FEDERAL UNIVERSITY OF CEARÁ

SCIENCE CENTER

COMPUTER SCIENCE DEPARTMENT

MASTER AND DOCTORATE PROGRAM IN COMPUTER SCIENCE

DOCTORATE IN COMPUTER SCIENCE

PEDRO ALMIR MARTINS DE OLIVEIRA

HEALFUL - INTERNET OF HEALTH THINGS PLATFORM TO MONITOR QUALITY OF LIFE

FORTALEZA

2023

PEDRO ALMIR MARTINS DE OLIVEIRA

HEALFUL - INTERNET OF HEALTH THINGS PLATFORM TO MONITOR QUALITY OF LIFE

Thesis presented to the Master and Doctorate Program in Computer Science of Science Center of Federal University of Ceará, as a requirement to obtain the title of Ph.D. in Computer Science. Area of Concentration: Information Systems

Supervisor: Ph.D. Rossana Maria de Castro Andrade
Co-supervisor: Ph.D. Pedro de Alcântara dos Santos Neto

FORTALEZA
2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

O49h Oliveira, Pedro Almir Martins de.
HEALFUL - Internet of Health Things Platform to Monitor Quality of Life / Pedro Almir Martins de Oliveira. – 2023.
172 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação , Fortaleza, 2023.
Orientação: Profa. Dra. Rossana Maria de Castro Andrade.
Coorientação: Prof. Dr. Pedro de Alcântara dos Santos Neto.

1. Internet das Coisas Médicas. 2. Qualidade de Vida. 3. Aprendizagem de Máquina. I. Título.
CDD 005

PEDRO ALMIR MARTINS DE OLIVEIRA

HEALFUL - INTERNET OF HEALTH THINGS PLATFORM
TO MONITOR QUALITY OF LIFE

Thesis presented to the Master and Doctorate Program in Computer Science of Science Center of Federal University of Ceará, as a requirement to obtain the title of Ph.D. in Computer Science. Area of Concentration: Information Systems

Approved on:

COMMITTEE

Ph.D. Rossana Maria de Castro Andrade (Supervisor)
Federal University of Ceará (UFC)

Ph.D. Pedro de Alcântara dos Santos Neto (Co-supervisor)
Federal University of Piauí (UFPI)

Ph.D. Miguel Franklin de Castro
Federal University of Ceará (UFC)

Ph.D. Ismayle de Sousa Santos
Ceará State University (UECE)

Ph.D. Tales Paiva Nogueira
University of the International Integration of the Afro-Brazilian Lusophony (UNILAB)

Ph.D. Wanderley Lopes de Souza
Federal University of São Carlos (UFSCar)

This thesis is dedicated to my beloved wife and son, Jayane and Hugo; my inspiring parents, Angelica and Matias; my great-aunt, Daluz; my sister, Juliana, and the close family of my wife, Brito, Umbelina, Jayro, Michelly, and Heitor. They provided unconditional support and deep inspiration in this journey.

ACKNOWLEDGEMENTS

What is the meaning of gratitude? According to the dictionary, gratitude means the quality of being grateful, recognition, the feeling of remembrance and gratitude for something received, the act of recognizing someone for an action or benefit achieved (FERREIRA, 2009). In psychology, this feeling is described as a reactive action generated by the capacity of an individual to feel satisfaction or a positive feeling due to a benevolent action of another (FREITAS *et al.*, 2009). For philosophy, gratitude is not only the richest of virtues but the mother of all others (Roman philosopher Marcus Tullius Cicero).

However, a better definition was written by Fr. Ederson Iarochevski, who says: **being grateful is a gesture of love!** He wrote, in Portuguese, that:

> "Dizer muito obrigado a uma pessoa é tão belo quando dizer eu te amo. Quando agradecemos, estamos nos desvencilhando de um dos vícios mais perigosos de nosso tempo, a auto-suficiência. Como é triste alguém pensar que não precisa de ninguém para viver. Quando o filho agradece a seus pais pelo esforço e carinho a ele dispensados, sela seu amor em relação aos pais. Quando o esposo diz muito obrigado à sua esposa pelo carinho e cuidado que a mesma tem com ele, está estabelecendo um laço ainda mais forte. Mesma situação quando a esposa reconhece a importância de seu esposo. Quando os amigos reconhecem o quanto são importantes uns para os outros, estão reafirmando uma relação que permanecerá no tempo, independentemente das intempéries que possam surgir. Mas, dizer muito obrigado em tempos de individualismos é um ato de coragem e receber um muito obrigado em tempos de auto-suficiências é um ato de humildade."

In this way, it is possible to conclude that giving thanks is a good exercise for the soul and, as stated in the epigraph of this work, if I saw further, I was undoubtedly supported by giants, not only in the professional and academic sphere but also in my personal life. Therefore, **I would like to thank everyone who directly or indirectly supported this arduous journey**. This victory is not mine! This victory is of all those who believed it was possible.

This time, I won't mention names, but everyone who reads this message will know through my actions the respect and gratitude I have for you. Finally, I thank God for everything. Sometimes I lost the way, but, in the end, the lessons learned will be highlighted, and I hope they can guide others on their journeys. Thanks for guiding my steps!

"If I have seen further it is by standing on the shoulders of Giants"

(Isaac Newton | Bernard of Chartres)

RESUMO

Avanços na Internet das Coisas (IoT), tais como miniaturização de sensores, expansão na capacidade de processamento de dados e aplicação de algoritmos inteligentes têm possibilitado avanços em vários domínios, incluindo a saúde. O termo Internet das Coisas Médicas (IoHT) é usado quando IoT é aplicada na saúde para prover soluções tais como o sensoriamento da Qualidade de Vida (QoL), detecção de quedas em idosos e análise de marcha. Assim, monitorar a Qualidade de Vida das pessoas tem atraído interesse devido aos benefícios associados, por exemplo, detecção de doenças e intervenções preventivas de promoção à saúde. Esses benefícios também possuem impacto individual no bem-estar dos pacientes, impacto econômico ao otimizar a relação custo-benefício dos recursos de saúde e impacto social ao promover melhores condições de vida. No entanto, a maioria dos instrumentos propostos para avaliar QoL são questionários, os quais tendem a ser custosos, invasivos e propensos a erros. Então, este trabalho apresenta uma solução para coleta de dados a partir de dispositivos inteligentes e aplicação de algoritmos de Aprendizagem de Máquina a fim de inferir a Qualidade de Vida dos usuários. Para alcançar essa solução, foi desenvolvida uma plataforma IoHT chamada Healful, a qual foi inspirada no loop de adaptação MAPE-K e fundamentada por duas revisões da literatura. Além disso, um estudo de caso com 44 participantes foi conduzido ao longo de seis meses nos quais dados de saúde foram coletados diariamente por meio de smartphones e dispositivos vestíveis. Esses participantes responderam o questionário WHOQOL-BREF semanalmente e os dados foram processados e compilados em dois datasets com 1.373 instâncias cada. Então, cinco modelos de Aprendizagem de Máquina foram construídos usando a técnica 10-fold cross-validation para estimar a Qualidade de Vida dos participantes. O Random Forest (RF) obteve os melhores resultados considerando a raiz do erro médio quadrático (RMSE). RF obteve um RMSE de 7,8616 para o domínio físico e 7,4591 para o domínio psicológico. Os resultados desta tese mostram que i) é possível usar IoHT para inferir QoL dos usuários, considerando uma margem de erro; ii) RF obteve performance aceitável para este problema, considerando os parâmetros estabelecidos na avaliação; e, iii) não foi encontrado um subconjunto decisivo para esse processo de inferência. Este último resultado reforça que a inferência da Qualidade de Vida usando IoHT não é trivial e apenas a combinação de um grande número de características pode dar insights relevantes para a inferência da Qualidade de Vida dos usuários.

Keywords: internet das coisas médicas; qualidade de vida; aprendizagem de máquina.

ABSTRACT

Advances in the Internet of Things (IoT), such as sensor miniaturization, efficient communication protocols, expansion in data processing capacity, and application of intelligent algorithms, have made possible advances in several domains, including healthcare. Internet of Health Things (IoHT) is the term used when IoT is applied to healthcare to provide solutions, for example, non-invasive Quality of Life (QoL) sensing, older adults' fall detection, and gait analysis. Monitoring people's QoL has attracted interest due to the health benefits of an accurate QoL analysis, such as disease detection and early healthcare interventions. These benefits also have individual impacts by increasing well-being, economic impacts by improving the cost-effectiveness of healthcare resources, and social impacts by promoting better living conditions. Although many instruments for QoL assessment have been proposed, most of them are questionnaires, and their application is time-consuming, intrusive, and error-prone. Based on that and using IoHT, this work proposes to collect data from Smart Devices and apply Machine Learning techniques to infer users' QoL. To achieve that, an IoHT platform called Healful was developed to monitor users' QoL. This platform was inspired by the MAPE-K loop and supported by two literature reviews. Also, a case study with 44 participants was conducted for six months, and during this evaluation, health data were collected through smartphones and wearables daily. These participants answered the WHOQOL-BREF questionnaire weekly, and these data were processed and compiled into two datasets with 1,373 instances each. Next, five Machine Learning models were built using 10-fold cross-validation to estimate participants' QoL. Random Forest (RF) had the best results considering the Root Mean Squared Error (RMSE). RF got an RMSE of 7.8618 for the physical domain and 7.4591 for the psychological domain. The thesis findings showed that: i) it is possible to use IoHT to infer users' QoL, considering a certain margin of error; ii) RF had a reasonable performance for this problem; and iii) a decisive subset of features for the inference process was not found. This last point reinforces that QoL inference using IoHT is not trivial, and only the combination of a large number of features can give relevant insights into users' Quality of Life.

Keywords: internet of health things; quality of life; machine learning inference.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

6LoWPAN	IPv6 over Low-Power Wireless Personal Area Networks
AI	Artificial Intelligence
ANN	Artificial Neural Networks
API	Application Programming Interface
BAN	Body Area Network
BLE	Bluetooth Low Energy
CI	Computational Intelligence
CNN	Convolutional Neural Networks
CRISP-DM	CRoss Industry Standard Process for Data Mining
EC2	Amazon Elastic Compute Cloud
EHR	Electronic Health Record
GAE	Google App Engine
HRQoL	Health-related Quality of Life
IBM	International Business Machines Corporation
IoHT	Internet of Health Things
IoT	Internet of Things
ISOQOL	International Society for Quality of Life Research
JSON	JavaScript Object Notation
LAN	Local Area Network
LDA	Latent Dirichlet Allocation
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multilayer Perceptron
MSE	Mean Squared Error
NHS	National Health Service
OECD	Organization for Economic Co-operation and Development
PaaS	Platform as a Service
PAN	Personal Area Network
PoC	Proof-of-Concept
QoL	Quality of Life

REM	Rapid Eye Movement
RMSE	Root Mean Squared Error
RQ	Research Questions
SOM	Self-Organizing Maps
SSL	Secure Sockets Layer
TAM3	Technology Acceptance Model 3
TAR	Technical Action Research
UN	United Nations
WHO	World Health Organization

CONTENTS

1 INTRODUCTION

Kevin Ashton stated that “Internet of Things (IoT) has the potential to change the World, just as the Internet did. Maybe even more so”. (ASHTON *et al.*, 2009). This paradigm allows physical objects to sense and act in a transparent way (ATZORI *et al.*, 2010) and domains, such as smart cities, agriculture, robotics, manufacturing, retail, smart grid, and healthcare have used IoT to innovate (ANDRADE *et al.*, 2020). A century ago, for example, patients with diabetes could not know their health status without going to the doctor. However, with the evolution of IoT technology in healthcare, continuous health monitors have become increasingly accessible (RODRIGUES *et al.*, 2018; THANGAM *et al.*, 2022).

Then, this brief summary exemplifies how IoT is able to remodel health and medical care, improving people’s Quality of Life (QoL), and this Chapter introduces the thesis presenting its context and motivation (Section 1.1), the problem, goal, and scope (Section 1.2), a running example (Section 1.3), research questions and hypothesis (Section 1.4), research methodology (Section 1.5), and the thesis outline (Section 1.6).

1.1 Context and Motivation

The Internet of Things (IoT) paradigm enables interaction and cooperation among physical things through the Internet to achieve common goals (ATZORI *et al.*, 2010). For Gubbi *et al.*, IoT environments enable the interaction between physical and logical objects to provide ubiquitous services to their users. These objects, like sensors and actuators, can be combined to improve their capabilities. These large objects - called Smart Devices or Smart Objects - are designed to help people accomplish tasks (KORTUEM *et al.*, 2009; BARRETO *et al.*, 2017).

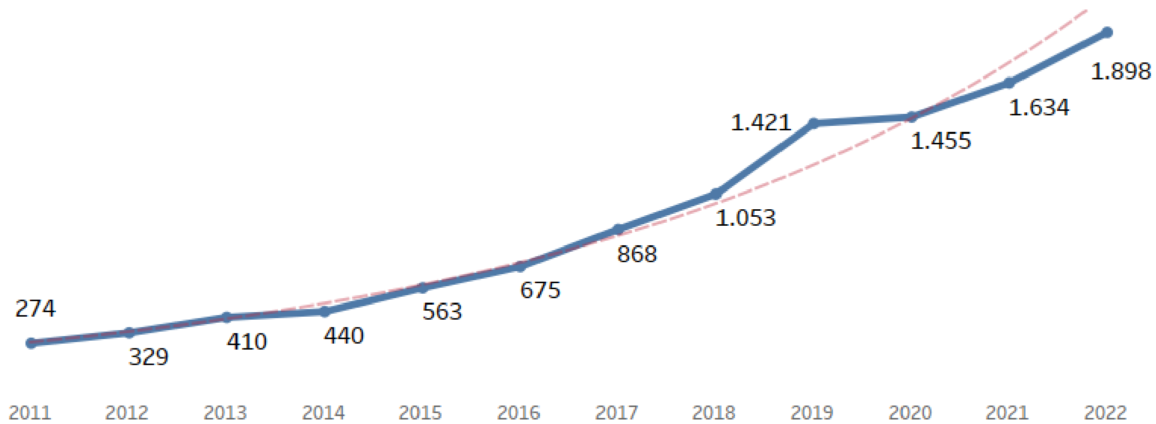
Since its emergence, IoT has been adapted and strengthened by advances, such as sensor miniaturization, efficient communication protocols, expansion in data processing capacity, and application of machine learning algorithms (ATZORI *et al.*, 2010). Due to these advances, many industries have used IoT solutions to address complex problems. Among them, it is possible to highlight the healthcare industry (ISLAM *et al.*, 2015).

IoT, when applied to healthcare, can be called of Internet of Health Things (IoHT), and it has the potential to reduce costs, increase QoL, and improve the user’s experience (RODRIGUES *et al.*, 2018). Thus, the Internet of Things has been widely applied in the healthcare domain (GUBBI *et al.*, 2013; ISLAM *et al.*, 2015). As examples of IoHT solutions, there are glucose sensing (ISTEPANIAN *et al.*, 2011), electrocardiogram monitoring (AGU *et al.*, 2013), ingestible sensors for measuring medication adherence (HAFEZI *et al.*, 2014), gait and posture analysis (JUNIOR *et al.*, 2021), older adults fall detection (ARAÚJO *et al.*, 2022), and non-intrusive Quality of Life monitoring (OLIVEIRA *et al.*, 2022b).

The advancement of IoHT is also closely related to population aging (NATIONS, 2019) and the increasing number of citizens in cities (NATIONS, 2018) because these phenomena have been putting much pressure on healthcare systems around the World. For example, the United Nations (UN) projected that, in 2050, there will be 2.1 billion of older persons (*i.e.*, approximately 20% of the world's population will be aged 65 or over. In 2022, this percentage is nearly 10%) and 68% of the world's population residing in urban areas (NATIONS, 2018; NATIONS, 2019). Due to these phenomena, healthcare systems must be adapted to reduce costs, provide better living conditions and improve people's QoL (MANO *et al.*, 2019).

Hence, many studies have been conducted to propose new healthcare solutions. For example, Figure 1 shows a sampling of the number of papers obtained in the Scopus[1] database, using the term "Quality of Life" as the search string and filtering for the Computer Science area. With this data, it is possible to observe exponential growth (represented by the red line and validated with a *p-value* < 0.0001 and an R-squared equals to 0.9888).

Figure 1 – Distribution of QoL-related papers in Computer Science from 2011 to 2022



Source: author.

Quality of Life has been studied for a long time, and in 1994, the World Health Organization (WHO) published a widely used definition. For WHO, QoL can be defined as the individual perception of life in a sociocultural context concerning goals, expectations, and

[1] Scopus website: scopus.com.

personal standards (WHOQoL Group, 1994). In addition, according to the WHO, it is crucial to measure QoL because it has a close relationship with the health status (ESTRADA-GALINANES; WAC, 2018; MATE, 2022), and it provides valuable data to medical practice (WHO, 1998). From this definition, many mechanisms to evaluate QoL have been proposed such as, SF-36 (JR, 1999), KIDSCREEN-52 for children and teenagers (RAVENS-SIEBERER *et al.*, 2005), EQ-SD from the EuroQoL Group (RABIN; CHARRO, 2001), RAND-36 (HAYS; MORALES, 2001), and others (ADAY; CORNELIUS, 2006; PEQUENO *et al.*, 2020).

One of the most mentioned instruments for QoL assessment is the WHOQOL-BREF questionnaire (SKEVINGTON *et al.*, 2004) due to its reliability and cross-cultural validity. The WHOQOL-BREF was evaluated in 23 countries and is available in 19 different languages. It has 26 questions distributed into four domains: physical, psychological, social, and environmental. The physical domain assesses motor facets such as daily living activities, medicine dependence, mobility, and sleep quality. The psychological domain relates to body image, negative and positive feelings, self-esteem, and other mental aspects. The social domain observes social relationships, and the environment domain aims to evaluate the environmental facets. Unfortunately, the continuous application of this kind of questionnaire is tedious, bothersome (SANCHEZ *et al.*, 2015; OLIVEIRA *et al.*, 2022a), and can also include a bias as the patient needs to actively provide data, making the patients' adherence hard (HAO *et al.*, 2017).

Thus, despite solid medical knowledge on measuring people's Quality of Life, continuous QoL monitoring is still an open problem (OLIVEIRA *et al.*, 2022b). First, there are many assessing instruments, and it is complex to select the right one due to its particularities (HYLAND, 2003; PEQUENO *et al.*, 2020). Second, most of the available instruments for this assessment are questionnaires. Thus, it is hard to engage patients in monitoring this indicator (SILQUEIRA, 2005; SANCHEZ *et al.*, 2015; PEQUENO *et al.*, 2020). Third, a literature review showed that although there are digital versions of the Quality of Life questionnaires (OLIVEIRA *et al.*, 2022a), such solutions inherit issues from their non-digital versions. Another reason to consider QoL continuous monitoring a problem to be explored is that there are many opportunities to study the relationship between IoHT data and individual behavior (this association has been called digital phenotyping) (ONNELA, 2021). Finally, QoL monitoring solutions that use inference models trained with smart device data need to deal with data collection and processing complexity and security and privacy factors (OLIVEIRA *et al.*, 2022a).

The relevance of this problem emerges from the health benefits generated from up-to-date and accurate Quality of Life data, such as disease detection and early healthcare interventions. DOHR *et al.* (2010) also reinforces that these benefits have individual impacts by increasing safety and well-being; economic impacts by improving the cost-effectiveness of limited healthcare resources; and social impacts by promoting better living conditions.

Concerning the cost-effectiveness relation, it is common to observe that current healthcare systems worldwide are mainly reactive, *i.e.*, they wait to act when the patient becomes ill (MARVASTI; STAFFORD, 2012). Thus, many studies are promoting a reengineering of "sick care" to real healthcare systems (HARKIN, 2004; MARVASTI; STAFFORD, 2012). To do that, investing in preventive health strategies is fundamental, but there is a long path to achieving the ideal scenario. For example, as Figure 2 shows, in 2017, a report of the Organization for Economic Co-operation and Development (OECD) pointed out that only 2.8% of health spending goes on prevention (GMEINDER *et al.*, 2017). In addition, a WHO report indicates that this percentage grew to 5% in 2020 (VRIJBURG; HERNÁNDEZ-PEÑA, 2020).

Figure 2 – Fraction of health expenditure applied in preventive health strategies

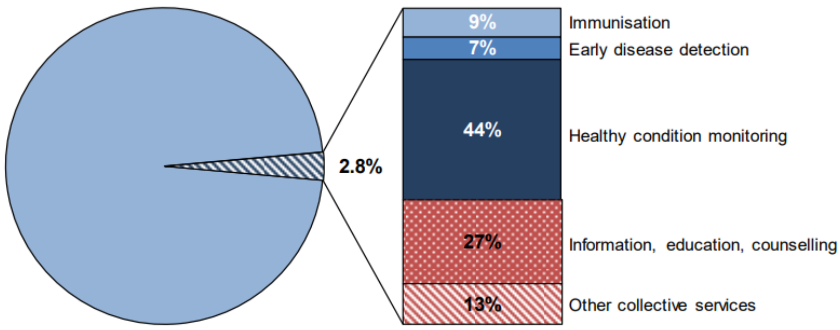


Image from OECD Health Statistics report (GMEINDER *et al.*, 2017).

To conclude this section, despite the number of initiatives to improve the citizens' QoL, there is room for opportunities, especially regarding continuous monitoring of QoL, which should collaborate in the early identification of health issues (CDC, 2018).

1.2 Problem, Goal, and Scope

Based on the aforementioned motivation, this thesis focuses on the problem related to the need for less-intrusive, continuous and ubiquitous Quality of Life monitoring. In general, the existing solutions are time-consuming and non-ubiquitous because users must actively interact with the Quality of Life assessment instrument (PEQUENO *et al.*, 2020; OLIVEIRA *et al.*, 2022a). Thus, it takes work to keep user engagement (SANCHEZ *et al.*, 2015), and the results can be biased (HAO *et al.*, 2017).

To address these issues, the goal of this thesis can be summarized as follows:

This work **aims to** develop an Internet of Health Things (IoHT) platform capable of inferring users' Quality of Life (QoL) ubiquitously, using users' Smart Devices and Machine Learning algorithms.

The **scope** of this thesis is focused on the physical and psychological Quality of Life domains, and the target profile is independent adults. As the expected impacts of this study include providing a solution capable of early detecting health issues, it was decided to focus on the independent adult profile. Other profiles, such as older adults, tend to live with chronic diseases, and these health issues could be alleviated with early care in adult life. On the other hand, children and teenagers depend on their parents or guardians for medical care (FACCHINETTI *et al.*, 2023; GOLDTHORPE *et al.*, 2019). Regarding the domains, the rationale for investigating these two domains was based on analyzing the facets incorporated in each domain and the availability of data to characterize each domain (WHOQoL Group, 1994; DOHERTY; GAUGHRAN, 2014; ARSENOVIĆ *et al.*, 2023).

Within the physical domain are facets such as activities of daily living, energy and fatigue, mobility, sleep and rest, and work capacity (WHOQoL Group, 1994). As for the data, devices used in this research allow for obtaining steps, daily calories, mobility by GPS, time in each sleep stage, and physical activities (OLIVEIRA *et al.*, 2022b).

For the psychological domain, although the facets are difficult to characterize, the medical literature points to the existence of a direct relationship between QoL in the physical and mental spheres (SCHUCH *et al.*, 2016). As for the data, the relationship between mental health and patients' behavior (measured using smartphones) was observed (DAHLBERG *et al.*, 2022).

In turn, it is **out of the scope** of this thesis the social and environmental domains that involve the collection of sensitive data (such as social network activities) and data that are difficult to collect, such as the characteristics of the user's environment (*e.g.*, pollution, noise, traffic, climate, quality of transport, accessibility, and security). However, it is an opportunity for future research to explore public data related to these last two domains.

To achieve the main goal, the following specific objectives are defined:

- To identify which data can be obtained from commercial IoHT devices to support QoL inference process and which health indicators can be defined from these data.

– To build Machine Learning (ML) models to infer QoL using IoT data.
– To design and develop a platform able to infer the user's QoL using Smart Devices data.

This thesis delivers three significant results:

- a ubiquitous solution able to infer the users' Quality of Life through IoHT data, reducing the time to get this health metric and improving users' engagement.
- a platform to support the development of IoHT systems that can i) deal with device heterogeneity and lack of interoperability; ii) reduce the cost to analyze QoL data; and iii) simplify the definition of digital healthcare interventions.
- a fully anonymized dataset to support the investigation of techniques to infer users' QoL.

1.3 Running Example

This section presents an illustrative COVID-19-related scenario to reinforce the relevance of this study. Nevertheless, this scenario could easily be generalized to contexts beyond the pandemic, and the proposed solution is not restricted by the COVID-19 context.

The COVID-19 pandemic significantly affected people's health (USHER *et al.*, 2020). This impact was direct in cases of those who died or were left with sequelae due to the coronavirus infection and indirect in those who developed other health issues from social isolation. Unfortunately, these impacts remain for years after the last significant wave (MUNSTER *et al.*, 2020). Observing this context, a health insurance company decided to develop a ubiquitous solution to continuously monitor people's Quality of Life, aiming to detect health issues related to physical and psychological aspects as soon as possible.

This scenario reveals many challenges. Suppose the company opts for a solution that uses digital versions of the Quality of Life questionnaires (PEQUENO *et al.*, 2020). In that case, it will have to deal with the difficulty of engaging users in answering the questions (PUKELIENE; STARKAUSKIENE, 2011) and with the bias introduced by the repeated application of this survey (HAO *et al.*, 2017). Furthermore, suppose that the users' profile is not homogeneous. In that case, there will be the issue of choosing the suitable questionnaire given the many options considering the profile (LAGADEC *et al.*, 2018; GERMAIN *et al.*, 2019) and pre-existing health problem (MOKHATRI-HESARI; MONTAZERI, 2020; ASADY *et al.*, 2021). Even so, if the company decides to use digital QoL surveys, it will have to consider the effort (measured by time) to be included in the user's routine (KOPEC; WILLISON, 2003; BOWLING, 2005) because, depending on the chosen questionnaire, the number of questions can grow significantly.

For sure, digital QoL surveys cannot be classified as ubiquitous based on Mark Weiser's definition (WEISER, 1999). For Weiser, ubiquitous computing promotes increasingly natural and less intrusive interactions, allowing users to benefit from services without realizing that they are interacting with these technologies (YONG *et al.*, 2022).

Aiming to find a ubiquitous solution, the company could use the Internet of Health Things to collect users' data and train Machine Learning algorithms to perform QoL inference. In this way, the challenges faced would be related to the pipeline from data collection to inference model building. Regarding collection, there is a heterogeneity challenge due to the number of IoHT devices available in the market (ANDRADE *et al.*, 2020) and the noise and data gaps inherent in this process (BERROCAL *et al.*, 2020). Furthermore, in analysis, it is necessary to find proper methods to clean and improve the data to avoid inconsistencies. Finally, in inference, there are issues related to model selection and the workflow that enables online learning based on user feedback (OLIVEIRA *et al.*, 2018; KHAN *et al.*, 2021).

After this initial analysis of the technical challenges, the health insurance company should select an employee to investigate how the proposed solution would impact his life. The chosen employee is named John, and his profile is as follows.

John is a 35-year-old software engineer, married, with one child. Before the COVID-19 pandemic, he was engaged in daily physical activities, regularly communicated with friends and family, and slept well. However, with the worsening global health crisis and the need for restrictive actions, John faced difficulties maintaining his healthy habits and gradually began having sleep and loneliness issues.

If a questionnaire were applied to assess John's Quality of Life in 2019, the physical and psychological scores would likely be high. Nevertheless, the reapplication of this instrument in 2021 would bring different results. While in 2019, the scores were good, in 2021, they indicated an alert situation. Keeping these indicators at low levels can lead to severe health problems such as depression and chronic anxiety.

The ideal scenario would be to continuously monitor these health indicators to alert John as soon as possible about the worsening in his indicators. However, as previously discussed, the continuous application of self-reported QoL questionnaires (non-ubiquitous monitoring, according to Weiser) makes engaging users challenging. This reinforces the need for an IoHT solution capable of monitoring users' Quality of Life to aid them in maintaining good health status. Therefore, the discussed scenario is used as a reference throughout this thesis document.

1.4 Research Questions and Hypothesis

According to WOHLIN; AURUM (2015), the Research Questions (RQ) formulation is critical since it determines the process followed by the investigation. Suitable RQ should guide the study to address relevant problems. Thus, three major RQ were defined in this thesis:

I. What prior knowledge is available about the IoHT and its application to Quality of Life?
Rationale: this RQ is fundamental for identifying prior knowledge about the Internet of Health Things and Quality of Life and how to monitor people's QoL automatically. The answer to this question is presented at the end of Chapter 2.

II. Which data can be ubiquitously obtained from commercial Internet of Health Things devices to represent users' health context?
Rationale: as this thesis seeks to address the inference of Quality of Life, it is necessary to investigate what data are available in IoHT environments for this task. In addition, the commercial term was included in this RQ to guide the investigation of solutions that the industry can adopt. The answer to this question is presented at the end of Chapter 3.

III. What requirements are important to design and implement an IoHT platform for ubiquitous QoL monitoring?
Rationale: this RQ concerns what is necessary to materialize a solution for the main problem investigated in this thesis. The answer to this question is presented in Chapter 5.

Together, these RQs guide this thesis to get the necessary background about IoHT and QoL, understand which data can be used to infer QoL, and develop the platform to integrate and deliver all knowledge produced in this study. Taking into account these questions, it is possible to raise the thesis hypothesis:

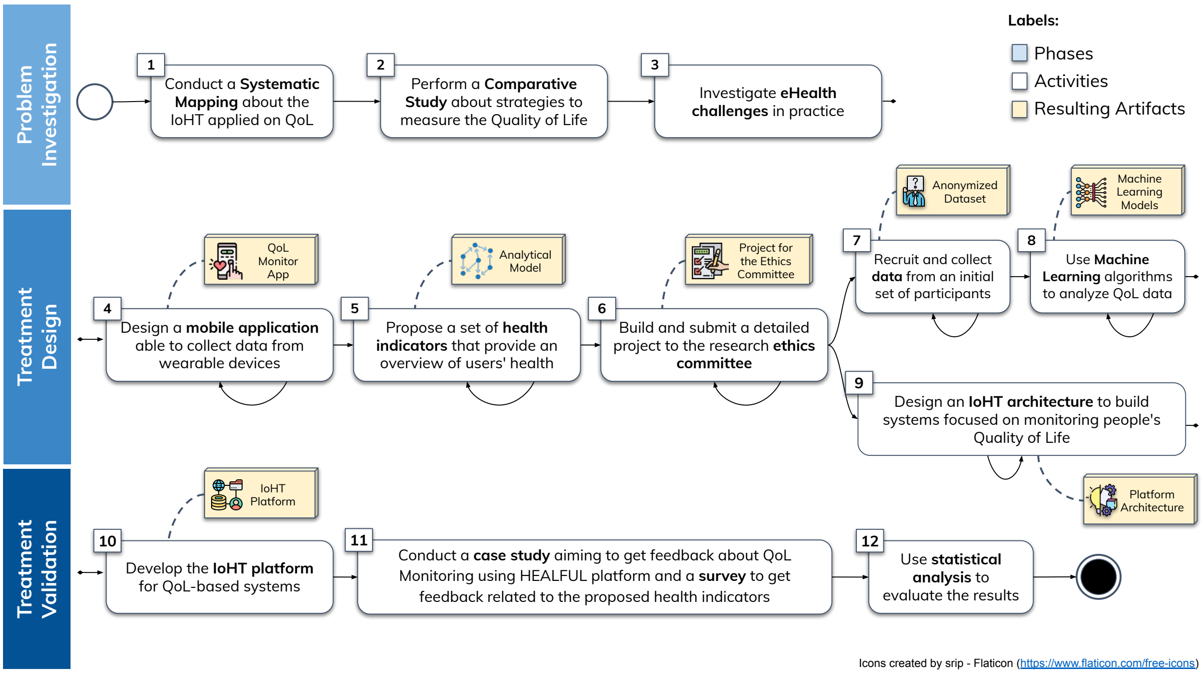
– H_0: an IoHT platform that uses Machine Learning algorithms can infer users' Quality of Life for the physical and psychological domains, making QoL monitoring less intrusive when compared with self-reported QoL questionnaires.

This hypothesis is discussed in Chapter 5 considering the intelligent model accuracy when compared with reference questionnaires, the analysis of eHealth professionals on the proposed indicators using a survey, and the user's perception of the system's ubiquitously.

1.5 Research Methodology

The methodology of this thesis uses the Technical Action Research (TAR) that consists of two cycles: design and engineering (WIERINGA, 2014a). The thesis research activities[2] are related to the first cycle (Design Cycle), as presented in Figure 3.

Figure 3 – Research methodology



Source: author.

Figure 3 presents the phases and activities highlighting the artifacts built throughout this investigation. The three phases are: i) **Problem Investigation** to study the concepts related to the Internet of Health Things (IoHT) and Quality of Life (QoL). It includes problem identification, motivation, and goal definition; ii) **Treatment Design** for proposing the artifacts to address the research problem; and, iii) **Treatment Validation** for conducting case studies focused on validating the proposed solution.

- ✓ **Activity 1**: to conduct a Systematic Literature Mapping to investigate state of the art about the Internet of Health Things applied to Quality of Life. This study made it possible to identify the challenges regarding developing QoL-based IoHT systems. This activity resulted in two papers (OLIVEIRA *et al.*, 2021; OLIVEIRA *et al.*, 2022a).
- ✓ **Activity 2**: to perform a comparative study about strategies to measure the Quality of Life. In the literature, there are many strategies to measure QoL. Thus, it is fundamental to find

[2] For detailed information on the research decisions made throughout this project, see Appendix A.

and compare these strategies to understand their context and applications. This review supports the choice of which strategy would be used as the standard throughout this work (OLIVEIRA *et al.*, 2022b).

✓ **Activity 3**: to investigate eHealth challenges in practice and analyze which data can be obtained from commercial IoHT devices to support QoL inference process. This activity is essential because, as previously mentioned, we intend to use commercial devices in the solution proposed in this thesis. The results of this activity were published at the 15th International Conference on Health Informatics (OLIVEIRA *et al.*, 2022).

✓ **Activity 4**: to design a mobile application able to collect data from wearable devices (*e.g.*, smartphones, smartwatches, smart bands) in order to validate QoL inference process. To assist the data collection process, it was decided to use studies previously developed in the research group, such as a lite version (only the context management module) of LoCCAM-IoT (ANDRADE *et al.*, 2020). As a result of this activity, it was developed the first version of the QoL Monitor app (see Chapter 4 Section 4.2 for more details).

✓ **Activity 5**: to propose health indicators to provide an overview of users' health as a complementary strategy to QoL inference. This activity is required because the result of the inference process is a number ($[0,100]$) that characterizes users' QoL. Therefore, it is essential to explain this result. At the end of this activity, an analytical model was produced with five health indicators (daily mobility, physical activity level, sleep quality, loneliness level, and social mobility level).

✓ **Activity 6**: to submit a detailed project to the research ethics committee of the Federal University of Ceará (UFC) requesting permission to conduct the study. This project[3], written in Portuguese, can be accessed through the link github.com/great-ufc/healful-thesis.

✓ **Activity 7**: to recruit and collect data from an initial set of participants over three months to obtain sufficient data for proposal validation. The data collection process was executed daily, including a weekly application of Quality of Life questionnaires. This activity started on March 14, 2022, with the recruitment of 20 participants, ending on June 14. The final result of this activity is an anonymized dataset with QoL-related information. This dataset is fundamental to building QoL inference models. Currently, this dataset is proprietary of the GREat lab and is not available for public use.

[3] The project (registered under the ID number 56153322.0.0000.5054) was approved on March 9, 2022, by the coordination of the UFC ethics committee (legal opinion number 5.282.056). This project is necessary since studies that require human participation must comply with CONEP regulations 466/2012 and 510/2016.

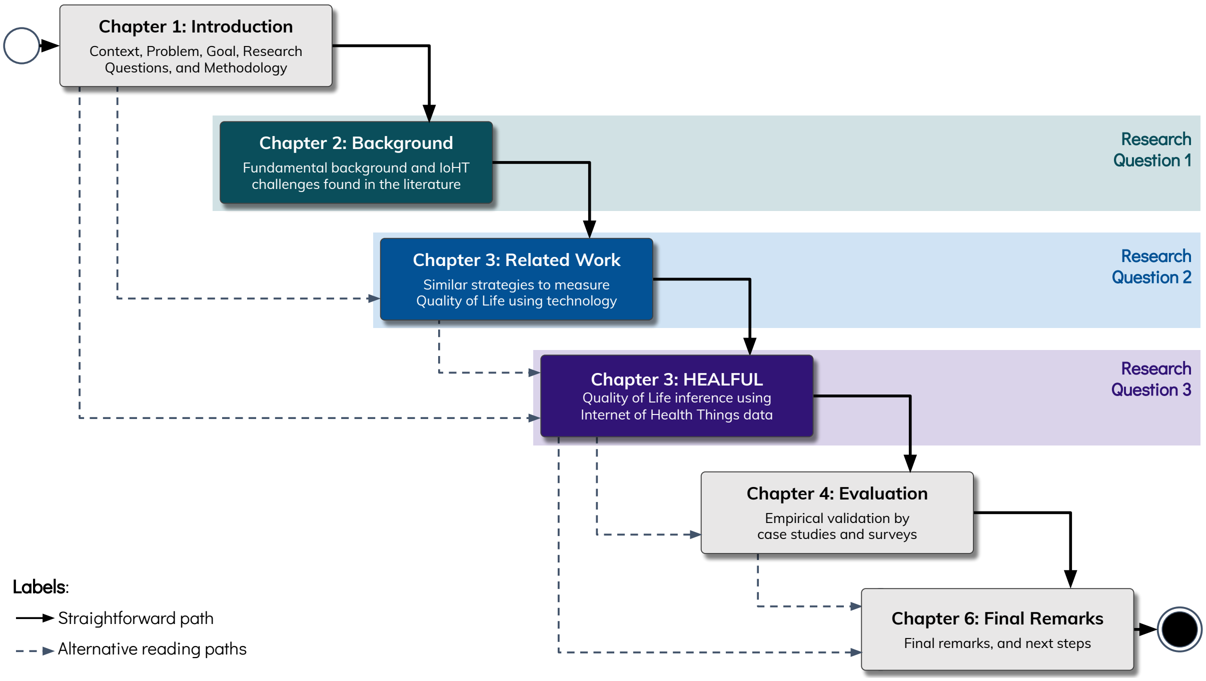
- ✓ **Activity 8**: to use Machine Learning algorithms to analyze the correlation among the data of smart environments sensors and QoL domains. As a result of this activity, it is expected that ML models can infer QoL from IoHT data.
- ✓ **Activity 9**: to design the proposed IoHT to build systems focused on monitoring and improving people's QoL, taking into account the critical requirements of healthcare applications such as reliability, data privacy, and security. This activity was performed parallel with activities 7 and 8 since data collection was passive. As a final result, an architecture for a platform was obtained.
- ✓ **Activity 10**: to develop a platform that integrates user data monitoring, the intelligent model for inferring the level of Quality of Life, and the application of strategies to enhance the user environment through smart objects.
- ✓ **Activity 11**: to conduct a case study to get feedback from users regarding a QoL monitoring system developed with the support of the proposed platform (OLIVEIRA *et al.*, 2023). To do that, a partnership was established with two educational institutions to recruit a significant number of participants. In addition, a survey was conducted with health professionals to get feedback about the proposed health indicators.
- ✓ **Activity 12**: to use statistical analysis to evaluate the results obtained in the previous evaluation methods. This activity is fundamental for artifact improvement.

1.6 Thesis Outline

The thesis chapters are structured as shown in Figure 4. Chapter 2 brings an overview of Health, Quality of Life, the Internet of Health Things, Software Engineering for IoHT, Machine Learning applied to Health, and the Internet of Health Things (IoHT) challenges found in the Systematic Literature Mapping. The concepts of these areas compound the work background and they provide the answer to the first Research Question (RQ1) regarding available prior knowledge about IoHT applied in QoL.

Then, Chapter 3 presents the related works and discusses the answer to the second Research Question (RQ2) concerning which data can be collected from commercial IoHT devices to represent users' health context. Next, in Chapter 4, the Healful platform is presented together with the answer to the third Research Question (RQ3), and the platform evaluation is discussed in Chapter 5. In the end, Chapter 6 brings the main deliverables, key findings and the open areas for further research.

Figure 4 – Work structure and possible reading paths



Source: author.

Figure 4 also visually represents the possible reading paths. The straightforward path starts with Chapter 1 and follows all others until its conclusion in Chapter 6. The solid line represented this straightforward path. However, suppose the reader is familiar with IoHT literature and background. In that case, it is possible to flow from the Introduction in Chapter 1 to the Healful platform in Chapter 4. The dashed lines represent alternative paths, such as the Introduction, Healful, Evaluation, and Final Remarks.

2 BACKGROUND

This chapter presents the fundamental theory for this thesis. Initially, there is the technological background concerning the Internet of Health Things (Section 2.1) and Machine Learning applications in healthcare (Section 2.2). Then, the WHO's definition of Quality of Life and the relevance of measuring people's QoL are discussed in Section 2.3. Next, Section 2.4 discusses IoHT challenges found in the literature[1] since they are relevant to highlight the gaps addressed by this thesis, and, finally, Section 2.5 summarizes the chapter content, providing an answer for the Research Question 1.

2.1 Internet of Health Things

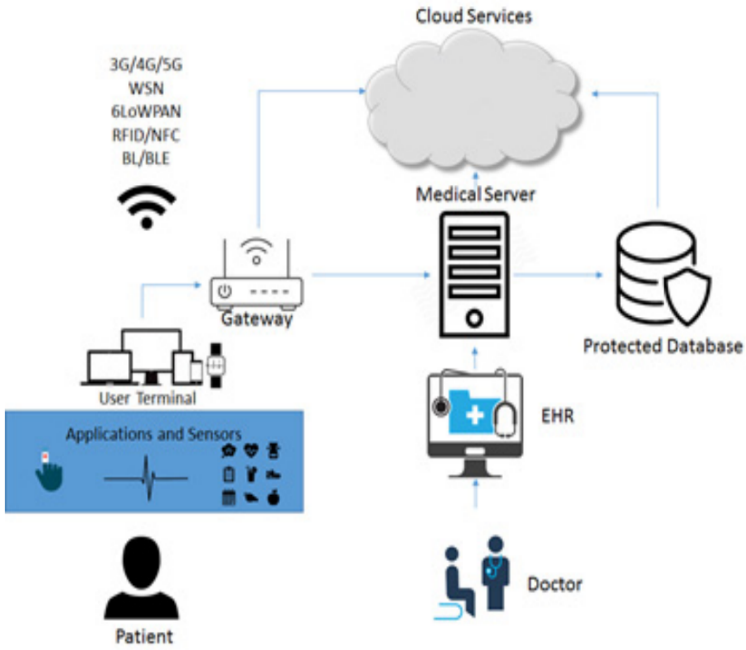
Applying IoT in healthcare created a new area called the Internet of Health Things (RODRIGUES *et al.*, 2018). IoHT has the potential to reduce costs, increase QoL, and improve the user's experience (ISLAM *et al.*, 2015). IoHT can also be seen as a specialization of IoT focused on connecting health objects to provide healthcare facilities, for example, systems to monitor vitals signs (LI *et al.*, 2017; DESHKAR *et al.*, 2017), ambient assisted living for elderly (DOHR *et al.*, 2010), detection of adverse drug reaction (JARA *et al.*, 2010), fall detection (ALMEIDA *et al.*, 2016), and fitness tracking (YONG *et al.*, 2018).

In general, IoHT systems have the following structure (Figure 5): many sensors are used to collect patient data. Then, due to these sensors' restrictions, the data is transmitted to a more robust node or a gateway using protocols like Bluetooth Low Energy (BLE) and IPv6 over Low-Power Wireless Personal Area Networks (6LoWPAN). After that, the data can be sent to cloud services or specific Electronic Health Record (EHR) systems. Finally, health data can be processed using Machine Learning algorithms to detect anomalies (RODRIGUES *et al.*, 2018).

According to QADRI *et al.* (2020), in the following years, IoHT will significantly impact the advancement of the healthcare industry. The authors also pointed out that, in the so-called Medicine 4.0, there will be an increased effort to develop platforms at the hardware and software levels. Nevertheless, despite the recent advances, this area still has faced challenges such as the standardization of devices, IoT platforms specialized in supporting the development of healthcare applications, quality assurance, data security, and privacy (ISLAM *et al.*, 2015; AGHDAM *et al.*, 2021; RAYAN *et al.*, 2021).

[1] It was decided not to be deep into the systematic mapping process because this research method is already well-known. However, all information is in the Appendix B.

Figure 5 – An example of IoHT-related architecture



Source: RODRIGUES *et al.* (2018).

2.2 Machine Learning applied on Health

TOPOL (2019) stated that the Artificial Intelligence (AI) in healthcare could provide a panoramic view of medical data to improve decision-making; to avoid errors (*e.g.*, misdiagnosis and unnecessary procedures); to support exam interpretation; and to recommend appropriate treatments. Most of these advances are becoming a reality due to a large amount of available health data and improvements in Machine Learning techniques.

Machine Learning is an AI subarea focused on the study of algorithms that use experience (*e.g.*, historical data) to improve their performance or to make accurate predictions (MOHRI *et al.*, 2018). These algorithms have been successful in several applications such as document classification, natural language processing, speech processing, computer vision applications, fraud detection, identification of network intrusion, learning how to play games, autonomous vehicle control, and, more recently, in healthcare applications (MOHRI *et al.*, 2018).

The tasks in which Machine Learning techniques are commonly applied can be summarized in five points (MOHRI *et al.*, 2018). Classification, when the problem is to assign a category for each data set item; regression, for problems that aim to predict a real value; ranking, when it is needed to order the items according to specific criteria; clustering, when the main goal is to build homogeneous data partitions; and, dimensionality reduction, when it is necessary to transform an initial representation of items into lower-dimensional representation.

Despite the significant advances in this area, using historical data as a learning strategy also brings ethical challenges (IAN; EIBE, 2005). Inappropriate data selection can lead to biased models. When data are related to people, the impact of these biases can be discriminatory behavior in resulting models. For example, if the data used to train an ML algorithm to decide when to grant loans has a racial bias, the model built from this training will tend to keep this behavior. SRINIVASAN; CHANDER (2021) present a taxonomy of possible AI biases along the AI pipeline and discuss strategies for reducing potential bias, such as incorporating domain-specific knowledge and using representative datasets.

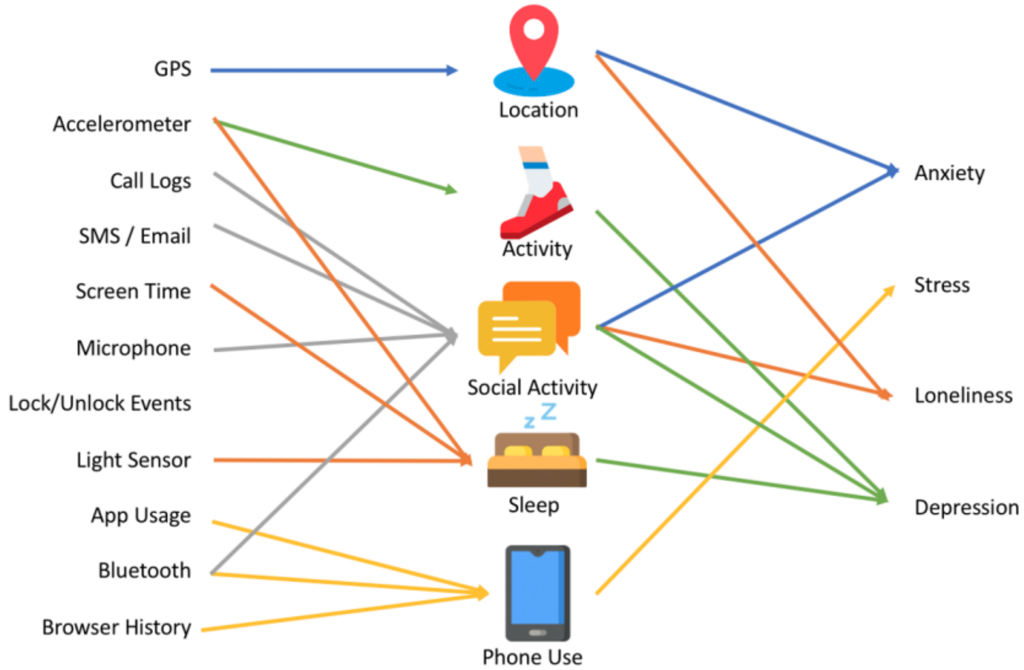
Once the ethical challenges are overcome, the correct application of Machine Learning can generate countless benefits for healthcare systems. For example, using this technology can potentially improve the cost-effectiveness of healthcare services, increasing the assertiveness of diagnoses and reducing the cost of this process (BEAM; KOHANE, 2016).

