A Demonstration of 4D Digital Avatar Infrastructure for Access of Complete Patient Information

Project acronym: MyHealthAvatar

Deliverable No. 6.4

Data reasoning utilities for decision support & evaluation report

Grant agreement no: 600929

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	MyHealthAvatar
Project Full Name:	A Demonstration of 4D Digital Avatar Infrastructure for Access of Complete Patient Information
Deliverable No.:	D6.4
Document name:	Data reasoning utilities for decision support & evaluation report
Nature (R, P, D, O)[1]	R

[1] **R**=Report, **P**=implementation, **D**=Demonstrator, **O**=Other

Dissemination Level (PU, PP, RE, CO)[2]	PU
Version:	Final
Actual Submission Date:	08/03/2016
Editor: Institution: E-Mail:	Youbing Zhao BED Youbing.Zhao@beds.ac.uk

ABSTRACT:

This deliverable reports deliverable 6.4 – Data reasoning utilities for decision support & evaluation report

KEYWORD LIST:

Data validation, event extraction, information summary, knowledge extraction, lifeloop generation, place annotation

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 600929.

---

[2] **PU**=Public, **PP**=Restricted to other programme participants (including the Commission Services), **RE**=Restricted to a group specified by the consortium (including the Commission Services), **CO**=Confidential, only for members of the consortium (including the Commission Services)

MODIFICATION CONTROL			
Version	**Date**	**Status**	**Author**
1.0	**15/12/2015**	**Draft**	Prof. Feng Dong
2.0	**20/02/2016**	**Draft**	Hong Qing Yu, Po Yang, Zhikun Deng, Farzad Parvinzamir, Youbing Zhao, Xu Zhang
3.0	**08/03/2016**	**Final**	Dr Youbing Zhao

List of contributors

- Youbing Zhao (BED)
- Po Yang (BED)
- Zhikun Deng (BED)
- Hong Qing Yu (BED)
- Xu Zhang (BED)
- Feng Dong (BED)

Contents

1 Executive Summary

This document reports deliverable 6.4 Project background

Owing to the highly fragmented health systems in European countries, gaining access to a consistent record of individual citizens that involves cross-border activities is very difficult. MyHealthAvatar is an attempt at a proof of concept for the digital representation of patient health status. It is designed as a lifetime companion for individual citizens that will facilitate the collection of, and access to, long-term health-status information. This will be extremely valuable for clinical decisions and offer a promising approach to acquire population data to support clinical research, leading to strengthened multidisciplinary research excellence in supporting innovative medical care.

MyHealthAvatar will be built on the latest ICT technology with an aim of engaging public interest to achieve its targeted outcomes. In addition to data access, it is also an interface to access integrative models and analysis tools, utilizing resources already created by the VPH community. Overall, it will contribute to individualized disease prediction and prevention and support healthy lifestyles and independent living. It is expected to exert a major influence on the reshaping of future healthcare in the handling of increased life expectancy and the ageing population in Europe. This complies with the priority and strategy of FP7 ICT for healthcare, and constitutes a preparatory action aiming at the grand challenge on a "Digital Patient", which is currently the subject of a roadmap in the VPH community.[3]

The MyHealthAvatar project focuses on research and demonstration actions, through which the achievability of an innovative representation of the health status of citizens, named 4D MyHealthAvatar, is explored. The 4D Avatar is anticipated as an interface that will allow data access, collection, sharing and analysis by utilizing modern ICT technology. It is expected to become the citizen's lifelong companion, providing long-term and consistent health status information of the individual citizen along a timeline representing the citizen's life, starting from birth. Data sharing will be encouraged, which will potentially provide to an extensive collection of population data to offer extremely valuable support to clinical research. The avatar will be equipped with a toolbox to facilitate clinical data analysis and knowledge discovery.

MyHealthAvatar can be described as a personal bag carried by individual citizens throughout their lifetime. It is a companion that will continually follow the citizen and will empower them to look after their own health records. This fits very well into the recent trend of developing patient-centred healthcare systems.

1.1 Data reasoning in MyHealthAvatar

This deliverable D6.4 - Data reasoning utilities for decision support & evaluation report – is a report of T6.5 Data reasoning (PM7=>PM33) in WP6 Data & repositories (PM2-PM36). This task deals with data reasoning based on the linked data in the RDF data repository, which help to discover inexplicit relationships between data. The data reasoning provide necessary supporting data for clinicians in diagnosis process and citizens in decision making for health related issues. Together with semantic reasoning, the work can be linked to visual data analytics in WP8 to show the usage of the data in decision supporting process. The object of the deliverable is to provide a detailed report of the work of the data reasoning tasks in MyHealthAvatar. This report describes the work of MyHealthAvatar

[3] MyHealthAvatar project, Description of Work (DoW) document.

data analysis on both the web and mobile platform, including data validation, event extraction and ranking, information summary and recommendation, knowledge extraction and reasoning, etc.

.

2 Introduction

2.1 Overview

Due to the rapid growth of wearable devices and mobile apps, it has become increasingly possible to remotely monitor a patient or citizen's health by connecting heterogeneous medical devices into a platform. A promising trend in healthcare fields appears that the sensor enabled technology is transforming traditional hubs of healthcare, such as clinics and hospitals, to personalized healthcare systems and especially mobile environments. Continuing monitor patients' conditions outside the hospital environment enables future healthcare to be delivered faster, safer and at lower cost, with enhanced sustainability.

MyHealthAvatar proposes and implemented a unified platform for collecting and storing the heterogeneous health data from a variety data sources, including wearable sensors, mobile apps, EHRs, etc. However, data repository alone is not the hallmark of MyHealthAvatar as there are also other health data repositories such as Microsoft HealthVault [HealthVault]. What makes MyHealthAvatar special is the integrated data analysis and visualization capability.

In this report the detailed features of both the MyHealthAvatar web and mobile platforms on data analysis is introduced, including data validation, event extraction and ranking, information summary and recommendation, knowledge extraction and reasoning, etc.

The work shows that MyHealthAvatar is not only a powerful health data repository but also a versatile data analysis platform. It provides data collection tools as well as data analysis tools for both the citizens and clinicians in knowledge acquisition and decision making on health related issues. With the visual analytics tools introduced in D8.2, more potential of the data analysis capability can be revealed and utilized which leads to more efficient and capable insight gaining.

2.2 Relevant Tasks

The related and reported task is T6.5 Data reasoning (PM7=>PM33) which deals with data reasoning based on the linked data in the RDF data repository and helps to discover inexplicit relationships between data. The data reasoning provide necessary supporting data for clinicians in diagnosis process and citizens in decision making for health related issues. Together with semantic reasoning, the work can be linked to visual data analytics in WP8 to show the usage of the data in decision supporting process.

3 Data Validation

3.1 Introduction

Technically and functionally sophisticated wearable devices such as Fitbit[Fitbit], Withings [Withings], iHealth[iHealth], Jawbone Up [JawboneUp], Nike+ [Nike+] and mobile applications such as Moves[Moves], Endomondo[Endomondo] can record a variety of longitudinal personal health information; including physical activity, weight, sleep quality, heart rate, blood pressure, etc. Among this data, physical activity is mostly well-observed due to the maturity of microelectromechanical systems (MEMS) based accelerometer technology as well as easily and openly accessible Global Position System (GPS). Numerous research works and commercial products have attempted to accurately monitor longitudinal physical activity and access activity patterns and intensity level, by using either dedicated wearable sensors or advanced machine learning algorithms. But these studies mostly consider performance optimization of single sensor or a combination of GPS and accelerometer by analyzing raw sensors' signals. In sensor-enabled personalized healthcare systems, physical activity data is mostly daily basis from globally heterogeneous third party devices. Traditional physical activity validation methods hardly deal with these scattered and heterogeneous data. Also, due to diversity and rapid change of personal life patterns and environmental impacts, personal physical activity data in e-healthcare systems has remarkable uncertainties. Effective validation of the physical activity data from heterogeneous devices is an essential but demanding task. The requirements of customization and longitudinal study in an e-healthcare environment make this task ever harder. This paper investigates the problem of effectively validating personal physical activity in a heterogeneous devices based internet enabled personalized healthcare environment. A general rule based adaptive physical activity validation model is proposed for eliminating irregular uncertainties and estimating physical activity data reliability in sensor enabled personalized healthcare systems. It enables data validation procedure in internet environments to be a dynamic standardized empirical analysis workflow with four layers including factors, methodologies, knowledge and actions. The factors impacting the validity of physical activity are categorized into device, personal and geographic. Each factor defines a longitudinal data analysis based investigation strategy. The validation rules are represented with a set of uncertainty threshold parameters and reliability indicators, which can be initiated by historical data and adaptively updated regarding the needs of a sensor enabled personalized healthcare system. To demonstrate the effectiveness of the model, a case study on MyHealthAvatar platform [MHAWeb] are carried out. The results reflect that the validation rules and action criteria delivered by the model effectively improve the validity of physical activity. The main contributions are below:

1. A rule based adaptive physical activity validation model is proposed for eliminating irregular uncertainties and estimating physical activity data reliability in Internet enabled personalized healthcare systems.

2. A series of validation rules representing with uncertainty threshold parameters and reliability indicators are defined and evaluated to improve the validity of physical activity data. The validation rules are capable of being adaptively and dynamically updated regarding the needs of an Internet enabled personalized healthcare system.

3.2 Related Work

As a major risk measure for chronic diseases, daily physical activity recognition and monitoring with wearable sensors have been investigated by a number of researchers. In [Paradiso2005, Parkka2006], authors carry out a study on recognizing and classifying physical activity by analyzing signal features from 3D (triaxial) accelerometers on hip and wrist and GPS data with a hybrid classifier of custom decision tree and neural networks. The results are reported a classification accuracy up to 89% for detecting 10 daily actions. ProeTex [Ghasemzadeh2011] project develops an algorithm that combines features of ECG and triaxial accelerometer in smart garments for detecting nice classes of physical activity with overall classification accuracy up to 88.8%. In [Curone2010, Atallah2011], researchers have integrated on-body sensors in a wireless network for the purpose of activity recognition and lifestyle monitoring. Authors in [Curone2010] utilize a network of five accelerometers to classify a sequence of 20 daily activities with accuracy of 84%. The system in [Atallah2011] that uses seven different sensors embedded in a single node, including microphone, phototransistor, 3D accelerometer, 2D compass, barometer, ambient light and digital humidity, to classify 12 movements with accuracy up to 90%. The outstanding achievement of all aforementioned work on daily physical activity recognition is high classification accuracy of recognizing multiple daily activity actions. But all of these studies rely on a collection of physical activity data as a raw accelerometers' signals. In internet based personalized healthcare systems, physical activity data comes mostly from globally heterogeneous third party devices. The traditional classification methods are infeasible to handle these scattered and heterogeneous physical activity data.

Recently, many commercial wearable products and mobile applications have been released for the long term record and collection of personal health information, particularly on physical activity. The most famous mobile apps, such as *Moves* [Moves], are based on smartphone 3D accelerometer data and GPS information which allows tracking user movement activities including location, distance and speed. The wearable products, such as *Fitbit Flex* [Fitbit], *Nike+ Fuelband* [Nike+], *Withings* [Withings], are all wristband devices that record steps count, distance, and calories burnt. These wearable devices communicate with mobile phone via Bluetooth employing relevant mobile applications. While above products have been proven its popularity among general users, their majority usages are limited in the fitness fields. It is due to diversity of life pattern and environmental impacts; personal physical activity data from individual wearable device exhibits remarkable uncertainty. The validating of these physical activity data in longitudinal healthcare cases is very challenging. Also, as the exponential growth of mobile healthcare market, numerous similar wearable products have been developed, which will significantly increase the heterogeneity and diversity of devices connected in internet enabled personalized healthcare systems. Effective validation of physical activity data from heterogeneous devices in internet enabled personalized healthcare environments becomes more difficult.

3.3 MyHealthAvatar Physical Activity Validation (PAV) Model

3.3.1 Model Ecosystem

The ecosystem is the theoretical cornerstone of validating of physical activity in an internet environment. Personal health data are accumulated and measured as a cube in three dimensions (3D): *Persons*, *Devices* and *TimeLine*. The increment in any dimension results in an expansion of the health data grid. The products like *Fitbit* [Fitbit] or *Moves* [Moves] occur on a 2D plane (Persons × TimeLine), which refer to scenarios that single device is used by increasing population over time. Similarly, physical activity recognition with sensor fusion appears on a 2D plane (Devices × TimeLine) for classifying individual person's activities with historical health data. To distinct from the above two categories of studies, the target of PAV model is a cube of rapid-growth historical raw health data (Physical Activity).

Ideally, the workflow of PAV model for validating physical activity is a dynamic recurrence by duration along the timeline. The validation rules are initiated by feeding a set of historical raw physical activity data in PAV model; and then are used to validate the current physical activity. After a period, historical raw physical activity data is expanded with more users or devices over time. The validation rules have to be dynamically changed and updated by feeding new historical physical activity data into the model. Also, PAV model provides a configuration to register the information on person and devices dimensions. It adaptively supports the need from different users or groups.

The idea of PAV model is to identify the influencing factors with detailed issues causing uncertainty of physical activity; and design a series of benchmarks and experimental study methods for qualitatively evaluating these influencing factors. Through these experiments, PAV model enables delivering a practically efficient validation strategy containing a set of validation principles, rules and actions. Fig.3 shows a conceptual diagram of PAV model. Three main objectives of PAV model are:

Data Validity and Reliability: PAV aims at providing an effective validation strategy containing a set of validation rules for an Internet healthcare system, which enables producing validated personal physical activity measurements of demonstrated quality with quantifiable uncertainties.

Genericness and Adaptivity: PAV is a generic conceptual model for supporting a variety of heterogeneous devices in Internet enabled healthcare environments. It will not be limited by certain type of wearable devices or mobile applications. The validation rules of model have to be adapted to fit to the Internet healthcare application situations.

Extendibility and Scalability: PAV has to be extendible and scalable for supporting emerging technological possibilities of devices in an Internet healthcare environment. New unidentified influencing factors can be added in the PAV model and investigated with a similar evaluation methodology.

