a resident or a visitor) and another component related to the profile data of the user (such as heart rate recorded) and preferences (preferred activities).

In [24], it is proposed a ontology-based context modelling approach for a home care assistance scenario where it is defined a Patient Personal Domain Ontology where it is identified different relevant context items related to patient physical data (such as biomedical acquired values), location and activity. These data is then used to automatically infer patient current health status and detect and alert problematic or dangerous situations and events.

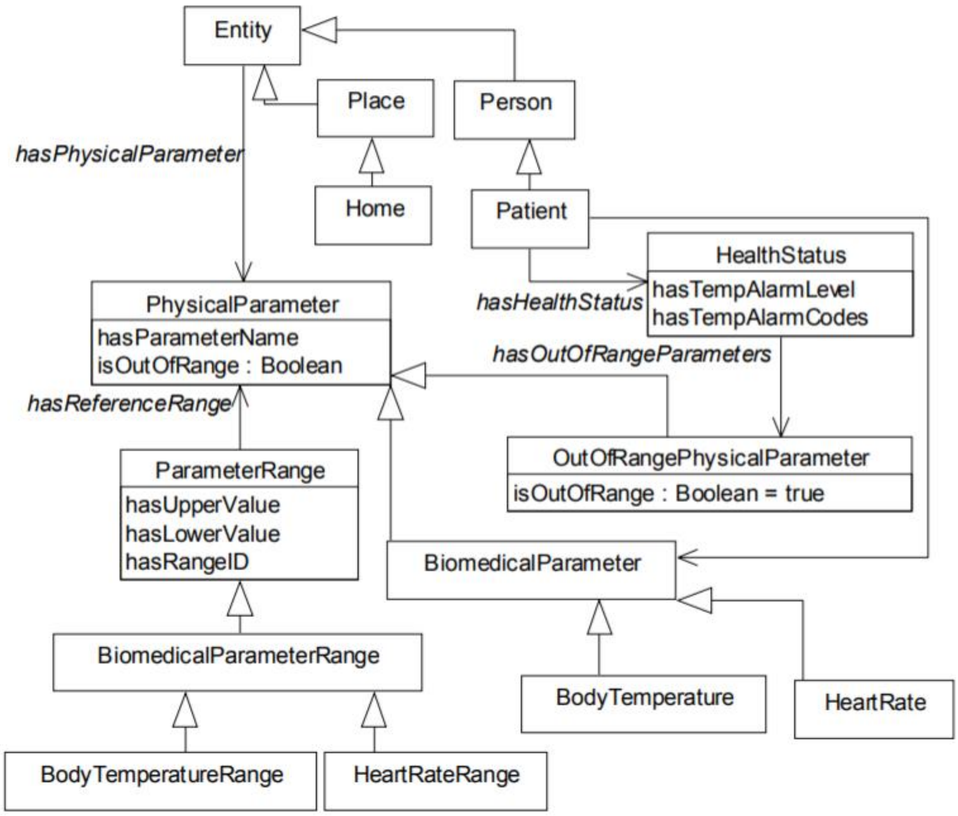

Figure 8 – Patient Personal Domain Ontology, adapted from [24]

2.2.3.2 Domain Ontologies

Domain specific ontologies allow the user to model domain specific concepts and relations. This type of ontology usually focusses on one specific modelling target or area of application, such as healthcare or assisted living. Domain ontologies allow the reuse of complex models that usually require extensive expertise input. Furthermore, domain ontologies can be easily combined since they use same semantic model. In [24], authors also propose the use of Home Domain Ontology which contains relevant context data related to the monitoring of environmental parameters (such as temperature and relative humidity) and then also detect dangerous environmental situations (for example, detect a gas leak or even a fire inside the home environment).

