



**Personal Healthcare
Empowerment**

SOTA

PUBLIC STATE-OF-THE-ART-DOCUMENT

Document requested by ITEA reviewers

Document Type	Document
Document Version	V0.3
Access Level	public
Editor	Robin De Croon (KU Leuven)
Contributors	All partners

Document History

Date	Version	Editors	Status
17/07/2019	0.1	KU Leuven	Draft
03/09/2019	0.3	KU Leuven	Draft

Table of Content

Executive summary	3
1 Heterogeneous data analytics.....	5
1.1 Personal health data monitoring	5
1.1.1 Existing PHR's and cloud-based data management solutions	5
1.1.2 Quantified self and wearable technologies.....	6
1.2 Respiratory and lung function monitoring	7
2 Learning persona-specific models.....	7
2.1 Risk analysis and absenteeism	8
3 User modelling definitions and state-of-the-art	9
3.1 User Characteristics.....	9
3.1.1 Domain Dependent Data (DDD).....	9
3.1.2 Domain Independent Data (DID).....	10
3.1.3 Common Characteristics	13
3.2 Techniques for User Modelling	14
3.2.1 Statistical Techniques.....	14
3.2.2 Non-Statistical Techniques.....	15
3.2.3 Ontologies	18
4 Personalized coaching.....	23
4.1 Recommender systems	23
4.2 Visualisation	26
4.3 Coaching solutions	26
4.3.1 Psychological approaches	26
4.3.2 Technology enabled coaching solutions	28
4.4 Preventive health solutions.....	28
4.5 Healthcare (and pain) apps	29

Executive summary

In this document, we will present the state-of-the-art with respect to tool support and usage in the prevention and increased wellbeing and functioning, as well as on the level of data analytics and coaching technologies applied in the context of the ITEA Personal Health Empowerment project. Within the 'Personal Health Empowerment' project we rely on trackers for physiological signals, sensors for behavioural signals and self-reported data.

One of the key challenges in this project consists of linking physiological and occupational parameters obtained from health records, wearables, devices and self-report tools to wellbeing and to translate these insights in appropriate personal health recommendations that suggest how to improve health and lifestyle. This implies several challenges in the areas of data-analytics and recommendation techniques.

Heterogeneous data analytics

A first challenge will be the treatment of the heterogeneous data (of mixed continuous and categorical type containing missing data and sparse feature vectors) that arises from different parameters stored in electronic health records and different types of measurements from wearables and self-report tools. These multimodal streams of data must be properly translated into actionable insights in the relationship between wellbeing and different work-related health parameters. For this purpose, we will build on expertise in statistical learning algorithms (multilevel hierarchical clustering models) to extend them with techniques that are able to deal with the heterogeneous data and that allow multimodal information fusion. In comparison to existing studies, these analytics are intended to lead to the discovery of novel patterns and relationships between the large number of different parameters provided by health records and the user himself and the occurrence of, for example, absenteeism.

Learning persona-specific models

A second challenge will be the difficulty to come to persona-specific models that enable to boost performance of general models that are averaged over larger populations such that more reliable estimation on risks may be obtained, allowing to sooner detect situations of elevated risks. Research must decide how to improve predictive accuracy and restrict bias of these models of absenteeism in a sequential and incremental manner where the user is able to provide the model with new information through wearables or his smartphone while allowing heterogeneous data sources and sparse feature vectors. An additional challenge is that in such adaptive and incremental learning approach, streams of incoming data from wearables and self-report tools could become large. Models should therefore keep a bounded memory footprint allowing for real-time prediction while avoiding storage issues. This implies the ability to discard useless or old information.

User modelling definitions

Personalization can only be realized with appropriate user modelling. In the past several years, modelling users became one of the business where big enterprises invest most resources and efforts. New trends shift to user's personal assistants. User modelling constitutes a global field of research that involves

several branches, some examples are modelling user behaviour, user preferences, and even user experience when using a system. Having an accurate profile of users becomes an asset when personalizing systems and services i.e. adaptation.

Personalised coaching

In addition, little research has been published on the application of recommendation techniques to support health recommendations. Although interesting ideas have been presented in the literature, there are several challenges that need to be tackled to deploy the techniques in real world applications. First and foremost, typical recommender systems act like a “black box”: i.e. they show suggestions without explaining the provenance of these recommendations ¹. Whereas the approach is suitable in typical e-commerce applications that involve little risk, transparency is a core requirement in higher risk application domains. Employees as well as health professionals need to understand why a recommendation is made in order to assess its value and importance. User trust is a key issue that needs attention.

Health related recommendations often need input from domain experts. In this project, we will research the use of interactive visualisations on top of recommender systems for both healthcare professionals and employees. The objective for health professionals is to provide input and feedback to steer the recommendation process. The objective for employees is to engage with the system and similarly provide additional input and feedback to improve recommendations. The approach will be researched to increase accuracy of recommendations, as well as to improve user engagement and motivation. *“Vocational rehabilitation [targeting prevention] is an active process that depends on the participation, motivation and effort of the individual, supported by the workplace and healthcare.”* By visualising recommendations and progress that has been made by employees, we aim to increase adherence and reduce drop-out of coaching apps. Hence, a next challenge consists of researching to what extent employees are accepting health related recommendations and how they should be represented to engage and empower the end-user. To address this challenge, we will rely on participatory action research ². Via this methodology, interactive recommendation techniques will be designed, selected and refined in collaboration with health professionals and employees. Based on the outcome of these studies, these interactive recommendation techniques will then be designed and prototyped in the coaching platform. As this proposal seeks an evidence-based approach, the effects will be validated via elaborated pilots (6-9 months).

1 He, C., Parra, D., Verbert, K. (2016). Interactive recommender systems: a survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, 56, 9-27.

2 Baum, F., MacDougall, C., & Smith, D. (2006). Glossary: Participatory action research. *Journal of Epidemiology and Community Health* (1979-), 60(10), 854-857.

1 Heterogeneous data analytics

1.1 Personal health data monitoring

Existing technologies for health monitoring are used to promote physical activity and a healthy lifestyle to keep users fit^{3,4}. These include wearables (e.g. Fitbit) and smartphone apps using e.g., inquiries, to monitor a variety of health-related modalities such as motion (through accelerometers), heart rate, blood pressure, skin temperature, calories and eating habits. Little or no technology exists for the discovery of patterns and relationships between the large number of different parameters provided by such health & fitness data and the occurrence of health issues. In this project, advanced learning algorithms will be developed to identify factors from personal data that influence absenteeism and that may vary from person to person and may change over time.

1.1.1 Existing PHR's and cloud-based data management solutions

The Medical Library Association defines a personal health record (PHR) as: *“A private, secure application through which an individual may access, manage, and share his or her health information. A PHR can include information that is entered by the consumer and/or data from other sources such as pharmacies, labs, and health care providers. A PHR may or may not include information from the electronic health record (EHR) that is maintained by the health care provider and is not synonymous with the EHR. PHR sponsors include vendors who may or may not charge a fee, health care organizations such as hospitals, health insurance companies, or employers.”* As indicated, a PHR does not eliminate the need for an EHR, which provides the official tool used for care documentation. On the other hand, in countries, where EHR systems and infrastructures are not yet well established, a PHR may be a substitute for an EHR or convey clinical information between healthcare organisations.

In many cases, a PHR is an integral part of a healthcare provider's electronic health record system or online patient portal, in which case it is referred to as a tethered PHR. The disadvantage of a tethered PHR is that it is bound to a certain healthcare service provider and available to the individual only if he or she continues to be a customer of the provider. A tethered PHR also lacks support for the continuity of care paradigm⁵. Alternatively, an interconnected PHR exposes open interfaces to other systems, which are trusted based on certain quality criteria. The advantage of interconnected PHRs is that they can, in principle, be connected to all EHR systems containing the individual's health data.