Regarding the application of ML on health, it is possible to find many examples for the most different health issues and user profiles. Recently, humanity has struggled against the effects of the COVID-19 pandemic. In this scenario, ML plays a fundamental role against the virus (KUSHWAHA *et al.*, 2020). KASSANIA *et al.* (2021) applied ML to detect COVID-19 in chest X-ray and CT images. In China, researchers developed intelligent applications to recognize people with fever quickly (FENG *et al.*, 2021). Other initiatives tried to find patterns in the spread of the pandemic (MALKI *et al.*, 2020) and to analyze the patient's behavior to find strategies to deal with future pandemics (PUNN *et al.*, 2020; OLIVEIRA *et al.*, 2021).

Another area of health that has received much attention is psychiatry. Recently, an increase in the prevalence of mental health problems was observed. Moreover, considering the post-COVID-19 pandemic context, humankind will undoubtedly face many challenges related to anxiety and depression (REINHART; REINHART, 2020). In this scenario, Big Data combined with Mobile Health, Machine Learning, and Deep Learning techniques can move the assessment of behavior, mood, and cognition from subjective clinical approaches to digital objective approaches (TOPOL, 2019).

Figure 6 represents how low-level mobile data such as GPS and accelerometer can be used to create high-level features (*e.g.*, social activity and sleep); and how these features can indicate the development of mental issues (*e.g.*, anxiety, stress, loneliness, and depression). However, to automate this process, it is necessary to use intelligent techniques. For example, PRIYA *et al.* (2020) use ML techniques to predict anxiety, depression, and stress.

Figure 6 – Correlation among mobile data and mental health (MELCHER *et al.*, 2020)



Source: MELCHER *et al.* (2020).

HUCKVALE *et al.* (2019) discuss opportunities in this area, such as the need to develop a data collection platform with a focus on equity, trust, and privacy; to propose shared data resources to accelerate collaborative research; to combine digital phenotyping with digital interventions; and to establish collaborations with healthcare professionals, providers, and computer science. These opportunities also motivate the development of this thesis.

2.3 Health and Quality of Life

Before any formal definition, having a better Quality of Life is probably the greatest desire of humankind. Naturally, this desire has driven the development of studies focused on improving people's QoL (BAKER *et al.*, 2017), mainly because there is a close relationship between health and QoL (GUYATT *et al.*, 1993).

Another factor that has been boosting the studies related to Quality of Life is the world's population aging (NATIONS, 2019). The United Nations projected that, in 2050, there will be 2.1 billion of older persons (*i.e.*, approximately 20% of the world's population will be aged 65 or over. In 2022, this percentage is nearly 10%). Although the aging population process is highly positive since it indicates the progress of society, new socioeconomic challenges emerge from the desire for a long life with health and well-being. Among them, it is possible to highlight the need for a healthcare system focused on preventive care (GMEINDER *et al.*, 2017).

The history of the Quality of Life term began a long time ago (ELKINTON, 1966; SPITZER, 1987). Even so, despite being discussed a lot, this term is confusing and can be observed from many perspectives (KARIMI; BRAZIER, 2016). For example, the Quality of Life can be related to the absence of chronic diseases, perception of loneliness, physical well-being, and, in the case of older adults, understanding of the aging and death process. In addition, some definitions include economic and political circumstances within QoL (KARIMI; BRAZIER, 2016). Thus, before presenting the Quality of Life definition adopted in this work, it is fundamental to detail the difference between health status, Health-related Quality of Life (HRQoL), and Quality of Life.

Health status can be defined as a state of complete well-being. This status goes beyond the absence of physical or psychological illnesses, including social well-being (WHO, 2014). In addition, the HRQoL considers those factors related to the individual's health (TORRANCE, 1987). Finally, although QoL has many definitions in the literature, most comprise objective and subjective assessments of physical, psychological, social, and environmental well-being (FELCE; PERRY, 1995).

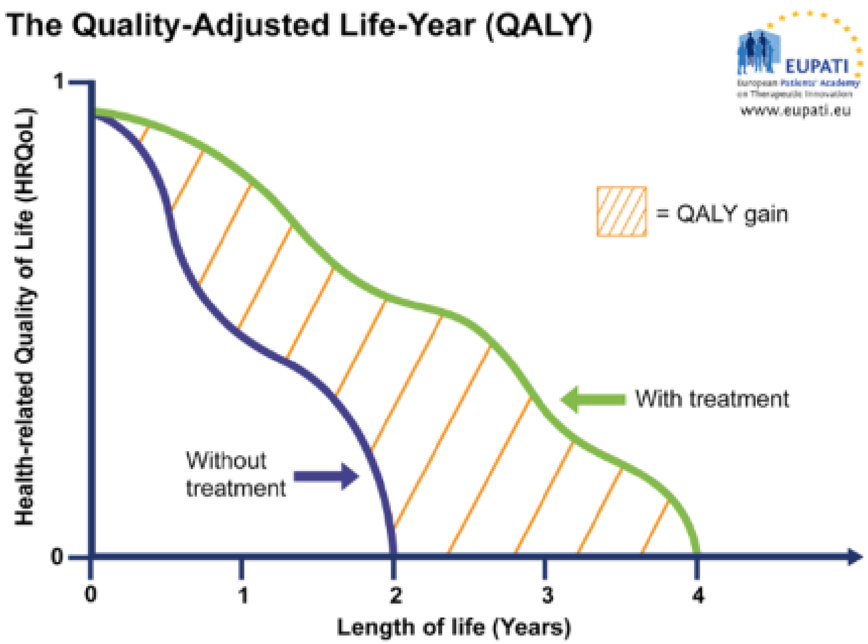
The World Health Organization's definition of Quality of Life was considered the primary reference in this work. For WHO, QoL can be described as the individual perception of life in a sociocultural context concerning goals, expectations, and personal standards (WHOQoL Group, 1994). This definition was adopted because of its relevance in the literature and because there is scientific evidence that contextual data collected through technological devices can draw a profile of physical, mental, and social health (HUCKVALE *et al.*, 2019).

From this definition, many mechanisms to evaluate QoL have been proposed (GROUP, 1990; JR; GANDEK, 1998). One of the most cited is the WHOQOL-BREF questionnaire (SKEVINGTON *et al.*, 2004) due to its reliability and cross-cultural validity. The WHOQOL-BREF was evaluated in 23 countries (including Brazil) and is available in 19 languages. It has 24 questions in four domains: physical, psychological, social, and environmental.

The “Physical" domain assesses motor facets such as daily living activities, medicine dependence, mobility, sleep quality, and work capacity. The “Psychological” domain relates to body image, negative and positive feelings, self-esteem, spirituality, and other mental health aspects. The “Social” domain observes personal relationships, social support, and sexual activity. Finally, the “Environment” domain aims to evaluate the environmental facets such as freedom, safety, security, participation in leisure activities, pollution, noise, traffic, and climate.

Unfortunately, the continuous application of this kind of instrument is tedious and bothersome (SANCHEZ *et al.*, 2015), which makes it challenging to engage the participants. Also, given the available techniques and tools, it is possible to state that seamless Quality of Life monitoring is complex, expensive, non-ubiquitous, and error-prone. It is complex because there are at least 150 strategies to measure QoL (GILL; FEINSTEIN, 1994), and it remains difficult to firmly define QoL (KARIMI; BRAZIER, 2016). It is expensive and non-ubiquitous because these strategies (usually questionnaires) should be self-rated by the patient or applied by a healthcare professional (BOWLING, 2005). Furthermore, it is error-prone since an active answer to a questionnaire can add a bias in the outcomes (CRANE *et al.*, 2016).

Figure 7 – QALY indicator proposed by EUPATI institution



Source: EUPATI website learning.eupati.eu.

The relevance related to QoL continuous monitoring emerges from the health benefits that can be achieved from up-to-date and accurate QoL information (*e.g.*, early interventions). Figure 7 shows the quality-adjusted life year (QALY) indicator proposed by the European Patients' Academy on Therapeutic Innovation (EUPATI), which is a multi-stakeholder public-private partnership committed to meaningful healthcare innovations. With this graph, it is possible to observe that the absence of treatments or medical interventions can reduce lifespan. Therefore, QoL monitoring is vital for a long and healthy life.

2.4 IoHT Challenges

A Systematic Literature Mapping (SLM) was conducted to get a comprehensive overview of IoHT applied to QoL (OLIVEIRA *et al.*, 2022a). This process resulted in 182 challenges grouped into eight categories. Each category is presented in a different section of this chapter. The main challenges of each category were highlighted in bold.

Intrinsic IoT Challenges

The Intrinsic IoT Challenges category, which had the most mentions (71) in the papers, encompasses well-known and intrinsic IoT challenges that have been studied for a long time. However, despite the advances, there is still no "killer solution" for them.

The **Lack of Interoperability** is a well-known difficulty for those who work with IoT. This challenge has many facets, and one of its major causes is the huge heterogeneity in IoT environments. It is possible to find devices from many different companies with specific protocols and data structures. Heterogeneity combined with the lack of widely accepted standards and vendor lock-in improves this problem. In this way, the usage of middleware platforms (*e.g.*, FIWARE[2]) and the health data standardization (such as, HL7[3] and FHIR[4]) are the most common solutions. ENSHAEIFAR *et al.* (2018) used a subset of FHIR to overcome the data interoperability in an IoT system for dementia care, and this kind of initiative has been strengthened by private sector initiatives like the Argonaut Project[5]. Other authors have proposed smart gateways capable of supporting syntactic, semantic, and technical interoperability among heterogeneous sensors (YACCHIREMA *et al.*, 2018) and solutions based on the idea of plug-and-play, extensible components (YAO *et al.*, 2018).

Trustworthiness is a critical issue in IoT systems, especially when these systems are responsible for people's healthcare. This aspect deals with the user's expectation of the service competency (ELMISERY *et al.*, 2019). Moreover, trustworthiness is a challenge in this context due to its close relationship with the data quality (LAURIE, 2019), privacy (ELMISERY *et al.*, 2019), and quality of network (CHUI *et al.*, 2019). The solution to this problem involves strengthening verification and validation techniques and fault tolerance strategies.

[2] FIWARE website: https://www.fiware.org
[3] HL7 website: https://www.hl7.org
[4] FHIR website: https://www.hl7.org/fhir
[5] Argonaut Project: https://argonautwiki.hl7.org/Main_Page.

In underdeveloped countries, medical services are restricted to wealthy people due to their high cost (AL-TURJMAN *et al.*, 2020). Then, the **Cost Efficiency** is the desired goal for IoT systems applied on remote health sections (ALBAHRI *et al.*, 2019). In this way, it is possible to observe many low-cost solutions. To name a few, FRANK; MENG (2016) designed a low-cost Bluetooth enables digital stethoscope to detect cardiac murmurs; AGUILERA-ASTUDILLO *et al.* (2016) created a low-cost 3D printed stethoscope that can be connected to a smartphone; KIM *et al.* (2012) developed a low-cost portable ultrasound system, and LUO *et al.* (2019) proposed a low-power sensor network to detect human activities.

Healthcare should be highly personalized (SHETH *et al.*, 2017). In this way, the **Personalized IoT-Health** is a challenge that has been more explored recently. AZIMI *et al.* (2017) stated, for example, that in general-propose elderly monitoring systems, several presumptions are made, and these presumptions can result in inefficiencies in the long term. Therefore, investigating these challenges can represent an opportunity to develop self-adaptive IoHT systems or reinforce the data analytics strategies to find the best setup for each user.

The **Standardization** is another well-discussed challenge in the IoT-related area. The definition of standards for data representation, data exchange, communication protocols, quality of service, development methodologies, and many others can potentially remove barriers to developing many IoHT solutions. In this way, there are initiatives such as the ISO/IEEE 11073 standards (YAO; WARREN, 2005) for point-of-care medical device communication; HL7, FHIR, and OpenEHR[6] standards for electronic health records; and the DICOM[7] for medical images.

Security and Privacy

Health data is extremely sensitive and private information about a patient. Thus, improper access can be harmful. The challenges found in this category are closely related to the three primary security goals: confidentiality, integrity, and availability (MERKOW; BREITHAUPT, 2014).

In this way, AL-TURJMAN *et al.* (2020) mentioned the **Data Security** as a critical requirement for IoHT systems, which is necessary both to develop new solutions to keep the data consistent and to train healthcare professionals to be aware of this criticality. Problems with the data can hinder decision-making regarding the treatment of a patient.

[6] OpenEHR website: https://www.openehr.org
[7] DICOM website: https://www.dicomstandard.org.

For the **Access control**, it was found mentions for the Identify Establishment and Capability-based Access Control (IECAC) protocol and the Elliptical Curve Cryptography (ECC) algorithm to protect the IoT from the man-in-the-middle, replay, and denial-of-service (DDoS) attacks (PAWAR; GHUMBRE, 2016). RAMU (2018) proposed a framework to preserve privacy in patient-doctor communication based on exchanging public and private keys.

For the **Confidentiality**, the PAWAR; GHUMBRE (2016) mentioned solutions using the Datagram Transport Layer Security (DTLS) protocol and cryptography based on symmetric encryption and elliptic curve. Other researchers have also investigated the usage of Blockchain for this purpose (KOTHA, 2020). There are also government initiatives to regulate data protection with laws. For example, the European Union has the General Data Protection Regulation law on data protection and privacy. Here in Brazil, there is the General Law nº 13.709/2018 on Protection of Personal Data (LGPD, in Portuguese "Lei Geral de Proteção de Dados Pessoais").

In addition, it was found papers mentioning specific issues, such as the problems with methods to de-identify data without introducing noise and re-identification attacks (ALTHOFF, 2017), and security issues in scenarios with heterogeneous resource-constrained devices (ONASANYA *et al.*, 2019). Considering this context, it is essential to adopt an expanded view of security and privacy in IoHT systems (ALKHATIB *et al.*, 2018).

Data Science

The Data Science category was mentioned 55 times, including challenges related to Big Data, Data Analytics, and the usage of Artificial Intelligence to support decision-making in healthcare. These three areas are fundamental for moving from reactive health, in which the diagnosis and treatment are defined in response to symptoms, to proactive health, which is focused on early warnings (data inference) using the data collected by sensors and other health objects (GOPAL *et al.*, 2019; ALTHOFF, 2017).

Regarding **Big Data**, it is estimated that 30% of the world's data volume is produced by the healthcare industry (FAGGELLA, 2018). In addition to this high volume of data, this data also has a great variety as it may involve medical images, monitoring vital signs, sleep data, location, medical notes, laboratory test results, and administrative health plan data. Furthermore, the veracity is related to the data quality and can impact clinical decision-making (FILHO; JUNIOR, 2017; FAGGELLA, 2018). Then, the volume, variety, velocity, and veracity are concerns that should be addressed during the IoHT development (KHODKARI *et al.*, 2018).

Concerning the **Data Analytic**, the challenge is to analyze the massive amount of data providing valuable information for the users. Usually, this involves data acquisition, filtering, cleaning and transformation, application of statistical methods and data mining algorithms, interpretation, and formatting of results (BANDODKAR *et al.*, 2016). In the results, it was found papers discussing new algorithms for data cleaning and filtering (CHEN *et al.*, 2009) and improvements in the mining approaches to deal with heterogeneous data (WU *et al.*, 2013). Also, LAURIE (2019) discussed the relevance to assure the quality of datasets, and ALTHOFF (2017) highlighted the data silos challenge.

The challenges related to **Artificial Intelligence (AI)** in healthcare can be summarized as the application of Machine Learning and Deep Learning to monitor patients, recognize user activities, and predict diseases. For example, WANG *et al.* (2014) has used logistic regression, and Artificial Neural Networks (ANN) for early detection of hypertension, HARISH *et al.* (2015) used a swarm intelligence method to diagnose an arrhythmia, SMITH; ROBERTS (2002) conducted a study with decision trees applied on the influenza treatment, SCHÄUBLIN *et al.* (1996) has used fuzzy systems to support mechanical ventilation during anesthesia, and Convolutional Neural Networks (CNN) been used in the treatment of brain tumor (PEREIRA *et al.*, 2016). However, as stated by AL-TURJMAN *et al.* (2020), there are still challenges to guarantee accuracy since false alerts or the absence of warnings in situations where the patient has a health issue are critical factors.

Network and Communication

Network and communication technologies are essential for developing any IoT-based systems. Naturally, challenges arise from this requirement. In the papers, it was found 50 mentions related to this category.

The most mentioned challenge was **Real-time** probably because latency is critical for healthcare systems. To tackle this issue, many strategies have been proposed, such as the usage of efficient processing units close to the sensors (DISI *et al.*, 2018), new system architectures to use concepts of edge and fog computing (SILVA; JUNIOR, 2018; YACCHIREMA *et al.*, 2018; ONASANYA; ELSHAKANKIRI, 2019), or adaptive data transmission policies using mist, fog, and cloud Computing (ASIF-UR-RAHMAN *et al.*, 2019). It is also expected that 5G technology will make profound changes in digital healthcare through its high-throughput, low-latency wireless connectivity (LATIF *et al.*, 2017).

Sensors and Wearable

This category of challenges has many opportunities due to the growing interest in wearables. ATHAVALE; KRISHNAN (2017) estimated that, in the future, wearables can represent 30% of healthcare, sports, and fitness tracking. The mapping result found 38 papers discussing challenges in this area, especially focused on resource constraints.

Regarding the **Device Resources Constraint**, energy consumption is the central issue (KHODKARI *et al.*, 2018). The limited power can restrict the transmission and processing capabilities, making seamless monitoring unfeasible. In addition, requesting the user to perform frequent energy charges can make it difficult to accept these technologies. As stated by CHONG *et al.* (2019), wearables should operate continuously with minimal human intervention, and adopting large-capacity batteries makes them uncomfortable. This context paves the way for developing protocols that support low energy consumption (FAHEEM *et al.*, 2019; AL-TURJMAN *et al.*, 2020) and for strategies to harvest energy(CHONG *et al.*, 2019).

Another relevant characteristic of wearables is their **User Acceptance** since this technology must be dressed. In specific cases, this requirement can be a barrier to its usage (*e.g.*, people with cognitive decline) (NEWCOMBE *et al.*, 2017). Thus, it is also essential to investigate less invasive methods to collect data (ATHAVALE; KRISHNAN, 2017).

Finally, it is worth mentioning that with the recent emergence of various fitness tracking and other medical devices, there is a demand for the development of approaches for **Testing and Validating** these devices and their mobile applications (GOPAL *et al.*, 2019).

Software Engineering

The most mentioned software engineering challenges were **reliability** (SILVA; JUNIOR, 2018), **reconfigurability** (GATOUILLAT *et al.*, 2018; RAHMANI *et al.*, 2015), **easy installation** (MARQUES; PITARMA, 2019), **legacy systems and technical debts** (GOPAL *et al.*, 2019), **system compatibility** (DAUWED *et al.*, 2018a), and **fault tolerance** (ALBAHRI *et al.*, 2019). In this way, many solutions have been proposed. For example, reliability and fault tolerance can be addressed with testing techniques and adaptive models for this kind of system. Reconfigurability requires design solutions; the legacy systems and technical debts can be mitigated by decoupling the data from the legacy system and adopting debt management processes. Finally, many solutions use middleware to overcome heterogeneity issues.

Human-Computer Interaction

Fourteen (14) papers mentioned the following Human-Computer Interaction challenges: **usability** (AL-TURJMAN *et al.*, 2020; ENSHAEIFAR *et al.*, 2018; HUYNH *et al.*, 2020), **non-invasive care technologies** (DOBRE *et al.*, 2019; GOMEZ-CARMONA *et al.*, 2018), **empowered users** (KOTHA, 2020), **engagement in health interventions** (TUN *et al.*, 2021), and **acceptance and learning of the elderly** regarding technologies (GKOUSKOS; BURGOS, 2017). As the adoption of IoHT solutions increases, interest in technologies that provide a better user experience should also grow, leading to studies focused on the impact of functional and non-functional requirements in that experience.

Cloud Computing

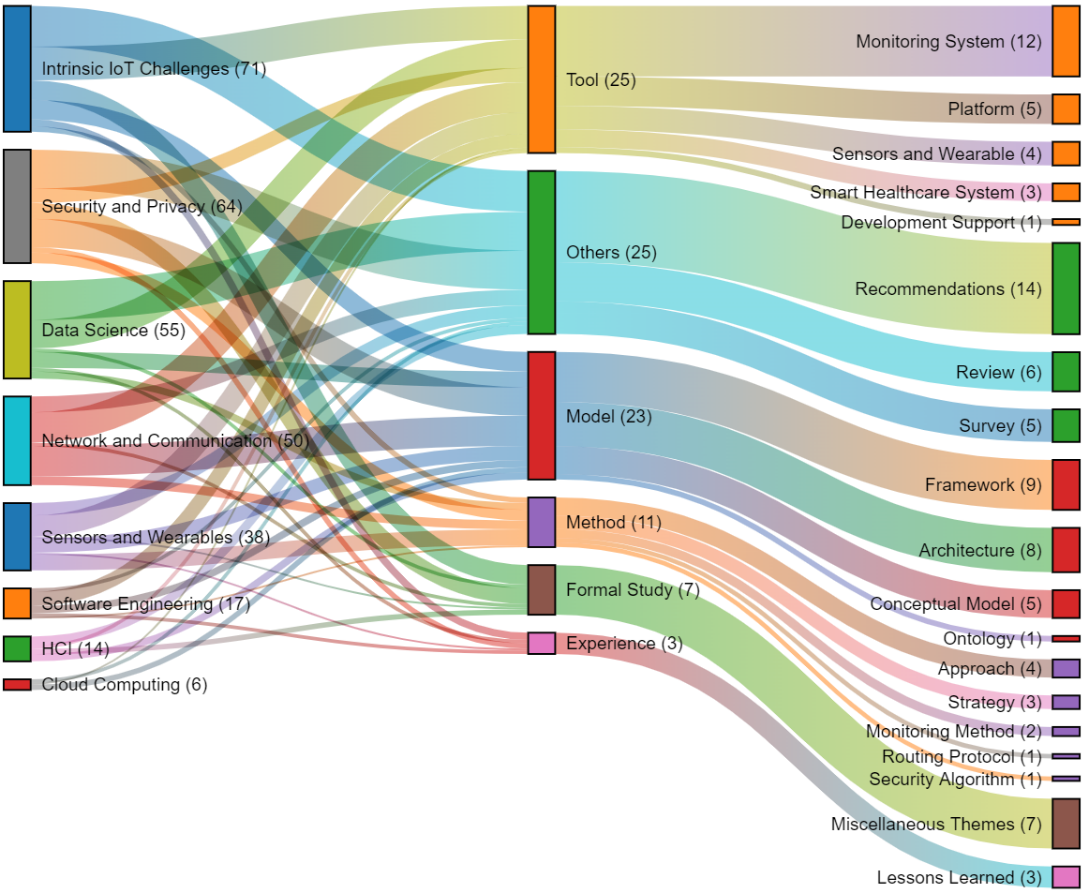
Six (6) papers have mentioned Cloud Computing challenges. For example, the **complexity of integration and management of different layers of Cloud and the Internet of Things** for healthcare systems (KHODKARI *et al.*, 2018), **delay** in cloud computing (PAZIENZA *et al.*, 2020), **offloading** (MANO *et al.*, 2019), the **usage of fog computing** in IoT-Health (PRIYADARSHINI *et al.*, 2018), and the **synchronization between different cloud vendors** (ABDELNAPI *et al.*, 2018). These issues can be solved with more effective system architectures, but it is fundamental to evolve the strategies to approximate the cloud capabilities to the users. This can be done by investigating fog, mist, and edge computing or even with offloading approaches.

2.5 Discussion

The chapter presented the main concepts and challenges associated with this thesis. To conclude the analysis regarding the challenges, a correlation is presented among the eight IoHT challenge categories, the contribution types (tool, model, method, formal study, experience, and others), and the proposed solutions.

Figure 8 shows a Sankey diagram (SCHMIDT, 2008), in which it is possible to observe this correlation considering the papers found. Most solutions are monitoring systems or platforms, but many papers present recommendations, reviews, and surveys. 24.5% of the works present models, 11.7% describe methods, 7.4% are formal studies on miscellaneous topics, and 3.2% bring reports of experiences.

Figure 8 – Correlation among the challenge, contribution type, and proposed solution



Source: author.

Finally, based on the investigation conducted to build this chapter, the answer to Research Question 1 (RQ1) can be drawn as follows:

RQ1: What prior knowledge is available about the IoHT and its application to Quality of Life?

Based on a Systematic Literature Mapping, it was found 182 challenges grouped into eight categories. These eight categories represent research areas with opportunities regarding QoL-based IoHT systems. They are Intrinsic IoT Challenges (71 mentions), Security and Privacy (64), Data Science (55), Network and Communication (50), Sensors and Wearables (39), Software Engineering (17), Human-Computer Interaction (14), and Cloud Computing (6).

Among this large set of challenges, it is possible to highlight an increased interest in personalized IoHT applications, data security, wearables to monitor patients, and Machine Learning to predict health issues. Also, the strengthening of mobile health is expected due to the cost reduction of devices and increasingly reliable solutions. Naturally, this strengthening will demand new software engineering methods, mainly focused on testing and systems' usability.

3 RELATED WORK

After defining the thesis focus and conducting an in-depth study into IoHT challenges for QoL, it was possible to conduct an investigation to identify studies related to this thesis. This chapter presents the methodology to find these studies (Section 3.1), the results (Section 3.2), and a final discussion about the related work (Section 3.3). This review was also published in the 15th International Conference on Health Informatics (OLIVEIRA *et al.*, 2022b).

3.1 Methodology

In order to compose the related work, a literature review was performed on papers indexed on the Elsevier Scopus database. This database was selected based on its coverage[1] of software engineering venues and relevant digital libraries such as ACM Digital Library, IEEE Xplorer, Science Direct, and Springer Digital Library. Thus, the selected papers represent a suitable sample to describe this study area. Also, it is important to mention that it was not included any date restriction.

The search string was composed of the following terms and their synonyms: "*smart quality of life, passive sensing, internet of health things, platform and machine learning*". The first two terms were included to ensure the retrieval of three control papers (previously identified by the authors). The other terms are directly associated with the final goal of this thesis: to build a platform to support the development of IoHT systems that use Machine Learning to monitor users' Quality of Life.

Initially, 122 papers were found, but in the end, only 13 were chosen after full reading. The eligibility criteria were: be a primary study, written in English, fully available on the Web, with more than five pages, published in conferences or journals, and present a discussion about IoHT solutions to monitor QoL.

During this literature review, the challenge of defining keywords that characterize studies focused on monitoring Quality of Life indicators was faced. This challenge arises because the term QoL is used in many contexts related to applying technology to specific health issues. For example, the paper presented by GREENE *et al.* (2016) discusses an IoT solution for fall detection. Throughout the paper, the authors mention that the proposed system can improve older adults' QoL. Although this paper has the terms IoT, Quality of Life, and monitoring, it does

---

[1] Scopus Coverage: elsevier.com/?a=69451.

not focus on monitoring users' QoL. Due to this context, most papers (104) were not accepted, because they did not discuss solutions to monitor QoL indicators. In addition, it was identified two short papers, one duplicate study, and two reviews. To sum up, 109 papers were rejected.

3.2 Results

Before starting the discussion of the related work, it is essential to reinforce the difficulty in building a search string capable of differentiating studies focused on using IoHT systems to infer users' QoL from the studies that bring specific solutions for health issues. Generally, both kinds of studies use the terms Quality of Life, and monitoring.

Based on this difficulty, the need for a keyword that can represent the research area targeted by this thesis is clear. For this purpose, the term "Smart Quality of Life" is a suitable candidate. However, the first authors[2] to use this expression in the context of QoL-based studies did not provide a formal definition for this term (QIU *et al.*, 2020). Then, a formal definition inspired by the WHO statement (WHOQoL Group, 1994) is presented as follows.

Smart Quality of Life can be described as a person's Quality of Life inferred from individual and contextual data acquired using ubiquitous and less-intrusive technologies.

Usually, this Smart Quality of Life indicator is built through Machine Learning. Over time, it can be used to early detect health issues, representing an important tool in medical practice. Now, in light of this definition, it is possible to discuss the related work.

Table 1 – Case studies selected in the literature review.

Work	Contribution	QoL Domain	Profile	# of Part.	Env.	Devices	Analysis
BADE *et al.* (2018)	Longitudinal study	Physical	Patients with Lung Cancer	30	Not specified	Fitbit Zip and Smartphone	Spearman rank correlation
KIM *et al.* (2019b)	Longitudinal study	Physical	People with spinal issues	22	Hospital	Fitbit Charge	Pearson correlation and regression analysis
LEE *et al.* (2019a)	Longitudinal study	Physical	Fibromyalgia patients	14	Not specified	Specific wearable built for this study	Statistical analysis
ANGTHONG; VELJKOVIC (2019)	Longitudinal study	All	Adults with foot- and ankle-related conditions	52	Not specified	Foot pod (Garmin)	Pearson's correlation
OLIVEIRA *et al.* (2021)	Longitudinal study	All	Lymphoma patients	16	Not specified	Microsoft Band 2	Statistical analysis
CONCHEIRO-MOSCOSO *et al.* (2021)	Study Protocol	Physical	Adults	11	Work	Xiaomi Mi Band 3	Statistical analysis
BRUDY *et al.* (2021)	Longitudinal study	All	Children with congenital heart disease	343	Not specified	Garmin vivofit Jr	Logistic regression

Source: author.

[2] Other authors have used this term, but in the context of smart cities.

Tables 1 and 2 summarize the 13 selected papers and the proposed work. Table 1 brings only the case studies and, although these case studies do not present a tool for monitoring Quality of Life as their primary contribution, they discuss approaches to correlate data obtained from IoHT devices with QoL indicators. Among the characteristics of these studies, it was highlighted their main contribution, QoL domain, patient profile, number of participants in the assessment, study environment, device used in data collection, and data analysis strategy.

The studies presented by BADE *et al.* (2018), KIM *et al.* (2019b), LEE *et al.* (2019a), ANGTHONG; VELJKOVIC (2019), OLIVEIRA *et al.* (2021), and BRUDY *et al.* (2021) were classified as longitudinal studies because they involve analyzing the participants' data through an extended period in order to prove the correlation between health data and the patient's QoL. Although these studies do not present software artifacts as the main contribution, their discussion is relevant to indicate strategies for evaluating solutions that use health data to infer the Quality of Life of their users. All these works (excluding only the study conducted by (LEE *et al.*, 2019a)) used commercial smart bands and their native applications. This decision is probably related to the costs of these devices. In general, devices with higher processing power that allow the development of native apps for their platforms are expensive. Another difficulty observed in these studies is the absence of APIs for data extraction, which makes it an arduous process.

Regarding the data analysis, all of these longitudinal studies present statistical analyses to validate their hypotheses. BADE *et al.* (2018) and OLIVEIRA *et al.* (2021) proved that there is a correlation between physical activity data and the QoL of people with cancer. Similarly, KIM *et al.* (2019b) showed this same correlation for hospitalized patients with spinal issues; LEE *et al.* (2019a) for patients with fibromyalgia; ANGTHONG; VELJKOVIC (2019) for adults with foot-ankle, and BRUDY *et al.* (2021) for children with congenital heart disease.

The results of these studies can be generalized to state that **it is possible to use data collected by smart objects to measure the Quality of Life of patients even with different QoL questionnaires and for different health conditions**. This opportunity has also been reinforced in renowned medical journals (HUCKVALE *et al.*, 2019). However, unfortunately, none of these studies made their datasets publicly available, which hinders the advancement in this study area.

To conclude this first group of works, CONCHEIRO-MOSCOSO *et al.* (2021) brings a protocol to assess the impact of stress on workers' QoL. Their main contribution is to present a guide on conducting studies that seek to correlate health data with QoL facets. In this investigation, they started by collecting socio-demographic data from the participants.

Then, to collect health indicators, participants will wear a wearable called Xiaomi Mi Band 3 for six months collecting steps, sleep, and heart rate data. Finally, five questionnaires will be periodically applied: EuroQol-5D-5L, to assess the Quality of Life; Pittsburgh Sleep Quality Index for sleep habits; State-Trait Anxiety Inventory for anxiety; Perceived Stress Scale-10 for stress; and, finally, the Stress Questionnaire to assess stress factors at work.

The second group of studies (listed in Table 2) brings models, methodologies, frameworks, systems, architectures, and platforms as their primary contribution. In addition to the characteristics selected for the first group of studies (Table 1), it was also decided to include aspects related to the IoHT challenges identified in the previous literature review (see Chapter 2): how they deal with device heterogeneity, the use of intelligent algorithms for data analysis, and if it is possible to define strategies for adapting users' environment.

Table 2 – Comparison between the platform proposed in this thesis (Healful) and the tools, methods, and models selected in the literature review.

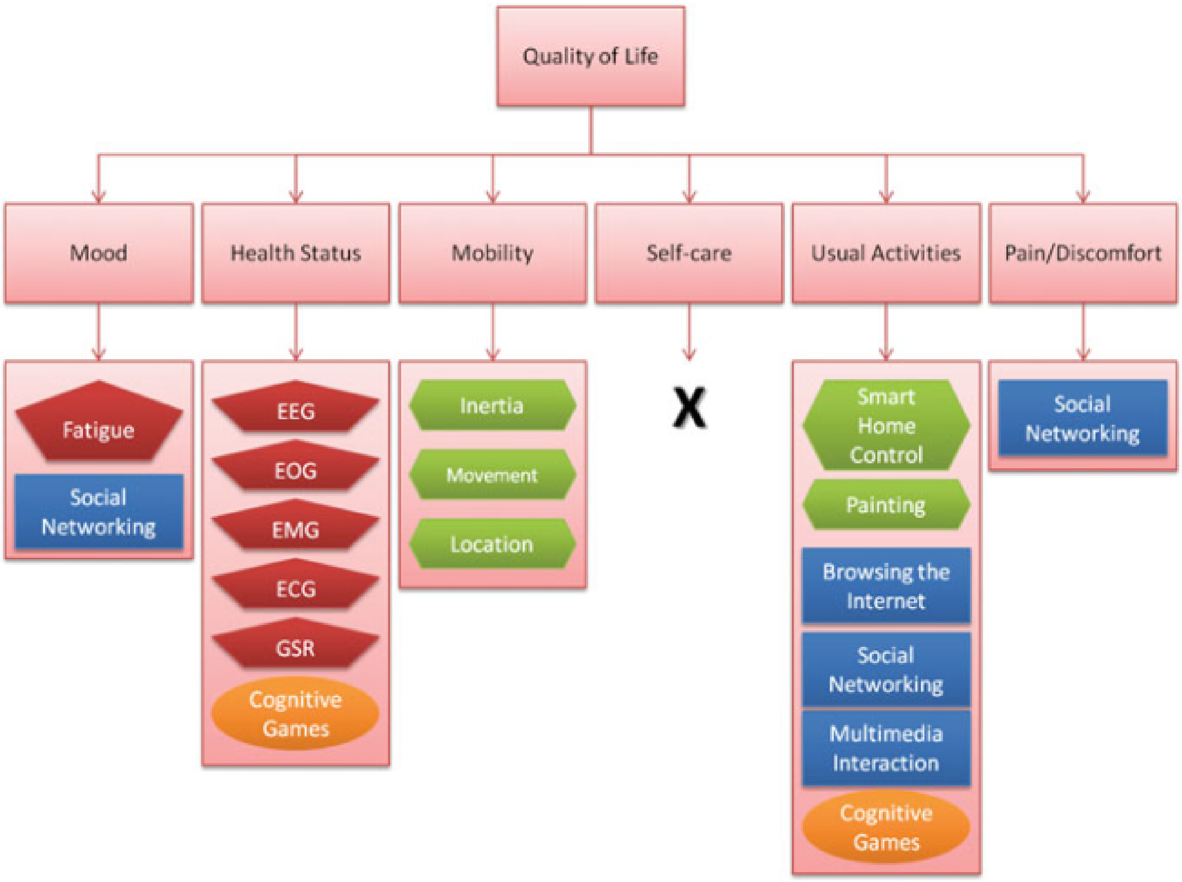
Work	**Contribution**	**Deal with hetero-geneity?**	**Provide AI for data analysis?**	**Allow environmental adaptation?**	**QoL Domain**	**Profile**	**# of Part.**	**Env.**	**Devices**	**Analysis**
MERILAHTI *et al.* (2012)	Model	No	Yes	No	Physical	Older adults	19	Not specified	Actigraphy, Bed sensor, Omron Walking Style II pedometer and Omron 705IT	Spearman correlation and k-means clustering
VARGIU *et al.* (2014)	Methodology	Yes	Yes	No	All	People with disabilities	N/I	Home	Brain/Neural Computer Interface (BNCI), inertial sensors, environmental sensors, smart home devices	C4.5 and k-NN
BONO-NUEZ *et al.* (2014)	System	No	Yes	No	Physical	Older adults	N/I	Smart Kitchen	Kitchen appliances, Zigbee sensors, RFID and portable devices	Self-Organizing Maps (SOM)
MASI *et al.* (2016)	Platform	No	Yes	No	All	Generic	N/I	Indoor and outdoor	Smartphone and Wearables	User data timeline
DOBRE *et al.* (2019)	Architecture	Yes	No	No	All	Older adults	N/I	Indoor and outdoor	Smartwatch, smart shoes, camera	Statistical Analysis
RADULESCU *et al.* (2019)	Framework	No	No	No	All	Older adults	17	Not specified	Not specified	Spearman correlation
Healful	Platform	Yes	Yes	Yes	Physical and Psycho-logical	Adults	44	Indoor and Outdoor	Smartphones and Wearables	Machine Learning

Source: author.

In MERILAHTI *et al.* (2012), the authors present a study about the performance of health monitoring technologies to estimate the physical function of older adults. In addition, they offer a hypothesis that health data would predict pre-clinical measures. Thus, nineteen (19) older adults were analyzed through eighty-four (84) days using wrist-worn activity monitors, bed sensors, pedometers, weight scales, and blood pressure monitors. The acquired raw data were transformed into sixteen (16) features, and they were analyzed using statistical correlation (Spearman rank correlation and linear regression) and clustering methods (K-Means method).

The results presented by MERILAHTI *et al.* (2012) were not promising, indicating only a correction with the daily steps. However, the authors bring interesting insights about which features (*e.g.*, daily steps, weight, systolic and diastolic blood pressure, mean of heart rate at night, breathing frequency, sleep duration) can be used in this type of investigation and points out issues in data collection. For example, they faced problems with automated data collection, such as data loss and missing values. This issue indicates that, in real scenarios, strategies must be used to overcome data collection and recording difficulties. Compared to this thesis, MERILAHTI *et al.* (2012) present only a specific model for the physical domain, which does not consider self-assessment questionnaires and does not concern itself with other aspects of IoHT (such as heterogeneity and environmental adaptation).

Figure 9 – Relation among user context and QoL aspects



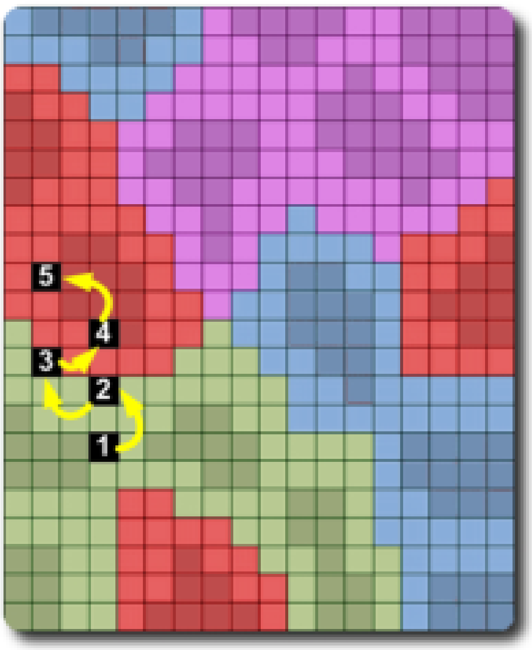
Source: VARGIU *et al.* (2014).

VARGIU *et al.* (2014) propose a context-aware methodology to telemonitor QoL concerning the physical and social autonomy of people with disabilities. Thus, they adjusted the EQ-5D-5L questionnaire to assess mood, health status, mobility, self-care, usual activities, and pain/discomfort. On the other hand, health data were collected by Brain/Neural Computer Interface (BNCI), inertial and environmental sensors, and smart home devices. The authors achieved good results using the C4.5 and k-NN algorithms. Nonetheless, synthetic data was used due to the lack of real data. Similar to this thesis, the authors seek to evaluate the relationship between the user's context and the Quality of Life assessment.

In Figure 9, VARGIU *et al.* (2014) present the relationship between QoL aspects and data collected from users. It is possible to observe that their methodology is quite extensive and ambitious because it requires a significant and varied amount of data to be validated. In addition, the authors differ from this thesis because it focuses only on data monitoring.

BONO-NUEZ *et al.* (2014) focused their contribution on creating a QoL evaluation system to support the work of caregivers. The idea is to periodically provide QoL assessments of older adults to help with the decision-making of caring actions. The authors reinforce the relevance of observing the QoL to assist decision-making about medical care and that these decisions are based on subjective observations, interviews, or complex metrics such as the International Classification of Functionalities. Given this problem, using IoHT data emerges as an excellent alternative.

Figure 10 – A SOM representing a health deterioration



Source: BONO-NUEZ *et al.* (2014).

Unlike other studies, BONO-NUEZ *et al.* (2014) did not choose a QoL questionnaire as a reference. Instead, they decided to cluster the data using Self-Organizing Maps (SOM). Nevertheless, the authors focused on smart kitchens, requiring the analysis of a domain expert to interpret the results obtained by the SOM. Figure 10 brings the interpretation of a user's health deterioration process. Initially, in the first two evaluations, their data were in the green cluster, representing good health status. In the following evaluations (3, 4, and 5), the result went to the red cluster representing cognitive problems. Finally, this study differs from this thesis because the Healful platform is able to create a model for the Quality of Life inference that does not rely on domain experts for result interpretation.

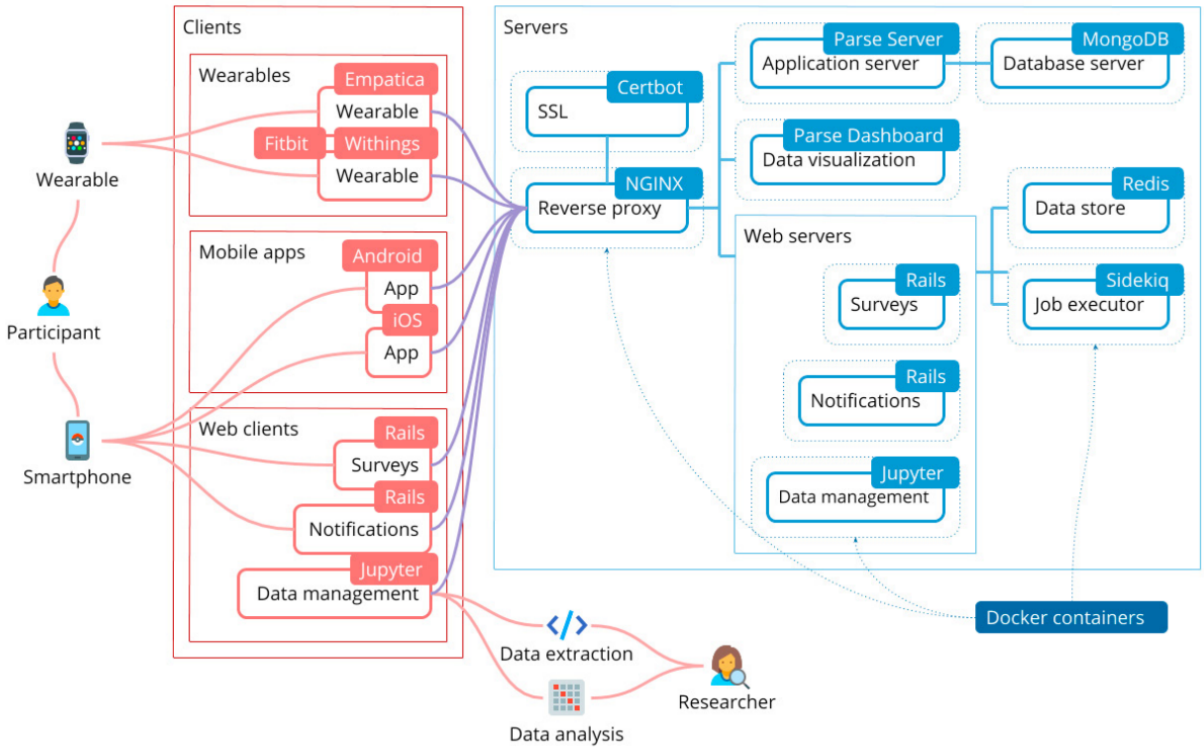


Figure 11 – mQoL Living Lab architecture.
Source: BERROCAL *et al.* (2020).

Based on the explored investigations, the work proposed by MASI *et al.* (2016) is probably the most related to the platform presented in this thesis. The significant difference is that MASI *et al.* (2016) built a platform to support interdisciplinary mobile health studies related to Quality of Life. In contrast, this thesis delivers a platform able to collect and analyze IoHT data, aiming to infer users' Quality of Life through Machine Learning algorithms. In addition, the Healful platform integrates built-in Machine Learning algorithms and adaptation rules for healthcare interventions.

Despite the limitations and differences regarding this thesis, studies conducted by QoL Lab[3] has insightful ideas about the inference of Quality of Life measures. MASI *et al.* (2016) present the first version of the mQoL Living Lab in their paper. However, this platform is not publicly available, and it was not possible to conduct a more in-depth evaluation. Even so, looking at other authors' studies, the mQoL Living Lab architecture (Figure 11) was found. The mQoL Living Lab architecture allows correlating data collected from smartphones and wearables with information from health questionnaires. However, in mQoL Living Lab, researchers need to code the data processing using the Jupyter[4] tool.

DOBRE *et al.* (2019) propose an IoT architecture to deliver non-intrusive monitoring and support older adults' healthcare. One of the most interesting points lies in the authors' concern about inherent Internet of Things challenges, such as interoperability. The architecture was designed with a modular structure and, similar to the work proposed by MASI *et al.* (2016), the data analysis module aims to conduct scientific studies. However, the authors do not address intelligent techniques to infer QoL and strategies to act in the environment.

Finally, the study proposed by RADULESCU *et al.* (2019) brings a framework to find a correlation of health parameters with QoL questionnaires. The authors deal with this problem using mathematical models. They selected the Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS), which uses the concept of "ideal" and "anti-ideal" solutions to find an overall health index for older adults. Nonetheless, the method was evaluated only with synthetic data, and its complexity makes its adoption challenging.

3.3 Discussion

Each of the studies analyzed in this chapter contributes to the search for methods for monitoring patients' QoL. This research area becomes increasingly critical as the need for disease detection and early healthcare interventions increases. In a transversal way, this investigation also helped to identify which data can be collected to understand users' health context.

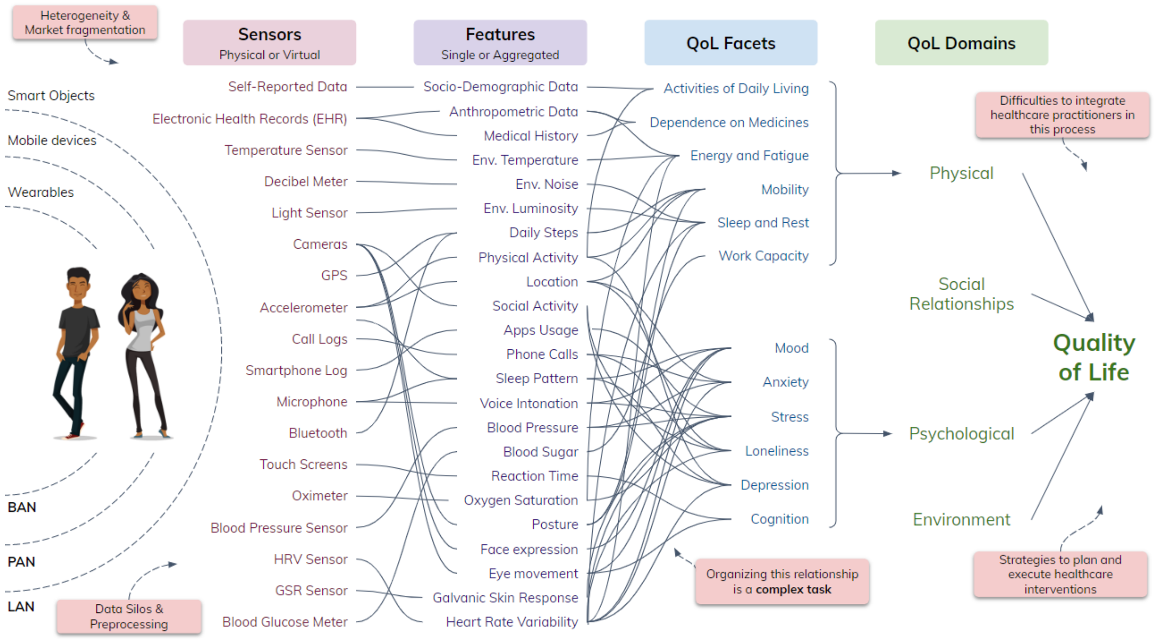
The first conclusion drawn from this analysis is the need for a keyword able to aggregate such investigations. Regarding this topic, it was found a suitable candidate, which is "Smart Quality of Life". As a result, Smart Quality of Life was defined using an adaptation of the WHO QoL definition (WHOQoL Group, 1994), aiming to include concepts of context-sensitivity,

[3] QoL Lab website https://www.qualityoflifetechnologies.com.
[4] Jupyter website: https://jupyter.org.

ubiquity, and intelligent algorithms. Thus, **Smart Quality of Life** can be described as a person's Quality of Life inferred from individual and contextual data acquired using ubiquitous and less-intrusive technologies.

