Improving Work Life Conditions via Portable Knowledge-Driven Recommender System

(Discussion Paper)

Maurizio Atzori[1], Alessandro Barenghi[2], Sara Comai[2], Mariagrazia Fugini[2], Diego Marcia[1], Gerardo Pelosi[2], Manuela Sanguinetti[1] and Vincenzo Scotti[2]

[1]*Dipartimento di Matematica e Informatica (DMI), Università di Cagliari, Via Ospedale 72, 09124, Cagliari (CA), Italy*
[2]*Dipartimento di Elettronica Informazione e Bioingegneria (DEIB), Politecnico di Milano, Via Ponzio 34/5 - Via Golgi 42, 20133, Milano (MI), Italy*

Abstract

WorkingAge is a EU H2020 Project (lasting 2019-2022) aiming at promoting healthy habits in working environments to improve quality of life of workers. IoT sensors are used to detect environmental and workers conditions. Raw sensors data are then transformed into interventions in the form of reminders or recommendations communicated to the user through a mobile application. This paper describes the main issues of data management in the project: the knowledge-based DSS that has been developed to infer the recommendations for the users, which is based on a set of probabilistic rules and takes into account users' preferences to promote users' well being, and the design choices to guarantee data security and privacy.

Keywords

Recommender System, Probabilistic Prolog, Privacy, Security, Risk Management, Quality of Life, Decision Support System

1. Introduction

This paper presents the overall approach and mid-term results of the WorkingAge (WA - Smart Working environments for all Ages) Project, which focuses on the use of innovative Human Computer Interaction methods, like augmented reality, virtual reality, gesture/voice recognition and gaze tracking. The purpose is to measure the user's health state at workplaces and provide recommendations about possible corrections of damaging or harmful states. The large scale introduction of social and technological innovation, such as e-health, mobile health, integrated care or independent living, can improve the efficiency of health and well being systems. Remote monitoring healthcare models, in particular those that include users actively in the design of

SEBD 2021: The 29th Italian Symposium on Advanced Database Systems, September 5-9, 2021, Pizzo Calabro (VV), Italy
atzori@unica.it (M. Atzori); alessandro.barenghi@polimi.it (A. Barenghi); sara.comai@polimi.it (S. Comai); mariagrazia.fugini@polimi.it (M. Fugini); diego.marcia@unica.it (D. Marcia); gerardo.pelosi@polimi.it (G. Pelosi); manuela.sanguinetti@unica.it (M. Sanguinetti); vincenzo.scotti@polimi.it (V. Scotti)
0000-0001-6112-7310 (M. Atzori); 0000-0003-0840-6358 (A. Barenghi); 0000-0002-9554-8815 (S. Comai); 0000-0002-0692-0153 (M. Fugini); 0000-0003-2576-8836 (D. Marcia); 0000-0002-3812-5429 (G. Pelosi); 0000-0002-0147-22080 (M. Sanguinetti); 0000-0002-8765-604X (V. Scotti)