3.3.2 Uncertainty Classification

PAV model is built upon a theoretical classification of influencing factors leading to uncertainty of physical activity data by specifying four layers and three components in an Internet healthcare environment. The uncertainty of physical activity here is categorized into two types:

Irregular uncertainty: Irregular Uncertainty (IU) in physical activity data occurs randomly and accidently. The causes of these uncertainties may include device malfunctions or faults, breakdown of third party server, misuse of mobile apps, sudden change of personal circumstance. The occurrence of irregular uncertainty in physical activity data will appreciably impact the efficiency and accuracy of assessing personal health. PAV model establishes a set of rules to identify and eliminate these uncertainties.

Regular uncertainty: Regular Uncertainty (RU) in physical activity data occurs frequently and persistently. The causes resulting in these uncertainties are mainly from some regular influencing issues, like intrinsic sensors' errors, differentiation of personal physical fitness and changes of environment. The occurrence of regular uncertainty in physical activity data is inevitable so that it is impossible to completely eliminate these uncertainties. PAV model is supposed to establish a set of rules to manage these regular uncertainties for reducing their impacts on utilizing physical activity data.

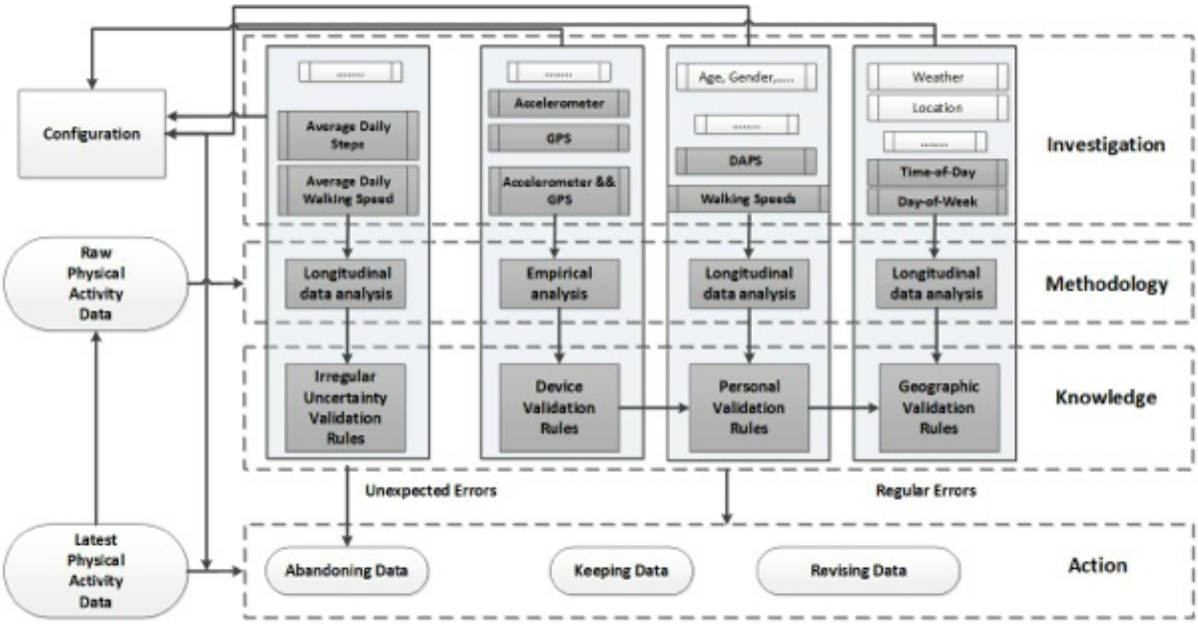

Figure 1 Diagram of PAV model

3.3.3 Impacting Factors Analysis and Matrix

While irregular uncertainties occur accidently and are hardly quantified by impacting factors, their occurrence frequency is relatively low over time. A statistical analysis in historical data can detect threshold parameters to filter them. Daily physical activity is mainly measured as daily steps (S_d), daily walking distance (D_{dw}) and daily average walking speed (V_{daw}) as it is shown in Table 1. It is believed that the majority of daily steps and daily average walking speed have to be in a specific range. Two threshold parameters (T_s and T_v) are defined to filter the irregular uncertainties regarding a probabilistic distribution.

For regular uncertainties, the impacting factors in PAV are categorized into three modules, which are device factors, personal factors and geographic factors. In the device factors module, existing popular wearable devices or mobile apps are classified by sensory technique into three types: GPS based, Accelerometer based, a combination of sensors based. The accuracy of these three sensory techniques for measuring step count and distance are quantified by Mean of relative error and Standard Deviation of relative error though a series of experiments.

The personal factors module studies if the differences of human demographic, anthropometric and fitness data give regular uncertainties to physical activity data. These differences usually include the age, gender, height, weight and medical history, etc. The information relies on users' efforts of manual input, which maybe incomplete. There is a need for a benchmark to represent a person's physical fitness from completed data sources. Here a walking speed related score is defined to represent a person's physical fitness, named as Daily Activity in Physical Space (DAPS). This score is inspired from work [Herrmann2011] that proposes a Movement and Activity in Physical Space score as a functional outcome measurement for encompassing both physical activity and environmental interaction. Currently, most third party APIs of wearable devices or mobile apps have provided the functions to assess the intensity of physical activity regarding walking speed. For instance, *Fitbit* [Fitbit] classifies the intensity of daily activities into Very Active, Moderately Active, Lightly Active and Sedentary; *Moves* [Moves] records a series of walking segments containing duration, distance and speed. Here, we classify the intensity of daily physical activity into *N* levels in terms of the ranges of walking speeds (V_1, V_2 …V_n). The DAPS formula is created by summing these different level walking speeds:

$$DAPS = \sum_{1}^{N} V_t \tag{1}$$

The geographic factors module aims at investigating the impacts of location specific information related contextual data on the accuracy of daily physical activity. This information can include Time (time of day, life events, i.e.), Location (country, part of city, "at work" etc.), Environmental factors (weather conditions, etc.). Considering the difficulty of establishing and recording completed user life and environment profiles, we only list three items in geographic factors: weather, hourly-change of physical activity, and weekly-change of physical activity. The changes of daily physical activity over these three issues are measured with statistical analysis in historical data. A few range and type of parameters are defined in Table 1. A reliability indicator (R) for estimating the overall impact of above three impacting factors is formulated below:

Table 1 Listing of Parameters and Indicators in PAV Model

		Parameters	*Descriptions*
Raw Physical Activity Data		S_d	Daily walking steps
		D_{dw}	Daily walking distance
		V_{daw}	Average daily walking speed
		C	Confidence interval for filtering historical data distribution
Irregular Uncertainty		T_s	Threshold parameter for filtering incorrect daily steps data
		T_v	Threshold parameter for filtering incorrect average daily walking speed
Regular Uncertainty	**Devices**	ES_mean	Mean of step count relative error
		ES_std	Standard Deviation of step account relative error
		ED_mean	Mean of measured distance relative error
		ED_std	Standard Deviation of measured distance relative error
	Personal	DAPS	Daily Activity in Physical Space score
		V_1, V_2 …V_n	Average Walking Speed regarding intensities of daily physical activity.
	Geographic	Sh(morning, afternoon, night)	Daily steps range in morning, afternoon and night
		Dh(morning, afternoon, night)	Daily walking distance range in morning, afternoon and night
		Swk(working, weekend)	Daily steps range in working days and weekend
		Dwk(working, weekend)	Daily walking distance range in working days and weekend
Reliability Indicator		D	Reliability dependent on device factors
		P	Reliability dependent on personal factors
		E	Reliability dependent on geographic factors
		R	Reliability Indicator for estimating physical activity data

$$R = D \times P \times E \qquad (2)$$

Where:

D: Reliability of device factors on physical activity

P: Reliability of personal factors on physical activity

E: Reliability of geographic factors on physical activity

3.3.4 Data Validation Strategy

Data validation strategy of the PAV model aims at conducting a set of validation rules for eliminating irregular uncertainties and reducing the impacts of regular uncertainties on physical activity data. This strategy is designed by using a combination of statistical analysis methods on longitudinal historical data and experimental analysis approaches. The workflow of data validation strategy is presented as 4-layers structure in Figure 1.

Investigation Level: it provides analysis and classification of detailed influencing items in each impacting factor module, also establishes corresponding uncertainty measurement matrix. A notable feature of influencing items level is extendibility that may add more items into the PAV for further investigation.

Methodology Level: it designs a set of investigation approaches for each impacting factor module regarding identified items and established matrix. The investigation approaches include statistical longitudinal data analysis and experimental based empirical analysis methods.

Knowledge Level: it conducts a series of validation rules and principles following the investigation approach. These rules and principles aim at quantitative removal of irregular uncertainties, and qualitative exploration of the relationship between impacting factors and regular uncertainty.

Action Level: It contains the options of executed actions on physical activity data regarding validation rules. Three main types of actions are given in the model: to abandon data, to keep data and to revise data. The main purpose of PAV model is to validate and verify physical activity data, so the action of revising data is not considered in this paper.

Following the four layers described above, the steps of data validation strategy in the PAV model are described below:

For removing irregular uncertainty:

1.To configure the information related to impacting factors and collect certain type of raw historical physical activity data.

2.To calculate the parameters S_d, $D_{dw,}$ V_{daw} with raw data.

3.To plot the data of S_d, $D_{dw,}$ V_{daw} in line and calculate the value of T_s and T_y to cover data with a confidence interval of 95%.

4.To use T_s and T_y for removal of irregular uncertainty physical activity data.

5.To circulate the above process in another period with updated raw data.

For device factors:

1.To list and classify typical wearable devices and mobile applications for physical activity data recording.

2.To design a set of evaluation experiments including daily activities, such as walking for measuring accuracy parameters of the devices: *Es_mean*, *Es_std*, *Ed_mean* and *Ed_std* (see Table 1 for definitions).

3.To conduct the experimental findings as validation rules and establish the equation for device reliability indicator D.

4.To circulate the above process with new types of devices.

For personal factors:

1.To calculate the value of V_1, V_2 …V_N with raw historical physical activity data by individual person.

2.To calculate the value of DAPS by summing up V_1, V_2 …$V_{N.}$

3.To calculate the value of Pearson Correlation between DAPS and S_d or D_{dw} by individual person.

4.To conduct the experimental findings as validation rules and establish the formula for personal reliability indicator P.

5.To circulate the above process with more subjects.

For geographic factors:

1.To classify and categorise physical activity data regarding weather, hourly-change and weekly-change parameters.

2.To plot the data of S_d, $D_{dw,}$ V_{daw} in line and calculate the range value of parameters to cover a confidence interval of 95%.

3.To conduct the experimental findings as validation rules and establish the formula for personal reliability indicator E.

4.To circulate the above process in another period with updated raw data.

3.3.5 Adaptability and Extendibility

The design of PAV model aims at generic utilization in Internet enabled personal healthcare systems. Two important features of PAV model are adaptability and extendibility.

Adaptability: Configuration is defined here in the PAV model for registering the information regarding devices factor, personal factor or geographic factor. By using this information, PAV model can adaptively adjust the values of parameters in validation rules to account for different requirements. PAV model is also able to adapt itself efficiently; it is fast in responding to changed settings or needs in an Internet enabled healthcare environment.

Extendibility: The impacting factors and investigation methods defined in PAV model are all capable of being extended and scaled to support some emerging technological possibilities of the devices or new unidentified influencing factors in an Internet healthcare environment. The data validation strategy can be utilized into an extended environment.

3.4 Experimental investigation

In theory, PAV model can be used for evaluating and validating physical activity in an Internet healthcare environment with any population, for any devices and at any time periods. We use two EU healthcare projects: MyHealthAvatar [MHA] and CARRE [CARRE] as case studies to verify PAV model. This section presents the establishment of validation rules with PAV model by MyHealthAvatar and CARRE projects. The evaluation of device factors modules include seven physical activity recorders used in CARRE project: *Fitbit Flex* (entitled as *Flex*) [FitbitFlex], *Fitbit One* (entitled as *One*) [FitbitOne], *iHealth AM3* (entitled as *iHealth*) [iHealth], *Medisana Vifit Connect* (entitled as *Medisana*) [Vifit], *Withings Pulse O2* (entitled as *Withings*) [Withings], *Jawbone UP24* (entitled as *Jawbone*) [JawboneUp], and *Moves* (entitled as *Moves*) [Moves]. The evaluation of irregular uncertainty, personal and geographic factors are based on MyHealthAvatar platform, which is an Internet enabled personal healthcare experiment platform connecting *Moves*, *Fitbit* and *Withings*. This platform enables user to transfer their physical activity data from these third party providers into MyHealthAvatar server, and then to be able to visualize and analyse this information for a better user understanding and experiences.

3.4.1 Irregular Uncertainty

Eliminating irregular uncertainties is the primary step of data validation strategy in PAV model. On MyHealthAvatar platform, we initially collect daily physical activity (Steps, Distance and Calories) of 7 users over 6 months by 3 types of wearable devices of recorders (Withings, One and Moves). All these 7 users (1 female and 6 male) are researchers in university, and their ages are in the range of 30-50 years old. The features of this raw activity data are:

- All 7 people use Moves, 2 of them additionally use Withings, and another 3 people use Flex.
- Missing data occurs frequently in Withings and Flex, because users easily forget wearing them.
- Some data in Flex shows lower steps, which is probably because users take off their wearable devices some time, or devices are out of battery.

Moves data are more completed than Flex or Withings, but with relatively high errors.

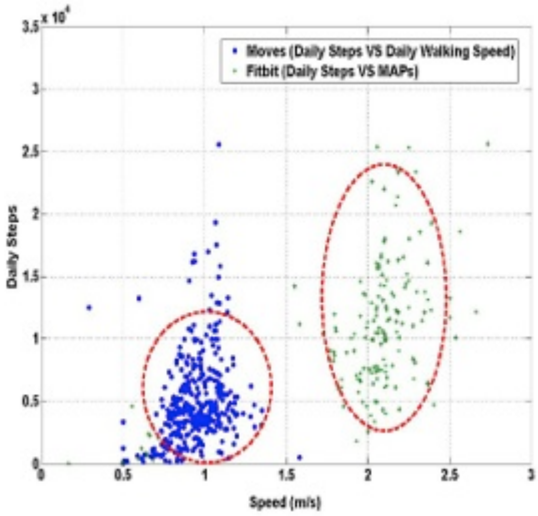

Figure 2 Distribution of Irregular Uncertainty

We calculate V_{daw}, and plot S_d and V_{daw} in 2D diagram as in Fig.4. In order to measure T_s and T_y to remove irregular uncertainty physical activity data, we use an eclipse equation (3) to cover 95% of data (C = 0.95).

$$\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1 \quad (3)$$

Where:

h : Average daily walking speed

k : Average daily walking steps

a : Error range of average daily walking speed

b : Error range of average daily walking steps

A noticeable issue here is that we only consider the lower limits of walking steps and the upper limits of walking speeds as threshold parameters. On some days users might walk distinctly more steps than usually, while the other days might be more sedentary. The threshold parameters are represented in equation (4):

$$\begin{aligned} T_y &= h+a \\ T_s &= k-b \end{aligned} \quad (4)$$

Rules:

- Daily steps of individual by Moves are about 4000 – 7000,
- Flex or Withings give daily steps about 6000 – 13000.
- Moves gave a lower measurement of daily steps than Flex or Withings on the same condition.
- Normal people should have a daily steps in the range 1000– 20000.
- Flex and Withings sometimes show daily steps below 1000.

