Position Papers of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 61–68
DOI: 10.15439/2024F9093
ISSN 2300-5963 ACSIS, Vol. 40

Conceptional Framework for the Objective Work-Related Quality of Life Measurement Through Multimodal Data Integration from Wearables and Digital Interaction

Jenny Voigt, Jakob Hohn
0009-0003-1881-7422
0009-0000-3055-2357
4K Analytics GmbH
Email: {jenny.voigt, jakob.hohn}@4k-analytics.de

Ekaterina Mut, Christian Hrach, Ulf-Dietrich Brauman
0009-0000-1220-3305
0000-0002-1643-188X
0000-0002-0987-4498
Institute for Applied Informatics (InfAI e. V.)
Email: {mut, hrach, braumann}@infai.org

Celine Schreiber, Carsta Militzer-Horstmann
0009-0009-4130-4039
0000-0003-4566-9755
University of Leipzig, Health Economics and Management
Email: celine.schreiber@uni-leipzig.de; Email: militzer-horstmann@wifa.uni-leipzig.de

Sophia Mareike Geisler, Pauline Sophia Pinta, Alisa Hamm, Franziska Stutzer
0000-0001-6300-4349
0009-0004-9522-0897
0009-0006-9574-2964
0009-0003-1850-8750
Scientific Institute for Health Economics and Health System Research (WIG2 GmbH)
Email: {mareike.geisler, pauline.pinta, alisa.hamm, franziska.stutzer}@wig2.de

Hamlet Kosakyan
0009-0000-5103-6677
Appsfactory GmbH
Email: hamlet.kosakyan@appsfactory.de

Hubert Österle
0009-0002-0084-998X
University of St. Gallen, Business Engineering (Professor Emeritus),
Email: hubert@oesterle.ch

Juliette-Michelle Burkhardt, Bogdan Franczyk
0009-0001-1363-6716
0000-0002-5740-2946
University of Leipzig, Information Systems
Email: juliette-michelle.burkhardt@mailbox.tu-dresden.de; Email: franczyk@wifa.uni-leipzig.de

***Abstract*—In the evolving domain of occupational health, assessment of Work-related Quality of Life (WrQoL) has gained critical importance, particularly with recent expedited developments of decentralized and digital work. Conventional methods relying on subjective questionnaires are limited by high drop-out rates and potential biases. This paper introduces a novel approach to evaluating WrQoL by leveraging data generated from digital office environments, wearable devices, and smartphone applications. Our methodology includes the collection of physiological data, analysis of digital interactions, and prosody analysis to construct a comprehensive model of WrQoL influences. Initial and weekly questionnaires as well as multiple daily self-reports of valence and arousal levels will serve to initially validate this model. Prospectively utilizing machine learning, we aim to predict WrQoL scores from aggregated data. This method presents a non-invasive alternative for assessing WrQoL, providing significant implications for both research and industry with the potential to enhance workplace conditions and employee well-being.**

***Index Terms*—job satisfaction, machine learning, Occupational Health, valence, sensors, multimodal data integration, Organizational studies-Behavior**

This project is co-funded by the European Union and is co-financed from tax revenues on the basis of the budget adopted by the Saxon State Parliament (ZX66X3CWB).

I. INTRODUCTION

A. Background: (work-related) quality of life

IN THE contemporary landscape of occupational health and well-being, the concept of Quality of Life (QoL) and, more specifically, Work-related Quality of Life (WrQoL) has gained paramount importance. While recent expedited developments of decentralized work due to the COVID-19 pandemic entailed the promise of a more seamless integration of work and private life, it has simultaneously fractured the traditional work-leisure divide, necessitating a nuanced examination of the impacts of this new reality [1]. These new work constructs not only encapsulate the general well-being of individuals but also highlight the critical interplay between their professional environments and their life satisfaction. Understanding and improving WrQoL is essential for fostering productive, healthy, and sustainable workplaces. Research shows

This project is co-financed from tax revenues on the basis of the budget adopted by the Saxon State Parliament.

Co-funded by the European Union

that high workplace stress severely impacts employees' mental and physical health. For instance, a Harvard University study found that insecure work environments increase poor health risks by about 50%, high job demands raise illness likelihood by 35%, and long working hours elevate mortality rates by nearly 20% [2]. This position paper presents the current research endeavor focused on the objective measurement of WrQoL, proposing a methodological advancement in the assessment and optimization of employee well-being beyond conventional measurement approaches.

QoL is a comprehensive concept encompassing overall well-being, reflecting both positive and negative life dimensions. It is inherently multidimensional, covering emotional, physical, material, and social aspects along with subdomains, such as income and wealth, jobs and earnings, housing, health status and work-life balance [3]–[5]. The World Health Organization (WHO) defines QoL as "an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards, and concerns," highlighting its subjective nature and measurement challenges (WHOQOL-BREF [3]).