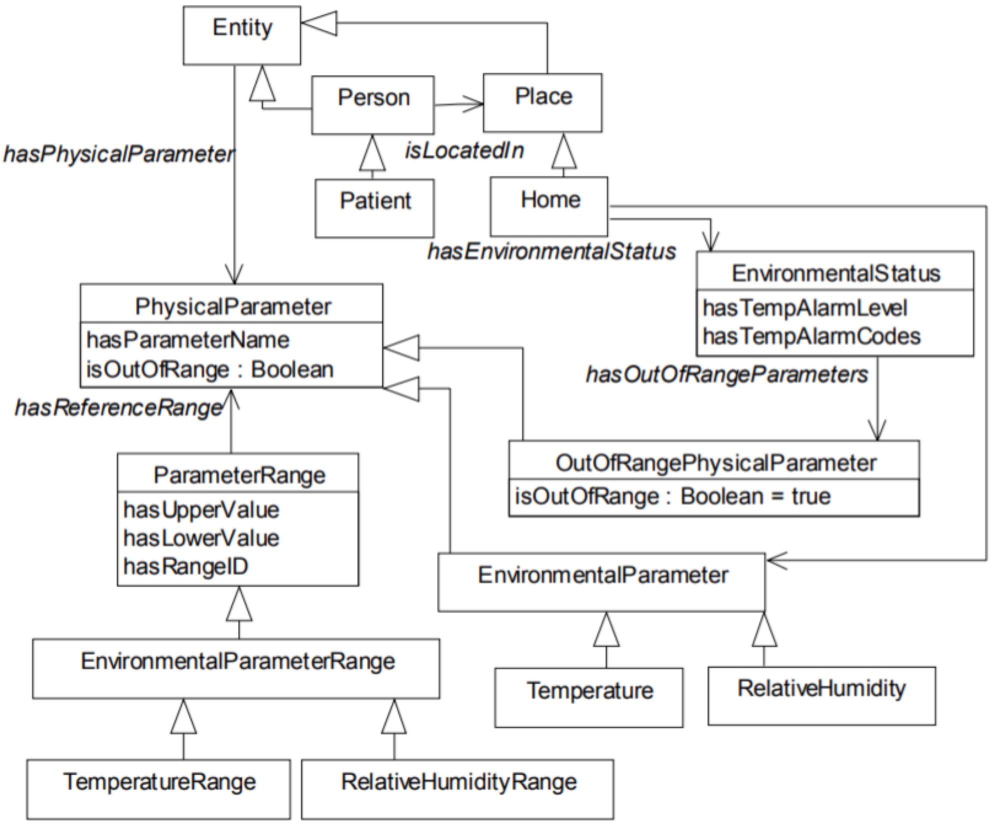

Figure 9 – Home Domain Ontology, adapted from [24]

In [22], it was also proposed a Health Domain Ontology which describes all the basic concepts required to model and support the daily treatment of a disease. The authors proposed a schema for which the knowledge base keeps the information provided to identify problematic situations and detect diseases which the inhabitants may suffer. This domain ontology considers four main classes which are: Disease (it Is modelled each disease the inhabitant may have and the level of gravity), Symptom (symptoms that may occur to the patient and that are relevant to identify a disease), Treatment (describes the type of treatment required to deal with the disease including medication, actions and measurements), and Restriction (restrictions associated to the disease which affects activities, environmental conditions, medication and nutrition).

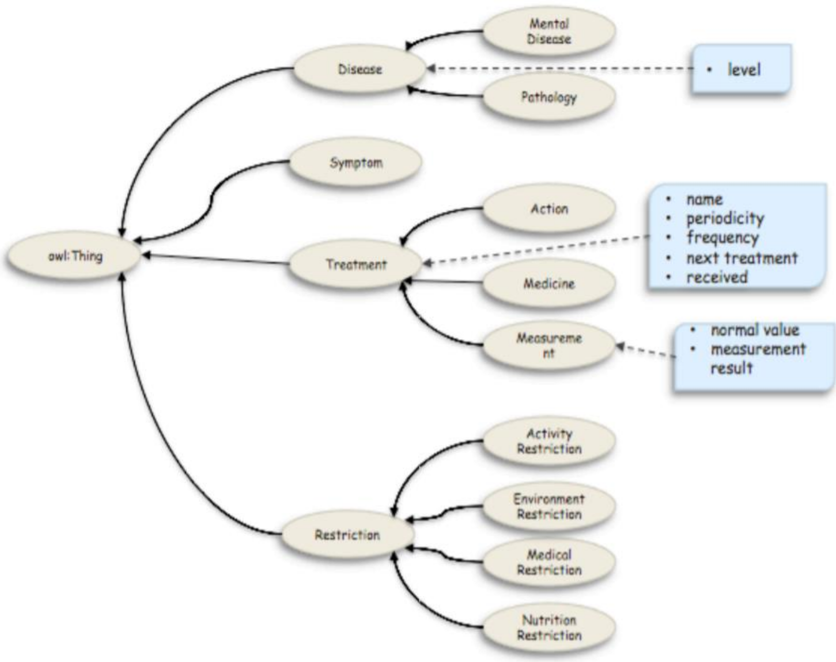

Figure 10 – Health Domain Ontology, adapted from [22]