- Following equation 4, we can get Ts = 68, and Ty = 0.56 for Moves, and Ts = 1329, and Ty = 1.67 for Flex.

3.4.2 Device Factor

The characteristic evaluation of wearable devices and apps accuracy is the second step of the PAV model. In this chapter, we present design and results of experimental investigation carried out in order to evaluate the accuracy of wearable equipment. A total of 6 devices (as mentioned before) were included in this study: *Flex, One, iHealth, Vifit, Withings, Jawbone*. All these devices are classified as an "accelerometer only" based physical activity trackers. They were chosen from the market as the suitable devices for long term physical activity monitoring due to low price, long battery life, compatibility with Android and the most importantly – API availability. The *Moves* app was included in the evaluation as it is the only piece of equipment employing both GPS and accelerometer technology with available API. Two more apps were included in the study as the "GPS only" equipment: Endomondo [Endomondo] (entitled as Endo), Google MyTracks [GoogleMyTracks] (entitled as Tracks). The same main criterion – API availability– was applied when choosing the GPS enabled apps.

The study was performed in two stages: the primary and final investigations. In both parts, some of the physical activity parameters available from the selected devices were measured on healthy volunteers and compared to the reference parameters. Some of the devices are suitable to wear on the wrist, others – on the waist or in the pocket and some provide the ability to choose how to wear them. The wrist wearing site was preferred during the experimentation since it allows easy and unobtrusive non-stop physical activity tracking. In the primary investigation, three variables were measured – steps taken, distance travelled and calories burned. All accelerometer based devices output these three parameters, while Moves outputs only the step count and the distance and the "GPS only" apps output only the distance and the calories. The reference method for measuring the step count consisted of raw accelerometer signals acquired by the custom physiological and kinematical signal recorder KTU BMII Cardiologer v6 [Gargasas2012] attached to the waist, and a semi-automatic peak detection algorithm implemented in Matlab. The reference method for measuring calories was indirect calorimetry implemented in portable calorimeter Cosmed K4b2 [COSMEDK462]. Since this calorimeter is enabled with a GPS module, it also was used as a reference method for measuring travelled distance. Four healthy volunteers participated in this part. The aim of this primary study was to define preliminary accuracy / error ranges for selected commercial devices. The experimentation protocol was the following for each participant:

A short walk within fixed distance of 160 m (80 meters back and forth with stopping) where only the step count was measured. "GPS only" devices were not included.

1. Calculation of the average step length using the distance and the step count from the reference method.
2. Update of the devices with personal information, such as birth date, height, weight, step length, running step length.
3. The approximate of 1000 meters long casual walking exercise via fixed rounded route. The participant was able to choose his/her own walking pace.

4. Jogging exercise of 200 m (100 m back and forth without stopping). "GPS only" devices not included.
5. Slow walking exercise of 200 meters (100 m back and forth without stopping). "GPS only" devices not included.
6. Stair climbing exercise (5 floors). "GPS only" devices not included.

The protocol can be summed up in two parts. One part includes the most frequent physical activity – walking (exercises 1, 4 and 6). The other part includes less frequent physical activity (exercises 5 and 7). The results from this primary evaluation are also divided into two parts respectively. The error ranges for each type of devices are presented in Table 2 and Table 3.

Table 2 The error ranges for walking exercises

	Error range (min – max), %		
Device type	**Steps**	**Distance**	**Calories**
Accelerometer	0,0 – 82,5	0,1 – 68,1	0,2 – 93,3
Accelerometer + GPS	4 – 56,4	N/A	N/A
GPS	N/A	0 – 5,4	2,4 – 45,8

Table 3 The error ranges for less frequent exercises

	Error range (min – max), %		
Device type	**Steps**	**Distance**	**Calories**
Accelerometer	0,0 – 74,6	0,7 – 72,4	6,4 – 80,6
Accelerometer + GPS	6,9 – 94,2	N/A	N/A
GPS	N/A	N/A	N/A

Even though the Moves app output the distance information, it was not accurately recorded by the operator. Therefore this data was discarded from the investigation. The GPS devices data was not acquired during the less frequent exercises. Although these results present only the preliminary error ranges of the devices, they create some guidance for further experimentation. The calories estimation from the accelerometer devices shows the worst performance, while accelerometer + GPS devices do not output such information at all. The Cosmed K4b2 calorimeter system is also very complicated for the participants to work with. Therefore the calories estimation comparison was excluded from the further experimentation. On the other hand, the GPS devices showed very good performance in measuring distance. It was decided to replace the reference GPS device with the GPS enabled app in the smartphone (Tracks). In order to simplify the exercises in the experimentation and due to some limitations (GPS not working inside the building), it was decided to exclude the less frequent exercises from the experimentation. The reference method for counting steps remained the same as in the primary investigation. 6 healthy volunteers participated in the second investigation. Consequently, the new simplified experimentation protocol was established as the following:

1. A short walk within fixed distance of 100 m (50 meters back and forth with stopping) where only the step count was measured.
2. Calculation of the average step length using the distance and the step count from the reference method.
3. Update of the devices with personal information, such as birth date, height, weight, step length, running step length.
4. The approximate of 1000 meters long casual walking exercise via fixed rounded route. The participant was able to choose his/her own walking pace. Step count and distance was measured.

The first short experiment shows the ability of the devices to accurately capture short episodes of physical activity (e.g. walking in the office). The long walk experiment shows the ability to accurately record the most frequent daily physical activity – casual walking (e.g. walking to/from work). The results as a mean of error and the STD of error are presented in Table 4 for each device and each measured variable separately.

Table 4 The accuracy of the devices

Device	Error in steps 100 m		Error in steps 1000 m		Error in distance 1000 m	
	Mean	STD	Mean	STD	Mean	STD
Fitbit Flex	-6,6%	17,7%	-8,5%	14,2%	-6,6%	26,3%
Fitbit One	0,2%	1,5%	0,0%	0,4%	-4,9%	8,2%
iHealth	-11,4%	19,9%	-0,8%	2,4%	-8,1%	6,4%
Vifit	-10,3%	11,7%	-2,8%	5,6%	-9,2%	4,3%
Withings	-1,3%	2,0%	-0,6%	2,0%	5,1%	9,8%
Jawbone	-7,8%	14,7%	4,7%	11,3%	-7,2%	20,5%
Moves	-7,2%	25,2%	-0,2%	3,0%	-5,6%	1,4%

These results show that devices based on the same accelerometer technology perform differently and could not be used interchangeably. It may seem that the wrist wearing site can cause problems as the Flex tracker has lower accuracy than One. On the other hand, we can see that Withings performs similarly to the One while also worn on the wrist. The error ranges were updated according to the results of the final investigation and are presented in Table 5.

Table 5 The updated error ranges for walking exercises

	Error range (min – max), %	
Device type	**Steps**	**Distance**
Accelerometer	0,0 – 47,5	1,0 – 41,2
Accelerometer + GPS	0,0 – 37,1	3,8 – 7,4

We can see that the actual ranges are lower than in primary investigation. Another observation is that Accelerometer + GPS devices have slightly lower error range for step count and significantly lower error range for distance estimation.

We propose that the device reliability factor should be separately calculated for each of the measured parameters. In this particular case with two parameters, the following two equations are introduced:

$$D_s = 0.5 \cdot \left[\left(1 - ES_{mean,100}\right) \cdot \left(1 - ES_{mean,1000}\right) \right] + 0.5 \cdot \left[\left(1 - ES_{STD,100}\right) \cdot \left(1 - ES_{STD,1000}\right) \right] \quad (5)$$

$$D_d = 0.5 \cdot \left[\left(1 - ED_{mean,1000}\right) \right] + 0.5 \cdot \left[\left(1 - ED_{STD,1000}\right) \right] \quad (6)$$

Where:

D_s : reliability of step counting for physical activity devices;

D_d : reliability of distance estimation for physical activity devices;

$ES_{mean,100}$: mean of error in step count in 100 m walk;

$ES_{mean,1000}$: mean of error in step count in 1000 m walk;

$ED_{mean,1000}$: mean of error in distance estimation in 1000 m walk;

$ES_{STD,100}$: STD of error in step count in 100 m walk;

$ES_{STD,1000}$: STD of error in step count in 1000 m walk;

$ED_{STD,1000}$: STD of error in distance estimation in 1000 m walk.

Then, the overall reliability of the device can be formulated as a combination of these separate reliability indicators (7):

$$D = \sum_n kD_n \quad (6)$$

Where:

D : overall reliability of the device for physical activity;

D_n : reliability of one of the parameters;

k : weight of each parameter reliability.

The calculated reliability factors (with the weight k = 0,5) are presented in Table 6. We can observe that One is the most reliable while Withings shows only slightly lower performance. The only GPS + Accelerometer equipment Moves performs similarly to Accelerometer only trackers worn on the wrist.

Table 6 The reliability factors of the devices

Device	D_s	D_d	D
Flex	0,781	0,879	0,830

One	0,990	0,968	0,979
iHealth	0,830	0,860	0,845
Vifit	0,853	0,896	0,874
Withings	0,971	0,964	0,968
Jawbone	0,818	0,891	0,854
Moves	0,826	0,846	0,836

3.4.3 Personal Factor

In terms of the definition of DAPS in PAV model, a person's physical fitness can be represented by a walking speed related score. Moves does not classify the intensity of physical activity regarding the walking speed, so its DAPS is equal to the Average Daily Walking Speed. Fitbit Flex physical activity data has been classified into the intensity of four types as so DAPS and its related walking speeds are measured. Each person has different physical activity characteristics, such as walking speed. The issue here is that individual physical characteristics will impact the accuracy of collected raw data. We measure the parameters like MAX, MIN, AVER and STDEV of 7 users historical raw data, as shown in Table.7. The features of these raw activity data are:

In *Moves*,

- 7 people average walking speed is 0.69 m/s ~ 1.26 m/s
- 7 people average step speed is 1.18 step/s ~ 1.60 step/s
- The figure using Moves segment (minute-by-minute) data is slightly lower than Moves summary (daily).

In *Flex*,

- people DAPS is 1.72 m/s ~ 2.07 m/s
- Active average step speed is 1.30 m/s ~ 1.50 m/s
- Moderate average step speed is 0.48 m/s ~ 5.07 m/s
- Slightly average step speed is 0.14 m/s ~ 0.16 m/s

Each person has different physical activity, but their daily speed or DAPS are in a similar range.

Regarding the international standard of human walking cadence and speed, female walking is roughly 1.95 steps/s in cadence and 1.85 m/s in speed; male waking is about 1.95 steps/s in cadence, and his average speed is 1.43 m/s. It appears that both *Flex* and *Moves* underestimate users' walking speed.

Table 7 Personal Factors Investigation

Moves		P1	P2	P3	P4	P5	P6	P7
Daily Walking Speed (V_{daw}) (m/s)	MAX	0.98	1.19	1.10	1.00	1.50	1.58	1.09
	MIN	0.50	0.29	0.69	0.51	0.69	0.82	0.50
	AVER	**0.68**	**1.00**	**0.99**	**0.85**	**1.26**	**1.09**	**0.84**
	STDEV	0.14	0.13	0.10	0.10	0.16	0.17	0.19
Walking Cadence (steps/s)	MAX	1.86	1.85	1.78	1.44	1.95	1.84	1.82
	MIN	0.67	1.13	1.12	0.82	1.35	1.13	0.67
	AVER	**1.24**	**1.54**	**1.50**	**1.18**	**1.60**	**1.53**	**1.31**
	STDEV	0.29	0.15	0.16	0.10	0.14	0.14	0.27
Fitbit								
DAPS (m/s)	MAX	2.17	2.18	2.40	1.93			
	MIN	0.17	0.55	0.62	1.82			
	AVER	**1.72**	**1.88**	**2.07**	**1.88**			
	STDEV	0.62	0.28	0.40	0.08			
Active Speed (m/s)	MAX	1.42	1.53	1.82	1.30			
	MIN	1.22	1.10	1.25	1.22			
	AVER	**1.30**	**1.27**	**1.50**	**1.25**			
	STDEV	0.05	0.12	0.13	0.06			
Moderate Speed (m/s)	MAX	0.67	0.65	0.65	0.50			
	MIN	0.33	0.41	0.46	0.47			
	AVER	**0.57**	**0.52**	**0.56**	**0.48**			
	STDEV	0.08	0.05	0.05	0.03			
Slightly Speed (m/s)	MAX	0.18	0.17	0.18	0.15			
	MIN	0.13	0.13	0.13	0.13			
	AVER	**0.16**	**0.14**	**0.15**	**0.14**			

	STDEV	0.01	0.01	0.01	0.001			
Pearson Correlation								
Fitbit Err_p (regular error by personal factors)	**DAPS**	0.736	0.120	0.380				
	Active	-0.159	-0.163	-0.196				
	Moderate	-0.073	0.467	0.092				
	Slightly	0.173	-0.068	0.069				
Moves Err_p (regular error by personal factors)	**DAPS (Daily Walking Speed)**	0.514	0.173	0.497	0.378	0.140	0.033	0.589
	Walking Cadence	0.190	0.120	0.094	0.199	0.112	-0.068	0.437

In order to estimate the impact of individual physical activity, we take an assumption that the regular errors generated by personal factors are supposed to linearly follow the Daily Steps. In other words, if a person walks more steps or distances, regular errors will be reflected more heavily in daily steps, as it is shown in Equation (8):

$$Err_p = \beta \times S_d \tag{8}$$

Pearson Correlation ($\beta = 0.01$) is used for measuring the relationship between DAPS or walking speeds and Err_p, as shown in Table 3. The Pearson Correlation results reflect variability among individual subjects, for instance, in *One* (DAPS vs Err), the physical fitness of Subject P1 may have a strong relationship with irregular errors, which gives a value up to 0.73; but for subjects P2 and P3, this relationship has only a value lower to 0.12. Similarly, in *Moves*, the value of Pearson Correlation differs among subjects in the range 0.173-0.589. So, the findings indicate that differences in physical fitness of personal factors will not generate significant regular errors in physical activity data. The rules are concluded below:

- The Pearson correlation between Daily Speed and Daily Steps for individual is diverse.
- No strong impact of daily speed or MAPS on daily steps. While each subject has different physical activity ability, but their speed or MAPs are within a range, and no correlation with daily steps was observed.
- Personal factors (for normal people) will not generate significant errors in physical activity data.

