Activity Recognition and Personalized Feedback Solution for Active and Healthy Ageing

Thanos Stavropoulos **Georgios Meditskos** **Stefanos Vrochidis** **Ioannis Kompatsiaris**
Centre for Research & Technology Hellas, Information Technologies Institute
6th Km Charilaou – Thermi, 57001, Thessaloniki, Greece
{athstavr, gmeditsk, stefanos, ikom}iti.gr

ABSTRACT
As the average lifespan increases, the care of diseases related to lifestyle and age, such as chronic and neurodegenerative ones, becomes costlier and less accessible, highlighting the need for self-management using technology. This paper proposes a pervasive system for autonomous healthy ageing, which integrates two layers of intelligence via semantic interpretation: multi-modal sensor fusion from smart devices, wearables and multimedia, and personalized spoken feedback based on context-sensing and user input. Aiming for a practical, acceptable system, the proposed architecture considers aspects of integration, security, privacy and cost. The currently implemented components include activity recognition and problem detection, complemented by end-user applications and personalized spoken feedback. The proof-of-concept implementation is evaluated both in a lab setting, for the more complex personalized feedback component, and in four real home environments, presenting efficient activity recognition, and improvement in several neuropsychological areas, such as mood, physical functional and cognitive condition of elders.

Author Keywords
Ambient Assisted Living; Active and Healthy Ageing; Sensors; Knowledge Management; Reasoning; Internet of Things; Personalized Feedback; Activity Recognition;

INTRODUCTION
The increase of the average lifespan across the world has been accompanied by an unprecedented upsurge in the occurrence of dementia, with high socio-economic costs, reaching 818 billion US dollars worldwide, in 2015[1]. Nevertheless, its prevalence is increasing as the number of people aged 65 and older with Alzheimer's disease may nearly triple by 2050, from 46.8 million to 131 million people around the world, the majority of which, living in an institution [1]. Dementia, as well as several other ailments such as depression, cardiovascular diseases, obesity and bad habits (like smoking), require consistent lifestyle changes, usually through interventions driven by experts. However, as the number of people in need of care together with the high costs as well as the inability in several regions for such high quality services prohibit in-person treatment.

However, assistive technologies promise to alleviate those barriers by providing low cost self-management or at least remote and efficient clinical care. In detail, several such technologies, employing Internet of Things (IoT) on the rise, are used to subjectively and intelligently enhance clinical diagnosis and decision making e.g. by efficiently sensing and estimating cognitive status and disease progression faster than standard neuropsychological tests [2]. Furthermore, assistive technology, also met as Ambient Assisted Living (AAL), is expected to play a critical role in improving quality of life, both on cognitive and physical level, by providing tailored interventions, advice and support, without the need of costly in-person care [3].

Still, current systems present many drawbacks, namely many of them target a single purpose (e.g. pharmacological treatment) or aspect (e.g. sleep quality or exercise). Other systems are still based on end-user interviews, leading to generic interventions. Even though remote monitoring of patients is a promising “patient-centered” management approach that provides specific and reliable data, enabling the clinicians to monitor daily function and provide adaptive and personalized interventions, these systems must provide a multi-modal view of several aspects combined and be complemented with self-management functionality to reduce the effort of clinicians.

Towards this direction, we propose a holistic approach for context-aware monitoring and personalized home care, together with intuitive end-user interfaces to autonomously prolong independent living. To begin with, the system integrates a wide range of sensor modalities and high-level analytics to support accurate monitoring of all daily life aspects including physical activity, sleep and activities of daily living (ADLs). After all gathered knowledge is represented in a universal format, semantic interpretation, via a hybrid reasoning scheme, is used for complex activity recognition from atomic events, emotional and well-being status and highlighting clinical problems.

The high-level meaningful information is presented in applications tailored to clinicians, but most importantly end-users themselves. They are also to be exploited further so as to automatically provide support and suggest interventions, combined with spoken user input as context. The proposed system is intended for real-life and wide-spread usage, hence, security and privacy aspects, cost, equipment, acceptance, integration and interoperability

[1] Dementia Statistics by Alzheimer’s Disease Internationl - https://www.alz.co.uk/research/statistics

aspects are also discussed. A proof-of-concept has been deployed and evaluated either in home settings focused in dementia, showing effective performance for activity recognition and long-term improved in several domains.

The system builds upon lessons learned from previous work: smart home, wearable, image and audio sensing integrated in an existing service-oriented middleware, together with semantic models for activity recognition [4]. The novelty of this paper is their further integration of sensed qualities as context, with an additional layer of intelligent spoken feedback, enabling autonomy and previously impossible self-management, in a platform that accounts for security, usability and cost aspects.

The following sections present: related work, overall aspects and requirements, the proposed architecture, sensing and analysis, the personalized spoken feedback method, the end-user interface, proof-of-concept evaluation, conclusions and future work.

RELATED WORK

Pervasive technology solutions have already been employed in several ambient environments, either homes or clinics, but most of them focus on a single domain to monitor, using only a single or a few devices. Such applications include wandering behavior prevention with geolocation devices, monitoring physical activity, sleep, medication and performance in daily chores [3] [2].

