Objective Measurement of Quality of Life: First Results of the Project "Machine Intelligence for the Objective Determination of Individual Quality of Life" - MI-LQ

OP20

S. Mareike Geisler[1], Celine Schreiber[1,2], Hamlet Kosakyan[3], Jakob Hohn[4], Jenny Voigt[4], Alisa Hamm[1], Tobias Schäffer[2], Ekaterina Mut[5], Franziska Claus[1], Axel Krinner[1], Dennis Häckl[2], Franziska Stutzer[1], Christian Hrach[5], Weronika Weber[2], Bogdan Franczyk[5,6], Hubert Österle[7], Ulf-Dietrich Braumann[5], Carsta Militzer-Horstmann[1,2]

Affiliations:
[1] Scientific Institute for Health Economics and Health System Research (WIG2 GmbH), Leipzig, Germany
[2] University of Leipzig, Health Economics and Management, Leipzig, Germany
[3] Appsfactory GmbH, Leipzig, Germany
[4] 4K ANALYTICS GmbH, Leipzig, Germany
[5] Institute for Applied Informatics (InfAI e.V.), Leipzig, Germany
[6] University of Leipzig, Information Systems, Leipzig, Germany
[7] University of St. Gallen, Business Engineering (Professor Emeritus), St. Gallen, Switzerland

Background

In an era of blurred work-life boundaries, understanding and enhancing job satisfaction, leading to a higher Work-Related Quality of Life (WrQoL), is crucial for productive and healthy workplaces. Our research employs multimodal data collection and analysis for an objective job satisfaction assessment, moving beyond conventional subjective methods. By proposing a model for identifying and measuring WrQoL indicators objectively, we aim to circumvent respondent bias and survey fatigue, offering a nuanced understanding to enrich scientific insights and practical applications in workplace optimization. The development of objective and automated WrQoL measurement methods combined with machine learning (ML) for prediction of a WrQoL is still in its infancy but represents an innovative approach due to the large amount of data generated by digital services, applications, and sensors.

Objectives

Our research focuses on developing, refining, and validating a model to identify critical factors influencing WrQoL. Centering on life-quality indicators within office and living environments, we aim to create a tool for WrQoL's objective measurement. This initiative, leveraging ML for data analysis, facilitates an empirical study to broaden the model's scope. Our goal is to enrich scientific insights into WrQoL and share findings across sectors, enhancing workplace optimization for improved employee well-being and organizational efficacy, condensed into a purpose-driven endeavor for a broader understanding and application.

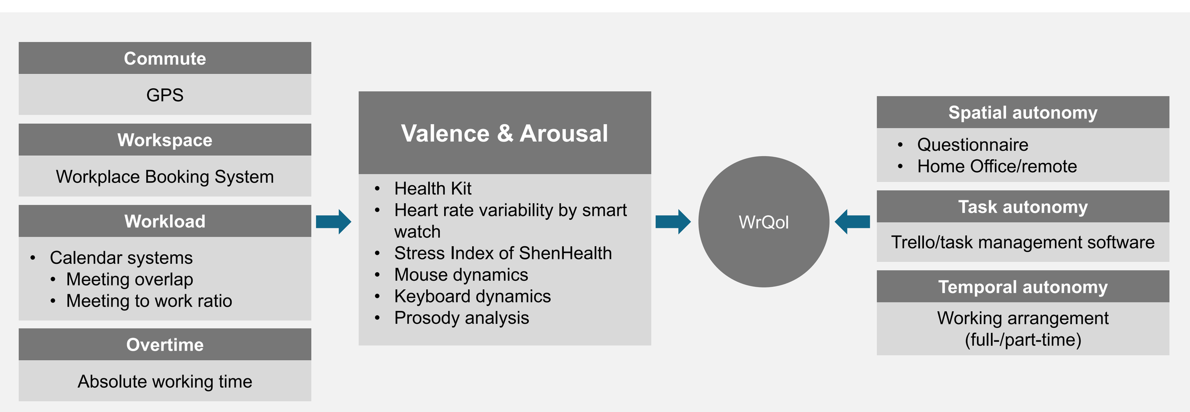

Fig. 1: Base model of WrQoL indicators and their objectifiability with different digital approaches

Related research

In the pursuit of understanding the dynamics of **valence and arousal** within office environments, recent studies have shed light on various methodologies and approaches. Our research extends beyond the scope of three notable contributions, each offering insights into the interplay of factors influencing stress levels in office settings.