³ Leijdekkers, P., & Gay, V. (2015, June). Improving User Engagement by Aggregating and Analysing Health and Fitness Data on a Mobile App. In International Conference on Smart Homes and Health Telematics (pp. 325-330). Springer International Publishing

⁴ Wang, J., Patten, S.B., Lam, R.W., Attridge, M., Ho, K., Schmitz, N., Marchand, A., & Lashewicz, B.M. (2016). The effects of an e-mental health program and job coaching on the risk of major depression and productivity in Canadian male workers: Protocol for a randomized controlled trial. *Journal of Medical Internet Research - Research Protocols*, 5(4).

⁵ Gorp, P. V., Comuzzi, M., Fialhom, A., Kaymak, U. (2012). Addressing Health Information Privacy with a novel Cloud-Based PHR System Architecture. 2012 IEEE International Conference on Systems, Man, and Cybernetics; 2012 Oct 14–17; Seoul, Korea: IEEE. Pp. 1841–6.

1.1.2 Quantified self and wearable technologies

Today's wearable technologies are offering novel opportunities to enable people to overcome their conditions. A great proportion of users has reported improved capability of self-awareness of their own health and lifestyle and has indicated that wearables induce to some extent temporary behaviour change, but there is still little evidence to support the implication of wearable technologies in healthy lifestyle promotion and behaviour change.

There is a growing movement of 'quantified self' across several domains, including medicine, sports, and learning. The focus of quantified self is on collecting traces that users leave behind and using those traces to improve their experiences. An important driver for this research is increasing motivation, autonomy, effectiveness and efficiency of users. Capturing relevant quantified self-data in a scalable way is challenging. Both manual and automated tracing methodologies can be used. Automatic tracing technologies capture traces from either hardware or software sensors. Although the data include more noise, they are less intrusive than the manual ones and the amount of missing data will be less significant.

The development of accurate algorithms and engaging coaching app within this project could support the validation on wearables-based approaches for the problem of addiction in the consumer market. At this stage wearables are used only as a monitoring tools and data analytics, connected to it, focuses mainly to improve visualization of information. The possibility to use wearables to identify trigger inducing addictive behaviour in daily life and consequently be instrumental for the prediction and reduction of relapses and worsening of the health condition, remains unexplored in real-life scenarios. Wearables offer a quantitative alternative and/or complementary solution to self-reports. Very little is known about the evolution of physiological functioning in combination with the behavioural phenotype of people with addiction. The use of highly accurate wearables for vital signs monitoring could then also offer new insights on the physiopathology of these disorders as never explored before.

The recognition of daily activities will be essential for the follow-up of people's lifestyle. Today's wearable watches and health integration platforms are primarily centred on motion or accelerometry (ACM) for the classification of activities. Moreover, underlying algorithms (based on machine learning) are trained in a supervised setting where labels of the data are required at forehand leading to patient independent models. Human activities, however, can be complex and very diverse among different individuals such that person independent models can perform suboptimal, especially when activities become more complex.

Current research efforts in alcohol addiction from the technological point of view are focusing on the development of wearable solutions for continuous alcohol consumption monitoring, using transdermal sensors, sweat analysis and near the skin ethanol vapor detection. Despite the known physiological measurable and macroscopic effects of alcohol abuse, technologies available on the market for the addiction treatment industry are still limited and their effectiveness is mainly based on anecdotal evidence.

1.2 Respiratory and lung function monitoring

The tool recently proposed by MEDIDA using respiratory sounds as spirometry surrogate, although with very encouraging results, still has limited clinical validation and lacks parameters characterizing the patient state and evolution with respect to last spirometry. The estimation of clinically relevant parameters from the sound signal during forced breathing is far from complete. The definition of new indices / parameters that reflect the patient's condition, namely indices of variation with respect to the last available laboratory respiratory function test, is a pertinent question to be explored. Even so, the biggest challenges are associated with the acquisition and pre-processing component. On the one hand, the algorithms have to be robust to variables that are difficult or even impossible to control, such as variations between devices in the technical specifications of the microphones, physical location in them, distance of the user to the microphone during the manoeuvre, position of the mouth during forced expiration. On the other hand, it is necessary to develop algorithms that deal adequately with the contamination of acquired signals and separate unwanted noise (e.g., coughing, ambient noise). Recognizing these limitations, MEDIDA has made small incremental improvements to the initial algorithms while continuing to collect data in its LabAIR clinical units that will serve as a foundation for developments.

Based on the team's experience in developing the automatic respiratory recording and analysis system, and the experience of a team member in DLA in patients with COD, this project aims to develop a new respiratory sound recording and analysis system, using the microphone of the smartphone as input. The pre-existing algorithms developed to analyse sound files acquired essentially with electronic stethoscopes also need to be improved to properly analyse the files acquired with the microphone.

2 Learning persona-specific models

Data gathering and analytics trends have increased during the past few years. The massive data collection from sensors and devices of all kind that will be processed in centralized data centres or in the cloud leads to a potential congestion of the data networks and to a most times unnecessary amount of processing data effort. This results in performance reduction and increased costs. Those effects can be minimized applying intelligence to today's data and IoT networks. By means of data analysis, what data sources or inputs are decisive or more important in a process can be determined, so the network itself can decide which sensing device input needs to give priority and which one can be ignored for a specific task. With this information, data networks can be re-configured dynamically, reducing the amount of data gathered, avoiding network congestion and processing data effort that is not needed as it can be inferred from the learning process.

Big data in healthcare is related to large, diverse and complex electronic health datasets, by nature difficult to manage and analyse with traditional software/hardware approaches. Predictive modelling and real-time analytics, but also small data, are important forms of data analytics that may contribute to collect, manage, and analyse raw data towards the improvement of the quality of healthcare, expectedly with lower costs.

While traditionally analytics focused on business intelligence, operational research and data mining, advanced analytics is focusing on descriptive and predictive tasks. Particularly, healthcare data analytics can be improved using different tools and techniques, for different tasks, including automated and semi-automated classification, regression, clustering, text mining, social network analysis, anomaly detection, time series/sequence analysis, advanced data visualization, among others.

There are many techniques and tools available for use with healthcare data. The volume of data is not necessarily related to better analytic results. Each case is a case and choosing the right analytics tools, for scenarios, may lead to better results.

2.1 Risk analysis and absenteeism

Statistical models to estimate risks mainly rely on correlation and (logistic) regression analysis⁶. Past research has also identified personal factors (e.g. work ability, individual health status, age, gender) as well as work-related factors (e.g. job demands, organisation changes, size of organisation, location)^{7,8}. Literature also indicates possible effects of pain experience, pain cognitions and stress symptoms on absenteeism⁹. However, little or no research exists that combines all these parameters to predict and characterise wellbeing issues¹⁰. Also, in a recent study only expectations on recovery have been shown to be a good predictor for absenteeism in the context of pain¹¹. Studies on the effect of personal factors on absenteeism as e.g. activity are mainly limited to the inclusion of self-reported measures that are acquired on a limited number of moments in time¹². The inclusion of parameters and trends therein that are continuously monitored, including vital sign data, and the effect of this inclusion on the accuracy and speed of prediction has not been studied yet.

Furthermore, existing models to estimate risks on absenteeism are obtained in a generic way, i.e. they are tuned to an average target population. The predictive power of these models, however, may not translate to individuals separately due to personal differences. An attempt to overcome this problem is to come to personalised models. Learning of personalised models is a well-recognised problem in machine learning

6 Sturman, M.C. (1996) Multiple Approaches to Absenteeism Analysis (CAHRS Working Paper #96-07). Ithaca, NY: Cornell University, School of Industrial and Labor Relations, Center for Advanced Human Resource Studies.

