Improving Sleeping Habits: Preliminary Experiments in Barcelona and Lleida

Eloisa Vargiu[1 **[0000-0002-3840-628X] *], Carme Zambrana[1 **[0000-0003-2625-1900]], Adriano Targa[2], and Ferran Barbe[2,3 [0000-0002-2340-8928]]

[1] Eurecat, Centre Tecnlogic, eHealth Unit, Barcelona, Spain
{name.surname}@eurecat.org
[2] Institut de Recerca Biomedica (IRBlleida), Lleida
febarbe.lleida.ics@gencat.cat, adrianotargads@gmail.com
[3] CIBERES, Madrid, Spain

Abstract. Due to intrinsic (e.g., daily-life habits) and extrinsic (e.g., environmental change) factors, people are far to have a healthy life and, thus, there is an increase of chronic diseases, mental disorders, and premature death. In this paper, we propose a solution to improve lifestyle habits in terms of sleeping activity through an intelligent system that monitors the sleeping and further habits together with environmental data and provides personalized recommendations and nudges. Our solution aims to study how to improve sleeping habits of citizens, investigating if and how the environment impacts the sleeping activity, as well as how it may be influenced by bad lifestyle habits. To perform a feasibility study of the proposed solution, on May we started recruiting volunteers in Barcelona (from Eurecat) and in Lleida (from IRBLleida). Volunteers were asked to wear their activity tracker 24/7, to weight once a week, and to answer the selected questionnaires at the end of the week. The day of the inclusion, we also took note of the comfort of the bed of each volunteer, as well as her/his address to take into consideration the environmental factors. Preliminary experiments clearly show the need for improving sleeping habits in the population and, thus, of intelligent solutions like the one presented in this paper.

Keywords: Sleeping activity · Life-style · Activity monitoring · Environmental monitoring.

1 Introduction

To live longer, healthier and more active, people at any age must follow simple and clear suggestions that cover the 3 main pillars of health: nutrition,

---

* Corresponding author.

** These authors contributed equally to the work

physical-, and sleeping-activity. Unfortunately, due to the intrinsic (e.g., daily-life habits) and extrinsic (e.g., environmental change) factors, people are far to have a healthy life and, thus, there is an increase of chronic diseases, mental disorders, and premature death. While a lot of effort has been done to support good habits on nutrition providing personalized diet and continuous follow-up and to encourage and monitor physical activity, sleeping activity is still an untaught and unconsidered problem. In this paper, we propose a solution aimed at improving sleeping habits through an intelligent system that monitors the sleeping activity together with environmental data and lifestyle habits and provides personalized recommendations and nudges. In particular, the proposed system focuses on the following research questions:

- Q1 How we may improve sleeping habits of citizens?
- Q2 How the environment influences and impacts the sleeping activity of citizens?
- Q3 How bad lifestyle habits influence and impact the sleeping activity of citizens?

The work presented in this paper is part of the CarpeDiem project (Collaborative and Adaptive Recommender for PErsonalized DIEt Management), aimed at providing intelligent and automatic support to people who want to follow a diet to lose weight, or to maintain a healthy lifestyle (in a comprehensive way, taking into account nutrition, physical activity, and sleep). CarpeDiem target users are elderly people who need to follow healthy habits, including sleeping better.

To study the feasibility of the solution and to start creating the machine-learning models that will allow to send personalized recommendations to the final users, on May we started collecting data from volunteers in Barcelona and Lleida (Spain). Accordingly, in this paper we present the overall idea of the solution, currently under development, and the preliminary experiments performed to study its feasibility.

The rest of the paper is organized as follow. Section 2 illustrates the underlying problem and its relevance for the population. In Section 3, we describe the overall solution for monitoring sleeping activity, environmental data, and lifestyle habits. Section 4 presents the preliminary experiments we made in Barcelona and Lleida with volunteers. Finally, Section 5 ends the paper with conclusions and future directions.

2 The Problem

The Center for Disease Control and Prevention (CDC) in the United States, within the Healthy People 2020 policy, has considered that sleeping little is a public health problem. In fact, according to a recent CDC study, more than a third of adult Americans do not get enough sleep on a regular basis [7]. Insufficient sleep is not exclusively a problem of the United States, it also affects other industrialized countries such as the United Kingdom, Japan, Germany, or Canada [1]. Regarding Catalonia, the average number of hours of the overall population is 7.1 hours at night, very close to the lower limit recommended by clinicians (i.e., 7 hours at night). At Catalan level, a survey made in 2015 [3] performed using the SATED questionnaire [2] confirms that sleeping duration is age dependent. Moreover, they found that the SATED score is lower for elderly people than citizens of other age groups, meaning that their sleep is worse.