Given the significant time spent at work, WrQoL is a crucial indicator of overall well-being and a core component and subdomain of QoL. WrQoL, which focuses on how the work environment affects an individual's overall QoL, is defined as "a multidimensional and dynamic psychological construct, directly related to individual and situational characteristics, which encompasses a set of worker characteristics and specific aspects of the organizational context" [6]. This definition highlights WrQoL's complexity, rooted in both personal attributes and workplace conditions. High WrQoL enhances job satisfaction, mental health, and personal fulfillment benefiting organizations with enhanced productivity, better employee retention, and reduced absenteeism [7]–[10]. Traditional WrQoL measurements rely on subjective self-report surveys, limited by respondent biases, mood fluctuations, and low response rates [11], [12]. This hampers both immediate data collection and longitudinal tracking. Hence, there is growing interest in developing more objective, reliable, and nuanced measurement tools to better understand WrQoL. The "Machine intelligence to objectively measure individual quality of life" (MI-LQ) project aims to objectively measure key indicators of WrQoL utilizing physiological data, digital interaction, and prosodic data, that will be incorporated into machine learning models. Data will be gathered in real office and residential environments as part of a pilot trial involving office workers. A mobile phone app prototype is developed serving as a user interface and a data relay system to a cloud-based platform for offline analysis of WrQoL metrics.

B. State of the art

Current research in the field of capturing and analyzing WrQoL includes a variety of methods that integrate both subjective and objective measurements. Subjective approaches such as questionnaires, interviews, and focus groups allow researchers to explore individual perceptions and interpretations of emotions in the workplace [13]. These predominantly qualitative methods provide important insights into the subjective experiences of employees and help to deepen our understanding of the complex dynamics of the work environment. The important insights include, for example, high workload, lack of support from superiors, job insecurity, inadequate work-life-balance or interpersonal conflicts [14]. Objective measurement approaches, however, rely on advanced technologies to capture and analyze physiological signals, behavioral data, and environmental factors. These advanced technologies can be organized based on the data sources they use, such as mouse, keyboard data, biosensors and mobile phone data.

Mouse and keyboard data

Behavioral data, such as keystrokes, mouse movements [15], and mobile phone activity [16], are used to identify behavioral patterns and derive emotional states. Recent studies by Naegelin *et al.* [17] and Shinde *et al.* [18] illustrate that the merging of physiological and behavioral data using machine learning models can lead to improved detection of workplace stress. Naegelin *et al.* [17] developed a machine learning method for stress detection based on multimodal data (mouse, keyboard, and cardiac data) and tested it in a simulated group office environment. They found that mouse and keyboard data detect stress in the office context better than cardiac data and that certain mouse movements and typing behavior are characteristic for specific stress predictions [17].

Biosensors

Wearable biosensors, such as heart rate monitors and skin conductance monitors, enable continuous data collection on workers' physical responses in different work situations [19]. Studies such as those by Shaffer and Ginsberg [20] and Ernst [21] provide a thorough analysis of heart rate variability metrics and their association with emotional states and work performance factors. Saganowski [22], on the other hand, discusses in particular the commercially available sensors for recording physiological data, signal processing techniques and deep learning architectures for the classification of emotions and the integration of emotion recognition technologies into everyday working life.

Mobile phone data

Furthermore, studies such as those by Burns *et al.* [23] and Hart *et al.* [24] show the usefulness of smartphone sensor technology for detecting depression and assessing well-being in the work context. For example, Burns *et al.* [23] developed a mobile phone application and supporting architecture like the cloud system and programming environment. This enabled the machine learning models to predict patients' mood, emotions, cognitive/motivational state, activities, environmental context, and social context based on at least 38 concurrent sensor readings from the phone (e.g., global positioning system, ambient light, recent calls) [23]. Contrarily, Hart *et al.* [24] investigated whether sparse motion-related sensor data can be used to train machine learning models capable of inferring individuals' states of work-related rumination, fatigue, mood, arousal, life engagement, and sleep quality. The participants' sensor data was collected via questionnaires on their smartphones [24].

Additionally, the objective measurement approaches can be categorized by target outcomes, such as:

Work-Life Balance

Research findings by Pawlicka *et al.* [25] and Gamage and Askana [26] contribute to the prediction of work-life balance and the detection of mental stress in IT work environments. For example, Pawlicka *et al.* [25] examined a machine learning tool to investigate correlations between employee-specific and job-related factors and the subjective feeling of work-life balance. They concluded that the relationship between the feeling of work-life balance and actual working hours was the most significant [25]. Moreover, Gamage and Asanka [26] have worked on a concept for a screening system that can predict mental health problems based on people's external characteristics. Supervised machine learning is used to identify workers at risk and refer them to professional help at an early stage [26].

Stress and emotion detection

Shinde *et al.* [18] developed the Real Time Employee Emotion Detection System (RTEED), which captures facial data in real time and uses machine learning to recognize the emotions of happiness, sadness, surprise, fear and disgust. The system helps companies to monitor the well-being of their employees and sends recognized emotions to the relevant employees to improve their work performance and lifestyle [18].