7 Cristofoli, D, Turrini, A and Valotti, G. (2011) Coming Back Soon: Assessing the Determinants of Absenteeism in the Public sector, *Journal of Comparative Policy Analysis: Research and Practice*, 13(1):75-89.

8 Van den Bergh, T.I, Elders, L.A., de Zwart B.C. and Burdorf, A. (2009). The effects of work-related and individual factors on the Work Ability Index: a systematic review. *Occupational and Environmental Medicine*, 66(4):211-220.

9 Adeniyi, P.O. (2015) Stress, a Major Determinant of Nutritional and Health Status, *American Journal of Public Health Research*, 3(1):15-20.

10 Meng, L. Robinson, K.T. and Smith, M.L. (2017) Factors associated with sickness absence among employees with chronic conditions, *Occupational Medicine*, 67(4):296-300

11 Kuijjer, W., Groothoff, J. W., Brouwer, S., Geertzen, J. H., & Dijkstra, P. U. (2006). Prediction of sickness absence in patients with chronic low back pain: a systematic review. *Journal of occupational rehabilitation*, 16(3), 430-458.

12 K. D. Ahern, M. J. Kreslake, and M. J. Phalen, What is ehealth (6): Perspectives on the evolution of ehealth research, *J Med Internet Res*, vol. 8, no. 1, Mar 2006, p. e4. ^[11]_{SEP}

applications such as activity recognitions and facial recognition^{13,14}. A large part of work in this area simply uses personal data from an individual to estimate a person specific model. This will not be possible in the prediction of absenteeism, however, as any relevant personal data may be too sparse and individual absence may be rare. Another attempt would be to come to personas or profiles that collect individuals with common underlying factors that influence absenteeism. A large part of the literature on the recognition of personas focuses on marketing purposes to outline representations of customer habits and interests and relies on unsupervised clustering and classification techniques¹⁵. In this project, amongst others, we aim to adapt these techniques to outline representations of different degrees of absenteeism. Existing literature on clustering identifies difficulties as incompleteness (missing data in electronic health records), mixed data of categorical and continuous type (e.g. gender or type of work is categorical, heart rate is continuous) and sparse features (e.g. self-reported features) that will arise simultaneously in the prediction of absenteeism¹⁶. Research must decide how to properly combine existing clustering methods in a unifying framework to treat these difficulties simultaneously. Furthermore, adaptive approaches will be investigated where predictions are updated according to novel relevant information from the user e.g., a new sickness or extra stress caused by a change in family situation.

3 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.

3.1 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.

3.1.1 Domain Dependent Data (DDD)

Domain dependent data is related with system responses tailored according to the domain knowledge of a user¹⁷. 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

13 Zen, G. Porzi, L., Sangineto, E., Ricci, E. and Sebe, N. (2016) Learning Personalized Models for Facial Expression Analysis and Gesture Recognition, *IEEE Transactions on Multimedia*, 18(4): 775-788.

14 Garcia-Ceja, E. (2015) Building Personalized Activity Recognition Models with Scarce Labeled Data Based on Class Similarities, In: García-Chamizo J., Fortino G., Ochoa S. (eds) *Ubiquitous Computing and Ambient Intelligence. Sensing, Processing, and Using Environmental Information*. Lecture Notes in Computer Science, vol 9454.

15 Blomstedt, P. Tang, J. Xiong, J. Grandlund, C. and Corander, J. (2015) A Bayesian Predictive Model for Clustering Data of Mixed discrete and Continuous Type. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37(3):489-498.

16 Xu, R and Wunsch D.C. (2010) Clustering algorithms in biomedical research: a review. *IEEE reviews in biomedical engineering*, 3:120-54.

17 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)

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.

3.1.1.1 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.

3.1.1.2 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.

3.1.2 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 ¹⁸. In other words, domain-independent knowledge corresponds to the psychological 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.

3.1.2.1 Myer-Briggs Type Indicator

Myer-Brigg Type Indicator model ¹⁹ 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 ²⁰ 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.

18 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)

19 Myers, I.B.: The Myers-Briggs Type Indicator: Manual (1962). (1962)

20 Jung, C.: Psychological types. Routledge (2016)

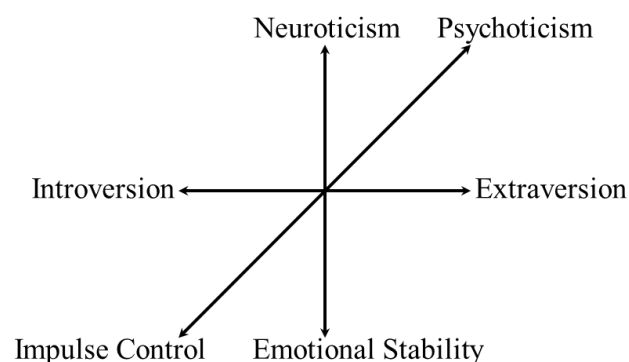


ISTJ Responsible Executors	ISFJ Dedicated Stewards	INFJ Insightful Motivators	INTJ Visionary Strategists
ISTP Nimble Pragmatics	ISFP Practical Custodians	INFP Inspired Crusaders	INTP Expansive Analizers
ESTP Dynamic Mavericks	ESFP Enthusiastic Improvisors	ENFP Impassioned Catalysts	ENTP Innovative Explorers
ESTJ Efficient Drivers	ESFJ Committed Builders	ENFJ Engaging Mobilizers	ENTJ Strategic Directors

Myer-Briggs Type Indicator

3.1.2.2 Eysenck's Pen Model

In 1950²¹, 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.



Eysenck's PEN Model

²¹ Eysenck, H.J.: Dimensions of personality. Transaction Publishers (1950)



3.1.2.3 *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 ^{22 23 24 25} 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.

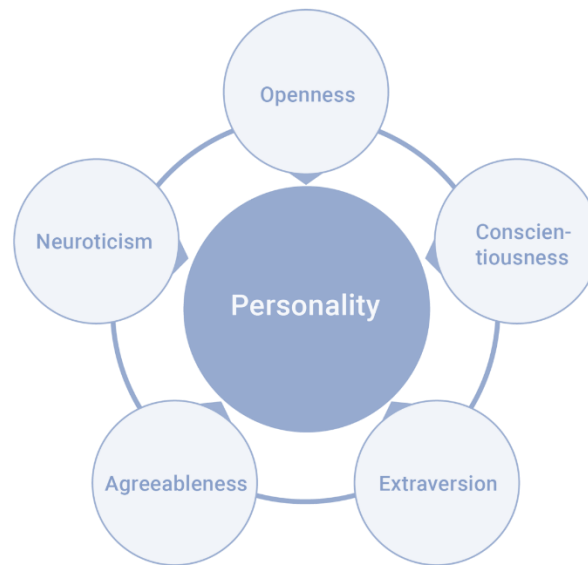
Neuroticism – Trait associated to characteristics such as sadness, moodiness, and emotional instability. People who have high neuroticism tend to experience mood swings, anxiety, irritability, and sadness.

22 Fiske, D.W.: Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology* 44, 329 (1949)

23 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)

24 McCrae, R.R., John, O.P.: An introduction to the five-factor model and its applications. *Journal of personality* 60, 175-215 (1992)

25 Goldberg, L.R.: An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology* 59, 1216 (1990)



Five Traits of Personality

3.1.3 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

3.2 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.

3.2.1 Statistical Techniques

3.2.1.1 Linear Modelling

Linear Modelling is a technique which takes the weighted sum of known values and predicts the value of an unknown quantity²⁶. 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.

3.2.1.2 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 an estimate and a confidence level²⁷. 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.

3.2.1.3 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 pairs²⁸. 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.