- **Banholzer et al. (2021)** conducted a field study in **real office environments** focusing solely on the usage of PC mice. Their investigation relied on correlation analysis to uncover **associations between mouse usage patterns and perceived stress levels**. Over a period of seven weeks, data was collected from 70 participants, monitored twice daily for 30 minutes each session. Through self-reported measures of stress, valence, and arousal, Banholzer et al. identified a discernible tradeoff between mouse speed and accuracy, implicating a connection with stress perception.
- Contrastingly, **Naegelin et al. (2023)** opted for a **simulated office setting**, employing AI models trained on keyboard and mouse dynamics data. Their study encompassed tasks such as transcribing handwritten reports, aggregating sales figures, and scheduling appointments. By comparing AI models utilizing keyboard/mouse dynamics with those incorporating **ECG-based HRV data**, Naegelin et al. concluded that the former exhibited superior predictive capabilities for stress, valence, and arousal. This finding underscores the significance of contextualized data in stress inference models.
- Similarly, **Androutsou et al. (2023)** pursued a comprehensive investigation integrating physiological markers and behavioral indicators. Employing **photoplethysmography (PPG)-based HRV data** alongside galvanic skin response (GSR) measurements, their study encompassed a diverse range of office tasks including text transcription, neuropsychological assessments, and stress induction exercises. By leveraging ML techniques, including artificial neural networks (ANN), Androutsou et al. demonstrated the **efficacy of combined physiological and behavioral data in stress inference**. Their findings suggest that a multi-modal approach outperforms reliance solely on keyboard/mouse dynamics, particularly in discerning stress levels.

These recent works provide critical insights into the complexities of stress assessment within office environments, emphasizing the importance of contextual factors and diverse data modalities. Further exploration of these methodologies promises to advance our understanding of workplace stress and inform targeted interventions for improved well-being.

For a comprehensive list of references, please refer to the supplementary material.

Methods

Systematic literature search (1)

A systematic literature search, still in progress, was initiated to identify indicators of WrQoL using SPICE framework to define the research question and scope. The PubMed, Web of Science, APA PsychNet and Scopus databases were used, yielding a total of 9,242 publications. After removing duplicates, 7,605 publications remained for title-abstract screening by 4 independent reviewers. After full-text screening, approximately 100 publications remained for data extraction.

Base model (2)

Based on indicators identified by the systematic literature search, a base model with objectively measurable WrQoL indicators was developed, including workload, work space, commute, overtime, autonomy, and valence and arousal (Fig. 1). Developed as a starting point for the subsequent implementation of the ML approach, the base model will be extended as the study progresses.

Data sources and data integration (3)

Computers, computer cameras, smartphones, wearables, microphones, keyboards, and mice will serve as objective measurement instruments. The use of these devices provide individual work-related data relevant for the defined WrQoL indicators, such as workload, appointment density, project management system assignments, mouse movements and keyboard strokes, prosody and heart rate variability (HRV). Additionally, for the assessment of baseline characteristics and subjective WrQoL, an initial questionnaire will be developed using existing validated questionnaires. A two-item measurement tool will be used for follow-up assessment of the persons' valence and arousal state. The subjective data will be used to annotate the sensoric data and serve as label in development of the classifier. An overview of data types, collection frequency, and data sources is given in Fig. 3.

Machine learning approach (4)

The collected user-specific data will be analyzed regarding the level of effect (positive, negative, neutral) of indicators on WrQoL and then used for data prediction using appropriate ML methods, e.g. classifiers based on ensemble learning with decision trees, or transformer models for prediction of the emotional state and SHapley Additive exPlanations (SHAP) analysis for the effect of the indicator. The focus is on design of the ML architecture, particularly data processing, to develop a model that can cope with small amounts of individual multimodal data, and on evaluation of reliability and relevance of the different data sources for model performance.

Development of an app prototype (5)

A mobile app prototype will be developed that serves as an input and output data platform. Based on the ML based model, the WrQoL of the mobile app user will be objectively calculated using the available data (Fig. 2).

Study with office workers (6)

To put theory into practice, office workers will be recruited and asked to provide work-related data by several digital tools and subjective questionnaires over a period of at least 4 weeks. This allows for a quality check of the selected WrQoL indicators and the creation of a dataset sufficient for ML prediction.

Preliminary results

Mobiles and Wearables for Heart Rate Variability Analysis

In the realm of health monitoring, the integration of mobile phones, wearable devices, and specialized trackers presents a diverse array of options for capturing biometric variables.

In our search for suitable wearables, we found that smartwatches, though popular, often lack comprehensive monitoring capabilities due to their emphasis on lifestyle features. Fitness trackers, specialized for physiological parameter tracking, were found to be more suitable for comprehensive monitoring, especially when equipped with PPG functionality and software development kit (SDK) support. Polar fitness trackers, identified through systematic search, deliver reliable measurements and systematic data quality and will be used to collect biometric data in our study with office workers. HRV parameters are valued for their objective assessment of both physical and mental states. From the abundance of HRV parameters, we selected single metrics (mean HR, mean NNI, RMSSD, CV) and combined metrics (PAPR, IVR, stress index) as appropriate HRV measures. PPG offers an efficient means of recording pulsometric data, facilitating HRV parameter extraction. A thorough analysis of HRV measures is essential for capturing work-related effects on valence and arousal throughout the office working day. By leveraging the capabilities of these devices, alongside thorough data analysis, we can assess physiological parameters such as HRV, that contribute to understanding both physical and mental states in various contexts, including the demands of the modern office environment.