In order to assess cognitive state, activity modelling and recognition appears to be a critical task, common amongst existing assistive technology. OWL has been widely used for modelling human activity semantics, reducing complex activity definitions to the intersection of their constituent parts. In most cases, activity recognition involves the segmentation of data into snapshots of atomic events, fed to the ontology reasoner for classification. Time windows [5] and slices provide background knowledge about the order or duration [6] of activities. In this paradigm, ontologies are used to model domain information, whereas rules, widely embraced to compensate for OWL's expressive limitations, aggregate activities, describing the conditions that drive the derivation of complex activities e.g. temporal relations.

Focusing on clinical care through sensing, the work in [7] has deployed infrared motion sensors in clinics to monitor sleep disturbances, limited, though, to a single sensor. Similarly, the work in [8] presents a sensor network deployment in nursing homes to continuously monitor vital signs of patients. Other systems employ environmental sensors to observe and assess activities [9] or security monitoring with actuators to control doors [10]. Nevertheless, it so far lacks the ability to fuse more sensor modalities such as sleep and ambient sensing, with limited interoperability. On the other hand, the proposed system offers a unified view of many life aspects, including sleep and activities, to automatically assess disturbances and their causes, to support end-users and clinicians.

Ontologies have been extensively used in natural language interfaces and information extraction [11][12], offering vocabularies and reasoning services to fuse contextual information [13] and solve disambiguation problems [14]. However, the ability to provide personalized responses requires not only language understanding, but also coupling profile and clinical knowledge. In our work, this coupling for personalized responses is realized through a combination of ontology reasoning and SPARQL.

THE PROPOSED SYSTEM

The scope of the system is to assemble a secure, compact solution, deployable for a wide audience and enabling self-management for healthy ageing, utterly reducing effort and cost of clinical dependence and early hospitalization. To do so, the system must not only employ reliable sensing, but also adaptive personalized and human-intuitive interaction, while clinical oversight and visits are becoming rarer. The proposed system addresses these requirements by integrating a set of best practices, the latest technological standards as well as valuable lessons learned from past research. The overall concept is shown on Figure 1. In general, the system incorporates two layers of intelligence: holistic multi-modal sensing interpretation and personalized spoken feedback. At the sensing layer, sensor and multimedia analysis are semantically combined to provide higher-level meaningful qualities. These constitute the user status after interpretation, namely physical, emotional, cognitive, medical and social state (e.g. via monitoring heart rate, stress, word utterance and medication). The second layer of intelligence capitalizes this information as context, which combined with spoken user input, can lead to further personalized feedback. On the other hand, clinicians are constrained to only providing the recommended set of interventions to the agent, perform seldom clinical visits, elicit end-user and clinical requirements and evaluating the system.

Moving from concept to implementation, the proposed system follows a multidisciplinary approach to integrate and bring into effect clinical expert knowledge in an AAL system. The system employs a synergy of the latest advances in sensor technologies, fusion and mining, knowledge representation and personalized feedback. The detailed architecture is shown on Figure 2. A sensing submodule integrates several heterogeneous devices and protocols in order to satisfy the variety of modalities mandated by the requirements. The modalities are retrieved by various lifestyle sensors and wearables. However, the sensing submodule also integrates more complex data format retrieval in the form of image and audio together with the corresponding specialized processing techniques.

After an early fusion and preprocessing takes place, information is unanimously stored in the Knowledge Base (KB) where it is interpreted and fused into higher-level meaningful information such as physiological, medical, cognitive, emotional and social aspects. Together with

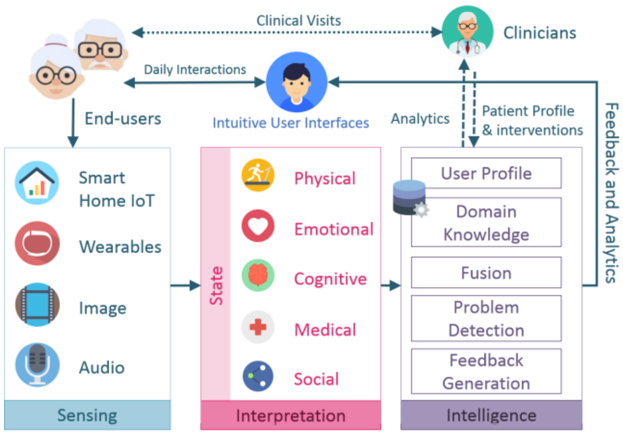

Figure 1. Overall system concept, sensing, interpretation, intelligent coaching and end-user interaction.

expert knowledge apart from showing the progress to end-users and clinicians, the system employs personalized feedback to provide advice and support for end-user self-management.

To maintain compatibility with global IoT solutions, the proposed system is built as a service-oriented middleware which can provide adapters to integrate with open IoT platforms. This process helps provide service discovery, matching and composition both internally to the system but also externally. The system can benefit from existing solutions for the lifecycle that open platforms provide and disseminate its own services to their wider ecosystems of solution providers. Some of the most popular IoT platforms to provide adapters for, are FIWARE, universAAL and the emerging and most relevant, ACTIVAGE[2].