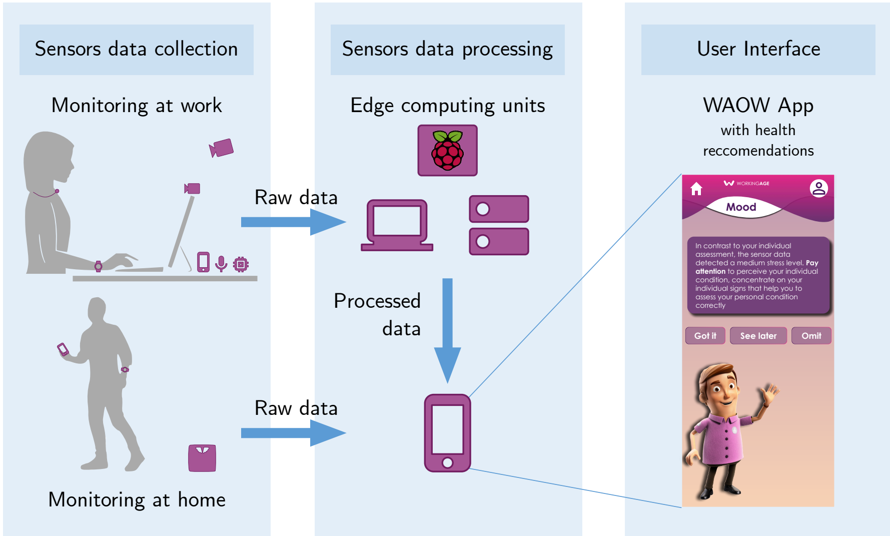

Figure 1: WA Tool concept.

health care systems, have shown clear benefits [1] and are included in recent EU actions [2]. We rely on the general concepts of the WA approach presented in [3, 4]. In particular, these papers describe the portion of the WA Tool that implements a Decision Support System (DSS) developed in the project, and discuss privacy and security aspects. The WA Project aims at creating a sustainable and scalable product that will empower users' comfort by easing work conditions and life, attenuating the impact of aging on their autonomy, health and well-being. A key point is the *adaptivity* of the WA tool to the user profile. Active user engagement is a focus, in order to ensure the match against user needs, safeguarding ethics, privacy, security and regulatory aspects of work premises. In the proposed DSS, recommendations take into account the effectiveness of previous advice on the monitored risks and the acceptance of the user of the given advice. A big issue in creating such a tool is privacy, since personal data are collected by several devices and elaborated to analyze the users features, observe wrong habits and give suggestions to improve wellness. We will show how privacy and security have been tackled in the project to face these issues.

The paper is organized as follows: Section 2 describes the issues related to data privacy and security that impacted the WA architecture. Section 3 describes the DSS we developed to infer recommendations for the users. Finally, Section 4 concludes the paper and outlines future work.

2. Data Privacy and Security

The success of the WA approach requires to collect a significant amount of personal information from different sensors both at work (through cameras, microphones, environmental sensors, etc.) and possibly also at home (through a smart band and a body scale), to provide a fitting set of suggestions to the WA tool user, as outlined in Fig.1. For example, if the camera detects a wrong posture, the worker receives an alert to change his/her position or a tip about a suggested physical exercise. In designing the solution in WA, we built on the considerable computing

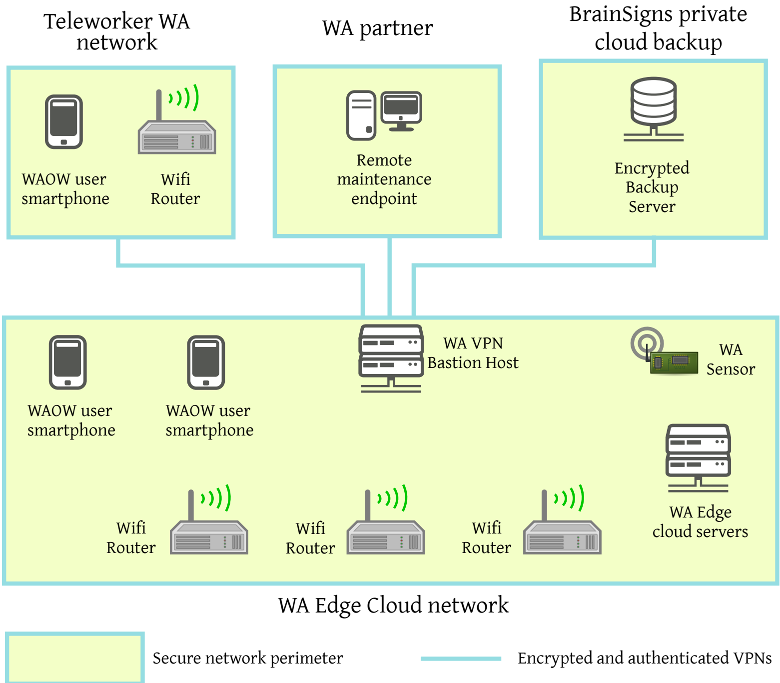

Figure 2: Logic view of the network perimeter within which data confidentiality with respect to external adversaries is provided