The second conclusion reached concerns the case studies found in this review. Based on the results of these studies, it is possible to affirm there is a direct correlation between users' physical data and their Quality of Life. Furthermore, such studies provide good examples of evaluating proposals related to the automation of QoL monitoring. On the other hand, they highlight the existing difficulties in performing investigations in this area. In general, longitudinal studies (a strategy commonly used in these cases) are long and expensive; real-world data collection (non-laboratory scenarios) is susceptible to many different issues, and there is a need for public datasets to train smart models.

Figure 12 – List of IoHT sensors and features related to QoL facets



Source: author.

The third conclusion identified refers to the opportunities in developing platforms that enable the collection of health data, the processing of these data using intelligent algorithms, the planning of interventions in risky situations, and the adaptation of the user environment to provide better living conditions. In short, a closed loop of healthcare that seeks to identify health issues as soon as possible. In addition, using the knowledge present in the analyzed papers and the discussion conducted by OLIVEIRA *et al.* (2022), it was possible to summarize a list of sensors, features, and their relation with QoL facets and domains.

Figure 12 shows these potential relationships among sensors, features, facets, and domains. Also, it is possible to observe a relationship with the challenges identified in the systematic mapping presented in Chapter 2. For sure, the relationship between the sensors, features, and QoL facets presented in the Figure is challenging to understand due to the high number of relationships. This reinforces the need to use intelligent techniques to deal with them. Finally, from the results of this investigation, it is possible to draw an answer for Research Question 2 (RQ2) as follows.

RQ2: Which data can be ubiquitously obtained from commercial Internet of Health Things devices to represent users' health context?

There is an extensive list of physical and virtual sensors present in Body Area Network (BAN), Personal Area Network (PAN), and Local Area Network (LAN) networks. These sensors generate data that can characterize socio-demographic context, anthropometric information, medical history, physical activities, location, app usage, sleep pattern, posture, gait pattern, heart rate variability, and many others. Furthermore, each of these data can be used to understand different QoL facets. For example, daily steps and user location are strongly related to daily living activities. In addition, sleep patterns and the usage of smartphone apps can be related to stress and anxiety.

4 HEALFUL PLATFORM

JAKUBCZYK; GOLICKI (2018) wrote that "*optimizing health services requires measuring health*". In light of this reasoning, it is possible to state that measuring health in a modern and highly dynamic society requires context awareness, ubiquity and intelligent algorithms. Thus, this thesis presents a solution based on the Internet of Health Things (IoHT) and Machine Learning, capable of inferring users' Quality of Life from Smart Devices data. This solution was titled Healful Platform, and the healful term was selected due to its relation to health-promoting.

This Chapter is then outlined as follows: Section 4.1 describes the analytical model developed to guide the thesis; Section 4.2 details the Healful platform and presents a running example; Section 4.3 puts light on how the QoL inference model was created and, to conclude this chapter, Section 4.4 presents how to calculate the health indicators associated with this work.

4.1 Research Analytical Model

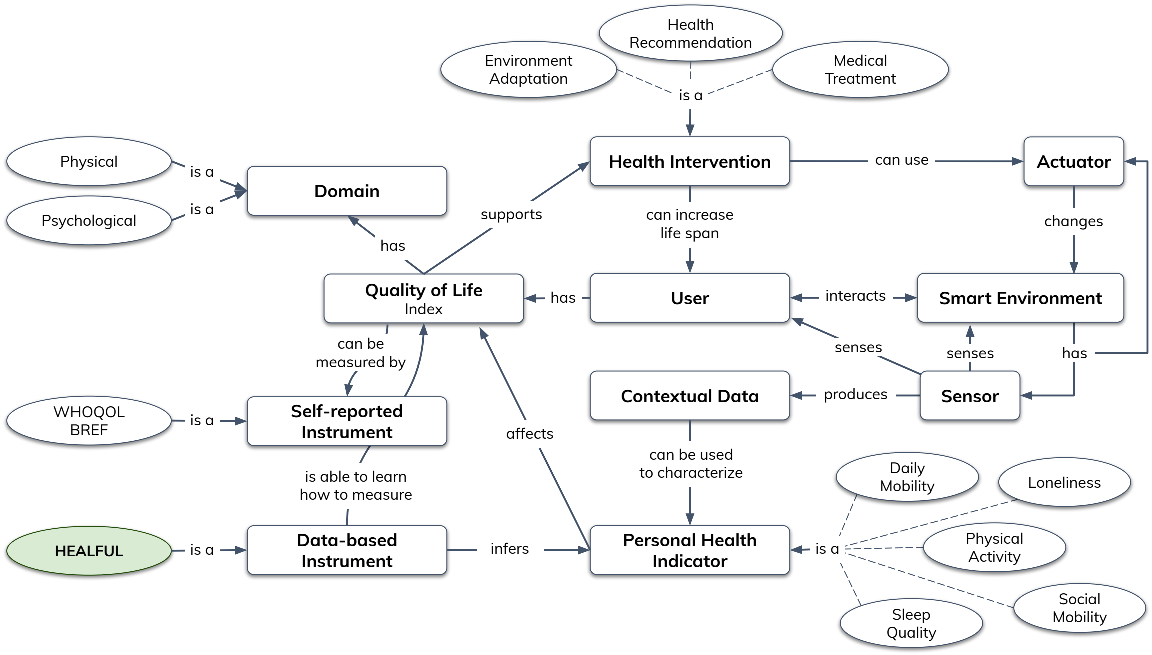
According to John von Neumann[1] "*a model is a human construct used to describe an observed phenomenon*" (AZUAJE; DOPAZO, 2005; MALEK, 2008) and these models "*can be used to understand complex real-world situations, providing a basis for effective problem solving*" (ABRAN, 2014). In this way, analytical models use logical reasoning to model entities of a system and specify their relationships (FRIEDENTHAL *et al.*, 2015).

This thesis includes an analytical model created to guide and delimit the research investigation, which was inspired by a previous study that brings a structured description for mHealth applications (OLIVEIRA *et al.*, 2023). Figure 13 shows this model representing entities (highlighted in the text using the underline) as rectangles and their relationships with labeled arrows. Also, instances of entities are represented as ellipses.

At the center of this model is the User entity, which can interact with the Smart Environment. In this scenario, Smart Environments are characterized by Sensors and Actuators used to enhance users' lives (OLSON *et al.*, 2015). A sensor produces Contextual Data from user and environment sensing. With these data, it is possible to represent many Personal Health Indicators. This work considers five indicators: daily mobility, physical activity level, sleep quality, loneliness level, and social mobility level (Section 4.4).

[1] John von Neumann's biography: mathshistory.st-andrews.ac.uk/Biographies/Von_Neumann

Figure 13 – Research Analytical Model

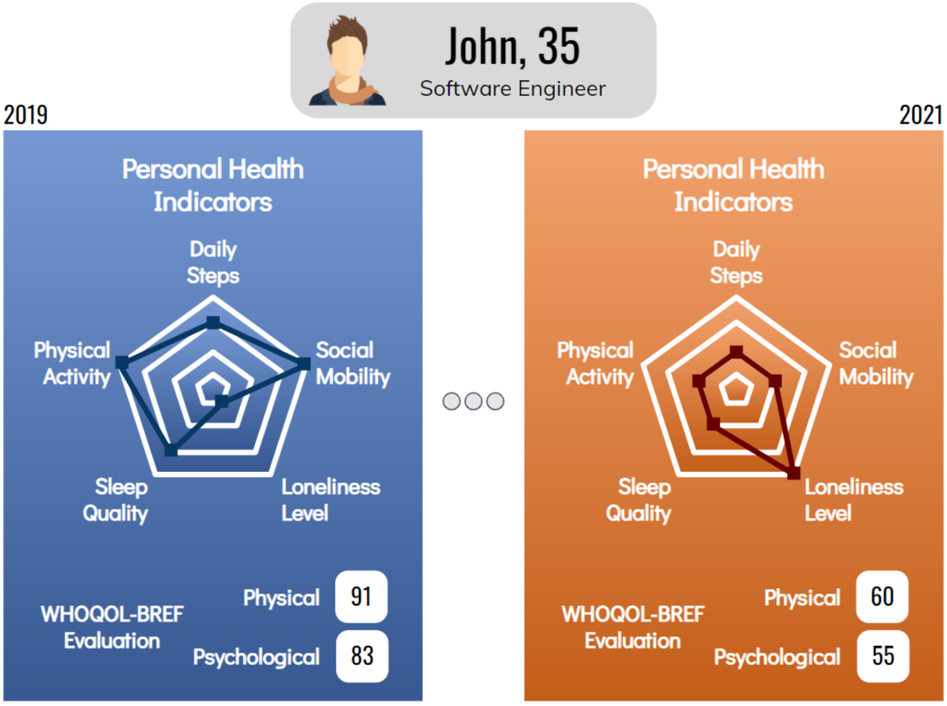


Source: author.

Continuing with the detailing of the analytical model, QoL has many Domains. However, this work focuses on two domains: physical and psychological. These two domains were selected due to their strong influence on the patient's health and the availability of data to characterize each of them. Based on this decision, QoL scores can be measured using instruments such as Self-reported Instrument (*e.g.*, WHOQOL-BREF (WHOQoL Group, 1994)). However, as an alternative strategy, it is proposed in this work a Data-based Instrument that uses Machine Learning algorithms to infer users' QoL from contextual data. Supported by the QoL scores, healthcare professionals can define Health Interventions such as adaptations to the user environment (when it is provided with smart actuators), periodic health recommendations (to promote changes in users' habits), or medical treatments (in critical cases). Such interventions can enhance user well-being, increasing life span.

Figure 14 provides a scenario to show the representativeness of the analytical model. In this scenario (previously presented in Chapter 1), John represents a 35-year-old software engineer engaged in physical activities, regularly communicated with friends and family, and slept well. However, after the COVID-19 pandemic, John faced difficulties maintaining his healthy habits and gradually began to have problems with sleep and loneliness. If a questionnaire had been applied to assess John's QoL in 2019, the physical and psychological scores would be high. Nevertheless, in 2021, it would bring different results.

Figure 14 – Illustrative scenario to show how the proposed analytical model can be applied



Source: author.

While in 2019, the scores were good, in 2021, they indicated an alert situation. Keeping these indicators at low levels can lead to severe health problems such as depression and chronic anxiety (KNIGHT *et al.*, 2020; DAHLBERG *et al.*, 2022). Thus, the ideal scenario would be to continuously monitor these health indicators to alert John as soon as possible about the worsening of his indicators. However, the continuous application of self-reported questionnaires (an intrusive and non-ubiquitous kind of monitoring) makes it challenging to engage the users (SILQUEIRA, 2005; SANCHEZ *et al.*, 2015; PEQUENO *et al.*, 2020).

This illustrative scenario can also be analyzed considering the perspective of public and private health managers. The growing demand for health services has forced searching for methods to optimize resources, promoting better healthcare services (FRØNSDAL *et al.*, 2010). John's worsening health impacts his use of health services (KARYANI *et al.*, 2016). Thus, continuous health monitoring could trigger early alerts to anticipate interventions, such as promoting preventive care.

This context reinforces the need for less-intrusive QoL monitoring to identify situations that tend to worsen with unhealthy habits. This need is addressed in this work by the Healful platform, which uses the analytical model of Figure 13 to systematize a service capable of inferring the user's QoL.

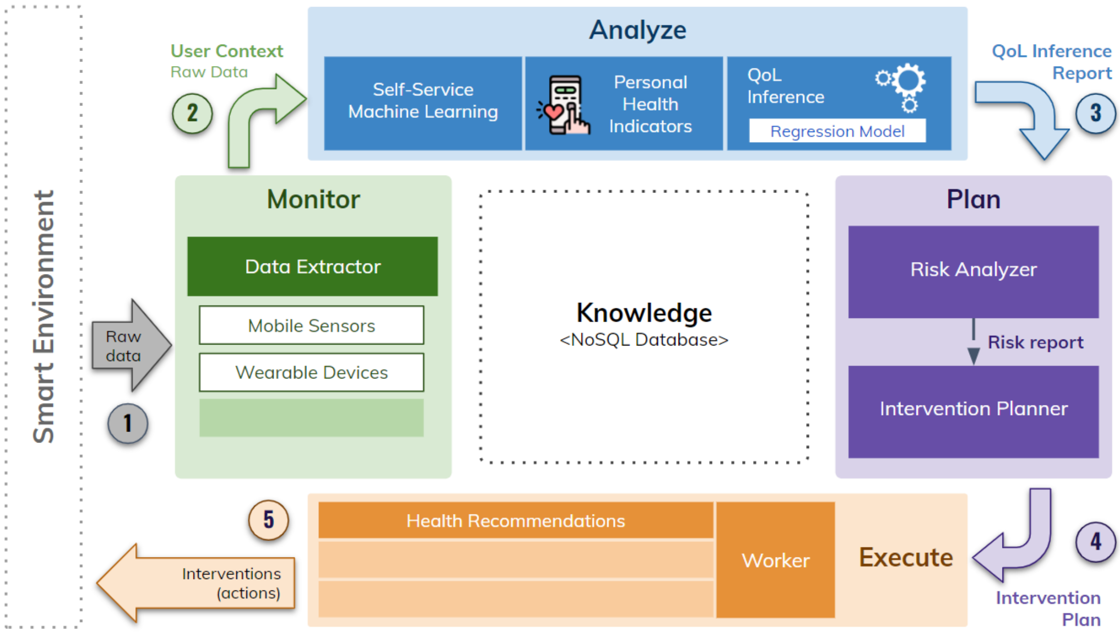
4.2 Healful platform for QoL-based systems

The Healful platform was developed to coordinate all software artifacts employed to achieve a less intrusive, continuous and ubiquitous Quality of Life monitoring. This section presents an overview of the platform structure, an architectural view, and details about the QoL Monitor app. Finally, Appendix E presents a running example of the platform.

Overview

The platform was inspired by the MAPE-K loop framework (IBM, 2005), which was presented in 2005 in a white paper by the International Business Machines Corporation (IBM) (see Figure 15), but, it remains widely used, especially in IoT scenarios (JAHAN *et al.*, 2020; CLELAND-HUANG *et al.*, 2022; MALBURG *et al.*, 2023).

Figure 15 – Healful platform design inspired on the MAPE-K loop



Source: author.

In MAPE-K, four stages were defined, aiming to sense and act in the environment using shared knowledge. In short, they are: i) monitor stage (1), which is focused on collecting, aggregating, and filtering data collected from the environment; ii) analyze stage (2) to observe and analyze situations to determine if some change needs to be made; iii) plan stage (3) to define a procedure able to conduct the required change; and, finally, iv) execute stage (4) to perform the necessary changes (IBM, 2005). Hence, the platform supports collecting and analyzing data, planning what actions should be taken based on the analysis, and acting in the environment.

The first stage – called **Monitor** – is responsible for obtaining raw data (1). These data can be collected from mobile devices and wearables (such as smartwatches and smart bands). Because of this responsibility, the behavioral pattern called Strategy (GAMMA *et al.*, 1993) with a generic interface (labeled Data Extractor) was implemented. Thus, a new strategy should be implemented for each new data source. These data extraction modules run in the background within a mobile application installed on a smartphone, and the context management was inspired by LoCCAM middleware (ANDRADE *et al.*, 2020).

The rationale for choosing users' smartphones as the edge node to monitor their behavior is that they are suitable to be the central point for obtaining data and acting in the Smart Environment (HARTMANN *et al.*, 2022). Also, as the target profile of this work is independent adults (as discussed in Chapter 1, Section 1.2), these devices are close to their owners for most of the day (HARARI *et al.*, 2017).

In the **Analyze** stage, the Healful platform provides the QoL inference report with a minimum frequency of a day. This report includes Quality of Life scores for the physical and psychological domains (calculated by the QoL Inference module using Regression Models) and the five health indicators (calculated by the Personal Health Indicators module). In addition, it is possible to add new attributes to the report from the systems created in Athena[2], which was included in the Self-Service Machine Learning module. Athena is a cloud-based tool to support the development of systems that require Computational Intelligence (CI) techniques. Athena uses the abstraction of visual modules to encapsulate CI algorithms allowing their interconnection to solve complex problems (OLIVEIRA *et al.*, 2018). With this QoL report, healthcare professionals can describe contexts of risk that should be monitored.

The **Plan** stage was designed with two modules: Risk Analyzer and Intervention Planner. For the risk analyzer, a simplified version of the method proposed by KIM *et al.* (2019a) was used as inspiration. KIM *et al.* (2019a) explore a What-If analysis to predict factors affecting adolescent obesity. The authors report that this approach was helpful in guiding health professionals to expand their skills in simulating scenarios based on specific factors. Thus, based on these characteristics, a mechanism to describe contexts and their associated risks was included in the Healful platform. Currently, this description is done through textual rules following the JEP[3] grammar of mathematical and logical expressions.

[2] Athena website: http://athenasystem.com.br.
[3] JEP website: http://sens.cse.msu.edu/Software/jep-2.23

For example, *Context* 1 defines a context in which the monitored user kept the Quality of Life indicator for the physical domain below 30. Thus, this context has a risk associated with obesity and cardiovascular issues. This illustrative example reflects how healthcare professionals can configure the risk analyzer.

```
# Context 1:
  - Check frequency: daily
  - Risk description: Obesity and Cardiovascular issues
  - Expression to check:
    * physical-qol < 30
```

The healthcare professional should create the intervention plan after defining the contexts and their associated risks. This plan must be built using if-then rules structured as follows: `if <context> then <action>`. The Healful platform was designed to deal with three kinds of actions: i) sending a periodic notification on the user's phone; ii) adapting the environment (when actuators are available), and iii) notifying the emergency or person in charge when a critical context that needs medical intervention is identified. Here, there are samples of the three kinds of rules:

```
Rule 1: if <Context 1> then <Send a periodic notification to user>
Rule 2: if <Context 2> then <Adapt the environment to reduce brightness>
Rule 3: if <Context 3> then <Call the emergency department>
```

After the rules definition, the last stage is enabled. The **Execute** stage is responsible for performing the planned interventions (actions). These actions are executed by the Worker module, which uses specific modules for each action, such as the Health Recommendations module for sending notifications about the user's health. In addition, it is worth mentioning that MAPE-K loops generate a vast amount of data, representing the **knowledge** acquired in that context. In the Healful platform, it is used MongoDB[4] (NoSQL database) to store the data.

To conclude this subsection, it is important to highlight three points:

1. Many other studies focused on collecting and analyzing users' health data can be conducted using the Healful platform. For example, i) the inference model would achieve better results using camera data; ii) it would be possible to reduce model error using machine

[4] MongoDB website: https://www.mongodb.com.

learning techniques not explored in this work, or iii) it would be interesting to correlate health data with different health questionnaires (*e.g.*, SF-36 (WARE *et al.*, 1994), EQ-SD from the EuroQoL Group (RABIN; CHARRO, 2001) and KIDSCREEN-52 for children and adolescents (RAVENS-SIEBERER *et al.*, 2005)). These possibilities reinforce the need for a structure to assist further investigations in this research area. Therefore, the Healful platform offers support to build QoL-based IoHT systems.

2. Although this subsection describes how the Healful platform was designed to support the four stages of MAPE-K, this work focuses on the first stages: monitor and analyze. The last stages (plan and execute) represent a promising opportunity for conducting a longitudinal study and this opportunity is discussed in future work.
3. Based on the scope of this work (discussed in Chapter 1, Section 1.2), the intervention plan of the current Healful version only deals with the first kind of intervention: send a periodic notification on the user's phone. However, the other types of action (such as adapting the environment when actuators are available and calling the emergency department) should be studied in depth in future work.

Architecture

Software architecture describes the system components and the relations between them (PRESSMAN, 2010). Thus, an architecture must represent how the system is structured and how its components work together to achieve the goal of the software. Based on this definition, this subsection discusses the architectural view of Healful modules.

Initially, it is essential to highlight the platform's prominent use cases and actors. The Healful platform was designed to serve three types of users: end-users, researchers, and health practitioners. Both share functions such as access to the home page, logging in, and registering new users. However, each type of user has specific functions shown in Figure 16. For end-users, the features are related to health monitoring, such as Connect with Google Fit and Get QoL Reports. For researchers, the platform has features related to the construction of adaptation loops, the definition of sensors used in data collection, and the construction of intelligent systems. Finally, health practitioners can define the risk contexts that the tool will monitor.

Furthermore, other features are under development, such as obtaining health indicators for end users and generating analytical reports for researchers and health practitioners. With such features, the platform will be able to propose changes in habits to avoid health issues.

Figure 16 – Healful Use Case Diagram with the main features

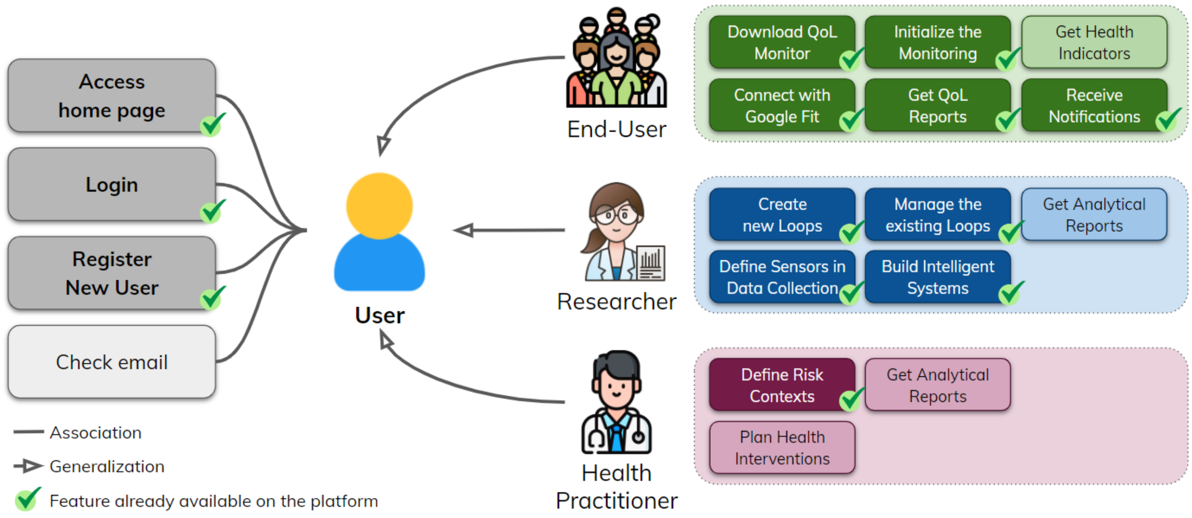
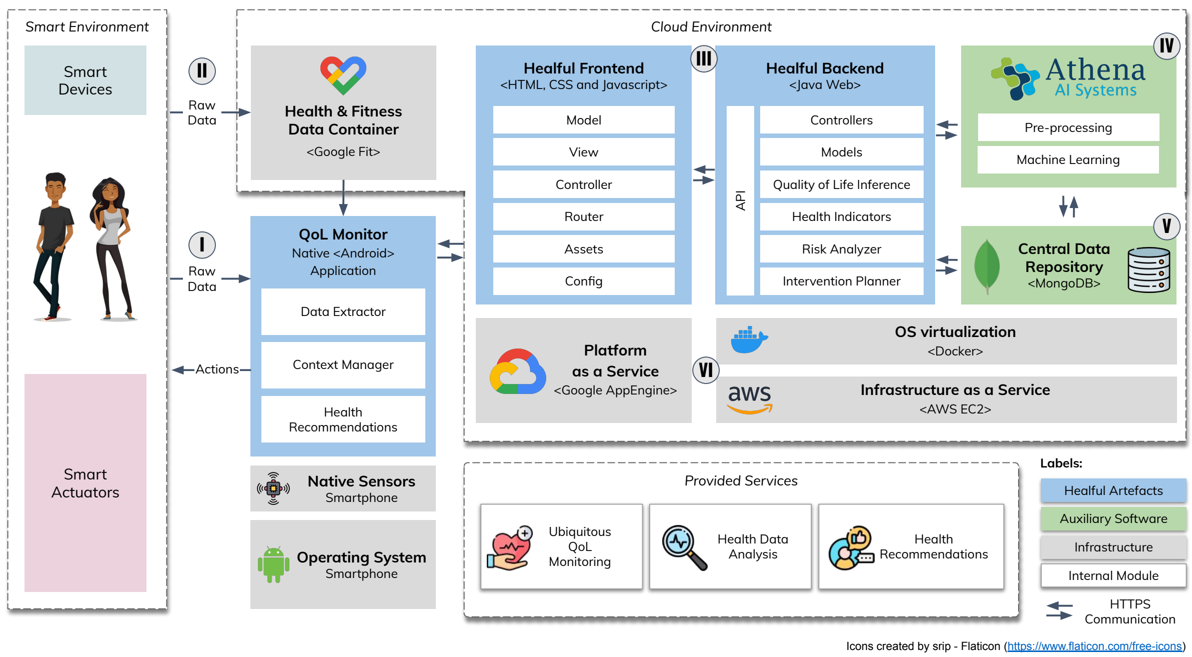


Figure 17 presents the platform software artifacts (represented in blue) and its internal modules (represented in white), the auxiliary software (represented in green), and the computing infrastructure (represented in gray) that supports its operation, such as Google App Engine, Docker Platform, and the Amazon Web Services. In addition, arrows represent communication between components, and all communications are performed using the HTTPS protocol to ensure the security of data transmitted over the network.

Figure 17 – Architectural view of the Healful platform modules



Note: Communication between the components (arrows) is performed using the HTTPS protocol.
Source: author

Data collection occurs directly from the native sensors of the user's smartphone (see point I in Figure 17) or through the Google Fit[5] (II). Google Fit was included in this process to overcome the challenges of heterogeneity and the absence of public Application Programming Interface (API) to extract wearable data. Without an API, data extraction became a complex problem, and if each wearable had a specific API, mapping all of them would be unfeasible.

As most commercial wearable devices allow users to synchronize data with their Google account automatically, these data can be collected in a unified standard using the Google Fit API. This decision allows Healful to deal with several models of wearables (*e.g.*, AmazFit, Fitbit, MiBand, Samsung Galaxy Watch, Polar Watches, Wahoo Watches, and all the others that use the Android Wear Operating System). The management of this collection is done by the application QoL Monitor[6]. Internally, this app has three main modules: Data Extractor, responsible for data extraction; Context Manager, responsible for user's context; and Health Recommendation, responsible for notifying users regarding their health status.

After a rigorous registration process that requires compliance with many privacy restrictions, the QoL Monitor was approved on Google Play Store[7] and is available to internal users. The app is also under registration by the National Institute of Industrial Property (ID: BR 512023003335-9).

In the cloud environment, Healful has two subsystems (III): Healful frontend, developed with Vue.JS (FILIPOVA, 2016); and Healful backend, developed using Java Servlet (HUNTER; CRAWFORD, 2001). The Healful frontend allows the researcher to configure how the loop should run, and this configuration includes defining which data should be collected, how data analysis should be performed, which contexts should be monitored, and which actions should be performed in each context. On the other hand, the Healful backend implements the platform main features, making them available through the API module. The backend is also responsible for communicating with Athena (IV) and MongoDB (V). MongoDB, in turn, was chosen for the central data repository due to its flexibility, allowing storage using JavaScript Object Notation (JSON) (BANKER *et al.*, 2016).

Regarding the infrastructure (VI), the frontend is hosted on Google App Engine (GAE)[8]. GAE selection was based on the simplicity of the deployment for Javascript applications.

---

[5] Google Fit website: https://www.google.com/fit.

[6] QoL Monitor website: https://www.qol-monitor.com.

[7] Play Store link of QoL Monitor for pre-selected users: play.google.com/apps/internaltest/4699945957479949489. To require access, please send an email to pedro.oliveira@ifma.edu.br.

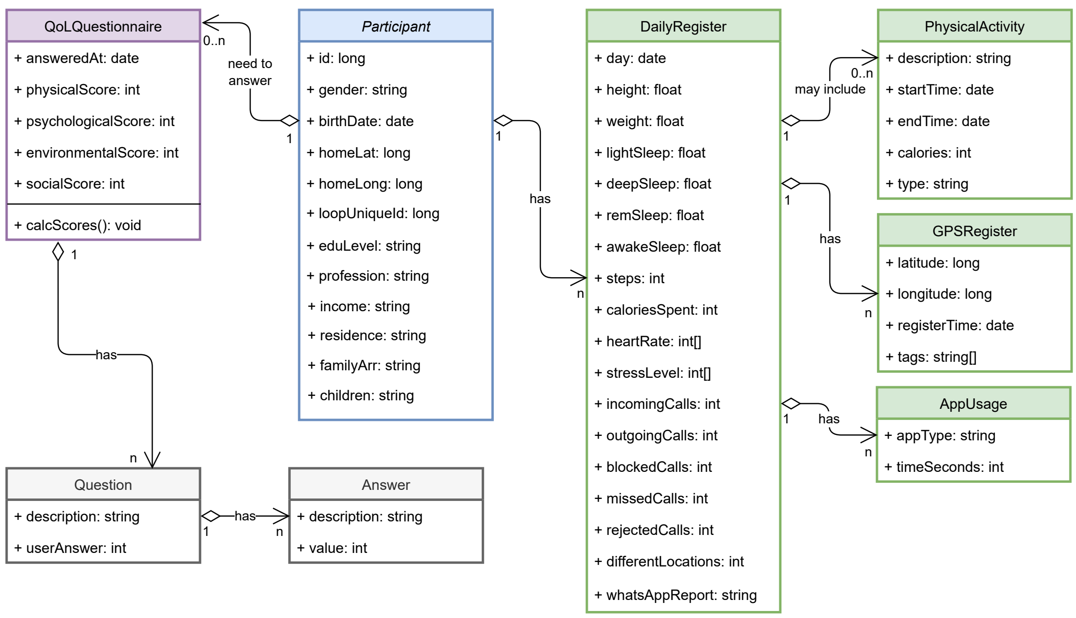
[8] GAE website: https://cloud.google.com/appengine.

In addition, GAE enables a quick association with custom domains and free Secure Sockets Layer (SSL) certificates. The backend, in turn, is hosted on the Amazon Elastic Compute Cloud (EC2)[9] service using Docker[10] containers. This choice was appropriate for the backend to avoid restrictions imposed on Platform as a Service (PaaS) models, such as using native CronJobs and Threads in Java. To conclude, it is important to mention that the first version of the Healful platform is available using the https://healful.life link.

QoL Monitor

This subsection provides details about the QoL Monitor app. This Android app is a fundamental part of the Healful architecture and it was developed to collect contextual and health data from users. Its current version collects sociodemographic and anthropometric data, sleep duration, daily steps, calories spent, physical activities, heart rate, location, apps usage time, and the number of calls made or received.

Figure 18 – QoL Monitor class diagram.



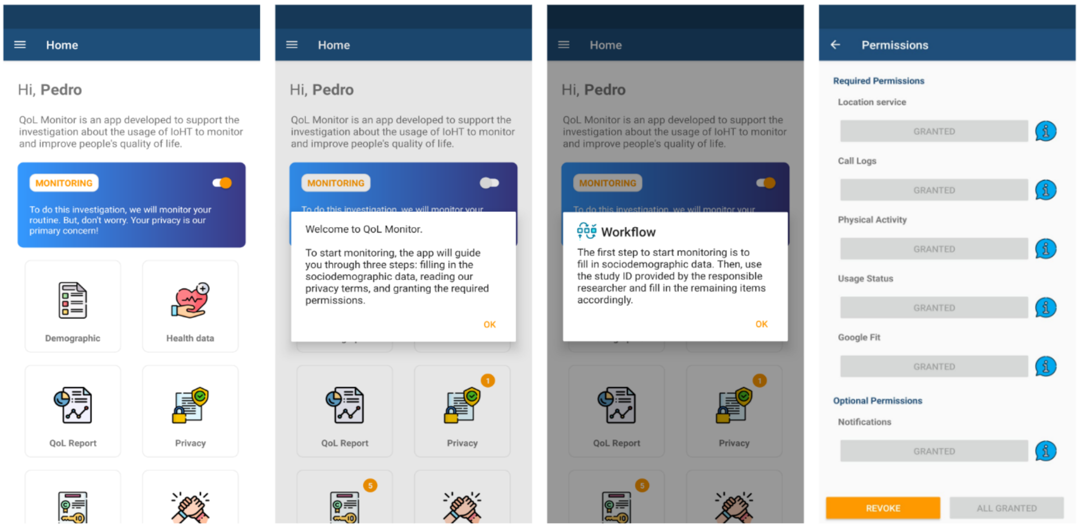
Source: author.

Figure 18 presents the central part of the class diagram used in modeling these data. To do this robust data collection, it was necessary to integrate the app with the Google Fit API. Thus, users can use different wearables as long as they are integrated with the Google Fit account.

[9] AWS EC2 website: https://aws.amazon.com/ec2.
[10] Docker website: https://www.docker.com.

The application was built as a native Android app using Kotlin language, given the need to access different sensors available on the user's smartphone. The app has a dashboard with quick shortcuts for sociodemographic data, privacy terms, and required permissions (Figure 19). Context management engine was implemented inspired by the LoCCAM-IoT middleware (ANDRADE *et al.*, 2020). That way, whenever new sensor data is available, it is stored locally, and, at the end of the day, the information is encrypted to send to the cloud. A symmetric algorithm (AES) and an asymmetric algorithm (RSA) were used to encrypt the user data. As an additional security effort, data is anonymized before submission.

Figure 19 – QoL Monitor dashboard and workflow up to granting permissions

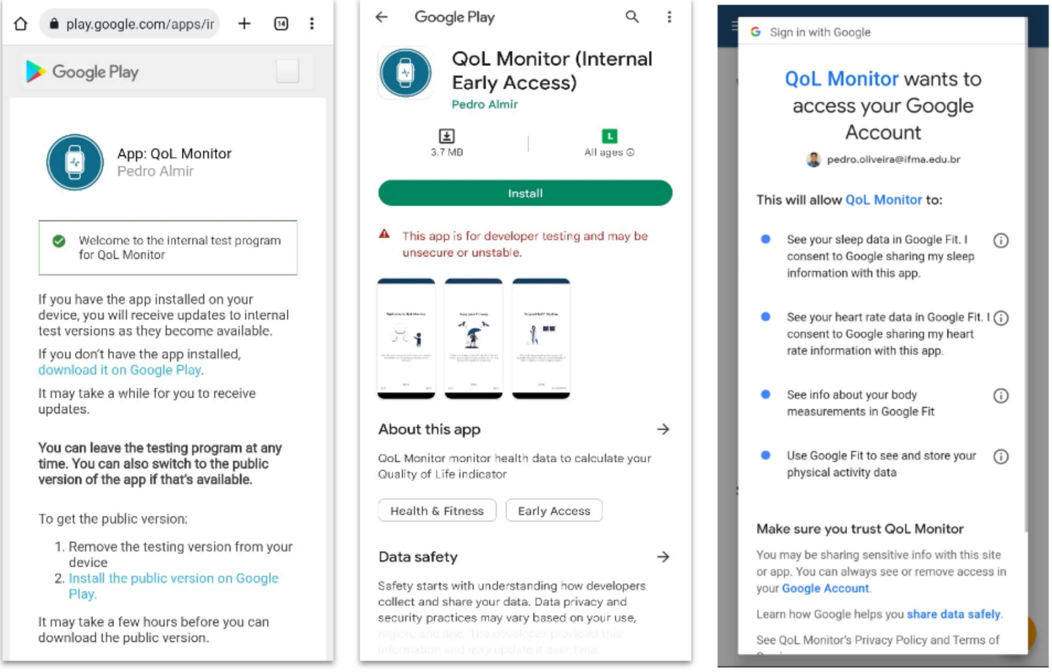


Source: author.

After enrollment, the app was incorporated into Google Play and already had a registered API in Google Cloud (Figure 20). Thus, users are assured that the app satisfies the app store's requirements. However, access is restricted to internal users while this work is not done. Figure 21 highlights other features of the app, such as stress perception, health data visualization, QoL report, and pending data listing. This latter feature guarantees that the data is not lost if, at the time of sending, the user does not have access to the Internet.

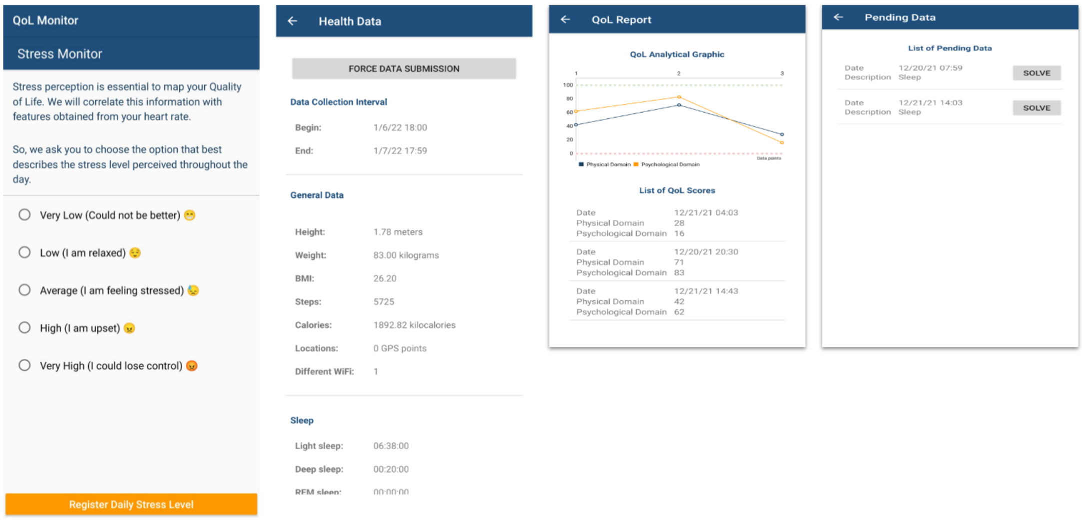
Regarding QoL questionnaires (Figure 22), QoL Monitor was designed to change the self-reported questionnaire easily. The questions are stored in a Google Spreadsheet, and using the Healthful platform; it is possible to adapt the questions or even completely change the questionnaire. So, suppose the researcher decides to work with another questionnaire, such as the SF-38. In that case, it is possible to inform the new questions on the platform, and the app will automatically be updated as the following tutorial videos discuss.

Figure 20 – QoL Monitor in Google Play and Google Cloud platform



Source: author.

Figure 21 – QoL Monitor additional features



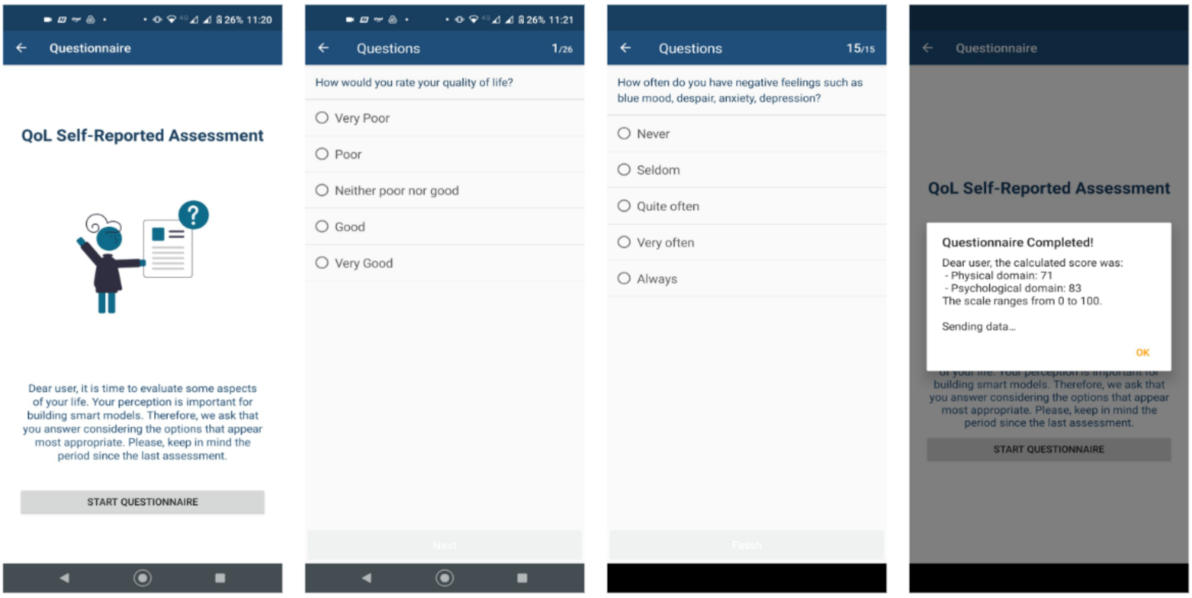
Source: author.

– Overview: youtube.com/watch?v=v3I7eDaQnlc
– Workflow: youtube.com/watch?v=LmxvG8Lgozg
– Permissions: youtube.com/watch?v=V6ejQqfGsco
– Google OAuth: youtube.com/watch?v=wZOFQs8UmMA

Finally, it is also possible to explore and test resulting regression models through a public API and in the project repository[11], there is a request sample in the Postman format.

[11] Request sample: github.com/great-ufc/healful-thesis/tree/main/qol-inference-service.

Figure 22 – QoL Monitor self-reported questionnaires



Source: author.

4.3 Healful Method to Infer Users' QoL

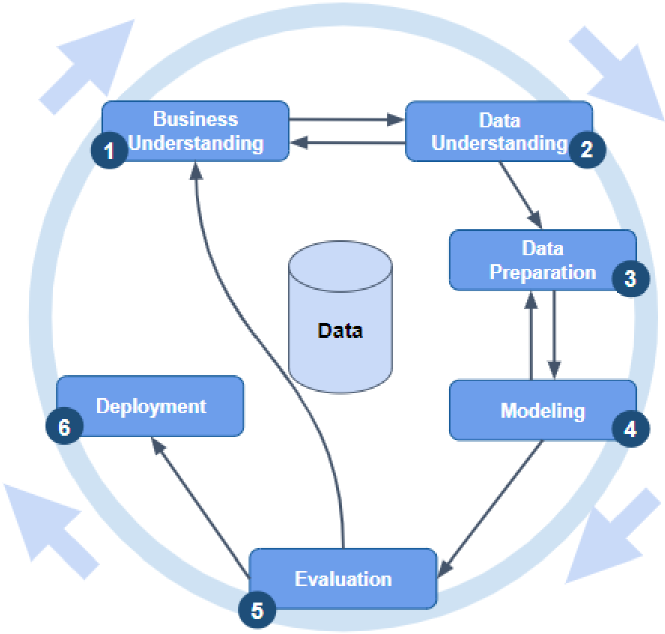
As stated by the International Society for Quality of Life Research (ISOQOL), investigations in the Quality of Life area can support health programs and policymakers to better allocate resource[12]. Due to this, many researchers are dedicated to this study area. However, as Chapter 3 discusses, there is room for less intrusive approaches. Thus, this work used commercial wearables to collect users' health data and Machine Learning techniques to infer QoL scores.

The QoL inference method uses Machine Learning algorithms trained with users' data. This method is an alternative to the self-reported QoL questionnaires. As this method is based on Machine Learning, some actions are required, such as defining which data are useful for training, selecting the algorithms with the best performance, and optimizing the parameterization of the selected model.

In order to guide the Machine Learning process, it was selected the CRoss Industry Standard Process for Data Mining (CRISP-DM) process (Figure 23) that was created to support researchers and practitioners in executing data mining projects (WIRTH; HIPP, 2000). This decision was based on process characteristics that allow iterative conduction, resulting in artifacts to be deployed, as well as on the good results obtained with this process in previous works (OLIVEIRA *et al.*, 2021). Although CRISP-DM is an old process (its release was in 2000), as pointed out by SCHRöER *et al.* (2021), it is still a *de-facto* standard in the data mining area.

[12] What is QoL by ISOQOL: https://www.isoqol.org/what-is-qol.

Figure 23 – CRoss Industry Standard Process for Data Mining activities



Source: author.

The CRISP-DM process has six steps. The first step – Business Understanding – focuses on understanding the goals and particularities of the target project. The second step – Data Understanding – involves the initial search for data to become familiar with it. The first two steps are closely related because understanding the business requires initial data, and this initial analysis of data can impact project goals. The third step – Data Preparation – encloses activities related to the dataset construction, such as attribute selection, data cleaning, building new attributes, and data transformation. In the fourth step – Modeling – a set of Machine Learning algorithms is selected to build intelligent models. The fifth step – Evaluation – uses statistical tests to identify the models' performance. Finally, in the sixth step – Deployment – the model with the best performance is usually made available as a service. This last step was not performed in this work because the inference model was embedded inside the Healful Platform. All these steps are described in the following subsections.

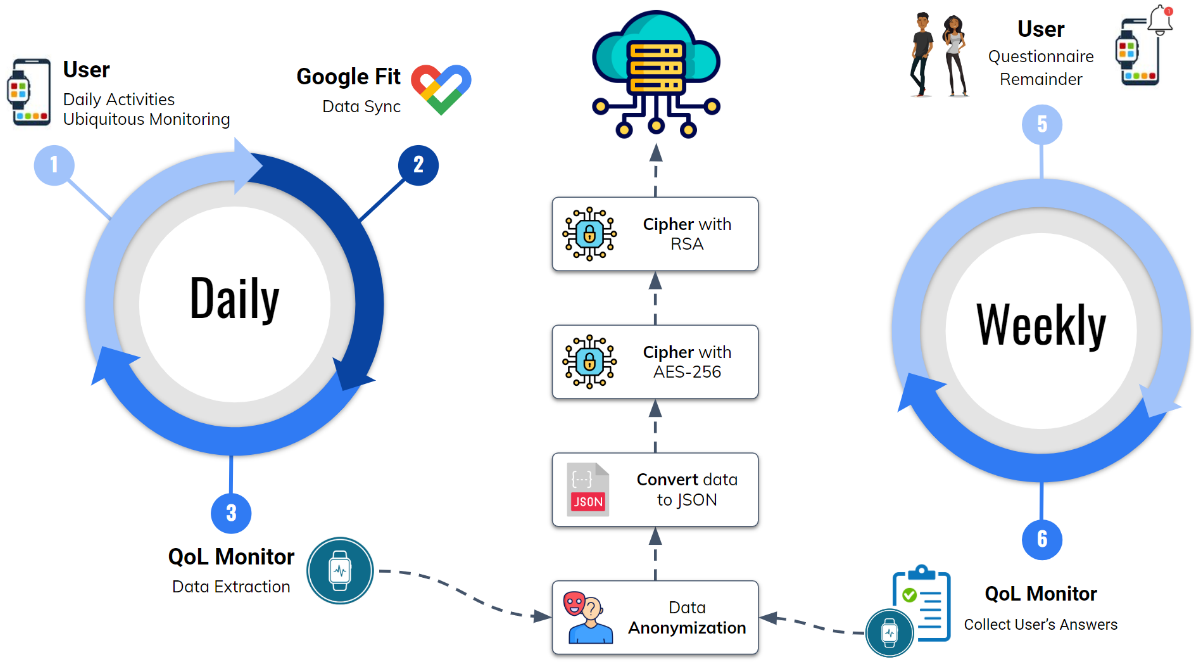
Business and Data Understanding

As previously mentioned, CRISP-DM first step is the business understanding, which is the Internet of Health Things applied to Quality of Life. This understanding was achieved through investigations in the IoHT literature as discussed in Chapter 2.

Regarding data understanding (step two), initially, a study to understand how to obtain health data in practice (*i.e.*, in the usual user environment) (OLIVEIRA *et al.*, 2022) was performed, and later, a multivocal review aiming to find existing datasets that could help the project (JUNIOR *et al.*, 2022). Both studies pointed to challenges, such as device heterogeneity, the absence of native API to obtain data directly from commercial wearables, and the absence of public datasets that correlate health measures (*e.g.*, daily steps and heart rate) with self-reported QoL questionnaires. Due to these challenges, it was necessary to develop a mobile application – called QoL Monitor (OLIVEIRA *et al.*, 2022b) – to obtain health data from "Health & Fitness" Data Containers (Google Fit was selected due to its slight integration complexity) and correlate this data with QoL questionnaires.

Figure 24 presents the data flow to create a dataset correlating health measures and self-reported QoL questionnaires. Initially, the user must connect the native app of their wearable to sync data with the Google Fit platform. This connection is necessary to overcome the device heterogeneity issue. Once the data is registered in Google Fit, it is possible to extract the data using the Google Fit public API. Then, daily, QoL Monitor extracts the data recorded in Google Fit, anonymizes, cipher (using AES-256 and RSA algorithms), and sends it to the cloud service. Finally, the app weekly requests the user to answer the QoL questionnaire to store it together with the health data.