Regarding the privacy, the proposed techniques for an AAL system extend beyond simple approaches, such as the removal or masking of the direct identifiers (i.e. names, identifications IDs, etc.), to mature technologies such as k-anonymity and methods such as differential privacy, syntactic anonymity, homomorphic encryption, secure search encryption and secure multiparty computation.

Regarding security, the proposed system employs standard enterprise protocols for secure authentication (OAuth) and transmission/retrieval (HTTP/SSL) from end-user sites to the cloud. Beyond most existing commercial cloud services, AAL systems should not only encrypt transmissions (to repel man-in-the-middle attacks) but also encrypt stored data and the passwords to decrypt them in the client side. This approach, namely Zero Knowledge cloud storage, is considered to be the most current and trustworthy method as the beholder (service provider) himself is unable to view (and further exploit) sensitive information. Depending on the installation, commercial cloud infrastructures can provide Zero Knowledge, but it can also be implemented easily in owned cloud infrastructure.

In the proposed architecture, the clinical process is continuously supported in two modes: the validation and the operation mode. During the validation mode, clinicians can transcode their expert knowledge for interventions into the system while they also perform clinical visits to the end-users for complimentary assessment and interventions. The system continuously supports them. End-user and clinician requirements periodically reform the system according to validation results. On the other hand, in the operational mode, interventions are more or less pre-decided, clinical visits are rare and the system is not reformed until critical updates, to allow smooth deployment and operation.

Since this proposed architecture is multi-layered and diverse to be thoroughly presented here, this paper focuses on the key components for personalized feedback. The following sections present an overview of sensing modalities and analysis, fusion and personalized feedback, which are then evaluated in a proof-of-concept implementation.

SENSOR DATA RETRIEVAL AND ANALYSIS

The sensing submodule includes two streams for data retrieval and processing, sensors and multimedia, as well as the modules for early fusion, i.e. preprocessing and transformation. After all information is unanimously stored in the knowledge base, it can be further interpreted and displayed to end-user or clinician interfaces. Further details for each of these modules is given below.

IoT Lifestyle and Wearable Sensing

The system currently integrates a wide selection of proprietary, low-cost, ambient or wearable devices, originally intended for lifestyle monitoring, repurposed to a medical context. This variety satisfies both the required modalities and the user needs according to context, always finding a balance between comfortability and functionality.

In detail, *Ambient depth cameras*[3] are collecting both image and depth data. The *Plug* sensors[4] are attached to electronic devices, e.g. to cooking appliances, to collect power consumption data. *Tags*[5] are attached to objects of interest, e.g. a drug-box or a watering can, capturing motion events and *Presence* sensors are modified Tags that detect people's presence in a room using IR motion. A selection of wearable *Wristwatches*[6] according to needs may measure physical activity levels in terms of steps, heart rate and

[2] https://www.fiware.org, http://universaal.sintef9013.com, http://www.activageproject.eu

[3] Xtion Pro (http://www.asus.com/Multimedia/Xtion_PRO/)

[4] Plugwise sensors (https://www.plugwise.nl/)

[5] Wireless Sensor Tag System (http://wirelesstag.net/)

[6] Jawbone UP24, FitBit Charge HR, Microsoft Band, == and Empatica E4

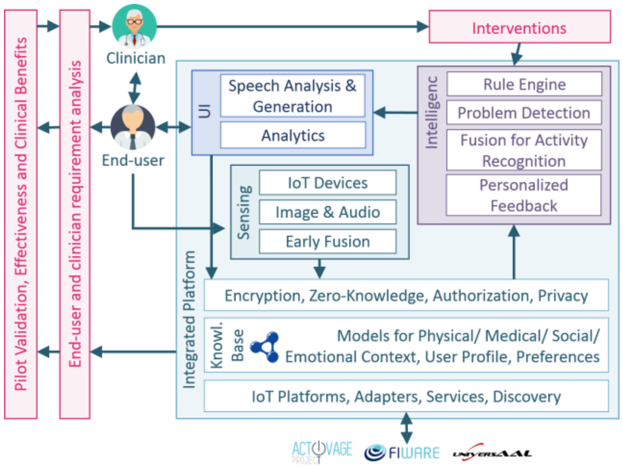

Figure 2. System architecture, end-user and clinician evaluation process.

EDA, while a pressure-based *Sleep sensors*[7] are placed underneath the mattress to record sleep duration, phases and interruptions.

Each device is integrated by using dedicated modules that wrap their respective API, retrieve data and process them accordingly to generate atomic events from sensor observations e.g. through aggregation. In the case of image data, computer vision techniques are employed to extract information about humans performing activities, such as opening the fridge, holding a cup or drinking [15]. Standard microphones are used to retrieve audio.

Fusion, Activity Recognition and Problem Detection

To obtain a more comprehensive image of an individual's condition, semantic fusion is used to transform atomic sensor events to complex ones, such as daily activities, and identify problematic situations. For this purpose the system employs a hybrid combination of OWL 2 reasoning and SPARQL.

