Abstract—Nowadays, the proliferation of wearable devices has enabled monitoring user behaviours and activities in a non-invasive, autonomous and straightforward way. Moreover, new trend analysis has been boosted by biosignal sensors from wearable trackers, such as inertial or heart rate sensors. The knowledge of such user activity presents a personalized monitoring to prevent any kind of physical or neurological disorders through the sensor evaluation by an expert. To this end, the definition of key indicators and the display of results and relevant analyses require agile and effective tools. Therefore, this proposal presents a novel web application where the data obtained from a Fitbit Ionic smartwatch wearable are collected, synchronized and compiled to present a summary of an user’s daily activity, which is defined by a linguistic description using fuzzy logic to represent the most relevant Health Key Indicators (HKI). Moreover, an analysis of the user’s behaviour over time is proposed by inferring relevant patterns from a fuzzy clustering algorithm.

I. INTRODUCTION

The control of different health parameters, thanks to the proliferation and accessibility of wearable devices, has allowed the vast majority of the population to improve their quality of life [1]. In this sense, these types of devices help to monitor not only the elderly [2] but also younger population groups [3] and allow the observance of health problems in a preliminary stage. Moreover, the wearable trackers have promoted the increasing of monitored physical activity improving the average health condition of population [4].

Therefore, the enormous amount of data generated by these devices must be processed to know the user’s status in time and form. To do that, there is a good number of proposals on the market as Gatt et al. summarized in [5]. One way to process this is to obtain a linguistic summary of the information implicit in this data, through an expertly guided analysis, where the significant information is described (see compilation in [6]).

Here, it is proposed to generate linguistic descriptions of the relevant information obtained from a Fitbit Ionic smartwatch user’s activity records. These linguistic summaries are closer to natural language by using Fuzzy Logic (FL) techniques and the Computing with Words (CW) paradigm [7].

By means of the collection and pre-processing of biosignal sensors, wearable trackers provide several measures from users’ daily activity, such as step counter, burned calories, heart rate, sleep states... In this work, we have focused on Heart Rate (HR) and sleep states (AS), since both parameters are commonly taken into account by health professionals in the prevention of heart problems or sleep chronic diseases, respectively. For instance, the National Sleep Foundation is studying these technologies to learn how to perform early interventions when there exist symptoms of any sleep disturbance [8].

According to this, a set of pre-defined Health Key Indicators (HKI), which have been supervised by experts, have been designed to generate linguistic summaries of the most relevant user states about heart rate and sleep in a specific period of time. To do that and as the first goal of this paper, the development of a web application that helps to infer and discover these HKIs has been proposed here. This system is designed to be synchronized automatically with the user’s wearable and to provide an interface to retrieve, under request, a daily summary of the main activities of the user.

Moreover, this set of HKIs can be used to search for patterns of user behaviour at a later stage. The pattern analysis process requires the classification of the user activity [9] which can lead us to the detection of potential health problems. The generation of these sets of HKIs to execute this functionality has been defined as the second goal of this application.

A brief study of previous research on this topic is included in section II. After that, the web application that synchronizes and analyzes the data provided by the wearable tracker is described in section III. The knowledge and linguistic model used to generate the HKI of the user and consequently the daily linguistic summaries are defined in section IV. To illustrate this, a case of study with real data is included in V. Finally, the paper conclusions are shown in section VI.

II. BACKGROUND

Data-to-text systems allow to extract useful information from data on the basis of the expert knowledge using natural language. An specific application of these systems is the Generation of Linguistic Description of Time Series (GLiDTS), which simulates expert knowledge to obtain a text summary that satisfies a final user requirements [6] from time-dependent information. These systems have been applied in the context of health care in the last decade thanks to the increase of control devices in the market. Some of the first proposals for
this purpose were performed by Gatt et al. [10], which was focused in the generation of summaries on Neonatal Intensive Care Units (NICU), or the Mikut et al.’s [11] which describes medical monitoring devices in a simple and interpretable manner. Recent reviews of these systems are provided by Gatt et al. in [5] and Marin et al. in [6].

An specific control device for health care, the Fitbit Ionic smartwatch, has been used to monitor some user activities here because of its advantages: multi-day battery life, heart rate accuracy, description of sleep stages, users datasets in cloud from data pre-processing, compatibility and GPS antenna [12]. This device monitors the following user features [13]:

**Activity:** Steps, distance, active minutes, calories, floors, heart rate, stationary time, cardio fitness score, and more.

**Sleep:** Sleep schedule, minutes asleep and awake, number of times you woke up, amount of time spent in bed, time spent in each sleep stage.