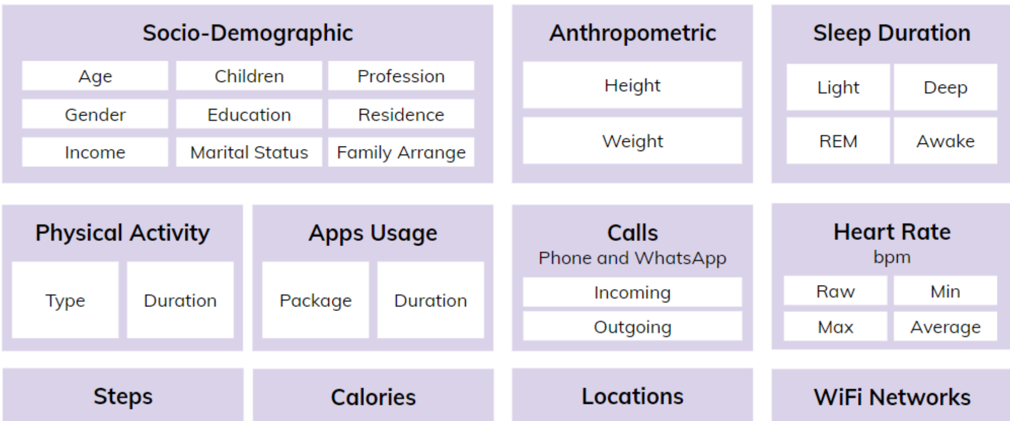
Figure 24 – Data flow to collect health measures and self-reported QoL questionnaires



Source: author.

Figure 25 highlights the data collected from users. Sociodemographic and anthropometric data are necessary to understand the characteristics of the users. The other raw data directly correlates with the health indicators chosen for this work. In addition, all of them can be obtained through common devices[13] such as smart bands and smartwatches. Additionally, it is worth mentioning that the location data only stores the number of points visited throughout the day, *i.e.*, the app does not record the specific places. The same logic was applied to the WiFi network identification. The app records the number of different WiFi Networks connected throughout the day. This strategy was adopted to preserve the users' privacy.

Figure 25 – Raw data collected from users



Source: author.

Data Preparation

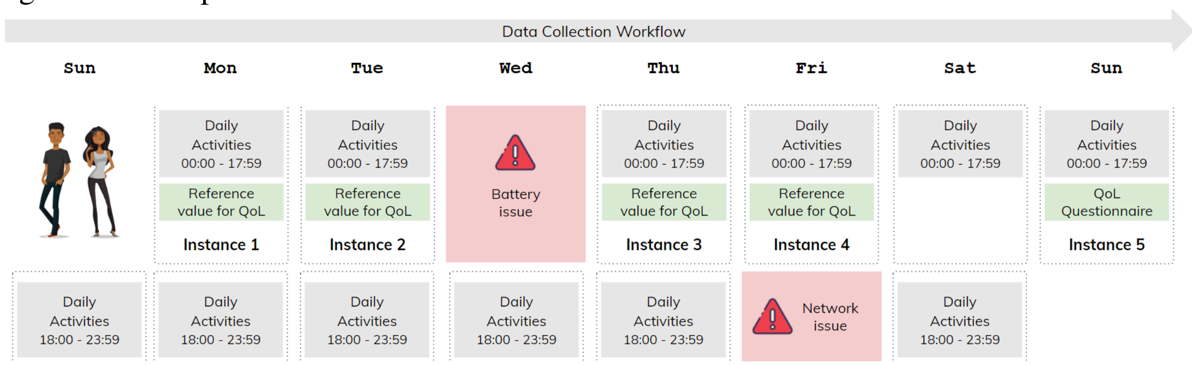
Data preparation covers the activities to build the dataset used to feed the modeling tools. These activities include defining data segmentation, attribute selection, data cleaning, and data transformation (WIRTH; HIPP, 2000).

Figure 26 puts light on how the instances are created. A sample has as predictors all data collected from 18:00 of the previous day to 17:59 of the current day. This time slot was empirically selected because the last night's sleep directly impacts the current day's activities. The value to be predicted is obtained after answering the questionnaire on Sundays. As the user must answer this questionnaire considering the past week, it is possible to use this value as a reference. Naturally, issues (*e.g.*, absence of network or battery issues) can arise during the data

[13] Some commercial devices may not support data collection of characteristics like heart rate. Thus, to use this method, it is required to consider devices that offer at least the data listed in Figure 25. The absence of some of these data can imply a reduction in the performance of intelligent models.

collection, and if the data is not recorded, such intervals do not generate new instances.

Figure 26 – A representation of how the instances are created



Source: author.

After obtaining the raw data, preprocessing activities are performed to prepare the dataset for the modeling stage. Among these activities are:

- removing inconsistencies (*e.g.*, duplicate entries);
- removing outliers, such as extremely high values for the daily steps. To remove these outliers, it was used three standard deviations below and above the mean;
- removing data gaps, for example, days without sleep or heart rate data;
- categorical variables encoding like sociodemographic data;
- data sync since users forgot to answer the QoL questionnaire on Sundays;
- computation of QoL scores based on the questionnaire responses;
- data transformation, such as summarizing time spent in each application category.

Finally, two datasets are obtained: i) a dataset in which the last column is the QoL score for the physical domain; and ii) a dataset in which the last column has the QoL score for the psychological domain. The last column changes because it is used as a reference for the learning process. Each final dataset contains 88 features described in Appendix C.

Modeling and Evaluation

After data preparation, there is the modeling step. In this step, Machine Learning algorithms are selected as candidates to build the intelligent model, and each model is evaluated in the evaluation step. For the Machine Learning algorithms modeling, the Scikit-learn toolbox (PEDREGOSA *et al.*, 2011) was selected due to its high acceptance in the scientific community and the consistency of its results (TANAKA *et al.*, 2019; GÉRON, 2019). Then, four algorithms were selected based on GÉRON (2019) guidelines: Linear Regression, Decision Tree Regressor, Random Forest Regressor, and GBoost Regressor. In addition, the algorithm called Extra Trees Regressor was included after applying the FLAML[14] Automated Machine Learning library on the final dataset. This library – maintained by Microsoft – was chosen due to its scientific relevance and ability to identify accurate Machine Learning models for different types of tasks (WANG *et al.*, 2021).

The first algorithm – Linear Regression – searches for linear relationships within the dataset. It is considered a simple model and an excellent choice to start investigating regression problems (IAN; EIBE, 2005). The second algorithm – Decision Tree Regressor – is robust compared to linear regression and can find nonlinear relationships in the data. The third algorithm – Random Forest Regressor – uses the concept of random forests to train multiple decision trees. This algorithm performs well for a wide variety of problems (PAUL *et al.*, 2018).

The fourth algorithm – GBoost Regressor – is an ensemble such as the Random Forest (FRIEDMAN, 2001). However, it uses the gradient descent method to minimize the error function. Finally, the last algorithm is called Extra (randomized) Trees Regressor (GEURTS *et al.*, 2006). This algorithm is also an ensemble, *i.e.*, it combines randomized decision trees (also known as extra-trees) to improve the predictive accuracy and control over-fitting[15]. The Extra Trees Regressor differs from the Random Forest regarding the selection of cut points for splitting the tree nodes. While Random Forest searches for the optimal points, Extra Trees makes a random selection. In this way, in many cases, Extra Trees Regressor achieves good results quickly (GEURTS *et al.*, 2006; TANG *et al.*, 2018).

Regarding model evaluation, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) measures were selected since the main task in this thesis is regression. These two metrics are probably the most used in this kind of task (IAN; EIBE, 2005). MAE represents the average magnitude of the individual errors without taking account of their sign (see equation 4.1 where p_n represents the nth predicted value and a_n means the nth actual value), and RMSE is the square root of Mean Squared Error (MSE), *i.e.*, the average of the square of the difference between the actual values and the predicted ones (see equation 4.2). Together, these metrics are proper for evaluating the performance of a given algorithm. In this case, the smaller, the better the model created by the algorithm.

[14] FLAML website: https://microsoft.github.io/FLAML.

[15] For additional information access the Sklearn documentation: scikit-learn.org/0.21/documentation.html.

$$\text{MAE (Mean Absolute Error)} = \frac{|p_1 - a_1| + ... + |p_n - a_n|}{n} \quad (4.1)$$

$$\text{RMSE (Root Mean Squared Error)} = \sqrt{\frac{(p_1 - a_1)^2 + ... + (p_n - a_n)^2}{n}} \quad (4.2)$$

Finally, the model with the best performance evaluation was selected for the final solution. The preparation, modeling and evaluation activities were conducted using the dataset collected from two case studies, and the results are discussed in Chapter 5.

4.4 Health Indicators

Five health indicators – daily mobility, physical activity level, loneliness, social mobility, and sleep quality – were included in this work because the inference mechanism presents only results similar to most QoL questionnaires: an integer value between 0 and 100. This score is helpful for medical practice since it represents how good the patient's Quality of Life is (MCNAUGHTON *et al.*, 2022). However, it is not self-descriptive. Thus, the patient needs additional information to understand the resulting QoL score. Therefore, the five health indicators were selected to support users in understanding which health aspects deserve attention after getting the physical and psychological QoL scores.

The health indicators were chosen based on the knowledge present in the literature that correlates each of them to physical and psychological health (targets of this work). For example, VETROVSKY *et al.* (2017) present an investigation correlating pedometer-based interventions with lower anxiety/depression and higher health-related QoL. Similarly, many other authors discuss the correlation between <u>daily mobility</u> and <u>physical activity level</u> with the patient's Quality of Life (VALLANCE *et al.*, 2016; KRAUS *et al.*, 2019; PANICKER; CHANDRASEKARAN, 2022).

Regarding <u>loneliness</u> and <u>social mobility</u> levels, there is evidence that the social component strongly influences psychological health (SANCHEZ *et al.*, 2015; DAHLBERG *et al.*, 2022). MUSHTAQ *et al.* (2014) state that "*satisfying social relationships are essential for mental and physical well-being*". Thus, these two indicators complement each other to observe how users interact with others and move outside their home environment. However, it is not part of this work to explore the quality of these social relationships, although this represents an interesting point for further investigation.

It is also important to mention that, in this work, social mobility is the level of the daily displacement of a person. For example, high social mobility can be described as someone who passes through different locations throughout the day (*e.g.*, work, supermarket, and friends' houses). On the other hand, loneliness is the opposite of social mobility but also includes communications features (*e.g.*, calls and messages).

The last indicator is sleep quality. Sleep is essential to restore the physical and psychological aspects of the human body (LUYSTER *et al.*, 2012; WORLEY, 2018; STOJANOV *et al.*, 2021). According to ARORA *et al.* (2020), deep sleep restores muscles and removes waste from the brain, while the Rapid Eye Movement (REM) re-energizes the mind.

These five health indicators are not an exhaustive list since there are many other possibilities (WHO, 2015; OPAS, 2018; COMMITTEE, 2020). However, based on the results of a survey (see Chapter 5) conducted to evaluate this set of indicators, it is possible to state that they are suitable for complementing the QoL inference process.

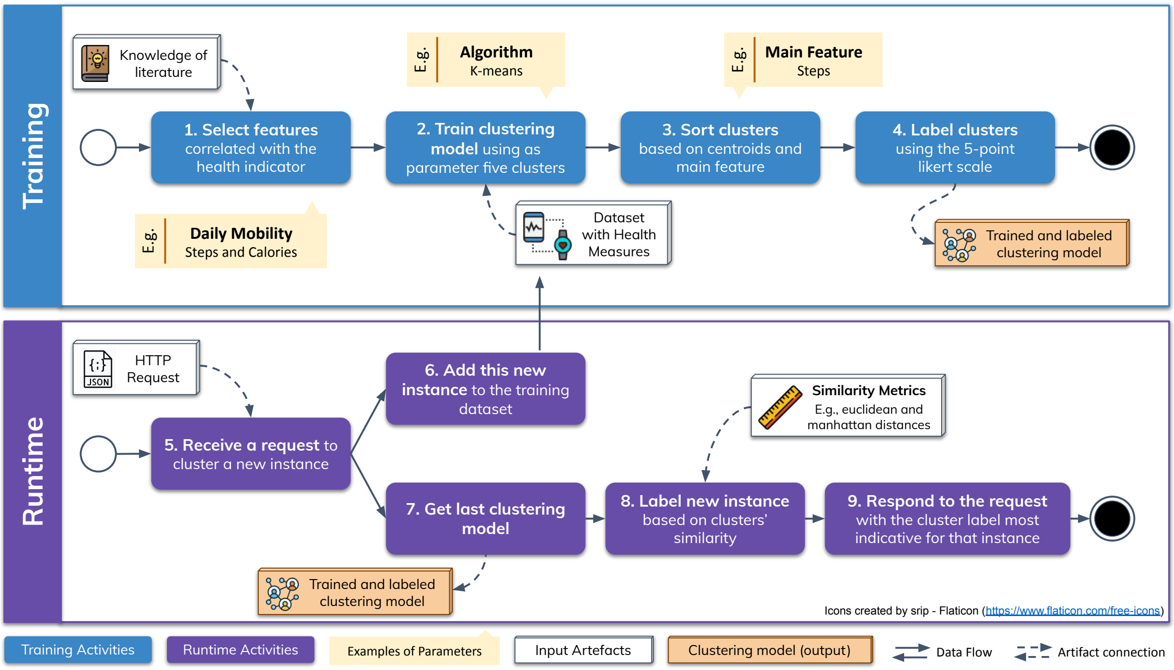
Clustering and labeling method for health indicators

While conducting the evaluation activity, it was realized the feasibility of inferring the users' QoL using data collected by Smart Devices. However, the need to make this inference explainable was also prominent. As with QoL questionnaires, the value returned for each QoL domain inferred varies between 0 and 100, and the user needs details about what affects this result. Therefore, inspired by how radar graphics communicate quantitative information about a given process, it was decided to calculate five indicators directly related to the QoL domains.

However, in the literature, it is complex to find a consensus on how such indicators should be calculated because they depend on many variables. Due to this issue, a clustering and labeling method was defined to calculate daily mobility, physical activity level, social mobility, and loneliness. The last indicator – sleep quality – is calculated based on an adapted deterministic solution presented in the next section.

The method for calculating the first four indicators – daily mobility, physical activity level, social mobility, and loneliness – was unified because the measurements on which they depend can be sorted and compared. For example, the greater the daily mobility, the better. In addition, this method uses a social component to classify indicator levels. In Figure 27, it is possible to realize that the method has two moments: training (represented in blue) and runtime (represented in purple).

Figure 27 – Clustering and labeling method for health indicators



Source: author.

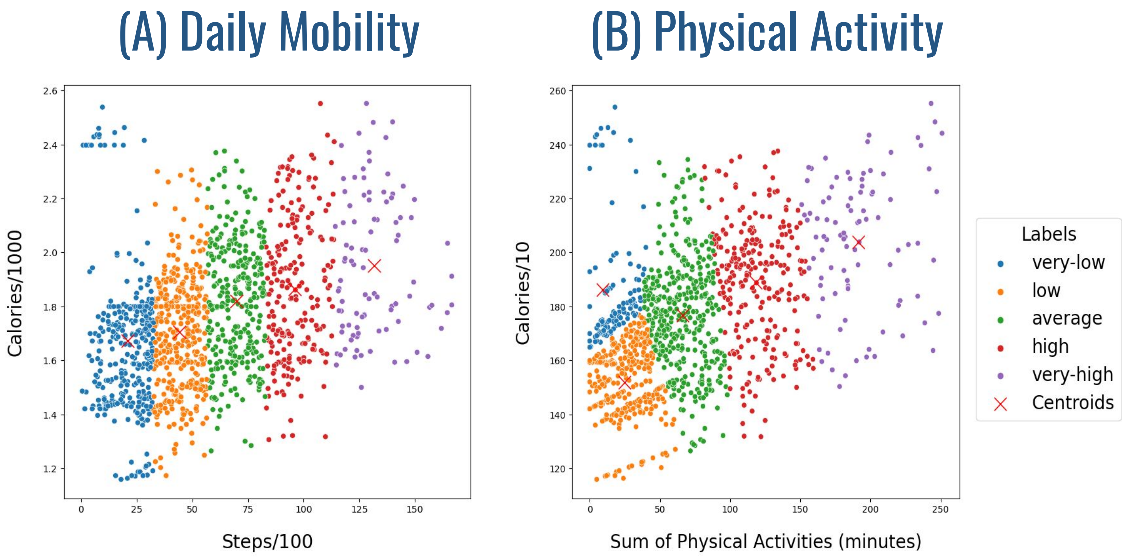
During training, four steps are performed. Initially, correlated features for each health indicator are defined based on the literature and available raw data (1). For example, in the case of daily mobility, the number of steps and the calories burned affect this indicator. Then, a training of the clustering model parameterized with five clusters is performed (2). Next, the resulting clusters are ordered (3) using the centroids and the main feature of the health indicator (*e.g.*, for daily mobility, the main feature is the number of daily steps). Finally, the clusters are labeled using a 5-point Likert scale (4).

Regarding the artifacts used during the training, in the first step, the input artifact to define the features applied in clustering is the knowledge present in the healthcare literature. In the second step, the input artifact for the clustering process is the dataset with the health measures collected by the QoL Monitor. Finally, the fourth step returns a trained and labeled clustering model as an output artifact.

Figure 28.A shows an example of clusters for daily mobility created using the K-means algorithm based on two features: steps and calories. It is possible to realize that the greater the number of daily steps and calories burned, the better the classification presented. However, this method depends on the data quality. Thus, it is also possible to observe situations where clustering does not show good results. For example, also in Figure 28.A, there is a set of users with high caloric consumption and a low number of steps. These users had their daily mobility

classified as "very low" because the main attribute is the number of steps. Figures 28.B, 29.C and 29.D brings clustering samples for physical activity, social mobility, and loneliness, respectively.

Figure 28 – Clustering and labeling for daily mobility (A) and physical activity (B)



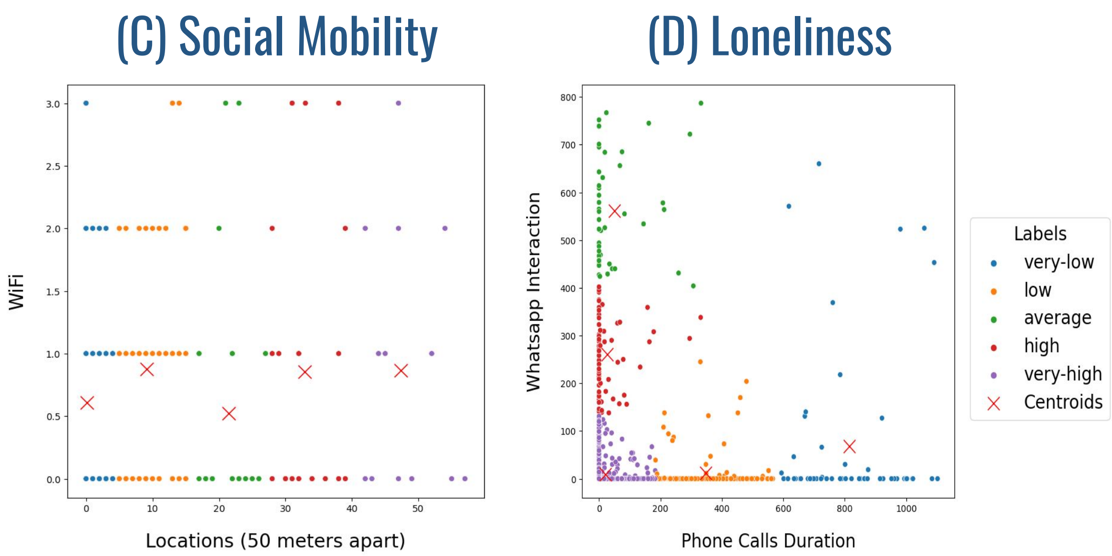
Source: author.

After the training stage, two actions are triggered in parallel when receiving a request to cluster a new data instance (5) at runtime: add this instance to the training dataset (6) and get the last trained model (7). Then, this instance is labeled based on cluster similarity (8), and the system responds with the level of the health indicator considering the Likert scale (9).

It is worth reflecting on a critical point related to the presented method to calculate these four indicators. The indicators are calculated based on raw measurements, and there is no consensus on how they should be calculated. Therefore, a clustering method was defined to calculate these indicators. This method is based on the ability of clustering techniques to group data according to their similarity (GIORDANI *et al.*, 2020). Such methods have a limitation concerning cluster labeling (TREERATPITUK; CALLAN, 2006). Thus, a labeling strategy was included considering that the data generated by these indicators can be ordered and classified using the Likert scale.

After a prolonged investigation of the healthcare literature, it is possible to state that finding an acceptable scale ("*silver bullet*") for these indicators (daily mobility, physical activity level, social mobility, and loneliness) is unfeasible. For example, there is no consensus regarding the number of daily steps required to ensure reasonable daily mobility (HOEGER *et al.*, 2008; TUDOR-LOCKE *et al.*, 2011; WATTANAPISIT; THANAMEE, 2017). In addition,

Figure 29 – Clustering and labeling for social mobility (C) and loneliness (D)



Source: author.

other methods like rule-based instruments would make rules management a complex and costly process (MCNEILL; THRO, 1994). The method presented in this work depends on the data quality to present good results. However, unlike other studies, i) it does not require human intervention after the deployment, ii) it can evolve as new data instances are collected, and iii) it has a low maintenance cost when compared to rule-based solutions.

Method for Sleep Quality

Sleep Quality is one of the most critical health indicators and should be constantly monitored. A night of proper sleep is responsible for muscle and mental recovery (ARORA *et al.*, 2020). On the other hand, poor quality sleep can generate health issues in the physical, emotional, and cognitive domains (STRINE; CHAPMAN, 2005; ALTEVOGT *et al.*, 2006). In addition, there are studies that correlate sleep disorders with symptoms of other diseases such as depression and anxiety (NUTT *et al.*, 2022; ABAD; GUILLEMINAULT, 2005).

The previous-presented method (clustering and labeling-based method) is not able to measure sleep quality because more sleep time does not imply higher sleep quality (on the contrary, too much sleep can indicate health problems). Therefore, after a literature analysis, it was decided to adapt the deterministic method proposed by ARORA *et al.* (2020). The adjustments were to remove the sleep onset latency measurement, as it was unfeasible to identify the difference between the time the user went to bed and the time the user indeed slept.

Furthermore, the scale proposed by ARORA *et al.* (2020) was inverted and mapped to the Likert scale in order to align with the results achieved by the clustering and labeling method. ARORA *et al.* (2020) method was chosen because it is deterministic and presents values smoothly adaptable to the Likert scale.

Table 3 – Sleep metrics based on ARORA *et al.* (2020) work.

ID	Metric	Unit	Equation	Transformation
1	Sleep Efficiency	Percentage	$SEff = \frac{Sleep\ Time}{Time\ in\ Bed} \rightarrow$	$\begin{cases} 1, & \text{if } SEff \leq 0.65 \\ 2, & \text{if } 0.65 < SEff \leq 0.75 \\ 3, & \text{if } 0.75 < SEff \leq 0.85 \\ 4, & \text{if } SEff > 0.85 \end{cases}$
2	Sleep Duration	Hours	$SDur = Light + Deep + REM \rightarrow$	$\begin{cases} 1, & \text{if } SDur \leq 5 \\ 2, & \text{if } 5 < SDur \leq 6 \\ 3, & \text{if } 6 < SDur \leq 7 \\ 4, & \text{if } SDur > 7 \end{cases}$
3	Sleep Disturbance	Minutes	$SDis = Awake \rightarrow$	$\begin{cases} 4, & \text{if } SDis \leq 20 \\ 3, & \text{if } 20 < SDis \leq 30 \\ 2, & \text{if } 30 < SDis \leq 40 \\ 1, & \text{if } SDis > 40 \end{cases}$
4	Deep Sleep	Percentage	$DPer = \frac{Deep}{Sleep\ Time} \rightarrow$	$\begin{cases} 1, & \text{if } DPer \leq 0.1 \\ 2, & \text{if } DPer > 0.1 \end{cases}$
5	REM Sleep	Percentage	$RPer = \frac{REM}{Sleep\ Time} \rightarrow$	$\begin{cases} 1, & \text{if } RPer < 0.2 \text{ or } RPer > 0.25 \\ 2, & \text{if } 0.2 \leq RPer \leq 0.25 \end{cases}$

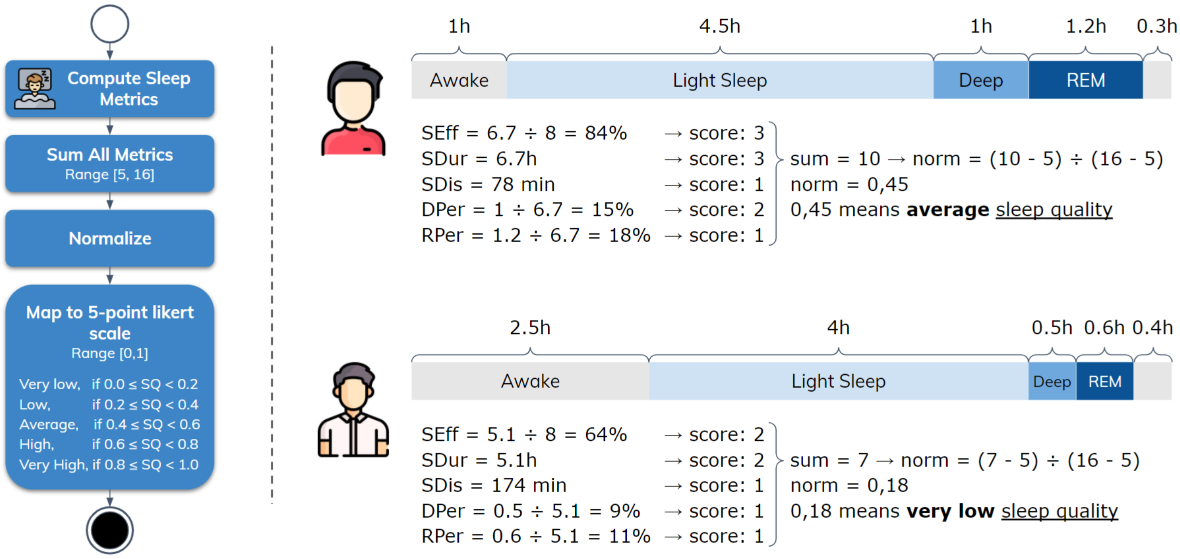
Source: author.

It is also essential to detail how the measurements are collected during the user's sleep. Through the sensors present in most commercial smart bands and smartwatches, it is possible to get an estimation for the time that the user spent awake in bed (awake time); the time relative to the initial stage of the sleep cycle, in which the muscles begin to relax (light sleep); the relative time to the restoration sleep stage (deep sleep); and the time relative to the Rapid Eye Movement phase, which is vital to re-energize the mind (REM sleep).

Table 3 presents the metrics calculated based on the user's sleep data. Considering these metrics, it is possible to use the method shown in Figure 30 for calculating sleep quality. After computing the metrics, they are all summed, normalized, and mapped to the Likert scale.

Also, in Figure 30, two applications of the method are presented. In the first case, 4.5 hours of light sleep, 1 hour of deep sleep, 1.2 hours of REM sleep, and 1.3 hours of awake sleep were recorded. After applying the method, it was found that the normalized value was equal to 0.45. This value corresponds to "average" sleep quality. In the second case, the user had a shorter time for all sleep stages (light, deep, and REM). Thus, the value found was equal to 0.18, meaning "very low" sleep quality.

Figure 30 – Method for calculating sleep quality and two examples of its application



Source: author.

Naturally, the selected method has limitations, such as focusing only on sleep time to assess its quality and using hard thresholds. In addition, each individual has different sleep needs; therefore, there is no ideal scale. Even so, the metrics proposed by ARORA *et al.* (2020) can help users to understand their sleep dynamics using less-intrusive devices, such as smart bands and smartwatches.

4.5 Discussion

Chapter 4 presented the Healful platform, which was built to infer users' QoL ubiquitously using IoHT and Machine Learning algorithms. With the knowledge acquired during the development of this platform, it was possible to draw an answer for the last Research Question (RQ3). Although many requirements can be listed to answer this question, four with high criticality were selected based on the decisions taken throughout the project.

RQ3: What requirements are important to design and implement an IoHT platform for ubiquitous QoL monitoring?

Architectural requirement: it is indispensable to include a strategy to deal with the lack of interoperability among the different devices in IoHT environments. The absence of public APIs for data extraction worsens this issue. Because of that, in the Healful platform, it was necessary to adopt the Health & Fitness Data Container called Google Fit. An alternative for the iOS ecosystem is the Research Kit framework. Furthermore, due to the need to store and process a large volume of data, the use of NoSQL storage and integration with data analysis tools is recommended. In the case of the Healful platform, MongoDB was adopted for data storage, and Athena as a support tool for data analysis.

Ubiquity requirement: it is crucial to ensure that data collection will occur calmly (*i.e.*, less intrusive as possible) during the user's daily activities. It should also be possible to define the best time to receive daily QoL reports and health promotion notifications. The aim is to guarantee that users' monitoring is integrated into their day-to-day, making it almost imperceptible. In this scenario, the Healful platform employs a native Android application capable of extracting data from wearable devices through the Google Fit platform. Every day, users receive a notification in the morning to analyze their QoL report.

Machine Learning requirement: the need for a large amount of data to comprehend user behavior must be observed, in addition to test different algorithms by applying hyperparameter tuning and feature selection. After obtaining a reference model, it is recommended to implement a pipeline that guarantees the model evolution as new data is collected.

Security and privacy requirement: this requirement is vital for this kind of application. For example, the registering process of the QoL Monitor on the Play Store was prolonged, given the need to satisfy all privacy restrictions imposed by Google. Due to the criticality of this requirement, Healful uses asymmetric (RSA) and symmetric (AES-256) encryption to strengthen the security of the data communicated between the frontend and backend, in addition to the use of SSL certificates in both subsystems.

5 EVALUATION

GHOSH *et al.* (2022) stated that "your smartphone knows you better than you think". This statement is based on the possibility of building an individual's digital footprint by analyzing IoHT data, and this same possibility was considered in the hypothesis of this work, as follows.

An IoHT platform that uses Machine Learning algorithms can infer users' Quality of Life for the physical and psychological domains, making QoL monitoring less intrusive when compared with self-reported QoL questionnaires.

To test this hypothesis (presented in Chapter 1), it is necessary to i) validate at which level Machine Learning algorithms can infer users' Quality of Life (focusing on the physical and psychological domains) and ii) check users' perceptions about the platform ubiquity level.

In addition, as this work included five health indicators to support the interpretation of inferred QoL scores, it is essential to get feedback from healthcare practitioners and eHealth professionals about these indicators.

Based on this context, this chapter is structured as follows: Section 5.1 presents a case study conducted to evaluate Machine Learning regressors for QoL; Section 5.2 presents a survey with end-users to get feedback about their perception after using the Healful platform; Section 5.3 presents another survey with healthcare practitioners and eHealth researchers to get feedback about the five health indicators included in this work; Section 5.4 discusses some challenges and limitations of the conducted evaluations; Section 5.5 summarizes the lessons learned from the mobile health monitoring in the "wild"; and, finally, Section 5.6 summarizes the impact of the achieved results.

5.1 Case Study for Machine Learning Regressors

As previously mentioned, the central hypothesis of this work seeks to validate at what level the Healful platform can infer users' QoL using Machine Learning models created with IoHT data when compared with self-reported questionnaires. Therefore, a case study was conducted to investigate this phenomenon in a real-life context (WOHLIN, 2012; KITCHENHAM *et al.*, 2015). The **purpose** of this evaluation was to analyze the QoL inference in the physical and psychological domains, using data collected from users' smartphones and wearables.

Materials and Methods

From the goal presented at the beginning of this selection, the Root Mean Squared Error (RMSE) was selected as the primary **metric**[1] to check the assertiveness level of the Healful regressors. RMSE is the most commonly used measure for regression tasks since it has the same dimensions as the predicted value itself, making the interpretation process easier (IAN; EIBE, 2005). However, defining a hard comparative threshold was necessary since the Healful platform models would be compared with the values obtained by the WHOQOL-BREF questionnaire, and a perfect fit is unfeasible. In the literature, no reference with similar thresholds was found (OLIVEIRA *et al.*, 2022a; OLIVEIRA *et al.*, 2022b). Therefore, the level of 10% of error was selected as a reference for the Quality of Life scores on a scale of values ranging from 0 (worst score) to 100 (best score). Thus, it was possible to build the null and alternative **hypotheses of the case study**:

- Null hypothesis H_0 (*inference*): an IoHT platform that uses Machine Learning algorithms (in this case, the Healful platform) can not infer users' Quality of Life for the physical and psychological domains, considering a Root Mean Squared Error (RMSE) less than or equal to 10% of the reference value (WHOQOL-BREF questionnaire).
  H_0 (*inference*): RMSE(Healful) > 10% of the actual WHOQOL-BREF score
- Alternative hypothesis H_1 (*inference*): an IoHT platform that uses Machine Learning algorithms (in this case, the Healful platform) can infer users' QoL for the physical and psychological domains considering a Root Mean Squared Error (RMSE) less than or equal to 10% of the reference value (WHOQOL-BREF questionnaire).
  H_1 (*inference*): RMSE(Healful) $\leq$ 10% of the actual WHOQOL-BREF score

As **independent variables**, five Machine Learning algorithms implemented using the Scikit-learn (PEDREGOSA *et al.*, 2011) were selected: Linear Regression, Decision Tree Regressor, Random Forest Regressor, GBoost Regressor, and Extra Trees Regressor. The choice of these algorithms followed the GÉRON (2019) guidelines, starting from a more straightforward algorithm to a more robust algorithm. For this evaluation, it was also applied a randomized search on hyperparameters[2] and a feature selection[3] based on their relevance.

[1] Although RMSE was chosen as the primary metric, Mean Absolute Error (MAE) and the training time in seconds were also collected to understand how far the predictions are from the correct output and the time needed to build each Machine Learning model.

[2] Randomize Search with Scikit-learn: scikit-learn.org/RandomizedSearchCV.

[3] Feature Selection with Scikit-learn: scikit-learn.org/SelectKBest.

Data collection was conducted in two moments. Initially, 20 participants were recruited for a three-month evaluation (starting on March 14 and ending on June 14, 2022). Then, 24 new participants (eight undergraduate students from the Federal University of Ceará, Brazil, and 16 undergraduate students from the Federal University of Piauí, Brazil) were recruited for another three-month evaluation (starting on October 10, 2022, and ending on January 10, 2023).

Regarding the **participants' selection**, it was applied the following criteria: i) age between 18 and 65 years old; ii) prior knowledge of the use of smartphones and/or smartwatches; iii) availability for continuous use of a specific wearable. In addition, the participants' invitation was based on convenience, prioritizing people who had a smart band or smartwatch. This criterion was essential to reduce device acquisition costs.

After accepting the invitation, it was conducted the **study starting procedure**, which is composed of six steps: i) reading and agreeing to the free and informed consent form; ii) answering the WHOQOL-BREF questionnaire in the presence of the responsible researcher to clarify possible issues about the questions; iii) configuring the smartwatch or smart band to sync data with Google Fit; iv) install the QoL Monitor application; v) grant the necessary permissions to monitor health data; and vi) effectively initiate monitoring. After this initial procedure, the participant was instructed to follow their personal daily activities typically.

As shown in Figure 24 (see Chapter 4, Section 4.3), IoHT data were collected daily and sent anonymously to the cloud. Weekly, the app warned the participant to answer the WHOQOL-BREF questionnaire (inside the QoL Monitor app) only with questions related to the physical and psychological domains. Thus, it was possible to build a dataset[4] with 1,373 instances after data pre-processing. In addition, it is essential to highlight that this investigation (registered under the ID number 56153322.0.0000.5054) was approved by the ethics committee of the Federal University of Ceará (UFC) on March 9, 2022 (legal opinion number 5.282.056).

Finally, after data processing, each Machine Learning algorithm was coded in Python using Jupyter notebooks[5] hosted on the Kaggle platform[6] and was run 30 times using 10-fold cross-validation, resulting in 300 fits. A read-only version of the codes can be accessed via the link github.com/great-ufc/healful-thesis/.

[4] This dataset is registered on the Kaggle platform and will remain private until the publication of a paper detailing how it was built and how it can be used.

[5] The Jupyter Notebook is an open-source web-based interactive computing platform. A notebook can be written in Julia, Python, or R and combines live code, equations, narrative text, and visualizations.

[6] Jupyter notebooks hosted on Kaggle run in a remote computational environment. Each running session has 12 hours of execution time for the CPU and 20 Gigabytes of auto-saved disk space. CPU Specifications: 4 CPU cores and 30 Gigabytes of RAM. This information is publicly available at kaggle.com/docs/notebooks.

Results and Discussion

Table 4 describes the participants' profiles considering sociodemographic aspects collected by the QoL Monitor application. In a nutshell, this profile is composed of thirty-three (33) men and eleven (11) women aged between 19 and 47 years. Approximately 77% of participants are single (34) and 23% are married (10). Most characterize themselves as university students. Regarding income, twenty-two (22) participants reported receiving between 0 and 1 minimum wages[7] and only one claimed to live in a rural area. Concerning the family arrangement, most participants live with 1 or 2 more people at home, and there are two large groups in terms of the number of children: those who do not have children (35 participants) and those who have 1 or 2 children (9 participants).

Table 4 – Participants' profile.

Gender	Female	11	25.00%
	Male	33	75.00%
Age	18-29	31	70.45%
	30-39	11	25.00%
	40-49	2	4.55%
Marital Status	Single	34	77.27%
	Married	10	22.73%
Educational Level	Secondary	10	22.73%
	Undergraduate	21	47.73%
	Graduate	13	29.54%
Group	Initial set	20	45.45%
	UFC	8	18.18%
	UFPI	16	36.37%
Wearable	Mi Band	37	84.09%
	AmazFit	4	9.09%
	Galaxy Fit	1	2.27%
	Galaxy Watch	1	2.27%
	P70-Pro	1	2.27%

Profession	Part-time worker	5	11.36%
	Self-employed	2	4.55%
	Student	24	54.55%
	Full-time worker	13	29.54%
Income	0 to 1	22	50.00%
	2 to 4	14	31.81%
	5 to 7	2	4,56%
	8 to 10	5	11.36%
	More than 10	1	2.27%
Children	None	35	79.55%
	1 to 2	9	20.45%
Residence	Rural	1	2.27%
	Urban	43	97.73%
Family Arrangement	Lives alone	3	6.82%
	Lives with more 1 or 2	18	40.91%
	Lives with more 3 or 4	16	36.36%
	Lives with 5 or more	7	15.91%

Source: author.

Concerning wearable devices, thirty-six (36) Xiaomi Mi Band 5 devices were acquired by the researcher and distributed to the participants. Then, the remaining participants (8) joined the study using their own devices.

[7] For this collection, Brazilian minimum wage was considered to be R$1,100.00 reais.

It is important to mention that, in an ideal scenario, it would be interesting for everyone to use the same kind of device. However, in the real world, this is unfeasible. Thus, it was decided to allow these participants to join the study, aiming to analyze the impacts caused by data acquired by sensors of different brands and models. As for the wearable acquisition by the researcher, the Xiaomi Mi Band 5 was chosen for its low cost (approximately 40 dollars in Brazil) and the possibility of syncing data with Google Fit in a native way (*i.e.*, without the need of middleware applications).

Table 5 summarizes the initial results achieved in this case study. In both datasets (physical and psychological), the training time tends to grow as the classifier complexity increases. Naturally, the errors (the metric to be minimized) tend to decrease using more robust regressors. Analyzing the data in Table 5, it is possible to observe that the Extra Trees regressor performed better considering the MAE and RMSE metrics. Regarding training time, the Extra Trees regressor is the shorter one among the three algorithms with the best error metrics.

Table 5 – Case study initial results regarding MAE, RMSE, and training time (measured in seconds) for the physical and psychological QoL datasets.

ML Techniques	**Physical Dataset**			**Psychological Dataset**		
	MAE	RMSE	Time	MAE	RMSE	Time
Linear Regression	9.5658	14.4308	0.7544	10.6868	17.6120	0.8286
Decision Tree	6.9889	10.4243	1.4479	6.8111	10.5715	1.5317
Random Forest	5.6870	8.0745	92.0384	5.4534	7.7493	98.3695
GBoost	6.0078	8.1860	528.8100	5.7768	8.0693	438.2732
Extra Trees	5.3672	7.4918	16.8884	5.1965	7.3320	16.7467

Note: The detailed results obtained for each dataset instance can be accessed in the thesis repository: github.com/great-ufc/healful-thesis.
Source: author.

However, more than analyzing the mean of MAE and RMSE metrics is needed to validate the hypotheses defined in this case study. Thus, it is crucial to conduct robust statistical tests to reject or accept the hypotheses. To perform these tests, Origin Pro 9.1 software was employed to guarantee consistency and reduce the bias (SEIFERT, 2014). Therefore, raw metrics were first submitted to the Anderson-Darling normality test (SCHOLZ; STEPHENS, 1987) to verify if they follow the Gaussian distribution. This result is essential for selecting appropriate hypothesis tests (OTT; LONGNECKER, 2015).

Table 6 brings the results of the Anderson-Darling test considering the RMSE metric (target of the hypothesis test described at the beginning of the section). Based on these results, it is possible to state – with a confidence level of 95% – that the RMSE of the Linear Regression, GBoost, and Extra Trees do not follow the Gaussian distribution. Hence, the hypotheses tests to verify which regressor obtained the best result need to be non-parametric (OTT; LONGNECKER, 2015) and the Kruskal-Wallis test was selected because it can check whether three or more independent samples come from the same population (MCKIGHT; NAJAB, 2010).

Table 6 – Normality check using the Anderson-Darling test in the RMSE metric.

Dataset	**Algorithm**	$\overline{X}$	$\sqrt{\sigma^2}$	α	ρ**-value**	**Decision**
Physical	Linear Regression	14.4308	0.0000	0.05	<0.0001	Reject normality
	Decision Tree	10.4243	0.1763	0.05	0.64241	Can not reject normality
	Random Forest	8.07451	0.0306	0.05	0.37311	Can not reject normality
	GBoost	8.1860	0.0000	0.05	<0.0001	Reject normality
	Extra Trees	7.4918	0.0000	0.05	<0.0001	Reject normality
Psychological	Linear Regression	17.6120	0.0000	0.05	<0.0001	Reject normality
	Decision Tree	10.5715	0.1909	0.05	0.96359	Can not reject normality
	Random Forest	7.7493	0.0357	0.05	0.51094	Can not reject normality
	GBoost	8.0693	0.0000	0.05	<0.0001	Reject normality
	Extra Trees	7.3320	0.0000	0.05	<0.0001	Reject normality

Note: $\overline{X}$ - average of metrics; $\sqrt{\sigma^2}$ - standard deviation; and α - significance level.
Source: author.

Kruskal-Wallis test found – with a confidence level of 95% – that RMSE samples come from different populations (ρ-value < 0.0001). As a posthoc analysis, it was applied Dunn's test (DINNO, 2015) for non-parametric pairwise comparisons. In the comparisons, the Random Forest and the Extra Trees have a significant difference of means when compared with Linear Regression, Decision Tree, and GBoost regressors (ρ-value < 0.0001 for both datasets). However, the mean of RMSE obtained by the Extra Trees and the Random Forest is not statistically different (ρ-value = 0.07017 for both datasets). This implies that both regressors have statistically the same performance (based on RMSE).

Nevertheless, after a meticulous analysis of the Extra Trees model, an over-fitting scenario (HAWKINS, 2004; YING, 2019) was observed even with the use of cross-validation, given that the R^2 metric for this model was equal to 1.0 (perfect fit). Therefore, the Random Forest regressor was chosen as the reference in this work.

After getting a good candidate for the model that would be included within the Healful platform, it was decided to apply methods for optimizing the settings of the Random Forest. These settings – commonly called hyperparameters - must be carefully chosen due to their impact on model performance (BISCHL *et al.*, 2023).

To optimize the Random Forest, the random search method (BERGSTRA; BENGIO, 2012) was employed to randomly select hyperparameters until a specific stopping condition (in this case, 30 executions of a 10-fold cross-validation, resulting in 300 fits). Although there are other methods for hyperparameter optimization, such as grid search, genetic algorithms, evolutionary algorithms, and Bayesian optimization, the random search was selected due to its ability to explore a vast space of possibilities without suffering from the "dimensionality curse"[8] since the number of possibilities grows exponentially (HUTTER *et al.*, 2019).

After applying the random search method, a 1.19% improvement in the RMSE metric was obtained for the physical dataset and 3.01% for the psychological dataset. The initial metric for the physical dataset was 8.0745 and became 7.9793, while the initial metric for the psychological dataset was 7.7493 and became 7.5162. The hyperparameters selected for each dataset are presented as follows. Although the Scikit Learn library documentation[9] details how each parameter operates, it is worth highlighting the high number chosen for n_estimators, which represents the number of trees in the forest (evidencing the high complexity of the models), and the low number selected for min_samples_leaf (*i.e.*, the minimum number of samples required to be at a leaf node) that reinforces the attempt to smooth the regression model.

Selected hyperparameters for physical dataset

```
{'warm_start': False,
 'n_estimators': 894,
 'min_samples_split': 5,
 'min_samples_leaf': 2,
 'max_features': 1.0,
 'max_depth': 10,
 'criterion': 'squared_error',
 'bootstrap': True}
```

Selected hyperparameters for psychological dataset

```
{'warm_start': False,
 'n_estimators': 1778,
 'min_samples_split': 2,
 'min_samples_leaf': 1,
 'max_features': 'sqrt',
 'max_depth': None,
 'criterion': 'poisson',
 'bootstrap': False}
```

To conclude refining the model, feature selection experiments were conducted to improve the RMSE metric of the Random Forest. In general, features can be relevant, irrelevant,

[8] Course of dimensionality by Encyclopedia of Machine Learning of Springer: Springer link.
[9] Random Forest: scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html.

or redundant (KARAGIANNOPOULOS *et al.*, 2004), and there are many methods focused on this task (JOVIĆ *et al.*, 2015). In this work, SelectKBest[10] was applied due to its relevance in practical problems (DESYANI *et al.*, 2020; RATMANA *et al.*, 2020). The experiment evaluated three scoring functions (f_regression, mutual_info_regression, chi2) suitable for regression tasks and several different numbers of features to be selected (starting with the default value of 10 and proceeding up to the maximum number of features [88] in a step of 10).

Table 7 presents the results – related to the feature selection experiments – in which improvements were observed for the Random Forest RMSE metric. The best improvement was achieved with the mutual_info_regression function selecting 70 of 88 features in the physical dataset (1.47%) and 50 of 88 features in the psychological dataset (0.76%). This scoring function measures the mutual dependence between two variables and relies on entropy estimation from k-nearest neighbors distances (KOZACHENKO; LEONENKO, 1987; ROSS, 2014). At this point, RMSE metric for the physical dataset was 7.9793 and became 7.8618, while the RMSE metric for the psychological dataset was 7.5162 and became 7.4591.

Table 7 – Results of feature selection experiments with SelectKBest.

Dataset	**Scoring function**	**Selected features**	**RMSE without Feature Selection**	**RMSE with Feature Selection**	**Improvement**
Physical	f_regression	80	7.9793	7.9223	0.71%
	mutual_info_regression	40	7.9793	7.9677	0.15%
		50	7.9793	7.8850	1.18%
		60	7.9793	7.8673	1.40%
		70	7.9793	7.8618	1.47%
		80	7.9793	7.8964	1.04%
Psychological	mutual_info_regression	20	7.5162	7.4996	0.22%
		40	7.5162	7.4999	0.22%
		50	7.5162	7.4591	0.76%
		60	7.5162	7.4947	0.29%
		70	7.5162	7.4794	0.49%

Source: author.

Table 8 presents the twenty (20) most relevant features (in both datasets) for Random Forest considering the accumulation of the impurity[11] decrease within each tree. Therefore, the greater the importance of the feature, the greater its ability to increase the purity of the leaves (LOUPPE, 2014). Thus, analyzing the data presented in Table 8, it possible to find insights,

[10] SelectKBest: scikit-learn.org/ stable/modules/generated/sklearn.feature_selection.SelectKBest.html.
[11] Feature importance evaluation: scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html

such as that these most relevant features can be classified into six groups: anthropometric and sociodemographic features (represented in light orange), features based on heart rate (represented in light green), features based on apps usage (represented in light yellow), features related to physical activities (represented in light red), sleep features (represented in light blue) and features related to social component (represented in light purple).

Table 8 – Twenty most relevant features according to Random Forest.

Physical Dataset			Psychological Dataset		
1	Weight	0.3075	1	Height	0.1072
2	Height	0.2928	2	The standard deviation of differences between adjacent RR-intervals (SDSD)	0.0962
3	Root mean square of successive RR interval differences divided by the mean of RR-intervals (CVSD)	0.0652	3	Weight	0.0829
4	Age	0.0599	4	Time in light sleep	0.0413
5	Time using shopping apps	0.0362	5	Median Absolute values of the successive differences between the RR-intervals	0.0407
6	Standard Deviation of Heart Rate	0.0275	6	The proportion of the number of interval differences of successive RR-intervals greater than 50 ms by the total number of RR-intervals (PNNI 50)	0.0392
7	Income	0.0148	7	Averageduration of incoming calls	0.0358
8	Average of Heart Rate	0.0130	8	Number of missed calls	0.0332
9	Time using art apps	0.0114	9	The square root of the mean of the sum of the squares of differences between adjacent NN-intervals	0.0304
10	Time in lifestyle apps	0.0096	10	Root mean square of successive RR interval differences divided by the mean of RR-intervals	0.0275
11	Time using social apps	0.0077	11	Time in lifestyle apps	0.0252
12	Calories	0.0076	12	Steps	0.0238
13	The mean of RR-intervals	0.0071	13	Number of rejected calls	0.0197
14	Number of different WiFi SSID connected	0.0070	14	Maximum of Heart Rate	0.0197
15	Time in REM sleep	0.0063	15	Calories	0.0186
16	Time in deep sleep	0.0053	16	The number of interval differences of successive RR-intervals greater than 20 ms by the total number of RR-intervals (PNNI 20)	0.0184
17	Average duration of incoming calls	0.0051	17	Number of incoming calls	0.0169
18	Steps	0.0049	18	Minimum of Heart Rate	0.0159
19	The proportion of the number of interval differences of successive RR-intervals greater than 50 ms by the total number of RR-intervals (PNNI 50)	0.0049	19	Number of Whatsapp incoming voice call	0.0153
20	Number of Whatsapp incoming voice call	0.0047	20	Time using social apps	0.0145

Note: the appendix C provides more details about each feature.
Source: author.