Regarding activity recognition, a simple pattern models the context of complex activities. Each activity context is described through class equivalence axioms that link them with lower-level observations of existing domain models (which can be found in [4]). The instantiation of this pattern is used by the underlying reasoner to classify context instances, generated during the execution of the protocol, as complex activities. The instantiation involves linking ADLs with class equivalence axioms. For example, given that the activity *PrepareTea* involves the observations *KettleOn*, *CupMoved*, *KettleMoved*, *TeaBagMoved* and *KettleOff*, its semantics are defined as:

[7] Withings Aura and Beddit

$$
\begin{aligned}
PrepareTea \equiv\ & Context \sqcap \exists contains.KettleOn \\
& \sqcap \exists contains.CupMoved \\
& \sqcap \exists contains.KettleMoved \\
& \sqcap \exists contains.TeaBagMoved \\
& \sqcap \exists contains.KettleOff
\end{aligned}
$$

According to clinical experts involved in the development so far, highlighting problematic situations next to the entire set of monitored activities and metrics would further facilitate and accelerate clinical assessment. This is addressed by a set of predefined rules (expressed in SPARQL) with numerical thresholds that clinicians can adjust and personalize to each of the individuals in their care. Furthermore, each analysis is invoked for a period of time allowing different thresholds for different intervals e.g. before and after a clinical intervention. Problematic situations supported so far regard night sleep (short duration, many interruptions, too long to fall asleep), physical activity (low daily activity totals), missed activities (e.g. skipping daily lunch) and reoccurring problems (problems for consecutive days). The following example illustrates a rule for a short sleep duration problem:

```
CONSTRUCT {
       ?new a :SleepDurationProblem;
              :duration ?D; :date ?date.
}
WHERE {
  ?activity a :Sleep; :startTime ?st;
    :endTime ?et.
  BIND(:duration(?st, ?et) as ?D)
  {
    SELECT ?_d ?ActivityType
    WHERE {
      ?p a :SleepDurationPattern;
        :hasDescription [
          :definesActivityType[
            :classifiesActivity :Sleep;
            :hasDurationDescription [
              time:seconds ?_d]]].
     }
  }
  FILTER(?D > _d)
  BIND (extract_date(?startTime) as ?date)
}
```

PERSONALISED SPOKEN FEEDBACK

Empowering and motivating people in need of guidance and care due to age-related conditions is essential in order to preserve elderly's ability to remain active and independent, with the highest quality of life. This second layer of intelligence in the system shows how monitored qualities can be exploited for personalized assistance, suggestions and recommendations, using again ontologies and rules. This closed loop between the elderly and the system is realized through a natural language interface. Users are able to ask questions about their daily activities and habits, getting feedback and suggestions about health-related problems and situations in return. The underlying processes entailed are:

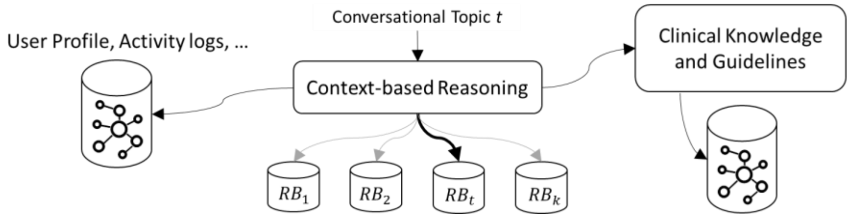

Figure 3 Conceptual dependencies among user profile, topics and clinical guidelines

1. **Automated Speech recognition:** In order to support the transformation of spoken language into text, we use a state-of-the-art ASR system[8] that employs statistical speech models for both acoustic and language modeling, specifically trained for basic and healthcare domains.
2. **Language analysis:** The language analysis consists in itself of four stages: a) surface-syntactic parsing, b) deep-syntactic parsing, c) frame-semantics parsing, and d) projection to ontological representations [16]. The output is a set of FrameNet-based structures projected to a DOLCE+DnS UltraLite[9] compliant representation.
3. **Question topic understanding:** This tasks is responsible for bringing conversational awareness into the system, recognizing the topic of the question based on the language analysis results and on a topic OWL 2 ontology.
4. **Question interpretation and reasoning:** Each concrete topic of the topic ontology is associated with a rule template that couples profile and clinical knowledge to derive the context of the response to a given question.
5. **Language generation:** The verbal communication capitalizes on the ontological representations returned from question interpretation, following the inverse cascade of processing stages described in language analysis.

In this paper, the focus is given on topic understanding, interpretation and reasoning (points 3 and 4). The rest of this section describes the ontologies used to capture topics, clinical knowledge and how knowledge derived from sensor monitoring can be combined with rules to provide personalized feedback and suggestions.

