NESTORE

D4.3.2 – Plans and recommendations based on profiles
2019.08.31

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

www.nestore-coach.eu

DELIVERABLE ID:	**WP4/D4.3.2/TASK NUMBER 4.3**
DELIVERABLE TITLE:	Plans and recommendations based on profiles
RESPONSIBLE PARTNER:	Eurecat
AUTHOR (S):	Silvia Orte, Paula Subías (Eurecat), Laura Fernández (FSiE), Filippo Palumbo, Michele Girolami (CNR-ISTI), Martin Sykora (LU-CIM)
CONTRIBUTORS:	Alfonso Mastropietro, Simone Porcelli, Giovanna Rizzo (CNR), Noemí Boqué, Lucia Tarro (Eurecat), Sabrina Guye, Christina Röcke (UZH), Petia Radeva (UB)
NATURE	Report
DISSEMINATION LEVEL:	PU
FILE:	DEL4.3_v3.5_PUBLIC
REVISION:	V3.5
DUE DATE OF DELIVERABLE:	2019.08.31
ACTUAL SUBMISSION DATE:	
CALL	European Union's Horizon 2020 Grant agreement: No 769643
TOPIC	SC1-PM-15-2017 Personalized coaching for well-being and care of people as they age

Document History

Revision	Date	modification	Author
0.0	2019.05.29	Creation of TOC for D4.3.2	EURECAT
1.0	2019.06.26	First draft of D4.3.2 taking the content from D4.3.1	EURECAT
1.1	2019.08.08	Content added	EURECAT
2.0	2019.08.19	Sections about indicators added	CNR-ISTI, LU-CIM
2.1	2019.08.20	Content about coaching events added	EURECAT
2.2-2.6	2019.08.23	Eurecat internal revision	EURECAT
2.7	2019.08.27	Section 6.3 and annexes updated	CNR-ISTI
3.0	2019.08.28	Final version after internal review	EURECAT
3.1	2019.08.29	Comments from CNR-ITB integrated	EURECAT
3.2	2019.09.03	Comments from CNR-ITB integrated	FSiE
3.3	2019.09.03	Overall review and edition	UB
3.4	2019.09.12	Public and private versions created	EURECAT
3.5	2019.09.17	Dates updated	EURECAT

Approvals

Date	Name	Organization	Role
2019.09.03	Giovanna Rizzo	CNR-ITB	Reviewer
2019.09.03	Alfonso Mastropietro	CNR-ITB	Reviewer
2019.09.03	Petia Radeva	UB	Reviewer
2019.09.04	Silvia Orte	EURECAT	WP Leader
2019.09.13	Giuseppe Andreoni	POLIMI	Scientific Coordinator

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

Short Abstract

This document represents the deliverable D4.3 (Plans and recommendations based on profiles) of WP4 (NESTORE Decision Support System) and contains a description of the research carried out in the context of task 4.3 "Algorithms for modelling and profiling individuals". In order to preserve the Intellectual Property generated in this task, we have split the D4.3. in two parts: a public and a private document. We encourage the reader to ask for the private version of D4.3. to get the complete understanding of the development carried out in D4.3.; this public version only outlines what has been done. The purpose of this report is to:

- characterize NESTORE users through Personas to create end-users' models;
- explain the scope of user profiling implementation in the DSS;
- demonstrate its functionality with examples;
- present the methodology followed;
- describe the personalization through recommendations;
- explain users' models;
- present the objectives and implementation details of a profiler simulator.

In chapter 1, we introduce the Decision Support System and its main components with the aim of presenting the user profiling as the main element of personalization. In chapter 2, the user profiling process that has been carried out is listed. In chapter 3, we describe why and how Personas are designed going through all domains of NESTORE interest, and we present the approach we followed to integrate this concept in the profiling system. Chapter 4 introduces and explains the tagging system, which is the core of the recommendation system. In chapter 5, we introduce the recommendations proposed by domain experts and depict their integration into the system. Finally, chapter 6 describes how users are modelled and simulated.

Key Words

User profiling, Decision Support System, Older Adults, Personalization, Attributes, Personas, Recommendations.

Table of Content

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

Table of Figures

List of Tables

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

1. Introduction

A *decision support system* (DSS) can be defined as a computerized information system used to support decision-making in which the characteristics of an individual are matched to a computerized knowledge base (Holsapple, Whinston, Benamati, & Kearns, 1996). DSS lets users sift through and analyse massive reams of data and compile information that can be used to solve problems and make better decisions. In NESTORE, the DSS is intended to help older people to compile useful information about their lifestyle in order to identify proper actions and make decisions to improve or maintain a healthy life.

One of the primary objectives in NESTORE project is to develop a DSS so that the users can obtain fast, reliable, personalised, and directly applicable advice. Suggestions are delivered in form of coaching plans, which are divided into pathways composed of different coaching events (refer to D5.1, section 5.2.4). The DSS and, concretely, a user profiling module is in charge of proposing the coaching plans and recommendations that better fit each user based on extracted attributes.
The information that is used in the personalization process comes from:

- models described in D2.1;
- recommendations and guidelines defined in D2.2;
- behavioural models and intervention techniques reported in D5.1;
- existing knowledge from domain experts and other evidence-based sources.

User profiling is one of the key steps in recommendation processes since it is essential for extracting user characteristics and predicting how much a user will like an item.
As depicted in Figure 1, the user profile and user preferences feed the DSS engine with the necessary inputs to select the most convenient coach plan for each user throughout all the weeks of the intervention phase.

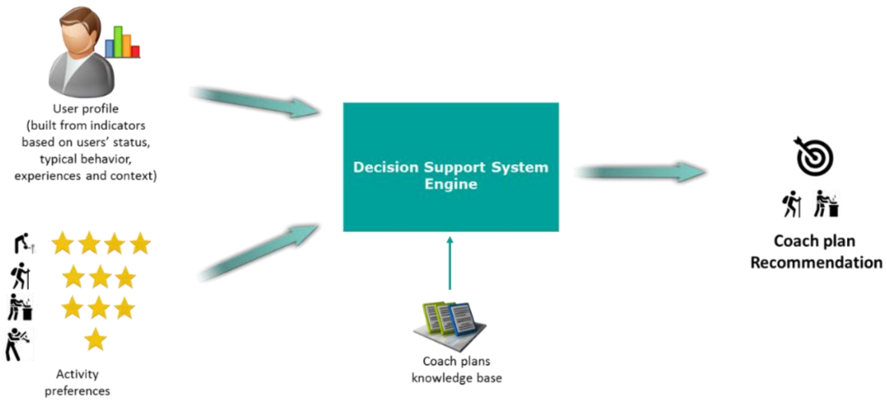

Figure 1. Conceptual view of DSS engine

In this deliverable, we focus on the personalization side of the DSS, mainly embodied in the user profiling component. It describes the way we profile the users with the final aim of selecting the recommendations and coaching plans that better fit the user.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

1.1. Relation with other work packages

The procedure and algorithms explained in this document belong to the work done under the tasks T4.2 ("Recognition of trends and user habits") and T4.3 ("Modelling and profiling individuals"). As depicted in Figure 2, the results of these two tasks conform the modelling of the user profile (both in short-term and long-term) and they are used by means of the designed and developed tagging system to adapt the recommendations given to the user in form of Coaching Events.

On one hand, user profiling mechanisms are fed by both the indicators explained in D4.1 and developed under T4.1, and the results of discovering routines and habits developed under T4.2 On the other hand, domain experts of NESTORE, explained in detail under WP2 deliverables the characteristics that the reasoning system and the user profiling should follow in order to fit domain's requirements.

WP6 developments are also used in these tasks, given that general user profile attributes and preferences, filled in by the user during the sign-in process, are stored in the cloud and accessed through the communication APIs developed under WP6.

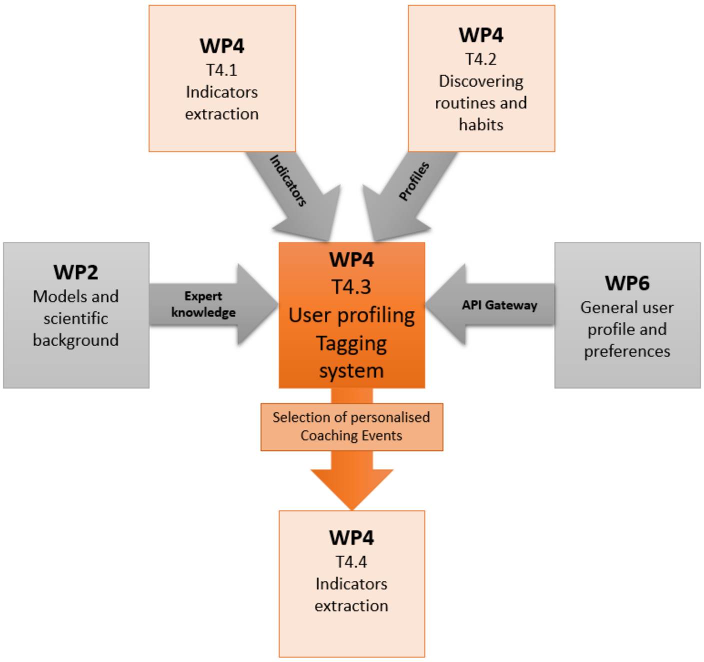

Figure 2. Graphical representation of the relationships among WP4 – T4.2 & T4.3, the activities of other NESTORE work packages, and the activities of other WP4 tasks

2. Methodology

User profiling can be defined as the process of identifying the data about a user interest domain. This information can be leveraged by the DSS to better understand the user needs and, thereby, provide personalized recommendations.
Various sensors and applications deployed in NESTORE platform generate the input data of the DSS. These data shapes the user profile, which is the element that leads the personalization process.
The process to build the user profiling is as follows:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

- *Step 1.* Personas are designed to analyse the different types of information that we need to personalize NESTORE recommendations. In section 3 of this document, we explain the process of developing NESTORE Personas.
- *Step 2.* The final set of Personas is analysed and a list of attributes is extracted from it.
- *Step 3.* The list of attributes is complemented with other items that domain experts believe that are important for the personalization procedure. In section 6.1.1 of this document, we list the static profile attributes extracted from the research carried out in steps 1, 2, and 3 of this process and we explain the need of developing a simulator of static profiles.
- *Step 4.* The data flow for recommending coaching plans is designed and different use cases where user profiling will be used are envisaged.
- *Step 5.* Different user profiling methods are analysed.
- *Step 6.* The decision of implementing a tagging system is taken after doing some experiments with data coming from a simulator.
- *Step 6.* User profiling module is implemented and integrated in the NESTORE system.