Regarding the group of anthropometric and sociodemographic features, there are many studies correlating them with QoL scores. Aspects such as height and weight influence the body mass index, which is related to the person's health. Similarly, there is also an influence of income and age on Quality of Life (GOODE *et al.*, 2016; SAINTILA *et al.*, 2020; TOZETTO *et al.*, 2021; FAN *et al.*, 2022; VÁZQUEZ-LORENTE *et al.*, 2023).

As for the features group based on heart rate, these features – calculated by the HRV analysis library[12] – are associated with users' stress level (KIM *et al.*, 2018) and this stress level is strongly correlated with physical and psychological Quality of Life because it reflects mental effort with a given activity (CIABATTONI *et al.*, 2017; BERDIDA; GRANDE, 2023; DOURIS *et al.*, 2023). However, it is worth noting that i) the HRV features and their association with stress level represent a broad research field; and ii) this work did not delve into calculating and selecting these features, including only those available in the HRV analysis library. Chapter 6, Section 6.3 discusses this research opportunity.

Concerning the group of features based on apps usage report, there is evidence that mobile phone data can be used as a predictor of health issues, such as loneliness and depression, which are strongly related to users' QoL (LI *et al.*, 2016; YANG *et al.*, 2023). For example, ADAM; ALHASSAN (2021) stated that "students addicted to mobile phone had significantly lower scores across all QoL domains". In addition, similar to the apps usage group, the social component group has features that strongly influence psychological health (SANCHEZ *et al.*, 2015; DAHLBERG *et al.*, 2022).

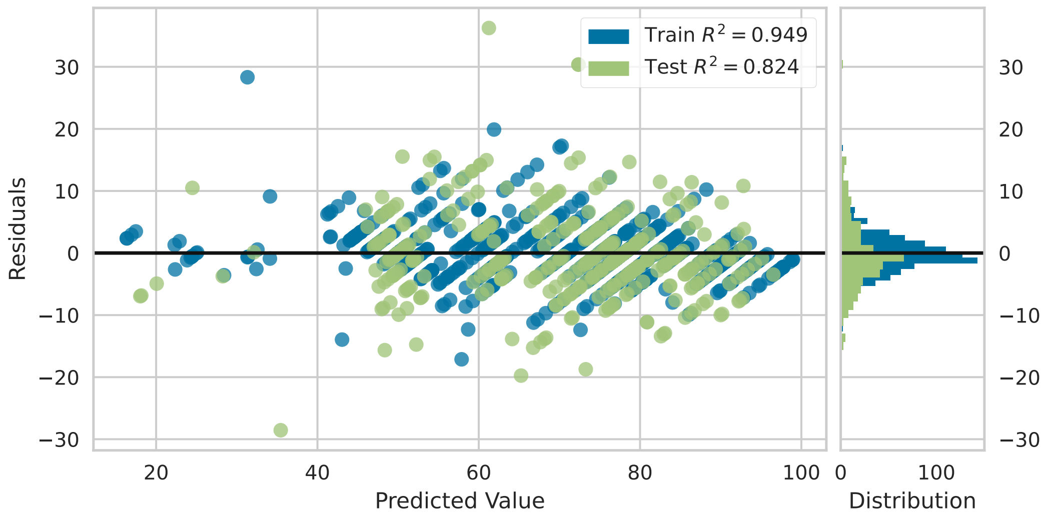
To conclude the analysis of feature groups, the last two groups – sleep and physical activities – have features strongly related to both datasets. For example, a relationship between QoL score and sleep quality was expected, given that deep sleep is considered the stage of restoring muscles and removing waste from the brain, and the REM stage re-energizes the mind (WORLEY, 2018; ARORA *et al.*, 2020; STOJANOV *et al.*, 2021). Regarding physical activities, features such as steps and calories can be used to calculate users' daily mobility, which is related to both physical and mental health (VALLANCE *et al.*, 2016; KRAUS *et al.*, 2019; PANICKER; CHANDRASEKARAN, 2022).

Although such features were selected by Random Forest as the most relevant, it is noteworthy that when analyzing Pearson's correlation between the dataset features and the value to be predicted (QoL score), all calculated correlations were less than 0.39, which represents a weak or negligible correlation (SCHOBER *et al.*, 2018). Therefore, based on these results, it is not possible to determine a decisive subset of features to infer users' Quality of Life.

[12] HRV analysis library: github.com/ Aura-healthcare/hrv-analysis.

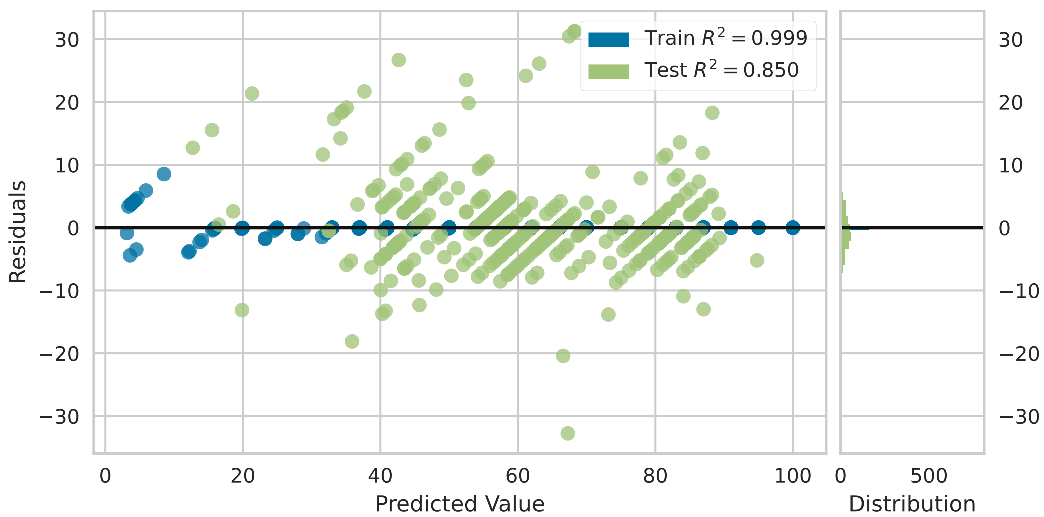
Finally, Figures 31 and 32 show residual plots for the Random Forest applied in physical and psychological datasets, respectively. In these graphics, residuals represent the difference between the actual value (obtained by the WHOQOL-BREF questionnaire) and the predicted value (obtained by the regressor). In this view, it is possible to see that the regressor error is concentrated in the range $[-10, +10]$. Even with outliers in the test subset, the error histogram is normally distributed around zero.

Figure 31 – Residual plot for Random Forest in the physical dataset



Source: author.

Figure 32 – Residual plot for Random Forest in the psychological dataset



Source: author.

Therefore, considering that it was proved the possibility to build a regressor (Random Forest) whose RMSE metric was, on average, 7.8618 for the physical dataset and 7.4591 for the psychological dataset, it is possible to state that this model can estimate users QoL considering an average error of approximately 8 points more or less. This interpretation is obtained from how the RMSE metric is calculated in the equation 4.2 and based on Machine Learning literature knowledge (IAN; EIBE, 2005). Furthermore, Figures 31 and 32 corroborate this interpretation by showing that the errors (residuals) are concentrated in the interval $[-10, +10]$.

Hence, the null hypothesis of this case study was rejected. Because of that, it is possible to state that an IoHT platform that uses Machine Learning algorithms can infer users' QoL for the physical and psychological domains considering an RMSE less than or equal to 10% of the reference value (in this case, WHOQOL-BREF questionnaire). This result implies that it is possible to use the proposal of this work as a complementary element in the medical practice, allowing daily QoL monitoring. Although the QoL score is influenced by model error, as this error is within a range of 10% more or less than the actual value, such inference can be used as an indicator to find healthier habits or even to contact medical care to verify critical situations, such as low-level scores for a long time.

5.2 Survey with End-Users

The hypothesis of this work – reiterated at the beginning of Chapter 5 – emphasizes the desire to make QoL monitoring less intrusive when compared to self-reported strategies. As stated by PEQUENO *et al.* (2020), self-reported questionnaires have challenges related to misunderstanding, biases inserted by the participants, and the difficulty of engaging the respondents. Therefore, at the end of each data collection period, a final survey (KITCHENHAM *et al.*, 2015) was conducted with the **purpose** of getting feedback about users' perceptions. These perceptions are essential to understand how intrusive the solution presented in this work is.

Materials and Methods

Based on the purpose mentioned earlier, a survey was created based on Technology Acceptance Model 3 (TAM3) (VENKATESH; BALA, 2008) to collect feedback about the study and the tool used to monitor users' QoL. TAM helps researchers and practitioners to understand aspects related to the adoption of new technologies considering two beliefs: perceived usefulness and perceived ease of use. The first belief reflects how much the user believes in the usefulness of the technological solution, and the second focuses on how much effort the user believes will be free when using the solution (VENKATESH; BALA, 2008).

The applied questionnaire was composed of 13 questions subdivided into four groups, each exploring an aspect present in TAM: perceived usefulness (I), perceived ease of use (II), self-efficacy when using the tool (III), and intention to use the tool (IV). For each question, five possible alternatives were defined based on the Likert scale: I fully agree (A); I partially agree (B); neutral (C); partially disagree (D); I totally disagree (E). Finally, to conclude the questionnaire, an open question was included to get other perceptions about the study.

It is worth mentioning that in the questions, the term QoL Monitor was used as a reference to the Healful platform. This is because the QoL Monitor application is the GUI that end-users have access to. Therefore, the functionalities of the Healful platform are delivered through this app. As previously discussed in Chapter 4, the QoL Monitor allows data collection to infer users' Quality of Life.

Results and Discussion

Figure 33 presents the quantitative results based on participants' answers. The survey was conducted anonymously using Google Forms, and only 38 of the 44 participants answered. The Cronbach's Alpha (FELDT *et al.*, 1987) for this survey was $\alpha = .72$[13]. This measure – Cronbach's Alpha – assesses the internal consistency of the survey, and the value obtained is classified as acceptable (TAVAKOL; DENNICK, 2011).

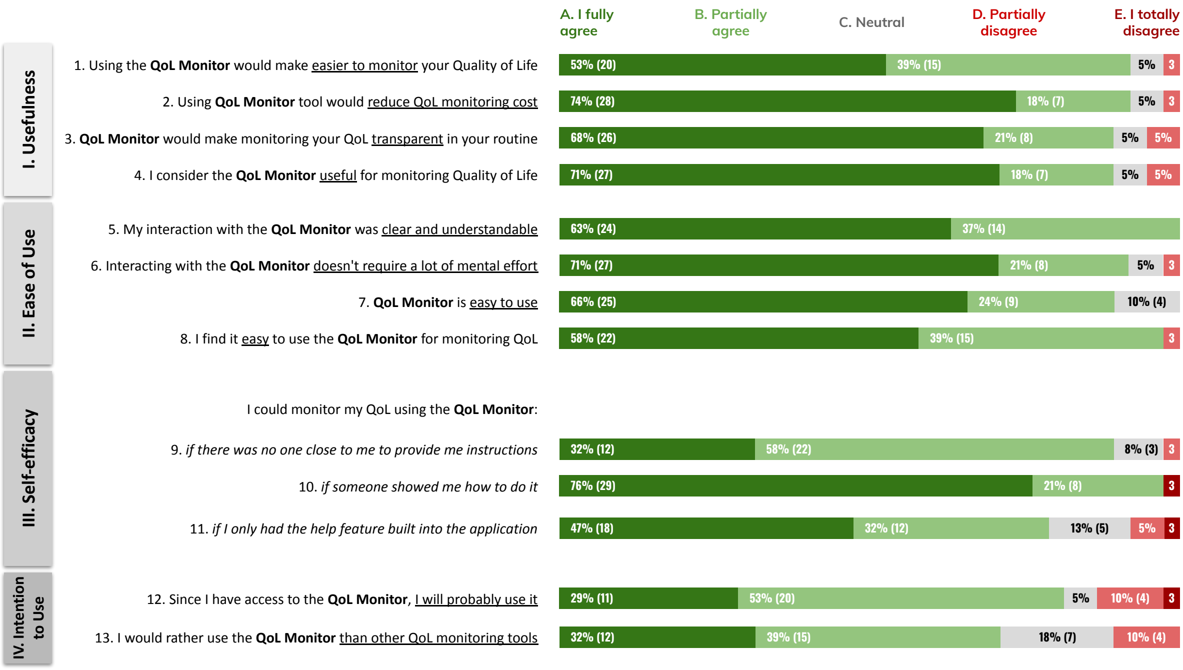
Regarding perceived <u>usefulness</u> (I), most participants fully or partially agreed that the QoL Monitor is helpful since it makes QoL monitoring easier and less costly. Also, most respondents fully or partially agreed that the QoL Monitor has a transparent interaction. This ability is critical for ubiquitous systems as these kinds of systems should be able to hide themselves, promoting interactions in a natural way (CARVALHO *et al.*, 2018). However, there was also disagreement in this group of questions, probably because the proposed solution requires the cost of purchasing smartwatches or smart bands, in addition to the fact that it requires users' feedback through a built-in questionnaire to train the intelligent models.

As for the perceived <u>ease of use</u> (II), most volunteers considered the interaction clear and did not require much mental effort. This result was expected because the QoL Monitor was

[13] Calculated using Pingouin library: pingouin-stats.org/build/html/generated/pingouin.cronbach_alpha.html.

designed to simplify user interactions. However, when analyzing the self-efficacy (III) aspect, which assesses the users' ability to use the tool in situations with little or no prior instruction, it was possible to observe that specific users disagreed about the possibility of monitoring their Quality of Life only with the support of the tool or having the help feature built into the app. This behavior shows that initial training is crucial for users to understand aspects related to QoL monitoring. Finally, to conclude the quantitative results, most participants stated that they would use QoL Monitor again instead of other similar tools.

Figure 33 – Results of the TAM questionnaire



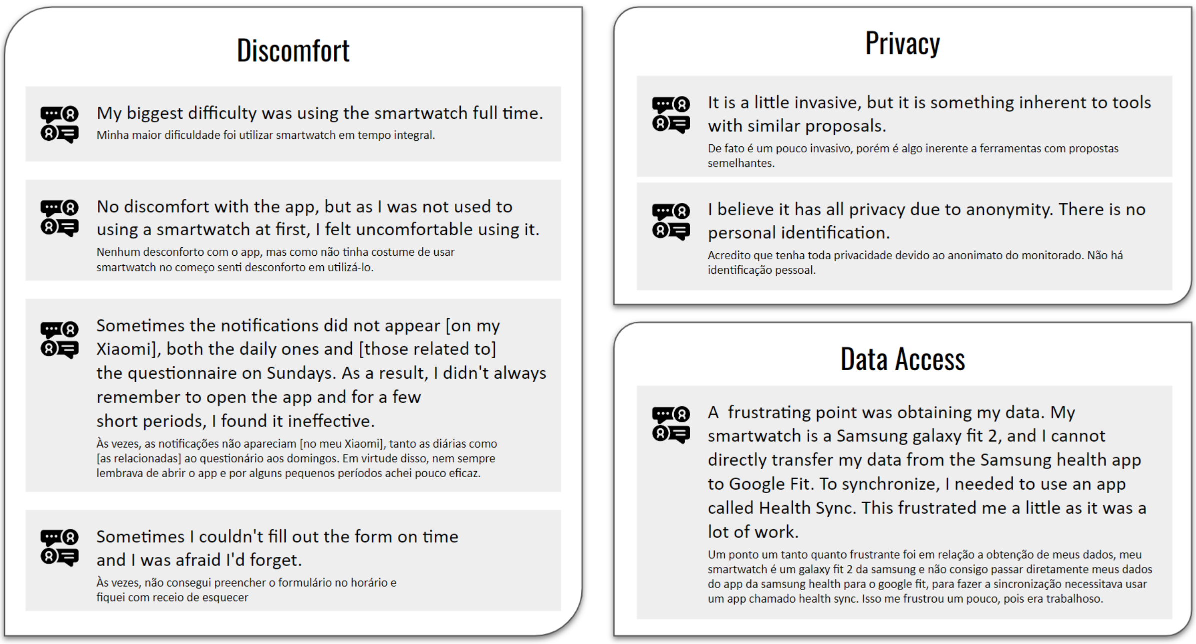
Note: A high-resolution version of this image can be accessed at the following link for better visualization of the questions. Link: github.com/great-ufc/healful-thesis/blob/main/images/6-eval/fig38-tam-evaluation.pdf.
Source: author.

In addition to the quantitative results, qualitative perceptions of the volunteers regarding the difficulties they faced throughout the study were summarized. Such perceptions were organized into three groups: discomfort, privacy, and access to data. Figure 34 presents some comments in Portuguese and its translation into English.

Regarding discomfort, the participants reported difficulties using devices such as smartwatches and smart bands uninterruptedly and keeping the routine of filling in the surveys. In addition, on specific devices, users reported not receiving notifications due to restrictive policies concerning background apps. Such discomforts can induce biases in the collected data, either because the user decided to remove the wearable at night, preventing sleep data from being registered, or the user did not answer the QoL questionnaire due to the absence of notifications. In this case, in addition to optimizing the QoL Monitor to deal with operational system restrictions, it is essential to point out to future users that there is a requirement to wear wearable devices continuously, but the absence of use only implies unreliable results. Upon resuming seamless use, the results should return to the expected quality level (at least 10% of the actual value).

Few participants reported that data collection was a bit invasive. This perception is probably related to the large amount of requested data and the need to grant many permissions. However, Machine Learning models would only perform satisfactorily with this massive data. Therefore, it was tried to comfort participants about data privacy using anonymization and encrypted requests. Finally, a set of participants reported issues in extracting the data from wearables. Usually, commercial wearables do not deliver methods to access their data using APIs, and due to this restriction, the Google Fit platform was selected as the central point to extract user health data. Nevertheless, specific wearables apps did not allow native integration with Google Fit, requiring a third-party app to extract this information. The complexity of this process frustrated volunteers who used Samsung devices, for example.

Figure 34 – Participants' comments about discomfort, privacy, and data access



Note: A high-resolution version of this image can be accessed at the following link for better visualization of the comments. Link: github.com/great-ufc/healful-thesis/blob/main/images/6-eval/fig39-tam-comments.png.
Source: author.

From the data analyzed in this survey, it is possible to state that the second part of the central hypothesis of this study was also satisfied because, in the users' perception, the QoL Monitor app (Healful main interface with end-users) is capable of transforming the QoL monitoring in a transparent process – based on CARVALHO *et al.* (2018) definition – in users' routine. Therefore, it becomes a less intrusive than answering self-reported questionnaires.

5.3 Survey with Health and eHealth Professionals

As discussed in Chapter 4, Section 4.4, the regressors created in this work return a Quality of Life score as an integer that varies between 0 and 100. Naturally, the higher the score, the better the QoL (the reciprocal is also valid). However, when obtaining the QoL score, users should comprehend which aspects of their health need attention.

Having this context as a premise, this work selected five health indicators (daily mobility, physical activity level, loneliness, social mobility, and sleep quality) and presented a clustering-based method to calculate the first four indicators. A survey with health professionals and eHealth researchers was conducted with the **purpose** of validating whether these health indicators and the clustering-based method are suitable as a complement for the QoL regressors. Thus, this section presents how this survey was conducted and its main results.

Materials and Methods

Based on the above-presented goal, a questionnaire with open and closed questions was designed to understand at what is the relationship level between the health indicators selected in this work and the QoL domains (physical and psychological) and at what level can the clustering-based method calculate proper health indicators (daily mobility, physical activity level, loneliness, and social mobility).

The survey sampling considered researchers and practitioners with experience in health or eHealth. Thus, to select the survey target, a probabilistic sampling of the authors of the papers selected in the literature review (OLIVEIRA *et al.*, 2022a) and a non-probabilistic sampling considering the work contacts of the researcher were applied.

Although the participation was entirely voluntary and users could withdraw at any time, a Xiaomi Mi Band 5 was raffled among those who completed the survey to encourage participation. Thus, two hundred and seventy-five (275) researchers and practitioners were contacted by email, but only 59 (21%) answered all the survey questions after two months and three kind reminders. In fact, this percentage is low, but as discussed by KITCHENHAM *et al.* (2015), surveys tend to have fewer responses, and a response rate of 21% can be considered acceptable (KITCHENHAM *et al.*, 2015).

Regarding the data collection instrument, the questionnaire was divided into three parts. The first refers to demographic data, while the second part had questions about the relationship between health indicators and QoL domains, in addition to the applicability of the clustering-based method. Before this part, both the health indicators and the clustering-based method were detailed to the participants. To proceed with the questionnaire, participants should declare whether they correctly understood the presented information. Finally, in the third part, respondents could bring suggestions or recommendations.

Results and Discussion

Figure 35 presents the participants' demographic profile. It is possible to observe that most respondents are Brazilian, although researchers from other countries have also collaborated (A). The educational level (B) is well distributed with at least ten respondents for each level, except "Master's in progress". Regarding professional or study experience (C), most participants have between 6 and 10 years of experience.

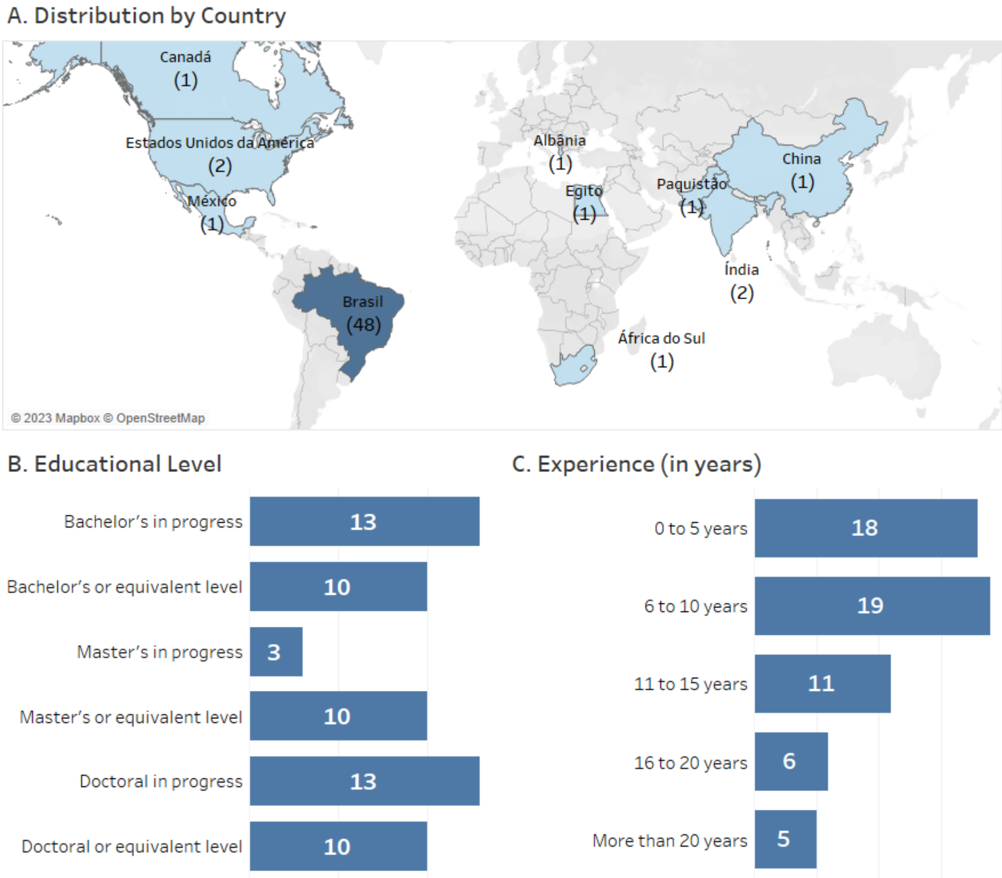
To conclude the survey demographic analysis, it is important to mention that most participants claimed to be a professional in the Information Technology area. Among the student participants, ten are Computer Science students, and seven indicated courses in the health area, such as medicine and biomedical imaging. As for participants who work professionally, there are 12 software engineers, 12 professors, eight researchers, two project managers, two tech leaders, two data scientists, one nurse, one data engineer, and one chair in Dementia Research. Finally, one participant claimed to be unemployed.

Participants' profile indicates that a highly specialized group were selected. Although most respondents are from the technology area (50), all indicated an understanding of the relationship between health indicators and QoL, in addition to the presented clustering-based method. Concerning the internal consistency, the Cronbach's Alpha (FELDT *et al.*, 1987) for this survey was $\alpha = .85$[14], which is classified as good (TAVAKOL; DENNICK, 2011).

Figure 36 shows heat maps with the relationship level between QoL physical domain and health indicators (A) and this same relationship considering QoL psychological domain

[14] Calculated using Pingouin library: pingouin-stats.org/build/html/generated/pingouin.cronbach_alpha.html.

Figure 35 – Survey demographic profile: country, educational level, and experience (in years)



Source: author.

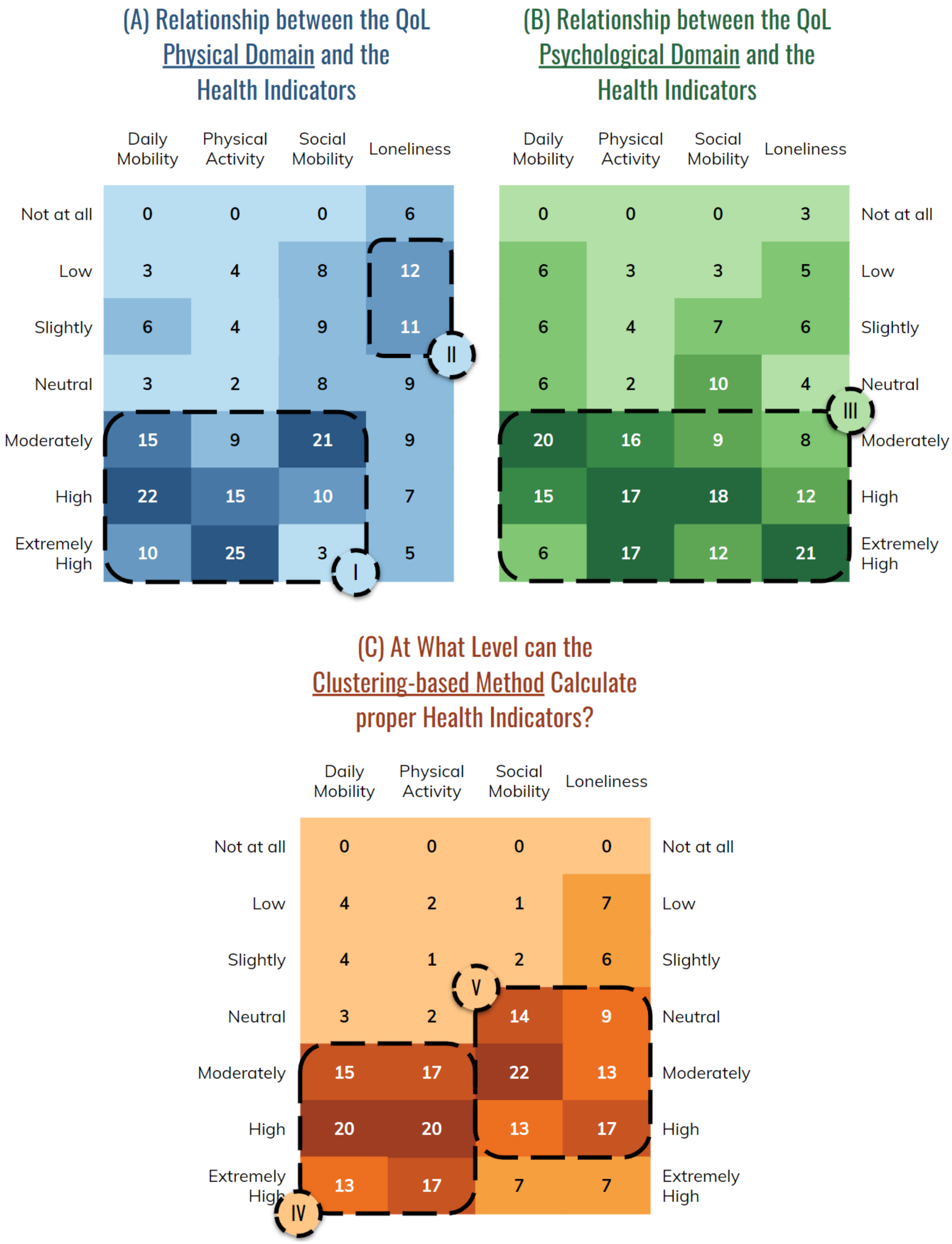
(B). In (C), participants' assessment concerning the ability of the clustering-based method to calculate the health indicators adequately is presented.

Regarding the relationship between the physical domain and the health indicators selected in this work, it is possible to observe (A.I) that most participants agreed that this relationship is from moderate to extremely high for daily mobility, physical activity, and social mobility. However, there was a divergence (A.II) about the loneliness indicator, for which respondents indicated that the relationship level was low. On the other hand, in (B.III), it is possible to observe that all indicators were at least moderately associated with the psychological domain. This result matches the state of the art because many papers indicate that physical health strongly impacts mental health (HERBERT, 2022; TAKIGUCHI *et al.*, 2023).

Therefore, the survey participants agree that the four healthcare indicators selected in this work have at least a moderate relationship with physical and psychological domains, except for loneliness and the physical domain. Hence, such indicators can help users understand which habits impact their QoL scores.

It is worth mentioning that the 5th indicator – sleep quality – was not included in this evaluation because i) it is calculated from a deterministic method already validated (ARORA *et al.*, 2020) and ii) there are many studies correlating sleep and QoL (SELLA *et al.*, 2023). However, it is a future opportunity to check with end-users how effective these indicators are.

Figure 36 – Relationship level between QoL domains (A - physical, B - psychological) and the health indicators, including the participants' analysis about the clustering-based method (C)



Source: author.

Concerning the participants' analysis of the clustering-based method, there was no complete agreement. In (C.IV), most agreed that such a method is at least moderately adequate to calculate daily mobility and physical activity. However, in (C.V), most marked the neutral position. Therefore, according to the experts contacted by this survey, there is room for improvement in the presented method. In future work, it would be interesting to observe the behavior of this method in a longitudinal study followed up by healthcare professionals.

Finally, it is worth mentioning that only 5% of the participants stated that they would not use the method presented in this work. Concerning the others, 30% reported that they could use it and 65% that they would definitely use it when available. In the open field, at the end of the survey, some participants suggested other indicators such as medical history, socioeconomic factors, tiredness, stress, eating habits, family health history, time spent with technology, and interaction level with social networks. The investigation of these other indicators represents a promising future research path, which can start with the report "The good indicators guide"[15] produced by the National Health Service (NHS) of United Kingdom.

5.4 Challenges and Limitations

Challenges and limitations are common in empirical studies (WOHLIN; AURUM, 2015). Their discussion seeks to reinforce the work's reliability since the author presents the main issues and what strategies were used to mitigate them (KITCHENHAM *et al.*, 2015). In addition, this discussion also represents a valuable contribution to researchers and practitioners who work or wish to work in this investigation area. Based on the topics exposed here, it is possible to anticipate or even avoid problems conducting new IoHT studies. This section has organized the discussion of challenges and limitations considering three significant phases of the case study (planning, conducting, and analyzing the results) and the issues faced in the surveys (with end-users and healthcare professionals).

In the case study planning, the first – and probably the most – significant issue faced was the need to build a formal **project for ethics committee**. In many areas, mainly health-related, the construction and submission of a project to the ethics committee is usual, and the researchers are familiar with it. However, in Computer Science, this process is not usual. Therefore, there are many doubts about terminology, study classification, population definition, selection criteria, risks, benefits, and other required documents (*e.g.*, consent form).

[15] The good indicators guide by NHS: fingertips.phe.org.uk/documents/The Good Indicators Guide.pdf.

Also, the submitted project must all comply with current regulations. For example, in Brazil, two resolutions from the National Health Council guide this process (466/2012 and 510/2016)[16]. Finally, the project needs to be registered on the *Plataforma Brasil*[17] system for analysis. To overcome this difficulty, the norms were carefully analyzed with the support of other researchers in the GREat lab. Thus, the project was approved quickly and without reservations. Because of that, an online version of this project was shared to assist future researchers (github.com/great-ufc/healful-thesis/).

The case study conducting phase had the most challenges as follows. The first was related to the **participants' recruitment**. Forty-four (44) adults between 18 and 65 years old with prior knowledge of using smartphones or smartwatches were recruited. Due to restrictions imposed by the COVID-19 pandemic, first, the GREat research group members who already own a smart band or smartwatch were invited. Thus, the first recruitment could be completely remote. However, only six participants met such restrictions. Then, it was necessary to expand the recruitment to close people (considering the researcher's social network). Even so, that number only increased to nine participants. Thus, it was necessary to purchase devices (Xiaomi Mi Band 5) and send them to interested participants. This purchase and shipping logistics delayed the start of data collection and increased the study costs.

The second round of recruitment was conducted after removing many of the COVID-19 pandemic restrictions, making it easier. This second recruitment was focused on Computer Science students at UFPI and UFC. However, it was also necessary to purchase devices to ensure that all recruits could participate. Finally, a few recruited participants withdrew after the initial presentation, citing a lack of time and others having a smartphone incompatible with the QoL Monitor app. Therefore, despite the efforts, the relatively low number of participants (44) in this case study is a limitation that can be addressed in future studies.

After recruiting participants for the case study, data collection started, and many issues related to **noise in data collection** were faced during this step. For example, the absence of an internet connection when sending data, smartphone or wearables unloaded, sensors turned off, sensors and devices with different levels of accuracy. These situations caused noise in the registry, making it difficult to clean the data. Regarding **failures in sending data to the cloud**, it was necessary to implement a mechanism to perform retries on the connection and, after five failed attempts, internally store the data for sending on the next day. As for the **disconnecting**

[16] CONEP laws: http://conselho.saude.gov.br/o-que-e-rss/92-comissoes/conep/normativas-conep
[17] Plataforma Brasil Website: plataformabrasil.saude.gov.br

sensors, participants were warned about the continuous use of the devices and about the need to keep at least Bluetooth and GPS active. However, it was received feedback from participants that the seamless GPS use increased battery consumption. Therefore, participants turned it off in moments of low battery. Consequently, it was necessary to filter out inconsistent data during data processing. Participants were guided to charge their smartphones daily and their wearables weekly regarding the **battery consumption**. However, when charging wearables, a data gap is created. Then, such gaps were removed from the study.

Another limitation faced in data collection is the **non-standardization in using the same kind of device** for monitoring. Although all devices purchased in this work were of the same brand and model (Xiaomi Mi Band 5), some participants were allowed to remain in the case study with their own devices to comprehend the behavior of the QoL regressors in a real scenario. In the real world, assigning the same device to all users is impossible. However, this decision could have messed up the data as different sensors can have different accuracy levels. Therefore, investigating the performance of regressors comparing different groups of devices is a possible future research path. Currently, the dataset built in this work has information about the wearable of the participants, but the groups are pretty unbalanced (see Table 4).

To conclude the challenges faced when conducting the case study, there are **issues in user engagement** and **real-world issues**. Achieving and maintaining user engagement in healthcare technologies is such a broad and complex challenge that many studies have been conducted to find proper strategies to keep users active in an organic way (WANG *et al.*, 2022; GANESH *et al.*, 2022). This study faced engagement challenges, as users had to follow a series of recurring actions, such as opening a wearable app daily to ensure data sync and weekly answering the QoL questionnaire. Even with the application's support to remind these actions, it was observed that at least one-third of the participants failed to answer the questionnaire.

When investigating what could be happening, participants reported that day-to-day activities made them forget. It was also clear that despite recognizing the benefits of daily health monitoring, many participants ignored these benefits and forgot to access the app's notifications. This challenge needs further investigation to understand the real reasons for this lack of engagement and what strategies can be used to overcome it.

Many other difficulties can be classified as **real-world issues**, for example, one participant reported that he lost the smart band during a bath in the sea; two participants reported that the smart band was causing a wrist allergy, and, therefore, they had to stop using it for a few days; another participant caught chikungunya, which affected his joints, preventing him from using the wearable for a few days. These problems are inherent to "in the wild" studies, and there are few strategies to avoid such situations. A specific approach was adopted for each of them, but, in general, all data gaps were eliminated.

In the results analysis phase of the case study, a data variability issue was faced. That happens because the profile of study participants has little variability when considering health issue criteria. Thus, many records have intermediate QoL scores and few high or low scores. Consequently, this impairs the regressors ability to generalize. Therefore, as future work, a new assessment with more participants (up to 100) can be conducted, varying the subjects' profiles.

In the survey with end-users, the main limitation is related to **misunderstandings by participants about the transparency** feature of ubiquitous systems. Although the survey form provided an overview explanation, the term transparency can be confusing. Furthermore, as the questionnaire was inspired by the TAM, other evaluation aspects were included in the questions, such as the intention of use. However, in future evaluations, it would be better to conduct semi-structured interviews to understand how intrusive the solution is in users' routines.

Finally, in the survey with health and eHealth professionals, the main challenge was to get a significant **number of responses**, especially from healthcare practitioners. Considering that the study target population has many professionals, only 59 responses do not have statistical strength for profound conclusions. In future evaluations, in addition to expanding the number of participants, it is essential to plan evaluations with end-users and health professionals to observe the actual applicability of the health indicators in users' daily lives.

5.5 Lessons Learned

This section presents and discusses ten lessons learned from conducting a case study (also classified as a longitudinal investigation) to monitor the Quality of Life using IoHT and Machine Learning. The lessons were organized with a title, a short description, and alternatives to overcome it. Finally, it is presented a 5W1H table to summarize this discussion.

- **Study design needs to be carefully validated**

The planning phase is crucial for adequately conducting health monitoring studies "in the wild". It is also decisive in the approval by the ethics committee. On the other hand, the absence of a rigorous planning process can invalidate the data collected and increase research costs. Thus, a possible alternative to validate the planning is to conduct pilot studies. According to TEIJLINGEN *et al.* (2010), pilot studies refer to mini versions of a full-scale investigation, and they can identify potential practical problems in the research procedure.

- **Data privacy must be a priority**

Currently, laws and regulations protect digital health users from mishandling data (PURTOVA *et al.*, 2015). In this sense, privacy must be prioritized to create trust with the volunteers. Moreover, from the feedback collected in the qualitative assessment, it became evident that participants will be hesitant to use an invasive technology without a robust process for keeping their data secure. In this regard, a good alternative is to use data anonymization (SNEHA; ASHA, 2017). Another option is to avoid using data that makes it possible to re-identify the user, such as location or Internet access data.

- **Volunteer engagement requires attention from beginning to end of the study**

Recruiting participants takes work; keeping their engagement is even more problematic. Thus, research involving health data monitoring has the significant challenge of finding volunteers. Regarding this challenge, a helpful strategy is establishing partnerships with universities or health centers and making key people in these organizations aware of the work's relevance. These people should become ambassadors to attract volunteers. In addition, it is crucial to consider strategies to keep volunteers committed until the end of the work, for example, rewarding students who remain active.

- **The technology discomfort can be a bias**

Monitoring health data requires sensors (RODRIGUES *et al.*, 2018). Such sensors can be wearable like smart bands and smart rings, personal devices such as smartphones, or even instruments fixed in the environment such as cameras and infrared sensors. During planning, the researcher should decide which sensors will be used and how to collect the data (using native apps, for example). For this decision, it must be taken into account possible discomforts for the users. For example, even commercial devices already established on the market, such as smart bands, can provoke wrist-related allergies. Thus, it is recommended to investigate whether the participants are already used to the selected technology to avoid discomfort and, consequently, bias in the data.

- **The project budget needs to be considered when selecting devices for monitoring**

As stated before, the researcher must select sensors for data collection during planning. Among the criteria for this selection are the number, variety, and accuracy of measurements, battery consumption, ease of use, market availability, data access, and durability. However, while conducting the case study, it was realized that the project budget is a vital criterion in this selection. In general, most volunteers do not have these devices, and even for those that have, there is the problem of non-standardization since different brands and models can cause inaccurate data. In addition, as this type of study requires many participants, the strategy selected was to opt for a low-cost device that would allow us to obtain the necessary measurements. Thus, it would be possible to include a more significant number of participants. Therefore, the Xiaomi Mi Smart Band was chosen, which costs approximately 40 dollars in Brazil.

- **Extracting data from wearables is complex**

A significant challenge for those who work with wearable devices is data extraction (OLIVEIRA *et al.*, 2022a). If the researcher chooses to build their own device, this new technology can face many additional issues due to the lack of maturity. On the other hand, few options for commercial wearables have methods for extracting data. Furthermore, those commercial wearables that share Web APIs to retrieve data tend to have a higher cost, such as smartwatches with Android Wear or Fitbit devices. An alternative is to look for devices that allow data synchronization with platforms such as Google Fit (for Android) and HealthKit (for iOS) (OLIVEIRA *et al.*, 2022). Such platforms were designed to centralize users' health information and have well-documented APIs for data extraction.

- **Data collected "in the wild" always has noise**

Monitoring patients in a controlled environment allows the researcher or professional to establish the required minimum parameters for the system. It is possible, for example, to guarantee that the devices will always have access to the Internet or even that there will be no lack of battery supply. On the other hand, collecting health data in real life (uncontrolled environment) implies noise in the data. For example, data gaps will be generated when removing devices for charging. Also, synchronization problems can make it impossible to record specific measures. Again, another situation that can occur in studies that include self-reported surveys is the user forgetting to answer the survey. In addition to these examples, a wide variety of other situations can occur; unfortunately, it is impossible to avoid them all. Therefore, a suitable strategy to deal with these issues is to intensify the effort dedicated to data cleaning and processing. This step is crucial to remove noise.

- **Constant Internet access cannot be a premise**

As stated by RODRIGUES *et al.* (2018), IoHT architecture for healthcare monitoring systems involves collecting sensor data and sending it to robust nodes for processing and analysis.

It is common for these nodes to be at the edge or in the cloud. However, the periodic sending of data cannot presuppose continuous access to the Internet. In uncontrolled environments, it is common to have unavailable access for a while, resulting in failures in sending data. In this way, it is essential to implement strategies for resending in case of failure or even temporary storage for later sending. This strategy should prevent information loss.

- **Getting feedback should be uninterrupted**

After recruiting the participants, a session was conducted to explain the study operation, configure the devices, clarify doubts, and sign the informed consent form. On this occasion, participants were notified that they could withdraw at any time and could send feedback throughout the monitoring period. Unfortunately, few volunteers kept the practice of continuous feedback. In this case, only using the final evaluation questionnaire could extract qualitative data, and details may have been lost. Thus, it is crucial to encourage volunteers to provide periodic feedback. For future studies, it is planned to leave an anonymous form open from the beginning to the end of the research and ask them to keep feeding whenever they face a positive or adverse situation.

- **Unexpected problems should arise**

Finally, the researcher must be prepared for unexpected issues. For example, a device being stolen from the user or even a volunteer getting sick and having to withdraw. Unfortunately, there is no specific approach to dealing with these problems. However, it is essential to keep the research team watchful to reverse them as soon as possible.

Table 9 summarizes the lessons learned using the 5W1H model (however, who and where were omitted because the research team is always who conducts these activities, and location is not applied in this scenario).

5.6 Discussion

To evaluate the Healful platform, Chapter 5 presented a case study conducted with 44 participants over six months. In addition, two surveys were conducted to assess how intrusive the platform's user interface is and how adequate the selected health indicators are to support the inference result, respectively.

The results showed that it is possible to build a Random Forest model capable of inferring users' Quality of Life with an average error (RMSE) of 7.8618 for the physical domain and 7.4591 for the psychological domain. In addition, according to the case study participants,

Table 9 – Summarized lessons learned.

What	When	Why	How
Study design needs to be validated	Planning	It can increase project costs	Conducting pilot studies
Data privacy must be a priority	Planning and Recruiting	It can hamper recruitment, in addition to legal issues	Anonymizing data and making privacy policies clear
Volunteer engagement requires attention	From the beginning to the end	It can lead volunteers to withdraw	Encouraging volunteer participation
The technology discomfort can be a bias	Planning, and Conduction	It can insert bias in data	Selecting usual technologies
Project budget needs to be considered when selecting devices	Planning	It can increase project costs	Weighing the cost against the device resources required by the study
Extracting data from wearables is complex	Conduction	Without data, there is no health monitoring	Using health data hub platforms like Google Fit and Health Kit
Data collected “in the wild” has noise	Conduction	Noise can lead to inaccuracies	Investing in cleaning and processing activities
Constant Internet access cannot be a premise	Planning, and Conduction	It can cause data loss	Implementing data sending retries and data staging
Getting feedback should be uninterrupted	From the beginning to the end	To avoid missing relevant feedback	Allowing continuous sending of anonymous feedback
Unexpected problems should arise	Conduction	To ensure proper study conduction	Keeping the research team on its toes

Source: author.

68% fully agreed and 21% partially agreed that the Health platform, through its end-user interface called QoL Monitor, makes QoL monitoring transparent in users’ routines. Finally, 59 health and eHealth specialists pointed out that the health indicators selected in this study have a strong relationship with the physical and psychological QoL domains. However, such specialists also stated that the clustering-based method (presented in Chapter 4 Section 4.4) can not be the most suitable for some indicators such as social mobility and loneliness.

6 FINAL REMARKS

Eric Topol said in his book Deep Medicine (TOPOL, 2019) that:

> The greatest opportunity offered by AI is not reducing errors or workloads, or even curing cancer: it is the opportunity to restore the precious and time-honored connection and trust – the human touch-between patients and doctors. The relationship between doctors and patients has eroded over recent decades, with minimal time and keyboards as the main culprits. Physicians-and all clinicians-are experiencing burnout at increasing rates, and superficial contact with patients is resulting in diagnostic errors and unnecessary tests and procedures.

Topol's quote is closely related to this work because the presented results indicate that it is possible to build intelligent models from data collected by IoHT devices to augment the health professional's perception of a given patient. The vast amount of data generated by users' devices makes it possible to build a digital phenotyping (HUCKVALE *et al.*, 2019), which enables individual, personalized and ubiquitous health monitoring.

In light of this scenario, this Chapter concludes this thesis by presenting, in Section 6.1, an overview of the work context, motivation, goal, research questions, and methodology. In addition, the main deliverables, key findings and the open areas for further research are discussed in Sections 6.2 and 6.3, respectively.

6.1 Overview

Quality of Life has been investigated for a long time (NUSSBAUM; SEN, 1993), and there is still significant interest in this research area (NANDASENA *et al.*, 2022; SELLA *et al.*, 2023) due to the benefits that can be obtained from continuous QoL monitoring. In the literature, many studies reveal the relationship between patients' Quality of Life and health issues (VOJTA *et al.*, 2001; RASPOVIC; WUKICH, 2014; ALAM *et al.*, 2022; SELLA *et al.*, 2023).

However, the most employed instruments to monitor people's QoL are self-reported questionnaires (OLIVEIRA *et al.*, 2022b), and they have many disadvantages, such as the complexity of participant engagement (SANCHEZ *et al.*, 2015) and participant tendency to distort responses (PEQUENO *et al.*, 2020). In addition, although there are digital versions of these questionnaires (OLIVEIRA *et al.*, 2022b), such solutions are intrusive, given the need to frequently answer a set of questions. This kind of user interaction is the opposite of the ideas described by Weiser for ubiquitous systems (WEISER, 1999).

Therefore, this thesis presents a data-driven approach (see Figure 13) to deal with the QoL monitoring challenge. The main goal was to develop an Internet of Health Things platform capable of ubiquitously inferring users' Quality of Life, using users' Smart Devices and Machine Learning algorithms. This goal supports defining three Research Questions to guide the work:

I. What prior knowledge is available about the IoHT and its application to Quality of Life?

II. Which data can be ubiquitously obtained from commercial Internet of Health Things devices to represent users' health context?

III. What requirements are important to design and implement an IoHT platform for ubiquitous Quality of Life monitoring?

To achieve the primary purpose of this research, the Technical Action Research method was selected as the reference methodology (WIERINGA, 2014b). Based on this reference, six phases were executed to build and evaluate software artifacts aiming to check the following hypothesis: the Healful platform can infer users' QoL for the physical and psychological domains using IoHT devices and Machine Learning algorithms, making QoL monitoring less intrusive when compared with self-reported questionnaires.