Topic Understanding

An important aspect is the ability to recognize the context of the question, so as to trigger the appropriate rule template and meaningfully extract and combine knowledge to respond to the user's inquiry. The topic ontology was designed collaboratively with the clinical experts that suggested both the needed topics and the underlying spoken language semantics that characterize each of them. The modelling follows a hierarchical decomposition of abstract topics into concrete ones, defining necessary and sufficient OWL 2 restrictions for class membership. Examples of such restrictions that capture domain knowledge and formulate the verbal vocabulary the system can understand in the form of a Description Logic theory model [17] are shown below:

- $PainTopic \equiv Context \sqcap \exists contains.(Pain \sqcup Hurt)$
- $HeadacheTopic \equiv PainTopic \sqcap \exists contains.Head$
- $BackPainTopic \equiv PainTopic \sqcap \exists contains.Back$
- $ActivityTopic \equiv Context \sqcap \exists contains.Activity$
- $SleepTopic \equiv ActivityTopic \sqcap \exists contains.Sleep$
- $NightSleepTopic \equiv SleepTopic \sqcap \exists contains.Night$
- $StressTopic \equiv Context \sqcap \exists contains.Stress$
- $DurationTopic \equiv Context \sqcap \exists contains.Duration$
- $ActivityDurationTopic$
$\equiv ActivityTopic \sqcap DurationTopic$

The root of the hierarchy is the $Context$ class that allows $contains$ property assertions to be defined for associating language analysis results. More specifically, the context of each question is represented as an instance of the $Context$ concept that is associated though multiple $contains$ assertions with language analysis frame entities (verbal domain model). For example, a context instance (i.e. user utterance) containing the domain elements $Night$, $Sleep$ and $Duration$ is automatically classified by the ontology reasoned in the topic $ActivityDurationTopic$. As described in the next section, such multiple classifications act as semantic annotations of utterances and are used to extract relevant information from the KB through rules.

Reasoning and Feedback

Like in domain models and spoken vocabulary, clinical experts were involved to define the logic behind coupling templates with user information. A context-aware knowledge extraction module using rule templated was developed to retrieve information. Intuitively, a rule template acts as a conceptual link between question topics, user profiling information and clinical logic (when needed). Topic detection triggers the execution of a rule base that defines the logic to extract, process and return information relevant to user's inquiry. The association of topics with templates is done in an abstract manner, exploiting the subsumption ontological hierarchy. Figure 3 graphically illustrates the concept behind rule templates.

One simple example of template-based reasoning is the ability to answer questions about the duration of certain

[8] http://www.vocapia.com/speech-to-text.html

[9] http://www.ontologydesignpatterns.org/ont/dul/DUL.owl

activities. Such functionality is useful both for care-recipients and care-givers that have the ability to get activity logging information with a natural way. In this context, the classification of utterances in the $ActivityDurationTopic$ class is used to trigger the corresponding rule base, without needing to couple clinical logic. This simple case can be handled by two template SPARQL rules (SPIN rule[10]). More specifically, Rule 1 is executed on top of the RDF graph with the activity logs to return the duration of an activity.

```
#Rule 1
CONSTRUCT {
  [] a :Duration; :value ?d .
}
WHERE {
  [] a :Variable; :name "ACTIVITY_TYPE";
    :value ?$ACTIVITY_TYPE .
  [] a :Variable; :name "DATE"; :value ?$DATE.
  ?activity a ?$ACTIVITY_TYPE; :start ?start ;
    :end ?end .
  FILTER (:match(xsd:date(?start) == ?$DATE)))
  BIND (:duration(?start, ?end) as ?d)
}
```

The variables starting with '$' denote template variables that are instantiated at runtime, based on the context captured in the topic. The custom SPARQL function $:match$ tests the equality of a date value ($?start$) against another symbolic date value (e.g. Yesterday). The $:duration$ function computes the duration given two date time values.

Assuming that the user asked the question: "*What was the duration of the prepare breakfast activity yesterday?*", the following context instance is generated (in Turtle format):

```
:ctx1 a :Context ;
  :contains [a :Duration];
  :contains [a :PrepareBreakfast];
  :contains [a :Yesterday] .
```

where $Duration$, $PrepareBreakfast$ and $Yesterday$ are domain concepts detected through language analysis. According to the domain model described earlier, $ctx1$ is automatically classified in the $ActivityDurationTopic$ class, which triggers the second rule of the rule base (Rule 2) to assert triples for the two variables used in Rule 1.

```
#Rule 2
CONSTRUCT {
  [] a :Variable; :name "ACTIVITY_TYPE";
    :value ?activityType .
  [] a :Variable; :name "DATE"; :value ?period .
}
WHERE {
  ?ad a :ActivityDurationTopic ;
    :contains [a ?activityType] .
  ?activityType rdf:subClassOf :Activity .
  FILTER (?activityType != :Activity)
  ?ad :contains [rdf:type ?period] .
  ?period  rdf:subClassOf :Period .
  FILTER (?period != :Period)
}
```

By inserting the two $Variable$ instances, the graph pattern in Rule 1 is now matched and a $Duration$ triple is returned and sent to the language generation. It is worth mentioning that the query for extracting the variables for Rule 1 requires the presence of a $Period$ concept, denoting the date of the activity (in this case $:Yesterday$). So, a question of the form: "*What was the duration of the prepare breakfast*" does not return any result, since the system misses information about the date when the activity should have taken place.