According to recent studies, the proportion of people who sleep less than the recommended hours (i.e., between 7 and 9) continuously increases and it is associated to lifestyle factors related to a modern society, such as psycho-social stress, unbalanced diet, lack of physical activity, and excessive use of electronic media, among others [8]. This is alarming since it has been found that lack of sleep is associated with negative social and health outcomes.

The association of lack of sleep with a series of negative social and health outcomes is becoming worrisome. Insufficient sleep duration has been linked to seven of the fifteen leading causes of death in the United States. The existing evidence suggests that the relationship between sleep time and health, outside of the normal window (7-9 hours), both below and above, is associated with a high risk of stroke, coronary heart disease, hypertension, obesity, type 2 diabetes and mortality [11] [12]. For instance, Gallicchio and Kalesan [5] report that the combined relative risk (RR) for all-cause mortality for sleep insufficiency is 1.10 (95% CI = [1.06, 1.15]).

The link between sleep duration and mortality represents that a citizen who sleeps on average less than 6 hours per night has a mortality risk 13% higher than an individual who sleeps between 7 and 9 hours, which is the window that is considered as the healthy amount of sleep. Furthermore, a person who sleeps between 6 and 7 hours per night has a 7% increased risk of death, including all causes of death such as fatal car accidents, strokes, cancer, or cardiovascular disease.

Finally, sleeping insufficient time reduces productivity in the workplace: workers who sleep less than 6 hours per day have an average productivity loss of more than 2.4 percentage points due to absenteeism, calculated on those workers who sleep between 7 and 9 per day. Those who sleep on average between 6 and 7 hours continue to generate a loss of productivity greater than 1.5 percentage points compared to those of seven to nine hours. To put these numbers in perspective, assuming that there are 250 business days in a given year, this means that a worker who sleeps less than 6 hours loses about 6 business days per year more than a worker who sleeps 7 to 9 hours. A person sleeping for 6 to 7 hours loses an average of 3.7 more business days per year [9].

3 The Proposed Solution

The aim of this work is improving sleeping habits in the population. Thus, we propose an automatic and intelligent system that empowers and supports the citizens giving them personalized recommendations and nudges. In order to ensure a clinical relevance, the actual set of recommendations and nudges will be defined by the clinical personnel. Personalization will be then reached by monitoring the citizens and studying their usual behavior.

Three main categories of factors that impact sleeping habits will be considered:

1. number of hours;
2. sleep efficiency, calculated as the ratio between the number of slept hours and the total number of hours in bed per night;
3. the satisfaction of the citizen about her/his sleep calculated through the standard questionnaire SATED.

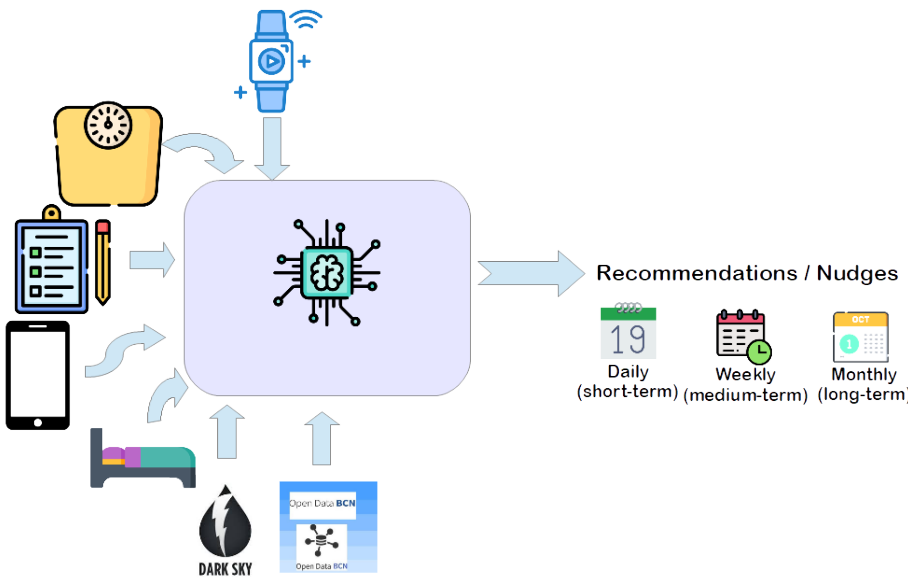

Fig. 1. The proposed solution at a glance.

The proposed solution, sketched in Figure 1, works with a set of data coming from different sources:

– The citizen wears a wristband 24/7 and the sleep activity is monitored in terms of: total number of sleeping hours, number of sleeping hours during the night, number of the sleeping hours during the day, number of naps, time when going to sleep, wake-up time, and sleep efficiency. Moreover, the data regarding the physical activity in terms of total number of steps and vigorous activity before going to sleep are constantly monitored.
– The citizen is asked to monitor her/his weight once a week in order to take under control her/his Body Mass Index (BMI) and to avoid reaching obesity.