3. Characterization of the user through the design of Personas

The Inmates Are Running the Asylum (Cooper, The inmates are running the asylum. Indianapolis, IA: SAMS, 1999), introduced the use of personas as a practical interaction design tool. Personas are hypothetical archetypes of end users. Although they are imaginary, they are defined with significant rigour and precision, and they help to base the potential users' descriptions in real cases to achieve more realism. The main aims of the Persona methodology are to:

- define simple and real personas' profiles in an effective way;
- create end users' models for representing their life, needs and preferences;
- build a new understanding about who is the end user to help team members feel connected to them, raise empathy and work with the same personas' cases.
- work in levels of complexity in function of the depth of definition of each model, for example from expert users to novice and advance their needs and requirements if it is possible;
- have a model to facilitate discussions in cognitive walkthroughs, storyboarding, role-playing, and other usability activities;
- create a collection of archetypes to help new team members learn about the characteristics of users' profile.

In this stage of the project and in this concrete task, the design of Personas helps us in:

- creating use cases to permit researchers to better analyse the problem;
- generating a structure of attributes with their ranges and domains;
- developing a complete and realistic profile simulator;
- approaching the *cold start* problem in machine learning methods.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

3.1. How were Personas designed?

The process of creating Personas was based not only on previous research projects prepared for the development of user profiles (Wöckl, et al., Basic senior personas: a representative design tool covering the spectrum of European older adults, 2012) (Wöckl, Yildizoglu, Buber-Ennser, Aparicio Diaz, & Tscheligi, Elderly Personas: A Design Tool for AAL Projects focusing on Gender, Age and Regional Differences, 2013) but also on an iterative process to facilitate the transversal cooperation between the different NESTORE partners and key agents. The entire process was based on the importance of reflecting the idiosyncrasies and realities to develop useful profiles for the implementation of the system.

The research was developed consulting the main European demographic public resources (United Nations) (European Comission) (European Parliamentary Research Service, 2014) (European Comission, 2019) to detect the core characteristics of elderly population, but also to be aware of the possible heterogeneity in this target group (see D2.1). Personas were designed by the co-design experts and piloting countries to include their privileged view of the real user's needs and preferences. There were also taken into account the co-design experts considerations to include their privileged view of the real user's needs and preferences. Domain experts' considerations were also taken into account in order to introduce valuable information to enrich the global understanding of the potential NESTORE users. Another valuable feedback was obtained from the Forum Advisory Stakeholders (FAS). Suggestions and questions pointed out by FAS members were reflected in a new version of users' profile and personas document. This fruitful cooperation had, as a result, a large list of profiles (n=24). This contribution aimed to reflect the heterogeneity from the European contest.

Three tools were created to help in the process of refining profiles.

Firstly, **a checklist with key questions to be asked to co-design experts and pilot teams** was created (see annex 1 – section 0). The main purpose was to select the final personas systematically and guide experts in the evaluation of each profile to detect those who have more capacity to be more informative or descriptive for technical researchers and developers. Secondly, **a document based on a table with two tabs, one for comparing and grouping the different profiles and a second tab for merging and defining 8 contexts** was produced. This tool helped to refine the status, preferences, and attributes. Finally, the third tool was **a diagram that presents three important aspects (personal and environmental characteristics and possible pathways)**. This schema was crucial to highlight the needs and preferences of personas' profiles, which will determine the possible elections of pathways of real users.

Finally, it was proposed to create a card template to reflect the main characteristics of each profile. This task helps to be systematic and gain the consistency to build profiles.
It was necessary to develop a refinement process to adjunt the profiles:

The use of the diagram tool (Figure 3) helped to define two main aspects to be included in the refinement of Personas, in accordance with previous projects and the Cooper definition (Cooper, Reimann, & Cronin, About face 3: the essentials of interaction design, 2007):

- **End goals:** motivational goals but based on their live preferences. These goals could be very effective to determine in some way the final acceptance or user perception of the usefulness of a product or service when it is achieved a convergence between real users' needs and product or service features to answer these needs. When these goals are reflected in profiles, it could help to understand the cognitive walkthroughs, personal contexts or "a day-in-the-life of" scenarios. In the NESTORE case, these goals were defined based on the project domains (physiological, nutrition, cognitive and mental, social interaction).

- **Life goals:** defined as the Persona's long-term desires, motivations, self-image attributes and personal aspirations. This description could help to explain why the user is trying to accomplish goals. The previous work developed in NESTORE co-design phase helps to build a better understanding of real-life facts of the elderly population and to add in the descriptions of each profile.

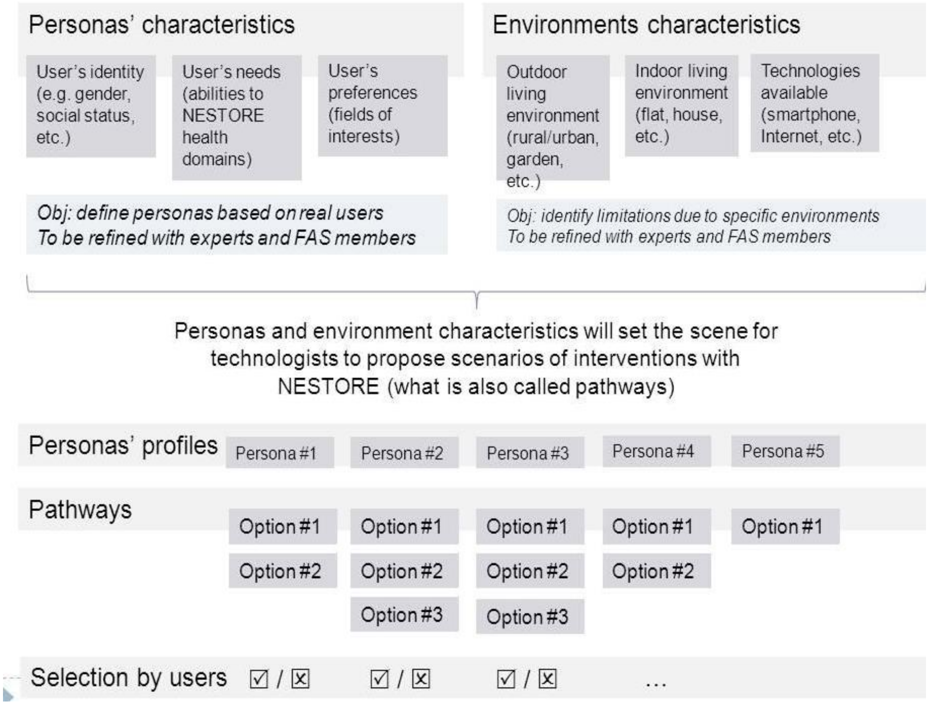

Figure 3. Diagram of Persona's profile template

In NESTORE's profiles, the end goals and life goals were indirectly suggested by means of the description of personas' daily activities and main interests; for example, in some profiles spending time with family, to be involved in cultural or voluntary movements, etc.

In Figure 4 and Figure 5, two examples are provided to illustrate the building process to define each profile.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

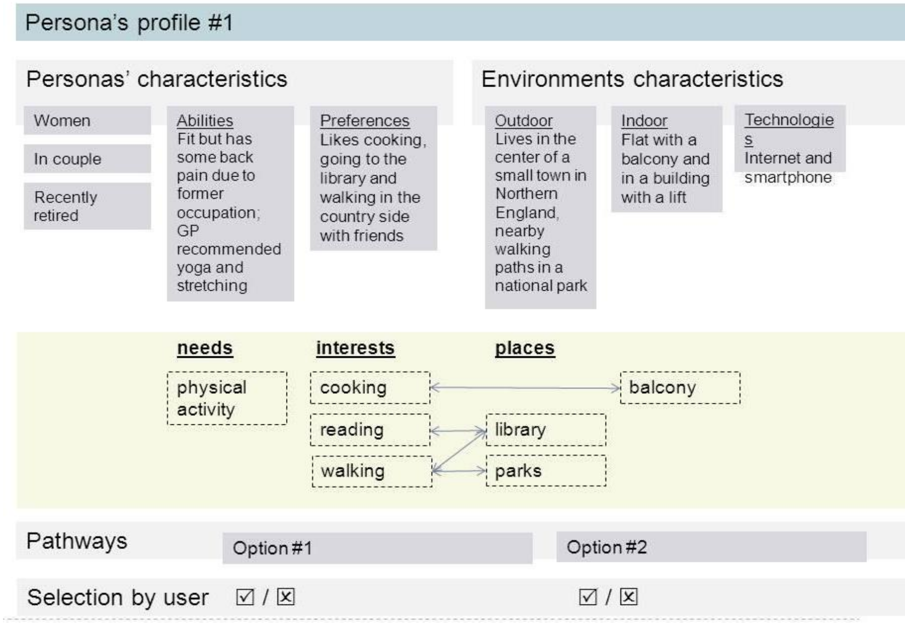

Figure 4. Diagram of Persona example 1.