Example of Spoken Feedback Based on Monitoring

One of the use cases, inspired by the KRISTINA project[11], involves the interaction of users with the system in order to acquire feedback about problems that may have. For example, the use may ask the system: *"Why does my head hurt?"*. This is a question that requires the coupling of clinical knowledge, in order to give as feedback potential causes of the headache, based on the user profile. For example, clinicians suggested that the sleep quality of the previous night should be checked (based on the results provided by the sleep sensor of the framework), together with the number of coffees the user had the last 24 hours. Both are implemented in terms of SPARQL rules that search the user activity monitoring graph to detect relevant patterns.

It is important to highlight that the detection of a question topic (e.g. $HeadacheTopic$ based on language analysis results $Hurt$ and $Head$) triggers a rule base that usually contains more than one rule. Intuitively, each topic is associated with a small rule-based application that tries to match graph patterns in the activity log of the user. As an example, we present the rule for checking the sleep quality of the previous night. If the value computed by the sleep sensors is lower than a threshold, then this fact is marked as a potential cause and returned as feedback to the user.

```
CONSTRUCT {
  [] a :Feedback; :value :NightSleep.
}
WHERE {
  ?activity a :NightSleep; :start ?s ;
    :quality ?q .
  FILTER (?q < 0.4)
  FILTER (:match(xsd:date(?start)==:Yesterday)))
}
```

If sleep quality is less than 0.4, the rule asserts a $Feedback$ instance in the RDF graph that is collected by another rule

[10] http://spinrdf.org/

[11] http://kristina-project.eu

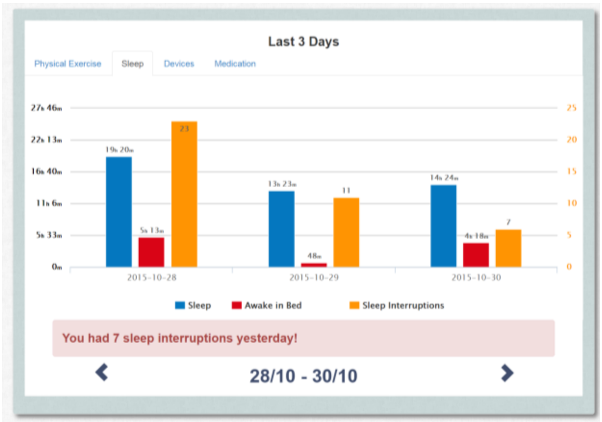

Figure 4. The end-user tablet application interface.

of the rule base to compile an RDF response graph, which is finally forwarded to language generation.

END-USER INTERFACE

Besides the personalized spoken feedback the users also have access to a tablet application with a Graphical User Interface (GUI), tailored to provide comprehensive, intuitive monitoring of their daily life aspects and feedback. In detail it provides a restricted, simplified view of the most important measurements so as to avoid overwhelming the users or even stressing them out. The displayed interval spans across three days of information regarding Physical Activity (daily steps and burned calories), Sleep, Usage of Appliances and Medication. Besides user status and trends, the application also provides feedback with regards to problems detected such as sleep problems. It also provides educational material, such as recipes or step-by-step instructions to perform routine tasks, and the ability to exchange messages between end-users and clinicians. Figure 5 shows an example view with a digested view of three-day trends of sleep aspects and a warning for many sleep interruptions. Overall, the application is designed to help patients feel confident and secure with the system they are using, but also their relatives and carers as well as to encourage social interaction between them. End-user and carer feedback for the application through questionnaires was so far positive.

PROOF-OF-CONCEPT EVALUATION

Due to its very high complexity and the limited scope of this paper, evaluation results are presented only for the most representative of components. Real-world pilot installations were used to test sensor data retrieval, analysis and interpretation, from which we present here the most sophisticated result, namely activity recognition. The other key aspect of this paper, spoken feedback, is a rather complicated component to evaluate as it also requires prior accurate information detection. In this cycle, the knowledge from pilots was used to test the component by experts instead of pilot users to avoid confusing them with an early prototype building the prototype. Finally, clinical aspects are also presented for the pilot installations.

Pilot Installations

The system was evaluated in four home installations, in the residences of individuals living alone (for clinical aspects please refer to clinical evaluation below), and maintained for four months. Two additional installations are still sustained for a one-year total duration study. Sensors and relevant home areas or devices of the installation were selected after a visit from the clinician to the participants and follow the placement guidelines of Table 1. The majority of deployed sensors covered the areas of kitchen, bathroom and bedroom, since these rooms are strongly linked with most activities.

Figure 5 shows a real-world installation, an image processing instance, collection of sensor events and aggregated information on an end-user tablet application (although the installation is real, a professional actor is depicted here to preserve the privacy of actual participants). These pilots were used so far to evaluate sensor data processing and interpretation performance such as activity recognition, stress detection, sleep problem detection and long-term clinical evaluation, some of which are presented below.

Activity Recognition Performance Evaluation

High-level activity recognition via ontology-based fusion has been evaluated from the four real-world pilot installations. Ground truth was obtained via annotation, based on images from ambient cameras. The metrics here are Recall (or True Positive Rate, TPR) and Precision (Positive Predicted Value. PPV), corresponding to activities recognized with respect those actually performed. Clinical experts suggested five activities, which are shown on. Table 2 shows the activities together with pertinent context dependency models.

