D4.5.1
Recommendation System

Work package	**WP4 Play Suite Development and Integration**
Task	**T4.5 Play Suite Recommendation System**
Editor	**Lucas Paletta (JRD)**
(co-)authors	**Martijn Niessen (MCR)**
Public / confidential	**public**

Project PLAYTIME

The research leading to these results has received funding from the AAL Programme of the European Union and by the Austrian BMVIT/FFG under the Agreement no 857334, ZonMw (the Netherlands) and funding from Brussels (Belgium). It reflects only the author's view and that the Union is not liable for any use that may be made of the information contained therein.

31/03/2018

PLAYTIME partner			organisation
01	JR		JOANNEUM RESEARCH Forschungsgesellschaft mbH DIGITAL – Institut für Informations- und Kommunikationstechnologien, 8010 Graz
02	FAM		FameL GmbH Steinbruchweg 20, A-8054 Seiersberg
03	SVD		Sozialverein Deutschlandsberg Kirchengasse 7, A-8543 Deutschlandsberg
04	GGZ		Geestelijke Gezondheidszorg Eindhoven en de Kempen Postbus 909, 5600 AX Eindhoven, The Netherlands
05	TIU		Stichting Katholieke Universiteit Brabant, Tilburg University PO Box 90153, 5000 LE Tilburg, The Netherlands
06	MCR		McRoberts BV. Raamweg 43, 2596 HN The Hague, The Netherlands
07	MBY		MindBytes F. Roosseveltlaan 348-349, B8, 90600 Ghent, Belgium
08	GEU		Ghent University Sint-Pietersnieuwstraat 25, 9000 Gent, Belgium

Contents

1 Executive Summary

The recommendation system for the PLAYSUITE serious game and playful training as well as for the interactive mat is outlined in this PLAYTIME Deliverable in terms of the potential for recommendation for personalization, such as, adjusting the difficulty level.

A concept is described for the structuring of a rule basis for the recommendation system.

The implementation plan is referring to the multiple stages of the development of the recommendation system in the frame of the second field trial and the final demonstrator at the end of the project.

2 Objectives of the recommendation system

The objective of the project PLAYTIME is to motivate dementia users to enter a positive feedback cycle of periodic training with sensors that enable indicative diagnostics on a daily basis, and to receive recommendations on the basis of these data that propose more personalized and better suited exercises for improved training.

User feedback in terms of continuous measurements about sensorimotor coordination and gaze behavior will provide diagnostics as a basis to determine personalised recommendations with the objective to optimise the user experience. Therefore the recommendation system is determined from analytics for the purpose of service personalization.

The integration of multimodal analytics, such as, from psychosocial, sensorimotor, gaze and emotion analytics, considers the definition of a state of human well-being, mental processes and performance and will provide the basis for diagnostic analytics as well as underlie the rule base of the recommender system (PLAYTIME project Task 4.5).

One subtask is to define the dementia relevant features for state definition features and a further subtask is to implement the interface to the front-end – for the parametrization of the rule base of the recommendation system - and back-end systems – for the archiving of the current recommendation parameters for further use.

3 Multimodal information based recommendations

3.1 Related work

3.1.1 Artifical intelligence

Recommender systems have recently obtained great success as an intelligent information system (Adomavicius & Tuzhilin, 2005). Content-based filtering approaches utilize a series of discrete characteristics of an item to recommend items with similar properties. In recent years, contexts, tags and social information have taken recommender systems (Ricci et al., 2011) into account.

Personality aspects were recently considered to personalize recommendations and enhance both recommendation quality and user experience (Nunes, 2009; HU & Pu, 2010). An important aspect of recommender systems is the sequential, long-term aspect of interaction and learning of recommendation strategies (Shani et al., 2005). The system learns the optimal strategy autonomously by observing the consequences of its actions on the users and also on the final outcome of the recommendation session (Mahmood & Ricci, 2007).

3.2 Recommender systems in mental health

Not much previous work on applying recommender system in health informatics or medicine exists. As of June 5th 2016, according to the review of Calero Valdez et al. (2017), only 17 articles are found when searching for the terms \recommender system health" in web of science. The oldest article is from 2007 and the most cited article has only 14 citations. Wiesner and Pfeifer (2014) distinguish between two scenarios: the first scenario targets health professionals as end-users of health recommender systems.

The second scenario targets patients as end-users. Health 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 Pfeifer (Wiesner & Pfeifer, 2014), lifestyle change recommendations (Farrell et al., 2012) and improving patient safety (Roitman et al., 2010). 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 (Rokicki et al., 2015; Elsweiler et al., 2015), exercising or sleeping behavior.