Moreover, Artificial Intelligence (AI) in speech analysis has the potential to become a crucial tool for measuring workplace stress as this technology can detect and evaluate stress-related strains in real time [27]. The work of Bromuri *et al.* [27] shows that a deep neural network trained for emotion recognition based on speech data can predict stress in call center employees with an accuracy of 80% in real time. This approach enables continuous and unobtrusive monitoring, which can lead to early warning systems and personalized training programs. The study by Baird *et al.* [28] investigates how language features can predict physiological stress markers. By using Long Short-Term Memory-Recurrent Neural Network (LSTM-RNN) and three German speech corpora, the study shows that speech features can effectively predict stress indicators such as cortisol levels, heart rate and respiration, opening up new possibilities for real-time, non-invasive stress monitoring.

These research findings contribute significantly to deepening our understanding of the physiological underpinnings of different aspects of WrQoL and support the development of more accurate models for predicting and assessing well-being in the workplace.

C. MI-LQ project: base model

Building upon our conceptual understanding of overall QoL, our focus is directed towards the nuanced factors shaping WrQoL, recognizing the pivotal role of workplace environment, job satisfaction, and occupational stress within this framework. The MI-LQ project is based on identifying key factors that influence WrQoL and translating them into a base model, which comprises the following indicators: 1) workload, 2) overtime, 3) workspace (office/home office), and 4) commute. All of these factors influence WrQoL through emotional experiences measured in two dimensions: 5) valence and arousal [29]. In addition, because of their strong direct influence on WrQoL, 6) spatial autonomy, 7) task autonomy, and 8) temporal autonomy are core components of the model (Fig. 1). The base model is conceptualized as an initial framework, established within office and residential settings to leverage digital behavioral data collection, crucial for refining and validating WrQoL indicators objectively. This conceptual framework guides the application of a digital assessment approach in a pilot study involving office workers for a minimum duration of four weeks, facilitating the initial validation and refinement of WrQoL assessment methods.

II. CONCEPTUAL FRAMEWORK

A. Data sources and model of data integration

The WrQoL indicators of the base model are measured by different objective measurement approaches. 1) Workload is assessed using calendar system-related data (e.g., Outlook) such as meeting overlap and meeting to work ratio. 2) Overtime is calculated as the difference between absolute hours worked, derived from computerized behavioral data, and self-reported contractual hours worked per week. 3) Workspace is intended to be assessed via workplace booking systems and 4) Commute via GPS. 5) Valence and arousal are determined by several approaches: a) work-related data from mouse and keyboard as well as software (e.g., Outlook, calendar) utilization b) physiological data application such as heart rate variability (HRV) from wrist-worn, c) stress index data based on plethysmography method obtained by ShenHealth application, and d) prosody analysis serving as speech-based emotion classification. 6–8) Autonomy data are mainly gathered by questionnaires and additionally by workplace booking systems (6), task management software, e.g., Trello (7) and working arrangement (8). Currently, the project is focused on gathering behavioral (5a), physiological (5b) and prosodic data (5d).

Features describing movements and actions of the mouse, speed of typing, and error correction of typed words have previously been studied as approach to predict work-related valence and arousal. Investigations have mostly been conducted in laboratory settings [15], [17], [30] and less in real office environments [31], with non-publicly available software.

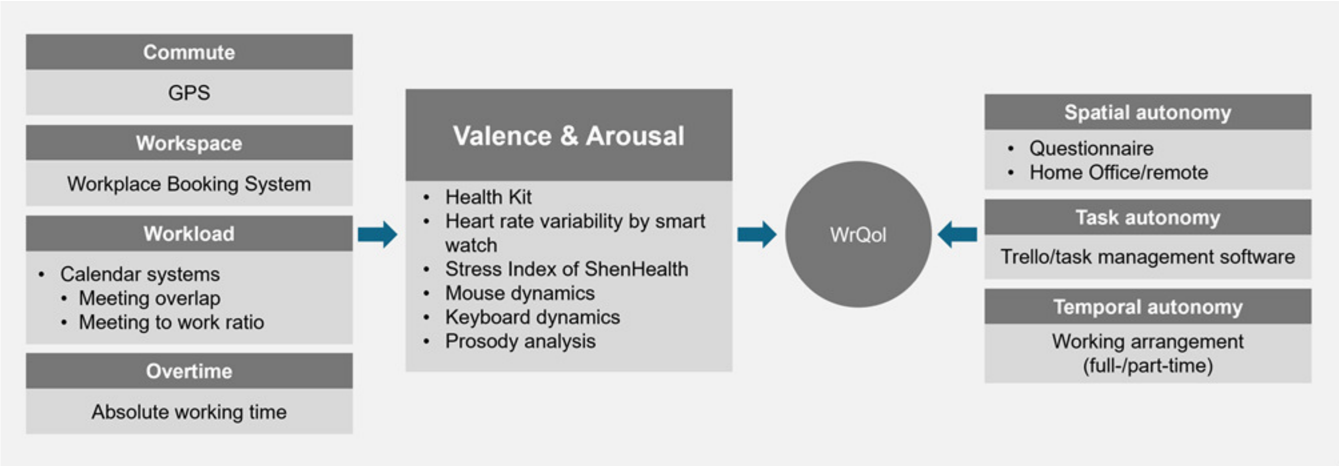

Fig. 1. Base model of WrQoL indicators and their objectifiability with different digital approaches