3.2.1.4 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 nodes²⁹. 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

26 Zukerman, I., Albrecht, D.W.: Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction* 11, 5-18 (2001)

27 Orwant, J.: For want of a bit the user was lost: Cheap user modeling. *IBM Systems Journal* 35, 398-416 (1996)

28 Zukerman, I., Albrecht, D.W.: Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction* 11, 5-18 (2001)

29 Zukerman, I., Albrecht, D.W.: Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction* 11, 5-18 (2001)

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.

3.2.1.5 Rule Induction Model

Rule Induction Model consists of learning sets of rules that predict the class of an observation from its attributes³⁰. 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.

3.2.2 Non-Statistical Techniques

3.2.2.1 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 base³¹.

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)³².

In this method, the user knowledge is related, layer to layer, to the Domain Model, producing the user knowledge model (Figure 1)³³. 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.).

30 Zukerman, I., Albrecht, D.W.: Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction* 11, 5-18 (2001)

31 Kass, R., Finin, T.: Modeling the user in natural language systems. *Computational Linguistics* 14, 5-22 (1988)

32 https://www.eelcoherder.com/images/teaching/usermodeling/03_user_modeling_techniques.pdf

33 Martins, C., Faria, L., De Carvalho, C.V., Carrapatoso, E.: User modeling in adaptive hypermedia educational systems. *Educational Technology & Society* 11, 194-207 (2008)

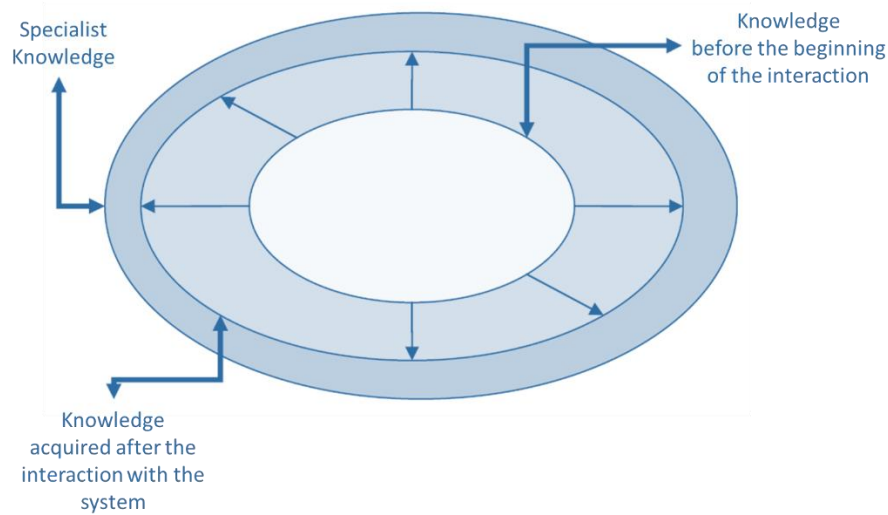


Figure 1 – Representation of the Overlay Model, adapted from [12]

3.2.2.2 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 knowledge³⁴.

Perturbation model represents learners as the subset of expert's knowledge plus their mal-knowledge³⁵.

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 2)³⁶. This method can be used to represent knowledge that is beyond the Domain Model defined by the specialist.

34 Kass, R., Finin, T.: Modeling the user in natural language systems. *Computational Linguistics* 14, 5-22 (1988)

35 Nguyen, L., Do, P.: Combination of Bayesian network and overlay model in user modeling. *International Conference on Computational Science* 5-14 (2009)

36 Martins, C., Faria, L., De Carvalho, C.V., Carrapatoso, E.: User modeling in adaptive hypermedia educational systems. *Educational Technology & Society* 11, 194-207 (2008)

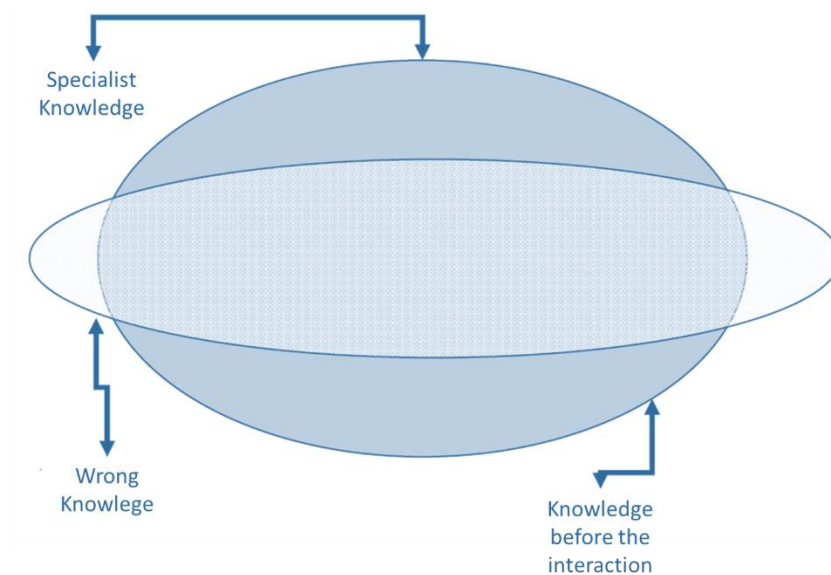


Figure 2 – Representation of the Perturbation Model, adapted from [12]

3.2.2.3 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.

3.2.2.4 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 users³⁷. The goal of this type of models is to develop a system that can accommodate 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.

3.2.2.5 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

³⁷ 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)

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.

3.2.2.6 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 38, 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	<i>exercise</i> <i>food</i>	ExercisePractitioner	<i>exercise</i> ₁	<i>exercise</i> ₂
		FoodEater	<i>food</i> ₁	<i>food</i> ₂
		FoodForExercise	<i>exercise</i> ₁	<i>food</i> ₁
Domain Attributes	<i>food</i>	FoodNutrient	<i>attribute</i> ₁ in <i>food</i>	<i>food</i> ₂
User Attributes	<i>user</i>	UserExerciser	People who have	Will like
			<i>attribute</i> ₂	<i>exercise</i> ₂

3.2.3 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. One of the most common definitions comes from Gruber which refer to ontologies as “an explicit specification of a conceptualization” 39. 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.

38 Rich, E.: User modeling via stereotypes. Cognitive science 3, 329-354 (1979)

39 Gruber, T.R.: A translation approach to portable ontology specifications. Knowledge acquisition 5, 199-220 (1993)

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 ⁴⁰ and OWL ⁴¹, with the last one being recognized as a standard by the W3C Consortium⁴². There are several advantages associated to the use of ontologies which are:

- Possibility to reuse existing ontologies, considering possible adaptations or extensions of knowledge base 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.

3.2.3.1 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 ⁴³. Using sharable data structures containing user's features and preferences will enable personalized interactions with different devices for the benefit of the users ⁴⁴. 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 ⁴⁵. Several works have been proposed in the literature regarding the definition and use of user ontologies. For example, in Zografistou et al. ⁴⁶ 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).