- To build the user's profiles, the day of the inclusion, the citizen is asked to answer to some questionnaires regarding her/his life-style habits. The following questionnaires will be used: SATED, about the satisfaction of the citizen regarding her/his sleep; Caffeine, regarding the number of cups of coffee or teas drunk on average during the week; Smoke, on the number of cigarettes smoked on average during the week; and Use of technology, concerning the number of minutes spent on average using the smartphone or a tablet before going to sleep.
- The smartphone of the citizen is used to detect the luminosity of the bedroom during the sleeping time.
- The citizen will be asked regarding the comfort of the bed, since especially vulnerable or at risk citizens often sleep on truly old or uncomfortable beds.
- To be aware regarding seasonality, daylight hours, and further environmental data, we will rely to the Dark Sky API[4]. It allows to look up the weather anywhere on the globe, returning (where available): weather conditions, minute-by-minute forecasts out to one hour, hour-by-hour and day-by-day forecasts out to seven days, hour-by-hour and day-by-day observations going back decades, weather alerts, and humidity.
- To monitor the data about the air quality and noise pollution in Barcelona, the Open Data service by the Barcelona Municipality (Open Data BCN) will be used[5].

The system will collect and fuse the heterogeneous data coming from all the data sources and, through machine learning techniques, will build the models to cluster and classify citizens according to their habits.

First, the profile of the citizen is calculated adopting data mining and machine learning approaches relying on Knowledge Data Discovery (KDD) models [4] as well as User Data Discovery (UDD) models [6]. A static profile will be created considering gender and age together with data coming from the questionnaires. That profile will be dynamically and continuously updated by considering the data gathered from the wristband, the changes in the BMI, if any, and the environmental data.

Once built the profile, the citizen will be automatically clustered according to fixed groups of habits defined considering the intersections among the three categories: number of sleeping hours (less than 7 hours, between 7 and 9 hours, more than 9 hours), sleep efficiency (less than 50%, between 50% and 95%, more than 95%), and satisfaction about the sleep (by using the 5 levels of the SATED questionnaire: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Usually, 5 = Always). For each cluster, the clinicians will define a suitable set of recommendations and nudges with the final goal of moving all the citizens in the healthiest cluster, corresponding to a number of hours between 7 and 9, a sleep efficiency higher than 95%, and a satisfaction of 5 (i.e., always satisfied). In the case a citizen does not belong to that cluster because s/he is not compliant with one or more

[4] https://darksky.net/dev

[5] https://opendata-ajuntament.barcelona.cat/en/

of the categories, the system asks to choose the category s/he wants to improve (monthly goal).

Recommendations and nudges will be sent to the citizen on daily (short-term), weekly (medium-term), and monthly (long-term) basis. Recommendations and nudges on daily basis will be based on the number of slept hours, the efficiency, the performed physical activity, and will take into account anomalies in environmental factors specific of the considered day (e.g., high temperatures). Taking particular reference to the category the citizen wants to improve, on weekly basis the system will send specific questions or request of filling a questionnaire in case of anomalies (e.g., in case of a decrease of the number of hours, the Caffeine questionnaire may be sent to be filled). Finally, monthly based ones will take into account the changes in habits during the month, the change from a cluster to another, if any, and will consider similarities among citizens in order to send recommendations and nudges according to a collaborative filtering approach [10]. At monthly basis the citizen will be also asked to confirm or change the category s/he want to improve, depending on the results of the past month.

From a technological point of view, the citizen will interact with the system through a simple and user-friendly app installed in her/his smartphone and that will automatically send recommendations and nudges as push notifications. Through the app, the citizen will be also able to set-up her/his profile, answer questionnaires, pair her/his smartphone and smart-scale (if any), or to manually input the weight (e.g., after weighting at the pharmacy). The app will also allow to capture the luminosity of the bedroom during the sleeping hours.

4 Preliminary Experiments

The success of the proposed solution strictly depends on the gathered data. In fact, profiling, recommendations and nudges definition, clustering, and personalization cannot be achieved without a preliminary phase in which data are collected and used to build all the models. Thus, we started on May recruiting volunteers in Barcelona (from Eurecat) and in Lleida (from IRBLleida) to profile the users and to study the feasibility of the proposed solution.

4.1 The Collected Dataset

A total of 30 volunteers has been recruited in Barcelona and Lleida. Volunteers were asked to wear their activity tracker 24/7, to weight once a week, and to answer the selected questionnaires at the end of the week. The day of the inclusion, we also took note of the comfort of the bed of each volunteer, as well as her/his address to take into consideration the environmental factors.