power available in modern mobile devices to shift the phase of data aggregation and user profiling entirely on a device which is owned by the user herself. This approach was guided by the provisions of the EU Regulation 2016/679, General Data Protection Regulation (GDPR), which states two fundamental principles: *i)* minimality of data retention and *ii)* security by design. Following the minimality of data retention approach led us to design a system where the raw data collected from the sensors, such as audio/video recordings, are processed as soon as possible by a set of dedicated machines on premise (i.e., the edge computing units in Figure 1). This early processing stage allows us to extract the relevant features from the collected data and prevent the long term storage of raw sensor datasets. Only data from the smart band and body scale that do not need preprocessing are directly sent to the mobile device. Instead, the dedicated set of processing machines supplies the DSS of the WA tool with small and highly informative pieces of data, allowing the decision support agent to be run on the mobile phone of the user. The security by design principle was embodied in the WA network and data processing infrastructure design mapping the legal basis for data treatment provided by the user consent onto cryptographically enforced access control. The first step in this sense is to provide a network infrastructure where data confidentiality is provided with respect to external adversaries. Figure 2 illustrates the WA network structure, highlighting its secure perimeter. We chose to employ well-established secure network standards for the WA edge cloud network, located on premise at the firms which host our test experiments. The network is formed by an IEEE 802.11n Wi-Fi network, employing the WPA2 key agreement and data encryption standard to provide the desired security guarantees. Extending the network to off-site locations to allow WA users to work from home, a highly desirable feature in an evolving work scenario, was performed by means of VPN tunnels relying on the standard Transport Layer Security (TLS) v1.3 secure transport protocol. The edge cloud network is thus extended via a remote VPN endpoint embedded in the WiFi router providing the teleworker WA network. The VPN endpoint connects to the WA VPN bastion host present in the WA edge cloud network placed in the hosting firm premises, allowing the sensing equipment present in the teleworker WA network to securely send the recorded data to the WA edge cloud servers. Finally the WA edge cloud network perimeter is extended, by means of VPN links, also to the WA partner workstations, which may be in need of performing maintenance on the WA edge cloud servers, and to a remote encrypted backup server, maintained by the partner BrainSigns. The latter provides off-site backups of encrypted data bundles, sent to it by both the edge cloud servers and the user smartphones. Having ensured data confidentiality and endpoint authentication for the entities connected to the WA edge cloud network, we complete the enforcement of the access control to the data by means of the standard OpenPGP hybrid encryption, combined with an appropriate certificate management. Each edge cloud server and each smartphone are endowed with a dedicated OpenPGP certificate, containing an RSA keypair and a textual identity. While the textual identity of the edge cloud servers is their common name, the identity of the smartphones is a Universally Unique IDentifier (UUID) version 4, i.e., a 128-bit wide random unique identifier. The WA sensors communicate with the WA edge cloud server which is in charge of performing the data processing to extract high level features either by means of a dedicated TLS connection, if a continuous data stream is required, or send asynchronously data bundles encrypting them with the public key corresponding to the server certificate. The WA edge cloud server in turn will send the processed data to the WA user smartphone encrypting them with the public key contained in the OpenPGP certificate bound to it. This approach effectively ensures that the consent provided by the user, which allows each specific partner to process a subset of the collected data is mapped onto the technically enforced impossibility of accessing data for which the private key required to decrypt them is not available. In addition, it is possible to provide fault resilience in terms of an off-site encrypted backup simply by storing the encrypted messages exchanged within the WA edge cloud network, with no further action. Indeed, only the entity in possession of the appropriate private key will be able to decrypt the backed-up data bundle. The WA smartphones also profit from the availability of the remote backup facility, storing on it a full backup of the DSS-derived user profile, encrypted with their own public key, and copy of the private key symmetrically encrypted under an emergency password derived key. This strategy allows the user, in case of theft or loss of the smartphone itself, to securely retrieve the backup simply providing the emergency password to a new smartphone where the WA app has been installed.