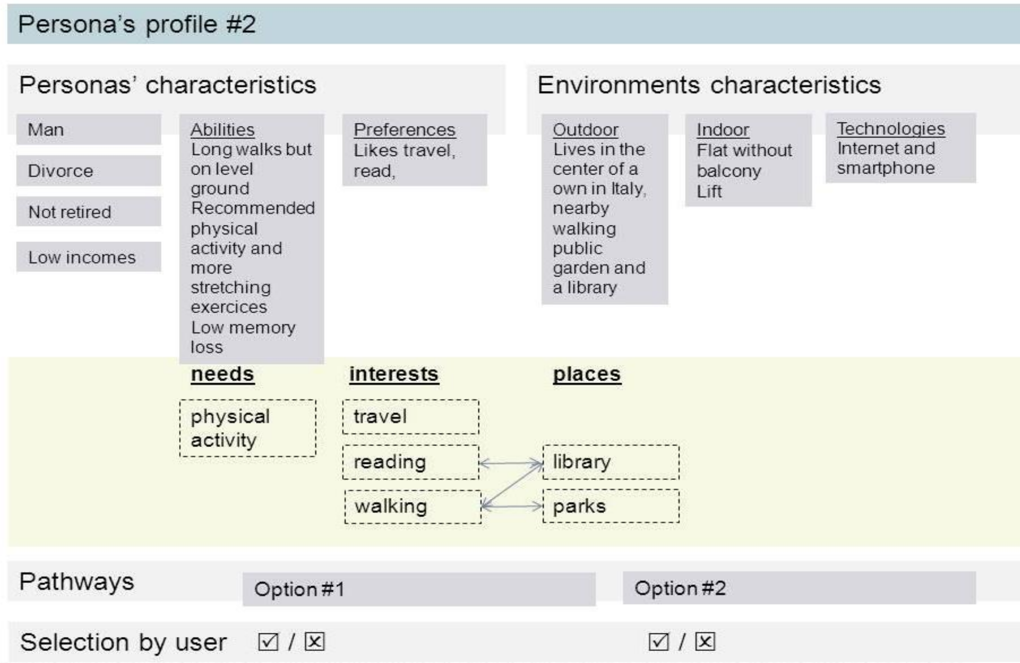

Figure 5. Diagram of Persona example 2.

3.2. Generated Personas

NESTORE's card model includes general information such as gender, age, country or socio-economic status, and more specific details about how many people live in the home, the main characteristics of the living space (size, existence of stairs, balcony or garden), where they live (urban or rural), web connection level, if they have a pet or not. Environmental information (weather and humidity that could affect their activities in daily living) is also provided. Finally, Personas' status in relation to the different domains is provided with a colour based scale codes associated to 3 levels (red when is in an unsatisfactory level of achievement, orange when is moderated and green when is satisfactory) and a definition of the status and target with a narrative description that includes information about preferences and values.

The first version of the card was more synthesized and included few aspects that could help to determine the selection of the pathway. Another version was designed including more information about Personas interests and needs. However, finally, the last version was improved to be more graphic, informative and comprehensive, in order to give important information in a short glance (Figure 6).

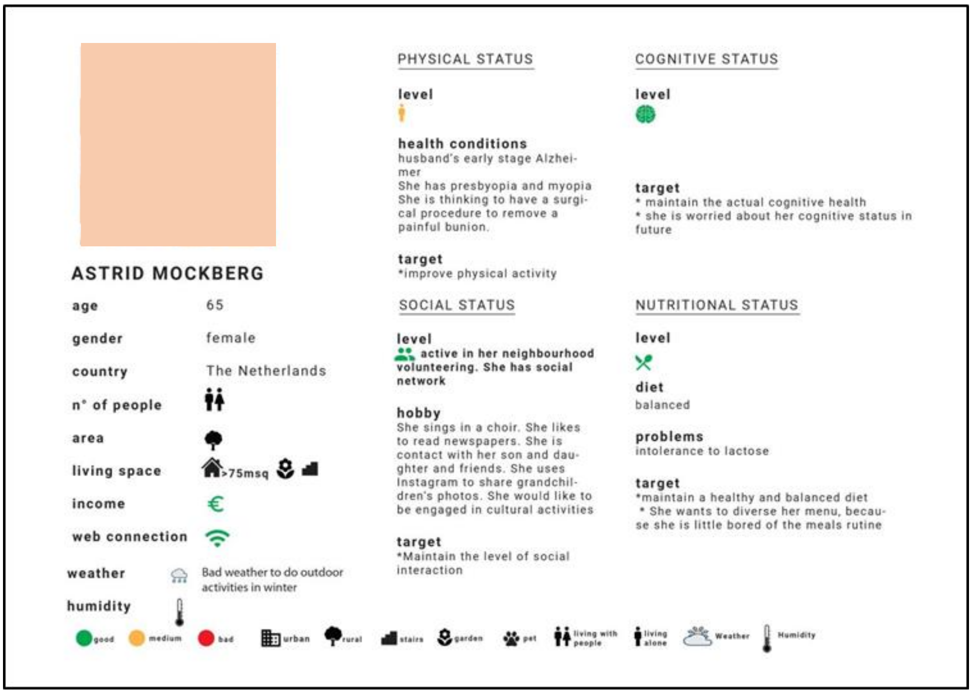

Figure 6. Final version card of NESTORE Personas.

3.3. Personas characterization

The final 10 Personas illustrate the main characteristics of target users trying to represent their lives and preferences. They are characterized as follows:

- NESTORE Personas represent 50% men and 50% women from 65 to 75 years old.
- Geographical representation of Personas is heterogeneous. The most representative nationalities are Italy, Spain, and the United Kingdom. However, there are also profiles from The Netherlands, Belgium, Germany, and France.
- Personas were designed to represent different social status: three couples, four singles, two living apart together, and one widowed. Only 2 personas have a pet.
- The economic status was represented with the distribution: four upper level, four middle, and two with low incomes.
- Living space, personas live: 2 lower than 50msq, 4 between 50-60msq, and 4 upper 60msq. 8 personas have a garden and 5 stairs. Five people live in rural area and five in urban area.

- Four Personas live in geographic area with bad weather (adverse and extreme climate conditions to do outdoor activities i.e too hot, too cold, too windy, etc.) on average, and 6 in areas with high humidity that in summer creates discomfort to do outdoor daily activities.
- Internet Connection is diverse, from good (n=7), medium (n=2) and bad (n=1).

3.4. Domains in personas' profiling

Personas present also diversity in relation to the different domains in order to have a wider spectrum that could enrich the views and understanding of potential needs and preferences of future end users.

3.4.1. Physiological status and physical activity domain

Although, NESTORE users are defined as healthy older people, it is relevant to include different type of health conditions (not severe chronic diseases) very common and prevalent in elderly population. There are acute illnesses or health conditions that could determine behaviours or affect system functionalities, and because of this, experts pointed out the need to consider in health status some conditions.

Since NESTORE pathways are based on user needs to maintain or improve a defined physiological status, Personas included profiles with different physical activity levels and several behavioural targets. According to this, Personas have a wide range of physical activity level and profiles with high (2 profiles), medium (5 profiles), and low activity (3 profiles) are included included based on the number of steps per day. Similarly, Personas include, for example, profiles who need to improve aerobic activities such as walking but do not need stretching exercise as well as aerobically fit subjects who need to increase the frequency of strength activities.

3.4.2. Social interaction domain

Personas were defined to describe different living conditions, even though the majority of profiles are characterized by medium or high levels of interaction. Personas are retired or working part-time, or taking care of grandchildren or other family members. In addition, there are very active profiles, involved in volunteering activities, hobbies (music, reading, travel, etc.), doing cultural or training activities. However, it was decided to also include perceptions of some loneliness in some profiles that could affect the perception of the quality of interactions with others.

We also defined the use of social networks. Pathways considered in the social interaction domain were defined to maintain or improve a person's social opportunities or skills.

3.4.3. Cognitive and emotional domain

Personas were described to include a broad range of cognitive and emotional status. The majority of profiles (n=7) have a good cognitive status, but they could be worried to maintain it, or they could be worried about future memory loss. Because of this, the personas' profiles have interest in pathways such as: "retain/improve broader thinking skills", "retain/improve memory" or "retain/ improve everyday mental skills".

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

3.4.4. Nutrition

Regarding nutrition domain, the majority of participants have a well-balanced diet, but they want to improve some aspects as the diversity of menus, introduce some foods and nutrients such as proteins or fibre from vegetables or fruits, or reduce others as cakes, fats, etc. Some of them need to increase the intake of water.

In addition, two profiles with digestive problems were included to help identify other needs and preferences that could affect the diet behaviour or food elections. Two Personas are overweight, but their target was defined to diversify menus and balance their diet because it is possible that in existence of overweight problems the user decides that he/she does not want to reduce body weight or fat mass. However, the NESTORE System will firstly encourage him/her to lose weight (explaining the benefits, risk factors, etc.). If users continue interested in diet, then NESTORE system will understand their needs and preferences in order to propose a pathway that includes some activities, which could encourage a behaviour change, if possible. In addition, four Personas have a different diet profile because one has food allergies, one has lactose intolerance, and two are vegan or vegetarian in order to introduce some diversification in profiling.

3.5. Clustering to mimic Personas

Following the Personas reasoning, we proceeded with clustering users and define personalized recommendations for each of the resulting clusters. We foresaw clusters where each of them would be characterized by a set of features, such as singles that live in a city or people in their 70s with low physical status. In the private version of D4.3. we explain in depth how this approach was executed and the reasons why we opted for a very distinct approach: a *tagging system* that helped us to personalize all recommendations.

4. The tagging system

Tagging is the process of assigning metadata to content in the form of keywords. Tags are used in NESTORE to characterize users and coaching events (CEs). The system uses the pre-set tags to look for the CEs that users can do fitting their interests the best.

In a broader sense, tagging can be seen as the action of connecting a relevant user-defined keyword to a document, image or video, which helps users to better organize and share their collections of interesting stuff (Song, Zhang, & Giles, 2011). With the rapid growth of Web 2.0, tagged data is becoming more and more abundant on the social network websites in a free and unlimited manner. Consequently, collaborative tagging has grown in popularity on the web, on sites where users can tag photographs, bookmarks and other content.

Collaborative tagging describes the process by which many users add metadata in the form of keywords to shared content. In contrast, traditionally, tagging took place under the concept of categorizing or indexing documents in their collections by an authority, such as a librarian or the authors of the documents. Collaborative tagging is most useful when there is nobody in the *librarian* role or there is too much content for a single authority to categorize.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

Many benefits favour the tagging approach. For instance, tags do not need any structure, which eases the creation of tags *ad hoc*. Tags help to establish relationships between content and the people connected to the content. A tagging system scales exceptionally well, thereby suiting the miscellany of digital space. The popularity of tagging systems is creating an Internet that is marked up in metadata, that is machine-readable, sortable and sharable.

Tagging has been broadly used for distinct purposes, such as in natural language processing (Màrquez & Rodríguez, 1998), or in social bookmarking tools (Lund, Hammond, Flack, Hannay, & NeoReality, 2005). Little academic research work has been invested in tagging systems to date, one of the most well-known studies is the one conducted by Golder and Huberman, who study the information dynamics in collaborative tagging systems discussing the information dynamics in such a system and its semantic difficulties in (Golder & Huberman, 2005).