Due to technical problems and adherence issues, 6 of the 30 volunteers dropped-out before finishing the first month of the study. Thus, the analysis of the data have been done using the data of 24 volunteers (38.43 ± 11.46 years old; 15 females; and 23.09 ± 3.66 BMI). The preliminary study started on May and will finish in December.

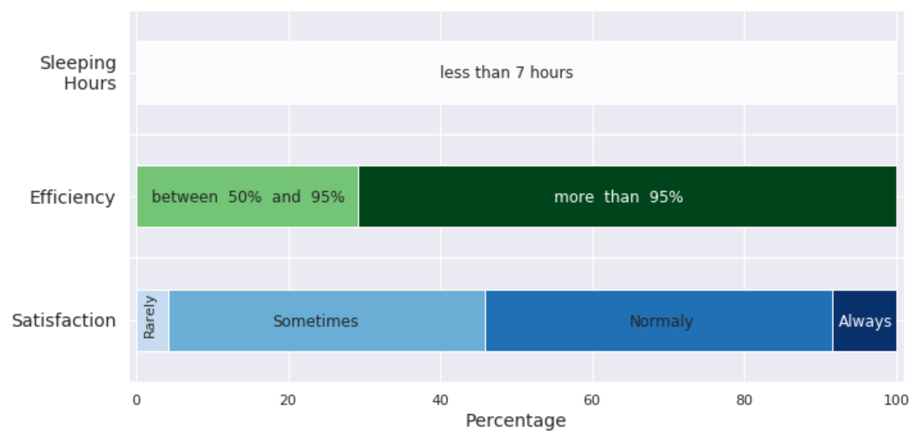

Fig. 2. Distribution of the volunteers in each group for the three categories (*Sleeping Hours*, *Efficiency* and *Satisfaction*) at inclusion time.

At inclusion time, we first analysed the distribution of our population among the three categories:

- *Sleeping Hours*
  1. less than 7 hours;
  2. between 7 and 9 hours (the healthiest group);
  3. more than 9 hours.
- *Efficiency*
  1. less than 50%;
  2. between 50% and 95%;
  3. more than 95% (the healthiest group).
- *Satisfaction*
  1. Never;
  2. Rarely;
  3. Sometimes;
  4. Normally;
  5. Always (the healthiest group).

Figure 2 shows the percentage of volunteers in each group for the three categories. As shown, all the volunteers slept less than 7 hours, which, as stated above, is risky for their health. Regarding the efficiency, the 70% of the volunteers are in the healthiest group, whereas the 30% are in the group between 50% and 95%. As for the satisfaction, the majority of the volunteers answered sometimes or normally with about 40% of the answers in each. The rest of the volunteers are in the *Always* group (about 8%) and in the *Rarely* one (only a 4%). It is worth noting that nobody belongs to the *Never* group. This first analysis performed at inclusion time shows the need of the proposed solution to improve sleeping

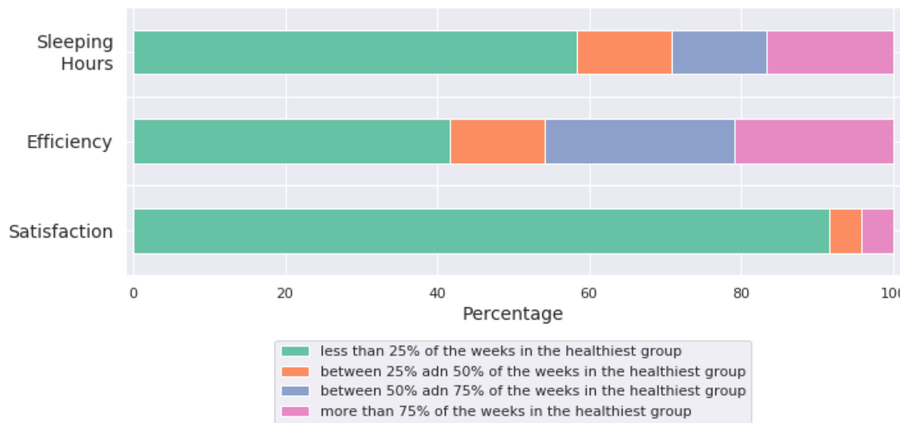

Fig. 3. Distribution of the volunteers in the following four sets: less than 25% of the weeks in the recommended group, between 25% and 50% of the weeks in the recommended group, between 50% and 75% of the weeks in the recommended group, and more than 75% of the weeks in the recommended group.

habits with the final goal to move the volunteers to the healthiest group of each category, specially in case of *Sleeping Hours* and *Satisfaction*.