In their research statement Fernandez-Luque et al. (2009) argue, that using recommender systems for personalized health education does not take advantage of the increasing amount of educational resources available freely on the web. As one reason, difficulties in finding and matching content is given. In a short review on health recommender systems by Sezgin & Ozkan (2013) provided at the EHB 2013, the authors emphasize the increasing importance of Health Recommender Systems (HRS). The authors argue that these systems are complementary tools used to aid decision making processes in all health care services. These systems show a potential to improve the usability of health care devices by reducing the information overload generated from medical devices and software and thus improve their acceptance.

The 2018 ACM Conference on Recommender Systems conference[1] is featuring a workshop on engendering health with recommender systems[2], where many of the topics from this topic will be discussed.

3.3 Implementation objective in PLAYTIME

PLAYTIME will particularly focus on the analysis of correlations in the data streams of the diagnostic toolbox and derive rules for the recommendation of useful exercises, such as, cognitive training or physical activities.

In particular, the recommender engine will be applied to control, in particular, the difficulty level of the cognitive tasks in an adaptive way.

Furthermore, JRD will implement a simple model of human motivation and analyse emotional and performance based data towards an estimation of a current or enduring level of motivation of the PwD.

In the following, aspects that provide a basis for a rule base, are discussed , and summarised with respect to the integration into the recommender system in Sec.3.7.

3.4 Recommendations from movement information

3.4.1 Personalization

3.4.1.1 Difficulty levels and accent

The difficulty levels of the movement oriented Playtime components can be based on the MoveTest or MoveMonitor outcomes.

The SPPB is chosen in Playtime as this contains 3 common activities in which people with dementia can experience difficulties (see D3.2.1). The scores of the SPPB on these 3 activities used to determine or evaluate difficulty levels of the Playtime components. The following difficulty levels can be chosen per activity (balance, gait or chair rise):

- Low: score of 0 or 1
- Medium: score of 2 or 3

[1] https://recsys.acm.org/recsys18/healthrecsys/

[2] https://healthrecsys.github.io/2018/

- High: score of 4

The amount of exercises offered to a subject can also be based on these levels. The better a subject scores, the less exercises he/she will be offered on this topic. In this way, the focus will be on the area where a subject has most difficulty. If a subject scores similar on 2 or all activities, the focus will be shared.

For the MoveMonitor (physical activity in daily life) there is currently no standard way of distributing subjects over difficulty levels. It is therefore suggested to use a similar approach as the SPPB: The data of the field trials in Playtime can be used as reference data. The outcome parameters can be divided into four quartiles (Q1-Q4). In this way, a 0-4 scale can be accomplished, comparable to the SPPB:

- Lower output than the subjects in Q1: 0 points
- Comparable output to subjects in Q1: 1 point
- Comparable output to subjects in Q2: 2 points
- Comparable output to subjects in Q3: 3 points
- Comparable output to subjects in Q4: 4 points

3.4.1.2 Base line

During Playtime, the MT (Short Physical Performance Battery) and MM are used to assess base line physical activity. In the game (both personal and group game), physical activity is incorporated by games and videos, as well as discussion scenario's and questions. The difficulty level of the movement components and assignments can be based on the (sub) scores of the SPPB and the physical activity in daily life (e.g. amount of steps).

3.4.2 Use and re-evaluation

3.4.2.1 During the use of PLAYTIME

Subjects play the game and during the game, feedback from the users can be asked on how they perceive the physical activity games, video's, discussions and questions. When a subject experiences the physical activity components as too easy, the level can be increased. If a subject is already in the highest level, the number of repetitions can be increased. When a subjects experiences difficulty and is already in the lowest difficulty category, the amount of repetitions can be lowered.

The time a subject takes to perform a task can also be used to increase or decrease difficulty. If a subject gets better (and thus faster) in certain tasks, the difficulty can be increased or other tasks (with the same focus) can be offered.

3.4.2.2 Multi-modality

During PLAYTIME, output of the other components can be used as well. The feedback from the serious game scenario's that are discussed with the (informal) caregiver can be used as input to increase or decrease the focus of PLAYTIME on physical activity, based on the needs of the subject.

3.4.2.3 Re-evaluation

At certain time intervals during the use of Playtime, the MoveTest and MoveMonitor can be used to objectively re-evaluate the appropriate difficulty levels of the game. These scores can not only be used to adjust the difficulty levels of the subject, but can also be used to assess the performance of the recommender engine, as proposed under section 1.3.3 and 1.3.4.