**Exercise:** Exercise history, exercise logs.

**Social:** Fitbit community groups you are a member of, challenges, badges, community group posts, and comments.

**Others:** Social Information, Coach workouts information, Corporate data (as challenges), Logs (Food, water, exercise, weight logs), Profile, Direct messages, female health, sleep score, friends, subscriptions.

The development platform of Fitbit provides an HTTPS Web API for accessing data to Fitbit wearable data from automatic activity logs. The API requires the use of OAuth 2.0 authorization framework that enables a third-party application to obtain limited access to an HTTP service, either on behalf of a resource owner by orchestrating an approval interaction between the resource owner and the HTTP service, or by allowing the third-party application (our application) to obtain access. OAuth 2.0 is an authentication standard defined in RFC 6749[14]. Also, an alternative API for the processing of fuzzy data and computing with words paradigm can be found in [15], [16].

### III. THE MONWATCH APPLICATION

Monwatch is a web application\(^1\) that allows users to monitor and obtain summaries of his/her activity by computing the collected data provided by the wearable Fitbit Ionic smartwatch. The basic functionality of this proposal consists of:

- Pre-processing of data collected from an user,
- Storage of the records in an internal database,
- Evaluation of the protoforms (HKI) and their truth degree,
- Display of a linguistic summary of the daily activity of the user,
- Exportation of the HKIs to generate a set of fuzzy clusters that represents the user behaviour.

To provide more detail about Monwatch’s utilities in the following subsections, its architecture, database structure and user interface used are explained.

\(^1\)Monwatch is available in [http://monwatch.frre.utm.edu.ar](http://monwatch.frre.utm.edu.ar) (user: test: testing1234).

### A. System architecture

The system architecture (see Figure 1) of Monwatch is described by two main operations that summarize the connectivity: i) the smartwatch collects the raw data from the user and sends them to the smartphone via Bluetooth using the Fitbit APP. ii) data are sent to the Fitbit cloud from the mobile device where, through an API Service, Monwatch extracts and analyzes them.

1) **Connectivity of the system:**

Monwatch allows to synchronize and view information from data collected by wearable trackers. Specifically, in this work, we focus on heart rate and sleep stages measures. Data are collected from the Fitbit Cloud through Fitbit API with Monwatch. It has been developed using Django framework version 2 as backend and Python 3.6. All data are synchronized via HTTPS (OAuth2) requests according to Fitbit API requirements and the returned data are stored locally in a MySQL Database.

Monwatch requires data permissions when redirecting the user to the authorization page by OAuth2 protocol. To obtain access, the application follows the steps listed below and in figure 2:

1) Monwatch selects the scope to be collected to ask for permissions.
2) Monwatch redirects the user to Fitbit’s authorization page to obtain the code grant as specified in RFC 6749.
3) Upon user consent, Fitbit redirects the user back to Monwatch’s redirected URL with an authorization code as a URL parameter.
4) The application exchanges the authorization code for an access token and refresh token and it stores the access token and refresh token. It uses the access token to make requests to the Fitbit API and the refresh token to obtain a new access token when the previous access token expires without having to re-prompt the user.

![Figure 1. Monwatch connection.](http://monwatch.frre.utm.edu.ar)
After the authorization process, the JSON Structures of sleep stages and heart rate data, returned by the Fitbit SDK, are shown in figures 3 and 4, respectively.

The synchronization of the data to Monwatch is developed using two scripts developed in Python, which are described in figure 5. The first script synchronizes the data, Heart Rate and Sleep stages, from the Fitbit website (via REST API) for a specific day. It runs daily in the early hours of the morning and it synchronizes the data of the day before the date the script is running, i.e., if the script is running on Dec, 3rd 2019 at 4:00 AM, it synchronizes the data of Dec, 2nd 2019. We note that current data cannot be synchronized until all day values are truly taken. The second script sends notifications in case the smartwatch that is collecting the data has not been synchronized in a set period of time. That is, there is no data. The script requests data from the last time the Fitbit device has performed a complete data synchronization. In case the number of days of last synchronization exceeds a certain threshold, the system notifies users via email of such inactivity, so that measures can be taken to prevent the loss of data on the device. Both scripts run automatically on the web server that hosts Monwatch. These routines are programmed to run in parameterized hours, once a day, thanks to the help of Cron (since Monwatch is hosted on a Linux server). Cron is a system process that is used to execute background tasks on a routine basis. Cron requires a file called crontab that contains the list of tasks to be executed at a particular time. All these jobs are
executed in the background in a specified time.