Tagging can also be embedded in a recommendation system, as the following works demonstrate by using collaborative filtering algorithms (Tao-Stutter, Marinho, & Schmidt-Thieme, 2008) , (Shepitsen, Gemmell, Mobasher, & Burke, 2008), and (Konstas, Stathopoulos, & Jose, 2009).

In NESTORE, we foster a restricted tag creation approach, in which only authors are allowed to set tags. Authors are defined as the domain experts and the tagging system developers, the former for the domain specific knowledge and the latter for the system know-how. Thus, tags are limited and their meaning is pre-set. Users' profiles are tagged automatically thanks to a given ontology (D2.3) and expert-driven criteria. This process is carried out by a case-based reasoning algorithm that describes how the profiles of users are translated to tags. For example, a lactose-intolerant person who lives in a city with beach, and enjoys swimming will be tagged with [lactose-intolerant, beach, swim]. CEs are tagged based on their requirements and specifications; e.g., the CE "Why don't you go to swim at the beach?" will be tagged with [swim, beach], and the CE "Add milk to your morning tea" will be tagged with [milk, morning]. In this example, the former CE could be suggested to this user, whereas the latter would be filtered out. This constraint-based system is combined with a hybrid recommendation system, which employs collaborative filtering (CF), content-based filtering (CBF) and a novel filtering technique to assure heterogeneity that will henceforth be referred to as log filtering (LG). These four techniques together form the complete tagging system (see Figure 7).

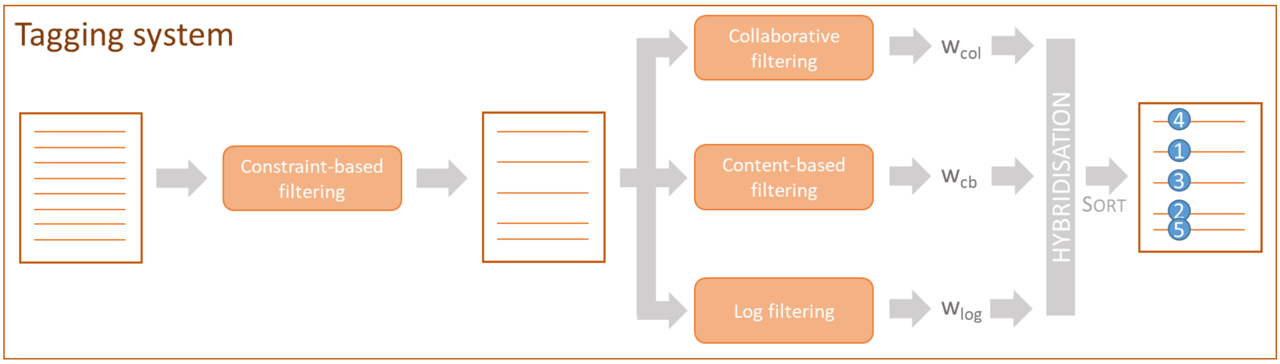

Figure 7. The tagging system and its modules.

4.1. Constraint-based filtering

The constraint-based filtering module deals with incompatibilities between users and CEs at two different levels: *user restrictions* and *availability*.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

- **User restrictions**. Restrictions take into consideration: refused foods; CEs requirements, such as bike, dog or stairs.
- **Availability**. CEs also contain time related tags when necessary. CEs tagged with "lunch" will only be recommended if users are available some days at lunchtime. This will be further explained in D4.4: Dynamic DSS for an Intelligent Coach.

4.2. Collaborative filtering

Collaborative filtering (CF) finds users in NESTORE that shared the same interests in the past to predict what the current user will be interested in (Brusilovsky, Kobsa, & Nejdl, 2007). There are many possible techniques to use that can be grouped in *memory-based* and *model-based* approaches. The former kind finds similar users based on cosine similarity or Pearson correlation and takes weighted average ratings. The main advantage of memory-based approaches is the ease of interpretation of their results but the price to pay is the reduction of performance when data is sparse. Model-based approaches use machine learning (e.g., Principal Component Analysis or Neural Networks) to find user ratings of unrated items. They deal with missing and sparse data but at great cost: inference is intractable due to hidden and latent factors. Since NESTORE follows a user-centric approach, we prioritize the fact of being able to explain the outcome over performance. Hence, we opted for a memory-based approach. In the private version of the deliverable you can find a detailed explanation of the steps carried out under the technical development.

4.3. Content-Based Filtering

Content-Based Filtering (CBF) generates recommendations from two sources: the tags associated with CEs and the ratings that a user has given to them. We treat each CE as a user-specific classification problem and learn a classifier for the user's likes and dislikes based on CEs' tags. Content-based systems are based on the idea that if you liked a certain item you are most likely to like something that is similar to it. It is able to recommend new and unpopular items, although it highly depends on the quality of the tags; meaningless or lacking tags would lead to an unnecessary set of CEs filtered out. In the private version of the deliverable you can find the explanation of the the technical approach followed.

4.4. Log filtering

The module *log filtering* of the tagging system aims to assure diversity when sending CEs' recommendations.

This module avoids the scenario in which the system always recommends the same based on the likings of the user. In this case, diversity is formulated in terms of specific CEs, but the system could also be enriched by reformulating diversity in terms of CEs main actions. For example, all running related activities could be grouped together in the physical activity domain. The technical details of this implementation are presented in the private version of D4.3.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

4.5. Hybridisation

The results from applying collaborative filtering, content-based filtering and log filtering are all mixed up in the hybridisation module. Here, weights are assigned to each of the modules and together create a total score that sorts the CEs in a personalized manner for each user. Thus, all of the system's capabilities are brought to bear on the recommendation process. The technical details of the hybridization module are explained in the private version of the deliverable.

5. Coaching events

A three-layer coaching timeline (refer to section 3 of D4.2) is proposed in NESTORE to adapt better to users' needs and preferences. This layered system allows them to 1) choose a general goal (pathway), 2) select the kind of activities that they prefer (coaching activity plans), and 3) accomplish their objective performing specific training scheduled by the system (coaching events). This makes NESTORE a user-friendly framework that converts general goals into specific actions supporting their accomplishment and, therefore, users' fulfilment.

A *Coaching Activity Plan* (CAP) is a category of activities that can be stratified into a set of specific activities, such as "run", which could be fragmented into "run in the park" or "run on the beach". Each pathway has many CAPs associated to tailor the coaching plan to users' preferences. *Coaching events* (CEs) are sets of activities scheduled by the system throughout the day/week. Its enjoyment and willingness to be repeated are assessed by the user using the five-level Likert scale. This feeds the users' profile helping the DSS to create personalised recommendations.

Experts came up with a set of CEs specific for each of the available pathways. Each CE has followed an enrichment process as it is exemplified in Figure 8. Version 0 contains the core content of the CE and it is enriched with contextual information, e.g., the name of a specific seniors' centre (see example in Figure 8). If information that is location-dependent is added, it appears the need of creating three different CEs, one per each pilot location, that will have language tags assigned so as to only recommend CEs that contain Dutch information to Dutch people.

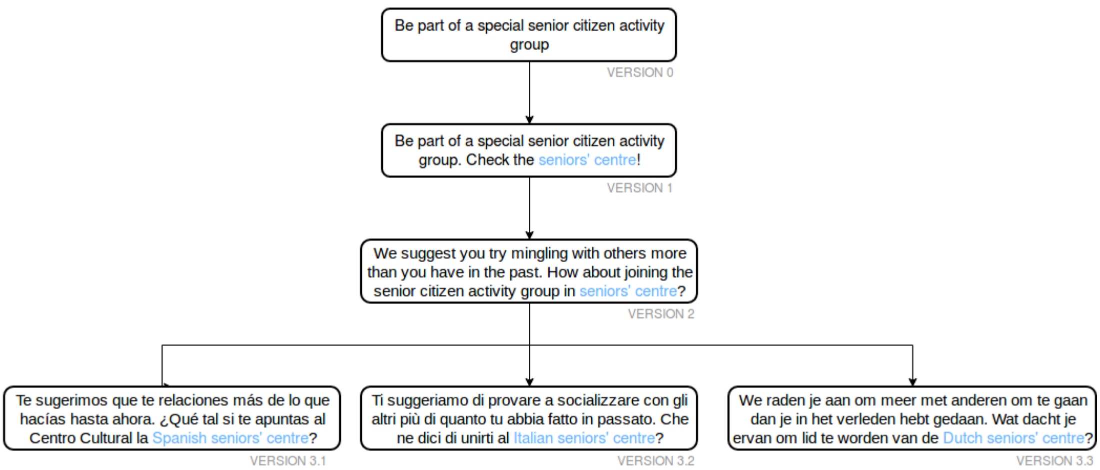

Figure 8. Visualization of the CEs' enrichment workflow.

In the private version of the document we present the list of CEs used and explain the kinds of tags we contemplate and its creation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

6. Modelling the user

By modelling the user we aim at building up a conceptual understanding of the user. The main goal of this technique is the customization and adaptation of the NESTORE system to the users' specific needs. In this section, we will explain how the system represents the user.

After analysing Personas and complementing the information with domain experts, a twofold user profile is proposed:

- **Static profile.** It is formed by the status and preferences of the user and it is characterized by containing non-varying attributes. Concretely it includes demographic characteristics, attributes regarding the context where the user lives, physical and physiological aspects and baseline data of the various domains.
- **Dynamic profile.** It is built dynamically while receiving data from sensors, applications and contextual APIs. It is foreseen to receive daily indicators about the different domains and also contextual information (i.e. current weather conditions). This profile is constructed in the context of tasks T4.1 and T4.2.

6.1. Static profiling

Static profiling is the process of analysing a user's static and predictable characteristics. As it has been said previously, users' static features comprehend factual data, such as the idiosyncrasy of their residence (e.g. do they live in a rural or in an urban area?), or their diet routines (e.g. is meat part of their diet?) as well as inter-individual differences in the other NESTORE domains (marital status and perceptions of loneliness, cognitive functioning, physical fitness, etc.). They also describe the environment and context of users.