As a second step, for each category, we computed the percentage of weeks that each volunteer belonged to the healthiest group. To better visualise the distribution, for each category we considered four sets (see Figure 3): *less than 25% of the weeks*, *between 25% and 50% of the weeks*, *between 50% and 75% of the weeks*, and *more than 75% of the weeks*. As shown, the majority of the volunteers are in the set *less than 25% of the weeks* for all the categories: about 58% for *Sleeping Hours*, 41% for *Efficiency*, and 91% for *Satisfactions*. Results clearly show that there is a urgent need for a solution to improve sleeping activity of citizens considering all the three categories. In fact, at least in our population, the percentage of people that sleeps the right number of hours, with a high efficiency, and is satisfied of their sleep is very low (about 16%, 20%, and 4% for *Sleeping Hours*, *Efficiency*, and *Satisfactions*, respectively).

Finally, we analysed the dynamics of the changes among groups in a monthly basis. To visualize it a Sankey Diagram is used: the left column display the groups from which the volunteer moved and the right column the groups to which the volunteer moved. The links between them are represented with arcs that have a width proportional to the number of times this transition occurred.

Figure 4 shows the dynamics of the *Sleeping Hours* category. As shown, all the volunteers are in the groups *less than 7h* and *between 7h and 9h*, being the majority of them in the first group (about 72%). The volunteers tend to stay in the same group. Only a small percentage changed from *less than 7h* to *between 7h and 9h* (about 11%), and vice versa (about 27%). In this case, the proposed

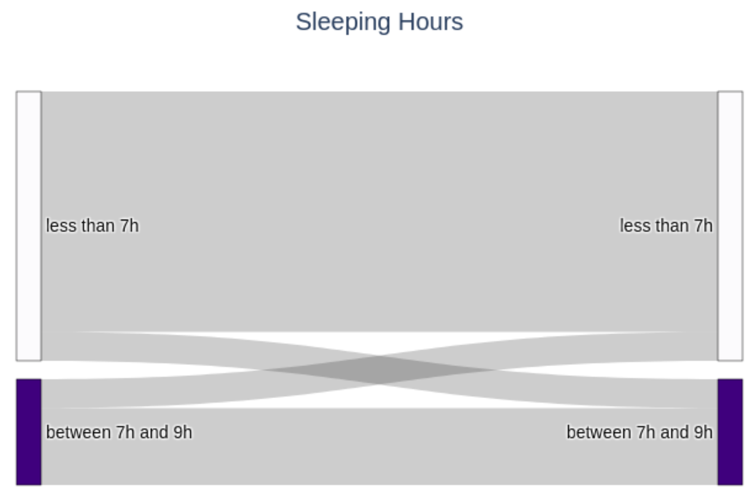

Fig. 4. Dynamics of the *Sleeping Hours* category.

solution will give support to move all the volunteers from the *less than 7h* to the *between 7h and 9h* group. It is worth noting that, once a volunteer reaches the healthiest group, the system will give support to do not change the group.

Figure 5 shows the dynamics of the *Efficiency* category. As shown, all the volunteers are in the groups *between 50% and 95%* and *more than 95%*, and they are balanced (about 51% and 49%, respectively). In this case, the tendency is either stay in the same group (about 60%) or change to the other one (about 40%). Seemly to the *Sleeping hours* category, our system will work to change this dynamic and move all the volunteers from *between 50% and 95%* group to the *more than 95%* one. Once a volunteer reaches the healthiest group, s/he will be supported by the system to stay in it.

Figure 6 shows the dynamics of the *Satisfaction* category. As shown, the majority of the volunteers are in the groups *Sometimes* and *Normally* (about the same number in each, i.e. 43% and 49%, respectively), and a small percentage in the group *Rarely* (about 8%). The volunteers from the *Rarely* group used to stay in the same group (about 67%) or to move to the *Normally* one (about 33%). On the contrary, volunteers belonging to *Sometimes* and Normally groups move to their consecutive groups: *Rarely*, *Sometimes*, and *Normally* (about 12%, 47% and 41%, respectively) and *Sometimes*, *Normally* and *Always* (about 26%,

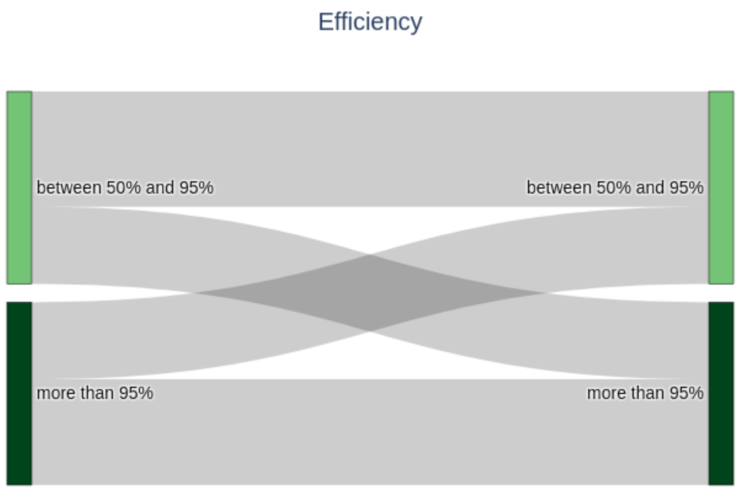

Fig. 5. Dynamics of the *Efficiency* category.