3. Decision Support System

The architecture of the system described in the previous section requires to adopt a hybrid data-driven/expert-driven approach to process the sensors data: for each type of sensor, raw data are processed with data-driven approaches based on Natural Language Processing (NLP), stochastic models, or Machine Learning (ML) to generate high-level information to be sent to the smartphone application. For example, from the video recording the posture is estimated

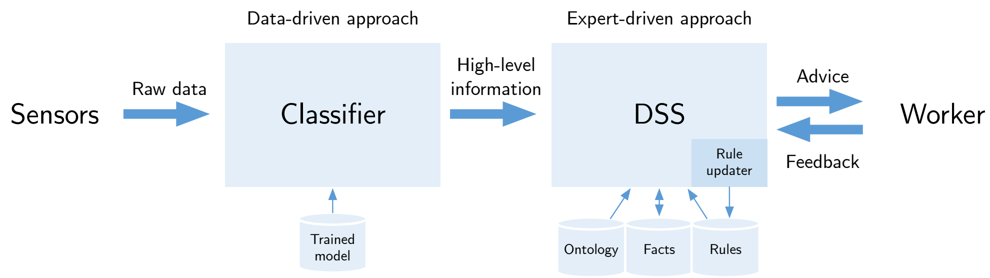

Figure 3: Overview of WA data processing approaches.

and a low/medium/high risk is computed. On the smartphone, a Decision Support System processes the high-level information to generate recommendations for the final users, based on an expert-driven approach. Figure 2 depicts the data processing flow. For coping with this "mismatch" between the data-driven classifiers and the model-driven DSS, we chose to adopt the ProbLog language [5], a probabilistic reasoning engine that allows not only inference of certainly-true facts (according to the usual declarative models), but also derives the level of confidence for the specific inferred fact.

3.1. Ontology

To determine the recommendation for the user, the DSS analyses the high-level information received from the sensors and described according to an Ontology, including eight main entities:

Worker is the core of the ontology, due to our user-centered design;

Profile holds the personal information about the worker (e.g., age, gender, etc.);

Task describes the job performed by the user (e.g., manager, clerk, etc.);

Sensor describes the High-level information collected about the worker (e.g., the sitting body posture risk can be low, medium or high; the facial expression can be positive, neutral or negative, etc.);

Smart Goal and Goal State represent what the WA tool suggests (e.g., a physical activity) and its degree of completion (goal reached or not);

Advice is the suggestion or feedback that provided to the worker (e.g., "Change pose" or "The current heartbeat is 150");

Feedback represents the attitude of the worker towards a piece of advice (e.g., the user accepts or rejects the advice).

Even though the user is the core of the Ontology, the DSS strongly depends on data coming from the sensors, classified into:

Body Sensors include wearable bio-metric devices (to gather ECG and GSR), cameras (to gather body postures, facial expressions and eye movements) and microphones (to collect voice recordings);

Environmental Sensors gather data about temperature, humidity, CO_2 concentration, noise level etc.;

Other Sensors include questionnaires that are periodically administered to understand workers current state.

3.2. Reasoning engine

The DSS developed for this project belongs to the class of knowledge-based DSSes [6] and has been implemented through an expert-driven approach. An expert defined a priori the rules we encoded in our engine. In our use case, the rules describe the *intervention strategy*, which is the set of recommendations to give to the workers. The whole DSS is based on stochastic-aware *Rules* and *Facts*. We employed fact f to state something about the worker or environment and we associated f with a truthfulness probability P_f (we see such probability value as a *reliability index* of f). In the same way, we associated rule $\mathcal{R}$ with a reliability index $P_{\mathcal{R}}$. We used a probabilistic representation to handle the uncertainty about the modelled system (including users and environment). The edge server computes the High-level information through probabilistic models and associates the predicted values with probabilities. To model this uncertainty within the DSS, we employed *ProbLog* [5]. ProbLog is a variant of ProLog, a declarative language and an inference engine, augmented with probabilistic descriptive and inference capabilities.