40 Brickley, D.: RDF vocabulary description language 1.0: RDF schema. <http://www.w3.org/TR/rdf-schema/> (2004)

41 McGuinness, D.L., Van Harmelen, F.: OWL web ontology language overview. W3C recommendation 10, 2004 (2004)

42 The World Wide Web Consortium: <https://www.w3.org/>

43 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)

44 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)

45 Jiang, X., Tan, A.-H.: Learning and inferencing in user ontology for personalized Semantic Web search. *Information sciences* 179, 2794-2808 (2009)

46 Zografistou, D.: Support for context-aware healthcare in ambient assisted living. Master's thesis (2012)

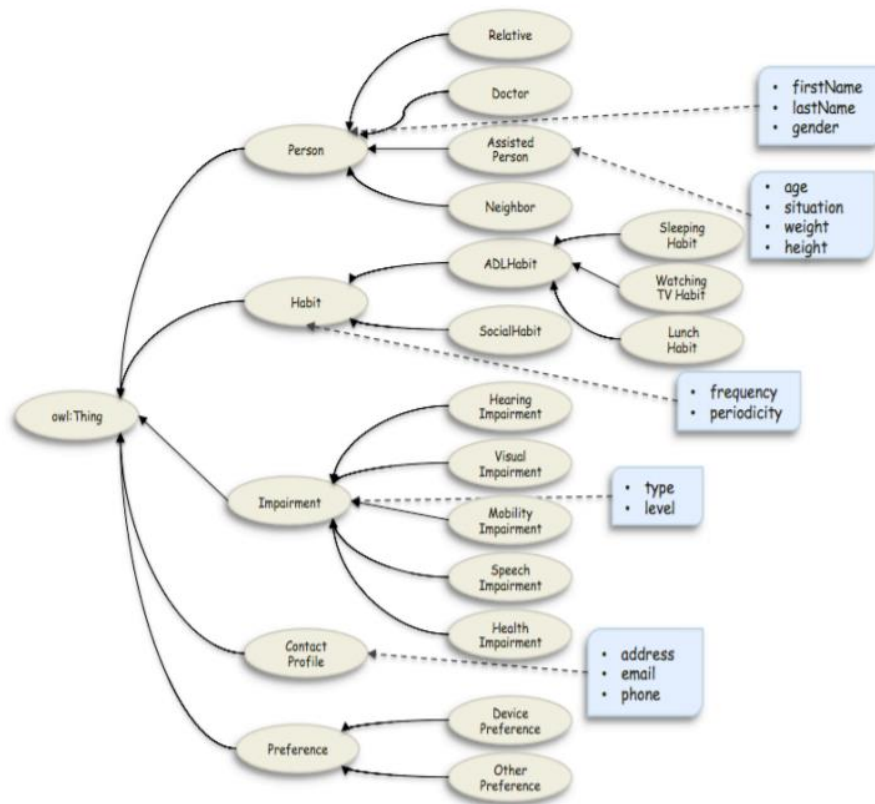


Figure 3 – Person Profile Ontology, adapted from [22]

In Ni et al. ⁴⁷, 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 Paganelli et al. ⁴⁸, 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 are then used to automatically infer patient current health status and detect and alert problematic or dangerous situations and events.

47 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)

48 Paganelli, F., Giuli, D.: An ontology-based context model for home health monitoring and alerting in chronic patient care networks. *null* 838-845 (2007)

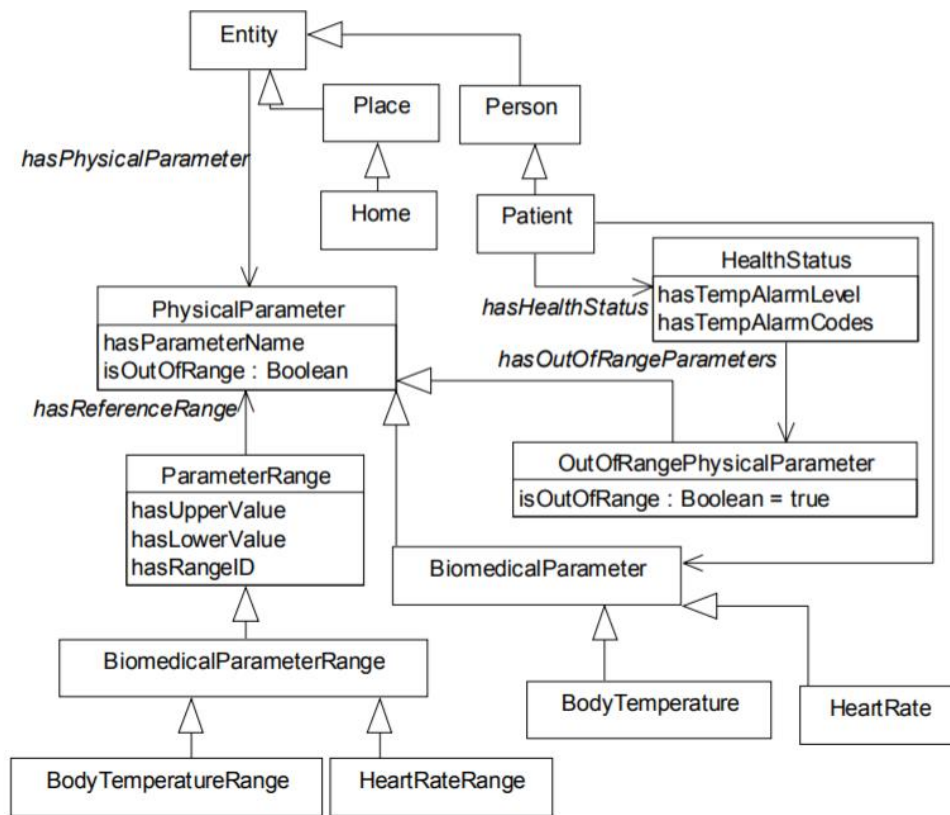


Figure 4 – Patient Personal Domain Ontology, adapted from [24]

3.2.3.2 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 ⁴⁹, 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).

49 Paganelli, F., Giuli, D.: An ontology-based context model for home health monitoring and alerting in chronic patient care networks. null 838-845 (2007)

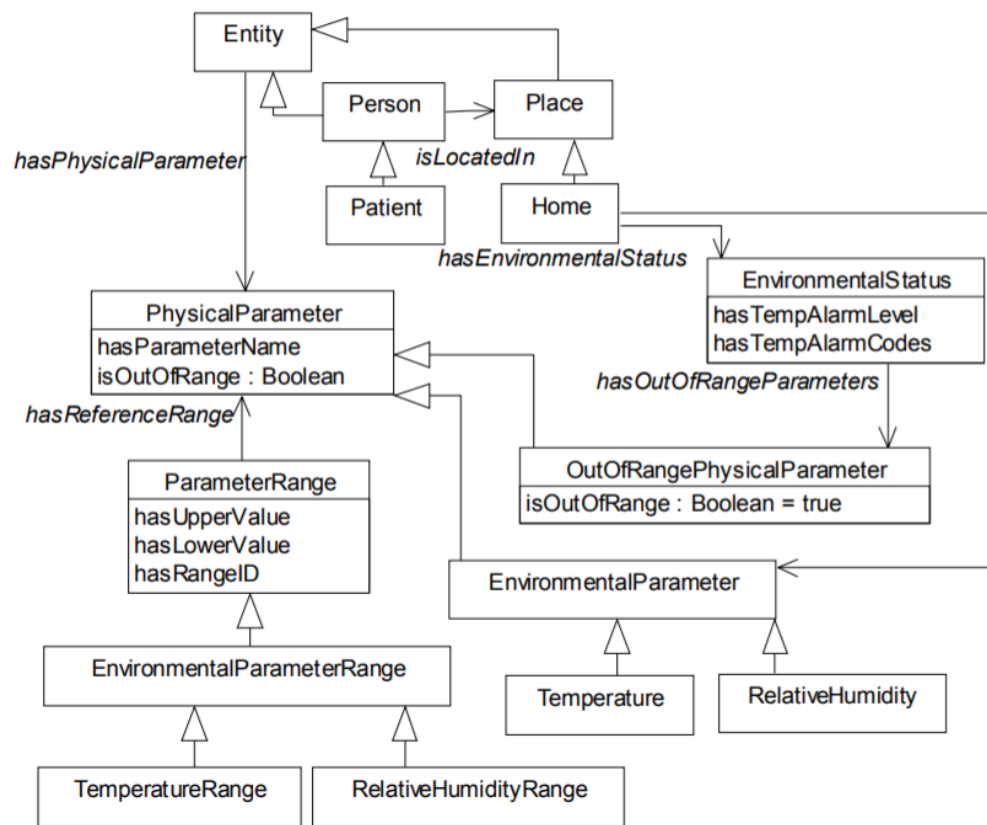


Figure 5 – Home Domain Ontology, adapted from [24]

In ⁵⁰, 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).