In terms of above findings, the reliability of estimating personal factors on physical activity can be measured by the difference of individual person's DAPS and a standard DAPS in a group of populations. If it is assumed that *M* subjects' DAPS data is recorded in the platform, the reliability of estimating personal factors on physical activity is formulated below:

$$\overline{DAPS} = \frac{\sum_{m=1}^{M} DAPS_m}{M} \tag{9}$$

$$P = 1 - \sqrt{\left(\frac{DAPS_i - \overline{DAPS}}{\overline{DAPS}}\right)^2} \tag{10}$$

3.4.4 Geographic Factors

The impact of geographic factor on irregular uncertainties is estimated by using empirical analysis methods on observed data of a small group of daily physical activity. We analysed Day-of-Week differences in this dataset including all three devices (*Fitbit One, Moves and Withings*) for both groups and individual. Figure 3 and Figure 4 respectively illustrate the distribution of Day-of-Week difference on group and individual daily physical activity. In Figure 5, the lines of (P1_m,..,P7_m) represent *Moves* users; the lines of (P1_f,…,P3_f) represents *Fitbit One* users; and the lines of (P4_w, P5_w) represent *Withings* users. Also, *Moves* provide time based walking segments data, we conduct the distribution of Time-of-Day difference on group based physical activity in Table 6Figure 6. In Figure 6, the physical activity at certain time-slot in a group of 7 users is summed as Distance, Steps and Durations. The features of this data are:

- For Day-of-Week difference, a similar trend line of group physical activity occurs in three devices. It shows that daily step appears stable in weekdays but decreases dramatically on weekend.
- The trend line of individual physical activity is fluctuated widely, but approximately follows the same trend of group physical activity.
- For Time-of the Day difference, the highest intensity of physical activity occurs from 7 am to 10 am. Then the intensity of physical activity keeps stable and slightly decreases in the Afternoon. At the night from 11 pm-12 pm, the intensity of physical activity increases bit. But it may be because users use their smartphone before sleep.

The rules are concluded below:

- People normally have stable physical activity in working day, but have much less physical activity on Sunday.
- People normally have an intensive physical activity in the morning session (7-10 am), and have moderate physical activity in other time of the day.

The reliability of estimating geographic factors on physical activity can be measured by the difference between individual daily steps and average daily steps in weekdays by devices in Fig.5. If it is assumed that *M* person wears one type device, his / her steps data in weekdays are recorded as Swk^t (t =1,..,7), the reliability of estimating geographic factors on physical activity is formulated below:

$$\overline{Swk^t(t=1,...,7)} = \frac{\sum_{1}^{M} Swk_m^t(t=1,..,7)}{M} \qquad (11)$$

$$E = 1 - \sqrt{\left(\frac{Swk_i^t(t=1,...,7) - \overline{Swk^t(t=1,...,7)}}{\overline{Swk^t(t=1,...,7)}}\right)^2} \qquad (12)$$

Where:

t : represents weekdays from Monday to Sunday. CASE STUDY AND performance EVALUATION

In this section, we discuss the performance evaluation of PAV model in a case study on MyHealthAvatar platform [MHAWeb], which is an Internet based healthcare project. MyHealthAvatar platform enables users to record, store and visualize their multi-dimensional health data by connecting wearable devices or mobile apps, like *Fitbit Flex*, *Moves*, *Withings*, Twitter and Facebook. The criteria of verifying our validation model are: efficiency of data validation, model adaptivity and extendibility.

We collected the empirical dataset by using MyHealthAvatar platform. The dataset includes 1 yearlong daily physical activity information from 14 subjects acquired with three devices: *Moves* was used by 14 users for 9 months; *Flex* was used by 5 users for 12 months; *Withings* was used by 3 users for 3 months. These people are healthy in the age range of 30-50 years. The evaluation methodology for verifying the efficiency of proposed model will interview the participants, and collect feedbacks on reflecting users' experiences on physical activity uncertainties through different devices. The feedbacks are used as a standard benchmark to compare the correctness of model.

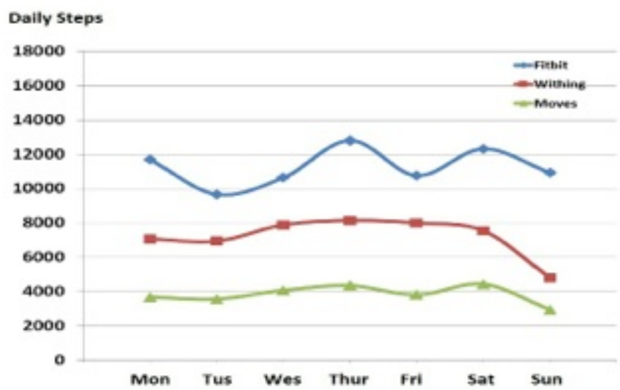

Figure 3 Distribution of Day-of-Week difference on group based daily physical activity

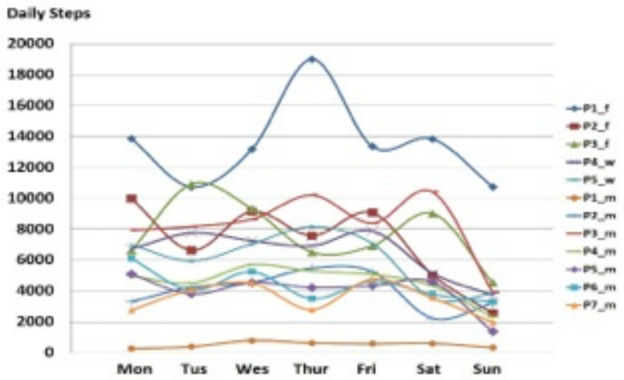

Figure 4 Distribution of Day-of-Week difference on individual based daily physical activity

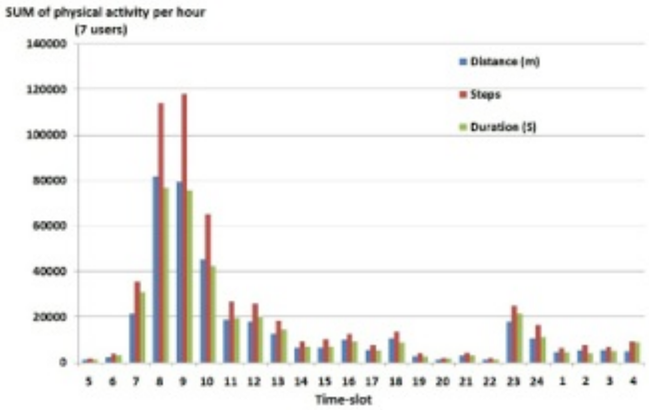

Figure 5 Distribution of Time-of-Day difference on group based physical activity

3.4.5 Data Validation Efficiency

In order to validate the accuracy of identifying IU, we follow equation (4) with a confidence interval of 95% to filter data from three different devices. We use the values (130, 1784, 884) of threshold parameter *Ts* respectively in *Moves, One* and *Withings*, for filtering incorrect daily steps data. The results are shown in Table 8.

Table 8 Removing irregular uncertainties (IU) by PAV

	Moves	Flex	Withings
Ts Daily Steps	130	1784	1267
T_Y DAPS Speed (m/s)	0.5	1.50	NA
Total number of People	14	5	3
Percentage of people with IU	43%	100%	100%
Number of IU occurrence	40	17	8
IU confirmed by User	40	15	5

Average number of IU occurrence per person (User Feedback)	6.6	5.4	2.7
Accuracy of identifying IU (95%)	100%	88.2%	62.5%
Accuracy of identifying IU (98%)	100%	100%	100%

Moves has much lower threshold parameters of Daily Steps and DAPS speed than *Flex* and *Withings* which are 130 and 0.5 m/s respectively (Table 8). This is because *Moves* has larger device uncertainties than *Withings* and *Flex* as we observed in section IV.C. Thus the GPS and smartphone internal sensors based App is not as accurate as accelerometer only based wrist wearable device. In terms of percentage of people having IU, *Moves* is much lower than *Withings* and *Flex*. It is probably because most of uncertainties from *Moves* have been classified into regular uncertainties, so its irregular uncertainties became less than for other two devices *Withings* and *Flex*. However, for average IU occurrence per subject, *Moves* has higher performance than other two devices (Table 8). The accuracy of identifying IU appears that on the condition with a confidence interval of 95%, the related value of threshold parameter *Ts* can successfully filter irregular uncertainty in *Moves*. So *Moves* have the best IU identification accuracy up to 100%, which means that the incorrect daily steps detected by PAV model in *Moves* have been all approved by users. *Flex* and *Withings* have accuracy up to 88.2% and 62.5% respectively, which implies that some correct daily steps are eliminated by PAV model.

If we increase the confidence interval up to 98%, and recalculate threshold parameters, the accuracy of identifying IU of three devices would increase to 100%. But, a noticeable issue here is that if we increase the confidence interval, some IU might be ignored and put into the procedure of dealing with regular uncertainties in PAV model. Similarly, in *Moves*, a high accuracy of identifying IU does not mean all the IU have been removed, probably some of IUs are considered as regular uncertainties in PAV model.

For validating reliability indicator of regular RU, we follow the strategies of PAV model and equations in Section IV to process the above dataset for getting average figures of the group of 14 people. Then we choose the data of one person (P1 in Table 7) who has three devices for estimating reliability indictor. The feedback from this person will assess the efficiency of our proposed reliability indictor. The criteria of interpreting the feedback contains five levels of agreement (Almost perfect, Substantial, Moderate, Fair, Slight). The results are shown in Table 9.

Table 9 Regular uncertainties Indicator by PAV

Reliability Indicator	Moves	Flex	Withings
D	83.6%	83.0%	96.8%
P	87.6%	96.7%	95.6%
E	78.6%	83.4%	87.4%
R	**57.5%**	**66.7%**	**80.9%**

User Feedback	Moderate	Substantial	Almost perfect

Table 9 reflects that using the regular reliability indicator of PAV model, the reliability estimation of collected physical activity data by three devices were approximately following the users' feedback. The data from *Moves* is estimated as reliability of 57.7%, and user believes this data are moderately accurate. The data from *Flex* and *Withings* are both more reliable than *Moves* regarding user's feedback. Especially, *Withings* is recognized by user as "almost perfect", which has a reliability value up to 80.9%. *Flex* is slightly less reliable than *Withings*, it is mainly from the difference of device factors. Above figures imply that the proposed reliability indicator of PAV model can be used as a quantitative analysis tool to estimate the reliability of personalized physical activity data collected from an Internet environment.

3.4.6 Model Adaptivity

For validating the adaptivity of PAV model, we consider the whole group of 14 subjects as one group due to the similar professions and backgrounds. We estimate the change of daily steps T_s and DAPS with different periods (from 1 month to 12 months) with a confidence interval of 95%. The results are shown in Figure 6 and Figure 7.

Figure 6 shows the parameter Daily Steps as the function of duration. The value of this parameter is lower for shorter time periods than for longer time periods. The value of this parameter also varies with different devices. For *Moves* and Withings, the value of this parameter over different periods is slightly growing, but for *Fitbit*, this parameter dramatically increases after 6 months. This effect may be influenced by the setting of confidence interval.

Figure 7 shows little variation of parameter DAPS in the PAV model when time period duration is changed. There are some mirror fluctuations of DAPS on both *Moves* and *Fitbit*. But in a long term, the value of DAPS is quite stable, which indicates that personal physical fitness does not have significant changes within this group of 14 people.

Figure 6 Average of daily steps as the function of time period duration

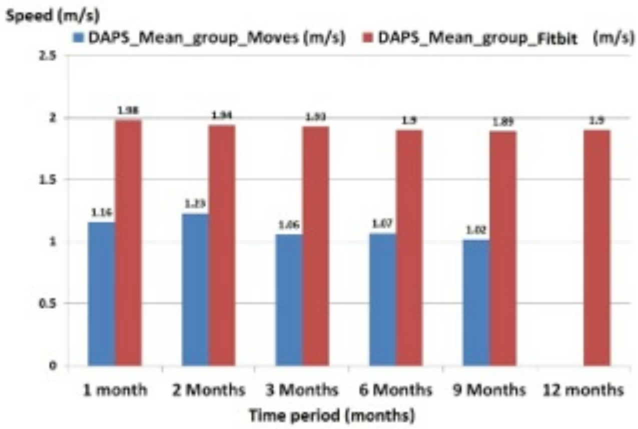

Figure 7 DAPS as the function of time period duration

3.5 Model Extendibility

Extendibility is one important feature of PAV model. PAV can be extended to improve the accuracy of other types of recorded health data related to physical activity.

As it was shown in 3.4.5, the experimental investigation of the devices showed poor performance of calorie estimation. It could be possible to enhance the accuracy of calorie estimation by taking in to account the specific personal characteristics (age, gender, weight, BMI, body fat). However, due to many variables in the measurement of this parameter, more reference data from different participants is required.

We propose to enhance the step count and the distance in the same manner, employing the results from our investigation described in Section IV B. These two measures incorporate less personal characteristic variables, such as only height and step length. The adjustment of the measures could be done as the following example for step count: if the mean of error for *Flex* is +X % and the mean of error for *Moves* is -Y %, the step count from *Flex* should be multiplied by the coefficient k_{Flex}:

$$k_{Flex} = \frac{1}{1 + \frac{X\%}{100\%}} \tag{13}$$

And the step count for Moves should be multiplied by k_{Moves}:

$$k_{Moves} = \frac{1}{1 - \frac{Y\%}{100\%}} \tag{14}$$

The qualitative proof of this adjustment could be accomplished by the feedback of a participant. The quantitative proof could be also possible including two devices for one participant in the experiment and comparing how well their output match before and after the adjustments.