Legend: planned, in progress, finished

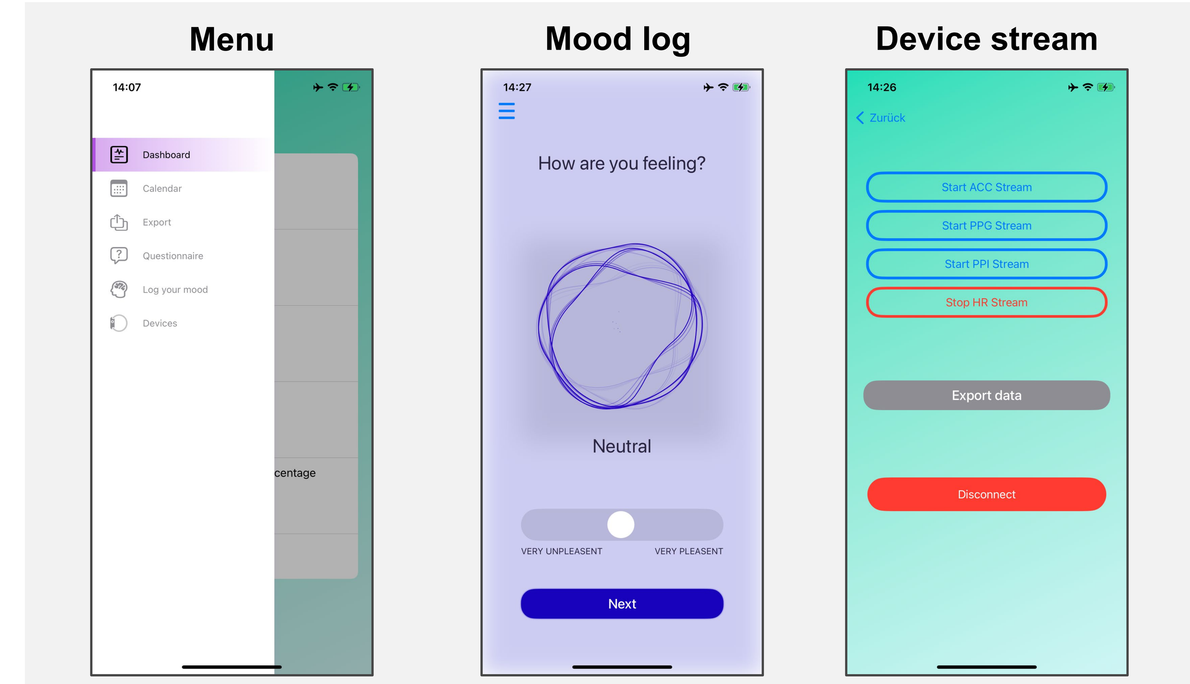

Fig. 2: Screenshots of the mobile app prototype

PC Software for Keyboard and Mouse Tracking

- Development of a software prototype designed to continuously monitor mouse and keyboard dynamics, as well as software type being used (e.g., browser, text editor, email, PowerPoint) to assess work-related valence and arousal
- Initiation of tracking and data collection after the user has launched the software and agreed to data collection upon every launch
- Collection of aggregated features related to emotion recognition, excluding the content of typing. This includes metrics such as the number of mouse movements, duration of pauses, mean dwell time, and number of pressed keys per minute, aimed at capturing the intensity and fluctuation of interaction in the office context
- Assessment of valence and arousal levels multiple times throughout the day through self-report via an integrated questionnaire, employing non-verbal pictorial emotion manikins based on the methodologies outlined by Banholzer et al. (2021) and Bradley and Lang (1994)
- Analysis of emotional data in the workplace context utilizing JAWS (Van Katwyk et al., 2000)