Therefore, a software prototype was developed to continuously monitor the mouse and keyboard dynamics by the office worker under real office conditions. For each keystroke, the time of press, the time of release and the type of the key pressed are documented, without content of typing. Privacy is ensured by distinguishing between two types of keys: delete keys (Delete, Backspace) and general use keys. Mouse dynamics include recording of mouse operation (move, scroll, click), time of movement, cursor speed, screen size in pixels, and x/y-coordinates for each mouse movement. The raw data collected are used to extract 25 mouse and 11 keyboard features per minute (Table I), representing the aggregated features with the highest predictive value [17]. In addition to the Naegelin *et al.* [17] metrics, two new metrics have been implemented: mean dwell time and SD dwell time to improve accuracy. The software prototype also monitors the applications and types of software being used by the office worker on a minute basis to extract the number of open windows without content analysis. Tracking and data collection is initiated after the user has launched the software and agreed to data collection upon every launch. To relate the aggregated mouse and keyboard data to the individual stress levels of office workers, respondents are asked to rate their current mood on a two-dimensional 5-point Likert scale several times a day via a software-integrated questionnaire using Self-Assessment Manikins (SAM) — non-verbal pictorial emotion manikins [31], [32]. The two dimensions describe valence and arousal, allowing any emotion to be described using the Circumplex model [29]. The Job-Related Well-Being Scale adapted this model to the work context [33].

B. Questionnaires

An initial questionnaire, based on existing validated WrQoL questionnaires, will be developed to assess baseline characteristics and self-reported WrQoL of office workers. Follow-up questionnaires will be administered throughout the study period to subjectively assess workload and emotional states for cross-checking with objective measurement data.

TABLE I.
EXAMPLES OF AGGREGATED DATA SOURCES OBTAINED BY MOUSE AND KEYBOARD INTERACTION

Mouse	Keyboard	Software/ Application
• Number of mouse movements per minute • Direct distance of a movement • Number of mouse pauses per minute • Mean/SD duration of a mouse pause • Mean/SD time between two clicks • Mean/SD Euclidean distance of a mouse movement • Mean/SD real distance of mouse movement • Mean/SD time duration of a mouse movement • Mean/SD average speed of a mouse movement per min • Mean/SD average angle of all angles in a movement • Mean/SD average distance of the real and straight line of a movement • Mean/SD sum of the difference between the real and straight line • Mean/SD number of direction changes in a mouse movement	• Number of pressed keys per minute • Error count • Mean/SD dwell time • Mean/SD digraph duration • Typing time • Mean/SD pause in typing per min • Keyboard pause count	• Average window count • Category of software (e.g., calendar, E-mail, presentation editor, text editor, messenger, programming environment etc.)

We evaluated four German validated self-report questionnaires and combined them into a comprehensive set of indicators of the base model and beyond, creating an initial and a shorter weekly questionnaire to be conducted in the MI-LQ app. The initial baseline questionnaire is designed to provide detailed information on participants' job satisfaction, autonomy dimensions, and work preferences. The shorter weekly check-in surveys via the MI-LQ app will be administered to capture changes over time on workload. The state of valence and arousal is asked using SAM at hourly intervals at the computer.

C. User interface

Our project's mobile phone app prototype (MI-LQ app) serves as both user interface and a data relay system to a cloud-based platform for offline analysis for WrQoL metrics. To enhance data collection, future iterations of the app will incorporate additional sensor categories, including work-related schedules and external influences. Ultimately, our aim is to provide participants with individualized WrQoL indicator scores and potential resources based on their analyzed data. The central aim of the application is to become a transparent point of data collection and data analysis presentation for the user, illustrated in Fig. 2.

D. Sensors and pulse data extraction and aggregation

After evaluation of various wearables from several manufacturers (Polar, Xiaomi, Apple, Samsung, Garmin, Fitbit), we decided to focus on Polar optical heart rate sensors: OH1 and Verity Sense. Those devices provide pulse-to-pulse intervals (PPI) extracted from photoplethysmography (PPG) signals. The Polar SDK for the iOS platform allows us to connect with a Polar device and perform live streaming measurements. The SDK provide pulse-to-pulse interval data in the following structure: 1) a PPI integer value which represents the interval between two pulses in milliseconds, 2) a heart rate (HR) value as calculated based on the PPI, 3) an error estimate integer value which represents an estimate of the expected absolute error of the PPI in milliseconds, 4) a block bit value which is set to 0 if PPI is considered valid, and otherwise set to 1, e.g., due to strong movement, and 5) a binary skin contact value which is 0 if there is no contact of the wearable to the skin detected, and 1 otherwise. Such additional information is important for a high-quality HRV analysis. Upon receipt from the sensor each PPI measurement is provided with a timestamp by the MI-LQ application and stored locally on the mobile phone in the following structure: Unix timestamp, PPI value, error estimate, block bit, skin contact bit. These data are accessible exclusively through the MI-LQ application, ensuring data security and privacy. The data is then processed and aggregated for analysis.