3.6 *Conclusions*

In this section, a rule based adaptive physical activity validation model, PAV, is proposed for eliminating irregular uncertainties and estimating data reliability in an Internet enabled personalized healthcare environment. It specifies four layers and three modules for evaluating the factors impacting the validity of physical activity. The validation rules are represented by defining a set of uncertainty threshold parameters and reliability indicators, which are initiated by historical raw data and adaptively updated regarding the needs of an Internet enabled personalized healthcare system. Following this model, a case study on an Internet enabled healthcare platform MyHealthAvatar connecting three state-of-the-art wearable devices and mobile apps was carried out. The results reflect that PAV model provides an efficient, adaptive and extendable solution for the validation of Internet environment based physical activity data.

4 Event & Activity Extraction and Ranking

4.1 *Activity Extraction*

The input data contains of sequences of locations and activities over time comes from Moves. We use a day event/activity matrix D_{ij}^{d} to record the activities that took place on the day d, where i ranges from 0 to 23 indicating the daily hours and j is one of the activity categories presented by LifeTracker. The matrix is:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,j} \\ \vdots & \ddots & \vdots \\ a_{i,1} & \cdots & a_{i,j} \end{pmatrix}$$

For simplicity, and without loss of generality, we present the locations of the activities in 13 categories as follows:

1 My home
2 My work/study place
3 Shop or shopping area
4 Restaurant, pub, bar, or other eating place
5 Transport
6 Healthcare (GP surgery, dentist, hospital, etc.)
7 Sports and Gyms
8 Entertainment (cinema, theatre, nightclub, museum, etc.)
9 Hotel
10 Governmental and public places (police, council, etc.)
11 Friends' homes
12 Children's school, college, etc
13 Unknown

While this categorisation could be extended, we believe that it represents a good balance between the complexity and usability.

It is relatively straightforward to complete the activity matrix from the input data. We calculate the activities based on their types, start and end time. An activity takes a value between 0-1 in the matrix D_{ij}^{d}. For instance, if the activity started at 09:30 and end at 10:15, then the corresponding entries are 0.5 for hour 9($i = 9$) and 0.25 for hour10 ($i = 10$), respectively.

Based on D_{ij}^{d}, we decide the visible activity at the i hour on day d by

$$VD_{i}^{d} = argmax_{j}D_{ij}^{d}$$

In other words, VD_{i}^{d} is used to calculate and then display the activity at the i hour on day d according to the major activity within the hour. Apart from this primary activity, the other activities are also stored and can be used in the calculation for activity ranking which will be depicted later in section 4.3.

4.2 Activity Summary

In this section the process of generating the activity summary will be portrayed.

Based on D_{ij}^{d}, we generate an overview for month m as follows:

$$M_{ij}^{m} = \sum_{d} D_{ij}^{d}$$

where the variables are as defined in Section 4.1. The visible activity at the i hour in month m is:

$$VM_{i}^{m} = argmax_{j}M_{ij}^{m}$$

Also, by using the above values, we can compute σ_{j}^{m}, which represents the measurement of the $j$activity at month m

$$\sigma_{j}^{m} = \sum_{d} M_{ij}^{m}$$

Similarly, we can work out yearly overviews as:

$$Y_{ij}^{y} = \sum_{d} M_{ij}^{m}$$

and the visible activity at the i hour in year y as

$$VY_{i}^{y} = argmax_{j}Y_{ij}^{y}$$

as well as the measurement of the j activity at year y

$$\sigma_{j}^{y} = \sum_{d} Y_{ij}^{y}$$

These calculations are useful for the visualisation of the activities and their summaries, which will be described in the following section.

4.3 Data and activity ranking

For activity ranking, we firstly rank each activity category j according to their monthly importance as follows:

$$CIMP\_MH_j^m = \sigma_j^m$$

which represents the importance of activity j in month m with a normalisation of:

$$\sigma_j^m = \frac{\sigma_j^m - \min(\sigma^m)}{\max(\sigma^m) - \min(\sigma^m)}$$

Correspondingly, we also rank each activity category j at year y according to its yearly importance, followed by normalising the values:

$$CIMP\_YR_j^y = \sigma_j^y$$

$$\sigma_j^y = \frac{\sigma_j^y - \min(\sigma^y)}{\max(\sigma^y) - \min(\sigma^y)}$$

which represents the importance of activity j in year y. Two lists $< HIMP\_MH >$, and $< HIMP\_YR >$ are formed to record the all the monthly and yearly important hours, correspondingly. These lists are created by scanning each day recorded in D_{ij}^d – if $CIMP\_YR_j^y$ or $CIMP\_MH_j^m$ is above an importance threshold, and D_{ij}^d is non-zero, then we find an important hour and record it in $< HIMP\_MH >$ or $< HIMP\_YR >$.

The ranking is performed via an on-the-fly process and can be repeated in case the data changed or added. The lists $< HIMP\_MH >$ and $< HIMP\_YR >$ include all of the hours during which the important activities occurred.

4.4 Future work and improvements

There will be number of improvements within the importance ranking which enhance the categories and algorithms by restructuring the categories to identify the available categories dynamically and also employing the tf-idf method to determine a better score for the logged events. Each of improvements can be implemented individually (modular) in order to enrich the quality of the final result.

Furthermore, the user preferences will be taken into consideration to tremendously improve the ranking algorithm based on what user define and get the conceivable match. Hence the following terms have been described as the parameters controlled by the user:

[1] W1: decide the importance of the Gradient
[2] W2: decide the importance of frequency
[3] W3: decide the importance of each event category

The overall progress would be as follow:

$$\mathbb{E}_{score} = (1 - w_1) \times (tf\text{-}idf) + (w_1 \times \mathbb{G})$$

where $\mathbb{E}$ is the score, $w_1$is the importance and $\mathbb{G}$ is the Gradient. The $tf\text{-}idf$ will be calculated as follow:

$$tf\text{-}idf = (0.5 + \log(tf))^{w_2} \times \log\left(\frac{\mathrm{N}}{df}\right)$$

Where the N is the period, the df is the occurrence of the term activity/place, and tf is the term frequency of each detected event. The Gradient $\mathbb{G}$ will be calculated as follow:

$$\mathbb{G} = \sqrt[2]{\sum_{i=0}^{n} (x_i - \overline{x})^2}$$

The above algorithms will support the importance ranking to be more accurate and also more intimate to the user's expectancy.

5 Information summary and recommendation on the MyHealthAvatar app

Since the MyHealthAvatar project is a proof of concept for the digital representation of patient's health status, it collects and shares the long term and consistent personal health data of its users through an integrated digital platform. Therefore, in order to achieve a better end-user experience, MyHealthAvatar app provides the health information, summary and recommendation as the output features for the healthcare functions.

5.1 Information Summary

The information summary aims to give a summary to users, in order to let the user check their progress towards the achievement of their goals. The app currently supports four different types of information summary reports such as daily, weekly, monthly goals summary, weekly activity summary and 12 weeks reduce weight program summary.

Daily, Weekly and Monthly Summary: Those summary reports are used to summaries the information which compares set goals to actual data. The implementation time is on 9:10 pm by default, and users can set a custom time to trigger this information summary report.

For the weekly and monthly summary reports, the app would check whether the current date is the last day of the week or the last day of the month, in order to trigger the specific summary report. If the given date satisfies both cases, then the system will trigger two different summary reports.

The detail of the summary items includes steps, distance, calories-burn, activity per mins, weight, BMI, blood pressure systolic, blood pressure diastolic and main location as shown in Figure 8.

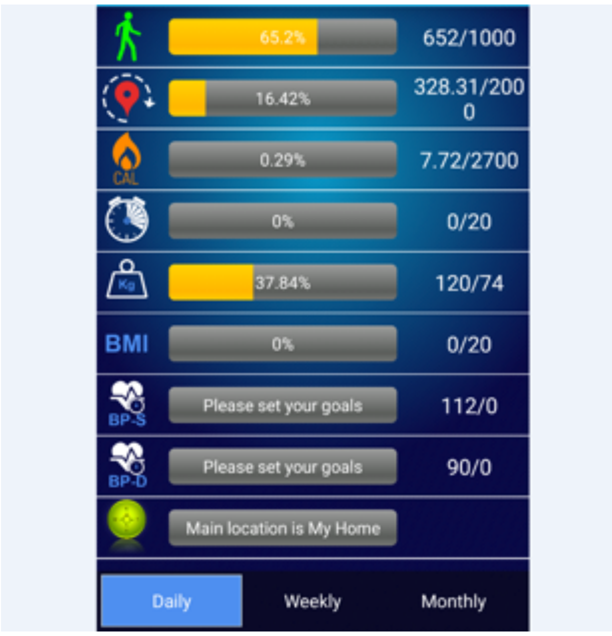

Figure 8 Daily, Weekly and Monthly Summary

There are four steps to generate the summary.

1. Once the information summary has been triggered, the system will generate the date range for the summary.
2. By using the given date range, the system will search in a local database and find out the dates D_1 with no records in the database. The reason why D_1 is important to find is to save the internet resource. The local database may have already stored a period of time as a record. The given date range may not be fully covered by the time range in database, thus to find the value of D_1 can save time and internet resource when retrieved the data from the API service.
3. Using Async-task to download the record from API service with the date D_1.
4. Calculate the average achievement for every item and display the detail in app. If the goal value is zero which means that user has not set the goal for the item, the system then sets the special text to remind the user to set the goals.

Main Location: the main location aims to provide a personal life pattern review to the users. According to the historical activity data, the app can calculate the main place which user would spent longest time during a given date range. Since the sleep event is not belong to a life pattern considering, the app will calculate the time that users have spent on sleep.

There are four steps to generate the daily, weekly and monthly main locations as shown in Figure 9.

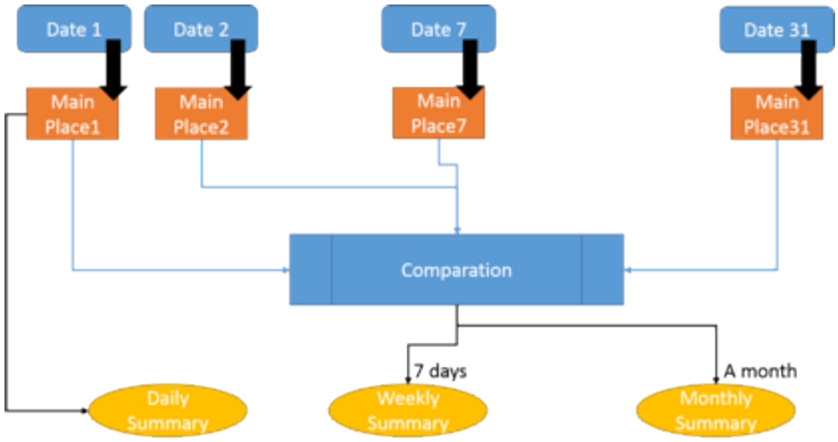

Figure 9 Summary generation flowchart

1. Generate the date ranges from the current date. For example, if it is a weekly summary, the app then generate the 7 days' date and save in a date list.
2. Find the main location from each date in the date list according to the time that user has spent in the places.
3. Find the weekly main location by comparing with other 6 days main location. As same as weekly summary, the monthly main location by comparing with other days in a month period.
4. List the main location in summary. if there are more locations which have the same time length, The main location will display more than one result in the summary dialog.

Weekly Activity Summary: This function aims to remind the user what has happened in the week before. The detail of the summary includes the special place, walking time period, running time period, cycling time period and transport time period as shown in Figure 10.

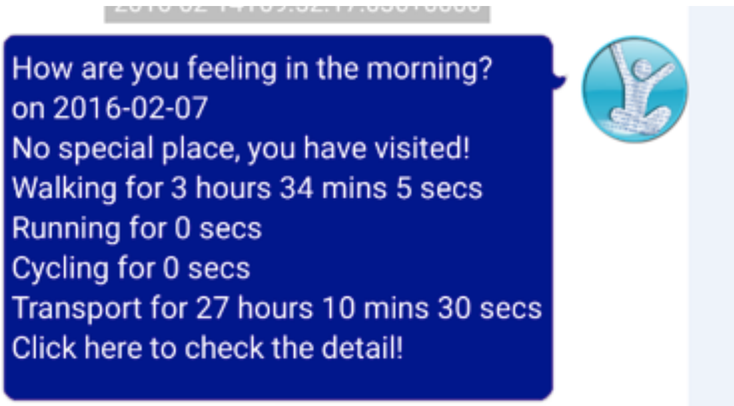

Figure 10 Weekly activity summary message

The special place means where a user has spent the most time in the week before. The walking, running, cycling and transport time period means the time the user has spent on the specific activity.

There are two steps to generate this summary:

1. Obtaining the date in the week before depends on the current date.

2. By using the given date, the app will search the data in the local database.

3. By using the Async-task, the app will download the data from the API service if no records exist in the database.

4. Find the special place, and the detail of the activities.

Moreover, user also can click the report message in the Journal, the app will display the day view of the date in week before.

12 week weight reducing program Summary: This program is used for overweight people to achieve the reduce weight purpose by using the app. Like the daily summary, the implementation time is set to 9:30 pm by default, and the user can set a custom time to trigger this information summary report.

Since the lifetime of this program will last for quite a long time, 12 weeks ideally, the summary report will be implemented every Sunday. The detail of this program summary items includes weekly calories-in, weight and activity per mins as shown in Figure 11.

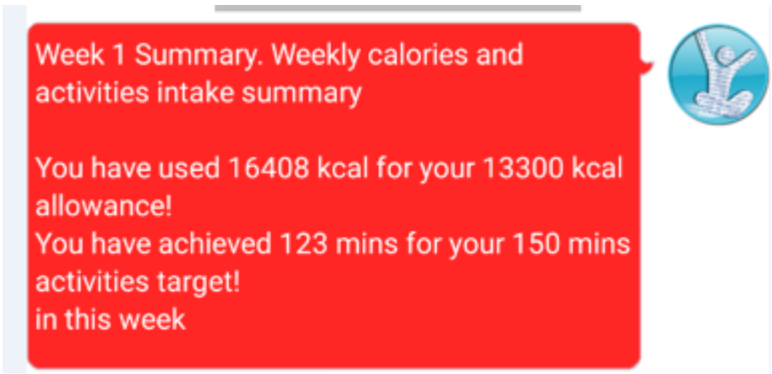

Figure 11 12-week weight reducing program summary

The program has default goal values for those items and all the 12 weeks data is store in the local database, therefore it is very easy to implement the summary report.