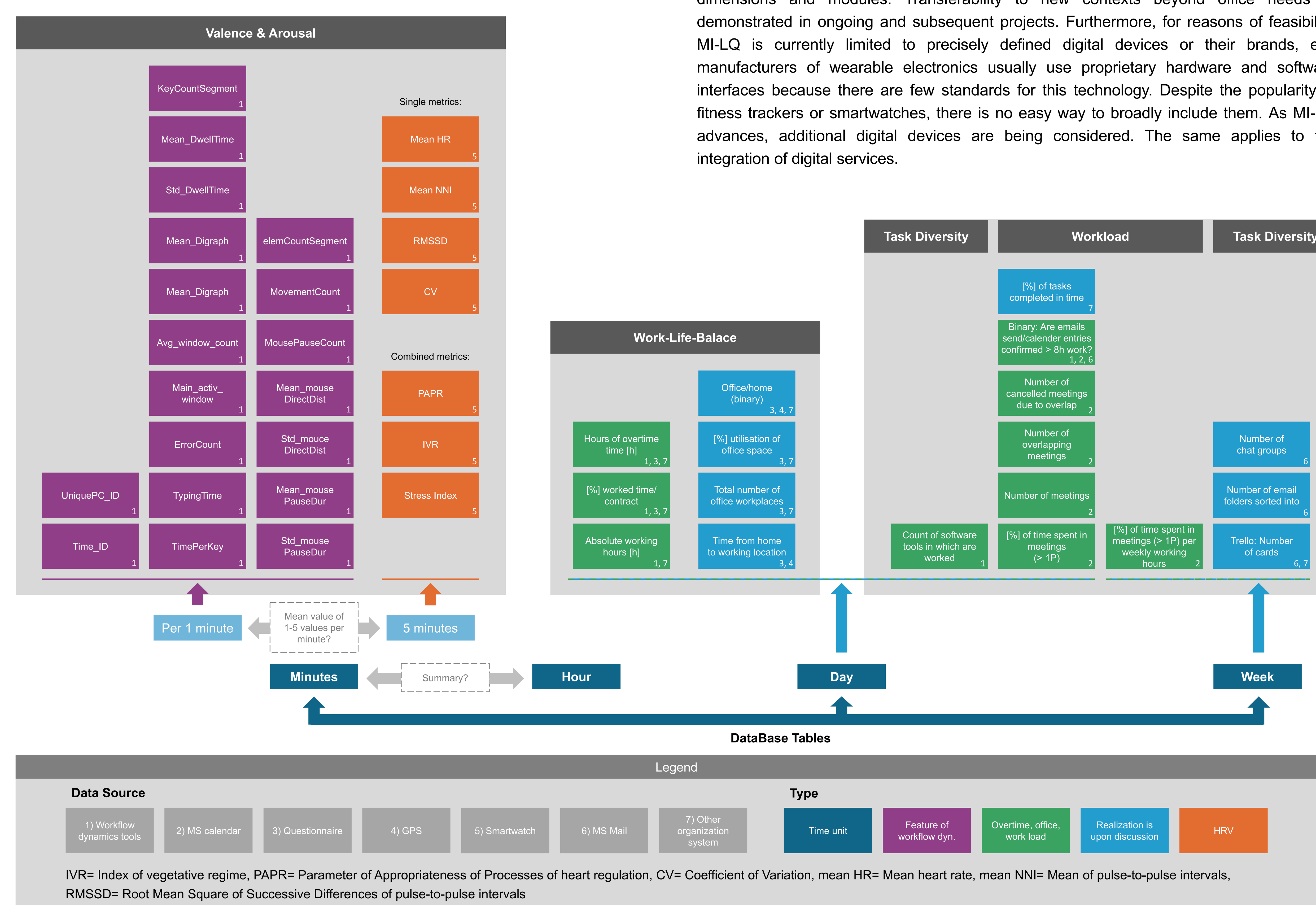

Fig. 3: Overview of data collection type and time point, datasets and data source

Mobile phone app prototype

Our project's mobile phone app prototype serves as both a user interface and a data relay system to a cloud-based platform for offline analysis of WrQoL metrics. Initially, participants will engage in regular check-ins (currently experimented with at hourly intervals), which will gradually be phased out as our model gains accuracy, mitigating participant fatigue while ensuring robust data collection. Participants will be presented with an initial baseline questionnaire and regular valence and arousal check-ins through the app interface (Fig. 2).

To enhance data collection, future iterations of the app will incorporate additional sensor categories, including work-related schedules and external influences. Additionally, both online and onboard computations of WrQoL measures will be integrated into the mobile app.

Ultimately, our aim is to provide participants with individualized WrQoL indicators and potential resources based on their data analysis.

Conclusions

The systematic literature search identified relevant indicators for the assessment of WrQoL, which were translated into a set of objectively measurable indicators in a base model. Different digital devices in combination with digital applications were shown to be suitable for measuring work-related indicators. The ML and mobile app architecture as well as a cloud system were established as a basic framework for data input, output and processing for the objective determination of WrQoL. This project provides an innovative approach to objectively measure WrQoL in the digital world. By understanding the relationship between digital service use, WrQoL, and mental health, the project may provide preliminary insights into strategies for preventing stress and burnout in the workplace.

Limitations

The MI-LQ project aims to objectively quantify individual quality of life (QoL) using digital devices and applications combined with a ML approach. However, this study focused mainly on WrQoL in the office context as a starting point for overall QoL quantification. The established models and approaches were built in a modular way to easily add new dimensions and modules. Transferability to new contexts beyond office needs to demonstrated in ongoing and subsequent projects. Furthermore, for reasons of feasibility, MI-LQ is currently limited to precisely defined digital devices or their brands, e.g. manufacturers of wearable electronics usually use proprietary hardware and software interfaces because there are few standards for this technology. Despite the popularity of fitness trackers or smartwatches, there is no easy way to broadly include them. As MI-LQ advances, additional digital devices are being considered. The same applies to the integration of digital services.

Disclosure:
This study was funded by Co-funded by the European Union. This project is co-financed from tax revenues on the basis of the budget adopted by the Saxon State Parliament.
Conflicts of interest:
All authors were fully independent in the study design, analysis, interpretation of results, and writing of the abstract/poster.