68% and 5%, respectively), respectively. Let us note that, only volunteers from the *Normally* group moves to the *Always* one (about 5%)[6].

4.2 Daily Monitoring

Recommendations and nudges on daily basis will be based on the number of slept hours, the efficiency, the performed physical activity, and will take into account anomalies in environmental factors specific of the considered day (e.g., high temperatures).

The aim of our system is to encourage the user to *Keep going* if s/he is improving with respect to "yesterday" or suggests to *Keep following the recommendation to improve* in case of worsening. Recommendations could be provided with respect to the sleeping habits (e.g., *You should go to sleep every night at the same hour and before 12 am*) and/or the performed physical activity (e.g., *Do not perform high-intensity exercises 2 hours before going to sleep*). Regarding the environmental factors, the system informs: when the city is suffering from high temperature and suggests to the user to keep the bedroom temperature to the recommended one (between 16-18°C); when the air quality is low and suggests the user to not perform vigorous activity outdoors; and when the noise level of her/his street is too high and suggests to keep closed the windows during night.

[6] The group *Always* appears only in the right column because the month which a volunteer was in the group *Always* was the last one of the study

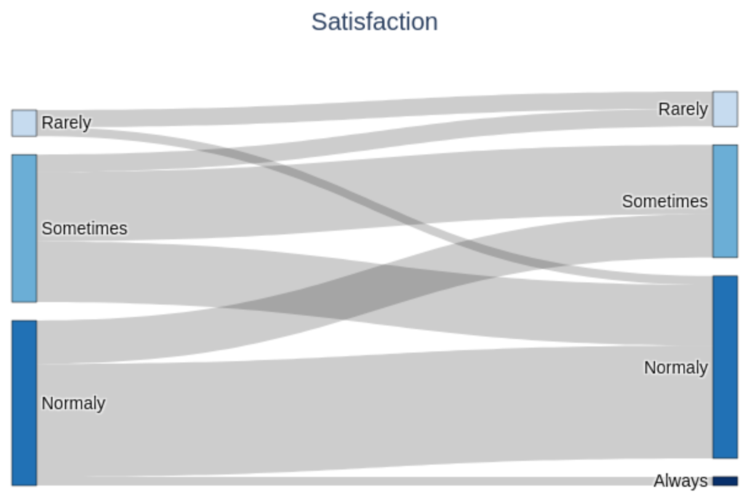

Fig. 6. Dynamics of the *Satisfaction* category.

Figure 7 shows an example of collected data of a volunteer (52 years old; male and who lives in Barcelona) and potential messages that the proposed system should send to the final user. Let us note that during the selected days (the week from June 24th to June, 30th) a high temperature was measured in Barcelona.

4.3 Weekly Check-Up

In order to support the user in changing her/his sleeping habits, the proposed system will compute a summary, weekly. Based on it, nudges and recommendations will be provided. In case a change greater than ±2 hours, 5% in efficiency, and/or ±3 points in satisfaction is detected, the system will send a requests to fill the questionnaires in order to recalculate the profile and find potential anomalies that could have affected the sleeping activity. Figure 8 shows an example of 3 consecutive weeks of a volunteer (26 years old; female): her evolution in terms of the three categories and the suggested messages to give her support.

4.4 Monthly Clustering

At the beginning of a new month, the system analyses all the collected data to cluster the users considering the three selected categories. In so doing, the system verifies if the user changed the cluster s/he belongs to, or not. In case of a positive change, the system sends an awards to the user and suggests the new goal to be