6.2 Deliverables and Key Findings

The main deliverable of this work is a ubiquitous solution (detailed in Chapter 4, Section 4.1) able to infer users' Quality of Life through IoHT data, reducing the effort to get this health metric and improving users' engagement. This solution was built in the second deliverable – called the Healful platform (detailed in Chapter 4, Section 4.2) – and the platform was developed to i) deal with device heterogeneity and lack of interoperability; ii) reduce the cost to analyze QoL data; and iii) simplify the definition of digital healthcare interventions. Then, the last deliverable is a fully anonymized dataset[1] (discussed in Chapter 4, Section 4.3) to support the investigation of techniques to infer users' QoL. Based on these deliverables, the answers to the three above-mentioned Research Questions were drawn and are summarized as follows.

Regarding the prior QoL-based IoHT knowledge (**RQ1**), a review of 378 papers from four scientific digital libraries was conducted to identify challenges and opportunities. In these papers, it was found 182 challenges that were grouped into eight categories, and among this large set of challenges, it is possible to highlight an increased interest in personalized IoHT applications, data security, wearables to monitor patients, and Machine Learning to predict health issues. Also, the strengthening of mobile health is expected due to the cost reduction of devices

[1] This dataset is registered on the Kaggle platform, but it will remain private until completing the publication of a paper detailing how it was built and how it can be used.

and increasingly reliable solutions. Naturally, this strengthening will demand new software engineering methods, mainly focused on testing and usability.

Concerning IoHT devices and data to represent people's health context (**RQ2**), there is an extensive list of physical and virtual sensors present in BAN, PAN, and LAN networks. These sensors generate data that can characterize socio-demographic context, anthropometric information, medical history, physical activities, location, app usage, sleep pattern, posture, gait pattern, heart rate variability, and many others. Furthermore, each of these data can be used to understand different QoL facets. For example, daily steps and user location are strongly related to daily living activities. In addition, sleep patterns and the usage of smartphone apps can be related to stress and anxiety.

As for the requirements (**RQ3**), there are architectural requirements to deal with the lack of interoperability among the different devices present in IoHT environments; ubiquity requirement to promote calmness interactions (*i.e.*, less intrusive as possible) during the user's daily activities; Machine Learning requirement to get a large amount of data to comprehend user behavior, in addition, to check different algorithms by applying hyperparameter tuning and feature selection; and, security and privacy requirement to ensure user privacy and anonymity.

Also, it is important to highlight other key findings:

– Smart Quality of Life can be described as a person's Quality of Life inferred from individual and contextual data acquired using ubiquitous and less-intrusive technologies.
– It was statistically proved that an IoHT platform that uses Machine Learning algorithms can infer users' QoL for the physical and psychological domains considering an RMSE less than or equal to 10% of the reference value (WHOQOL-BREF questionnaire).
– The Random Forest regressor achieved an RMSE of 7.8618 for the physical domain and 7.4591 for the psychological domain, implying that this model can estimate users' QoL considering an average error of approximately 8 points more or less.
– According to case study participants, 68% fully agreed and 21% partially agreed that the Health platform, through its end-user interface called QoL Monitor, makes QoL monitoring transparent in users' routines.
– Fifty-nine (59) health and eHealth specialists pointed out that the health indicators selected in this study have a strong relationship with the physical and psychological QoL domains. However, the presented clustering-based method needs further validation to check its applicability in the real world.

Finally, Table 10 presents in chronological order the papers produced during this doctoral research, and Table 11 shows the innovation projects in which the author worked.

The first five papers and the twelfth paper in Table 10 were produced in partnership with other researchers of the GREat research lab and were focused on eHealth and self-adaptive systems. After these papers, the author conducted eight more publications as the first author involving the areas of Data Mining, Machine Learning, and the Internet of Health Things. Finally, it is important to mention that a set of these publications is associated with the innovation projects presented in Table 11.

Table 10 – Papers already published by the author and their relation with this thesis.

ID	**Title**	**Venue**	**Relation with this Thesis**
1	Towards a Taxonomy to Older Adults Healthcare Applications (ARAÚJO *et al.*, 2020)	Hawaii International Conference on System Sciences (HICSS)	Co-author in the study into a thesis area (eHealth)
2	Mobile applications for elderly healthcare: a systematic mapping (PAIVA *et al.*, 2020)	PLOS One journal	Co-author in the study into a thesis area (eHealth)
3	Dorsal: Ferramenta para Geração de Modelos de Dados para Aplicações voltadas a Saúde e Cuidado de Idosos (OLIVEIRA *et al.*, 2020)	20th Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)	Author in the study into a thesis area (eHealth)
4	Lessons Learned from the Development of Mobile Application for Fall Detection (LINHARES *et al.*, 2020)	International Conference on Global Health	Co-author in the study into a thesis area (eHealth)
5	Multifaceted infrastructure for Self-Adaptive IoT Systems (ANDRADE *et al.*, 2020)	Information and Software Technology journal	Co-author in paper focused on IoT challenges
6	Software Development During COVID-19 Pandemic: an Analysis of Stack Overflow and GitHub (OLIVEIRA *et al.*, 2021)	3rd ICSE Workshop on Software Engineering for Healthcare (SEH)	Author of the paper focused on Data Mining
7	Issue Auto-Assignment in Software Projects with Machine Learning Techniques (OLIVEIRA *et al.*, 2021)	8th International Workshop on Software Engineering Research and Industrial Practice (SERIP)	Author of the paper focused on Machine Learning
8	IoT-Health Platform to Monitor and Improve Quality of Life in Smart Environments (OLIVEIRA *et al.*, 2021)	8th IEEE International Workshop on Medical Computing (MediComp) - COMPSAC	Author in paper directly related with this thesis
9	Ten Years of eHealth Discussions on Stack Overflow (OLIVEIRA *et al.*, 2022)	15th International Conference on Health Informatics (HEALTHINF)	Author in paper directly related with this thesis
10	Internet of Health Things for Quality of Life: Open Challenges based on a Systematic Literature Mapping (OLIVEIRA *et al.*, 2022a)	15th International Conference on Health Informatics (HEALTHINF)	Author in paper directly related with this thesis
11	Towards an IoT platform to monitor QoL indicators (OLIVEIRA *et al.*, 2022b)	15th International Conference on Health Informatics (HEALTHINF)	Author in paper directly related with this thesis
12	Where is the Internet of Health Things Data? (JUNIOR *et al.*, 2022)	24th International Conference on Enterprise Information Systems (ICEIS)	Co-author in the study into a thesis area (eHealth)
13	Lessons Learned from Health Monitoring in the Wild (OLIVEIRA *et al.*, 2023)	16th International Conference on Health Informatics (HEALTHINF)	Author in paper directly related with this thesis
14	Mobile Health in a Developer's Perspective (accepted for publication)	SN Computer Science	Author in paper directly related with this thesis
15	Big Data Fortaleza: Plataforma Inteligente para Políticas Públicas Baseadas em Evidências (SANTOS *et al.*, 2023)	Workshop de Computação Aplicada em Governo Eletrônico (WCGE)	Co-author in the study into a thesis area (BigData)

Source: author.

Table 11 – Innovation projects in which the author has worked.

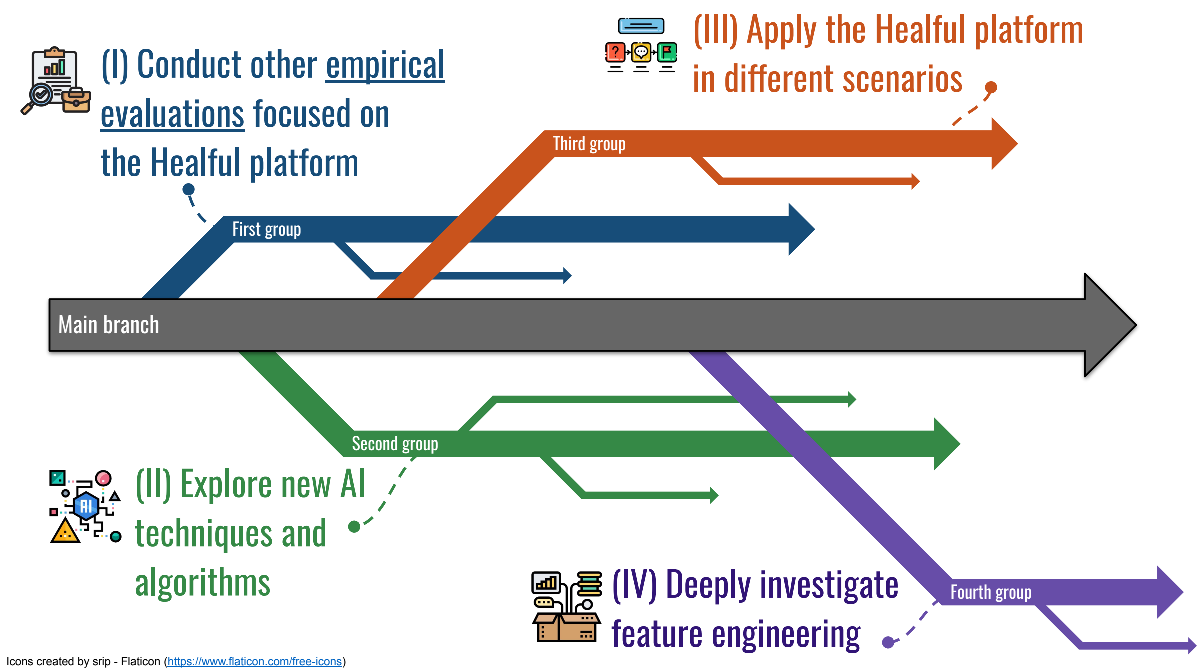
Short name	Partner	Begin	End	Role	Papers (ID)
Elderly Health Care	Fiocruz-CE	April, 2019	March, 2020	eHealth Researcher	1, 2, 3, 4
LG Avantte	LGE-SP	April, 2020	March, 2021	Machine Learning Researcher	7
Huawei ADaaS	Huawei Brazil	June, 2021	March, 2022	Cloud Computing Researcher	-
Dell Scope	Dell	April, 2022	October,2022	Machine Learning Researcher	-
BigDataFor	IPLANFOR	August, 2022	In progress	Data Scientist	15

Source: author.

6.3 Open Areas for Further Research

To maintain the science gears flowing, it is critical to further investigate the research opportunities that emerged during the development of this work. The natural path (main branch) to keep evolving this work is to assess health indicators in a real scenario, as they were only evaluated by domain experts. However, there are alternatives that can be organized into four branches, as presented by Figure 37: i) to conduct other empirical evaluations; ii) to explore new AI techniques; iii) to apply the Healful in different scenarios; and, iv) to investigate feature engineering.

Figure 37 – Alternatives for future work grouped into four branches



Source: author.

– **Conduct other empirical evaluations**: empirical studies performed in this work were essential to comprehend how Machine Learning algorithms can infer people's QoL. However, the number of participants was restricted, and the study duration was short. Therefore, conducting further assessments is a promising research opportunity, as new findings can be obtained from them. As a starting point for this branch, conducting longitudinal studies in partnership with health institutions or even with health insurance would be interesting. For example, in KAMAN *et al.* (2023), 2319 families had their Quality of Life monitored for two years, and in ZHANG *et al.* (2021), 1614 participants from Spain, Greece, Croatia, Netherlands, and United Kingdom were followed by 12 months.

– **Explore new AI techniques and algorithms**: in this work, the choice of AI techniques followed a well-known approach: starting with simple algorithms and increasing the complexity until finding a suitable algorithm. In addition, the Random Search was applied for the hyperparameter optimization and the SelectKBest strategy for the feature selection. However, many other algorithms and strategies can be explored. For example, considering the large number of features that can be generated in this scenario, a promising alternative to Random Forest would be Deep Learning (DONG *et al.*, 2021). As for hyperparameter tuning, genetic algorithms, evolutionary algorithms, or Bayesian optimization can be investigated as alternatives to random search.
  Thus, a good starting point for this branch would be the book “Deep Learning Techniques for Biomedical and Health Informatics”, which presents state-of-the-art approaches for deep-learning-based biomedical and health-related applications (DASH *et al.*, 2020).

– **Apply the Healful platform in different scenarios**: this work focuses on adults’ physical and psychological QoL domains. However, the platform was built to allow the construction of monitoring systems for different scenarios. Thus, there is a large room for opportunities regarding applying the Healful platform in different scenarios. For example, it is known that anxiety and depression have become critical health issues worldwide (CHANG *et al.*, 2021). Then, the Healful could be configured to work in this scenario by collecting and analyzing data from teenagers, for example. Thus, it would be possible to verify if intelligent models can support the diagnosis of such problems.
  A starting point for this branch would be a literature review on diseases that have questionnaires as the primary diagnostic tool and then searching for data collected by IoHT devices related to these diseases. In psychiatry, such an approach has been employed to study autism spectrum disorders, schizophrenia, perinatal psychiatry, mood and sleep disorders, suicide prevention, and others (PRAKASH *et al.*, 2021).

– **Deeply investigate feature engineering**: although Chapter 3 presents data related to the physical and psychological QoL domains, there is too much to be investigated. For example, the features used in this work were defined based on data availability. However, an alternative strategy would be to ask physicians and other healthcare professionals which features they believe would be significant for each domain. This would support selecting features to create a concise group, improving the interpretability of resulting models. In addition, other devices, such as cameras and infrared devices, could be included.
  For this branch, it is possible to adopt two approaches: reviewing all the features presented in this work and using domain experts to remove unnecessary features or starting the feature engineering through interviews with healthcare professionals without considering any previous feature set.

BIBLIOGRAPHY

ABAD, V. C.; GUILLEMINAULT, C. Sleep and psychiatry. **Dialogues in Clinical Neuroscience**, v. 7, n. 4, p. 291–303, dez. 2005. ISSN 1958-5969.

ABDELNAPI, N. M. M.; OMRAN, N. F.; ALI, A. A.; OMARA, F. A. A survey of internet of things technologies and projects for healthcare services. In: **2018 International Conference on Innovative Trends in Computer Engineering (ITCE)**. Aswan: IEEE, 2018. p. 48–55. ISBN 9781538608791.

ABRAN, A. (Ed.). **SWEBOK**: guide to the software engineering body of knowledge. Version 3.0. Los Alamitos, CA: IEEE Computer Society, 2014. ISBN 9780769551661.

ADAM, I. O.; ALHASSAN, M. D. The effect of mobile phone penetration on the quality of life. **Telecommunications Policy**, v. 45, n. 4, p. 102109, maio 2021. ISSN 03085961.

ADAY, L. A.; CORNELIUS, L. J. **Designing and conducting health surveys**: a comprehensive guide. 3rd ed. ed. San Francisco: Jossey-Bass, 2006. ISBN 9780787975609.

AGHDAM, Z. N.; RAHMANI, A. M.; HOSSEINZADEH, M. The Role of the Internet of Things in Healthcare: Future Trends and Challenges. **Computer Methods and Programs in Biomedicine**, v. 199, p. 105903, fev. 2021. ISSN 01692607.

AGU, E.; PEDERSEN, P.; STRONG, D.; TULU, B.; HE, Q.; WANG, L.; LI, Y. The smartphone as a medical device: Assessing enablers, benefits and challenges. In: IEEE. **IEEE International Workshop of Internet-of-Things Networking and Control**. Orleans, LA, USA, 2013. p. 48–52.

AGUILERA-ASTUDILLO, C.; CHAVEZ-CAMPOS, M.; GONZALEZ-SUAREZ, A.; GARCIA-CORDERO, J. L. A low-cost 3-D printed stethoscope connected to a smartphone. In: **2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. Orlando, FL, USA: IEEE, 2016. p. 4365–4368. ISBN 9781457702204.

AHMAD, M.; AMIN, M.; HUSSAIN, S.; KANG, B.; CHEONG, T.; LEE, S. Health fog: a novel framework for health and wellness applications. **Journal of Supercomputing**, v. 72, n. 10, p. 3677–3695, 2016.

AHMED, S.; ILYAS, M.; RAJA, M. Internet of things: Applications in smart healthcare. **ICSIT 2018 - 9th International Conference on Society and Information Technologies, Proceedings**, p. 19–24, 2015.

AL-FUQAHA, A.; GUIZANI, M.; MOHAMMADI, M.; ALEDHARI, M.; AYYASH, M. Internet of things: A survey on enabling technologies, protocols, and applications. **IEEE communications surveys & tutorials**, IEEE, v. 17, n. 4, p. 2347–2376, 2015.

AL-TURJMAN, F.; NAWAZ, M. H.; ULUSAR, U. D. Intelligence in the Internet of Medical Things era: A systematic review of current and future trends. **Computer Communications**, ELSEVIER, 150, p. 644–660, jan. 2020. ISSN 0140-3664.

ALAM, R.; PATEL, H. D.; SU, Z. T.; CHEAIB, J. G.; GED, Y.; SINGLA, N.; ALLAF, M. E.; PIERORAZIO, P. M. Self-reported quality of life as a predictor of mortality in renal cell carcinoma. **Cancer**, v. 128, n. 3, p. 479–486, fev. 2022. ISSN 0008-543X, 1097-0142.

ALBAHRI, O. S.; ALBAHRI, A. S.; ZAIDAN, A. A.; ZAIDAN, B. B.; ALSALEM, M. A.; MOHSIN, A. H.; MOHAMMED I, K.; ALAMOODI, A. H.; NIDHAL, S.; ENAIZAN, O.; CHYAD, M. A.; ABDULKAREEM, K. H.; ALMAHDI, E. M.; SHAFEEY, G. A. A.; BAQER, M. J.; JASIM, A. N.; JALOOD, N. S.; SHAREEF, A. H. Fault-Tolerant mHealth Framework in the Context of IoT-Based Real-Time Wearable Health Data Sensors. **IEEE Access**, IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 7, p. 50052–50080, 2019. ISSN 2169-3536.

ALKHATIB, S.; WAYCOTT, J.; BUCHANAN, G.; BOSUA, R. Privacy and the Internet of Things (IoT) Monitoring Solutions for Older Adults: A Review. In: Cummings, E and Ryan, A and Schaper, LK (Ed.). **Connecting the System to Enhance the Practitioner and Consumer Experience in Healthcare**. Netherlands: IOS PRESS, 2018. (Studies in Health Technology and Informatics, 252), p. 8–14. ISBN 978-1-61499-890-7; 978-1-61499-889-1. ISSN 0926-9630.

ALMEIDA, R. L.; MACEDO, A. A.; ARAÚJO, Í. L. de; AGUILAR, P. A.; ANDRADE, R. M. Watchalert: An evolution of the f-alert app for detecting falls on smartwatches. In: SBC. **Extended Proceedings of the XXII Brazilian Symposium on Multimedia and Web Systems**. Porto Alegre: SBC, 2016. p. 124–127.

ALTEVOGT, B. M.; COLTEN, H. R. *et al.* **Sleep disorders and sleep deprivation**: an unmet public health problem. Washington, DC: National Academies Press, 2006. ISBN 9780309101110.

ALTHOFF, T. Population-scale pervasive health. **IEEE Pervasive Computing**, v. 16, n. 4, p. 75–79, 2017.

ANDRADE, R. M.; JUNIOR, B. R. A.; OLIVEIRA, P. A. M.; MAIA, M. E.; VIANA, W.; NOGUEIRA, T. P. Multifaceted infrastructure for self-adaptive iot systems. **Information and Software Technology**, Elsevier, p. 106505, 2020.

ANGTHONG, C.; VELJKOVIC, A. Relationships among subjective patient-reported outcome, quality of life, and objective gait characteristics using wearable foot inertial-sensor assessment in foot–ankle patients. **European Journal of Orthopaedic Surgery & Traumatology**, Springer, v. 29, n. 3, p. 683–687, 2019.

ARAÚJO, I.; PEREIRA, M. B.; SILVA, W.; LINHARES, I.; MARX, V.; ANDRADE, A. M.; ANDRADE, R. M.; CASTRO, M. F. de. Machine learning and cloud-enabled fall detection system using data from wearable devices: Deployment and evaluation. In: SBC. **Anais do XXII Simpósio Brasileiro de Computação Aplicada à Saúde**. Porto Alegre, 2022. p. 413–424.

ARAÚJO, L. D.; JUNIOR, E. C.; DUARTE, P.; SANTOS, I. D. S.; OLIVEIRA, P. A. M. D.; MENDES, C. M. O.; ANDRADE, R. M. D. C. Towards a Taxonomy for the Development of Older Adults Healthcare Applications. **Proceedings of the 53rd Hawaii International Conference on System Sciences**, USA, n. 1, 2020.

ARORA, A.; CHAKRABORTY, P.; BHATIA, M. Analysis of data from wearable sensors for sleep quality estimation and prediction using deep learning. **Arabian Journal for Science and Engineering**, Springer, v. 45, n. 12, p. 10793–10812, 2020.

ARSENOVIĆ, S.; TRAJKOVIĆ, G.; PEKMEZOVIĆ, T.; GAZIBARA, T. Association of health literacy with physical and mental health in people with chronic diseases. **Revue d'Épidémiologie et de Santé Publique**, Elsevier, v. 71, n. 1, p. 101419, 2023.

ARULANANTHAN, C.; HANIFA, S. M. Smart Health - Potential and Pathways: A Survey. In: **International Conference on Materials, Alloys and Experimental Mechanics (ICMAEM-2017)**. England: IOP PUBLISHING LTD, 2017. (IOP Conference Series-Materials Science and Engineering, 225). ISSN 1757-8981.

ASADY, E.; GHANIMA, W.; JELSNESS-JORGENSEN, L.-P.; KLOK, F.; KAHN, S. R.; STROMME, H.; WIK, H. S. Health-related quality-of-life questionnaires for deep vein thrombosis and pulmonary embolism: A systematic review on questionnaire development and methodology. **Research and practice in thrombosis and haemostasis**, Wiley Online Library, v. 5, n. 5, p. e12556, 2021.

ASHTON, K. *et al.* That 'internet of things' thing. **RFID journal**, Hauppauge, New York, v. 22, n. 7, p. 97–114, 2009.

ASIF-UR-RAHMAN, M.; AFSANA, F.; MAHMUD, M.; KAISER, M. S.; AHMED, M.; KAIWARTYA, O.; JAMES-TAYLOR, A. Toward a heterogeneous mist, fog, and cloud-based framework for the internet of healthcare things. **IEEE Internet of Things Journal**, v. 6, n. 3, p. 4049–4062, 2019.

ATHAVALE, Y.; KRISHNAN, S. Biosignal monitoring using wearables: Observations and opportunities. **Biomedical Signal Processing and Control**, v. 38, p. 22 – 33, 2017. ISSN 17468094.

ATZORI, L.; IERA, A.; MORABITO, G. The internet of things: A survey. **Computer networks**, Elsevier, v. 54, n. 15, p. 2787–2805, 2010.

AZIMI, I.; RAHMANI, A. M.; LILJEBERG, P.; TENHUNEN, H. Internet of things for remote elderly monitoring: a study from user-centered perspective. **Journal of Ambient Intelligence and Humanized Computing**, SPRINGER HEIDELBERG, GERMANY, 8, n. 2, SI, p. 273–289, APR 2017. ISSN 1868-5137.

AZUAJE, F.; DOPAZO, J. **Data analysis and visualization in genomics and proteomics**. Chichester: Wiley Online Library, 2005. ISBN 9780470094419 9780470094396.

BADE, B. C.; BROOKS, M. C.; NIETERT, S. B.; ULMER, A.; THOMAS, D. D.; NIETERT, P. J.; SCOTT, J. B.; SILVESTRI, G. A. Assessing the correlation between physical activity and quality of life in advanced lung cancer. **Integrative cancer therapies**, SAGE Publications Sage CA: Los Angeles, CA, USA, v. 17, n. 1, p. 73–79, mar. 2018. ISSN 1534-7354, 1552-695X.

BAJENARU, O.-L.; CUSTURA, A.-M. Enhanced framework for an elderly-centred platform: Big data in monitoring the health status. In: **2019 22nd International Conference on Control Systems and Computer Science (CSCS)**. Bucharest, Romania: IEEE, 2019. p. 643–648. ISBN 9781728123318.

BAKER, S. B.; XIANG, W.; ATKINSON, I. Internet of things for smart healthcare: Technologies, challenges, and opportunities. **IEEE Access**, IEEE, v. 5, p. 26521–26544, 2017.

BALAKRISHNA, C.; RENDON-MORALES, E.; AVILES-ESPINOSA, R.; DORE, H.; LUO, Z. Challenges of Wearable Health Monitors : A Case study of Foetal ECG Monitor. In: **2019 Global IoT Summit (GIoTS)**. Aarhus, Denmark: IEEE, 2019. p. 1–6. ISBN 9781728121710.

BANDODKAR, A.; JEERAPAN, I.; WANG, J. Wearable chemical sensors: Present challenges and future prospects. **ACS Sensors**, v. 1, n. 5, p. 464–482, 2016.

BANKER, K.; BAKKUM, P.; HAWKINS, T.; VERCH, S.; GARRETT, D. **MongoDB in Action**. 2nd ed. ed. Saintmpford, LaVergne: Manning Publications Company Ingram Publisher Services [distributor], 2016. ISBN 9781638353560.

BARRETO, F. M.; DUARTE, P. A. D. S.; MAIA, M. E. F.; ANDRADE, R. M. D. C.; VIANA, W. CoAP-CTX: A Context-Aware CoAP Extension for Smart Objects Discovery in Internet of Things. In: **2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)**. Turin: IEEE, 2017. p. 575–584. ISBN 9781538603673.

BEAM, A. L.; KOHANE, I. S. Translating artificial intelligence into clinical care. **Jama**, American Medical Association, v. 316, n. 22, p. 2368–2369, 2016.

BELEN, R. de; FAVERO, D. D.; BEDNARZ, T. Combining mixed reality and internet of things: An interaction design research on developing assistive technologies for elderly people. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 11593 LNCS, p. 291–304, 2019.

BELESIOTI, M.; CHOCHLIOUROS, I.; VANYA, S.; ORAVEC, V.; THEOLOGOU, N.; KOUTLI, M.; TRYFERIDIS, A.; TZOVARAS, D. E-health services in the context of iot: The case of the vicinity project. **IFIP Advances in Information and Communication Technology**, v. 520, p. 62–69, 2018.

BERDIDA, D. J. E.; GRANDE, R. A. N. Academic stress, covid-19 anxiety, and quality of life among nursing students: The mediating role of resilience. **International Nursing Review**, Wiley Online Library, v. 70, n. 1, p. 34–42, 2023.

BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. **Journal of machine learning research**, v. 13, n. 2, 2012.

BERROCAL, A.; MANEA, V.; MASI, A. D.; WAC, K. mQoL Lab: Step-by-Step Creation of a Flexible Platform to Conduct Studies Using Interactive, Mobile, Wearable and Ubiquitous Devices. **Procedia Computer Science**, v. 175, p. 221–229, 2020. ISSN 18770509.

BIDMESHKI, M.-M.; JAFARI, R. Low power programmable architecture for periodic activity monitoring. In: **Proceedings of the ACM/IEEE 4th International Conference on Cyber-Physical Systems**. Philadelphia Pennsylvania: ACM, 2013. p. 81–88. ISBN 9781450319966.

BISCHL, B.; BINDER, M.; LANG, M.; PIELOK, T.; RICHTER, J.; COORS, S.; THOMAS, J.; ULLMANN, T.; BECKER, M.; BOULESTEIX, A.-L. *et al.* Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. **Wiley Interdisciplinary Reviews, Data Mining and Knowledge Discovery**, Wiley Online Library, v. 13, n. 2, p. e1484, 2023.

BLEI, D. M.; NG, A. Y.; JORDAN, M. I. Latent dirichlet allocation. **Journal of machine Learning research**, v. 3, n. Jan, p. 993–1022, 2003.

BONO-NUEZ, A.; BLASCO, R.; CASAS, R.; BRÍO, B. Martín-del. Ambient intelligence for quality of life assessment. **Journal of Ambient Intelligence and Smart Environments**, IOS Press, v. 6, n. 1, p. 57–70, 2014.

BOWLING, A. Just one question: If one question works, why ask several? **Journal of Epidemiology & Community Health**, v. 59, n. 5, p. 342–345, maio 2005. ISSN 0143-005X.

BREIVOLD, H. P. Internet-of-Things and Cloud Computing for Smart Industry: A Systematic Mapping Study. In: **2017 5th International Conference on Enterprise Systems (ES)**. Beijing: IEEE, 2017. p. 299–304. ISBN 9781538609361.

BROWN, S.; HINE, N.; SIXSMITH, A.; GARNER, P. Care in the community. **BT Technology Journal**, v. 22, n. 3, p. 56–64, 2004.

BRUDY, L.; MEYER, M.; OBERHOFFER, R.; EWERT, P.; MÜLLER, J. Move more–be happier? physical activity and health-related quality of life in children with congenital heart disease. **American Heart Journal**, Elsevier, v. 241, p. 68–73, 2021.

CALHOUN, B.; LACH, J.; STANKOVIC, J.; WENTZLOFF, D.; WHITEHOUSE, K.; BARTH, A.; BROWN, J.; LI, Q.; OH, S.; ROBERTS, N.; ZHANG, Y. Body sensor networks: A holistic approach from silicon to users. **Proceedings of the IEEE**, v. 100, n. 1, p. 91–106, 2012.

CARBONARO, A.; PICCININI, F.; REDA, R. Integrating heterogeneous data of healthcare devices to enable domain data management. **Journal of E-Learning and Knowledge Society**, v. 14, n. 1, p. 45–56, 2018.

CARLETTA, J. Assessing agreement on classification tasks: The kappa statistic. **Computational Linguistics**, MIT Press, Cambridge, MA, v. 22, n. 2, p. 249–254, 1996.

CARVALHO, R. M.; ANDRADE, R. M. de C.; OLIVEIRA, K. M. de. Aquarium-a suite of software measures for hci quality evaluation of ubiquitous mobile applications. **Journal of Systems and Software**, Elsevier, v. 136, p. 101–136, 2018.

CDC, N. C. for C. D. P. **HRQOL Concepts**. 2018. https://www.cdc.gov/hrqol/concept.htm. Accessed on November 10, 2022.

CELESTI, A.; FAZIO, M.; MáRQUEZ, F.; GLIKSON, A.; MAUWA, H.; BAGULA, A.; CELESTI, F.; VILLARI, M. How to develop iot cloud e-health systems based on fiware: A lesson learnt. **Journal of Sensor and Actuator Networks**, v. 8, n. 1, 2019.

CHANG, J.-J.; JI, Y.; LI, Y.-H.; PAN, H.-F.; SU, P.-Y. Prevalence of anxiety symptom and depressive symptom among college students during covid-19 pandemic: A meta-analysis. **Journal of Affective Disorders**, Elsevier, v. 292, p. 242–254, 2021.

CHEN, M.; LEUNG, V. C.; MAO, S. Directional controlled fusion in wireless sensor networks. **Mobile Networks and Applications**, Springer, v. 14, n. 2, p. 220–229, 2009.

CHONG, Y.-W.; ISMAIL, W.; KO, K.; LEE, C.-Y. Energy harvesting for wearable devices: A review. **IEEE Sensors Journal**, v. 19, n. 20, p. 9047–9062, 2019.

CHUI, K.; LIU, R.; LYTRAS, M.; ZHAO, M. Big data and iot solution for patient behaviour monitoring. **Behaviour and Information Technology**, v. 38, n. 9, p. 940–949, 2019.

CIABATTONI, L.; FERRACUTI, F.; LONGHI, S.; PEPA, L.; ROMEO, L.; VERDINI, F. Real-time mental stress detection based on smartwatch. In: **2017 IEEE International Conference on Consumer Electronics (ICCE)**. Las Vegas, NV, USA: IEEE, 2017. p. 110–111. ISBN 9781509055449.

CLELAND-HUANG, J.; AGRAWAL, A.; VIERHAUSER, M.; MURPHY, M.; PRIETO, M. Extending MAPE-K to support human-machine teaming. In: **Proceedings of the 17th Symposium on Software Engineering for Adaptive and Self-Managing Systems**. Pittsburgh Pennsylvania: ACM, 2022. p. 120–131. ISBN 9781450393058.

COBAN, S.; GOKALP, M. O.; GOKALP, E.; EREN, P. E.; KOCYIGIT, A. Predictive Maintenance in Healthcare Services with Big Data Technologies. In: **2018 IEEE 11th Conference on Service-Oriented Computing and Applications (SOCA)**. Paris: IEEE, 2018. p. 93–98. ISBN 9781538691335.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and psychological measurement**, Sage Publications Sage CA: Thousand Oaks, CA, v. 20, n. 1, p. 37–46, 1960.

COMMITTEE, N. A. S. E. M. **Leading Health Indicators 2030**: Advancing Health, Equity, and Well-Being. Washington, D.C.: National Academies Press, 2020. ISBN 9780309671873.

CONCHEIRO-MOSCOSO, P.; GROBA, B.; MARTÍNEZ-MARTÍNEZ, F. J.; MIRANDA-DURO, M. d. C.; NIETO-RIVEIRO, L.; POUSADA, T.; QUEIRÓS, C.; PEREIRA, J. Study for the design of a protocol to assess the impact of stress in the quality of life of workers. **Int. journal of environmental research and public health**, Multidisciplinary Digital Publishing Institute, v. 18, n. 4, p. 1413, 2021.

CRANE, M.; RISSEL, C.; GREAVES, S.; GEBEL, K. Correcting bias in self-rated quality of life: an application of anchoring vignettes and ordinal regression models to better understand qol differences across commuting modes. **Quality of life research**, Springer, v. 25, n. 2, p. 257–266, 2016.

DAHLBERG, L.; MCKEE, K. J.; FRANK, A.; NASEER, M. A systematic review of longitudinal risk factors for loneliness in older adults. **Aging & Mental Health**, Taylor & Francis, v. 26, n. 2, p. 225–249, 2022.

DASH, S.; ACHARYA, B. R.; MITTAL, M.; ABRAHAM, A.; KELEMEN, A. (Ed.). **Deep Learning Techniques for Biomedical and Health Informatics**. Cham: Springer International Publishing, 2020. v. 68. (Studies in Big Data, v. 68). ISBN 9783030339654 9783030339661.

DAUWED, M.; YAHAYA, J.; MANSOR, Z.; HAMDAN, A. Determinants of internet of things services utilization in health information exchange. **Journal of Engineering and Applied Sciences**, v. 13, n. 24, p. 10490–10501, 2018.

DAUWED, M.; YAHAYA, J.; MANSOR, Z.; HAMDAN, A. Human factors for iot services utilization for health information exchange. **Journal of Theoretical and Applied Information Technology**, v. 96, n. 8, p. 2095–2105, 2018.

DESHKAR, S.; THANSEEH, R.; MENON, V. G. A review on iot based m-health systems for diabetes. **International Journal of Computer Science and Telecommunications**, v. 8, n. 1, p. 13–18, 2017.

DESYANI, T.; SAIFUDIN, A.; YULIANTI, Y. Feature Selection Based on Naive Bayes for Caesarean Section Prediction. **IOP Conference Series, Materials Science and Engineering**, v. 879, n. 1, p. 012091, jul. 2020. ISSN 1757-8981, 1757-899X.

DINNO, A. Nonparametric pairwise multiple comparisons in independent groups using dunn's test. **The Stata Journal**, SAGE Publications Sage CA: Los Angeles, CA, v. 15, n. 1, p. 292–300, 2015.

DISI, M. A.; DJELOUAT, H.; KOTRONI, C.; POLITIS, E.; AMIRA, A.; BENSAALI, F.; DIMITRAKOPOULOS, G.; ALINIER, G. Ecg signal reconstruction on the iot-gateway and efficacy of compressive sensing under real-time constraints. **IEEE Access**, v. 6, p. 69130–69140, 2018.

DOBRE, C.; BAJENARU, L.; MARINESCU, I. A.; TOMESCU, M. Improving the Quality of Life for Older People: From Smart Sensors to Distributed Platforms. In: **2019 22nd International Conference on Control Systems and Computer Science (CSCS)**. Bucharest, Romania: IEEE, 2019. p. 636–642. ISBN 9781728123318.

DOHERTY, A. M.; GAUGHRAN, F. The interface of physical and mental health. **Social psychiatry and psychiatric epidemiology**, Springer, v. 49, p. 673–682, 2014.

DOHR, A.; MODRE-OPSRIAN, R.; DROBICS, M.; HAYN, D.; SCHREIER, G. The internet of things for ambient assisted living. In: IEEE. **2010 Seventh International Conference on Information Technology New Generations**. Las Vegas, NV, USA, 2010. p. 804–809.

DONATI, M.; CELLI, A.; RUIU, A.; SAPONARA, S.; FANUCCI, L. A Telemedicine Service System Exploiting BT/BLE Wireless Sensors for Remote Management of Chronic Patients. **TECHNOLOGIES**, SWITZERLAND, 7, n. 1, JAN 18 2019. ISSN 2227-7080.

DONG, S.; WANG, P.; ABBAS, K. A survey on deep learning and its applications. **Computer Science Review**, Elsevier, v. 40, p. 100379, 2021.

DOURIS, P. C.; HALL, C. A.; JUNG, M.-K. The relationship between academic success and sleep, stress and quality of life during the first year of physical therapy school. **Journal of American College Health**, Taylor & Francis, v. 71, n. 3, p. 830–835, 2023.

ELBASANI, E.; LEE, H.; CHOI, J. Wsn/rfid indoor positioning and tracking based on machine learning: A health care application. **Lecture Notes in Electrical Engineering**, v. 536 LNEE, p. 446–452, 2020.

ELKINTON, J. R. Medicine and the quality of life. **Annals of Internal Medicine**, v. 64, p. 711–714, 1966.

ELMISERY, A.; RHO, S.; ABORIZKA, M. A new computing environment for collective privacy protection from constrained healthcare devices to iot cloud services. **Cluster Computing**, v. 22, p. 1611–1638, 2019.

ENSHAEIFAR, S.; BARNAGHI, P.; SKILLMAN, S.; MARKIDES, A.; ELSALEH, T.; ACTON, S. T.; NILFOROOSHAN, R.; ROSTILL, H. The Internet of Things for Dementia Care. **IEEE Internet Computing**, IEEE Computer Society, USA, 22, n. 1, p. 8–17, JAN-FEB 2018. ISSN 1089-7801.

ESTRADA-GALINANES, V.; WAC, K. Visions and Challenges in Managing and Preserving Data to Measure Quality of Life. In: **2018 IEEE 3rd International Workshops on Foundations and Applications of Self* Systems (FAS*W)**. Trento: IEEE, 2018. p. 92–99. ISBN 9781538651759.

FACCHINETTI, G.; PETRUCCI, G.; ALBANESI, B.; MARINIS, M. G. D.; PIREDDA, M. Can smart home technologies help older adults manage their chronic condition? a systematic literature review. **International Journal of Environmental Research and Public Health**, Multidisciplinary Digital Publishing Institute, v. 20, n. 2, p. 1205, 2023.

FAGGELLA, D. Where healthcare's big data actually comes from. **Tech Emerg**, v. 11, 2018.

FAHEEM, M.; BUTT, R. A.; RAZA, B.; ALQUHAYZ, H.; ABBAS, M. Z.; NGADI, M. A.; GUNGOR, V. C. A Multiobjective, Lion Mating Optimization Inspired Routing Protocol for Wireless Body Area Sensor Network Based Healthcare Applications. **Sensors (Basel)**, v. 19, n. 23, Nov 2019.

FAN, Y.-J.; FENG, Y.-J.; MENG, Y.; SU, Z.-Z.; WANG, P.-X. The relationship between anthropometric indicators and health-related quality of life in a community-based adult population: A cross-sectional study in southern china. **Frontiers in Public Health**, Frontiers, v. 10, p. 955615, 2022.

FELCE, D.; PERRY, J. Quality of life: Its definition and measurement. **Research in developmental disabilities**, Elsevier, v. 16, n. 1, p. 51–74, 1995.

FELDT, L. S.; WOODRUFF, D. J.; SALIH, F. A. Statistical inference for coefficient alpha. **Applied psychological measurement**, Sage Publications Sage CA: Thousand Oaks, CA, v. 11, n. 1, p. 93–103, 1987.

FENG, C.; WANG, L.; CHEN, X.; ZHAI, Y.; ZHU, F.; CHEN, H.; WANG, Y.; SU, X.; HUANG, S.; TIAN, L. *et al.* A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected covid-19 pneumonia in fever clinics. **MedRxiv**, Cold Spring Harbor Laboratory Press, p. 2020–03, 2021.

FERREIRA, A. B. d. H. Novo dicionário aurélio da língua portuguesa. In: **Novo dicionário Aurélio da língua portuguesa**. Brasil: [*S. n.*], 2009. p. 2120–2120.

FILHO, I. B.; JUNIOR, G. de A. Iot-based healthcare applications: A review. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 10409 LNCS, p. 47–62, 2017.

FILHO, I. D. M. B.; AQUINO G.S., J. D. Proposing an iot-based healthcare platform to integrate patients, physicians and ambulance services. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 10409 LNCS, p. 187–202, 2017.

FILIPOVA, O. **Learning VueJS 2**. [*S. l.*]: Packt Publishing Ltd, 2016.

FRANK, P.-W. L.; MENG, M. Q.-H. A low cost Bluetooth powered wearable digital stethoscope for cardiac murmur. In: **2016 IEEE International Conference on Information and Automation (ICIA)**. Ningbo, China: IEEE, 2016. p. 1179–1182. ISBN 9781509041022.

FREITAS, L. B. de L.; SILVEIRA, P. G.; PIETA, M. A. M. Um estudo sobre o desenvolvimento da gratidão na infância. **Revista Interamericana de Psicología/Interamerican Journal of Psychology**, Sociedad Interamericana de Psicología, v. 43, n. 1, p. 49–56, 2009.

FRIEDENTHAL, S.; MOORE, A.; STEINER, R. Chapter 18 - integrating sysml into a systems development environment. In: FRIEDENTHAL, S.; MOORE, A.; STEINER, R. (Ed.). **A Practical Guide to SysML (Third Edition)**. Third edition. Boston: Morgan Kaufmann, 2015, (The MK/OMG Press). p. 507–541. ISBN 978-0-12-800202-5.

FRIEDMAN, J. H. Greedy function approximation: a gradient boosting machine. **Annals of statistics**, JSTOR, p. 1189–1232, 2001.

FRØNSDAL, K. B.; FACEY, K.; KLEMP, M.; NORDERHAUG, I. N.; MØRLAND, B.; RØTTINGEN, J.-A. Health technology assessment to optimize health technology utilization: using implementation initiatives and monitoring processes. **International journal of technology assessment in health care**, Cambridge University Press, v. 26, n. 3, p. 309–316, 2010.

GAMMA, E.; HELM, R.; JOHNSON, R.; VLISSIDES, J. Design Patterns: Abstraction and Reuse of Object-Oriented Design. In: **ECOOP' 93 — Object-Oriented Programming**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993. v. 707, p. 406–431. ISBN 9783540571209 9783540479109.

GANESH, D.; BALAJI, K. K.; SOKKANARAYANAN, S.; RAJAN, S.; SATHIYA-NARAYANAN, M. Healthcare Apps for Post-COVID Era: Trends, Challenges and Potential Opportunities. In: **2022 IEEE Delhi Section Conference**. New Delhi, India: IEEE, 2022. p. 1–7. ISBN 9781665458832.

GARDNER, M. J.; LUTES, J.; LUND, J.; HANSEN, J.; WALKER, D.; RINGGER, E.; SEPPI, K. The topic browser: An interactive tool for browsing topic models. In: WHISTLER CANADA. **Nips workshop on challenges of data visualization**. Canada, 2010. v. 2.

GAROUSI, V.; FELDERER, M.; MÄNTYLÄ, M. V. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. **Information and Software Technology**, Elsevier, v. 106, p. 101–121, 2019.

GATOUILLAT, A.; MASSOT, B.; BADR, Y.; SEJDIć, E.; GEHIN, C. Building IoT-Enabled Wearable Medical Devices: An Application to a Wearable, Multiparametric, Cardiorespiratory Sensor:. In: **Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies**. Funchal, Madeira, Portugal: SCITEPRESS - Science and Technology Publications, 2018. p. 109–118. ISBN 9789897583070.

GERMAIN, N.; ABALLÉA, S.; TOUMI, M. Measuring health-related quality of life in young children: how far have we come? **Journal of Market Access & Health Policy**, Taylor & Francis, v. 7, n. 1, p. 1618661, 2019.

GÉRON, A. **Mãos à Obra**: Aprendizado de Máquina com Scikit-Learn & TensorFlow. [*S. l.*]: Alta Books, 2019.

GEURTS, P.; ERNST, D.; WEHENKEL, L. Extremely randomized trees. **Machine learning**, Springer, v. 63, p. 3–42, 2006.

GHOSH, S.; LÖCHNER, J.; MITRA, B.; DE, P. Your smartphone knows you better than you may think: Emotional assessment 'on the go'via tapsense. **Quantifying Quality of Life - Incorporating Daily Life into Medicine**, Springer Nature, p. 209, 2022.

GILL, T. M.; FEINSTEIN, A. R. A critical appraisal of the quality of quality-of-life measurements. **Jama**, American Medical Association, v. 272, n. 8, p. 619–626, 1994.

GIORDANI, P.; FERRARO, M. B.; MARTELLA, F. **An Introduction to Clustering with R**. Singapore: Springer Singapore, 2020. v. 1. (Behaviormetrics: Quantitative Approaches to Human Behavior, v. 1). ISBN 9789811305528 9789811305535.

GKOUSKOS, D.; BURGOS, J. I'm in! Towards participatory healthcare of elderly through IOT. **Procedia Computer Science**, v. 113, p. 647–652, 2017. ISSN 18770509.

GMEINDER, M.; MORGAN, D.; MUELLER, M. How much do oecd countries spend on prevention? **OECD Health Working Papers**, OECD, n. 101, 2017. https://www.oecd-ilibrary.org/content/paper/f19e803c-en. Accessed on September 1, 2023.

GOLDTHORPE, J.; EPTON, T.; KEYWORTH, C.; CALAM, R.; ARMITAGE, C. Who is responsible for keeping children healthy? a qualitative exploration of the views of children aged 8–10 years old. **BMJ open**, British Medical Journal Publishing Group, v. 9, n. 5, p. e025245, 2019.

GOMEZ-CARMONA, O.; CASADO-MANSILLA, D.; GARCIA-ZUBIA, J. Health Promotion in Office Environments: A Worker-Centric Approach Driven by the Internet of Things. In: **INTELLIGENT ENVIRONMENTS 2018**. AMSTERDAM, NETHERLANDS: IOS PRESS, 2018. (Ambient Intelligence and Smart Environments, 23), p. 355–363. ISBN 978-1-61499-874-7; 978-1-61499-873-0. ISSN 1875-4163.

GOODE, R. W.; YE, L.; SEREIKA, S. M.; ZHENG, Y.; MATTOS, M.; ACHARYA, S. D.; EWING, L. J.; DANFORD, C.; HU, L.; IMES, C. C. *et al.* Socio-demographic, anthropometric, and psychosocial predictors of attrition across behavioral weight-loss trials. **Eating Behaviors**, Elsevier, v. 20, p. 27–33, 2016.

GOPAL, G.; SUTER-CRAZZOLARA, C.; TOLDO, L.; EBERHARDT, W. Digital transformation in healthcare - architectures of present and future information technologies. **Clin. Chem. Lab. Med.**, v. 57, n. 3, p. 328–335, 02 2019.

GREENE, S.; THAPLIYAL, H.; CARPENTER, D. IoT-Based Fall Detection for Smart Home Environments. In: **2016 IEEE International Symposium on Nanoelectronic and Information Systems (iNIS)**. Gwalior, India: IEEE, 2016. p. 23–28. ISBN 9781509061709.

GROUP, T. E. Euroqol-a new facility for the measurement of health-related quality of life. **Health policy**, Elsevier, v. 16, n. 3, p. 199–208, 1990.

GUBBI, J.; BUYYA, R.; MARUSIC, S.; PALANISWAMI, M. Internet of things (iot): A vision, architectural elements, and future directions. **Future generation computer systems**, Elsevier, v. 29, n. 7, p. 1645–1660, 2013.

GUYATT, G. H.; FEENY, D. H.; PATRICK, D. L. Measuring health-related quality of life. **Annals of internal medicine**, American College of Physicians, v. 118, n. 8, p. 622–629, 1993.

HAFEZI, H.; ROBERTSON, T. L.; MOON, G. D.; AU-YEUNG, K.-Y.; ZDEBLICK, M. J.; SAVAGE, G. M. An ingestible sensor for measuring medication adherence. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 62, n. 1, p. 99–109, 2014.

HAO, T.; WALTER, K. N.; BALL, M. J.; CHANG, H.-Y.; SUN, S.; ZHU, X. Stresshacker: towards practical stress monitoring in the wild with smartwatches. In: AMERICAN MEDICAL INFORMATICS ASSOCIATION. **AMIA Annual Symposium Proceedings**. United States, 2017. v. 2017, p. 830.

HAOYU, L.; JIANXING, L.; ARUNKUMAR, N.; HUSSEIN, A. F.; JABER, M. M. An IoMT cloud-based real-time sleep apnea detection scheme by using the SpO2 estimation supported by heart rate variability. **FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE**, ELSEVIER, NETHERLANDS, 98, p. 69–77, SEP 2019. ISSN 0167-739X.