Figure 5. Synchronization process.

2) Database structure:
Monwatch stores the information from the JSON files in an MySQL database whose structure is illustrated in figure 6. Tables that represent data related with heart rate, store the records in BPM (beats per minute). Sleep levels, which are semantically described in section IV, are represented in the database by the following values:

\[
\text{T.I. LEVELS CHOICE} = (\\n\quad (\text{W}, \text{Wake}),\\n\quad (\text{L}, \text{Light}),\\n\quad (\text{D}, \text{Deep}),\\n\quad (\text{R}, \text{REM}),\\n\quad (\text{T}, \text{Restless}),\\n\quad (\text{A}, \text{Asleep}))
\]

3) Interface:
Materialized UI presentation layer (CSS responsive framework) is used together with JavaScript and HTML to obtain an attractive, portable and functional web application. An screenshot of the final result of this application is shown in figure 7. This web application consists of:

- an authentication page,
- a side menu with access to the main functionalities,
- a synchronization page,
- a page with a summary of daily activities.

IV. GENERATION OF LINGUISTIC DESCRIPTIONS OF DATA

In a first step for analyzing wearable data from user activity, a linguistic approach to analyze data from HKIs is applied. This proposal uses Fuzzy Logic techniques and the Computing with Words paradigm to represent such knowledge as described in section II. Thus, in the context of monitoring an user behaviour using wearable trackers, the data analyzed are represented by two linguistic variables \( (V_r; r = 1, \cdots, n) \):

- **Activity status** (AS) represents the information related with the sleep/rest of the user. It is defined by means of 6 discrete values or terms:

  - *Awake (WK)*. The user is awake.
  - *Restless (RS)*. The user is not resting (unease, discomfort).
  - *Light sleep (LS)*. It is a non-restful sleep.
  - *Deep sleep (DS)*. It is related with the relax of the body (sleep when breathing slows and muscles relax).
  - *REM phase (RP)*. It is a deep sleep with the brain.
  - *Asleep (A)*. It is when the user is not sleeping but resting.

- **Heart rate** (HR) represents the beats per minute (bpm) in the human range \([40,220]\). This variable has been fuzzyfied into four summarizers, i.e., *high*, *high-intensity*, *normal* and *low* \([17]\) and four fuzzy sets which have been defined by the trapezoidal functions shown in figure 8.

These variables \( (V_r; r = 1, \cdots, n) \) are defined by the measures collected as a data stream by the wearable. Formally, each measure, i.e., HR or AS, is defined by a 2-tuple value \( s_i = \{s_i, t_i\}; i = 1, \cdots, m \) where \( s_i \) defines a given value of the \( V_r \) in time-stamp \( t_i \). Thus, the long-term information of an specific variable from a given user is composed of a data stream \( S^r = \{s_{0}^{r}, \cdots, s_{r}^{r}, \cdots, s_{m}^{r}\} \).

A. Knowledge representation

The formalism chosen to represent the knowledge from the raw data is related to the protoforms defined by Zadeh in \([7]\). In this application, the following protoform has been proposed
to define the main HKIs that summarizes the activity of the user,

$$Q_k A_r T_j$$

where:

- $A_k$ is a fuzzy linguistic term that has been defined in the context of the linguistic variable $V_r$.
- $T_j$ is a fuzzy temporal window (FTW) which is defined by a fuzzy temporal linguistic term and a fuzzy set (illustrated in figure 9). A FTW aggregates the values of fuzzy linguistic term $A_k$ in a temporal period $\Delta t_r = t_0 - t^{r}_i; t_0 < t^{r}_i$. The aggregation function that allows to compute the membership degree of each $A_k$ to each FTW is defined by [18]:

$$A_k \cup T_j(S^r) = \bigcup_{s^r_i \in S^r} A_k \cap T_j(s^r) \in [0,1]$$

where the $A_k \cap T_j(S^r)$ are computed by the fuzzy weighted average equation [19].
- $Q_k$ is a fuzzy quantifier [20] that evaluates the aggregation of the value $A_k$ within the FTW $T_j$. The set of quantifiers defined in this domain are represented by the fuzzy sets shown in figure 10. The quantifier applies a transformation $\mu_{Q_k} : [0,1] \rightarrow [0,1]$ to the aggregated degree of $\mu_{Q_k}(A_k \cup T_j(S^r))[21].$
In this paper, for our study case, five HKIs (Health Key Indicators) have been defined by a supervisor that accurately describe the source streams of HR or AS from the Fitbit data. These HKIs are instances of the protoform defined above:

1) **Most HR is high between 2 and 4 hours**.
2) **At least half of HR is high intensity between 1/2 and 1 hour AND most AS activity is awake between 1/2 and 1 hour**.
3) **At least half of sleep is deep between 2 and 4 hours**.
4) **A few of sleep is deep between 2 and 6 hours AND most sleep is light between 2 and 4 hours**.
5) **A few of sleep is restless between 1/2 and 2 hours AND most sleep is asleep between 1/2 and 1 hour**.