Objective Measurement of Quality of Life: First Results of the Project "Machine Intelligence for the Objective Determination of Individual Quality of Life" - MI-LQ

OP20

S. Mareike Geisler[1], Celine Schreiber[1,2], Hamlet Kosakyan[3], Jakob Hohn[4], Jenny Voigt[4], Alisa Hamm[1], Tobias Schäffer[2], Ekaterina Mut[5], Franziska Claus[1], Axel Krinner[1], Dennis Häckl[2], Franziska Stutzer[1], Christian Hrach[5], Weronika Weber[2], Bogdan Franczyk[5,6], Hubert Österle[7], Ulf-Dietrich Braumann[5], Carsta Militzer-Horstmann[1,2]

Affiliations:
[1] Scientific Institute for Health Economics and Health System Research (WIG2 GmbH), Leipzig, Germany; [2] University of Leipzig, Health Economics and Management, Leipzig, Germany; [3] Appsfactory GmbH, Leipzig, Germany; [4] 4K ANALYTICS GmbH, Leipzig, Germany; [5] Institute for Applied Informatics (InfAI e.V.), Leipzig, Germany; [6] University of Leipzig, Information Systems, Leipzig, Germany; [7] University of St. Gallen, Business Engineering (Professor Emeritus), St. Gallen, Switzerland

Background

In an era of blurred work-life boundaries, understanding and enhancing job satisfaction, leading to a higher Work-Related Quality of Life (WrQoL), is crucial for productive and healthy workplaces. Our research employs multimodal data collection and analysis for an objective job satisfaction assessment, moving beyond conventional subjective methods. By proposing a model for identifying and measuring WrQoL indicators objectively, we aim to circumvent respondent bias and survey fatigue, offering a nuanced understanding to enrich scientific insights and practical applications in workplace optimization. The development of objective and automated WrQoL measurement methods combined with machine learning (ML) for prediction of a WrQoL is still in its infancy but represents an innovative approach due to the large amount of data generated by digital services, applications, and sensors.

Objectives

Our research focuses on developing, refining, and validating a model to identify critical factors influencing WrQoL. Centering on life-quality indicators within office and living environments, we aim to create a tool for WrQoL's objective measurement. This initiative, leveraging ML for data analysis, facilitates an empirical study to broaden the model's scope. Our goal is to enrich scientific insights into WrQoL and share findings across sectors, enhancing workplace optimization for improved employee well-being and organizational efficacy, condensed into a purpose-driven endeavor for a broader understanding and application.

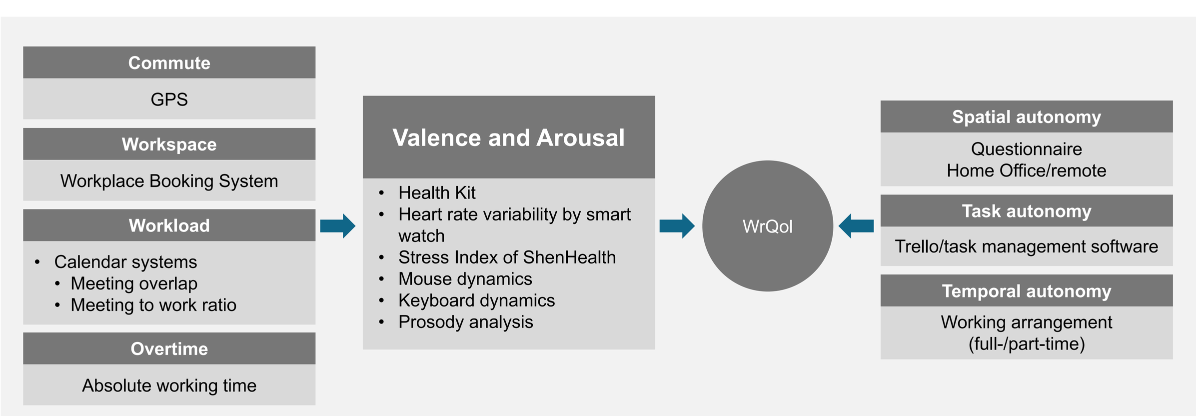

Fig. 1: Base model of WrQoL indicators and their objectifiability with different digital approaches

Related research

In the pursuit of understanding the dynamics of **valence and arousal** within office environments, recent studies have shed light on various methodologies and approaches. Our research extends beyond the scope of three notable contributions, each offering insights into the interplay of factors influencing stress levels in office settings.

- **Banholzer et al. (2021)** conducted a field study in **real office environments** focusing solely on the usage of PC mice. Their investigation relied on correlation analysis to uncover **associations between mouse usage patterns and perceived stress levels**. Over a period of seven weeks, data was collected from 70 participants, monitored twice daily for 30 minutes each session. Through self-reported measures of stress, valence, and arousal, Banholzer et al. identified a discernible tradeoff between mouse speed and accuracy, implicating a connection with stress perception.
- Contrastingly, **Naegelin et al. (2023)** opted for a **simulated office setting**, employing AI models trained on keyboard and mouse dynamics data. Their study encompassed tasks such as transcribing handwritten reports, aggregating sales figures, and scheduling appointments. By comparing AI models utilizing keyboard/mouse dynamics with those incorporating **ECG-based HRV data**, Naegelin et al. concluded that the former exhibited superior predictive capabilities for stress, valence, and arousal. This finding underscores the significance of contextualized data in stress inference models.
- Similarly, **Androutsou et al. (2023)** pursued a comprehensive investigation integrating physiological markers and behavioral indicators. Employing **photoplethysmography (PPG)-based HRV data** alongside galvanic skin response (GSR) measurements, their study encompassed a diverse range of office tasks including text transcription, neuropsychological assessments, and stress induction exercises. By leveraging ML techniques, including artificial neural networks (ANN), Androutsou et al. demonstrated the **efficacy of combined physiological and behavioral data in stress inference**. Their findings suggest that a multi-modal approach outperforms reliance solely on keyboard/mouse dynamics, particularly in discerning stress levels.