5.2 Recommendation

The recommendation aims to give a proper recommendation message to user according to his settings, profile, life style etc. Currently, the recommendation function has involved in Journal Daily Goals, NHS news, Activity sensor, questionnaires, location detection, and 12 weeks weight reduce program.

Journal Daily Goals: This recommendation offers daily goals for the user, in order to remind the user to achieve the goals as shown in Figure 12.

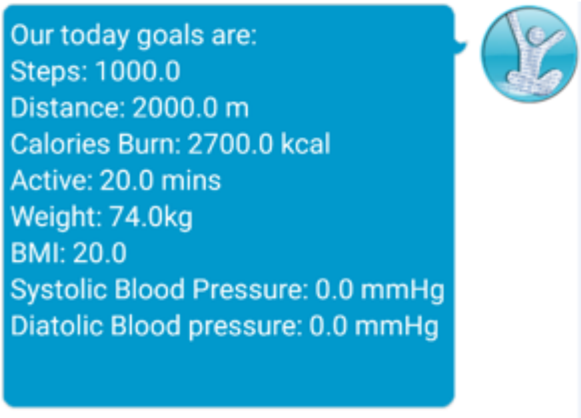

Figure 12 Daily goals message

There are two steps to generate this recommendation:

1. The goals record is retrieved from the API service, and system then updates the goals in the local database.
2. Read the goals from the local database and display them in the Journal page.

NHS News: This function offers news articles acquired from the NHS website to the user, generally according to a user's medical profile. Once a user has created a new medical record and saved it in the database, the app finds a proper NHS news topic to filter the news list. Therefore, a user could easily have the proper NHS news that is more related to his own body situation. For example, as shown in Figure 13, as user has already set their medical record to "Cancer", then the app recommends to the user to have the "Cancer" news.

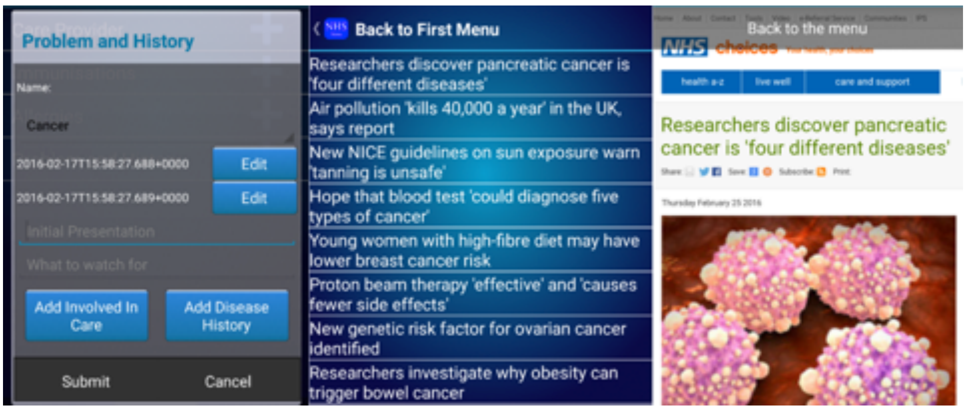

Figure 13 NHS News

The mapping between mobile app collection data and NHS news topics are show in Table 10.

Table 10 NHS news mapping

App Collection Data	NHS news topics
Gout	Diet and nutrition

Inguinal Hernia	
Congestive Heart Failure	
Coronary Artery Disease	
Glaucoma	Lifestyle and environment
Diabetes Mellitus, Type 2	Diabetes
Insomnia	Lifestyle and environment
Hypertension	Obesity and weight loss
Major Depressive Disorder	Lifestyle and environment
Neurology and dementia	
Cancer	
Mental Health	
12-week-reduce weight Program	Obesity and weight loss
Age (>65)	Older people and ageing
Gender (male)	Heart and lungs

The green colour items represent the NHS news topics, the black colour items represent the data collected from user "Medical Profile", the red colour items represent the data collected from user "General Profile" and the blue on represent the data collected from app program. As described before, it makes sense to map items from left to right side. Depending on the app's collection, the corresponding NHS topic will be highlight and set in NHS news function.

Web Data Scraping: The web data scraping aims to retrieve the health information from NHS office website, in order to offer user recommendation.

The web page has five top-level menus, and there are submenus under each the top menu. The news are classified depends on the different health issues. This method uses node.js to parse the web page and then retrieves the correspondent information.

Although the web page is marked as html data structure, it does not have strict tag syntax e.g. xml. Therefore, it is very difficult to parse the content in text format. The Cheerio module of node.js uses the V8 JavaScript engine parse the data in DOM, so that the data can be retrieved more accuracy from the website.

The process of retrieving the NHS data to JSON is shown in Figure 14.

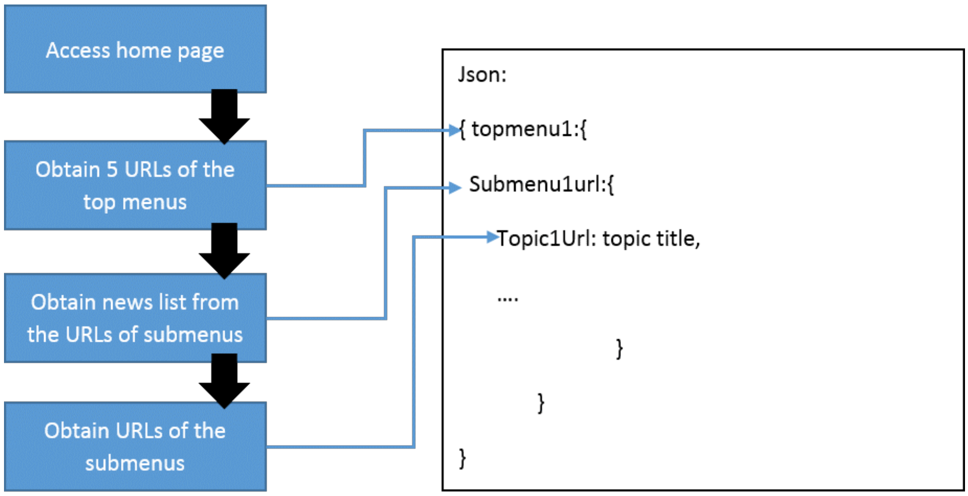

Figure 14 Web Data Scraping

The program first requests the home page to obtain the 5 top menus and their URLs. The URLs are used as the keys in the JSON object. The value of the keys are the URLs of submenus. The news bullets are embedded into the value of submenus objects in the JSON objects. Each item of the news is requested and the contents and the image URLs are stored into the JSON object.

Activity Sensor: This function offers a recommendation in the Journal page, if a user has been sitting for a quite a long time period. Actually, this function is recommended by researchers from the sports department in the University of Bedfordshire, due to walking being a great way to improve or maintain people's overall health [HealthWalking1, HealthWalking2]. Therefore. In the MyHealthAvatar app, the activity sensor provides the walk recommendation message according to user's steps and GPS location as shown in Figure 15.

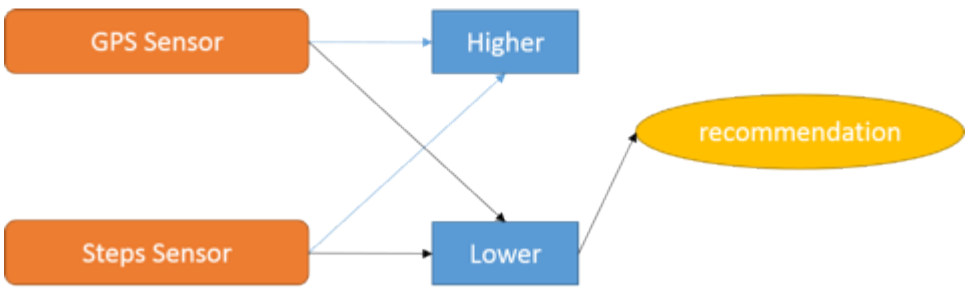

Figure 15 Activity Sensor

If a user's GPS location has been changed in a given time period, and steps changes are quite low, this means that user status is "travel" (e.g. car, train).

If a user's GPS location has been changed in a given time period, and steps changes are quite high, this means that the user is walking or running.

If a user's GPS location has not changed too much in a given time period, and steps changes are quite high, this means that user has been walking around in a small area.

If a user's GPS location has not changed too much in a given time period, and steps changes are quite low, this means that user has stayed in a place for quite a long time, and they need to take a walk. Then app will trigger the recommendation message to notify the user.

Questionnaires and Location Detection: Questionnaires are a nice tool to detect health problems. However, user may not have time, or are too bored to answer the questions due to the large amount questions in the Questionnaires. MyHealthAvatar app has thought about this issue and found out a better solution that depends on the current location to recommend that a user answer the question.

According to location detection function, app obtains the current location information and reasonably map it in a category as mentioned in Section 4.1. The correspondence between the activity categories and questionnaires is described in Table 11.

Table 11 correspondence between the activity categories and questionnaires

MyHealthAvatar activity Category		Questionnaires
My home	It is a very good time to ask some questions.	VF-14 QOL Questionnaire Life-Space Assessment
My work/study place	It is a good time to ask some questions when user arrived at the work or study place.	VF-14 QOL Questionnaire
Shop or shopping area	It is not a good time to ask some questions.	
Restaurant, pub, bar, or other eating place	It is not a good time to ask some questions due to user may not like to be disturbed.	
Transport	It is very good time to ask some questions.	Life-Space Assessment
Healthcare (GP surgery, dentist, hospital, etc.)	It is not a good time to ask some questions due to user may be sad and not like to answer question.	
Sports and Gyms	It is not a good time to ask some questions.	
Entertainment (cinema,	It is not a good time to ask	

theatre, nightclub, museum, etc.)	some questions due to user may not like to be disturbed.	
Hotel	It is very good time to ask some questions.	VF-14 QOL Questionnaire Life-Space Assessment
Governmental and public places (police, council, etc.)	It is not a good time to ask some questions due to user may not like to be disturbed.	
Friends' homes	It is not a good time to ask some questions due to user may not like to be disturbed.	
Children's school, college, etc	It is not a good time to ask some questions due to user may not like to be disturbed.	
Unknown		

Depends on those categories, the recommendation of questions at current location would be different, e.g. current location category is "My work/Study place", the recommendation questionnaire would be "VF-14 QOL Questionnaire" which are related to eyes, reading. Moreover, user could choose answer the questions any time through manually trigger the recommendation of question appear in Journal page.

12 Weeks Weight Reduce: As mentioned earlier, this program is used by overweight people to achieve the reduce weight purpose by following the recommendations provided by app. Once the user turns on the 12 weeks weight reduce program in Settings, the recommendation time will be set.

There are three times a day to trigger the recommendation tips, which includes cooking recipes, food, sports, and social interaction recommendation etc. Since the recommendation time can be customised by the user, the user can choose his dining time, sport time to trigger the tips.

Moreover, 12 weeks weight reduce program also supports the location detection and manually triggered. For example, when user goes to a Restaurant, pub, bar, or other eating place, although it is not a good time to ask questions as mentioned in Table 11, the app can give the reduce weight tips to the users.

Based on the previous description, there are three basic method to trigger the recommendation message to appear in the app as shown in Figure 16.

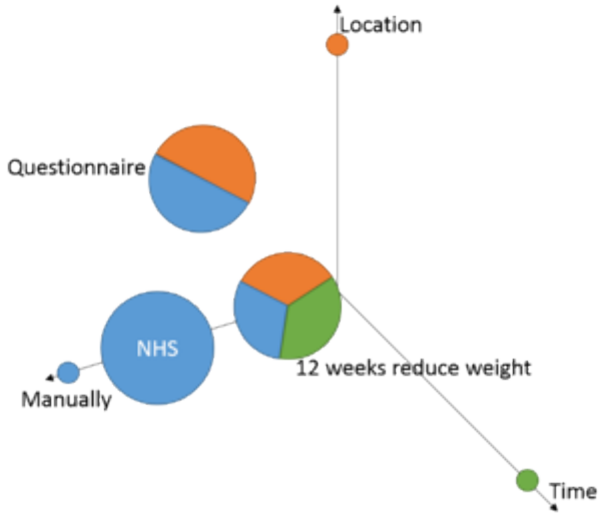

Figure 16 Three basic method to trigger the recommendation message

NHS is only based on manual user triggers. The questionnaire is based on two dimensions, manual and location. 12 weeks reduce weight is based on three dimensions, manual, location and time.

5.3 Performance Evaluation

The performance of summary and recommendation are list in Table 12.

Table 12 performance evaluation

Feature	Expected Functionality	Actual
Daily, Weekly and Monthly Summary	The summary completely and accurately displays in the summary page	Data are correctly display in the summary page ISSUE: If user set a higher goal value, the percentage of the result may become 0, due to the system only supports accuracy of 2 decimal points.
Main Location	The main location completely and accurately displays in the summary page	Data are correctly display in the summary page
Weekly Activity Summary:	The summary completely and accurately displays in the Journal page	Data are correctly display in the Journal page
12-week weight reducing program Summary	The summary completely and accurately displays in the Journal page	Data are correctly display in the Journal page
Journal Daily Goals	The information completely and accurately displays in the Journal page	Data are correctly display in the Journal page

		ISSUE: some users report that they have confused between weekly activity summary and Journal daily goals, due to they are quite similar and close to each other. SOLUTION: use different background colour for weekly activity summary and Journal daily goals.
NHS News	The specific Topic has completely and accurately displays in the NHS page depends on the user's setting.	NHS news function works fine.
Web Data Scraping	The function completely and accurately obtain the information from NHS news Website	Web Data Scraping function works fine.
Activity Sensor:	The information completely and accurately displays in the Journal page	Data are correctly display in the Journal page
Questionnaires and Location Detection:	The information completely and accurately displays in the Journal page	Data are correctly display in the Journal page ISSUE: Due to this function need mobile to check GPS position in a higher frequency, the user may need to face the power consumption issue.
12 Weeks Weight Reduce:	The information completely and accurately displays in the Journal page.	Data are correctly display in the Journal page

6 Knowledge extraction and reasoning

Based on previous results of activity summarisations, activity ranking and semi-automatic annotations, the knowledge extraction and reasoning process contain three components of RDF repository, semantic lifting engine and semantic reasoning engine as Figure 17 shows:

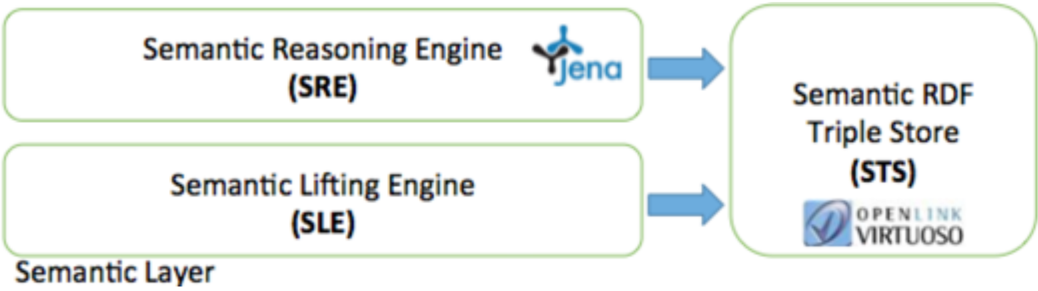

Figure 17 Knowledge extraction and reasoning process

The semantic lifting engine aims to compose all related mining data from lower level NoSQL repository together as input for mapping and lifting the data to the RDF triple storage (Virtuoso RDF repository) according to MyHealthAvatar semantic ontology. The Jena RDF reasoning framework is applied to perform the reasoning tasks. The In D4.3 and D4.4, the major PHR-based knowledge extraction process has been detailed illustrated and explained. In this report, we focus on the activity knowledge extraction and the reasoning. Importantly, the technologies used are compliant with D4.3 and D4.4 which allows the efficient integration with partners' work.