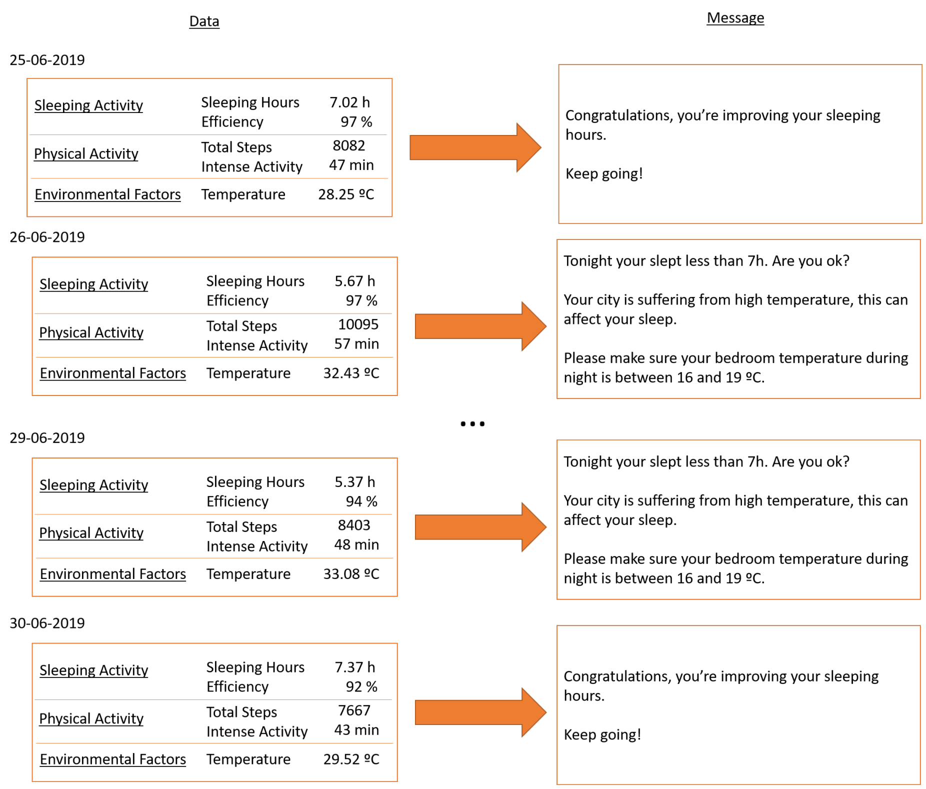

Fig. 7. Example of daily monitoring during a week in which in Barcelona a high temperature was measured.

followed during then next month. In case of a negative change or not changing, the system recommends to still work on the same goal during the next month. In both cases, the users can accept or reject the received suggestion and choose the goal to be achieved. Figure 9 shows an example of 2 consecutive months for a volunteer (26 years old; female) in which efficiency has been improved while the satisfaction still needs to be enhanced.

5 Conclusions and Future Work

In this paper, we presented a novel solution aimed at improving sleeping activities. Although the research and development is at its early stage, the feasibility study performed right now shows how the research questions listed in the Introduction will be addressed.

Q1. How we may improve sleeping habits of citizens? The proposed solution is a recommender system that will be built according to clinical evidence and with the supervision of the clinical team of the Sleep Unit at the IRB

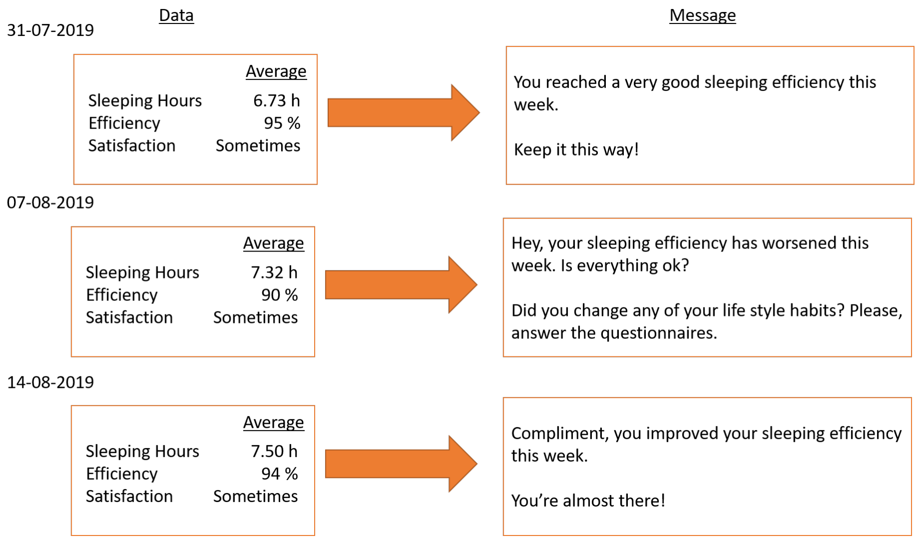

Fig. 8. Example of weekly check-up and messages during 3 consecutive weeks.

in Lleida. The overall set of recommendations and nudges will be defined by clinicians of that Unit and will be fully personalized according to the typology of the final user. The system will automatically send the recommendations and nudges to the citizens on daily base (short-term recommendations), on weekly base (medium-term recommendations), and on monthly base (long-term recommendations).

Q2: How the environment influences and impacts the sleeping activity of citizens? It has been shown in the literature that the main environmental factors that influence and impact the sleeping activity are: daylight hours, seasonality, luminosity, humidity, pollution, noise, and comfort of the bed. The proposed solution will consider all these factors, their influence on sleeping, and will define recommendations taking into account the peculiarities of each citizen according to the environment in which s/he lives.