Hereafter we reported an example of code written with ProbLog.

```
1 1.0::time(morning).
2 0.7::state(user,stressed).
3
4 0.9::candidate_suggestion(take_a_short_pause)
5     :- time(morning), state(user, stressed).
```

Listing 1: Rules and advice example.

Lines *1* and *2* contains two facts: the former states the it is `morning` with a 100% reliability (certainty), the latter states that the `user` is `stressed` with a 70% of confidence. Lines 4 and 5 represent a rule: it suggests to `take_a_short_pause` with a 90% of reliability. The probability of displaying advice $\mathcal{A}$ is given by:

$$P_{\mathcal{A}} = P_{\mathcal{R}} \cdot \prod_{f \in F_{\mathcal{R}}} P_f \ . \tag{1}$$

while, for instance, the probability to show the advice `take_a_short_pause` is:

$$P_{\mathcal{A}} = P_{\mathcal{R}} \cdot P_{\texttt{time(morning)}} \cdot P_{\texttt{state(user,stressed)}} = 0.9 \cdot 1 \cdot 0.7 = 0.63 \ . \tag{2}$$

ProbLog is available as a Python package; as such, it can not run directly in an Android environment. To cope with this issue we adopted *Chaquopy*[1], a Python interpreter for Android, and we developed a wrapper around the whole DSS, to offer the main functionalities through Java. In this way the WA App that represents the user interface can easily interact with the DSS.

[1]https://chaquo.com/chaquopy/

3.3. Rule adaptation

For the scope of this project, we have defined two rule adaptation strategies, to adapt the probabilities of the recommendations on the basis of user preferences and system effectiveness. The former adaptation is referred to as *short-term*, the latter as *long-term* adaptation. For both long- and short-term rules adaptation, we relied on a simple, yet effective technique: the Exponentially Weighted Moving Average (EWMA) [7]. We set the initial probability of a rule $\mathcal{R}$ to a value of 1 at time $\tau = 0$; then, at a given time step $\tau \geq 1$, the chosen adaptation strategy updates the rule probability (reliability in this case), with the EWMA strategy. Short-term adaptation is achieved through user feedback on the recommendations (e.g., the user is (not) going to apply the suggested advice). We paired each $\mathcal{R}$ with a *feedback log* $\mathbf{f}_{\mathcal{R}} \in \mathbb{1}^n$. This feedback log is a sequence of Boolean values representing whether the i-th time $\mathcal{R}$ was triggered (with $i \in [1, n] \in \mathbb{N}$) the user accepted ($\mathbf{f}_{\mathcal{R},i} = 1$) or rejected ($\mathbf{f}_{\mathcal{R},i} = 0$) the advice. Each user has his own feedback log within her/his WA App. Given a feedback log $\mathbf{f}_{\mathcal{R}}$, we computed the short-term update of probability $P_{\mathcal{R}}$ as prescribed by Equation (3)

$$P_{\mathcal{R}}^{(\tau)} = \eta_s \cdot \frac{\sum_{i=1}^{n} \mathbf{f}_{\mathcal{R},i}}{n} + (1 - \eta_s) \cdot P_{\mathcal{R}}^{(\tau-1)} \ . \tag{3}$$

Long-term adaptation is achieved though the measures of effectiveness of the WA tool. In particular, we considered a subset M of such measures, so that by the end the of a time period Δt each measure $m_i \in M$ reports an improvement. In this context, we derived the effectiveness of rule $\mathcal{R}$ as the correlation (during Δt) between the number of times advice $\mathcal{A}$, generated by rule $\mathcal{R}$, collected a positive feedback and the subset of measures $M_{\mathcal{R}} \subseteq M$ such that the Pearson's correlation test between $\mathcal{R}$ and $m_i \in M_{\mathcal{R}}$ results in a p-value ≤ 0.05. On this premise, we defined with R the set of rules in the DSS and with $\rho_{\mathcal{R},M_{\mathcal{R}}}$ the average correlation between $\mathcal{R}$ and its set of measures $M_{\mathcal{R}}$. Thus, we computed the long-term update of probability $P_{\mathcal{R}}$ at time $\tau \geq 1$ as prescribed by Equation (4)