6.1 *MyHealthAvatar Ontology extension*

The semantic ontology used for the lifting and reasoning is shown in Figure 18.

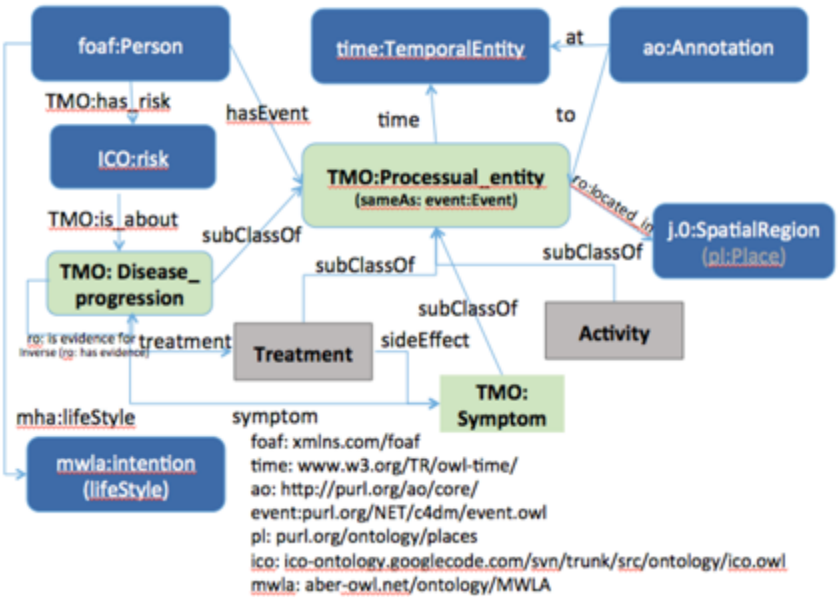

Figure 18 MyHealthAvatar semantic ontology

The core concepts of MyHealthAvatar ontology includes 10 major terms. The ontology majorly extends TMO terminologies with some existing semantic concepts from well-known domain ontologies and our defined personal activity and treatment terms.

1. **Event** defined as same as **TMO.Processual_entity** is a super concept to classify an interesting event that related to the health of an individual user. The event is the supper class of (discover a) **Symptom**, (taking a) **Treatment**, (diagnosed a) **TMO.Disease_progression** and (Having a) **significant**

activity. Each event associates to a particular time point on user's time.TemporalEntity. In addition, Event is the central point of the whole ontology, which can be detected from data mining layer.

2. **Person** is the concept to describe a MyHealthAvatar user using FOAF ontology. The FOAF ontology includes all possible aspects about a general profile of a person such as health history, gender and height. In our proposed framework, the semantic layer will nominalized user's name and address information that will be stored in lower level NoSQL database with more secured data management infrastructure and we will not discuss the security topic in this paper.
3. **Significant activity** is a subclass of **Event** concept to identify the activity that is more significant to the user rather than includes all daily activities. In general, all the significant event should be related to understand user's health situation or life style. The activity type can be grouped by the exercise type such as “Running”, “Driving”, and “Shopping” but also can be categorized by the places and social activity type. Each significant activity should also records the time duration, places and possibly with distances, calories consumptions and steps.
4. **Symptom** imported from TMO is a subclass of Event concept to present the unusual health related condition that are detected and concluded from the user's data. As same as all other events, the symptoms have to have a time stamp and places. Currently, the subclasses of Symptom includes low/high blood pressure, unusual heart rate, unwell sleeping and significant weight/fat changes. Other unsenserable symptoms can also be added but have to rely on user's manually inputs.
5. **Annotation** defined in AO (Annotation Ontology) is used as a semantic vocabulary link to an event. The annotations should use controlled vocabularies or semantic identifiers to define the meaning. The annotation can be automatically added through linked data annotation engine e.g. DBpedia spotlight or can be added by users via annotation tools from MyHealthAvatar platform.
6. **Treatment** is a subclass of Event concept for recording the treatments that have been taken by the user from medical health organization or user-self. The treatment refers to any medical actions that have been done to the user such as taking dugs, operation and physical and mind therapies. In addition, the treatment requires identifying the exact time point on user's timeline.
7. **Disease_progression** reused from TMO is a subclass of Event concept and presents the medical situations that were diagnosed in pass according the user's timeline or will be a potential risk for the user. The Health Condition, Treatment and Symptom concepts structure a triangle relation that could be a very valuable knowledge for individual user or a group of users.
8. **Risk defined by ICO** is used as a concept to evaluate the possibility or progress levels to a particular health condition.

However, we are only concentrate on the activity event and Symptom event (we also defined it as **Physiologic event**) knowledge extraction in this document. The ranking process of activity importance has been explained in section 4. At the same time, the general physiologic knowledge of the person can be easily detected through well-defined math formulars such as BMI changes, weights changes, high/low blood pressure and unwell sleeping and so on. which allows the semantic lift engine to composed the activity data with following information as two specific Json inputs: {activity type, ranking value, activity duration, location annotation, distance, steps} for activity event and {symptom type, value, time, location}.

6.2 Semantic lifting

The semantic mapping includes three steps of domain mapping, property mapping and range mapping.

- Domain mapping:

In the first step the domain matching algorithm is applied to identify the domain element that is the most simple algorithm in these three steps. According to our Json structure of composed summary data analysis, only activity type or symptom name elements are suitable candidate for the domain that can be lifted as subject elements of the instance RDF triples. If the element is under activities Json structure, then there will be URI generated and specified as an Activity class defined in the OWL ontology. For example, if the activity classified as running, then the URI will be

`myhealthavatar.org/onotology#running` and the definition triple will be

```
<http://myhealthavatar.org/semantic/0.1/activity/e40b922f-acbe-
476c-8c91-df880507e9b5>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://myhealthavatar.org/ontology#running>
```

The similar process will be generated for the symptom event. For example, the hypertension symptom event can be defined by the following triples:

```
<http://myhealthavatar.org/user/8d94-25ab-267a-4db5-94b5-2ac3-
653c-3112>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://xmlns.com/foaf/0.1/Person>

<http://myhealthavatar.org/ontology/symptom/e042b8b4-8a1e-4628-
ba11-161c9305098c>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type/>
<http://myhealthavatar.org/ontology/symptom#hypertension>

<http://myhealthavatar.org/ontology/symptom/e042b8b4-8a1e-4628-
ba11-161c9305098c>
<http://myhealthavatar.org/ontology/time/>
<http://myhealthavatar.org/ontology/timeinstance/2015_10_28_21:15>
```

```
<http://myhealthavatar.org/user/8d94-25ab-267a-4db5-94b5-2ac3-
653c-3112>

<http://myhealthavatar.org/ontology/symptom/>

<http://myhealthavatar.org/ontology/symptom/e042b8b4_8a1e_4628_ba1
1_161c9305098c>
```

- Property and range mapping:

According to the Json input structure, the property and range mapping are performed together based on the following pre-defined mapping Table 13 and Table 14:

Table 13 Property and range mapping 1

Json Structure Syntax	Mapped Property defined in the Ontology	Range value
ranking value	MyHealthAvatar: rank	(0,1]
duration,	MyHealthAvatar: time	Seconds
Destination	MyHealthAvatar: located in	Annotation text or unknown
Distance	MyHealthAvatar: distance	Meters
Step	MyHealthAvatar: step	Count
Activity_group	MyHealthAvatar: hasEvent	Activity type

Table 14 Property and range mapping 2

Json Structure Syntax	Mapped Property defined in the Ontology	Range value
Value	MyHealthAvatar: hasValue	Text value with unit
time	MyHealthAvatar: time	Time spot/date information
Location	MyHealthAvatar: located in	Annotation text or unknown

6.3 Semantic reasoning

The final goal to have the data lifted into the semantic repository is to enable digging the data further to discovery hidden knowledge of the user by getting the benefits from the smaller but more machine understandable data representations – RDF triples based on well-defined semantic ontology. The specific objectives of the our semantic reasoning process are:

- Understanding the user's lifestyles. 25 lifestyles have been proposed by Medical Web Lifestyle Aggregator (MWLA)[MedAggregator] ontology developed by other EC founded research project CARRE [CARRE]. Some interesting lifestyle should be able to be reasoned from our reasoning engine such as travel, sports, social (people and society), shopping, healthy, food_drink and career, although not all of them.
- The semantic relations between the user's activities and symptoms as well as the links between the life style and the health conditions.
- In the long-term, the user's life-long health situation can be concluded to help the disease prevention and prediction.

In order to achieve the reasoning objectives, we developed a semantic reasoning engine using Jena semantic framework that supports the SPARQL and SWRL (Semantic Web Rule Language) programming and can be integrated into Virtuoso RDF repository. One is based on SPARQL query that can define the reasoning formulas on ontology level (T-box). The other one is based on SWRL rule that cannot be specified on ontology level but instances level (A-box). For clearly explaining these two different reasoning processes. We represent two reasoning scenarios in the remaining section to illustrate how reasoning can be applied for inference of life style pattern and links to the certain health condition or symptoms.

Example 1: Travel/long commute lifestyle for the past month

Definition: Long transport (more than 2 hours) activity events have been lifted in to RDF at least four times in the last month. The reasoning process will be:

- Construct (SPARQL query) the last month activity event RDF memory-model based on the ontology retrieved from the triple storage.
- Specify the SWRL rule or SPARQL query for reasoning. In this example, the SPARQL query will be used as Script 1:

```
PREFIX MyHealthAvatar: <http://myhealthavatar.org/ontology/>
SELECT (COUNT(?numberOflongTravelling) AS ?howMany ?p)
WHERE  { ?p  MyHealthAvatar:hasEvent ?e .
         ?e  rdf:type MyHealthAvatar:transport .
         ?e  MyHealthAvatar:time ?d .
              FILTER (?d >= 7200)
}
HAVING ( ?howMany > 4 )
```

Script 1 Lifestyle query

Script 1 specifies a query that will be return transport times (COUNT ?howMany) a month if the transport activity events have been detected more than 4 times (use ASK query to filter ?howMany > 4) in a targeted month.

- If the query return a value, then it means the user satisfy the defined reasoning query, then we can construct the semantic links between the person to the Travel term defined in MWLA.

The steps of (2) and (3) can be integrated by combining Construct update query to the model as one step (Script 2).

```
PREFIX MyHealthAvatar: <http://myhealthavatar.org/ontology/>
PREFIX MyHealthAvatar: < http://purl.bioontology.org/ontology/MWLA>

CONSTRUCT (?p MyHealthAvatar:lifestyle mwla:Travel)
SELECT (COUNT(?numberOflongTravelling) AS ?howMany ?p)
WHERE  { ?p  MyHealthAvatar:hasEvent ?e .
         ?e  rdf:type MyHealthAvatar:transport .
         ?e  MyHealthAvatar:time ?d .
           FILTER (?d >= 7200)
}
 HAVING ( ?howMany > 4 )
```

Script 2 Combine Construct update query to the model

The similar processes can be applied to detect other lifestyle such as lack of activity or lack of exercise and so on.

Example 2: SWRL based health condition risk alarm reasoning.

Definition: If a person lacks of activity and age > 60, then it will be a risk of high blood pressure

The rule can be defined as Script 3 in Jena rule engine:

```
[rule: (?p MyHealthAvatar:lifestyle mwla:noActivity), (?p foaf:age ?i),
greaterThan(?i, 60) -> (?p tom:has_risk ?x), (?x tom:is_about ?d), (?d
rdf:type tom:High_Blood_Pressure)]
```

Script 3 SWRL based health condition risk alarm reasoning

7 Future Work

7.1 *Activity Summarisation using activity loops*

7.1.1 Introduction

Life pattern mining are especially useful for lifestyle detection, health monitoring, crime analysis, traffic management, urban planning, etc. A daily storyline is the full sequence of places and activities taken by a person in his one day life. A loop is an important life pattern in daily activity storylines. To recognize the daily activity patterns, in the first stage loops in the daily activity storylines may be recognized and analysed. A lifeloop is one of the special activity patterns which starts and ends at the same place. The definition of a loop type can be specified by the user. Similarity definition is the key for loop definition and recognition. There may be other types of activity patterns which can be specified by the user for the future work.

The objectives of Lifeloop include:

- Detect all the loops from the movements data based on user specification.
- Define an event as a loop by detecting the "semantics" of the loop, together with its granularity (the granularity is used to define the detail of the events)
- Allow queries for the events from temporal and spatial dimensions
- Define similarity between different events (namely the loops)
- Clustering of the events based on the similarity
- Outlier detection of the events
- Statistics of the events (Similarity-based)
- Importance ranking of the events

7.1.2 Lifeloop Generation

7.1.2.1 Data Source

In MyHealthAvatar the current loop generation is based on annotated Moves data, but in the future it can also be extended to other data format, such as Google timeline data [GoogleTimeline] which is similar to Moves data.