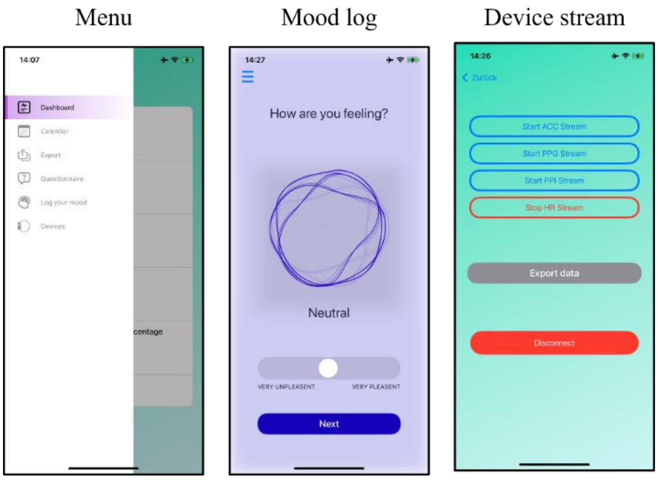

Fig. 2. Screenshots of the mobile app prototype

Basically, we can also integrate devices such as wearables into our application that can measure and provide health-related information. The workflow for those devices is slightly different:

1) Use device-specific application provided by manufacturers.

2) Synchronize health-related data from the manufactory's app into HealthKit (Apple, Cupertino, California, USA)

3) Grant permission for MI-LQ application to read data from HealthKit.

4) Prepare the received data for further aggregation and analysis.

E. Calculation of HRV metrics

HRV parameters have long been valued for their objective assessment of physical and mental status [34]. Within our preparative work we have seen more than 30 aggregated HRV-related indices based on Polar PPG sensor-derived PPI values. Although the manufacturers of so-called fitness trackers or pulse watches clearly have noticed the huge potential of the HRV framework for their consumer products, we rarely see reliable ready-to-use HRV-based measures, and some solutions applying a stress scale (from 0 to 100 percent) are presently not convincing: Calculations from different manufacturers are not transparent and cannot be verified; and each manufacturer uses its own scale, which is not comparable and not reproducible. Currently, we consider seven HRV-related indices (either statistical ones or combined ones according to Baevsky and Chernikova [34]) to be significant and meaningful as well as computationally effective:

1) Mean HR: mean of Heart Rate, i.e., average number of heart beats per minute

2) Mean NNI: mean of PPG-based PPIs

3) CV: Coefficient of Variation, i.e., mean of standard deviation of PPIs divided by the respective mean PPI (mean NNI)

4) RMSSD: Root Mean Square of Successive Differences of PPIs

5) SI: Stress Index according to Baevsky and Chernikova [34], i.e., ratio of amplitude of the modal value in the PPI histogram and the doubled product of modal value itself weighted by the min-max-span of PPIs

6) IVR: Index of Vegetative Regime is characterizing the ratio between sympathetic and parasympathetic influences on the heart rhythm, i.e., amplitude of the modal value in the PPI histogram divided by the standard deviation of PPIs

7) PAPR: Parameter of Appropriateness of Processes of heart Regulation, i.e., amplitude of the modal value in the PPI histogram divided by the respective modal value

These HRV parameters provide information about the autonomic nervous system during working hours and they allow to assess the effects of physiological responses to work stress.

F. AI-driven analysis of speech for emotion classification

In future work, we aim to develop a proprietary dataset for valence and arousal speech analysis based on the SAM scale [31], [32]. One of the objectives is to minimize aleatoric uncertainty, and therefore, we aim to achieve high entropy between the distinct classes. The available datasets do not meet our study's specific requirements, necessitating the decision to collect our own data. They lack German language compatibility, essential annotations for valence, and SAM, and access to datasets solely annotated with valence and arousal ratings has been denied despite requests. Moreover, the accessible dataset contains machine-generated annotations and an insufficient sample size (less than 200). Therefore, to ensure the completeness and accuracy of our research, independent data collection is imperative. Questions are asked under artificial induction of emotions according to Almazrouei *et al.* [35].

Once the dataset is created, we will develop separate classification models for valence and arousal respectively. These models will leverage prosodic features extracted from the speech data, such as pitch, shimmer, jitter (which play a crucial role when it comes to predicting stress out of the voice [36]), and Mel-Frequency Cepstral Coefficients (MFCC). According to the findings of Li *et al.* [36] these features play a critical role in the accurate classification of emotional states. For example, does a higher pitch directly correlate with emotions like anger.