HARARI, G. M.; MÜLLER, S. R.; AUNG, M. S.; RENTFROW, P. J. Smartphone sensing methods for studying behavior in everyday life. **Current opinion in behavioral sciences**, Elsevier, v. 18, p. 83–90, 2017.

HARISH, N.; MANDAL, S.; RAO, S.; PATIL, S. Particle swarm optimization based support vector machine for damage level prediction of non-reshaped berm breakwater. **Applied Soft Computing**, Elsevier, v. 27, p. 313–321, 2015.

HARKIN, S. T. Health care, not sick care. **American Journal of Health Promotion**, SAGE Publications Sage CA: Los Angeles, CA, v. 19, n. 1, p. 1–2, 2004.

HARTMANN, M.; HASHMI, U. S.; IMRAN, A. Edge computing in smart health care systems: Review, challenges, and research directions. **Transactions on Emerging Telecommunications Technologies**, Wiley Online Library, v. 33, n. 3, p. e3710, 2022.

HAWKINS, D. M. The problem of overfitting. **Journal of chemical information and computer sciences**, ACS Publications, v. 44, n. 1, p. 1–12, 2004.

HAYNES, R. B.; SACKETT, D. L.; RICHARDSON, W. S.; ROSENBERG, W.; LANGLEY, G. R. Evidence-based medicine: How to practice & teach ebm. **Canadian Medical Association. Journal**, Joule Inc, v. 157, n. 6, p. 788, 1997.

HAYS, R. D.; MORALES, L. S. The rand-36 measure of health-related quality of life. **Annals of medicine**, Taylor & Francis, v. 33, n. 5, p. 350–357, 2001.

HERBERT, C. Enhancing mental health, well-being and active lifestyles of university students by means of physical activity and exercise research programs. **Frontiers in public health**, Frontiers, v. 10, p. 849093, 2022.

HOEGER, W. W.; BOND, L.; RANSDELL, L.; SHIMON, J. M.; MERUGU, S. One-mile step count at walking and running speeds. **ACSM's Health & Fitness Journal**, LWW, v. 12, n. 1, p. 14–19, 2008.

HUCKVALE, K.; VENKATESH, S.; CHRISTENSEN, H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. **NPJ digital medicine**, Nature Publishing Group, v. 2, n. 1, p. 1–11, 2019.

HUNTER, J.; CRAWFORD, W. **Java servlet programming**: Help for server-side Java developers. [*S. l.*]: O'Reilly Media, Inc., 2001.

HUTTER, F.; KOTTHOFF, L.; VANSCHOREN, J. **Automated machine learning**: methods, systems, challenges. [*S. l.*]: Springer Nature, 2019.

HUYNH, Q. T.; NGUYEN, U. D.; TRAN, B. Q. A Cloud-Based System for In-Home Fall Detection and Activity Assessment. In: **7th International Conference on the Development of Biomedical Engineering in Vietnam (BME7)**. Singapore: Springer Singapore, 2020. v. 69, p. 103–108. ISBN 9789811358586 9789811358593.

HYLAND, M. E. A brief guide to the selection of quality of life instrument. **Health and quality of life outcomes**, Springer, v. 1, p. 1–5, 2003.

IAN, H. W.; EIBE, F. **Data mining**: practical machine learning tools and techniques. 1st ed. ed. Burlington, MA: Morgan Kaufmann Publishers, 2005. ISBN 9780123748560.

IBM. **An Architectural Blueprint for Autonomic Computing**. United States, 2005. https://api.semanticscholar.org/CorpusID:16909837. Accessed on April 13, 2016.

ISLAM, S. R.; KWAK, D.; KABIR, M. H.; HOSSAIN, M.; KWAK, K.-S. The internet of things for health care: a comprehensive survey. **IEEE access**, IEEE, v. 3, p. 678–708, 2015.

ISTEPANIAN, R. S.; HU, S.; PHILIP, N. Y.; SUNGOOR, A. The potential of internet of m-health things “m-iot” for non-invasive glucose level sensing. In: IEEE. **2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. United States, 2011. p. 5264–5266.

JAGADEESWARI, V.; SUBRAMANIYASWAMY, V.; LOGESH, R.; VIJAYAKUMAR, V. A study on medical Internet of Things and Big Data in personalized healthcare system. **HEALTH INFORMATION SCIENCE AND SYSTEMS**, SPRINGER, NEW YORK, USA, 6, SEP 20 2018. ISSN 2047-2501.

JAHAN, S.; RILEY, I.; WALTER, C.; GAMBLE, R. F.; PASCO, M.; MCKINLEY, P. K.; CHENG, B. H. MAPE-K/MAPE-SAC: An interaction framework for adaptive systems with security assurance cases. **Future Generation Computer Systems**, v. 109, p. 197–209, ago. 2020. ISSN 0167739X.

JAKUBCZYK, M.; GOLICKI, D. Estimating the Fuzzy Trade-Offs Between Health Dimensions with Standard Time Trade-Off Data. In: **Advances in Fuzzy Logic and Technology 2017**. Cham: Springer International Publishing, 2018. v. 642, p. 266–277. ISBN 9783319668239 9783319668246.

JARA, A. J.; BELCHI, F. J.; ALCOLEA, A. F.; SANTA, J.; ZAMORA-IZQUIERDO, M. A.; GÓMEZ-SKARMETA, A. F. A pharmaceutical intelligent information system to detect allergies and adverse drugs reactions based on internet of things. In: IEEE. **2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)**. Germany, 2010. p. 809–812.

JOVIĆ, A.; BRKIĆ, K.; BOGUNOVIĆ, N. A review of feature selection methods with applications. In: IEEE. **2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO)**. Croatia, 2015. p. 1200–1205.

JR, J. E. W. Sf-36 health survey. **Spine**, Lawrence Erlbaum Associates Publishers, v. 25, n. 24, p. 3130–3139, 1999. ISSN 0362-2436.

JR, J. E. W.; GANDEK, B. Overview of the sf-36 health survey and the international quality of life assessment (iqola) project. **Journal of clinical epidemiology**, Elsevier, v. 51, n. 11, p. 903–912, 1998.

JUNIOR, E. C.; ANDRADE, R.; VENCESLAU, A.; OLIVEIRA, P.; SANTOS, I.; OLIVEIRA, B. Where Is the Internet of Health Things Data?:. In: **Proceedings of the 24th International**

Conference on Enterprise Information Systems. Online Streaming: SCITEPRESS - Science and Technology Publications, 2022. p. 39–49. ISBN 9789897585692.

JUNIOR, E. C.; ANDRADE, R. M. D. C.; ROCHA, L. S. Development Process for Self-adaptive Applications of the Internet of Health Things based on Movement Patterns. In: **2021 IEEE 9th International Conference on Healthcare Informatics (ICHI)**. Victoria, BC, Canada: IEEE, 2021. p. 437–438. ISBN 9781665401326.

KAMAN, A.; ERHART, M.; DEVINE, J.; REISS, F.; NAPP, A.-K.; SIMON, A. M.; HURRELMANN, K.; SCHLACK, R.; HÖLLING, H.; WIELER, L. H. *et al.* Two years of pandemic: the mental health and quality of life of children and adolescents: Findings of the copsy longitudinal study. **Deutsches Ärzteblatt International**, Deutscher Arzte-Verlag GmbH, v. 120, n. 15, p. 269, 2023.

KANG, S.; KANG, S. Smart care to improve health care for the elderly. **Lecture Notes in Electrical Engineering**, v. 450, p. 54–58, 2017.

KARAGIANNOPOULOS, M.; ANYFANTIS, D.; KOTSIANTIS, S.; PINTELAS, P. Feature selection for regression problems. **Educational Software Development Laboratory, Department of Mathematics, University of Patras, Greece**, 2004.

KARAMITSIOS, K.; ORPHANOUDAKIS, T.; DAGIUKLAS, T. Evaluation of IoT-based Distributed Health Management Systems. In: **Proceedings of the 20th Pan-Hellenic Conference on Informatics**. Patras Greece: ACM, 2016. p. 1–6. ISBN 9781450347891.

KARIMI, M.; BRAZIER, J. Health, health-related quality of life, and quality of life: what is the difference? **Pharmacoeconomics**, Springer, v. 34, n. 7, p. 645–649, 2016.

KARYANI, A. K.; RASHIDIAN, A.; SEFIDDASHTI, S. E.; SARI, A. A. Self-reported health-related quality of life (hrqol) and factors affecting hrqol among individuals with health insurance in iran. **Epidemiology and health**, Korean Society of Epidemiology, v. 38, 2016.

KASSANIA, S. H.; KASSANIB, P. H.; WESOLOWSKIC, M. J.; SCHNEIDERA, K. A.; DETERSA, R. Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: a machine learning based approach. **Biocybernetics and Biomedical Engineering**, Elsevier, v. 41, n. 3, p. 867–879, 2021.

KEIM, D.; ANDRIENKO, G.; FEKETE, J.-D.; GöRG, C.; KOHLHAMMER, J.; MELANçON, G. Visual Analytics: Definition, Process, and Challenges. In: **Information Visualization**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. v. 4950, p. 154–175. ISBN 9783540709558 9783540709565.

KHAN, O.; BADHIWALA, J. H.; WITIW, C. D.; WILSON, J. R.; FEHLINGS, M. G. Machine learning algorithms for prediction of health-related quality-of-life after surgery for mild degenerative cervical myelopathy. **The Spine Journal**, Elsevier, v. 21, n. 10, p. 1659–1669, 2021.

KHAREL, J.; REDA, H. T.; SHIN, S. Y. Fog Computing-Based Smart Health Monitoring System Deploying LoRa Wireless Communication. **IETE TECHNICAL REVIEW**, TAYLOR & FRANCIS LTD, ENGLAND, 36, n. 1, p. 69–82, JAN 2 2019. ISSN 0256-4602.

KHODKARI, H.; MAGHREBI, S. G.; ASOSHEH, A.; HOSSEINZADEH, M. Smart Healthcare and Quality of Service Challenges. In: **2018 9th International Symposium on Telecommunications (IST)**. Tehran, Iran: IEEE, 2018. p. 253–257. ISBN 9781538682746.

KIM, C.; COSTELLO, F. J.; LEE, K. C.; LI, Y.; LI, C. Predicting factors affecting adolescent obesity using general bayesian network and what-if analysis. **International journal of environmental research and public health**, Multidisciplinary Digital Publishing Institute, v. 16, n. 23, p. 4684, 2019.

KIM, D. H.; NAM, K. H.; CHOI, B. K.; HAN, I. H.; JEON, T. J.; PARK, S. Y. The usefulness of a wearable device in daily physical activity monitoring for the hospitalized patients undergoing lumbar surgery. **Journal of Korean Neurosurgical Society**, The Korean Neurosurgical Society, v. 62, n. 5, p. 561, 2019.

KIM, G.-D.; YOON, C.; KYE, S.-B.; LEE, Y.; KANG, J.; YOO, Y.; SONG, T.-K. A single fpga-based portable ultrasound imaging system for point-of-care applications. **IEEE transactions on ultrasonics, ferroelectrics, and frequency control**, IEEE, v. 59, n. 7, p. 1386–1394, 2012.

KIM, H.-G.; CHEON, E.-J.; BAI, D.-S.; LEE, Y. H.; KOO, B.-H. Stress and heart rate variability: A meta-analysis and review of the literature. **Psychiatry investigation**, Korean Neuropsychiatric Association, v. 15, n. 3, p. 235, 2018.

KITCHENHAM, B. A.; BUDGEN, D.; BRERETON, P. **Evidence-based software engineering and systematic reviews**. Boca Raton: CRC press, 2015. v. 4. (Chapman & Hall/CRC innovations in software engineering and software development, v. 4). ISBN 9781482228656.

KNIGHT, M. J.; LYRTZIS, E.; BAUNE, B. T. The association of cognitive deficits with mental and physical quality of life in major depressive disorder. **Comprehensive psychiatry**, Elsevier, v. 97, p. 152147, 2020.

KOLDIJK, S.; KRAAIJ, W.; NEERINCX, M. A. Deriving Requirements for Pervasive Well-Being Technology From Work Stress and Intervention Theory: Framework and Case Study. **JMIR MHEALTH AND UHEALTH**, JMIR PUBLICATIONS, INC, TORONTO, CANADA, 4, n. 3, JUL-SEP 2016. ISSN 2291-5222.

KOPEC, J. A.; WILLISON, K. D. A comparative review of four preference-weighted measures of health-related quality of life. **Journal of clinical epidemiology**, Elsevier, v. 56, n. 4, p. 317–325, 2003.

KORTUEM, G.; KAWSAR, F.; SUNDRAMOORTHY, V.; FITTON, D. Smart objects as building blocks for the internet of things. **IEEE Internet Computing**, IEEE, v. 14, n. 1, p. 44–51, 2009.

KOTHA, M. Tech care: An efficient healthcare system using iot. **Advances in Intelligent Systems and Computing**, v. 1054, p. 655–667, 2020.

KOZACHENKO, L. F.; LEONENKO, N. N. Sample estimate of the entropy of a random vector. **Problemy Peredachi Informatsii**, Russian Academy of Sciences, Branch of Informatics, Computer Equipment, v. 23, n. 2, p. 9–16, 1987.

KRAUS, W. E.; JANZ, K. F.; POWELL, K. E.; CAMPBELL, W. W.; JAKICIC, J. M.; TROIANO, R. P.; SPROW, K.; TORRES, A.; PIERCY, K. L.; COMMITTEE, . P. A. G. A. *et al.* Daily step counts for measuring physical activity exposure and its relation to health. **Medicine and science in sports and exercise**, NIH Public Access, v. 51, n. 6, p. 1206, 2019.

KROENKE, K.; SPITZER, R. L.; WILLIAMS, J. B. The phq-9: validity of a brief depression severity measure. **Journal of general internal medicine**, Wiley Online Library, v. 16, n. 9, p. 606–613, 2001.

KUCHER, K.; KERREN, A. Text visualization techniques: Taxonomy, visual survey, and community insights. In: **2015 IEEE Pacific Visualization Symposium (PacificVis)**. Hangzhou, China: IEEE, 2015. p. 117–121. ISBN 9781467368797.

KUSHWAHA, S.; BAHL, S.; BAGHA, A. K.; PARMAR, K. S.; JAVAID, M.; HALEEM, A.; SINGH, R. P. Significant applications of machine learning for covid-19 pandemic. **Journal of Industrial Integration and Management**, World Scientific, v. 5, n. 04, p. 453–479, 2020.

LAGADEC, N.; STEINECKER, M.; KAPASSI, A.; MAGNIER, A. M.; CHASTANG, J.; ROBERT, S.; GAOUAOU, N.; IBANEZ, G. Factors influencing the quality of life of pregnant women: a systematic review. **BMC pregnancy and childbirth**, Springer, v. 18, n. 1, p. 1–14, 2018.

LATIF, S.; QADIR, J.; FAROOQ, S.; IMRAN, M. A. How 5g wireless (and concomitant technologies) will revolutionize healthcare? **Future Internet**, Multidisciplinary Digital Publishing Institute, v. 9, n. 4, p. 93, 2017.

LAURIE, G. T. Cross-sectoral big data: The application of an ethics framework for big data in health and research. **Asian Bioeth Rev**, England, v. 11, n. 3, p. 327–339, 2019.

LEE, J.; PARK, S.-H.; JU, J. H.; CHO, J. H. Application of a real-time pain monitoring system in korean fibromyalgia patients: A pilot study. **International journal of rheumatic diseases**, Wiley Online Library, v. 22, n. 5, p. 934–939, 2019.

LEE, U.; HAN, K.; CHO, H.; CHUNG, K.-M.; HONG, H.; LEE, S.-J.; NOH, Y.; PARK, S.; CARROLL, J. Intelligent positive computing with mobile, wearable, and iot devices: Literature review and research directions. **Ad Hoc Networks**, v. 83, p. 8–24, 2019.

LI, C.; HU, X.; ZHANG, L. The iot-based heart disease monitoring system for pervasive healthcare service. **Procedia computer science**, Elsevier, v. 112, p. 2328–2334, 2017.

LI, Z.; SHI, D.; WANG, F.; LIU, F. Loneliness Recognition Based on Mobile Phone Data. In: **Proceedings of the 2016 International Symposium on Advances in Electrical, Electronics and Computer Engineering**. Guangzhou, China: Atlantis Press, 2016. ISBN 9789462521810.

LINHARES, I.; ANDRADE, R.; JUNIOR, E. C.; OLIVEIRA, P. A.; OLIVEIRA, B.; AGUILAR, P. Lessons learned from the development of mobile applications for fall detection. **GLOBAL HEALTH**, p. 18–25, 2020.

LIU, Y.; HASSAN, K. A.; KARLSSON, M.; WEISTER, O.; GONG, S. Active plant wall for green indoor climate based on cloud and internet of things. **IEEE Access**, v. 6, p. 33631–33644, 2018.

LJUBOJEVIC, M.; ZORIC, M.; SIMIC, M.; BABIC, Z. Quality of Life Context Influence Factors Improvement Using Houseplants and Internet of Things. In: IEEE. **2016 IEEE INTERNATIONAL BLACK SEA CONFERENCE ON COMMUNICATIONS AND NETWORKING (BLACKSEACOM)**. NEW YORK: IEEE, 2016. (International Black Sea Conference on Communications and Networking). ISBN 978-1-5090-1925-0. ISSN 2375-8236.

LOUPPE, G. **Understanding Random Forests**: From Theory to Practice. 2014. https://arxiv.org/abs/1407.7502. Accessed on September 1, 2023.

LUO, F.; POSLAD, S.; BODANESE, E. Kitchen Activity Detection for Healthcare using a Low-Power Radar-Enabled Sensor Network. In: **ICC 2019 - 2019 IEEE International Conference on Communications (ICC)**. Shanghai, China: IEEE, 2019. p. 1–7. ISBN 9781538680889.

LUYSTER, F. S.; STROLLO, P. J.; ZEE, P. C.; WALSH, J. K. Sleep: a health imperative. **Sleep**, Oxford University Press, v. 35, n. 6, p. 727–734, 2012.

MALBURG, L.; HOFFMANN, M.; BERGMANN, R. Applying MAPE-K control loops for adaptive workflow management in smart factories. **Journal of Intelligent Information Systems**, v. 61, n. 1, p. 83–111, ago. 2023. ISSN 0925-9902, 1573-7675.

MALEK, M. The Art of Creating Models and Models Integration. In: **Model-Based Software and Data Integration**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. v. 8, p. 1–7. ISBN 9783540789987 9783540789994.

MALKI, Z.; ATLAM, E.-S.; HASSANIEN, A. E.; DAGNEW, G.; ELHOSSEINI, M. A.; GAD, I. Association between weather data and covid-19 pandemic predicting mortality rate: Machine learning approaches. **Chaos, Solitons & Fractals**, Elsevier, v. 138, p. 110137, 2020.

MANO, L. Y.; BARROS, V. A.; NUNES, L. H.; SAWADA, L. O.; ESTRELLA, J. C.; UEYAMA, J. Enlace: A combination of layer-based architecture and wireless communication for emotion monitoring in healthcare. **Mobile Information Systems**, Hindawi, v. 2019, 2019.

MARQUES, G.; PITARMA, R. Mhealth: Indoor environmental quality measuring system for enhanced health and well-being based on internet of things. **Journal of Sensor and Actuator Networks**, v. 8, n. 3, 2019.

MARTíNEZ-CARO, E.; CEGARRA-NAVARRO, J.; GARCíA-PéREZ, A.; FAIT, M. Healthcare service evolution towards the internet of things: An end-user perspective. **Technological Forecasting and Social Change**, v. 136, p. 268–276, 2018.

MARVASTI, F. F.; STAFFORD, R. S. From "sick care" to health care: reengineering prevention into the us system. **The New England journal of medicine**, NIH Public Access, v. 367, n. 10, p. 889, 2012.

MASI, A. D.; CIMAN, M.; GUSTARINI, M.; WAC, K. mQoL smart lab: quality of life living lab for interdisciplinary experiments. In: **Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct**. Heidelberg Germany: ACM, 2016. p. 635–640. ISBN 9781450344623.

MATE, K. K. V. Using New Technologies in Quality of Life Assessment. In: KASSIANOS, A. P. (Ed.). **Handbook of Quality of Life in Cancer**. Cham: Springer International Publishing, 2022. p. 123–131. ISBN 9783030847012 9783030847029.

MATTSSON, S.; PARTINI, J.; FAST-BERGLUND, Evaluating Four Devices that Present Operator Emotions in Real-time. **Procedia CIRP**, v. 50, p. 524–528, 2016. ISSN 22128271.

MCKIGHT, P. E.; NAJAB, J. Kruskal-wallis test. **The corsini encyclopedia of psychology**, Wiley Online Library, p. 1–1, 2010.

MCNAUGHTON, C. D.; MCCONNACHIE, A.; CLELAND, J. G.; SPERTUS, J. A.; ANGERMANN, C. E.; DUKLAS, P.; TROMP, J.; LAM, C. S.; FILIPPATOS, G.; DAHLSTROM, U. *et al.* Quality of life assessed 6 months after hospitalisation for acute heart failure: an analysis from report-hf (international registry to assess medical practice with longitudinal observation for treatment of heart failure). **European Journal of Heart Failure**, Wiley Online Library, 2022.

MCNEILL, F. M.; THRO, E. **Fuzzy logic**: a practical approach. Boston: AP Professional, 1994. ISBN 9780124859654.

MCRAE, M. P.; SIMMONS, G.; WONG, J.; MCDEVITT, J. T. Programmable Bio-nanochip Platform: A Point-of-Care Biosensor System with the Capacity To Learn. **Accounts of Chemical Research**, v. 49, n. 7, p. 1359–1368, jul. 2016. ISSN 0001-4842, 1520-4898.

MEA, V. D.; POPESCU, M. H.; GONANO, D.; PETAROS, T.; EMILI, I.; FATTORI, M. G. A Communication Infrastructure for the Health and Social Care Internet of Things: Proof-of-Concept Study. **JMIR MEDICAL INFORMATICS**, JMIR PUBLICATIONS, INC, 130 QUEENS QUAY E, STE 1102, TORONTO, ON M5A 0P6, CANADA, 8, n. 2, FEB 2020.

MEI, Q.; CAI, D.; ZHANG, D.; ZHAI, C. Topic modeling with network regularization. In: **Proceedings of the 17th international conference on World Wide Web**. Beijing China: ACM, 2008. p. 101–110. ISBN 9781605580852.

MEKKI, N.; HAMDI, M.; AGUILI, T.; KIM, T.-H. Scenario-based Vulnerability Analysis in IoT-based Patient Monitoring System:. In: **Proceedings of the 14th International Joint Conference on e-Business and Telecommunications**. Madrid, Spain: SCITEPRESS - Science and Technology Publications, 2017. p. 554–559. ISBN 9789897582592.

MELCHER, J.; HAYS, R.; TOROUS, J. Digital phenotyping for mental health of college students: a clinical review. **Evidence-Based Mental Health**, Royal College of Psychiatrists, v. 23, n. 4, p. 161–166, 2020.

MERILAHTI, J.; PäRKKä, J.; KORHONEN, I. Estimating Older People's Physical Functioning with Automated Health Monitoring Technologies at Home: Feature Correlations and Multivariate Analysis. In: **Grid and Pervasive Computing Workshops**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. v. 7096, p. 94–104. ISBN 9783642279157 9783642279164.

MERKOW, M. S.; BREITHAUPT, J. **Information security**: principles and practices. Second edition. Indianapolis, Indiana: Pearson, 2014. ISBN 9780789753250.

MILOVICH, M.; BURLESON, D. Quality of life: Older adults and the role of social media. **AMCIS 2017 - America's Conference on Information Systems - A Tradition of Innovation**, Boston, USA, v. 2017-August, 2017.

MITTELSTADT, B. Designing the health-related internet of things: Ethical principles and guidelines. **Information (Switzerland)**, v. 8, n. 3, 2017.

MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN, D. G. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. **Annals of Internal Medicine**, Am Coll Physicians, v. 151, n. 4, p. 264–269, 2009.

MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. **Foundations of machine learning**. Second edition. Cambridge, Massachusetts: The MIT Press, 2018. (Adaptive computation and machine learning). ISBN 9780262039406.

MOKHATRI-HESARI, P.; MONTAZERI, A. Health-related quality of life in breast cancer patients: review of reviews from 2008 to 2018. **Health and quality of life outcomes**, Springer, v. 18, n. 1, p. 1–25, 2020.

MUNSTER, V. J.; KOOPMANS, M.; DOREMALEN, N. van; RIEL, D. van; WIT, E. de. A novel coronavirus emerging in china—key questions for impact assessment. **New England Journal of Medicine**, Mass Medical Soc, v. 382, n. 8, p. 692–694, 2020.

MUSHTAQ, R.; SHOIB, S.; SHAH, T.; MUSHTAQ, S. Relationship between loneliness, psychiatric disorders and physical health? a review on the psychological aspects of loneliness. **Journal of clinical and diagnostic research (JCDR)**, JCDR Research & Publications Private Limited, v. 8, n. 9, p. WE01, 2014.

NABUCO, R.; RIBEIRO, A.; PEREIRA, L. Health-centered care based on co-designed cyber-physical system. **Smart Innovation, Systems and Technologies**, v. 135, p. 691–701, 2019.

NANDASENA, H.; PATHIRATHNA, M.; ATAPATTU, A.; PRASANGA, P. Quality of life of covid 19 patients after discharge: Systematic review. **PloS one**, Public Library of Science San Francisco, CA USA, v. 17, n. 2, p. e0263941, 2022.

NATIONS, U. World urbanization prospects 2018. **United Nations, Department of Economic and Social Affairs, Population Division**, 2018.

NATIONS, U. World population prospects: the 2019 revision. **United Nations, Department of Economic and Social Affairs, Population Division**, 2019.

NEWCOMBE, L.; YANG, P.; CARTER, C.; HANNEGHAN, M. Internet of Things Enabled Technologies for Behaviour Analytics in Elderly Person Care: A Survey. In: **2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)**. Exeter: IEEE, 2017. p. 863–870. ISBN 9781538630662.

NUSSBAUM, M.; SEN, A. **The Quality of life**. New York: Clarendon Press ; Oxford University Press, 1993. (WIDER studies in development economics). ISBN 9780198283959 9780198287971.

NUTT, D.; WILSON, S.; PATERSON, L. Sleep disorders as core symptoms of depression. **Dialogues in clinical neuroscience**, Taylor & Francis, 2022.

OLIVEIRA, A.; SILVA, E.; AGUIAR, J.; FARIA, B. M.; REIS, L. P.; CARDOSO, H.; GONÇALVES, J.; SÁ, J. Oliveira e; CARVALHO, V.; MARQUES, H. Biometrics and quality of life of lymphoma patients: A longitudinal mixed-model approach. **Expert Systems**, Wiley Online Library, v. 38, n. 4, p. e12640, 2021.

OLIVEIRA, P.; ANDRADE, R.; NETO, P. d. A. S. Lessons learned from health monitoring in the wild. In: INSTICC. **Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023)**. Portugal: SCITEPRESS - Science and Technology Publications, 2023. p. 155–166. ISBN 978-989-758-631-6.

OLIVEIRA, P.; ANDRADE, R. M. C.; BARRETO, I.; NOGUEIRA, T. P.; BUENO, L. M. Issue Auto-Assignment in Software Projects with Machine Learning Techniques. In: **2021 IEEE/ACM 8th International Workshop on Software Engineering Research and Industrial Practice (SER&IP)**. Madrid, Spain: IEEE, 2021. p. 65–72. ISBN 9781665444767.

OLIVEIRA, P.; ARAúJO, I.; JUNIOR, E.; DUARTE, P.; SANTOS, I. S.; ANDRADE, R. M. C.; BARRETO, I. C. H. C.; ANDRADE, L. O. M. Dorsal: Ferramenta para Geração de Modelos de Dados para Aplicações voltadas a Saúde e Cuidado de Idosos. In: **Anais Estendidos do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2020)**. Brasil: Sociedade Brasileira de Computação (SBC), 2020. p. 1–6.

OLIVEIRA, P.; JUNIOR, E. C.; ANDRADE, R.; SANTOS, I.; NETO, P. Ten Years of eHealth Discussions on Stack Overflow:. In: **Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies**. Online Streaming: SCITEPRESS - Science and Technology Publications, 2022. p. 45–56. ISBN 9789897585524.

OLIVEIRA, P.; NETO, P. S.; BRITTO, R.; RABELO, R.; BRAGA, R.; SOUZA, M. Ciaas-computational intelligence as a service with athena. **Computer Languages, Systems & Structures**, Elsevier, v. 54, p. 95–118, 2018.

OLIVEIRA, P. A.; ANDRADE, R.; NETO, P. S. Lessons Learned from mHealth Monitoring in the Wild:. In: **Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies**. Lisbon, Portugal: SCITEPRESS - Science and Technology Publications, 2023. p. 155–166. ISBN 9789897586316.

OLIVEIRA, P. A. M.; ANDRADE, R. M. C.; NETO, P. A. S. IoT-Health Platform to Monitor and Improve Quality of Life in Smart Environments. In: **2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)**. Madrid, Spain: IEEE, 2021. p. 1909–1912. ISBN 9781665424639.

OLIVEIRA, P. A. M.; ANDRADE, R. M. C.; NETO, P. S. N.; OLIVEIRA, B. S. Internet of health things for quality of life: Open challenges based on a systematic literature mapping. In: INSTICC. **Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies**. Online Streaming: SCITEPRESS - Science and Technology Publications, 2022.

OLIVEIRA, P. A. M.; ANDRADE, R. M. C.; NETO, P. S. N.; OLIVEIRA, B. S. Towards an ioht platform to monitor qol indicators. In: INSTICC. **Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies**. Online Streaming: SCITEPRESS - Science and Technology Publications, 2022. p. 438–445.

OLIVEIRA, P. A. M.; NETO, P. A. S.; SILVA, G.; IBIAPINA, I.; LIRA, W. L.; ANDRADE, R. M. C. Software Development During COVID-19 Pandemic: an Analysis of Stack Overflow and GitHub. In: **2021 IEEE/ACM 3rd International Workshop on Software Engineering for Healthcare (SEH)**. Madrid, Spain: IEEE, 2021. p. 5–12. ISBN 9781665444583.

OLSON, N.; NOLIN, J. M.; NELHANS, G. Semantic web, ubiquitous computing, or internet of things? a macro-analysis of scholarly publications. **Journal of Documentation**, Emerald Group Publishing Limited, 2015.

ONASANYA, A.; ELSHAKANKIRI, M. Smart integrated iot healthcare system for cancer care. **Wireless Networks**, 2019. ISSN 10220038.

ONASANYA, A.; LAKKIS, S.; ELSHAKANKIRI, M. Implementing iot/wsn based smart saskatchewan healthcare system. **Wireless Networks**, v. 25, n. 7, p. 3999–4020, 2019.

ONNELA, J.-P. Opportunities and challenges in the collection and analysis of digital phenotyping data. **Neuropsychopharmacology**, Springer International Publishing Cham, v. 46, n. 1, p. 45–54, 2021.

OPAS. **Indicadores de Saúde**: Elementos conceituais e práticos. Santiago, Chile: Organização Pan-Americana da Saúde, 2018. https://iris.paho.org/handle/10665.2/49057. Accessed on September 29, 2023.

OTT, R. L.; LONGNECKER, M. T. **An introduction to statistical methods and data analysis**. United States: Cengage Learning, 2015. ISBN 9780495017585.

PAI, M.; MCCULLOCH, M.; GORMAN, J. D.; PAI, N.; ENANORIA, W.; KENNEDY, G.; THARYAN, P.; JR, J. M. C. Systematic reviews and meta-analyses: an illustrated, step-by-step guide. **The National medical journal of India**, v. 17, n. 2, p. 86–95, 2004.

PAIVA, J. O.; ANDRADE, R. M.; OLIVEIRA, P. A. M. de; DUARTE, P.; SANTOS, I. S.; EVANGELISTA, A. L. d. P.; THEOPHILO, R. L.; ANDRADE, L. O. M. de; BARRETO, I. C. d. H. Mobile applications for elderly healthcare: A systematic mapping. **PloS one**, Public Library of Science San Francisco, CA USA, v. 15, n. 7, p. e0236091, 2020.

PANICKER, R. M.; CHANDRASEKARAN, B. “wearables on vogue”: a scoping review on wearables on physical activity and sedentary behavior during covid-19 pandemic. **Sport Sciences for Health**, Springer, p. 1–17, 2022.

PAUL, A.; MUKHERJEE, D. P.; DAS, P.; GANGOPADHYAY, A.; CHINTHA, A. R.; KUNDU, S. Improved random forest for classification. **IEEE Transactions on Image Processing**, IEEE, v. 27, n. 8, p. 4012–4024, 2018.

PAWAR, A. B.; GHUMBRE, S. A survey on IoT applications, security challenges and counter measures. In: **2016 International Conference on Computing, Analytics and Security Trends (CAST)**. Pune, India: IEEE, 2016. p. 294–299. ISBN 9781509013388.

PAZIENZA, A.; MALLARDI, G.; FASCIANO, C.; VITULANO, F. Artificial intelligence on edge computing: A healthcare scenario in ambient assisted living. **CEUR Workshop Proceedings**, v. 2559, p. 22–37, 2020.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEQUENO, N. P. F.; CABRAL, N. L. d. A.; MARCHIONI, D. M.; LIMA, S. C. V. C.; LYRA, C. d. O. Quality of life assessment instruments for adults: a systematic review of population-based studies. **Health and quality of life outcomes**, BioMed Central, v. 18, n. 1, p. 1–13, 2020.

PEREIRA, S.; PINTO, A.; ALVES, V.; SILVA, C. A. Brain tumor segmentation using convolutional neural networks in mri images. **IEEE transactions on medical imaging**, IEEE, v. 35, n. 5, p. 1240–1251, 2016.

PETERSEN, K.; FELDT, R.; MUJTABA, S.; MATTSSON, M. Systematic mapping studies in software engineering. In: **12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12**. Italy: BCS Learning & Development Ltd., 2008. p. 1–10.

PETTICREW, M.; ROBERTS, H. **Systematic reviews in the social sciences**: a practical guide. Malden, MA ; Oxford: Blackwell Pub, 2008. ISBN 9781405121101 9781405121118.

POPENTIU-VLăDICESCU, F.; ALBEANU, G. Big data quality for reliable industrial internet of things based systems. **International Journal of Control and Automation**, v. 12, n. 5, p. 546–552, 2019.

PRAKASH, J.; CHAUDHURY, S.; CHATTERJEE, K. Digital phenotyping in psychiatry: When mental health goes binary. **Industrial Psychiatry Journal**, Wolters Kluwer–Medknow Publications, v. 30, n. 2, p. 191, 2021.

PRESSMAN, R. S. **Software engineering**: a practitioner's approach. 7th ed. ed. New York: McGraw-Hill Higher Education, 2010. ISBN 9780073375977.

PRIYA, A.; GARG, S.; TIGGA, N. P. Predicting anxiety, depression and stress in modern life using machine learning algorithms. **Procedia Computer Science**, Elsevier, v. 167, p. 1258–1267, 2020.

PRIYADARSHINI, R.; BARIK, R.; DUBEY, H. Deepfog: Fog computing-based deep neural architecture for prediction of stress types, diabetes and hypertension attacks. **Computation**, v. 6, n. 4, 2018.

PUKELIENE, V.; STARKAUSKIENE, V. Quality of life: Factors determining its measurement complexity. **Engineering Economics**, v. 22, n. 2, p. 147–156, 2011.

PUNN, N. S.; SONBHADRA, S. K.; AGARWAL, S. Covid-19 epidemic analysis using machine learning and deep learning algorithms. **MedRxiv**, Cold Spring Harbor Laboratory Press, 2020.

PURTOVA, N.; KOSTA, E.; KOOPS, B.-J. Laws and regulations for digital health. In: **Requirements Engineering for Digital Health**. Cham: Springer, 2015. p. 47–74.

QADRI, Y. A.; NAUMAN, A.; ZIKRIA, Y. B.; VASILAKOS, A. V.; KIM, S. W. The future of healthcare internet of things: a survey of emerging technologies. **IEEE Communications Surveys & Tutorials**, IEEE, v. 22, n. 2, p. 1121–1167, 2020.

QIU, L.; TONG, Y.; YANG, Q.; SUN, N.; GONG, Y.; YIN, X. Reliability and validity of a smart quality of life scale for patients with tuberculosis. **Journal of Public Health**, Springer, v. 28, n. 5, p. 575–582, 2020.

QUINN, S.; MURPHY, N.; SMEATON, A. F. Tracking Human Behavioural Consistency by Analysing Periodicity of Household Water Consumption. In: **Proceedings of the 2019 2nd International Conference on Sensors, Signal and Image Processing**. Prague Czech Republic: ACM, 2019. p. 1–5. ISBN 9781450372435.

QURESHI, F.; KRISHNAN, S. Wearable Hardware Design for the Internet of Medical Things (IoMT). **Sensors (Basel)**, v. 18, n. 11, Nov 2018.

RABIN, R.; CHARRO, F. d. Eq-sd: a measure of health status from the euroqol group. **Annals of medicine**, Taylor & Francis, v. 33, n. 5, p. 337–343, 2001.

RADULESCU, C. Z.; ALEXANDRU, A.; BAJENARU, L. Health parameters correlation in an IoT monitoring, evaluation and analysis framework for elderly. In: **2019 23rd International Conference on System Theory, Control and Computing (ICSTCC)**. Sinaia, Romania: IEEE, 2019. p. 531–536. ISBN 9781728106991.

RAFFERTY, J.; MEDINA-QUERO, J.; QUINN, S.; SAUNDERS, C.; EKERETE, I.; NUGENT, C.; SYNNOTT, J.; GARCIA-CONSTANTINO, M. Thermal Vision Based Fall Detection via Logical and Data driven Processes. In: **2019 IEEE International Conference on Big Data, Cloud Computing, Data Science & Engineering (BCD)**. Honolulu, HI, USA: IEEE, 2019. p. 35–40. ISBN 9781728108865.

RAHMANI, A.-M.; THANIGAIVELAN, N. K.; Tuan Nguyen Gia; GRANADOS, J.; NEGASH, B.; LILJEBERG, P.; TENHUNEN, H. Smart e-Health Gateway: Bringing intelligence to Internet-of-Things based ubiquitous healthcare systems. In: **2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)**. Las Vegas, NV, USA: IEEE, 2015. p. 826–834. ISBN 9781479963904.

RAMU, G. A secure cloud framework to share ehrs using modified cp-abe and the attribute bloom filter. **Education and Information Technologies**, v. 23, n. 5, p. 2213–2233, 2018.

RASPOVIC, K. M.; WUKICH, D. K. Self-reported quality of life and diabetic foot infections. **The Journal of Foot and Ankle Surgery**, Elsevier, v. 53, n. 6, p. 716–719, 2014.

RATMANA, D. O.; SHIDIK, G. F.; FANANI, A. Z.; Muljono; PRAMUNENDAR, R. A. Evaluation of Feature Selections on Movie Reviews Sentiment. In: **2020 International Seminar on Application for Technology of Information and Communication (iSemantic)**. Semarang, Indonesia: IEEE, 2020. p. 567–571. ISBN 9781728190662 9781728190686.

RAVENS-SIEBERER, U.; GOSCH, A.; RAJMIL, L.; ERHART, M.; BRUIL, J.; DUER, W.; AUQUIER, P.; POWER, M.; ABEL, T.; CZEMY, L. *et al.* Kidscreen-52 quality-of-life measure for children and adolescents. **Expert review of pharmacoeconomics & outcomes research**, Taylor & Francis, v. 5, n. 3, p. 353–364, 2005.

RAYAN, R. A.; TSAGKARIS, C.; IRYNA, R. B. The Internet of Things for Healthcare: Applications, Selected Cases and Challenges. In: MARQUES, G.; BHOI, A. K.; ALBUQUERQUE, V. H. C. D.; K.S., H. (Ed.). **IoT in Healthcare and Ambient Assisted Living**. Singapore: Springer Singapore, 2021. v. 933, p. 1–15. ISBN 9789811598968 9789811598975.

REDA, R.; PICCININI, F.; CARBONARO, A. Semantic modelling of smart healthcare data. **Advances in Intelligent Systems and Computing**, v. 869, p. 399–411, 2018.

REINHART, C.; REINHART, V. The pandemic depression: The global economy will never be the same. **Foreign Aff.**, HeinOnline, v. 99, p. 84, 2020.

REN, Y.; TAN, S.; ZHANG, L.; WANG, Z.; WANG, Z.; YANG, J. Liquid level sensing using commodity wifi in a smart home environment. **Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies**, v. 4, n. 1, 2020. ISSN 24749567.

RODRIGUES, J. J.; SEGUNDO, D. B. D. R.; JUNQUEIRA, H. A.; SABINO, M. H.; PRINCE, R. M.; AL-MUHTADI, J.; ALBUQUERQUE, V. H. C. D. Enabling technologies for the internet of health things. **IEEE Access**, IEEE, v. 6, p. 13129–13141, 2018.

RODRIGUES, J. M.; OLIVEIRA, F.; RIBEIRO, C. P.; SANTOS, R. C. Mobile mental health for depression assistance: Research directions, obstacles, advantages, and disadvantages of implementing mhealth. **Digital Therapies in Psychosocial Rehabilitation and Mental Health**, IGI Global, p. 21–40, 2022.

ROSS, B. C. Mutual information between discrete and continuous data sets. **PloS one**, Public Library of Science San Francisco, USA, v. 9, n. 2, p. e87357, 2014.

RYKOV, Y.; THACH, T.-Q.; BOJIC, I.; CHRISTOPOULOS, G.; CAR, J. *et al.* Digital biomarkers for depression screening with wearable devices: cross-sectional study with machine learning modeling. **JMIR mHealth and uHealth**, JMIR Publications Inc., Toronto, Canada, v. 9, n. 10, p. e24872, 2021.

SAEB, S.; LATTIE, E. G.; KORDING, K. P.; MOHR, D. C. *et al.* Mobile phone detection of semantic location and its relationship to depression and anxiety. **JMIR mHealth and uHealth**, JMIR Publications Inc., Toronto, Canada, v. 5, n. 8, p. e7297, 2017.

SAINTILA, J.; LÓPEZ, T. E. L.; MAMANI, P. G. R.; WHITE, M.; HUANCAHUIRE-VEGA, S. Health-related quality of life, blood pressure, and biochemical and anthropometric profile in vegetarians and nonvegetarians. **Journal of nutrition and metabolism**, Hindawi, v. 2020, 2020.

SALAMA, U.; YAO, L.; PAIK, H.-Y. An internet of things based multi-level privacy-preserving access control for smart living. **Informatics**, v. 5, n. 2, 2018.

SANCHEZ, W.; MARTINEZ, A.; CAMPOS, W.; ESTRADA, H.; PELECHANO, V. Inferring loneliness levels in older adults from smartphones. **Journal of Ambient Intelligence and Smart Environments**, IOS Press, v. 7, n. 1, p. 85–98, 2015.

SANTOS, I.; OLIVEIRA, P.; OLIVEIRA, V.; NOGUEIRA, T.; DANTAS, A.; MENESCAL, L.; BATISTA Élcio; ANDRADE, R. Big data fortaleza: Plataforma inteligente para políticas públicas baseadas em evidências. In: **Anais do XI Workshop de Computação Aplicada em Governo Eletrônico**. Porto Alegre, RS, Brasil: SBC, 2023. p. 200–211. ISSN 2763-8723.

SCHÄUBLIN, J.; DERIGHETTI, M.; FEIGENWINTER, P.; PETERSEN-FELIX, S.; ZBINDEN, A. M. Fuzzy logic control of mechanical ventilation during anaesthesia. **British journal of anaesthesia**, Elsevier, v. 77, n. 5, p. 636–641, 1996.

SCHMIDT, M. The sankey diagram in energy and material flow management: part ii: methodology and current applications. **Journal of industrial ecology**, Wiley Online Library, v. 12, n. 2, p. 173–185, 2008.

SCHOBER, P.; BOER, C.; SCHWARTE, L. A. Correlation coefficients: appropriate use and interpretation. **Anesthesia & analgesia**, Wolters Kluwer, v. 126, n. 5, p. 1763–1768, 2018.

SCHOLZ, F. W.; STEPHENS, M. A. K-sample anderson–darling tests. **Journal of the American Statistical Association**, Taylor & Francis, v. 82, n. 399, p. 918–924, 1987.

SCHRöER, C.; KRUSE, F.; GóMEZ, J. M. A Systematic Literature Review on Applying CRISP-DM Process Model. **Procedia Computer Science**, v. 181, p. 526–534, 2021. ISSN 18770509.

SCHUCH, F. B.; VANCAMPFORT, D.; ROSENBAUM, S.; RICHARDS, J.; WARD, P. B.; STUBBS, B. Exercise improves physical and psychological quality of life in people with depression: A meta-analysis including the evaluation of control group response. **Psychiatry research**, Elsevier, v. 241, p. 47–54, 2016.

SEIFERT, E. Originpro 9.1: scientific data analysis and graphing software-software review. **Journal of chemical information and modeling**, v. 54, n. 5, p. 1552, 2014.

SELLA, E.; MIOLA, L.; TOFFALINI, E.; BORELLA, E. The relationship between sleep quality and quality of life in aging: A systematic review and meta-analysis. **Health psychology review**, Taylor & Francis, v. 17, n. 1, p. 169–191, 2023.

SHAFI, U.; MUMTAZ, R.; ANWAR, H.; QAMAR, A. M.; KHURSHID, H. Surface Water Pollution Detection using Internet of Things. In: **2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT)**. Islamabad: IEEE, 2018. p. 92–96. ISBN 9781538683538 9781538683545.

SHETH, A.; JAIMINI, U.; THIRUNARAYAN, K.; BANERJEE, T. Augmented Personalized Health: How Smart Data with IoTs and AI is about to Change Healthcare. **RTSI**, v. 2017, Sep 2017.

SIEVERT, C.; SHIRLEY, K. LDAvis: A method for visualizing and interpreting topics. In: **Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces**. Baltimore, Maryland, USA: Association for Computational Linguistics, 2014. p. 63–70.

SIGNORELLI, G. R.; LEHOCKI, F.; FERNáNDEZ, M. M.; O'NEILL, G.; O'CONNOR, D.; BRENNAN, L.; MONTEIRO-GUERRA, F.; RIVERO-RODRIGUEZ, A.; HORS-FRAILE, S.; MUNOZ-PENAS, J.; DALMAU, M. B.; MOTA, J.; OLIVEIRA, R. B.; MRINAKOVA, B.; PUTEKOVA, S.; MURO, N.; ZAMBRANA, F.; GARCIA-GOMEZ, J. M. A Research Roadmap: Connected Health as an Enabler of Cancer Patient Support. **Journal of Medical Internet Research**, v. 21, n. 10, p. e14360, out. 2019. ISSN 1438-8871.

SILQUEIRA, S. M. d. F. **O questionário genérico SF-36 como instrumento de mensuração da qualidade de vida relacionada a saúde de pacientes hipertensos**. Tese (Doutorado) – Universidade de São Paulo, 2005.

SILVA, C. A. D.; JUNIOR, G. S. D. A. Fog Computing in Healthcare: A Review. In: **2018 IEEE Symposium on Computers and Communications (ISCC)**. Natal: IEEE, 2018. p. 01126–01131. ISBN 9781538669501.

SKEVINGTON, S. M.; LOTFY, M.; O'CONNELL, K. A. The world health organization's whoqol-bref quality of life assessment psychometric properties and results of the international field trial. **Quality of life Research**, Springer, v. 13, n. 2, p. 299–310, 2004.

SMITH, K. J.; ROBERTS, M. S. Cost-effectiveness of newer treatment strategies for influenza. **The American journal of medicine**, Elsevier, v. 113, n. 4, p. 300–307, 2002.

SNEHA, S.; ASHA, P. Privacy preserving on e-health records based on anonymization technique. **Global Journal of Pure and Applied Mathematics**, v. 13, n. 7, p. 3367–3380, 2017.

SPITZER, W. O. State of science 1986: quality of life and functional status as target variables for research. **Journal of chronic diseases**, Elsevier, v. 40, n. 6, p. 465–471, 1987.

SRINIVASAN, K.; CURRIM, F.; RAM, S.; LINDBERG, C.; STERNBERG, E.; SKEATH, P.; NAJAFI, B.; RAZJOUYAN, J.; LEE, H.-K.; FOE-PARKER, C.; GOEBEL, N.; HERZL, R.; MEHL, M. R.; GILLIGAN, B.; HEERWAGEN, J.; KAMPSCHROER, K.; CANADA, K. Feature Importance and Predictive Modeling for Multi-source Healthcare Data with Missing Values. In: **Proceedings of the 6th International Conference on Digital Health Conference**. Montréal Québec Canada: ACM, 2016. p. 47–54. ISBN 9781450342247.

SRINIVASAN, R.; CHANDER, A. Biases in ai systems. **Communications of the ACM**, ACM New York, NY, USA, v. 64, n. 8, p. 44–49, 2021.

STOJANOV, J.; MALOBABIC, M.; STANOJEVIC, G.; STEVIC, M.; MILOSEVIC, V.; STOJANOV, A. Quality of sleep and health-related quality of life among health care professionals treating patients with coronavirus disease-19. **International Journal of Social Psychiatry**, SAGE Publications Sage UK: London, England, v. 67, n. 2, p. 175–181, 2021.