One of the uses of static profiling will be the cluster of users, the resulting groups of which will be inputted into the DSS to make thoughtful recommendations. Considering that real data will not be available until the pilots take off, a data simulator will be implemented to cope with the absence of data creating, thus, solid fundamentals for the clustering process and the recommendation system. Getting into detail, the static profile simulator will generate a population of users who will be described by its fact-based properties.

The static profile simulator will be only useful if it closely mirrors real-world outcomes. In order to achieve that, the target population of the NESTORE project will be simulated. In other words, the profile of people from 65 to 75 years old, living on their own, mainly retired or recently retired will be imitated. The only restriction applied will be territorial: the static profile simulator will only focus on European citizens since the pilots will only take place in Europe; thereby, the final results will be more accurate.

Firstly, the static features which define the users, as well as the range of those variables, will be defined. That task was already carried out by the experts and it is tackled in section 6.1.1. Next and lastly, the approach to build the static profile simulator assuring the reliability of the created users will be presented in section 6.1.2.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

6.1.1. Static profile variables

To build the static profile of a user, not only the four well-known well-being domains (physical activity, cognitive, social and nutrition) need to be considered. The user's context is a quite new feature in user profiling that will help to characterize the situation of the user. There are different types of contexts or contextual information that can be modelled within a user profile (Goker & Myrhaug, 2002), but we will focus our attention on the environmental and the demographic context.

The information presented below, comes from:

- models described in D2.1;
- recommendations and guidelines defined in D2.2;
- existing knowledge from domain experts and other evidence-based sources.

After deciding on the obtainable information to profile the users, a collection of variables with the values they can take has been defined. Those have been split per kind of contextual feature (demographic and environmental), and per well-being domain (physiological status and physical activity behaviour, nutrition, cognitive and mental status, and social behaviour). Besides, a category called activities has been added to include user's routines and preferences.

Demographic information is to a great degree relevant to group people according to their culture and generation. Due to the scope of the DSS, there is no need of depicting the participant's culture. Table 1 shows the variables which best characterize users' demographic context.

Table 1. Demographic variables.

Variable	Domain
Age	[65,75]
Gender	F, M

The environmental context captures the entities that surround the user. These entities can, for instance, be services, temperature, light, humidity, noise, and people (Schiaffino, 2009). Table 2 displays a compendium of variables that provide contextual information about the environment of users.

Table 2. Environmental variables.

Variable	DOMAIN
Living area	Urban, rural
Stairs	Yes, no
Garden	Yes, no
Pet - Dog	Yes, no

The factual data that best describes the physiological status of users mainly comes from their anthropometric characteristics, presented in Table 3.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

Table 3. Variables related to the physiological status.

Variable	Domain
Body height	[m]
Body weight	[kg]
Body mass index (BMI)	[kg/m^2]
Fat mass	[%]
Fat-free mass	[%]

The nutritional domain will basically be characterized by the dynamic profile. Only the list of refused foods will be considered to create the nutritional static profile, as it is presented in Table 4.

Table 4. Variables related to nutritional domain.

Variable	Domain
Refused foods	Text

The cognitive and social domain will be characterized by the dynamic profile. Only one variable is missing to complete the static profile: the activity preferences of the user, contained in Table 5.

Table 5. Variables related to user preferences.

Variable	Domain
Activity preferences	Walk, bike, swim, dance, extreme sport…

The list of variables presented here differs from the one presented in the previous version of this deliverable, D4.3.1. Actually, it is shorter. Now that the totality of the system is clearly sketched, the content and scope of the static and dynamic profile have evolved to what it is explained here.

6.1.2. Static profile simulator

Now that the variables and their domain have been defined, we can proceed to the introduction of the static profile simulator. The simplest approach would be to create users assigning a value randomly drawn for each of the variables. However, that could lead to very unreal scenarios; consequently, that line of action is ruled out.

To assure the reliability of the created users, the methodology used to develop the static profile simulator is twofold: restrictions are applied to avoid nonsensical scenarios, and statistical indicators are used to mirror the population of older European citizens.

The target population of NESTORE can be characterized by statistical indicators such as the ones provided by *Eurostat*. Eurostat is the statistical office of the European Union, whose mission is to provide high quality statistics for Europe. Eurostat offers a compilation of indicators ranging from general and regional measures to economy and finance indexes, to population and social conditions indicators.

Demographic indicators are drawn to simulate the population by age and gender per country (Nations, 2017). A summary of this information is presented in Table 6.

Table 6. Demographic indicators per country. It depicts the distribution of older people (age of 65 to 75) per country in Europe. Source: (Nations, 2017).

Country	**Older population of Europe (%)**
Bulgaria	1.63
Czech Republic	2.26
Hungary	1.91
Poland	6.37
Romania	3.47
Slovakia	0.89
Denmark	1.26
Estonia	0.24
Finland	1.23
Iceland	0.05
Ireland	0.71
Latvia	0.39
Lithuania	0.53
Norway	0.94
Sweden	2.09
United Kingdom	12.44
Croatia	0.81
Greece	2.07
Italy	12.71
Malta	0.09
Portugal	2.15
Slovenia	0.38
Spain	8.39
Austria	1.69
Belgium	1.99
France	11.95
Germany	16.32
Luxembourg	0.08
Netherlands	3.38
Switzerland	1.56

It has not been possible to find statistic indicators which describe the specific interests or activities carried out by older European citizens. However, little information about the digital activities performed by that population (Eurostat, Internet use and activities., 2017) has been found and used as a baseline to define the activity profile of users.

Little information has been found as of today in relation to the variables defined for each of the well-being domains. Regarding the physiological profile of users, statistic indicators which describe BMI distributions (Eurostat, Body mass index (BMI) by sex, age and educational attainment level, 2014) have been used. In relation to the nutritional behaviour, none of the variables from Table 4

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

has been found. The static variables whose distribution cannot be supported by statistical indicators, have been randomly sampled using thresholds defined by experts.

6.2. Dynamic profiling

Dynamic user models allow a more up to date representation of users. Interactions with the system are noticed and influence the user model. The model can thus be updated and take the newest users' data into account. The dynamic profile is formed by various layers presented as subsections.

6.2.1. Weather

The DSS uses the weather forecast to assess whether CEs can take place at a given day.

6.2.2. Users' preferences

This comes by the ratings given by users to CEs. All the ratings assign to a CE are saved, meaning that we have the log of ratings for a CE. In order to properly fit users' change of preferences, the final rating of a CE is computed as the mean of the latest rating and the average of the previous ones following the same logic, as it is described in the following equation:

$$
\begin{gathered}
r_t' = avg(r_t, r_{t-1}') \\
r_{t-1}' = avg(r_{t-1}, r_{t-2}') \\
\vdots
\end{gathered}
$$

Thus, the newest data weights more.

6.2.3. Daily routines

A routine is a habit or sequence that does not vary over time. In NESTORE, we aim at identifying daily routines to tailor the system to users. The system initializes setting a default value for daily routine events based on the residence country of users. As Table 7 shows, a value is assigned for each of the well-known daily routine events. These values are updated thanks to the data gathered by the multiple sensors integrated in the NESTORE system.

Table 7. Default values for routine daily events.

EVENT	ITALY	SPAIN	THE NETHERLANDS
Awakening	08:00	08:00	07:00
Breakfast	08:15	08:15	07:15
Morning	09:00	09:00	08:00
Lunch	13:00	13:00	12:00
Afternoon	14:00	14:00	13:00
Snack	17:00	17:00	15:00
Evening	18:00	18:00	16:00
Dinner	20:30	20:30	18:00
Night	21:00	21:00	21:00

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

These data is used by the DSS to propose CEs at the most approppiate time. An example is presented in the private version of the document.

6.2.4. Indoors / outdoors indicator

The detection of users' indoor behaviour comes from environmental sensors. This information is used when scheduling messages, mainly during the first two weeks when the system is gathering data from the user who is supposed to fill in multiple tests and questionnaires.

Five environmental Bluetooth Low Energy (BLE) beacons are deployed in the user's home to give information about the indoor mobility of users and their interaction with relevant Point of Interests (PoIs) of the house. During the installation of the NESTORE system, the five BLE beacons are deployed in the most used areas of the house (i.e., kitchen, living room, and bedroom) and on commonly used furniture, like the door of the fridge and the bathroom door.

The BLE beacons, leveraging the capabilities offered by the radio propagation of the Bluetooth signal, provide information about the proximity of the data-gathering device (i.e., the wristband) worn by the user with the beacon itself, therefore the position of the user in the area in which the beacon is installed. Furthermore, the beacons embed an accelerometer that is activated when it is moved, therefore when the furniture is used.

In the private version of D4.3. we explain in depth which are the indicators calculated and how are the computations done.

6.2.5. User habits

The environmental sensor network gathers the information to infer user habits. By monitoring the activations of sensors, indicating the opening/closing of doors and room occupation of users during their time spent at home, it is possible to retrieve heterogeneous and multivariate time-series over long periods. These time-series can be used to learn recurrent behaviours of users in their daily activities by analysing the time variations of several parameters like the room occupied by the user and qualitative activity level.

In the literature, many applications, which serve mainly to support diagnosis, effectively deal with temporal sequences, encouraging the development of the related "time series mining" research field (see (Antunes & Oliveira, 2001) and (Roddick & Spiliopoulou, 2002) for overviews). Discovery algorithms aim at extracting important pattern such as similarities, trends, or periodicity, with the aim of recurrent pattern description or prediction (Nanopoulos, Alcock, & Manolopoulos, 2001). Pattern discovery in time series is useful for temporal sequences synthesis (Hong & Huang, 2000) as well as for learning tasks like association rules mining (Das, et al., 1998) and (Höppner, 2001), classification (Keogh & Pazzani, 1998), and clustering (Vlachos, Kollios, & Gunopulos, 2002).

Encouraging results in building behavioural profiles of a person living in a smart home are highlighted in (Duchêne, Garbay, & Rialle, 2007) where a feature mining algorithm is presented. Also in (Mahmoud, Lotfi, & Langensiepen, Behavioural pattern identification in a smart home using binary similarity and dissimilarity measures, 2011) and (Mahmoud, Lotfi, & Langensiepen, Abnormal behaviours identification for an elder's life activities using dissimilarity measurements, 2011) behavioural pattern identification methods are proposed using binary similarity and dissimilarity

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

measures on data generated from occupancy sensors including door and motion sensors in a smart home.