Q3: How bad lifestyle habits influence and impact the sleeping activity of citizens? Having bad lifestyle habits has an impact on the overall sleeping activity. In particular, clinicians recommend to not perform the following tasks just before going to bed: take a coffee, smoke, eat, perform vigorous physical activity, and use the smartphone or a tablet. In addition, citizens have to control their weight, in order to avoid obesity which could cause sleeping problems. The final system will send recommendations and nudges on this kind of tasks to improve citizens behavior and will sporadically monitor coffee intake, smoking, eating, and smartphone usage, in case of detected anomalies with respect to the usual sleeping behavior. Moreover, physical activity will be continuously monitored and suitable suggestions given accordingly.

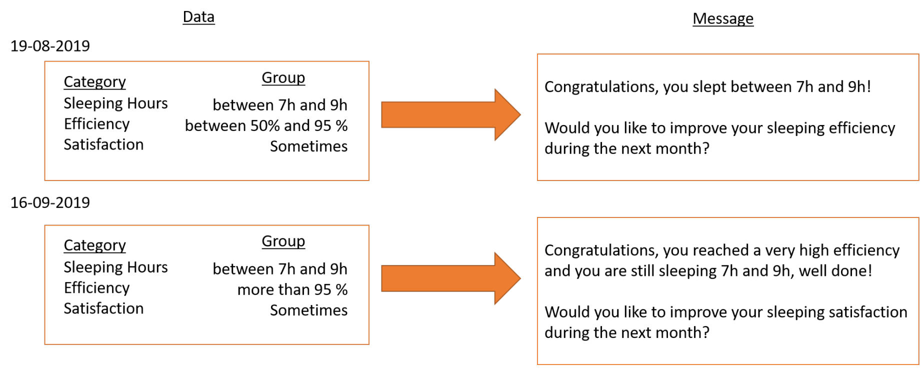

Fig. 9. Example of monthly clustering during 2 consecutive months.

As for the future work, as soon as 6 months of data will be collected from all the volunteers, we will build the models for sending daily, weekly, and monthly recommendations and nudges. Clinicians in IRBLleida are currently working on the definition of them. At the beginning of next year, the system will be integrated in the CarpeDiem app and the pilot will start with new volunteers from both Barcelona and Lleida, as well as further towns in the Catalonia Region.

6 Acknowledgments

This work was financially supported by the Catalan Government through the funding grant ACCI-Eurecat (CarpeDiem project).

References

1. Buman, M.P., Phillips, B.A., Youngstedt, S.D., Kline, C.E., Hirshkowitz, M.: Does nighttime exercise really disturb sleep? results from the 2013 national sleep foundation sleep in america poll. Sleep Medicine **15**(7), 755–761 (2014)
2. Buysse, D.J.: Sleep health: can we define it? does it matter? Sleep **37**(1), 9–17 (2014)
3. Dalmases, M., Benítez, I.D., Mas, A., Garcia-Codina, O., Medina-Bustos, A., Escarrabill, J., Saltó, E., Buysse, D.J., Roure, N., Sanchez-de-la Torre, M., et al.: Assessing sleep health in a european population: Results of the catalan health survey 2015. PloS one **13**(4), e0194495 (2018)
4. Fawcett, T., Provost, F.J.: Combining data mining and machine learning for effective user profiling. In: KDD. pp. 8–13 (1996)
5. Gallicchio, L., Kalesan, B.: Sleep duration and mortality: a systematic review and meta-analysis. Journal of sleep research **18**(2), 148–158 (2009)
6. Kanoje, S., Girase, S., Mukhopadhyay, D.: User profiling trends, techniques and applications. arXiv preprint arXiv:1503.07474 (2015)
7. Liu, S., Lamaze, A., Liu, Q., Tabuchi, M., Yang, Y., Fowler, M., Bharadwaj, R., Zhang, J., Bedont, J., Blackshaw, S., et al.: Wide awake mediates the circadian timing of sleep onset. Neuron **82**(1), 151–166 (2014)

8. Roenneberg, T.: Chronobiology: the human sleep project. Nature **498**(7455), 427 (2013)
9. Rosekind, M.R., Gregory, K.B., Mallis, M.M., Brandt, S.L., Seal, B., Lerner, D.: The cost of poor sleep: workplace productivity loss and associated costs. Journal of Occupational and Environmental Medicine **52**(1), 91–98 (2010)
10. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recommender systems. In: The adaptive web, pp. 291–324. Springer (2007)
11. Sperry, S.D., Scully, I.D., Gramzow, R.H., Jorgensen, R.S.: Sleep duration and waist circumference in adults: a meta-analysis. Sleep **38**(8), 1269–1276 (2015)
12. Wang, Q., Xi, B., Liu, M., Zhang, Y., Fu, M.: Short sleep duration is associated with hypertension risk among adults: a systematic review and meta-analysis. Hypertension Research **35**(10), 1012 (2012)