3 Proposed User Model for PHE

In this section is the described the proposed user model for PHE, where all features for each use-case (Healthy Workplaces and CORD Management) are identified.

Healthy Workplaces use case focus on work related illnesses and unwellness which is a global epidemic suffered by billions of workers all around the world. In addition, workforce is ageing due to demographic changes, and health conditions and chronic diseases are becoming increasingly common. Thus, all these factors, absenteeism, employers' medical expenses and employee productivity is a major burden for enterprises and forcing business to improve the management of their companies health and wellbeing. All over the world, many corporations have implemented different welfare initiatives for disease prevention. However, wellness programs in companies are in the process of expansion, transformation and improvement. Corporate wellness is no longer reduced to health scanners and gym membership reimbursement, currently major trends in corporate wellness are driven by technology. The inclusion of digital portals opens the possibility of data collection, online reporting and many other different features that are driving the growth of Corporate Wellness. In spite of this, the solutions currently on the market tend to provide generalized recommendations. Regarding personalised and preventive healthcare solutions, there are new technologies emerging that PHE intends to introduce into its Coaching Framework.

CORD is a public health problem with increasing demands on healthcare systems and there is a growing market demand for solutions which can help to reduce costs, while maintaining quality of care. Patients with CORD are continuously at risk of deterioration of health, requiring regular medical check-ups and monitoring of their health status. Traditionally health care is delivered through clinicians' face-to-face interaction. With the growing prevalence of CORD and continuous pressure from healthcare authorities/insurance companies, an increasing number of patients is being managed at home in their own environment and most of the time being left alone with traditional self-management materials (books, leaflets, videos, and web-based technology). Coaching solutions appear to be an ideal platform to deliver both simple and effective self-management interventions, while maintaining/improving quality of care and reducing costs. mHealth technologies for CORD should involve monitoring and managing signs and symptoms of the disease, empowering patients to recognize the early signs of exacerbations and to develop skills to better manage their disease.

3.1 User Characteristics

This section contains an outline of the characteristics considered within the Healthy Workplace and CORD Management use cases. There are characteristics related to both Domain Dependent and Independent Data. Domain Independent Data refers to information the user needs to self-report to the PHE system, such as personal information, healthy habits, etc. On the other hand, Domain Dependent Data is extracted directly from the system and connected sensors, such as user navigation, environment, activity tracking, etc.

Regarding Healthy Workplaces use case, the Occupational Health Department was involved in the definition of the main characteristics needed to develop the user model.

Regarding CORD Management use case, it is considered Domain Independent Data that includes characteristics regarding the generic profile of the user which are related to personal information, demographic data, patient background, deficiencies and patient knowledge, and it has been identified characteristics regarding the psychological profile of the user which are his/her cognitive capabilities and preferences. Besides that, it is also considered Domain Dependent Data that includes characteristics related to objectives, user navigation, knowledge acquired, context, aptitude and task preferences.

The description and examples of each mentioned characteristic for the two use cases is provided in **¡Error! No se encuentra el origen de la referencia.**.