Our approach will use these prosodic features to make predictions about the SAM ratings, providing a detailed analysis of how these features correlate with subjective emotional assessments. We will implement cross-validation techniques to evaluate the performance of the valence and arousal speech models, ensuring that they are rigorously tested and validated against diverse speech data collected through a website set up for this purpose.

This future work aims to contribute to the field of affective computing by providing a robust dataset and a validated methodology for speech-based stress classification, which could have wide-ranging applications in areas such as human-computer interaction and mental health monitoring.

G. Machine learning concept

Being time series data, PPI data as well as mouse and keyboard dynamics data allow tracking changes over time. Therefore, we use Long Short-Term Memory (LSTM) models based on the Keras framework to predict hourly valence and arousal class values, respectively. LSTM is a deep learning, sequential neural network that can learn hidden patterns in temporal sequences and retain information from previous time points [37]. The training data includes minute-by-minute data with up to 60 samples of HRV, mouse and keyboard features, along with information on defined active windows per target. These models are individually trained to generalize across different individuals, considering variations in physiological responses and interaction patterns, and are evaluated using classification report metrics to predict valence and arousal classes while capturing human emotional states over time.

III. Next Steps

Our research develops instruments to objectively assess workload through valence and arousal using smartphones and wearables in tandem with a software prototype for keyboard and mouse tracking. In the developmental stage of the technology, questionnaire responses from study participants are used to annotate and check data but will later be phased out when validity is reached.

Next steps in the MI-LQ project cover the establishment and implementation of the AI-driven analysis of speech-based emotion classification, an additional and supporting approach to detect and annotate stress-related strain at work to physiological and computer interaction-derived data. In addition, data derived from outlook, calendar, and project management system will be integrated into the MI-LQ app. Following this, the pilot study we aim to conduct will assess the reliability, validity, and feasibility of the digital framework for objective WrQoL assessment, involving office workers for a minimum of four weeks. This study will also enable the collection of work-related data within an authentic office environment, facilitating the assembly of a dataset sufficient for machine learning annotation.

IV. Limitations

While this study provides valuable insights into WrQoL and new approaches on how to measure it more objectively, several limitations should be acknowledged. First, due to the limited standards for wearable electronics, MI-LQ currently uses well-defined, brand-specific digital devices with proprietary hardware and software interfaces such as Polar optical sensors combined with iOS platform. Currently, there is no easy way to broadly integrate fitness trackers and smartwatches to reliably quantify QoL, as they lack comprehensive monitoring capabilities due to their lifestyle focus, limiting the number of potential users. HRV-related indices can only be calculated using PPI data, but most wearables only provide heart rate data and not PPI. This is not the case with Polar's optical sensors, which offer reliable measurement and data quality, and the great advantage of live streaming measurements when combined with the Polar iOS-SDK. As MI-LQ evolves, the integration of a wider range of digital devices and services will be considered, as well as the expansion to the Android platform. Second, in the initial phase, the MI-LQ project focused on assessing WrQoL specifically within the office context, recognizing the necessity to start with a defined population subset. Importantly, the modular design of the base model enables the seamless incorporation of new dimensions, indicators, and technologies as the study progresses. However, it's crucial to acknowledge the limitation of generalizability beyond the office context, necessitating further validation and demonstration of transferability in subsequent projects. This highlights the project's iterative nature and the ongoing refinement required to extend its applicability to diverse occupational settings such as construction workers, laboratory technicians, and beyond. Third, the data annotation process and subsequently the applied ML model depends on the provision of numerous reliable and continuous individual physiological and computational data and is therefore susceptible to patient dropouts. One countermeasure to avoid critical attrition rates that will be implemented in the pilot is incentives in the form of gift cards or expense reimbursements. Fourth, Likert scales, which will be part of the initial and weekly questionnaires need to be implemented carefully in terms of the number of items and a neutral position. Survey precision requires a careful balance; too few items risk imprecision, while an excess can hinder responses. Pre-test observations revealed a tendency to avoid extreme positions on the Likert scale.

V. Conclusion

WrQoL has emerged as a powerful indicator for industry and research into workplace conditions and employee well-being, revealing areas for improvement as well as levels of employee stress and burnout. As realized in the innovative MI-LQ approach, leveraging and linking the large amounts of data generated by wearable biosensors and computer interaction offers the opportunity to make WrQoL objectively measurable. In a base model, important physiological and WrQoL indicators have been considered and linked to measurement variables and instruments, set up in a modular way, with the possibility to expand the base model to more indicators and variables. Brand-specific digital devices and applications, ML and a mobile app architecture were found to be suitable tools as a basic framework for data input, output, processing, and prediction aimed at WrQoL assessment. Two wearable biosensors with satisfactory data granularity and quality were identified to provide the PPI needed to calculate the identified seven meaningful and computationally effective HRV indices. HRV data combined with mouse and keyboard dynamics and mood tracking data will be used to train LSTM models to predict particularly stressful working periods and emotional states. Opening to a greater variety of digital devices within the current model as well as expansion to other indicators or dimensions of WrQoL and/or QoL will be considered as MI-LQ evolves.