These recent works provide critical insights into the complexities of stress assessment within office environments, emphasizing the importance of contextual factors and diverse data modalities. Further exploration of these methodologies promises to advance our understanding of workplace stress and inform targeted interventions for improved well-being.

For a comprehensive list of references, please refer to the supplementary material.

Methods

Systematic literature search

1

A systematic literature search, still in progress, was initiated to identify indicators of WrQoL using SPICE framework to define the research question and scope. The PubMed, Web of Science, APA PsychNet and Scopus databases were used, yielding a total of 9,242 publications. After removing duplicates, 7,605 publications remained for title-abstract screening by 4 independent reviewers. After full-text screening, approximately 100 publications remained for data extraction.

Base model

2

Based on indicators identified by the systematic literature search, a base model with objectively measurable WrQoL indicators was developed, containing workload, work space, commute, overtime, autonomy, and valence and arousal (Fig.1). Developed as a starting point for the subsequent implementation of the ML approach, the base model will be extended as the study progresses.

Data sources and data integration

3

Computers, computer cameras, smartphones, wearables, microphones, keyboards, and mice will serve as objective measurement instruments. The use of these devices provide individual work-related data relevant for the defined WrQoL indicators, such as workload, appointment density, project management system assignments, mouse movements and keyboard strokes, prosody and heart rate variability (HRV). Additionally, for the assessment of baseline characteristics and subjective WrQoL, an initial questionnaire will be developed using existing validated questionnaires. A two-item measurement tool will be used for follow-up assessment of the persons' valence an[illegible] state[illegible]ject[illegible] will be us[illegible] to annotate the sensoric data and s[illegible] as labe[illegible] in [illegible]elopment [illegible] the c[illegible]fier. [illegible] overview of data types, collection freq[illegible]cy, and data so[illegible] is given [illegible] Fig. 3.

Machine learning approach

The collected user-specific data will be[illegible]rega[illegible]leve[illegible](positi[illegible] negative, neutral) of indicators on WrQoL and then used for data prediction using appropriate ML methods, e.g. classifiers based on ensemble learning with decision trees, or transformer models for prediction of the emotional state and SHapley Additive exPlanations (SHAP) analysis for the effect of the indicator. The focus is on design of the ML architecture, particularly data processing, to develop a model that can cope with small amounts of individual multimodal data, and on evaluation of reliability and relevance of the different data sources for model performance.

Development of an app prototype

5

A mobile app prototype will be developed that serves as an input and output data platform. Based on the ML based model, the WrQoL of the mobile app user will be objectively calculated using the available data (Fig.2).

Study with office workers

6

To put theory into practice, office workers will be recruited and asked to provide work-related data by several digital tools and subjective questionnaires over a period of at least 4 weeks. This allows for a quality check of the selected WrQoL indicators and the creation of a dataset sufficient for ML prediction.

Preliminary results

Mobiles and Wearables for Heart Rate Variability Analysis

In the realm of health monitoring, the integration of mobile phones, wearable devices, and specialized trackers presents a diverse array of options for capturing biometric variables.

In our search for suitable wearables, we found that smartwatches, though popular, often lack comprehensive monitoring capabilities due to their emphasis on lifestyle features. Fitness trackers, specialized for physiological parameter tracking, were found to be more suitable for comprehensive monitoring, especially when equipped with PPG functionality and software development kit (SDK) support. Polar fitness trackers, identified through systematic search, deliver reliable measurements and satisfactory data quality and will be used to collect biometric data in our study with office workers. HRV parameters are valued for their objective assessment of both physical and mental states. From the abundance of HRV parameters, we selected single metrics (mean HR, mean NNI, RMSSD, CV) and combined metrics (PAPR, IVR, stress index) as appropriate HRV measures. PPG offers an efficient means of recording pulsometric data, facilitating HRV parameter extraction. A thorough analysis of HRV measures is essential for capturing work-related effects on valence and arousal throughout the office working day. By leveraging the capabilities of these devices, alongside thorough data analysis, we can assess physiological parameters such as HRV, that contribute to understanding both physical and mental states in various contexts, including the demands of the modern office environment.