Table 3 - User Characteristics in User Modelling for CORD Management Use Case

Model Profile	Characteristics	Descriptions/Examples	Tools to Collect Data	CORD	HW
Domain Independent Data	Personal Information	Name, Email, Password, etc.	User input	Y	Y
	Demographic Data	Age, Height, Weight, Sex, etc.	User input	Y	Y
	Patient Background	Smoker, Pregnant, eating habits, sleep habits, exercise habits…	User input	Y	Y
	Diagnosis	Respiratory disease (asthma, COPD)	User input, Healthcare Records	Y	NR
	Domain of application	Geographic localization of the user, etc.	Smartphone Sensors (GPS)	Y	Y
	Inheritance of the characteristics	Creation of stereotypes that allow to classify the user	Data Mining, Clustering Algorithms	Y	Y
	Knowledge (Background Knowledge)	A collection of knowledge translated in concepts. Possibility of a qualitative, quantitative or probabilistic indication of concepts and knowledge acquired for the user	User input	Y	Y
	Cognitive Capabilities	Emotional State (Anxiety, Depression, Stress, etc,)	Psychological Exams User input	Y	Y
Domain Dependent Data	Objectives	User objectives regarding the use of the system	User input	Y	Y
	Personal Preferences	Classifications of Recommendations Provided (Useful, Not Useful), Interests (Hobbies, Routines)	User input	Y	Y
	Complete description of the navigation	Kept register of each page accessed Definition of the capacity to use the system Definition of the individual preferences with the objectives to adapt the navigation and contents	User Input (Clicks, Likes, Navigation History, etc.) User Performance while using the application (correct vs incorrect actions) Adaptative Interfaces	Y	Y
	Knowledge acquired	A collection of knowledge translated in concepts.	Expert input	Y	Y
	Medication Intake and Health Status	Data related to patient intake of medication; inhalations; record symptoms and exacerbations (SOS Medication, Hospitalization)	User Input, Computerised Respiratory Auscultation, Healthcare Records	Y	NR
	Context model	Data related with the environment of the user (Localization of the user?); Existence of Caregiver or Isolated User	External Resources (Public APIs)	Y	Y
	Activity Tracking	Kept register of end users daily activity	External Resources (Smartphones, Google Fit platform)	Y	Y

3.2 Techniques for User Modelling

Healthy Workplace User Modelling will be based on different user-stereotypes. They will be designed depending on the information provided by the end-user and some general information related to significant Healthy illnesses and compiled by the Occupational Department.

All these stereotypes will be designed and based on different types of variables determined by Health Experts and the Occupational Health Department.

As a result, users will be aggregated in clusters. Some may be more general used for delivering recommendations to maintain a good health condition. And other more specific including users who need to improve their health, carry out challenges or acquire healthy lifestyle habits.

Additionally, each user model will be also fed with extra information provided by the user or directly extracted from the app or wearable connected. The user model will be recalculated periodically in order to re-define and increase the level of personalisation of the Coach Engine.

According to the State of the Art on User Modelling Techniques presented before (Section 2.2) the following techniques are being considered for being used in the Healthy Workplaces User Modelling.

- Beta Distribution (Section 2.2.1.2): it is considered in order to measure the level of acceptance of the suggestions provided (by tracking the number of likes and no-likes of the suggested recommendations).
- Bayesian Networks (Section 2.2.1.4): this algorithm would allow us to establish causal relationships and make predictions on several of the variables. It might help us to predict the appropriate type of recommendation for a user.
- Rule-Based Model or Machine Learning Model (Section 2.2.2.5): Using learning algorithms it is possible to identify which recommendations and challenges work best. Initially it is planned to develop a model based on rules designed by health experts in order to model the current habits of the user or even predict possible unhealthy behaviours.
- Stereotypes (Section 2.2.2.6): This technique describes how to perform a successful stereotyping. It will allow us to group users by their state of health and/or needs on changes in lifestyle, as well as upgrade users from one cluster to another depending on their evolution and change of habits.

The above-mentioned techniques are the ones considered currently as most feasible. Anyhow, other more evolved algorithms may be also taken into account on further SW developments.

User Modelling for CORD Management Use Case will be performed through the use and definition of different stereotypes depending on the characteristics of each user in the system. For this, it will be used the information available in the clinical decision matrix which is proposed in Task T3.3 of PHE (Intelligent Coaching Engine) in which it is specified all the information regarding recommendations provided to the user. This information includes the specification of different high-level and low-level variables which characterize the user and his/her current health condition and the user surrounding environment. For example, the following variables could be considered:

- **BMI** – Body Mass Index, which corresponds to the value derived from the mass (weight) and height of an individual;
- **CPAP** – Use of Continuous Positive Airway Pressure which is a form a form of positive airway pressure ventilator that continuously applies mild air pressure to keep the airways open in people who are not able to breathe spontaneously on their own;

- **OSA** – Obstructive Sleep Apnea which is caused by complete or partial obstructions of the upper airway during sleep;
- **AQI** – Air Quality Index, which indicates how polluted the air currently is or how polluted it is forecast to become.
- **Age**;
- **Adherence to medication**;
- Etc.