References

[1] J. H. Greenhaus, K. M. Collins, and J. D. Shaw, "The relation between work–family balance and quality of life," *JOURNAL OF VOCATIONAL BEHAVIOR*, vol. 63, no. 3, pp. 510–531, 2003, doi: 10.1016/S0001-8791(02)00042-8.

[2] J. Goh, J. Pfeffer, and S. A. Zenios, "Workplace Stressors & Health Outcomes: Health Policy for the Workplace," *Behavioral Science & Policy*, no. 1, pp. 43–52, 2015.

[3] WHOQOL Group, "Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group," *Psychological medicine*, vol. 28, no. 3, pp. 551–558, 1998, doi: 10.1017/s0033291798006667.

[4] Eurostat, "Final report of the expert group on quality of life indicators," 2017. Accessed: Mar. 1 2024. [Online]. Available: https://ec.europa.eu/eurostat/documents/7870049/7960327/KS-FT-17-004-EN-N.pdf/f29171db-e1a9-4af6-9e96-730e7e11e02f

[5] M. Durand, "The OECD Better Life Initiative: How's Life? and the Measurement of Well-Being," *Review of Income and Wealth*, vol. 61, no. 1, pp. 4–17, 2015, doi: 10.1111/roiw.12156.

[6] C. Mendes and H. Pereira, "Assessing the Impact of COVID-19 on Work-Related Quality of Life through the Lens of Sexual Orientation," *Behavioral sciences (Basel, Switzerland)*, vol. 11, no. 5, 2021, doi: 10.3390/bs11050058.

[7] A. Jaiswal and A. Mahila, "Quality of work life," *Journal of Business Management & Social Sciences Research*, 02.2014, pp. 83–87, 2014. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2cfb960da043838669f9f75bda049c193e60e3a8

[8] J. Leitão, D. Pereira, and Â. Gonçalves, "Quality of Work Life and Contribution to Productivity: Assessing the Moderator Effects of Burnout Syndrome," *Int. J. Environ. Res. Public Health*, vol. 18, no. 5, 2021, doi: 10.3390/ijerph18052425.

[9] S. Wang, D. Kamerade, B. Burchell, A. Coutts, and S. U. Balderson, "What matters more for employees' mental health: Job quality or job quantity?," *Camb. J. Econ.*, vol. 46, no. 2, pp. 251–274, 2022, doi: 10.1093/cje/beab054.

[10] D. J. Horst, E. E. Broday, R. Bondarick, L. F. Serpe, and L. A. Pilatti, "Quality of Working Life and Productivity: An Overview of the Conceptual Framework," *International Journal of Managerial Studies and Research*, 07.2014, pp. 87–98, 2014.

[11] C. Seale, Ed., *Researching society and culture*, 3rd ed. Los Angeles, Calif.: Sage, 2011.

[12] M. L. Patten, *Questionnaire Research: A Practical Guide*, 4th ed. Los Angeles: Taylor and Francis, 2014. [Online]. Available: http://gbv.eblib.com/patron/FullRecord.aspx?p=4710457

[13] A. Alberdi, A. Aztiria, and A. Basarab, "Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review," *Journal of biomedical informatics*, vol. 59, pp. 49–75, 2016, doi: 10.1016/j.jbi.2015.11.007.

[14] N. Nappo, "Job stress and interpersonal relationships cross country evidence from the EU15: a correlation

analysis," *BMC PUBLIC HEALTH*, vol. 20, no. 1, p. 1143, 2020, doi: 10.1186/s12889-020-09253-9.

[15] T. Androutsou, S. Angelopoulos, E. Hristoforou, G. K. Matsopoulos, and D. D. Koutsouris, "Automated Multimodal Stress Detection in Computer Office Workspace," *Electronics*, vol. 12, no. 11, p. 2528, 2023, doi: 10.3390/electronics12112528.

[16] E. Sükei, A. Norbury, M. M. Perez-Rodriguez, P. M. Olmos, and A. Artés, "Predicting Emotional States Using Behavioral Markers Derived From Passively Sensed Data: Data-Driven Machine Learning Approach," *JMIR mHealth and uHealth*, vol. 9, no. 3, e24465, 2021, doi: 10.2196/24465.

[17] M. Naegelin *et al.*, "An interpretable machine learning approach to multimodal stress detection in a simulated office environment," *Journal of biomedical informatics*, vol. 139, p. 104299, 2023, doi: 10.1016/j.jbi.2023.104299.

[18] V. R. Shinde, K. R. Sonawane, A. D. Bhawar, and H. D. Walam, Eds., *Real time-Employee Emotion Detection system (RTEED) using Machine Learning*. Zenodo, 2023.