Legend: planned, in progress, finished

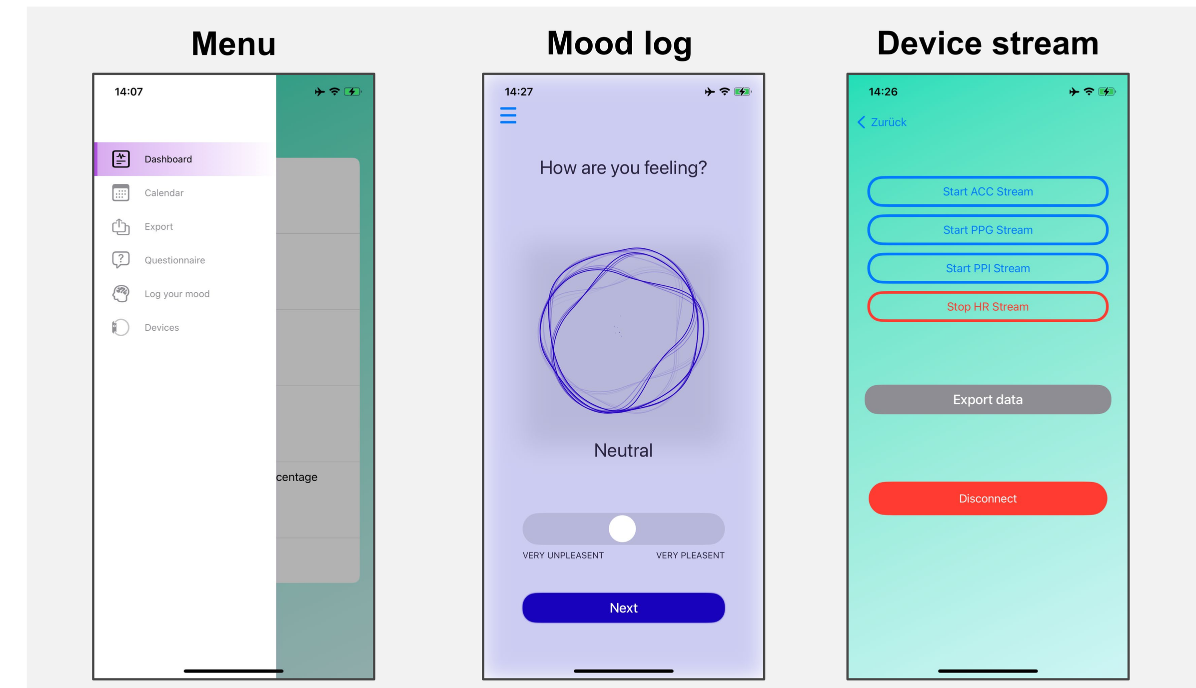

Fig. 2: Screenshots of the mobile app prototype

PC Software for Keyboard and Mouse Tracking

- Development of a software prototype designed to continuously monitor mouse and keyboard dynamics, as well as software type being used (e.g., browser, text editor, email, PowerPoint) to assess work-related valence and arousal
- Initiation of tracking and data collection after the user has launched the software and agreed to data collection upon every launch
- Collection of aggregated features related to emotion recognition, excluding the content of [illegible]ing. Thi[illegible]clude[illegible]suc[illegible]s the[illegible] movements, duration of pa[illegible]es, mea[illegible]well[illegible]e, and num[illegible] of [illegible]essed key[illegible]er minute, aimed at capturing the[illegible]tensity a[illegible] fluc[illegible]tion of intera[illegible] the office c[illegible]ext
- As[illegible]ssment [illegible]vale[illegible] and arousa[illegible]ve[illegible]ultiple ti[illegible]s throughout the day through sel[illegible] an int[illegible]ues[illegible]naire[illegible]mployi[illegible] non-verbal pictorial emotion manikins based on the methodologies outlined by Banholzer et al. (2021) and Bradley and Lang (1994)
- Analysis of emotional data in the workplace context utilizing JAWS (Van Katwyk et al., 2000)