⁵⁰ Zografistou, D.: Support for context-aware healthcare in ambient assisted living. Master's thesis (2012)

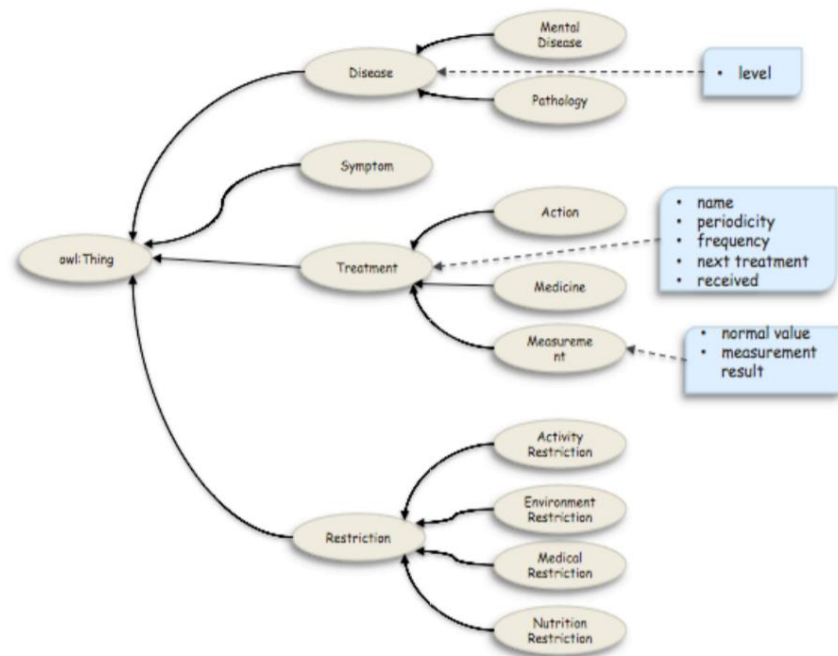


Figure 6 – Health Domain Ontology, adapted from ⁵¹

4 Personalized coaching

4.1 Recommender systems

Coaching solutions will be developed by researching the use of recommendation techniques on top of data that is tracked from employees to support personalised coaching. Such personalised coaching is key in prevention strategies and is lacking in current web-based therapy solutions as well as existing pain management apps that mostly focus on tracking of pain levels/intensity without actionable feedback.

Recommendation techniques are already widely used in online e-commerce or music listening applications. These applications aim to predict and recommend that item that is likely to be relevant to

⁵¹ Zografistou, D.: Support for context-aware healthcare in ambient assisted living. Master's thesis (2012)

the user. Several recommendation techniques, such as content-based⁵², knowledge-based⁵³, collaborative filtering⁵⁴ and their hybridisations⁵⁵, are discussed in several state-of-the-art surveys⁵⁶.

In collaborative filtering (CF), which is the most used and most mature technique, actions of multiple users are compared. Based on similarities with other users (user-based CF) or similarities in content-items (item-based CF) recommendations are made. The most basic recommender technique simply relies on the fact that an item that was valued by many users is probably also valued more by another user. A slightly more advanced CF technique looks for action patterns, and then recommends items that were also preferred by users with similar action patterns. Amazon.com is exemplary for this approach, when suggesting “Users who bought this item also bought that item”. Different algorithms are used to measure similarities between users, patterns and items. However, the technique that is providing the most useful recommendations, often depends on the application domain and the specific set of data at hand⁵⁷.

Content-based filters recommend items that are like other items preferred by the specific user. This kind of filters first built up a user profile, and next, they match descriptions of new items with the past items that were preferred by the user⁵². This approach uses only information provided by the individual user, along with the information of the items available. In contrast with collaborative filters, it ignores information provided by other users in the system.

Knowledge-based filters incorporate knowledge by logic inferences. A knowledge representation is needed to indicate how an item is preferred by a specific user. This knowledge can be provided by domain experts or can be built automatically based on previous cases. Case-based reasoning is often used to extract knowledge out of examples.

Finally, hybrid recommender systems combine multiple filter techniques to increase the accuracy of the recommendation systems⁵⁵. The advantages and disadvantages of the different techniques have been discussed in the literature⁵⁸.

52 M. J. Pazzani and D. Billsus. Content-Based Recommendation Systems. In P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, *The Adaptive Web*, volume 4321 of LNCS, pages 325–341. Springer-Verlag, Springer, 2007.

53 Shari Trewin. Knowledge-based recommender systems. *Encyclopedia of Library and Information Science: Volume 69-Supplement 32*, page 180, 2000.

54 J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J.T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22(1):5–53, 2004.

55 R. Burke. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User Adapted Interaction*, 12(4):331–370, 2002.

56 G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17:734–749, 2005.

57 Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., Duval, E. (2011). Dataset-driven research for improving TEL recommender systems. In : 1st International Conference on Learning Analytics and Knowledge. LAK11: International Conference on Learning Analytics and Knowledge. Banff, Canada, 27 Feb-1 March 2011 (pp. 44-53). ACM.

58 Manouselis, N., Drachsler, H., Verbert, K., Duval, E. (2012). Recommender systems for learning. SpringerBriefs in Electrical and Computer Engineering, The Netherlands: Springer.

The application of recommendation techniques in health scenarios has attracted increased attention in recent years. This year, the third edition of the Health Recommender Systems workshop is organised at ACM RecSys – the leading conference on recommender systems. The objective of the workshop is to “empower people to successfully adopt a healthy lifestyle [...] by offering specific, tailored suggestions for behavioural changes such as nutrition, physical activity, sleep and many more.”⁵⁹ In a recent book chapter, some project partners reviewed existing work on the application of recommendation techniques⁶⁰ in healthcare. Current approaches either target healthcare professionals or patients. Healthcare professionals can benefit from recommender systems to retrieve additional information for a certain case, such as related clinical guidelines or research articles. The second scenario focuses on delivering high quality, evidence based, health-related content to end-users. Most other articles that we have reviewed target patients as end-users. Objectives include delivering relevant information to end-users that is trustworthy, as in the work of Wiesner and Pfeifer⁶¹, lifestyle change recommendations⁶² and improving patient safety⁶³. The latter category for instance includes research on how to use recommender systems to suggest relevant information about interactions between different drugs, in order to avoid health risks. Lifestyle change recommendations focus among others on suggesting users how to improve their eating^{64,65}, exercising or sleeping behaviour. An example is to use recommendations as a basis to algorithmically derive balanced meal plans that meet nutritional guidelines for the user. These recommendations are based on a personal profile (gender, height, weight, physical activity, etc.) and estimate an individual's basal metabolic rate and daily kilocalorie requirements.

Whereas promising ideas and potential applications have been presented in the literature, these application areas also impose new challenges. Recommendations that influence the health status of a user need to be either liable or accompanied by domain experts⁵⁹. To make the recommender liable, complex domain specific user models need to be created. In this project, we will build on research of project partners, e.g., KU Leuven on user interfaces for recommender systems that support **transparency** and **user interaction** with recommender systems to enable end-users to **steer the recommendation**

59 Workshop on Engendering Health with Recommender Systems <https://recsys.acm.org/recsys16/healthrecsys/>

60 Valdez, A. C., Ziefle, M., Verbert, K., Felfernig, A., & Holzinger, A. (2016). Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives. In *Machine Learning for Health Informatics* (pp. 391-414). Springer International Publishing.