Understanding the behavioural profile of a user is extremely important to detect behavioural changes possibly related to a deterioration of the user physical and psychological status. This is an emerging research topic addressed in several works for supporting the independent living of older people. In (Lotfi, Langensiepen, Mahmoud, & Akhlaghinia, 2012), the authors describe a solution based on home automation sensors, including movement sensors and door entry point sensors. By monitoring the sensor data, important information regarding anomalous behaviour is identified using supervised approaches to predict the future values of the activities for each sensor in order to inform the caregiver in case anomalous behaviour is predicted.

Within the scope of the NESTORE project, we intend to address the unsupervised detection of these forms of behavioural anomalies, since collecting ground truth information for a long period in a real house can be very obtrusive for the user. For this reason, we will focus on motif search on sensory data collected in the test sites, represented as time series, by exploiting the results obtained in the field of time series motifs discovery (Fernández, Llatas, Benedi, García, & Gómez, 2013) (Van der Aalst, et al., 2003). Time series motifs are approximately repeated patterns found within the data. Such motifs have utility for many data mining tasks, including rule-discovery, novelty-detection, summarization and clustering. Since the formalization of the problem and the introduction of efficient linear time algorithms, motif discovery has been successfully applied to many domains, including medicine, motion capture, robotics and meteorology.

In the NESTORE scenario, physical displacements of users in their vital environments can offer information about the change of their individual behaviour, capturing all the areas (rooms in the home) visited by the user over time. In this domain, useful insights are given by (Lotfi, Langensiepen, Mahmoud, & Akhlaghinia, 2012) in the field of the representation of sensor data for further analysis of behaviour deviations detection. Authors propose two different techniques for the summarization of data: combined activity of daily living signal as a time series and start time and duration. The first method involves the use of a signal assuming different levels for each activity of daily living, where each level of the combined signal represents one of the sensors triggered by the user. In the second one, the signal is represented by the start time and the duration of an event representing the user entering a room and the duration that she stays in a specific location.

Different motif discovery approaches have been proposed in the literature: in (Fernández, Llatas, Benedi, García, & Gómez, 2013) authors present a solution based on process mining (also known as workflow mining) (Van der Aalst, et al., 2003) and (van der Aalst, 2011). It allows workflow inference from event or activity logs. A workflow is a formal representation of a process designed to be automatized. That means that process mining technology can be used to infer graphs understandable by human experts (workflows) using the daily actions collected by ambient intelligence environments. This allows the experts to understand the behaviour process of the individual and to compare it with previous inferences in order to detect specific behaviour changes and patterns.

Another important technique used in the field of motif discovery is represented by “stigmergy”. This is a term derived from the research on the foraging behaviour of ants, which communicate with each other exchanging information through the modification of the environment and the information can only be accessed when an ant visits the place marked by another ant. Several works used this technique in order to infer motifs in time series related to different fields, from DNA and

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 769643

biological sequences (Yang, Liu, & Chuang, 2001) (Bouamama, Boukerram, & Al-Badarneh, 2010) to intrusion detection systems (Cui, Beaver, Potok, & Yang, 2011).

Classical process mining algorithms, like Parallel Activity-based Log Inference Algorithm (PALIA) (Fernández-Llatas, Meneu, Benedi, & Traver, 2010), the alpha algorithm (Van der Aalst, et al., 2003), heuristic miner (Weijters & Ribeiro, 2011) or the genetic process mining algorithm (Medeiros, Weijters, & Aalst, 2007), have been tested in laboratory conditions in previous works (Weijters & Ribeiro, 2011) (Fernández, Lázaro, & Benedí, Workflow mining application to ambient intelligence behavior modeling, 2009). A helpful technique has been proposed by (Yankov, Keogh, Medina, Chiu, & Zordan, 2007) based on uniform-scaling invariant motif discovery. This approach overcomes one of the biggest limitations of existing state-of-the-art techniques in pattern discovery regarding the possibility of discovering pattern occurrences having the same time length, failing to capture similarities when the occurrences are uniformly scaled along the time axis. For this reason, in NESTORE, an in depth analysis of these techniques have been performed, in order to propose a cost effective technique to detect novel patterns in user's room occupation.

7. Bibliography

Antunes, C. M., & Oliveira, A. L. (2001). Temporal data mining: An overview. *Proceedings of the KDD workshop on temporal data mining*, 1-13.

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C., & Vaughn, B. (2012). The aasm manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specification*.

Bouamama, S., Boukerram, A., & Al-Badarneh, A. F. (2010). Motif finding using ant colony optimization. *Swarm Intelligence*, 464-471.

Brusilovsky, P., Kobsa, A., & Nejdl, W. (2007). Methods and Strategies of Web Personalization. *Lecture Notes in Computer Science, 4321*.

Cooper, A. (1999). *The inmates are running the asylum. Indianapolis, IA: SAMS.* Macmillan.

Cooper, A., Reimann, R., & Cronin, D. (2007). *About face 3: the essentials of interaction design.* John Wiley and Sons.

Cui, X., Beaver, J., Potok, T., & Yang, L. (2011). Visual Mining Intrusion Behaviors by Using Swarm Technology. *Proceedings of the 44th Hawaii International Conference on System Sciences*, 1-7.

Das, G., Lin, K. I., Mannila, H., Renganathan, G., & Smyth, P. (1998). Rule Discovery from Time Seriess. *KDD, 98*, 16-22.

Duchêne, F., Garbay, C., & Rialle, V. (2007). Learning recurrent behaviors from heterogeneous multivariate time-series. *Artificial intelligence in medicine, 39*(1), 25-47.

European Comission. (2019). People in Europe. Retrieved September 1, 2019, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals

European Comission. (n.d.). Eurostat's statistics on information and communication technologies (ICTs). Retrieved December 4, 2017, from http://ec.europa.eu/eurostat/web/information-sciety/overview

European Comission. (n.d.). People in Europe: statistics on the ageing society. Retrieved December 4, 2017, from http://ec.europa.eu/eurostat/statistics-explained/index.php/People_in_the_EU_%E2%80%93_statistics_on_an_ageing_society#The_elderly_living_alone

European Comission. (n.d.). Statistics on an ageing society: the Elderly living alone. Retrieved December 4, 2017, from http://ec.europa.eu/eurostat/statistics-explained/index.php/People_in_the_EU_%E2%80%93_statistics_on_an_ageing_society#The_elderly_living_alone

European Parliamentary Research Service. (2014). Older people in Europe. Retrieved from http://www.europarl.europa.eu/RegData/bibliotheque/briefing/2014/140811/LDM_BRI(2014)140811_REV1_EN.pdf

Eurostat. (2014). *Body mass index (BMI) by sex, age and educational attainment level.* Retrieved July 20, 2018, from EU-EHIS survey code: hlth_ehis_bm1e: http://ec.europa.eu/eurostat/product?code=hlth_ehis_bm1e&language=en&mode=view

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

Eurostat. (2017). *Distribution of population aged 65 and over by type of household.* Retrieved July 19, 2018, from Income and Living Conditions. EU-SILC survey code: ilc_lvps30: http://ec.europa.eu/eurostat/web/income-and-living-conditions/data/database?p_p_id=NavTreeportletprod_WAR_NavTreeportletprod_INSTANCE_CEM7npyJJgVL&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1#

Eurostat. (2017). *Distribution of population over 18 years by most frequent activity status, age group, and sex.* Retrieved July 19, 2018, from Income and Living Conditions. SILC survey code: ilc_lvhl902: http://ec.europa.eu/eurostat/web/income-and-living-conditions/data/database?p_p_id=NavTreeportletprod_WAR_NavTreeportletprod_INSTANCE_CEM7npyJJgVL&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1#

Eurostat. (2017). *Frequency of getting together with family and relatives or friends by sex, age and educational attainment level.* Retrieved July 19, 2018, from Income and Living Conditions: http://ec.europa.eu/eurostat/web/income-and-living-conditions/data/database?p_p_id=NavTreeportletprod_WAR_NavTreeportletprod_INSTANCE_CEM7npyJJgVL&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1#

Eurostat. (2017). *Immigration by age and sex*. Retrieved from http://ec.europa.eu/eurostat/en/web/products-datasets/-/MIGR_IMM8

Eurostat. (2017). *Internet use and activities.* Retrieved July 20, 2018, from Code: isoc_bde15cua, 2017.: http://ec.europa.eu/eurostat/web/products-datasets/-/isoc_bde15cua

Eurostat. (2017). *Participation in formal or informal voluntary activities or active citizenship by sex, age and educational attainment level.* Retrieved July 19, 2018, from Income and Living Conditions. EU-SILC survey code: ilc_scp19.: http://ec.europa.eu/eurostat/web/income-and-living-conditions/data/database?p_p_id=NavTreeportletprod_WAR_NavTreeportletprod_INSTANCE_CEM7npyJJgVL&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=1#

Eurostat. (2018, July 18). *Product Datasets: Immigration by age and sex*. Retrieved from http://ec.europa.eu/eurostat/en/web/products-datasets/-/MIGR_IMM8

Fernández, C., Lázaro, J. P., & Benedí, J. M. (2009). Workflow mining application to ambient intelligence behavior modeling. *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, 160-167.

Fernández, C., Llatas, J., Benedi, M., García, J., & Gómez, V. T. (2013). Process mining for individualized behavior modeling using wireless tracking in nursing homes. *Sensors, 13*(11), 15434-15451.

Fernández-Llatas, C., Meneu, T., Benedi, J. M., & Traver, V. (2010). Activity-based process mining for clinical pathways computer aided design. *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Societe*, 6178-6181.

Goker, A., & Myrhaug, H. (2002). User context and personalization. *Proceedings of ECCBR Workshop on Case Based Reasoning and Personalization*.

Golder, S., & Huberman, B. A. (2005). The structure of collaborative tagging systems. *arXiv prepring cs/0508082*.

Harvey, A. G., Stinson, K., Whitaker, K. L., Moskovitz, D., & Virk, H. (2008). The subjective meaning of sleep quality: a comparison of individuals with and without insomnia. *Sleep, 31*(3), 383-393.

Holsapple, C. W., Whinston, A. B., Benamati, J. H., & Kearns, G. S. (1996). Decision support systems: A knowledge-based approach. 161-193.