$$P_{\mathcal{R}}^{(\tau)} = \eta_l \cdot \frac{\rho_{\mathcal{R},M_{\mathcal{R}}}}{\max\limits_{\mathcal{R}' \in R} \rho_{\mathcal{R}',M_{\mathcal{R}'}}} + (1 - \eta_l) \cdot P_{\mathcal{R}}^{(\tau-1)} \tag{4}$$

In Equations (3) and (4), the parameters $\eta_s \in (0, 1]$ and $\eta_l \in (0, 1]$ are the learning factors, the parameters that control the updates. Higher values of η_s and η_l make the update system rely more on "newer" information in the update, instead of the previous one; lower values have the opposite effect.

4. Conclusion and Future Work

In this paper, we have presented the WA Tool based on the collection of sensors data and a DSS for the generation of recommendations to improve the quality of life of workers. The WA Tool is conceived to analyze the users' health and behavior at work (or, possibly, telework and smart working) and outside work, assisting with reminders and risks avoidance recommendations. Currently, we are continuing the research stream presented here by designing a platform to define safe working environments for human-robots teams (HRTs) for collaborative embodied (physical) AI. Current work has the aim of keeping people away from unsafe and unhealthy jobs in HRTs and engaging and empowering workers, regardless of their gender, age or background. This platform will enable a strong 'prevention through design' approach that integrates human factors and worker-centred design. The adaptive solutions proposed by the WA project are currently analyzing three different real working settings and living environments, by experimenting the profile of around 90 workers > 50 years old, from the one side, and the working place requirements from the other side. We are currently defining which information has to be collected and stored to improve and increase Robot capabilities and interactions with humans.

Acknowledgments

This work has been supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 826232, project WorkingAge (Smart Working environments for all Ages)

References

[1] A. Dixon, The united nations decade of healthy ageing requires concerted global action, Nature Aging 1 (2021) 2–2.

[2] G. D. Perkins, J.-T. Graesner, F. Semeraro, T. Olasveengen, J. Soar, C. Lott, P. Van de Voorde, J. Madar, D. Zideman, S. Mentzelopoulos, et al., European resuscitation council guidelines 2021: Executive summary, Resuscitation (2021).

[3] M. Fugini, A. Barenghi, S. Comai, G. Pelosi, R. Tedesco, M. van Gasteren, C. A. Catalina, E. Arribas Leal, R. Losada Durán, R. Martins de Almeida, et al., Workingage: providing occupational safety through pervasive sensing and data driven behavior modeling, in: 30th European Safety and Reliability Conference (ESREL 2020), Research Publishing, 2020, pp. 1–8.

[4] R. M. Resende de Almeida, A. Grau Aberturas, Y. Bueno Aguado, M. Atzori, A. Barenghi, G. Borghini, C. A. Catalina Ortega, S. Comai, R. Losada Durán, M. Fugini, H. Gunes, B. Musleh Lancis, G. Pelosi, V. Ronca, L. Sbattella, R. Tedesco, T. Xu, Decision Support Systems to Promote Health and Well-Being of People During Their Working Age: The Case of the WorkingAge EU Project, in: 20th Int. Conf. on Embedded Computer Systems: Architectures, Modeling and Simulation, volume 12471 of *Lecture Notes in Computer Science*, Springer Cham, 2020, pp. 336–347.

[5] L. De Raedt, A. Kimmig, H. Toivonen, Problog: A probabilistic prolog and its application in link discovery., in: Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, volume 7, Hyderabad, Morgan Kaufmann Pub., 2007, pp. 2462–2467.

[6] D. J. Power, Decision support systems: concepts and resources for managers, Greenwood Publishing Group, 2002.

[7] S. Roberts, Control chart tests based on geometric moving averages, Technometrics 42 (2000) 97–101.