61 Wiesner, M., Pfeifer, D.: Health recommender systems: concepts, requirements, technical basics and challenges. *International journal of environmental research and public health* 11(3) (2014) 2580-2607

62 Farrell, R.G., Danis, C.M., Ramakrishnan, S., Kellogg, W.A.: Intrapersonal retrospective recommendation: lifestyle change recommendations using stable patterns of personal behavior. In: *Proceedings of the First International Workshop on Recommendation Technologies for Lifestyle Change (LIFESTYLE 2012)*, Dublin, Ireland, Citeseer (2012) 24

63 Roitman, H., Messika, Y., Tsimerman, Y., Maman, Y.: Increasing patient safety using explanation-driven personalised content recommendation. In: *Proceedings of the 1st ACM International Health Informatics Symposium, ACM (2010)* 430{434

64 Rokicki, M., Herder, E., Demidova, E.: Whats on my plate: Towards recommending recipe variations for diabetes patients. *Proc. of UMAP 15* (2015)

65 Elsweiler, D., Harvey, M., Ludwig, B., Said, A.: Bringing the healthy into food recommenders. *DRMS Workshop* (2015)

process^{66,67}. The overall objective is to increase **user trust** into a recommendation by explaining why a recommendation is made. In addition, user control is researched to incorporate additional input and feedback from both healthcare professionals and employees as a basis **to improve recommendations and to support empowerment**. A coaching editor will be elaborated that enables healthcare professionals to personalise coaching programmes based on data collected from employees. In addition, visualisation techniques will be researched to provide feedback to employees on progress. The overall objective of the latter techniques is to improve adherence to the therapy.

4.2 Visualisation

Information visualisation is a powerful means of making sense of an abundance of data that has emerged from research in human-computer interaction, computer science, graphics, visual design, psychology, and quantitative data analysis. The main intent of information visualisation is to represent an abstract information space in a dynamic way, to facilitate human interaction for exploration and understanding. In recent years, a variety of so-called dashboard applications use information visualisation techniques to support Personal Health Informatics. These dashboards aim to enable users to have better control over what they do, much in the same way that a car helps the driver to not drive faster than the speed limit. Existing health dashboards provide a concise overview of relevant health metrics, such as their weight, blood pressure, sleep quality activities and their environment. Typical representations of these metrics are histograms, bar charts, line charts, traffic lights, and scatterplots. In practice, however, these dashboards are often meaningless as they lack contextual information and are not actionable. Our ambition is to investigate **innovative visual representations** that can help the user **explore the large amount of health-related data** in an interactive way, presenting an overview **combined with actionable information and recommendations for self-improvement**. The visualisation will also help to **understand the link between behaviour and health**. An important objective is to **explain recommendations** to increase user trust and acceptance of these recommendations. These should encourage reflection on their day-to-day living, by creating an interface that will give a concise overview of all individual and relevant metrics, in an actionable way that furthermore supports exploration of various recommendations to improve lifestyle and health.

4.3 Coaching solutions

4.3.1 Psychological approaches

While mass media campaigns are useful in targeting simple behaviour, they are less effective for the complex behaviours that are the hallmark of health habits. Health promotion interventions need to ensure that people gain sufficient motivation and abilities to go through with lifestyle changes and are able to overcome the social and environmental barriers that discourage healthy behaviours. Intervention content and techniques should be personalized to account for the patients' clinical risk factors and biomarkers,

66 He, C., Parra, D., Verbert, K. (2016). Interactive recommender systems: a survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, 56, 9-27.

67 Verbert, K., Parra, D., Brusilovsky, P. (2016). Agents vs. users: visual recommendation of research talks with multiple dimensions of relevance. *ACM Transactions on Interactive Intelligent Systems*, 6(2), 1-42.

existing lifestyles and habits, values, preferences, self-efficacy, awareness, social influences, life situation, environmental conditions etc. (see Figure 1.).

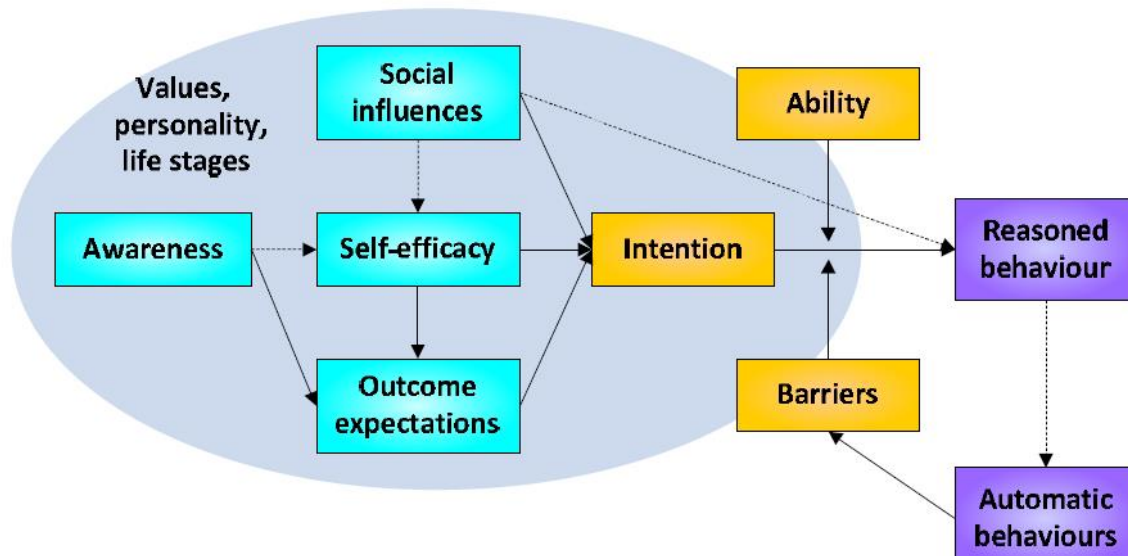


Figure 1. Environmental contexts

Information and communication technology can be developed with an intention to change people’s attitudes and behaviours. So-called behaviour change support systems (BCSSs) provide content and functionalities that engage users with new behaviours, making them easy to perform and embed in their everyday lives.

Recently, wearable technologies have significantly eased the burden of monitoring one’s behaviour. The behaviour change technique these trackers often employ is self-monitoring. The devices and systems used for collecting personal data are referred to as personal informatics and defined as “*systems [...] that help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge.*” However, it is relatively unclear how to create such designs that encourage and enable reflection. Reflection needs time and it is a developmental process that has different levels. Technology solutions may support reflection on these different levels by simply recording events; asking reflective questions; and giving new, multiple perspectives on issues. The Persuasive Systems Design model (PSD) defines seven postulates that would apply for all BCSSs and constitutes a wide applicable tool to define the BCSS and evaluate its persuasiveness. According to PSD resources such as self-monitoring tools, personalization and tailoring of information, contents and services, reminders, offering suggestions, trustworthiness and verifiability of the information, among others, should be considered. Automated support Health BCSSs have been investigated, regarding aspects such as the number of reminders and the feedback presentation to the user. It has been noticed

that automated support messages are often written by the researchers themselves and not based by on behaviour change theories, rather than health care professionals. new references^{68 69 70}

4.3.2 Technology enabled coaching solutions

Technology enabled coaching solutions for healthcare are seen as a very important enabler for establishing cost reductions in the health market. Various platforms exist which integrate communication between professional therapists and patients. Most existing solutions provide an online SaaS-based eHealth platform, enabling patients to manage a personal diary for self-management, to participate in inquiries or to digitally share information with therapists. Typical providers in The Netherlands are Karify, MindDistrict, and VitalHealth e-Vita. These platforms require the patient to take the initiative for providing information, e.g. by filling in online forms.

Other companies focus on the provision of smartphone solutions which use sensor data to enable self-management or personal coaching capabilities. Some solutions provide active engagement strategies: they involve the user by means of a smartphone application, through notifications or by applying other messaging mechanisms. This is the type of solution is provided by Sense Health. Its digital health application Goalie focuses on aspects like mood, sleeping patterns, eating behaviour and physical exercise. Therapists and patients can work jointly on configuring the patient's goals and realize more efficient therapy sessions.