The evaluation dataset spans over 31 days, in July 2015. As observed on Table 3, the more atomic and continuous an activity is, the more accurate the detection. In practice, *BathroomVisit*, the activity most accurately detected, is never interleaved to do something else. On the contrary, cooking is a long-lasting activity interrupted by instances of other events (e.g. watching TV) and influenced by uncertainty and the openness of the environment. *WatchTV* and *PrepareTea* are fairly short in duration, causing less uncertainty and interleaved events in-between, yielding decent precision and recall rates.

Personalized Spoken Feedback Evaluation

After evaluating the accuracy of sensor recordings and high-level activity recognition, the personalized spoken feedback component was separately examined, as it is not yet part of the pilots. Instead, internal IT and clinical staff was invited to test the current implementation of the component. The assessment was performed in accordance with Good Clinical Practice (GCP) following the procedure: (i) Informed consent was given for recording voice and filling a questionnaire, after a briefing of how data will be used internally. (ii) The participants conducted a guided conversation with the system. (iii) They filled in the questionnaire with assistance of personnel. The process took about one hour per participant and the questionnaires revealed the following critical aspects:

Speech recognition and language analysis: Topic detection depends solely on the speech and language analysis output, which is used by the reasoner to classify utterance contexts in the topic hierarchy. The current implementation is not able to handle missing information and uncertainty. Therefore, the absence of a domain descriptor from the input, e.g. a missing $Hurt$, preventing the detection of the correct topic (e.g. $Headache$) to trigger the corresponding rule base.

Level of detail of the topic hierarchy: The rules for collecting and returning feedback are based on abstract dependencies. For example, it is assumed that for an utterance with $Hurt$ and $Head$ the system should always return possible causes of headache. However, the same concepts are detected when the user just says: “My head hurts”, for which recommendation on how to stop the pain would be more relevant than possible causes. Currently, the system is not able to distinguish such cases, since the topic hierarchy defines very abstract dependences between topics and concepts, while language analysis does not provide the type of utterance (e.g. statement, question, etc.).

Sensor	Placement area or object
Camera	Kitchen, Living room, Hall
Plugs	TV, Iron, Vacuum, Cooking device,
Tags	TV remote, Iron, Fridge door, Drug
Presence	Kitchen, Bathroom, Living room
Wristwatch	Worn on the wrist
Sleep	Under the mattress

Table 1. Sensors in home installation.

Activity	Context dependency set
PrepareDrugBox	DrugBoxMoved, DrugCabinetMoved,
Cooking	TurnCookerOn, KitchenPresence
PrepareTea	TurnKettleOn, TeaBagMoved,
WatchTV	TurnTvOn, RemoteControlMoved,
BathroomVisit	BathroomPresence, TurnBathrLightsOn

Table 2. Context dependency models for the evaluation.

Activity	TPV	PPV	Activity
PrepareDrugBox	0.86	0.89	PrepareDrugBox
Cooking	0.61	0.68	Cooking
PrepareTea	0.81	0.86	PrepareTea
WatchTV	0.87	0.80	WatchTV
BathroomVisit	0.91	0.94	BathroomVisit

Table 3. Precision and recall for activity recognition.

Improvement of Clinical Condition

Six individuals living in six separate homes participated: five female (four Amnestic MCI, one mild dementia – AD) and one male (mild dementia – AD). Besides regular approximately weekly clinical visits, the system supported clinical objective insights as well as the participants and their family. Significant improvement was found in post-pilot clinical assessment in multiple domains, utterly bringing about positive change in mood and cognitive state, measured objectively via neuropsychological tests. In detail, the first participant has overcome insomnia ($p=0.001$) and neglecting daily chores ($p=0.000$), the second has shown improvement in sleep ($p=0.000$) and lingering in the bath ($p=0.03$), while the other two have been benefited with respect to sleep interruptions ($p<0.02$), lack of sleep ($p<0.08$) and medication. Physical exercise increased for participant three ($p=0.0$) while TV watching reduced ($p=0.03$) and personal hygiene improved for participant four (p-0.001). While this information was derived from statistical processing of system knowledge, neuropsychological assessment post pilot showed statistically significant improvement (pair sample t-test) in scales: Rivermead Behavioral Memory Test ($p=0.03$), MMSE (p=004), Hamilton depression scale ($p=0.01$), MoCA ($p=0.004$) and Rey Auditory Verbal Learning Test ($p=0.04$). Two participants converted from aMCI to SCI, with no pathological depression or anxiety symptoms, and one with moderate dementia switched to mild. For the last participant, state was unchanged but symptoms related to Parkinson’s were highlighted, showing the multi-domain coverage of the system. Detailed clinical aspects are available in [18].