Hong, P., & Huang, T. (2000). Learning to extract temporal signal patterns from temporal signal sequence. *Proceedings of the 15th International Conference on Pattern Recognition, 2*, 648-651.

Höppner, F. (2001). Discovery of temporal patterns. *Principles of Data Mining and Knowledge Discovery*, 192-203.

Huang, Z. (1997). Clustering large data sets with mixed numerical and categorical values. *Proceedings of the First Pacific Asia Knowledge Discovery and Data Mining Conference*, 21-34.

Keogh, E. J., & Pazzani, M. J. (1998). An Enhanced Representation of Time Series Which Allows Fast and Accurate Classification. *Clustering and Relevance Feedback, 98*, 239-243.

Konstas, I., Stathopoulos, V., & Jose, J. M. (2009). On social networks and collaborative recommendation. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval* (pp. 195-202). ACM.

Lastella, M., Roach, D., Halson, S. L., & Sargent, C. (2015). Sleep/wakebehaviours of elite athletes from individual and team sports. *European Journal of Sport Science, 15*(2), 94-100.

Lotfi, A., Langensiepen, C., Mahmoud, S. M., & Akhlaghinia, M. J. (2012). Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. *Journal of Ambient Intelligence and Humanized Computing, 3*(3), 205-218.

Lund, B., Hammond, T., Flack, M., Hannay, T., & NeoReality, I. (2005). Social bookmarking tools (II). *D-Lib magazine, 11*(4), 1-1.

Mahmoud, S. M., Lotfi, A., & Langensiepen, C. (2011). Abnormal behaviours identification for an elder's life activities using dissimilarity measurements. *Proceedings of the 4th International Conference on PErvasive Technologies Related to Assistive Environments*, 1-25.

Mahmoud, S. M., Lotfi, A., & Langensiepen, C. (2011). Behavioural pattern identification in a smart home using binary similarity and dissimilarity measures. *Proceedings of the 7th IEEE international conference on intelligent environments (IE)*, 55-60.

Màrquez, l., & Rodríguez, H. (1998). Part of speech tagging using decision trees. In *European Conference on Machine Learning* (pp. 25-36). Springer.

Medeiros, A. K., Weijters, A. K., & Aalst, W. M. (2007). Genetic process mining: an experimental evaluation. *Data Mining and Knowledge Discovery, 14*(2), 245-304.

Miano, S., Paolino, M. C., Castaldo, R., & Villa, M. P. (2010). Visual scoring of sleep: A comparison between the rechtschaffen and kales criteria and the american academy of sleep medicine criteria in a pediatric population with obstructive sleep apnoea syndrome. *Clinical neurophysiology, 121*(1), 39-42.

Nanopoulos, A., Alcock, R., & Manolopoulos, Y. (2001). Feature-based classification of time-series data. *International Journal of Computer Research, 10*(3), 49-61.

Nations, U. (2017). *World Population Prospects - Population Division 2017*. Retrieved July 18, 2018, from Population by Age Groups: https://esa.un.org/unpd/wpp/Download/Standard/Population/

Nedelec, M., Aloulou, A., Duforez, F., Meyer, T., & Dupont, G. (2018). Thevariability of sleep among elite athletes. *Sports Medicine - Open, 4*(1), 34.

Ohayon, M., Wickwire, E. M., Hirshkowitz, M., Albert, S. M., Avidan, A., Daly, F. J., . . . al, e. (2017). National sleepfoundation's sleep quality recommendations: first report. *Sleep Health, 3*(1), 6-19.

Rechtschaffen, A., & Kales, A. (1968). A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects. *Electroencephalography and Clinical Neurophysiology, 26*.

Roddick, J. F., & Spiliopoulou, M. (2002). A survey of temporal knowledge discovery paradigms and methods. *IEEE Transactions on Knowledge and Data Engineering, 14*(4), 750-767.

Schiaffino, S. A. (2009). Intelligent User Profiling. *Artificial Intelligence An International Perspective*, 193-216.

Shepitsen, A., Gemmell, J., Mobasher, B., & Burke, R. (2008). Personalized recommendation in social tagging systems using hierarchical clustering. In *Proceedings of the 2008 ACM Conference on Recommender systems* (pp. 259-266). ACM.

Song, Y., Zhang, L., & Giles, C. L. (2011). Automatic tag recommendation algorithms for social recommender systems. *ACM Transactions on the Web (TWEB), 5*(1), 4.

Spriggs, W. H. (n.d.). Essentils of polysomnography. 2014.

Tao-Stutter, K. H., Marinho, L. B., & Schmidt-Thieme, L. (2008). Tag-aware recommender systems by fusion of collaborative filtering algorithms. In *Proceedings of the 2008 ACM symposium on Applied Computing* (pp. 1995-1999). ACM.

United Nations. (n.d.). Demographic profile of the older population. *World ageing 1950-2050*. Retrieved Agoust 4, 2917, from http://www.un.org/esa/population/publications/worldageing19502050/pdf/90chapteriv.pdf

van der Aalst, W. M. (2011). Discovery, Conformance and Enhancement of Business Processes. *Springer-Verlag Berlin Heidelberg*.

Van der Aalst, W. M., van Dongen, B. F., Herbst, J., Maruster, L., Schimm, G., & Weijters, A. J. (2003). Workflow mining: A survey of issues and approaches. *Data & knowledge engineering, 47*(2), 237-267.

Vlachos, M., Kollios, G., & Gunopulos, D. (2002). Discovering similar multidimensional trajectories. *Proceedings of the IEEE 18th International Conference on Data Engineering*, 673-684.

Weijters, A. J., & Ribeiro, J. T. (2011). Flexible heuristics miner (FHM). *Proceedings of 2011 IEEE Symposium on Computational Intelligence and Data Mining*, 310-317.

Welk, G. J., Schaben, J. A., & Morrow, J. A. (2004). Reliability of Accelerometry-Based Activity Monitors: A Generalizability Study. *Medicine & Science in Sports & Exercise, 36*(9), 1637-1645.

Wöckl, B., Yildizoglu, U., Buber-Ennser, I., Aparicio Diaz, B., & Tscheligi, M. (2013). Elderly Personas: A Design Tool for AAL Projects focusing on Gender, Age and Regional Differences. *In Proceedings of the 3rd AAL Forum: Partnerships for Social Innovations in Europe*, 26-28. Retrieved December 2018, from http://www.aal-europe.eu/involving-end-users

Wöckl, B., Yildizoglu, U., Buber-Ennser, I., Aparicio Diaz, B., Krujiff, E., & Tscheligi, M. (2012). Basic senior personas: a representative design tool covering the spectrum of European older adults. *In Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12)*, 25-32.

Yang, C. H., Liu, Y. T., & Chuang, L. Y. (2001). DNA motif discovery based on ant colony optimization and expectation maximization. *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 1-6.

Yankov, D., Keogh, E., Medina, J., Chiu, B., & Zordan, V. (2007). Detecting time series motifs under uniform scaling. *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 844-853.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

8. Annex 1 - Personas

- Checklist to refine the final personas selection

Dimension	Persona	Questions to answer	Yes	No
PERSONAL DEMOGRAPHICS AND LIVING CONDITIONS	We assume they are all retired, unless when specified otherwise (E.g. Bertie) I think it is very good to have a persona in a couple but not living together (Antoni R), we might want to set up at least a woman in the same situation.	• Fields are well-defined? (i.e we know if are they retired, still part-time working or full-time work or other paid economic activities, and unpaid activities (i.e volunteering, doing a PhD?) • Age is in the fixed range? (from 65-75 years old) • Gender is balanced? (50%males and 50% females) • Country (i.e we need to search some geographic balance in personas profiles) • Number of people at home[1] • Living area[2] • Living space[3] (i.e we know the size of the space, if there are any barrier, or if they have a pet, or a garden or balcony, Etc.) • Income level (i.e it is interesting to search heterogeneity) • Web connection (from bad to good. Heterogeneity is better)	X	
PHYSICAL STATUS	Good example in the description of physical conditions, as they encompass a specific target/diseases they aim at: Carles Pau, Ana Garcia, Mildred, Louise, Mario, Marcus	The personas profile is informative because: • Helps to identify the physical level, health status or exercise preferences?[4] • What aspects they want to improve or maintain? • The description is helping to identify a pathway? • Is giving us any information that could help us to understand some change, challenge or goal that they want to face?		

[1] It is interesting to search some heterogeneity in personas' profiles. i.e. people living alone (singles, widowing, people living a part, etc.), people living with a couple, etc.

[2] It is interesting to have heterogeneity in urban and rural areas

[3] If we have personas' profiles very descriptive, it would be better.

[4] It is important to have the two targets population (those who wants to improve their status and those who wants to maintain their status)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

SOCIAL STATUS	Good examples of well described social status: Mayte Bo, Astrid, Antoni R, Ignacio Fe, Marcus, etc. the ones that are more extensive in their descriptions.	The personas profile is informative because: • Helps to identify the personas' social situation, beliefs, political convictions and behaviour? • What aspects regarding her/his social life they want to improve or maintain? • The hobbies information is enough to understand their preferences and how they are arranging their leisure time? • The description is helping to identify a pathway? • Is giving us any information that could help us to understand some change, challenge or goal that they want to face?		
COGNTIVE STATUS	Ignacio Fe, Carles Pau, Lia Andrew, Marta Gandini	The personas profile is informative because: • Helps to identify the cognitive status? • What aspects regarding her/his social life they want to improve or maintain? • The description is helping to identify a pathway? • Is giving us any information that could help us to understand some change, challenge or goal that they want to face?		
NUTRITIONAL STATUS		The personas profile is informative because: • Helps to identify her/his diet patterns? • What aspects regarding her/his nutrition they want to improve or maintain? • The description is helping to identify a pathway? • Is giving us any information that could help us to understand some change, challenge or goal that they want to face?		
IN CONCLUSION	They are very well done! Those who could provide more pieces of information have been those where the connections with NESTORE could be thought of more clearly. We are unsure which features should be prevalent, in case they will results in conflicting advices to be given to the users, but we will sort this out when it will apply. We imagine NESTORE will be able to make the right suggestions based on health parameters first, e.g. not suggesting doing physical activities if you suffer from low	This persona profile is helping us to better understand: • Her/his values • Her/his preferences or priorities (What is important to them and what is driving the change?) • Her/his limitations (if exist) • Her/his context • Is giving at least one aspect (in any of the domains) that could be maintained or improved by means of NESTORE' system Do you think this persona profile is contributing to have enough information (about his/her context, preferences, resources our opportunities, etc.) to design/define a concrete pathway? Do you think this final selection of personas is a good representation of cases (in the sense of a broad spectrum of cases (in terms of variability in real situations, living conditions or health status or domain interest) that finally could help us to define the NESTORE' system based on the agreed user' profile?		