A minority of the currently existing solutions take the step towards predictive analytics based on sensor data in order to provide early warnings to patients and/or therapists. Technologies which combine smartphone-based e-Coaching or mHealth with working predictive analytics and on-the-fly strategy adaptation do not exist at this moment. Other important features, which are not yet provided by existing solutions, are the handling of comorbidity situations, and further integration with social and professional support systems, e.g. through continuous communication capabilities.

4.4 Preventive health solutions

The use of smartphones for the study of physical activity is a recent research field. There are few studies considering the validity of a smartphone-based assessment of physical and behaviour activity. Smartphone applications can be efficacious in promoting physical activity and mobile message interventions may provide benefits in self-management of long-term illnesses. There is a range of opportunities to use smartphones in engaging intervention strategies for disease self-management and to promote healthy practices. The concept of disease self-management aided by information technology

⁶⁸ Kelders, Saskia M. & Oinas-Kukkonen et al Health Behavior Change Support Systems as a research discipline; A viewpoint., *International Journal of Medical Informatics*, Volume 96, December 2016, Pages 3-10.

⁶⁹ Oinas-Kukkonen H. (2013) A foundation for the study of behavior change support systems. *Ubiquit Comput* 17: 1223. <https://doi.org/10.1007/s00779-012-0591-5>

⁷⁰ Oinas-Kukkonen H, & Harjumaa M (2009) Persuasive systems design: key issues, process model, and system features. *Commun Assoc Inf Syst* 2, Article 28:485–500

is a concept, which has been shown to be effective⁷¹ and applied to e.g. smoking cessation⁷² and stress management⁷³.

However, the literature is sparse and modest considering reported intervention effects and so additional studies are needed to better understand the capabilities on the use of smartphones for health prevention.

4.5 Healthcare (and pain) apps

Early intervention plays an important role in preventing pain chronification and, as key influencers in the management of patients with acute pain, it is critical that both the patients and the health professionals are equipped with the necessary awareness, education and skills to manage pain patients appropriately⁷⁴. To prevent chronification of pain, personalised treatment tailored to the pain perceptions and needs of the patient is necessary⁷⁵.

Although multiple pain management applications exist, most of these apps are limited to tracking the level/intensity of pain⁷⁶. Research has shown that several other factors play a role in the chronification of pain⁷⁷, including the perception and cognitions of pain, how people deal with pain, self-efficacy, expectations of recovery or return-to-work⁷⁸, fear⁷⁹ etc. Expectations of recovery have been found as predictor for the decision to return to work. Patients with higher expectations had less sickness absence at the moment of follow-up measurement⁷⁸. Fear of pain has also been hypothesised to result in avoidance behaviour and has been described as an obstacle to recovery in populations of patients with low back pain⁷⁹. Fear of movement has for instance been shown to be linked to long term pain⁸⁰. Such

71 KR Lorig et al. "Internet-Based Chronic Disease Self-Management: A Randomized Trial", *Journal of Medical Care*, 2006

72 Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, Heffner JL. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. *Drug Alcohol Depend.* 2014 Oct 1;143:87–94. doi: 10.1016/j.drugalcdep.2014.07.006.

73 Ly KH, Asplund K, Andersson G. Stress management for middle managers via an acceptance and commitment-based smartphone application: A randomized controlled trial. *Internet Interventions.* 2014 Jul;1(3):95–101. doi: 10.1016/j.invent.2014.06.003.

74 Morlion, B., Coluzzi, F., Aldington, D., Kocot-Kepska, M., Pergolizzi, J., Mangas, A. C., ... & Kalso, E. (2018). Pain chronification: what should a non-pain medicine specialist know?. *Current medical research and opinion*, 1-10.

75 Azevedo, R., Bernardes, M., Fonseca, J., & Lima, A. (2015). Smartphone application for rheumatoid arthritis self-management: cross-sectional study revealed the usefulness, willingness to use and patients' needs. *Rheumatology international*, 35(10), 1675-1685.

76 Lalloo, C., Jibb, L. A., Rivera, J., Agarwal, A., & Stinson, J. N. (2015). There's a Pain App for That. *The Clinical journal of pain*, 31(6), 557-563.

77 Pincus, T., Burton, A. K., Vogel, S., & Field, A. P. (2002). A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine*, 27(5), E109-E120.

78 Kuijjer, W., Groothoff, J. W., Brouwer, S., Geertzen, J. H., & Dijkstra, P. U. (2006). Prediction of sickness absence in patients with chronic low back pain: a systematic review. *Journal of occupational rehabilitation*, 16(3), 430-458.

79 Pincus, T., Vogel, S., Burton, A. K., Santos, R., & Field, A. P. (2006). Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 54(12), 3999-4010.

80 Vlaeyen, J. W., de Jong, J., Geilen, M., Heuts, P. H., & van Breukelen, G. (2002). The treatment of fear of movement/(re) injury in chronic low back pain: further evidence on the effectiveness of exposure in vivo. *The Clinical journal of pain*, 18(4), 251-261.

factors are, however, not monitored in existing apps, and are therefore not available for predictions of absenteeism and coaching apps.

In addition, current apps mostly focus on tracking data and creating an overview for the caregiver. However, few applications also analyse the tracked data to provide additional insights for end-users. Such analysis is needed for giving timely and relevant information, support, strategies and personalised coaching⁷⁶.

As indicated by the systematic review of pain management apps by Lalloo et al.⁷⁶, “*there continues to be **major limitations** in the field of patient-focused pain self-management apps. Namely, current pain apps are characterized by a lack of: (a) integrated features allowing for personalisation through goal-setting, self-monitoring, self-care skills development, education, and social support; (b) involvement of healthcare professionals in development and evaluation; (c) foundation in current research or behavioural theories; and (d) scientific evaluation either through feasibility or effectiveness testing. This last finding is in agreement with the results of three recent systematic reviews of e-health and m-health interventions, which failed to identify any large-scale randomized controlled trials that evaluated the effectiveness of pain apps on health outcomes.*”

In addition, although several applications have been developed and tested in pilots, there are several barriers that hinder the adoption of these apps into real health care practice, including low usability and lack of transparency on the reliability and trustworthiness⁸¹. There is also a lack of standardisation for interoperability between various parts of the systems, including different monitoring components that track behaviour of end-users⁸².

In this project, we will research the fundamental building blocks to overcome these barriers and enable adoption of personal health applications in real practice. A first building block is the use of monitoring and analytics techniques that measure a more comprehensive set of factors that play a role in prediction of absenteeism related to pain. A second building block are recommendation that provide actionable feedback to employees: in contrast to many existing apps to simply present some data statistics, the objective is to present feedback to end-users that enables them to change their behaviour. Visualisation techniques will be researched to explain the rationale of these suggestions as a basis to increase trust and user engagement. Such engagement is key to support behaviour change and will be a core focus of the project. Visualisations will also play an important role in presenting behaviour of employees to healthcare professionals as a basis to revise and personalise coaching programmes. Finally, elaborate user studies will be conducted with a large number of employees to gather empirical evidence on the impact of health apps on health and wellbeing.

⁸¹ Kulyk, O., Akker, R., Klaassen, R., & Gemert-Pijnen, L. (2014). Personalized Virtual Coaching for Lifestyle Support: Principles for Design and Evaluation. *International journal on advances in life sciences*, 6(3-4), 300-309.

⁸² K. D. Ahern, M. J. Kreslake, and M. J. Phalen, “What is ehealth (6): Perspectives on the evolution of ehealth research,” *J Med Internet Res*, vol. 8, no. 1, Mar 2006, p. e4.