3.3 Proposed Architecture

The proposed architecture for Healthy Workplaces use case is divided in three main parts:

- The User Model describes the current user lifestyle, wellbeing habits and status. This component allows obtaining conclusions on the user characteristics.
- The Domain Model is a set of domain concepts that have different functions. Each concept is related to other concepts generating something similar to a semantic network and providing a structure that allows the representation of the user data set. The value of each concept will be calculated and expressed quantitatively, qualitatively or in probabilistic form depending on the final objective.
- The Interaction Model defines the interaction between the user and the system. This information is used to derive users' characteristics in order to update and validate the user model established. It will therefore include evaluation, adaptation and inference mechanisms.

The proposed architecture for CORD Management use case is divided in three main phases as can be seen in Figure 11.

The first phase is the Data Acquisition phase where different tools will be used to collect both domain dependent and domain independent data (user characteristics). In the second phase, the data processing phase, all the relevant variables associated to the user will be processed and analysed and different groups (or clusters) will be identified. For this we can consider the use of different clustering algorithms such as k-means clustering, hierarchical clustering, fuzzy clustering, biclustering, etc. [25, 26]. In the last phase, the Data Evaluation, each identified cluster will then correspond to a stereotype and describe a group of users according to certain characteristics.

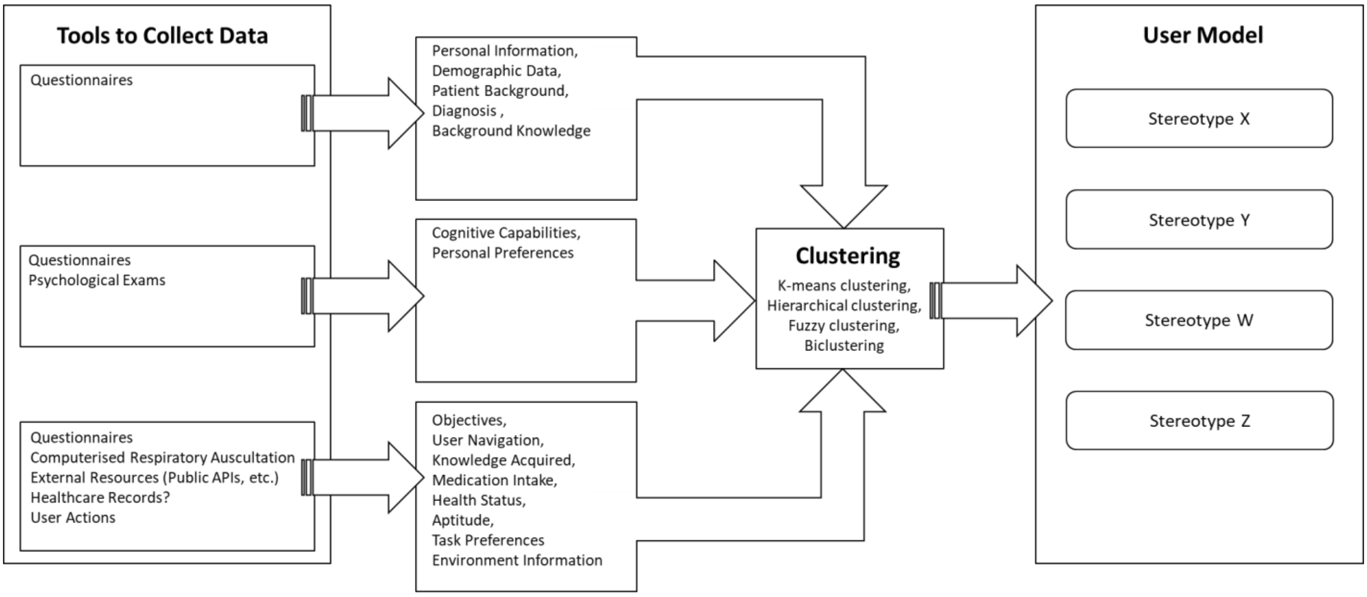

Figure 11 - CORD Management User Modelling Architecture

4 PHE as a whole

Coaching application will provide assistance for individuals in collaboration with suggestions to maintain healthier life. In the meantime, of supporting individuals, their psychological attitudes will be taking place in following parameters. Situation of workplaces as one of the most important effects in psychological attitudes will be analysed.