These protoforms can be interpreted as follows. An example is **Most HR is high between 2 and 4 hours**, where \( Q_k \) is Most, \( A_k \) is HR and \( T_j \) is between 2 and 4 hours.

### B. Linguistic description

In a GLiDTS system [22], the generation of protoforms allows to represent with words the knowledge extracted from the data and consequently it is the first step in the process of building a linguistic description. However, the production of a linguistic summary from the protoforms generated requires a further processing to transform them into a text that better suits to the user necessities. To do that, we have proposed a set of linguistic expressions closer to natural language for each HKI discovered. These expressions improve the semantic expressiveness of each HKI (or protoform activated) and shorten its length, which is always a desirable quality in a summary. These expressions are described in Table I.

### V. STUDY CASE

The functionality of this application has been tested on more than 270000 measurements recorded between September, 1st and November, 26th of 2019 by a single device and used by a single person. These measurements present the following characteristics:

- Each HR measurement involves the check out of the previous registered AS value which is assumed that it did not change and it is registered together with their date time afterwards.
- Each AS measurement involves the check out of the previous HR value, to be registered afterwards.

The first goal of this application consists in summarizing the user daily activity. To do so, the matching between defined protoforms or HKIs and the wearable data is computed for each minute in the complete timeline, generating a protoform degree over the timeline. Next, the maximal aggregation operator is applied to the protoform degrees to compute the maximal protoform degree per day. This degree is highly representative of the matching of the protoforms in a given day, e.g., *has the user done high intensity exercise?, has the user had a low heart frequency for a long time?*.

As an example of proof, we can see the resulting linguistic daily descriptions of the first days of the period in table II, where rules/HKIs with a degree of truth higher than 0.7 have been activated. These rules are shown in the daily summary of the Monwatch application, where, for example on September 1st, the activation of these 3 HKIs: **High HR, sport activity** and **light-long sleep** are shown in the calendar in the figure 7. This summary is equivalent to its extended version: **Most HR is high between 2 and 4 hours, at least half of HR is high intensity between 1/2 and 1 hour and a few of sleep is deep between 2 and 6 hours and most sleep is light between 2 and 4 hours**.

Next, the second objective of the Monwatch application, the generation of user behaviour patterns in terms of HR and AS, is obtained from the activated HKIs. A fuzzy clustering process is applied on the maximal degree per day for all protoforms, to analyze the most representative patterns over different days. The obtained clusters also provide a high representation of daily activities due to relate the degrees between protoforms. For example, a cluster shows days in which exercise is done without a nap, or days without activity and doing a nap.

An analysis of each HKI is performed to obtain the behaviour pattern of the user in a day. This clustering process is really relevant to collect the difference between behaviours for a same user. The proposed clustering method developed is a fuzzy clustering algorithm: fuzzy C-means. In figure 11, we show the obtained clusters from the dataset (with K=3). The main advantage of including a fuzzy clustering in this step is the increase in the interpretability of results from

<table>
<thead>
<tr>
<th>Day</th>
<th>Ling. Descr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sep, 1st</td>
<td>High HR, sport activity and light-long sleep.</td>
</tr>
<tr>
<td>Sep, 2nd</td>
<td>Light-long sleep.</td>
</tr>
<tr>
<td>Sep, 3rd</td>
<td>High HR, low HR awake, sport activity, light-long sleep and nap.</td>
</tr>
<tr>
<td>Sep, 4th</td>
<td>Sport activity and light-long sleep.</td>
</tr>
<tr>
<td>Sep, 5th</td>
<td>Low HR awake, sport activity and light-long sleep.</td>
</tr>
<tr>
<td>Sep, 6th</td>
<td>High HR, sport activity and nap.</td>
</tr>
<tr>
<td>Sep, 7th</td>
<td>…</td>
</tr>
</tbody>
</table>