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

	blood pressure and it is very warm outside.			

- Final Personas

LEGEND

INCOME LEVEL

- €€€ good
- €€ medium
- € low

WEATHER CONDITIONS

- high humidity for outdoor activity
- high humidity for outdoor activity
- good weather for outdoor activity
- bad weather for outdoor activity

LIVING FEATURES

- living alone
- living with other people
- living in urban area
- living in rural area
- presence of stairs at home
- garden property
- living with pet
- 50msq home dimension

WEB CONNECTION

- good
- medium
- low

GENERAL ASSESSMENT

- good
- medium
- low

FRANCOIS PINOT

age	68
gender	male
country	france
income	€€
weather	
living features	60msq
web connection	

PHYSICAL STATUS

HEALTH CONDITIONS
He has stomach problems and high cholesterol

TARGET
*to do more physical activity
*to introduce walking

COGNITIVE STATUS

TARGET
*to maintain the actual cognitive health

SOCIAL STATUS

He is a part time teaching for an additional five years. He sets in different cultural projects to keep himself active

HOBBY
He likes gardening, playing table games, reading and travelling around Europe with his camper van. He regularly meets friends for a meal or just a drink. Regularly attends exhibitions and he likes to collects 20th century design pieces

TARGET
*to improve her social life

NUTRITIONAL STATUS

DIET
Qualitative and quantitative unbalanced

PROBLEMS
He has caloric assumption problems

TARGET
*to encourage a healthy and balanced diet
*to reduce the quantity of saturated fat and loss of weight

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

GIUSEPPE FORTI

age 74

gender male

country italy

income €€€

weather

living features 40msq

web connection

PHYSICAL STATUS

HEALTH CONDITIONS
Cataract, rheumatism

TARGET
*to do more physical activity

COGNITIVE STATUS

TARGET
*to improve memory through exercises
*to improve the mood

SOCIAL STATUS

HOBBY
He works part time on the farm for another 10 years. Father of 2 adult children, one lives locally and the other overseas. He has a small close circle of friends, with whom he will socialise with.
He would like to be involved in volunteering activities of his church.

TARGET
*to increase the level of social interaction

NUTRITIONAL STATUS

DIET
Balanced

PROBLEMS
He has followed a low fat healthy diet for many years due to genetic heart disease.

TARGET
*to maintain a healthy and balanced diet
*he would likes to monitoring his weight

MARTA GANDINI

age 74

gender female

country italy

income €€€

weather

living features 60msq

web connection

PHYSICAL STATUS

HEALTH CONDITIONS
Light Incontinence. Due to it she has reduced her physical activity.
She was diagnosed of a mild glaucoma one year ago

TARGET
*to do more physical activity
*introduce walking

COGNITIVE STATUS

TARGET
*to improve the level of mental focus through exercises

SOCIAL STATUS

HOBBY
She plays piano, gardening, cooking. She likes reading books and watching TV

She spends time with her grand-daughter one day a week
She would like to do voluntee-ring, but she doesn't know how to engage herself in a NGO

TARGET
*to improve her social life

NUTRITIONAL STATUS

DIET
Qualitative unbalanced (low proteins intake)

PROBLEMS
Constant weight decreasing last years
lack of regular water intake

TARGET
*to increase caloric intake
*to regulate the water assumption

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

LAURA KIMMICH

age	72
gender	male
country	germany
income	€€€
weather	
living features	100ms
web connection	

PHYSICAL STATUS

HEALTH CONDITIONS
Low blood pressure

TARGET
*to introduce soft gym practice. She walks 7500 steps/day

COGNITIVE STATUS

TARGET
*to maintan the actual cognitive level through exercises

SOCIAL STATUS

HOBBY
She likes gardening, dressmaking, cooking with large group of friends with similar interests locally and nationally.
She Regularly visits exhibitions around the country. She travel-ling on a regular basis to visit children and friends around the country. She sells her own work. She speaks German, English and she would like to learn French

TARGET
*to maintain the actual level of social interaction

NUTRITIONAL STATUS

DIET
balanced

PROBLEMS
Food asssumption related to blood pressure

TARGET
*she needs a specific diet to control the blood pressure problems

CYRIL BUYTEN

age	69
gender	female
country	belgium
income	€€€
weather	
living features	90msq
web connection	

PHYSICAL STATUS

HEALTH CONDITIONS
Allergies (dust and sweet grass). It gets worse in spring

TARGET
*to maintain the actual level of physical activity. She use to walk 9000 steps per day and she cycle twice a week.
* She does streching exercises

COGNITIVE STATUS

TARGET
*to maintain the actual cognitive health
*ito mprove the mood through social activi-ties

SOCIAL STATUS

She has close friends but she feels loneliness in some moments. She would like to do some volunteer activity. She is working in her flat making some reforms.

HOBBY
She likes gardening, cinema, theater,reading. She uses Facebo-ok and she has started to use Twitter two weeks ago. She is well-informed about politics.

TARGET
*to increase the level of social interaction

NUTRITIONAL STATUS

DIET
Balanced. She is vegetarian

PROBLEMS
Food allergies (dried fruits)

TARGET
* to maintain the her vegetarian diet

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

LIA ANDREW

age	67
gender	female
country	scotland
income	€
weather	
living features	70msq
web connection	

PHYSICAL STATUS

HEALTH CONDITIONS
She has some problems due to medium-high blood pressure as bad sleeping and recurrent headaches

TARGET
*to do more physical activity as yoga, pilates.
She walks 8000 steps/day

COGNITIVE STATUS

TARGET
*to improve the mood through social activities

SOCIAL STATUS

The divorce has impacted on friendship groups and is adjusting to the changes. Sociable through work and enjoys the interaction

HOBBY
She likes painting. She doesn't like gyms and she has had a sendentary job. She likes to cook and experiments with food through new recipes. She would like to participate in environmental activism

TARGET
*to increase the level of social interaction

NUTRITIONAL STATUS

DIET
Qualitative unbalanced

PROBLEMS
She needs to control the assumption of some type of food as cookies and cakes

TARGET
*to control the level of salt in food
*to increase the assumption of vegetables

ASTRID MOCKBERG

age	65
gender	female
country	netherlands
income	€€€
weather	
living features	80msq
web connection	

PHYSICAL STATUS

HEALTH CONDITIONS
She has presbyopia and myo-pia.She is thinking to have a surgical procedure to remove a painful bunion.

TARGET
*to improve physical activity

COGNITIVE STATUS

TARGET
* to maintain the actual cognitive health
* she is worried about her cognitive status in future

SOCIAL STATUS

Active in her neighbourhood volunteering. She has social network. Her husband has an early stage Alzheimer.

HOBBY
She sings in a choir. She likes reading newspapers. She is in contact with her son and daughter and friends. She uses Instagram to share grandchil-dren's photos. She would like to be engaged in cultural activities

TARGET
*to maintain her social life

NUTRITIONAL STATUS

DIET
Qualitaive unbalanced (low protains intake)

PROBLEMS
Intolerance to lactose

TARGET
*to maintain a healthy and balanced diet
*she wants to differentiate her menu, because she is little bored about her meals ruotine

MARCUS RASHFORD

age 66

gender male

country uk

income €€€

weather

living features 90msq

web connection

PHYSICAL STATUS

HEALTH CONDITIONS
Backpain and knee pain

TARGET
*to do more physical activity

COGNITIVE STATUS

TARGET
*to maintain the actual cognitive health

SOCIAL STATUS

He is an active volunteer, working part-time across different community centres. He has children who live away from paternal home but in contact with computer

HOBBY
He likes travelling to unusual places around the world on holiday. He enjoys dancing and eating out and drinking craft beer

TARGET
*to mantain the level of social interaction

NUTRITIONAL STATUS

DIET
Balanced

PROBLEMS
-

TARGET
*to maintain a healthy and balanced diet

ANA GARCIA

age 73

gender female

country spain

income €€€

weather

living features 60msq

web connection

PHYSICAL STATUS

HEALTH CONDITIONS
Osteopenia. Recurrent cystitis. Due to her overweight she feels fatigated after walking. She losts her motivation in walking. She is tired to be uphill and downhill all the day.

TARGET
*more physical activity
*re-introduce walking

COGNITIVE STATUS

TARGET
*to improve the level of mental focus through exercises and the mood with social activities

SOCIAL STATUS

She is grandmother and she is engaged in taking care of her grandchildren. She is engaged in a Third Age University. She receives support from her relatives. She has good relationship with her neighbors

HOBBY
She sings in a chorus. She Loves music and read. She is a keen user of Whatsapp to be in contact with her daughter and friends.

TARGET
*to maintain the level of social interaction

NUTRITIONAL STATUS

DIET
Quantitative unbalanced

PROBLEMS
She has high food intakes with low variations. She will need to diversify her menu.

TARGET
*to encourage an healthy and balanced diet
*to increase protein and water intake

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643

ANTONY ROVIRA

age	67
gender	male
country	belgium
income	€€€
weather	
living features	50msq
web connection	

PHYSICAL STATUS

HEALTH CONDITIONS
He has stomach problems, presbyopia and recurrent constipation

TARGET
*to maintain physical activity

COGNITIVE STATUS

TARGET
*maintain the actual cognitive health

SOCIAL STATUS

Engaged in Volunteer Association for professional business orientation, Golf Club, OMNIUM (cultural association). Highly concerned & worried for the political situation in Europe

HOBBY
He plays golf (18 holes) three times a week. He likes Skiing

TARGET
*to mantain the level of social interaction

NUTRITIONAL STATUS

DIET
Qualitative unbalanced. He eats too much cured cheeses and he is trying to reduce coffee intake

PROBLEMS
He has a peptic ulcer and he wants to change his diet

TARGET
*to encourage a healthy and balanced diet

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769643