3.5 Recommendations from attention information

3.5.1 Personalization

3.5.1.1 Difficulty levels

In the MIRA attention games, PwD are engaged in gaze based control of serious games with narratives that aim at involving the interest of the PwD. There are several parameters that define the difficulty level of the MIRA games:

- Antisaccade game:
    - The eye movement features, such as, dwell time on the virtual agent cue, determine the triggering of the activities of the virtual agents. The threshold on the time needed to dwell on the cue in order to invoke a triggering event is related to difficulties of concentration and also how much the triggering event accidentally could be invoked by the PwD. A current guideline would recommend a dwell time of max. 500 ms to trigger.
    - The number of antisaccade tests that configure a single game – currently, 10 tests – has an impact on the difficulty of the game, since more tests means an increased duration the PwD has to be concentrated on the task.
- Spot-the-difference
    - The level of visual detail that is necessary to compare two otherwise identical images determines the difficulty level of this game.
    - The number of visual errors to be found determines the difficulty level of this game since increasing number of visual detail to be searched correlates with the duration the PwD has to be concentrated on the task.

In summary, based on the recommended difficulty level, the MIRA attention game engine is able to parametrise the difficulty of the game experience.

3.5.1.2 Baseline

During PLAYTIME, the MIRA attention games are used to assess base line cognitive activity of the PwD. In the playful training scenarios, the difficulty level of the cognitive components and assignments can be based on the (sub) scores of the eye movement features received from the eye tracking component. However, currently the eye tracking component is exclusively active during the MIRA attention game suite. SPPB and the physical activity in daily life (e.g. amount of steps).

3.5.1.3 Mental health defficiency

Based on the data collected in the second field trial, correlations between eye movement features from the attention games will be performed with the outcome measures of specific tests in the neuropsychological test battery.

We expect to form an estimator on the mental state of the PwD from these correlations by applying machine learning on the time-synchronized data and from this provide outcome measures like probability of ‘good’, ‘medium’ and ‘low’ level of mental state.

These classifications will be applied with the support of social science experts from Mental health Center Eindhoven and Sozialverein Deutschlandsberg.

The output of these estimators could be used, most probably on a week or month basis, to recommend a certain play mode in a forthcoming time period on the basis of attention experience.

3.5.2 Adjustment

3.5.2.1 Feedback and direct adjustment

Parameters of the attention game that were described in Sec. 3.4.1.1 could be adjusted directly by the PwD user. For example one could imagine to offer an interaction opportunity (i.e., buttons) to select the acceptance of the current game level by the user herself and consequently adjust the level either to a higher or lower degree of difficulty.

3.5.2.2 Multi-modality

The impact of the resulting analytics from other components, such as, the amicasa playful training or the socio-emotional serious game SERES; could as well trigger a change of difficulty level on the MIRA suite. For example would a negative experience in amicasa due to a high cognitive load when playing challenge cognitive games trigger a decrease of the difficulty level in MIRA so that the overall cognitive load would not be excessive in the same training session.

3.6 Recommendations from psychology and health care

There are several observations from mental health caregivers that will be taken into account for calibrating a first baseline performance. Firstly, positive emotions from success and pleasant game experiences always seem to support the motivation of the PwD to play, and then, also increase the self-esteem and the interest in further playing. However, there is also the observation that PwD tend to prefer to play more and more easy games and finalise in a state of complete under-stimulation that finally will not provide a ground to progress in terms of the objectives of the playful training intervention. Therefore, the difficulty level needs to be balanced on a daily basis with respect to several reasons. Firstly, the mental capacity as well as the motivation of the PwD is changing from day to day, and even during one single day. Secondly, a challenging level of play needs to be maintained in order to (i) nurture the further interest of the PwD who wants to have success in a challenging task, and (ii) keep up the stimulation for the PwD and therefore follow the guidelines of the intervention suite.

From a psychological point of view, positive emotions only serve to some limit for further stimulation, motivation and cognitive performance (see PLAYTIME Deliverable D3.1.2 for details). Even negative emotions seem to nurture better cognitive performance, however, only within a certain bandwidth and not exceedingly beyond it. Very negative emotions seem to hinder any cognitive performance level, while slightly negative emotions seem to support top-down attention performance which is highly relevant in cognitive control.

In summary, the parametrization on the basis of the recommender has to consider a balance between emotionally positive and negative user experience, as well as react to frequent changes in the personal perception of the difficulty of the game.

3.7 Integrated mental health recommender system

Figure 1 depicts a schematic sketch of dependencies between PLAYTIME components, analytics and the recommender system. The analytics engines of the individual components generate meta-features that contribute to a Hi / Mid / Lo classification with respect to the cognitive / mental state, the physical and the socio-emotional state. Depending on these configurations, the rule base in the recommender will trigger respective parametrisations according to a behavioral matrix for providing a recommendation to the professionist and the tuning of the individual components.