### Table I

**LINGUISTIC DESCRIPTION OF EACH HKI.**

<table>
<thead>
<tr>
<th><strong>Linguistic description</strong></th>
<th><strong>HKI</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>High HR</td>
<td>Most HR is high between 2 and 4 hours</td>
</tr>
<tr>
<td>Sport activity</td>
<td>At least half of HR is high intensity between 1/2 and 1 hour</td>
</tr>
<tr>
<td>Low-HR-awake</td>
<td>At least half of HR is low between 2 and 4 hour AND most activity is awake between 1/2 and 1 hour</td>
</tr>
<tr>
<td>Deep sleep</td>
<td>At least half of sleep is deep between 2 and 4 hours</td>
</tr>
<tr>
<td>Light-long sleep</td>
<td>A few of sleep is deep between 2 and 6 hours AND most sleep is light between 2 and 4 hours</td>
</tr>
<tr>
<td>Nap</td>
<td>A few of sleep is restless between 1/2 and 2 hours AND most sleep is asleep between 1/2 and 1 hour</td>
</tr>
</tbody>
</table>

### Table II

**DAILY LINGUISTIC DESCRIPTIONS.**

<table>
<thead>
<tr>
<th>Day</th>
<th>日常语文描述</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sep, 1st</td>
<td>高心率，体育活动和长时睡眠</td>
</tr>
<tr>
<td>Sep, 2nd</td>
<td>长时睡眠</td>
</tr>
<tr>
<td>Sep, 3rd</td>
<td>高心率，低心率清醒，体育活动，长时睡眠和小憩</td>
</tr>
<tr>
<td>Sep, 4th</td>
<td>体育活动和长时睡眠</td>
</tr>
<tr>
<td>Sep, 5th</td>
<td>低心率清醒，体育活动和长时睡眠</td>
</tr>
<tr>
<td>Sep, 6th</td>
<td>高心率，体育活动和清醒</td>
</tr>
<tr>
<td>Sep, 7th</td>
<td>…</td>
</tr>
</tbody>
</table>
the evaluated HKI. So each cluster determines a degree of relevance for each HKI within the cluster, which provides an intuitive representation for the expert to determine the relevance of the HKI in the day-pattern. For example, from Figure 11, which describes the most relevant clusters from the dataset, we can observe three different patterns in the user activity:

- **Cluster 1**: Days where user developed a normal sport activity were related to nap sleeps at afternoon. It represents a leisurely pattern.
- **Cluster 2**: Days where user developed low-none sport activity are related to no nap sleeps at afternoon but light sleeps at night. It relates the lack of sport activity with lower quantity of sleep in the user.
- **Cluster 3**: Days where user developed high sport activity are related with lack of low intensity in HR and to no nap sleeps. It relates the high sport with a higher activation in the behaviour of the user through the whole day.

VI. Conclusions

In this proposal, we have presented a web application that monitors the heart rate and the sleep activity of an user through the use of the Fitbit Ionic smartwatch wearable. The architecture of the system has been described here along with a technical description of the connection and implementation mechanism. This application basically allows us to summarize the user’s daily activity in reference to health indicators which are defined in a linguistic way. To do this, the activation of any of the six rules or HKI (Health Key Indicators), which have been defined through experts, lets us get an linguistic description of the user’s activity in a day.

Moreover, a fuzzy clustering operation can be carried out by means of the HKIs generated in this application to discover the most relevant user general behaviour patterns and relation between protoforms according with the HR and AS collected by the wearable device.

Since Monwatch application is presented in a preliminary stage, its functionality is planned to be extended by the analysis of other measures from the wearable device, such as sports activity, or social activity. At the same time, it is planned to carry out a deeper analysis of the clustering process with a greater number of devices and users, where patterns of behaviour can be profiled by sectors of the population in order to evaluate whether these patterns can lead us to diagnose any user disorder in a preventive manner with the collaboration of experts. Such analysis is planned to be displayed in the web application together with the local results of an user.

ACKNOWLEDGMENTS

Funding for this research is provided by EU Horizon 2020 Pharaoh Project Pilots for Healthy and Active Ageing, Grant agreement no. 85718 and the Spanish Ministry of Science, Innovation and Universities and the European Regional Development Fund - ERDF (Fondo Europeo de Desarrollo Regional - FEDER) under project PGC2018-096156-B-I00 Recuperación y Descripción de Imágenes mediante Lenguaje Natural usando Técnicas de Aprendizaje Profundo y Computación Flexible.

Moreover, this contribution has been supported by the Andalusian Health Service by means of the research project PI-0387-2018 and the Action 1 (2019-2020) no. EI_TIC01 of the University of Jaén.

REFERENCES


Figure 11. Resulting patterns of the user with the number of fuzzy clusters N=3.