7.1.2.2 Event Definition

Event definition is the key in Lifeloop visual analysis. In Lifeloop the event will follow the rule of five Ws [Zhang2013] and will be highly user customizable:

Who: whose's home, who did you meet, group analysis, etc

When : e.g. time range, frequency, duration, etc

Where : e.g. place name, from – to, distance, etc

What : event category, etc

Why : purpose, etc

7.1.2.3 Place Definition

Though Moves provides the suggested places, it is not always accurate and sufficient for customizable hierarchical Lifeloop analysis. Customisable place definition with the following functions is desired:

- Choose only the desired types of places as the Start/End point of the Loop
- Manually define place category hierarchies
- Same place Definition

7.1.2.4 Loop Generation Algorithm

The current loop generation is based on the storyline with place information:

1. Extract all the place information from moves storyline to generate a placeline

2. In the placeline, for each place, backtrace all the places to the beginning in the array, if a first same place is found, a loop is generated:

3. Slice all place nodes between the two place nodes

4. Generate a loop node to replace all the removed nodes

5. Repeat step 1-4 until the end of the placeline array is reached.

The loops generated based the above algorithm may be nested.

The ongoing work of Lifeloop visualization is introduced in section 11 of D8.2.

7.2 Semi-supervised Place and Activity Annotation

7.2.1 An Automatic Place Annotation method using Bayesian framework

7.2.1.1 Notations:

Pl: name of the places annotated

Co: coordinate of a place

Day: Day (From Monday to Sunday)

P(a): The probability of a

P(a|b): the probability of *a* conditioned on *b*

7.2.1.2 Problem

Given a set of {Co1, Co2 ….), each of which has a list of candidate place names from Four-Square, as well as a set of manually annotated place create the annotation {Pl1, Pl2….) for these coordinates.

7.2.1.3 Algorithm

using EM method to compute P(Pl | Co, Day) for each possible Pl and choose the one with the maximum probability for the annotation of Co.

Initialization:

P(Pl) = a flat rate

P(Co | Pl) = inversely proportional to the distance value from the Four-Square annotation, and to the distance to the closest place that was manually annotated as Pl

P(Day | Pl) = a random value

Then repeat between the following E and M steps:

E-Step:

Compute:

$$P(Pl\,|Co, Day) = \frac{P(Co, Day|Pl)\times P(Pl)}{\sum_{Pl'} P(Co, Day|Pl')\times P(Pl')}$$

M-Step:

Compute:

$$P(Co, Day|Pl) \sim n(Co, Day)\times P(Pl\,|Co, Day)$$

$$P(Pl) \sim \sum_{Co\ Day} n(Co, Day)\times P(Pl\,|Co, Day)$$

n(Co, Day) is the number of the occasions that Co and Day co exist

7.2.2 GPS location accuracy analysis for place annotation

GPS hardware give 10m horizontal accuracy since SA (selective availability) is turned of according to (Bill Clinton 1996) policy directive in May 2000 (it was 100m when SA was on). Vertical accuracy is usually about 3 times worse than horizontal. In 2006 low-cost unit commonly include Wide Area Augmentation System (WAAS) receivers whose accuracy is 7.6m worst case, 1-2 meter in best case. As of early 2015, high-quality FAA grade Standard Position Service (SPS) GPS receivers provide horizontal accuracy of better than 3.5 meters. [GPSSelAvail]

According to [GPSAccuracy], people has recorded a routinely 2 meters accuracy with certain Android phones like Motorola Droid.

Take into consideration: weather, surrounding high buildings and device on the move, it make sense to assume the Moves place location has an accuracy of 10 meters on its supported phones.

According to [WGS, AndroidLocation] World Geodetic System 1984 (WGS84) ellipsoid is used by GPS system and Android system. The WGS84 datum surface is an oblate spheroid with major radius a = 6 378 137m at the equator and flattening f = 1/298.257 223 563. The polar semi-minor axis b then equals a times (1-f), or 6 356 752.3142m.

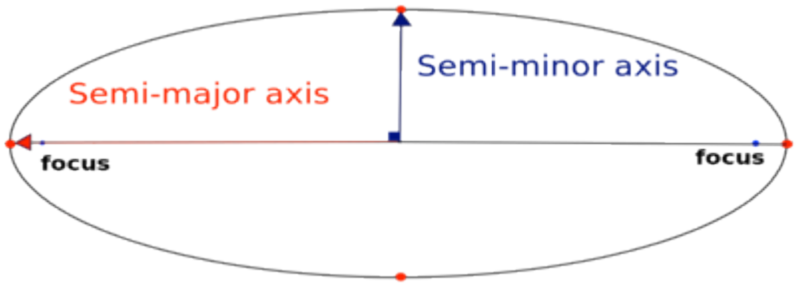

Figure 19 WGS84 datum surface

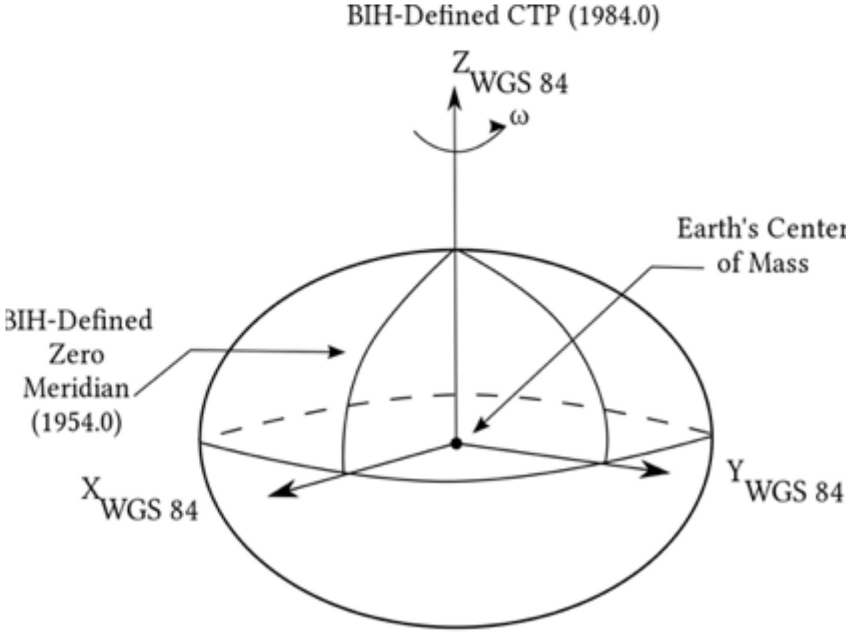

Figure 20 WGS 84 reference frame

According to [GeoCoordSys], at Greenwich longitudinal length equivalents of 0.0001 degree is about 6.95 meter, which is about the range of the accuracy 10 meters of GPS. A place of small shop, friends home, etc normally have a box size of 10m * 10m, however university, train stations can have a box size larger than 30m * 30m. Take into consideration the Moves data return a location data like: {lat: 37.**9743**830596, lon: 23.**7310**623093}. It make sense to consider a box of place within **±0.0002 ~ ±0.0007** (approximately range of 30 meter in longitude to 100m in longitude).

Based on the GPS accuracy analysis, we assume that within 30m * 30m will annotate use most visited place, out of 100m * 100m is considered not related to the given place.

Other considerations include :

User annotation consideration: public annotation, private annotation

Auto annotation should have a flag to note it is annotated automatically for user.

8 Conclusion

This deliverable reports work of data reasoning utilities for decision support with evaluation reports. The ultimate goal of data collection is to make use of data. With the data reasoning tools, MyHealthAvatar is not only a health data collecting tool and repository but also a powerful data analysis platform.

In this report the following work from both MyHealthAvatar web and mobile platforms is reported:

- Sensor data validation based on MyHealthAvatar PAV model
- Event and activity extraction and ranking based on MyHealthAvatar tf-idf model
- Information summary and recommendation model in the MyHealthAvatar mobile app
- Knowledge extraction and reasoning based on MyHealthAvatar ontology extension
- Ongoing work on LifeLoop recognition and analysis
- Ongoing work on Semi-supervised place/activity recognition and annotation

The task deals with data reasoning based on the linked data in the RDF data repository which can help to discover inexplicit relationships between data. The data reasoning will provide necessary supporting data for clinicians in diagnosis process and citizens in decision making for health related issues. Together with semantic reasoning, the work can be linked visual data analytics in WP8 to show the usage of the data in decision supporting process.

The report shows that MyHealthAvatar is not only a powerful health data repository but also a versatile data analysis platform which provides data collection tools as well as data analysis tools for both the citizens and clinicians in knowledge acquisition and decision making on health related issues. The data analysis tools introduced in this deliverable report are integrated with the visual analytics tools introduced in D8.2, which leads to more efficient and capable insight gaining.

9 References

[AndroidLocation] http://developer.android.com/reference/android/location/Location.html

[Atallah2011] L. Atallah, B. Lo, R. King and G. Z. Yang, "Sensor positioning for activity recognition using wearable accelerometers," IEEE Trans. Bio. Cir & Sys, vol. 5, no. 4, pp. 320-329, 2011.

[CARRE] CARRE Project, http://www.carre-project.eu

[COSMEDK462] "COSMED K462," COSMED, 2015. http://www.cosmed.it/en/products/cardio-pulmonary-exercise-testing/k4-b2-mobile-cpet.

[Curone2010] D. Curone, A. Tognetti, E. L. Secco, G. Anania, N. Carbonaro, D. DeRossi and G. Magenes, "Heart rate and accelerometer data fusion for activity assessment of rescuers during emergency interventions," IEEE Trans. Info. Tech. Bio, vol. 14, no. 3, pp. 702-710, 2010.

[Deng2015] Zhikun Deng, Po Yang, Youbing Zhao, Xia Zhao, Feng Dong: Life-Logging Data Aggregation Solution for Interdisciplinary Healthcare Research and Collaboration. CIT/IUCC/DASC/PICom 2015: 2315-2320

[Endomondo] "Endomondo," Endomondo Ltd, 2015. [Online]. Available: https://www.endomondo.com/.

[Fitbit] Fitbit: http://www.fitbit.com/

[FitbitFlex] "Fitbit Flex," Fitbit Ltd, http://www.fitbit.com/flex

[FitbitOne] "Fitbit One," Fitbit Ltd, Available: http://www.fitbit.com/one

[Fluxstream] Fluxstream: https://fluxtream.org/

[Gargasas2012] L. Gargasas, V. Jurkonis, L. Bikuliene, A. Zvironiene and S. Daukantas, "Functional state evaluation system with distributed intellect for elderly and disabled persons," Technologies for Computer Control, vol. 13, pp. 57-52, 2012.

[GeoCoordSys] https://en.wikipedia.org/wiki/Geographic_coordinate_system

[Ghasemzadeh2011] H. Ghasemzadeh and R. Jafari, "Physical movement monitoring using body sensor networks: a phonological approach to construct spatial decision trees," IEEE Trans. Ind. Informat, vol. 7, no. 1, pp. 66-77, Feb, 2011.

[GoogleMyTracks] Google MyTracks, (no longer available after April 30, 2016)

[GoogleTimeline] Google Maps Timeline, https://www.google.com/maps/timeline.

[GPSAccuracy] http://stackoverflow.com/questions/1567443/how-accurate-is-android-gps

[GPSSelAvail] https://en.wikipedia.org/wiki/Global_Positioning_System#Selective%5Favailability

[iHealth] iHealth: http://www.ihealthlabs.com/

[HealthVault] Microsoft HealthVault: https://www.healthvault.com/

[HealthWalking1] https://www.betterhealth.vic.gov.au/health/healthyliving/walking-for-good-health

[HealthWalking2] http://www.healthyalberta.com/656.htm

[Herrmann2011] S. D. Herrmann, E. M. Snook, M. Kang, C. B. Scott, M. G. Mack, T. P. Dompier and B. G. Ragan, "Development and validation of a movement activity in physical space score as a functional outcome measure," Arch. Phys. Med. Rehabil, vol. 18, no. 5, pp. 1636-1646, 2011.

[JawboneUp] "JawboneUp," Jawbone. Available: https://jawbone.com/up.

[MedAggregator] Medical Web Lifestyle Aggregator, http://aber-owl.net/ontology/MWLA

[MHA] MyHealthAvatar Project, http://www.myhealthavatar.eu/

[MHAWeb] MyHealthAvatar Platform, http://myhealthavatar.org

[Moves] Moves: https://www.moves-app.com/

[Nike+] "Nike+ Fuelband," Nike. Available: http://www.nike.com/gb/en_gb/c/nikeplus-fuelband.

[Paradiso2005] R. Paradiso, G. Loriga and N. Taccini, "A wearable healthcare system based on knitted integrated sensors," IEEE Trans. Info. Tech. Bio, vol. 9, no. 3, pp. 337-344, 2005.

[Parkka2006] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola and I. Korhonen, "Activity classification using realistic data from wearable sensors," IEEE Trans. Info. Tech. Bio, vol. 10, no. 1, pp. 119-128, 2006.

[Sebestyen 2013] G. Sebestyen, A. Hangan, S. Oniga and Z. Gal, "An evaluation of wearable activity monitoring devices," in The 1st ACM international workshop on personal data meets distributed multimedia, Barcelona, Spain, 2013.

[Spanakis2014] E. G. Spanakis, D. Kafetzopoulos, P. Yang, K. Marias, Z. K. Deng, M. Tsikanakis, F. Dong and V. Sakkalis, "MyHealthAvatar: Personalized and empowerment health services through internet of things technologies," in EAI 4th International Conference on Wireless Mobile Communication and Healthcare, Athens, 2014.

[Vift] "Vift," Medisana, 2015. [Online]. Available: http://www.medisana.com/en/Sport/Activity-Tracker/ViFit-Activity-Tracker.html.

[WGS] https://en.wikipedia.org/wiki/World_Geodetic_System

[Withings] Withings: http://www.withings.com

[Zhang2013] Zhang Z, Wang B, Ahmed F, Ramakrishnan I, Zhao R, Viccellio A, Mueller K., The Five W's for Information Visualization with Application to Healthcare Informatics. TVCG. 19(11), 2013

[ZhaoY2015] Y Zhao, F Parvinzamir, X Zhao, Z Deng, F Dong, N Ersotelos, G Clapworthy, Web-based Visual Analytics of Lifestyle Data in MyHealthAvatar, MOBIHEALTH 2015, DOI: 10.4108/eai.14-10-2015.2261796