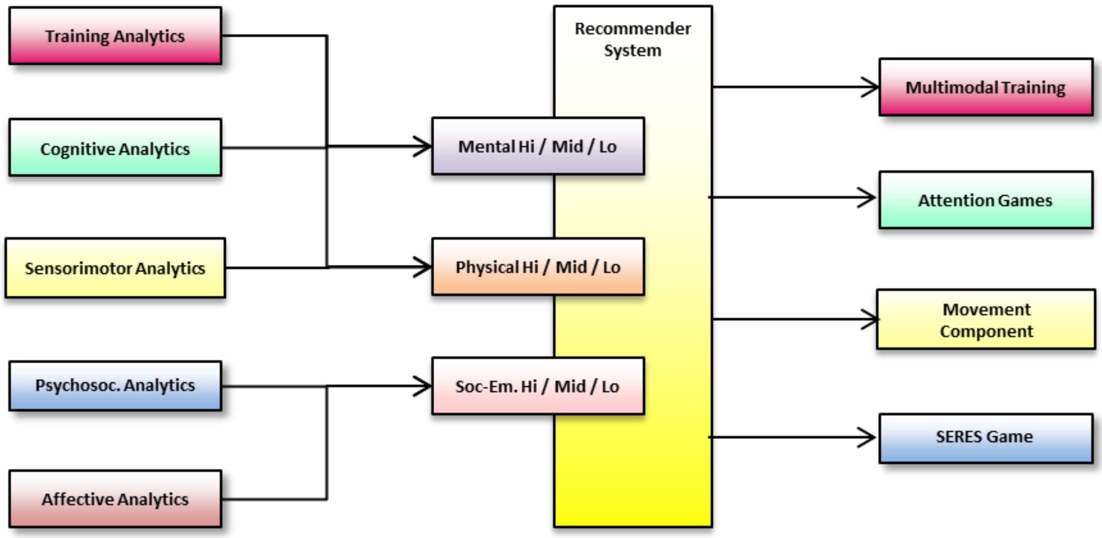

Figure 1. Schematic sketch of dependencies between PLAYTIME components, analytics and the recommender system. The analytics engines of the individual components generate meta-features that contribute to a Hi / Mid / Lo classification with respect to the cognitive / mental state, the physical and the socio-emotional state. Depending on these configurations, the rule base in the recommender will trigger respective parametrisations according to a behavioral matrix for providing a recommendation to the professionist and the tuning of the individual components.

4 Implementation plan

4.1 Prototype for the field trials

The prototype of the recommender system will include an interface that reads out relevant data from the central database (CD), referring to the user profile and actual data from the various PLAYTIME suite components. Most probably it will refer to the amicasa training component, i.e., referring to the performance in past weeks, and fro this recommend to switch to a lower or higher level of difficulty in the amicasa component.

4.2 Advanced prototype based on field trials

The final version of the recommender system will incorporate an interface that will include estimates about the cognitive, mental, sensorimotor and socio-emotional status of the user within previous weeks. It will be promising to argue from the data generated during game play that

5 Conclusions and Outlook

The objective of PLAYTIME Task 4.5 is to implement a recommendation system that enables personalised playful experience and training being adaptive to daily changes in mood and performance. This Deliverable D4.5.1 scanned the conditions under which PLAYTIME suite components could contribute to the recommender system and how they would benefit from a rule base that would decide upon various levels of difficulty in operating the individual component.

In the following Deliverable D4.5.2 we will outline the concrete interface to the central database and the rule base that would map to recommendations in those components that would be related to the rule base in the prototype for the second field trial.

6 Abbreviations

Table 1. *Abbreviations.*

Abbreviation	Description
PwD	Person with dementia

7 Bibliography

(**Calero Valdez, 2017**) André Calero Valdez, Martina Ziefle, Katrien Verbert, Alexander Felfernig, Andreas Holzinger (2017). Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives, in, Holziner, A., Machine Learning for Health Informatics, pp. 391-414, Springer, Heidelberg.

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

(**Farrell et al., 2012**) 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.

(**Fernandez-Luque et al., 2009**) Fernandez-Luque, L., Karlsen, R., Vognild, L.K.: Challenges and opportunities of using recommender systems for personalized health education. In: MIE. (2009) 903-907.

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

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

(**Sezgin & Ozkan, 2013**) Sezgin, E., Ozkan, S.: A systematic literature review on health recommender systems. In: E-Health and Bioengineering Conference (EHB), 2013, IEEE (2013) 1-4.

(**Wiesner & Pfeifer, 2014**) 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