Herein, these are some of the occurred factors in workplace;

- Workload
- Stress
- Fear of losing job
- Mobbing
- Personal Features and Characteristics
- Deadline Processes
- Infectious Diseases

Coaching application will be used in how these factors affect the efficiency of work. Besides, how the psychological and psychical healthcare effect workplace will be clarified. Making suggestions for recovery of individual's healthcare classification of diseases (chronical diseases, cardiac diseases, diabetes, etc.) with using Personal Coaching Application.

References

1. Durrani, Q.S.: Cognitive modeling: a domain independent user modeling. Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on 1, 217-220 (1997)
2. Myers, I.B.: The Myers-Briggs Type Indicator: Manual (1962). (1962)
3. Jung, C.: Psychological types. Routledge (2016)
4. Eysenck, H.J.: Dimensions of personality. Transaction Publishers (1950)
5. Fiske, D.W.: Consistency of the factorial structures of personality ratings from different sources. The Journal of Abnormal and Social Psychology 44, 329 (1949)
6. Norman, W.T.: Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. The Journal of Abnormal and Social Psychology 66, 574 (1963)
7. McCrae, R.R., John, O.P.: An introduction to the five-factor model and its applications. Journal of personality 60, 175-215 (1992)
8. Goldberg, L.R.: An alternative" description of personality": the big-five factor structure. Journal of personality and social psychology 59, 1216 (1990)
9. Zukerman, I., Albrecht, D.W.: Predictive statistical models for user modeling. User Modeling and User-Adapted Interaction 11, 5-18 (2001)
10. Orwant, J.: For want of a bit the user was lost: Cheap user modeling. IBM Systems Journal 35, 398-416 (1996)
11. Kass, R., Finin, T.: Modeling the user in natural language systems. Computational Linguistics 14, 5-22 (1988)
12. Martins, C., Faria, L., De Carvalho, C.V., Carrapatoso, E.: User modeling in adaptive hypermedia educational systems. Educational Technology & Society 11, 194-207 (2008)
13. Nguyen, L., Do, P.: Combination of Bayesian network and overlay model in user modeling. International Conference on Computational Science 5-14 (2009)
14. Bushey, R., Mauney, J.M., Deelman, T.: The development of behavior-based user models for a computer system. UM99 User Modeling, pp. 109-118. Springer (1999)
15. Rich, E.: User modeling via stereotypes. Cognitive science 3, 329-354 (1979)
16. Gruber, T.R.: A translation approach to portable ontology specifications. Knowledge acquisition 5, 199-220 (1993)
17. Brickley, D.: RDF vocabulary description language 1.0: RDF schema. http://www. w3. org/TR/rdf-schema/ (2004)
18. McGuinness, D.L., Van Harmelen, F.: OWL web ontology language overview. W3C recommendation 10, 2004 (2004)
19. Andrejko, A., Barla, M., Bielikova, M.: Ontology-based user modeling for web-based information systems. Advances in Information Systems Development, pp. 457-468. Springer (2007)
20. Gouardères, G., Conté, E., Mansour, S., Razmerita, L.: Ontology based user modeling for personalization of grid learning services. 1st International ELeGI Conference on Advanced Technology for Enhanced Learning 8 (2005)
21. Jiang, X., Tan, A.-H.: Learning and inferencing in user ontology for personalized Semantic Web search. Information sciences 179, 2794-2808 (2009)
22. Zografistou, D.: Support for context-aware healthcare in ambient assisted living. MasterÆs thesis (2012)

23. Ni, Q., Pau de la Cruz, I., García Hernando, A.B.: A foundational ontology-based model for human activity representation in smart homes. Journal of Ambient Intelligence and Smart Environments 8, 47-61 (2016)
24. Paganelli, F., Giuli, D.: An ontology-based context model for home health monitoring and alerting in chronic patient care networks. null 838-845 (2007)
25. Jain, A.K., Dubes, R.C.: Algorithms for clustering data. (1988)
26. Hartigan, J.A.: Clustering algorithms. (1975)