CONCLUSION AND FUTURE WORK

This work has showcased a proposed system for self-managing healthy ageing, based on multi-modal sensor data analysis and personalized spoken feedback. The proposed architecture has considered not only the required functionality, but also interoperability, acceptability, cost, security and privacy aspects. Evaluation was carried out through four real-world pilots, assessing activity recognition effectiveness and clinical condition improvement, while the more complex spoken feedback was evaluated by experts.

To reach its full potential the system has yet to breach many barriers, especially regarding feedback. Further language semantics and handling uncertainty can provide for more use cases of feedback, as shown in evaluation. Additional pervasive and human intuitive modalities can be added such as an expressive compassionate avatar in hand with emotion sensing. Finally, decision making and acting on the environment, e.g. setting home lighting, teleconferencing with friends and step-by-step guidance, can have much greater positive impact in people's health.

ACKNOWLEDGMENTS

This work has been funded by the projects H2020-ICT-645012 KRISTINA and H2020-IOT-732679 ACTIVAGE.

REFERENCES

[1] M. Prince, A. Wimo, M. Guerchet, G. Ali, Y. T. Wu, and M. Prina, "World Alzheimer Report 2015. The global impact of dementia. An analysis of prevalence, incidence, cost and trends," *Alzheimer's Dis. Int. London*, 2015.

[2] P. N. Dawadi, D. J. Cook, M. Schmitter-Edgecombe, and C. Parsey, "Automated assessment of cognitive health using smart home technologies.," *Technol. Health Care*, vol. 21, no. 4, pp. 323–43, 2013.

[3] C. Kerssens *et al.*, "Personalized Technology to Support Older Adults With and Without Cognitive Impairment Living at Home," *Am. J. Alzheimers. Dis. Other Demen.*, vol. 30, no. 1, pp. 85–97, 2016.

[4] T. G. Stavropoulos, G. Meditskos, and I. Kompatsiaris, "DemaWare2: Integrating sensors, multimedia and semantic analysis for the ambient care of dementia," *Pervasive Mob. Comput.*, 2016.

[5] G. Okeyo, L. Chen, H. Wang, and R. Sterritt, "Dynamic sensor data segmentation for real-time knowledge-driven activity recognition," *Pervasive Mob. Comput.*, vol. 10, pp. 155–172, 2014.

[6] T. Patkos, I. Chrysakis, A. Bikakis, D. Plexousakis, and G. Antoniou, "A reasoning framework for ambient intelligence," in *Artificial Intelligence: Theories, Models and Applications*, Springer, 2010, pp. 213–222.

[7] R. Suzuki, S. Otake, T. Izutsu, M. Yoshida, and T. Iwaya, "Monitoring Daily Living Activities of Elderly People in a Nursing Home Using an Infrared Motion-Detection System," *Telemed. e-Health*, vol. 12, no. 2, pp. 146–155, Apr. 2006.

[8] Y.-J. Chang, C.-H. Chen, L.-F. Lin, R.-P. Han, W.-T. Huang, and G.-C. Lee, "Wireless sensor networks for vital signs monitoring: Application in a nursing home," *Int. J. Distrib. Sens. Networks*, vol. 2012, 2012.

[9] S. Helal, W. Mann, J. King, Y. Kaddoura, and E. Jansen, "The gator tech smart house: A programmable pervasive space," *Computer (Long. Beach. Calif).*, vol. 38, no. 3, pp. 50–60, 2005.

[10] S. G. Bonner, "Assisted interactive dwelling house," in *Proc. 3rd TIDE Congress: Technology for Inclusive Design and Equality Improving the Quality of Life for the European Citizen*, 1998, vol. 23, p. 25.

[11] J. Sleeman and T. Finin, "Type prediction for efficient coreference resolution in heterogeneous semantic graphs," *Semant. Comput. (ICSC), 2013 IEEE*, 2013.

[12] R. Prokofyev, A. Tonon, M. Luggen, and L. Vouilloz, "SANAPHOR: ontology-based coreference resolution," *Int. Semant.*, 2015.

[13] S. Dourlens, A. Ramdane-Cherif, and E. Monacelli, "Multi levels semantic architecture for multimodal interaction," *Appl. Intell.*, vol. 38, no. 4, pp. 586–599, Jun. 2013.

[14] D. Damljanović, M. Agatonović, and H. Cunningham, "Improving habitability of natural language interfaces for querying ontologies with feedback and clarification dialogues," *Web Semant. Sci.*, 2013.

[15] K. Avgerinakis, A. Briassouli, and I. Kompatsiaris, "Recognition of Activities of Daily Living for Smart Home Environments," in *2013 9th International Conference on Intelligent Environments*, 2013, pp. 173–180.

[16] L. Wanner *et al.*, "Kristina: A knowledge-based virtual conversation agent," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10349 LNCS, pp. 284–295.

[17] M. Ballesteros, B. Bohnet, S. Mille, and L. Wanner, "Data-driven sentence generation with non-isomorphic trees." ACL (Association for Computational Linguistics), 2015.

[18] I. Lazarou *et al.*, "A Novel and Intelligent Home Monitoring System for Care Support of Elders with Cognitive Impairment," *J. Alzheimer's Dis.*, vol. 54, no. 4, pp. 1561–1591, Oct. 2016.