[19] S. Majumder, T. Mondal, and M. J. Deen, "Wearable Sensors for Remote Health Monitoring," *Sensors (Basel, Switzerland)*, vol. 17, no. 1, 2017, doi: 10.3390/s17010130.

[20] F. Shaffer and J. P. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," *Front. Public Health*, vol. 5, p. 258, 2017, doi: 10.3389/fpubh.2017.00258.

[21] G. Ernst, "Hidden Signals-The History and Methods of Heart Rate Variability," *Front. Public Health*, vol. 5, p. 265, 2017, doi: 10.3389/fpubh.2017.00265.

[22] S. Saganowski, "Bringing Emotion Recognition Out of the Lab into Real Life: Recent Advances in Sensors and Machine Learning," *Electronics*, vol. 11, no. 3, p. 496, 2022, doi: 10.3390/electronics11030496.

[23] M. N. Burns *et al.*, "Harnessing context sensing to develop a mobile intervention for depression," *Journal of medical Internet research*, vol. 13, no. 3, e55, 2011, doi: 10.2196/jmir.1838.

[24] A. Hart, D. Reis, E. Prestele, and N. C. Jacobson, "Using Smartphone Sensor Paradata and Personalized Machine Learning Models to Infer Participants' Well-being: Ecological Momentary Assessment," *Journal of medical Internet research*, vol. 24, no. 4, e34015, 2022, doi: 10.2196/34015.

[25] A. Pawlicka, M. Pawlicki, R. Tomaszewska, M. Choraś, and R. Gerlach, "Innovative machine learning approach and evaluation campaign for predicting the subjective feeling of work-life balance among employees," *PloS one*, vol. 15, no. 5, e0232771, 2020, doi: 10.1371/journal.pone.0232771.

[26] N. S. Gamage and D. P. Asanka, "Machine Learning Approach to Predict Mental Distress of IT Workforce in Remote Working Environments," in *2022 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Colombo, Sri Lanka, 2022, pp. 211–216.

[27] S. Bromuri, A. P. Henkel, D. Iren, and V. Urovi, "Using AI to predict service agent stress from emotion patterns in service interactions," *J. Serv. Manage.*, vol. 32, no. 4, pp. 581–611, 2021, doi: 10.1108/JOSM-06-2019-0163.

[28] A. Baird *et al.*, "An Evaluation of Speech-Based Recognition of Emotional and Physiological Markers of Stress," *Front. Comput. Sci.*, vol. 3, 2021, doi: 10.3389/fcomp.2021.750284.

[29] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, pp. 1161–1178, 1980, doi: 10.1037/h0077714.

[30] L. Pepa, A. Sabatelli, L. Ciabattoni, A. Monteriu, F. Lamberti, and L. Morra, "Stress Detection in Computer Users From Keyboard and Mouse Dynamics," *IEEE Trans. Consumer Electron.*, vol. 67, no. 1, pp. 12–19, 2021, doi: 10.1109/TCE.2020.3045228.

[31] N. Banholzer, S. Feuerriegel, E. Fleisch, G. F. Bauer, and T. Kowatsch, "Computer Mouse Movements as an Indicator of Work Stress: Longitudinal Observational Field Study," *Journal of medical Internet research*, vol. 23, no. 4, e27121, 2021, doi: 10.2196/27121.

[32] M. M. Bradley and P. J. Lang, "Measuring emotion: the Self-Assessment Manikin and the Semantic Differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994, doi: 10.1016/0005-7916(94)90063-9.

[33] P. T. van Katwyk, S. Fox, P. E. Spector, and E. K. Kelloway, "Using the Job-Related Affective Well-Being Scale (JAWS) to investigate affective responses to work stressors," *J. Occup. Health Psychol.*, 1939-1307(Electronic),1076-8998(Print), pp. 219–230, 2000, doi: 10.1037/1076-8998.5.2.219.

[34] R. M. Baevsky and A. G. Chernikova, "Heart rate variability analysis: physiological foundations and main methods," *Cardiometry*, no. 10, pp. 66–76, 2017, doi: 10.12710/cardiometry.2017.10.6676.

[35] M. A. Almazrouei, R. M. Morgan, and I. E. Dror, "A method to induce stress in human subjects in online research environments," *BEHAVIOR RESEARCH METHODS*, vol. 55, no. 5, pp. 2575–2582, 2023, doi: 10.3758/s13428-022-01915-3.

[36] X. Li *et al.*, "Stress and Emotion Classification using Jitter and Shimmer Features," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing: ICASSP 2007] ; Honolulu, HI, 16 [i.e. 15] - 20 April 2007*, Honolulu, HI, 2007, IV-1081-IV-1084.

[37] D. M. Ahmed, M. M. Hassan, and R. J. Mstafa, "A Review on Deep Sequential Models for Forecasting Time Series Data," *Applied Computational Intelligence and Soft Computing*, vol. 2022, pp. 1–19, 2022, doi: 10.1155/2022/6596397.