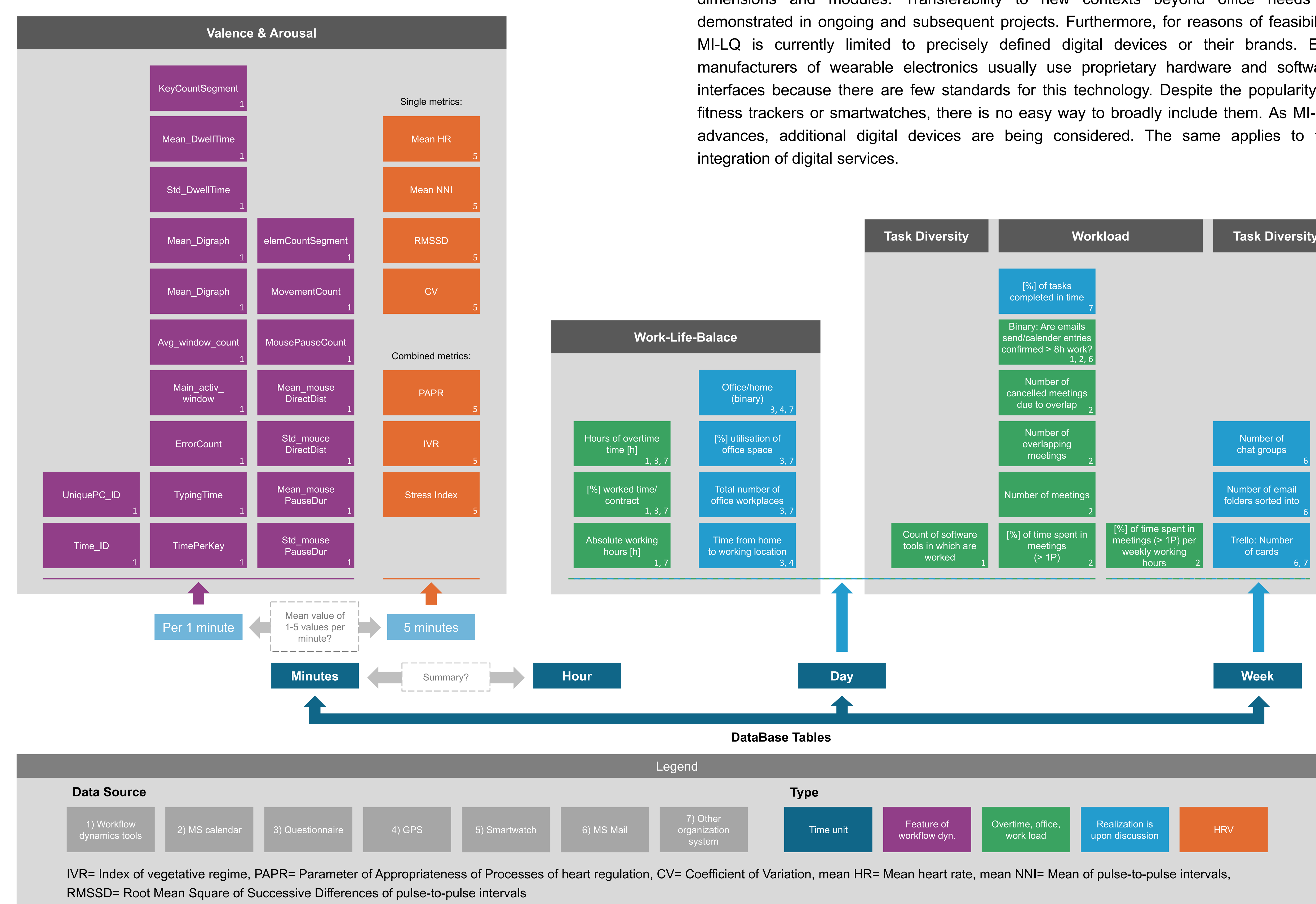

Fig. 3: Overview of data collection type and time point, datasets and data source

Mobile phone app prototype

Our project's mobile phone app prototype serves as both a user interface and a data relay system to a cloud-based platform for offline analysis of WrQoL metrics. Initially, participants will engage in regular check-ins (currently experimented with at hourly intervals), which will gradually be phased out as our model gains accuracy, mitigating participant fatigue while ensuring robust data collection. Participants will be provided with an initial baseline questionnaire and regular valence and arousal check-ins through the app interface (Fig. 2). To enhance data collection, future iterations of the app will incorporate additional sensor categories, including work-related schedules and external influences. Additionally, both online and onboard computations of WrQoL measures will be integrated into the mobile app.

Ultimately, our aim is to provide participants with individualized WrQoL indicators and potential resources based on their data analysis.

Conclusions

The systematic literature search identified relevant indicators for the assessment of WrQoL, which were translated into a set of objectively measurable indicators in a base model. Different digital devices in combination with digital applications were shown to be suitable for measuring work-related indicators. The ML and mobile app architecture as well as a cloud system were established as a basic framework for data input, output and processing for the objective determination of WrQoL. This project provides an innovative approach to objectively measure WrQoL in the digital world. By understanding the relationship between digital service use, WrQoL, and mental health, the project may provide preliminary insights into strategies for preventing stress and burnout in the workplace.

Limitations

The MI-LQ project aims to objectively quantify individual quality of life (QoL) using digital devices and applications combined with a ML approach. However, this study focused mainly on WrQoL in the office context as a starting point for overall QoL quantification. The established models and approaches were built in a modular way to easily add new dimensions and modules. Transferability to new contexts beyond office needs to demonstrated in ongoing and subsequent projects. Furthermore, for reasons of feasibility, MI-LQ is currently limited to precisely defined digital devices or their brands. E.g. manufacturers of wearable electronics usually use proprietary hardware and software interfaces because there are few standards for this technology. Despite the popularity of fitness trackers or smartwatches, there is no easy way to broadly include them. As MI-LQ advances, additional digital devices are being considered. The same applies to the integration of digital services.

Disclosure:
This study was funded by Co-funded by the European Union. This project is co-financed from tax revenues on the basis of the budget adopted by the Saxon State Parliament.
Conflicts of interest:
All authors were fully independent in the study design, analysis, interpretation of results, and writing of the abstract/poster.