STRINE, T. W.; CHAPMAN, D. P. Associations of frequent sleep insufficiency with health-related quality of life and health behaviors. **Sleep medicine**, Elsevier, v. 6, n. 1, p. 23–27, 2005.

TAKIGUCHI, Y.; MATSUI, M.; KIKUTANI, M.; EBINA, K. The relationship between leisure activities and mental health: The impact of resilience and covid-19. **Applied Psychology, Health and Well-Being**, Wiley Online Library, v. 15, n. 1, p. 133–151, 2023.

TANAKA, K.; MONDEN, A.; ZEYNEP, Y. Effectiveness of auto-sklearn in software bug prediction. **Computer Software**, Japan Society for Software Science and Technology, v. 36, n. 4, p. 46–52, 2019.

TANG, C.; GARREAU, D.; LUXBURG, U. von. When do random forests fail? **Advances in neural information processing systems**, v. 31, 2018.

TAVAKOL, M.; DENNICK, R. Making sense of cronbach's alpha. **International journal of medical education**, IJME, v. 2, p. 53, 2011.

TEIJLINGEN, E. V.; HUNDLEY, V. *et al.* The importance of pilot studies. **Social research update**, v. 35, n. 4, p. 49–59, 2010.

TEWELL, J.; O'SULLIVAN, D.; MAIDEN, N.; LOCKERBIE, J.; STUMPF, S. Monitoring meaningful activities using small low-cost devices in a smart home. **Personal and Ubiquitous Computing**, v. 23, n. 2, p. 339–357, 2019.

THANGAM, D.; MALALI, A. B.; SUBRAMANIAN, G.; PARK, J. Y. Transforming healthcare through internet of things. In: **Cloud and Fog Computing Platforms for Internet of Things**. Amsterdam, Netherlands: Chapman and Hall/CRC, 2022. p. 15–24.

TOPOL, E. J. **Deep medicine**: how artificial intelligence can make healthcare human again. First edition. New York: Basic Books, 2019. ISBN 9781541644632.

TORRANCE, G. W. Utility approach to measuring health-related quality of life. **Journal of chronic diseases**, Elsevier, v. 40, n. 6, p. 593–600, 1987.

TOZETTO, W. R.; LEONEL, L. d. S.; BENEDET, J.; DUCA, G. F. D. Quality of life and its relationship with different anthropometric indicators in adults with obesity. **Fisioterapia em Movimento**, SciELO Brasil, v. 34, 2021.

TREERATPITUK, P.; CALLAN, J. Automatically labeling hierarchical clusters. In: **Proceedings of the 2006 national conference on Digital government research**. San Diego, California: ACM Press, 2006. p. 167.

TUDOR-LOCKE, C.; CRAIG, C. L.; BROWN, W. J.; CLEMES, S. A.; COCKER, K. D.; GILES-CORTI, B.; HATANO, Y.; INOUE, S.; MATSUDO, S. M.; MUTRIE, N. *et al.* How many steps/day are enough? for adults. **International Journal of Behavioral Nutrition and Physical Activity**, BioMed Central, v. 8, n. 1, p. 1–17, 2011.

TUN, S. Y. Y.; MADANIAN, S.; MIRZA, F. Internet of things (IoT) applications for elderly care: a reflective review. **Aging Clinical and Experimental Research**, v. 33, n. 4, p. 855–867, abr. 2021. ISSN 1720-8319.

USHER, K.; DURKIN, J.; BHULLAR, N. The covid-19 pandemic and mental health impacts. **International Journal of Mental Health Nursing**, Wiley-Blackwell, v. 29, n. 3, p. 315, 2020.

VALLANCE, J.; EURICH, D.; GARDINER, P.; TAYLOR, L.; JOHNSON, S. Associations of daily pedometer steps and self-reported physical activity with health-related quality of life: results from the alberta older adult health survey. **Journal of aging and health**, Sage Publications Sage CA: Los Angeles, CA, v. 28, n. 4, p. 661–674, 2016.

VALLEE, T.; SEDKI, K.; DESPRES, S.; JAULANT, M.-C.; TABIA, K.; UGON, A. On Personalization in IoT. In: **2016 International Conference on Computational Science and Computational Intelligence (CSCI)**. Las Vegas, NV, USA: IEEE, 2016. p. 186–191. ISBN 9781509055104.

VARGIU, E.; FERNáNDEZ, J. M.; MIRALLES, F. Context-Aware Based Quality of Life Telemonitoring. In: **Distributed Systems and Applications of Information Filtering and Retrieval**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. v. 515, p. 1–23. ISBN 9783642406201 9783642406218.

VÁZQUEZ-LORENTE, H.; HERRERA-QUINTANA, L.; MOLINA-LÓPEZ, J.; LÓPEZ-GONZÁLEZ, B.; PLANELLS, E. Sociodemographic, anthropometric, body composition, nutritional, and biochemical factors influenced by age in a postmenopausal population: A cross-sectional study. **Metabolites**, MDPI, v. 13, n. 1, p. 78, 2023.

VENKATESH, V.; BALA, H. Technology acceptance model 3 and a research agenda on interventions. **Decision sciences**, Wiley Online Library, v. 39, n. 2, p. 273–315, 2008.

VETROVSKY, T.; CUPKA, J.; DUDEK, M.; KUTHANOVA, B.; VETROVSKA, K.; CAPEK, V.; BUNC, V. *et al.* Mental health and quality of life benefits of a pedometer-based walking intervention delivered in a primary care setting. **Acta Gymnica**, Acta Gymnica, v. 47, n. 3, p. 138–143, 2017.

VICINI, S.; BELLINI, S.; ROSI, A.; SANNA, A. An internet of things enabled interactive totem for children in a living lab setting. **2012 18th International Conference on Engineering, Technology and Innovation, ICE 2012 - Conference Proceedings**, Munich, Germany, 2012.

VIJAYAKUMAR, K.; BHUVANESWARI, V. A Ubiquitous first look of IoT Framework for Healthcare Applications. In: **2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)**. Vellore, India: IEEE, 2020. p. 1–7. ISBN 9781728141428.

VOJTA, C.; KINOSIAN, B.; GLICK, H.; ALTSHULER, L.; BAUER, M. S. Self-reported quality of life across mood states in bipolar disorder. **Comprehensive psychiatry**, Elsevier, v. 42, n. 3, p. 190–195, 2001.

VRIJBURG, K.; HERNÁNDEZ-PEÑA, P. Global spending on health: Weathering the storm 2020. **World Health Organization Working paper**, n. 19.4, 2020.

WALLACE, R. B.; HORSFALL, F.; GOUBRAN, R.; EL-HARAKI, A.; KNOEFEL, F. The Challenges of Connecting Smart Home Health Sensors to Cloud Analytics. In: **2019 IEEE Sensors Applications Symposium (SAS)**. Sophia Antipolis, France: IEEE, 2019. p. 1–5. ISBN 9781538677131.

WANG, A.; AN, N.; XIA, Y.; LI, L.; CHEN, G. A Logistic Regression and Artificial Neural Network-Based Approach for Chronic Disease Prediction: A Case Study of Hypertension. In: **2014 IEEE International Conference on Internet of Things(iThings), and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom)**. Taipei, Taiwan: IEEE, 2014. p. 45–52. ISBN 9781479959679.

WANG, C.; WU, Q.; WEIMER, M.; ZHU, E. Flaml: A fast and lightweight automl library. **Proceedings of Machine Learning and Systems**, v. 3, p. 434–447, 2021.

WANG, Z.; XIONG, H.; ZHANG, J.; YANG, S.; BOUKHECHBA, M.; ZHANG, D.; BARNES, L. E.; DOU, D. From personalized medicine to population health: A survey of mhealth sensing techniques. **IEEE Internet of Things Journal**, IEEE, 2022.

WARE, J.; KOSINSKI, M.; KELLER, S. **SF-36 physical and mental health summary scales**: A user's manual. 5th. ed. Boston, MA: Health Assessment Lab, 1994.

WATTANAPISIT, A.; THANAMEE, S. Evidence behind 10,000 steps walking. **Journal of Health Research**, v. 31, n. 3, p. 241–248, 2017.

WEISER, M. The computer for the 21st century. **ACM SIGMOBILE mobile computing and communications review**, ACM New York, NY, USA, v. 3, n. 3, p. 3–11, 1999.

WHO. **WHOQoL**: Measuring quality of life. 1998. https://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/index3.html. Accessed on November 05, 2020.

WHO. Constitution of the world health organization. **Basic documents of the World Health Organization**, Geneva, Switzerland, 2014.

WHO. **Global reference list of 100 core health indicators**. Switzerland, 2015. https://score.tools.who.int/fileadmin/uploads/score/Documents/Enable_data_use_for_policy_and_action/100_Core_Health_Indicators_2018.pdf. Accessed on September 1, 2023.

WHO. Comprehensive mental health action plan 2013–2030. **World Health Organization**, Geneva, 2021.

WHOQoL Group. The Development of the World Health Organization Quality of Life Assessment Instrument (the WHOQOL). In: **Quality of Life Assessment International Perspectives**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994. p. 41–57. ISBN 9783642791253 9783642791239.

WIERINGA, R.; MAIDEN, N.; MEAD, N.; ROLLAND, C. Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. **Requirements engineering**, Springer, v. 11, n. 1, p. 102–107, 2006.

WIERINGA, R.; MORALı, A. Technical Action Research as a Validation Method in Information Systems Design Science. In: **Design Science Research in Information Systems. Advances in Theory and Practice**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. v. 7286, p. 220–238. ISBN 9783642298622 9783642298639.

WIERINGA, R. J. Technical Action Research. In: **Design Science Methodology for Information Systems and Software Engineering**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. p. 269–293. ISBN 9783662438381 9783662438398.

WIERINGA, R. J. Technical Action Research. In: **Design Science Methodology for Information Systems and Software Engineering**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. p. 269–293. ISBN 9783662438381 9783662438398.

WIRTH, R.; HIPP, J. CRISP-DM: Towards a standard process model for data mining. In: SPRINGER-VERLAG LONDON, UK. **Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining**. Manchester, UK, 2000. p. 29–39.

WOHLIN, C. **Experimentation in software engineering**. New York: Springer, 2012. ISBN 9783642290435.

WOHLIN, C.; AURUM, A. Towards a decision-making structure for selecting a research design in empirical software engineering. **Empirical Software Engineering**, Springer, v. 20, n. 6, p. 1427–1455, 2015.

WORLEY, S. L. The extraordinary importance of sleep: the detrimental effects of inadequate sleep on health and public safety drive an explosion of sleep research. **Pharmacy and Therapeutics**, MediMedia, USA, v. 43, n. 12, p. 758, 2018.

WU, X.; ZHU, X.; WU, G.-Q.; DING, W. Data mining with big data. **IEEE transactions on knowledge and data engineering**, IEEE, v. 26, n. 1, p. 97–107, 2013.

YACCHIREMA, D.; SARABIA-JáCOME, D.; PALAU, C.; ESTEVE, M. System for monitoring and supporting the treatment of sleep apnea using iot and big data. **Pervasive and Mobile Computing**, v. 50, p. 25–40, 2018.

YANG, Z.; AZIMI, I.; JAFARLOU, S.; LABBAF, S.; BORELLI, J.; DUTT, N.; RAHMANI, A. M. **Loneliness Forecasting Using Multi-modal Wearable and Mobile Sensing in Everyday Settings**. [*S. l.*], 2023. http://medrxiv.org/lookup/doi/10.1101/2023.06.08.23291165. Accessed on September 01, 2023.

YAO, J.; WARREN, S. Applying the iso/ieee 11073 standards to wearable home health monitoring systems. **Journal of clinical monitoring and computing**, Springer, v. 19, n. 6, p. 427–436, 2005.

YAO, L.; SHENG, Q.; BENATALLAH, B.; DUSTDAR, S.; WANG, X.; SHEMSHADI, A.; KANHERE, S. Wits: an iot-endowed computational framework for activity recognition in personalized smart homes. **Computing**, v. 100, n. 4, p. 369–385, 2018.

YING, X. An Overview of Overfitting and its Solutions. **Journal of Physics Conference Series**, v. 1168, p. 022022, fev. 2019. ISSN 1742-6588, 1742-6596.

YONG, A.; RANA, M. E.; SHANMUGAM, K. Improved Shopping Experience Through RFID Based Smart Shopping System. In: **2022 International Conference on Decision Aid Sciences and Applications (DASA)**. Chiangrai, Thailand: IEEE, 2022. p. 635–644. ISBN 9781665495011.

YONG, B.; XU, Z.; WANG, X.; CHENG, L.; LI, X.; WU, X.; ZHOU, Q. Iot-based intelligent fitness system. **Journal of Parallel and Distributed Computing**, Elsevier, v. 118, p. 14–21, 2018.

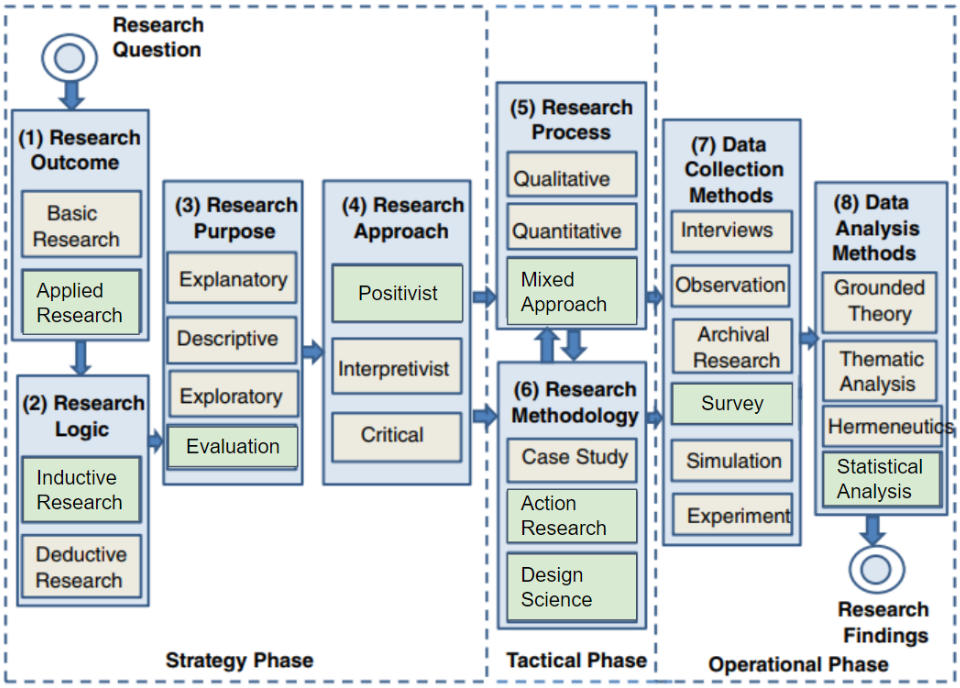
ZHANG, X.; TAN, S. S.; FRANSE, C. B.; ALHAMBRA-BORRÁS, T.; VERMA, A.; WILLIAMS, G.; GRIEKEN, A. V.; RAAT, H. Longitudinal association between physical activity and health-related quality of life among community-dwelling older adults: a longitudinal study of urban health centres europe (uhce). **BMC geriatrics**, BioMed Central, v. 21, n. 1, p. 1–11, 2021.

ZHOU, C.; HU, H.; WANG, C.; ZHU, Z.; FENG, G.; XUE, J.; YANG, Z. The effectiveness of mhealth interventions on postpartum depression: a systematic review and meta-analysis. **Journal of telemedicine and telecare**, SAGE Publications Sage UK: London, England, v. 28, n. 2, p. 83–95, 2022.

APPENDIX A – RESEARCH DECISION MAKING PROCESS

This appendix brings detailed information on the decisions made throughout this thesis. The process for constructing research designs proposed by WOHLIN; AURUM (2015) was adopted to support the choices made in this thesis. Figure 38 presents the research decision-making structure, as well as the options at each point and the decisions made in this work (represented in light green).

Figure 38 – Research decision-making structure



Source: image adapted from WOHLIN; AURUM (2015) to include the decisions of this work.

Regarding the research outcomes and logic, this study is classified as an applied investigation using inductive logic because the researcher is trying to solve a practical problem by moving from specific to general arguments (bottom-up strategy). Furthermore, since this study aims to determine the impact of the proposed solution through empirical data, it is also classified as positivist evaluation research.

In the tactical phase, a mixed approach was selected to use both qualitative and quantitative data, considering the TAR methodology, which combines Action Research and Design Science aspects (WIERINGA; MORALı, 2012). TAR proposes a way to validate software artifacts scaling up from laboratory conditions to the unprotected context of practice (WIERINGA, 2014b). Finally, in the operational phase, the methods to collect data were surveys and case studies, which will be analyzed using statistical methods.

APPENDIX B – SYSTEMATIC LITERATURE MAPPING

This appendix presents detailed information about the systematic mapping discussed in Chapter 2, including the protocol, detailed results, selected studies, and final systematic map. Such information is helpful to researchers who wish to reproduce this kind of study.

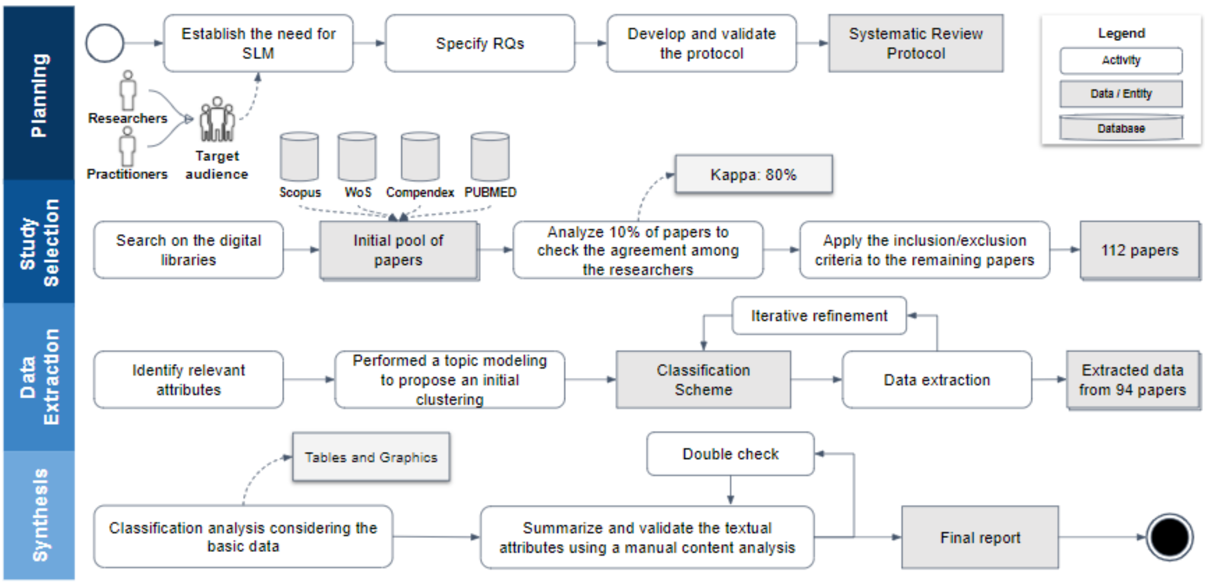
Mapping Protocol

This section presents the protocol built to conduct the systematic mapping discussed in Chapter 2, including the research questions (RQ), search strategy, eligibility criteria, study selection, data extraction, and synthesis strategy.

Overview

As previously mentioned, the process built for this study was adapted from the PETERSEN *et al.* (2008) framework, and it follows the guidelines proposed by KITCHENHAM *et al.* (2015) and GAROUSI *et al.* (2019).

Figure 39 – The process built for this systematic mapping



Source: author.

Figure 39 shows a diagram flow with SLM phases and activities. The first phase, called Planning, was responsible for establishing the need for systematic mapping, specifying the RQs, developing and validating the protocol. This validation was performed by employing pilot tests in the search string and other researchers evaluating the final protocol. The second phase, called Study Selection, aims to find relevant papers in the literature. For this, a search on digital libraries was conducted. Then, it was necessary to apply the eligibility criteria. In this work, two researchers participated in the selection process. Thus, in order to make it possible to parallelize the analysis, a Kappa agreement check (COHEN, 1960) was performed in 10% of the studies. Both researchers made the selection in this set, and the divergences were resolved in meetings.

After the study selection, the Data Extraction phase comes. There are two initial activities in this phase: i) to identify relevant attributes and ii) to perform topic modeling to propose an initial clustering for the papers. As a result, it was built a classification scheme. This scheme was also refined during data extraction.

Finally, in the Synthesis phase, a classification analysis was performed considering the basic data (*e.g.*, year, publication source, and authors' country) and a manual content analysis into the textual attributes of the classification scheme. The following sections detail the critical points of this process.

Research Questions

The need for this systematic literature mapping relies on the increasing interest in IoHT and the absence of studies focused on the systematization of IoHT applied to Quality of Life. Thus, to investigate this area, three Research Questions (RQ) were defined:

RQ1 What is the context of the papers published in the QoL-related IoHT literature?

RQ2 What are the challenges and opportunities related to IoHT for Quality of Life?

RQ3 What is the evidence that IoHT can monitor and improve people's Quality of Life?

RQ1 seeks to understand the context of the works published in the area under investigation. This context is complex because it can involve many aspects. Here, it was analyzed eleven (11) aspects: publication year, country of the first author, venue type, research type (adapted from (WIERINGA *et al.*, 2006)), contribution (BREIVOLD, 2017), type of solution (monitoring, acting, or both), empirical validation strategy (adapted from WOHLIN (2012)), user profile (fetus, child, young, adult, elderly, or anyone), the technology used, health issue addressed, and QoL data (domain and facet).

RQ2 aims to identify the challenges and opportunities in this area and how they are being addressed. These challenges were manually coded using excerpts from the papers, and they can point out interesting gaps for new studies.

RQ3 aims to investigate the strategies used to monitor and improve people's Quality of Life. This question was defined from the hypothesis that there are a significant number of works proposing IoT solutions for healthcare and stating that these solutions promote a better QoL, but only a few studies are concerned with analyzing the QoL in a holistic way.

Search Strategy

To define a suitable search strategy, it is necessary to select proper data sources and build an embracing search string with a high recall (relation between the selected studies and the studies that should be selected) and precision (relation between the studies accepted and the selected studies) (KITCHENHAM *et al.*, 2015). In this way, four scientific databases were chosen based on their representativeness for the health and technology areas: Scopus[1], Web of Science (WoS)[2], Compendex[3], and PubMed[4]. These databases provide access to the most relevant digital libraries for this research, such as IEEE, ACM, Elsevier, Wiley, Springer, and MEDLINE (KITCHENHAM *et al.*, 2015).

Regarding the search string, there are several strategies to elaborate one for systematic mappings (PETTICREW; ROBERTS, 2008). This work applied the PICO methodology (HAYNES *et al.*, 1997; PAI *et al.*, 2004) due to its wide acceptance in the literature. According to PICO, the search string must be composed of four parts: Population, Intervention, Comparison, and Outcome. For each of these parts, it must be defined as a key term. Then, it is necessary to select synonyms for all key terms and connect them using the logic connector "OR". Finally, all parts must be connected using an "AND" connective.

The Population was described as "IoT papers", the Intervention as "Quality of Life", and the Outcome was defined as "Challenges". Below are the synonyms used for each term, and Table 12 shows the final strings used in the databases (for future replications).

– **IoT**: OR Internet of Things OR Internet of Health Things OR Internet of Medical Things OR Cyber-physical System OR Ubiquitous System OR Pervasive System; AND
– **Quality of Life**: OR Health OR eHealth OR Smart Health OR Telemedicine OR Health Promotion OR Well-being OR Wellness; AND
– **Challenges**: OR Barrier OR Opportunity OR Open Issue OR Trend OR Open Questions

[1] Scopus Website: https://www.scopus.com.
[2] Web of Science website: http://www.webofknowledge.com.
[3] Compendex website: http://engineeringvillage.com.
[4] PubMed website: https://pubmed.ncbi.nlm.nih.gov.

Table 12 – Search strings applied in the scientific databases on June 3, 2020.

Database	**Search String**	**# of Papers**
Scopus	*TITLE-ABS-KEY (({iot} OR {internet of things} OR {internet of medical things} OR {internet of health things} OR "cyber-physical system*" OR "ubiquitous system*" OR "pervasive system*") AND ({smart health} OR {telemedicine} OR {health promotion} OR {health} OR {ehealth} OR {e-health}) AND ({quality of life} OR {well-being} OR {wellness}) AND (challenge* OR barrier* OR opportunity* OR "open issue*" OR trend* OR "open question*")) AND (LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))*	147
WoS	*TS=(IoT OR "internet of things" OR "internet of medical things" OR "internet of health things" OR "cyber-physical system" OR "ubiquitous system*" OR "pervasive system*") AND TS=("smart health" OR "telemedicine" OR "health promotion" OR "health" OR "ehealth" OR "e-health") AND TS=("quality of life" OR "well-being" OR "wellness") AND TS=(challenge* OR barrier* OR opportunity* OR "open issue*" OR trend* OR "open question*")*	94
Compendex	*(((((({iot} OR {internet of things} OR {internet of medical things} OR {internet of health things} OR "cyber-physical system*" OR "ubiquitous system*" OR "pervasive system*") WN KY) AND (({smart health} OR {telemedicine} OR {health promotion} OR {health} OR {ehealth} OR {e-health}) WN KY)) AND (({quality of life} OR {well-being} OR {wellness}) WN KY)) AND ((challenge* OR barrier* OR opportunit* OR "open issue*" OR trend* OR "open question*") WN KY))*	120
PubMed	*((({iot} [Title/Abstract] OR {internet of things} [Title/Abstract] OR {internet of medical things} [Title/Abstract] OR {internet of health things} [Title/Abstract] OR "cyber-physical system*" [Title/Abstract] OR "ubiquitous system*" [Title/Abstract] OR "pervasive system*"[Title/Abstract]) AND ({smart health} [Title/Abstract] OR {telemedicine} [Title/Abstract] OR {health promotion} [Title/Abstract] OR {health} [Title/Abstract] OR {ehealth} [Title/Abstract] OR {e-health}[Title/Abstract])) AND ({quality of life} [Title/Abstract] OR {well-being} [Title/Abstract] OR {wellness}[Title/Abstract])) AND (challenge* [Title/Abstract] OR barrier* [Title/Abstract] OR opportunit* [Title/Abstract] OR "open issue*" [Title/Abstract] OR trend* [Title/Abstract] OR "open question*"[Title/Abstract])*	17
Total		**378**
Total without duplicates		**187**

Source: author.

Eligibility Criteria

The eligibility criteria are critical because they transform the study's objective into filters capable of selecting the relevant works to answer the research questions. They also help increase the selection's precision since the search strings tend to be very comprehensive (prioritizing recall) to avoid losing relevant studies. In this way, the eligibility criteria balance these two metrics: precision and recall.

This study selection process uses both the inclusion and exclusion criteria. Thus, it was included papers that discuss IoHT solutions, challenges, or open questions focused on Quality of Life. Concerning the exclusion criteria, they are detailed below:

- Do not discuss IoHT focused on Quality of Life
- Do not be written in English
- Do not be available on the web
- Be a short paper (four pages or less)
- Be available only in the form of abstracts or presentations or expanded summary
- Do not be published in a workshop, conference, journal, magazine, or newspaper

Study Selection

The study selection was performed in three steps by two researchers: (A) the author of this work and (B) a volunteer student. This student was included in order to mitigate the researcher's bias during the study selection, improving the result reliability (WOHLIN, 2012). Thus, after getting the papers using the search string on the scientific databases, it was executed a selection process with the following three steps:

i. **Read the title and abstract**: in this step, the studies were selected considering the reading of titles and abstracts. In this step, an agreement analysis was conducted with the Kappa test considering 10% of the papers. This agreement is essential to check the protocol consistency and enable parallel analysis of the other 90% of primary studies. The agreement level achieved was considered good (CARLETTA, 1996), with a Kappa value of 0.8. In case of incompatibility in the results, meetings were conducted for perspective alignment and review.

ii. **Full reading**: after reading only titles and abstracts, a full reading step was conducted to deepen the understanding of the works and verify the relevance to answer the research questions. This step was carried out for a month with weekly review meetings.

iii. **Final review**: this final step is responsible for resolving doubts about the inclusion of a specific work and reviewing the whole selection process.

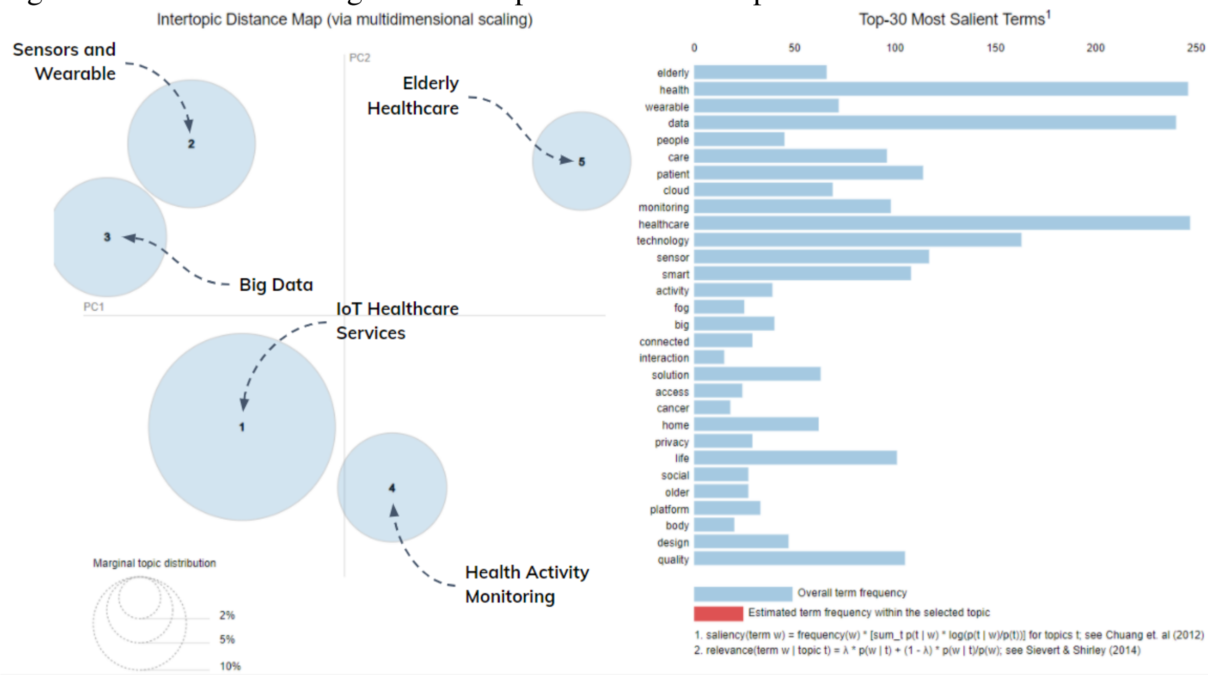
Data Extraction

The data extraction is responsible for extracting valuable data that can be used to characterize the papers' population. For this, initially, it is necessary to identify relevant attributes to be extracted. This identification was made during the full reading of the papers. It included the following fields: goals, research questions, the problem in health and computer science perspectives, challenges, proposed solutions, QoL data, technologies, research type, contribution type, type of solution, user profile, and empirical validation. Also, topic modeling was performed to propose an initial clustering for the studies. As a result of these activities, a classification scheme was proposed, and then it was iteratively refined throughout data extraction.

Regarding topic modeling, this method is widely used to get insights on textual data from the identification of semantically coherent word sets (*i.e.*, the topics) (MEI *et al.*, 2008). In this work, it was adopted one of the most used topic modeling algorithms called

Latent Dirichlet Allocation (LDA) (BLEI *et al.*, 2003), which was implemented[5] using the default parameters and English stopwords provided by the Scikit-learn library (PEDREGOSA *et al.*, 2011). Furthermore, to support the interpretation of the resulting topics, the LDAvis method (SIEVERT; SHIRLEY, 2014) was chosen because it enables an interactive web-based visualization for the two-dimensional topics' distribution, their prevalence in the whole *corpus*, and the most significant terms for each topic.

Figure 40 – Initial clustering with five topics and their interpretations



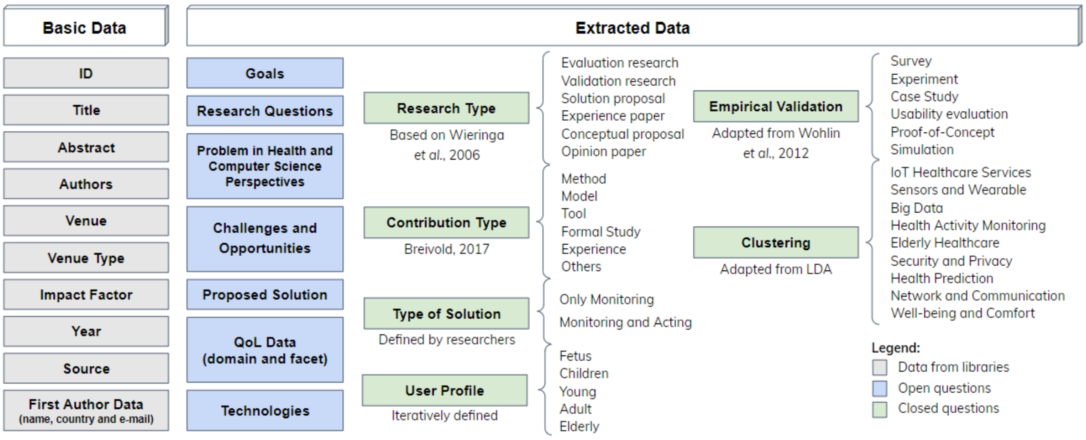
Source: author.

Figure 40 presents the initial clustering with five topics and their interpretation. This number of five topics was empirically defined after experiments with numbers from 3 to 15. Two researchers performed the interpretation of each topic separately, and the divergences were resolved in a meeting. Indeed, the five initial clusters labeled as IoT Healthcare Services, Sensors and Wearable, Big Data, Health Activity Monitoring, and Elderly Healthcare are suitable to describe the papers' population. Even so, during the paper's full reading, it was identified that the LDA algorithm did not clearly detect four more specific clusters. They are Security and Privacy, Health Prediction, Network and Communication, and Well-being and Comfort. After this adjustment in the classification scheme, all papers were reviewed to ensure reliable results.

Figure 41 presents the final classification scheme. This figure shows the basic data obtained directly from the scientific databases, the fields filled by open questions in blue, and the

[5] The LDA implementation is on github.com/great-ufc/healful-thesis repository.

Figure 41 – Final classification scheme



Source: author.

fields with closed options in green.

Synthesis Strategy

After Data Extraction, synthesis was performed to build the systematic map. This phase has two activities: i) a classification analysis considering the basic data and the fields with closed options, and ii) a summarization of the textual attributes using a manual content analysis (KITCHENHAM *et al.*, 2015).

The first activity is more direct, as it only involves analyzing graphs. Thus, it is common to use data visualization techniques to support the evaluation of exploratory studies' results. These techniques combine human strengths, and electronic data processing extends human capabilities to observe insights into the data (KEIM *et al.*, 2008). GARDNER *et al.* (2010) stated that proper visualizations are essential for extracting information and identifying trends in data. The survey proposed by (KUCHER; KERREN, 2015) presents a taxonomy with some text visualization idioms - approaches for creating and manipulating visual representations. This taxonomy helped filter possible visualization strategies by following criteria like analytical and visualization tasks, domain, and data source. Finally, it was used the Tableau software[6] to support this activity.

In the second activity, analyzing the excerpts removed from the papers was necessary to characterize the open fields, such as challenges and solutions. These excerpts were iteratively grouped according to their semantic interpretation to provide valuable and intelligible results. As this activity requires attention, it was carefully performed over two months.

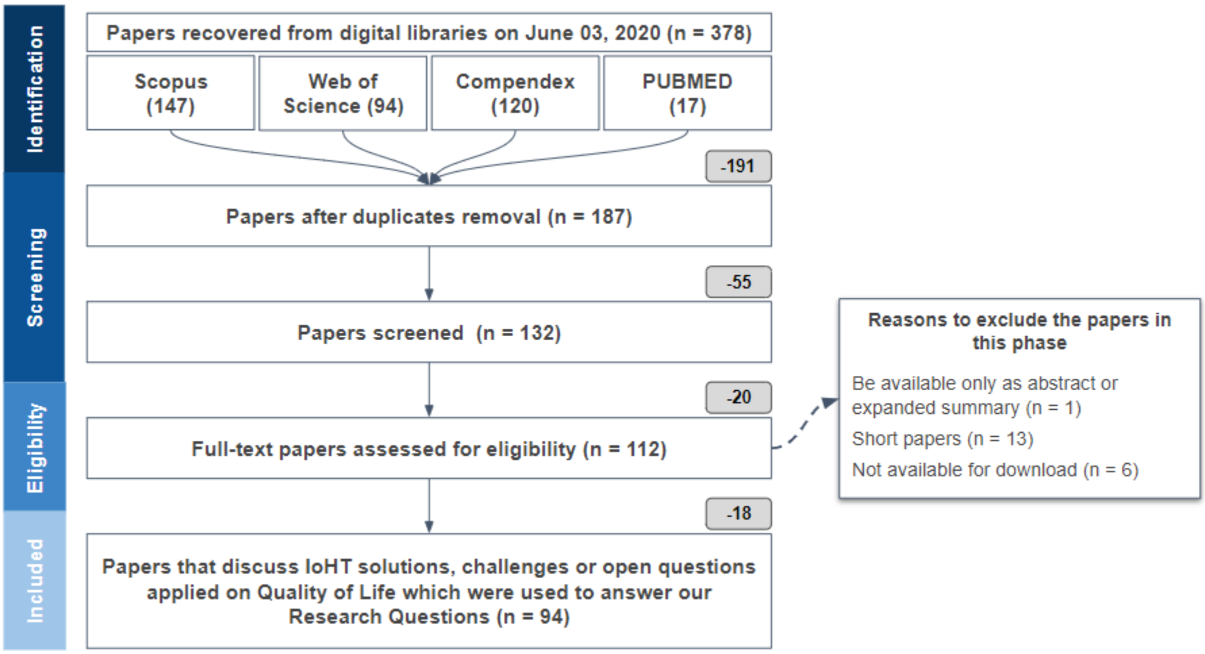
[6] Tableau website: https://www.tableau.com.

Detailed Results

Initially, it was recovered 378 papers from four scientific digital libraries on June 3, 2020[7]. After duplicate removal, 187 studies remained. Then, 55 were removed while reading titles and abstracts, and another 20 were excluded because they were only available as abstract or expanded summaries (1), were short papers (13), and were not available for download (6). Finally, 18 were excluded for not meeting the inclusion criteria. Thus, 94 relevant works were selected to answer the research question.

Figure 42 presents the whole process using a PRISMA Flow Chart (MOHER *et al.*, 2009). The reasons for excluding some papers are also detailed. In that way, this chapter presents the results of this work, discusses their implications, and answers the research questions.

Figure 42 – PRISMA flow chart with the selected and removed papers by each phase



Source: author.

All data extracted for this research and the data visualizations are stored in the github.com/great-ufc/healful-thesis repository. Besides, Appendix B brings tables with the most relevant fields.

[7] The date to conduct the initial search was chosen to include the maximum number of works already published until the beginning of this research. This decision certainly impacts the results of 2020 but does not prevent the historical analysis of this research area. Even so, the search strings were made available (see Table 12) to allow replications or extensions of this study.

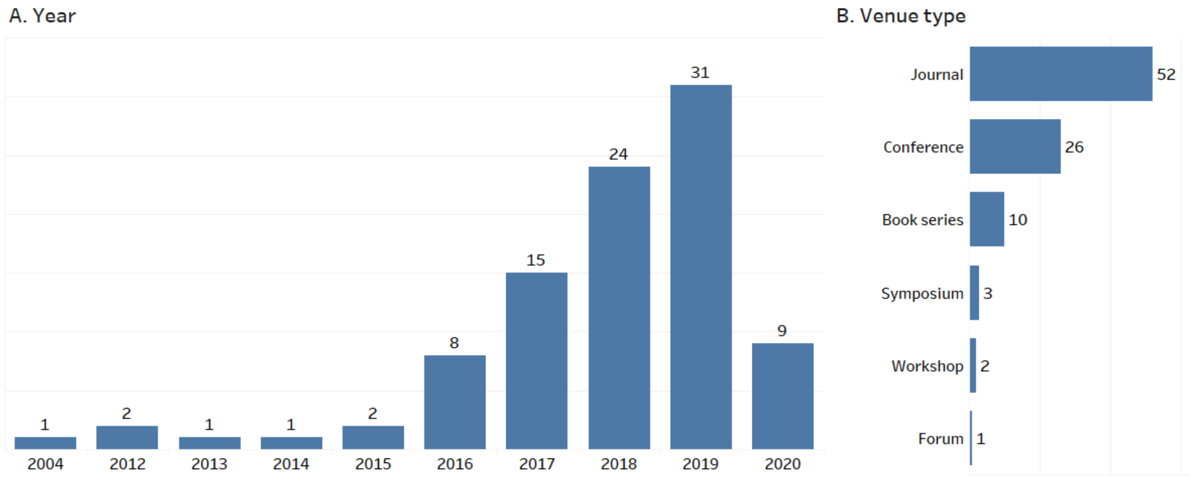
The context of the papers

Analyzing the papers' context is one of the most relevant aspects of understanding a research area. However, this context is complex because it involves many aspects. In order to provide greater detail without increasing complexity, this work considered eleven (11) aspects for the context: year of publication, country of the first author, venue, research type (adapted from (WIERINGA *et al.*, 2006)), contribution (BREIVOLD, 2017), empirical validation strategy (adapted from WOHLIN (2012)), type of solution (only monitoring or monitoring and acting), user profile (fetus, child, young, adult, elderly, or anyone), the technology used, health issue addressed, and QoL data (domain and facet). Among these aspects, the first three were obtained directly from scientific bases, two were defined by the author of this work (the type of solution and user profile), the research type, contribution, and empirical validation strategy used taxonomies already proposed in the literature, and the others were extracted as open fields. Together, these aspects assist in building a robust answer for RQ1. However, for didactic reasons, in this section, RQ1 was divided into eight (8) secondary research questions:

- **SRQ1**: What is the spatiotemporal distribution of papers?
- **SRQ2**: What are the hot topics in this area?
- **SRQ3**: What kind of research has been conducted?
- **SRQ4**: What are the empirical validation strategies used to validate the proposals?
- **SRQ5**: What technologies have been discussed?
- **SRQ6**: Which user profile is of most interest?
- **SRQ7**: What health issues have been addressed?
- **SRQ8**: Which QoL domains and facets have been investigated?

Regarding the spatiotemporal distribution, Figure 43 presents a graph with (A) the number of works published over the years and (B) the distribution per venue type. It is possible to observe an increase in the number of publications except for 2020. However, this low number in 2020 is due to the period in which the search was performed (June 3, 2020). It is highly probable (given the analysis of the works already accepted for publication but not yet available in the databases) that this number in 2020 will exceed the other years. Moreover, 55% of the studies were published in journals. In general, papers published in this type of venue tend to have more consolidated results, given the review's rigor. Papers published in conferences, symposiums, and workshops are usually initial works that bring interesting proposals for discussion at these events.

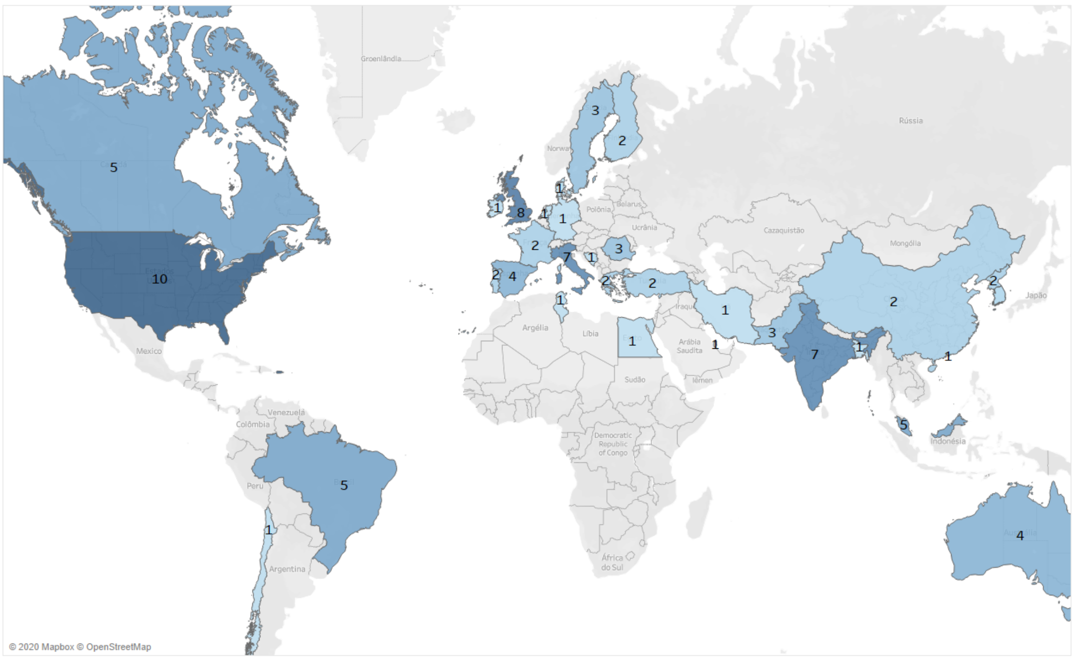
Figure 43 – A panel with (A) the distribution of papers over the years and (B) the number of studies per venue type



Source: author.

In the set of works retrieved in this mapping, despite the large number of papers published in journals and conferences, none of them had more than four studies. This can indicate a knowledge dispersion in different databases. The link https://github.com/great-ufc/healful-thesis can be used to access the complete set with the venue and number of published papers.

Figure 44 – Paper distribution per country considering the first author affiliation



Source: author.

To conclude the analysis of SRQ1, Figure 44 presents a world map with the paper distribution per country considering the first author affiliation. Obviously, this distribution is influenced by the restriction of papers written in English. Even so, considering this language's strength for scientific publications, this distribution represents a sample of the countries that have invested in this area. The United States of America, the United Kingdom, Italy, and India stand out. Thus, these countries' research institutes can represent good options for partnerships and knowledge exchanges in IoHT projects.

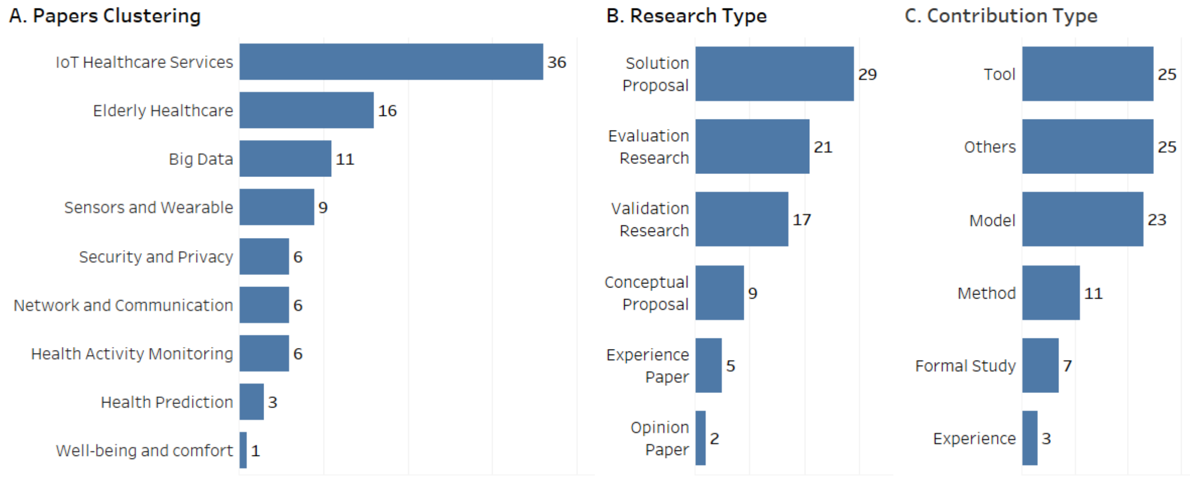
The SRQ1 answer is that since 2004 the number of studies about the Internet of Health Things has been growing with strong interest from North American, European, and Indian researchers. Moreover, about 83% of papers were published in journals and conferences, but none (journal or conference) have a significant concentration of works.

Concerning the hot topics, five clusters were initially proposed based on the LDA results. After the paper's full reading, this clustering was expanded with four new clusters. Then, Figure 45.A shows the distribution of papers by each cluster. In this graphic, it is possible to observe a high interest in IoT Healthcare Services (38%), followed by Elderly Healthcare (17%), Big Data (11%), Sensors, and Wearable (9%). The Security and Privacy, Network and Communication, and Health Activity Monitoring clusters have six studies each, and the last two categories are Health Prediction (3%) and Well-being and Comfort (1%).

These results indicate an interest in developing IoT services for healthcare, activity monitoring, and disease prediction with a special focus on the elderly. Other research areas have also been strengthened to support the development of these services, such as Big Data, Sensors and Wearable, Security and Privacy, and Network and Communication. Another interesting point to highlight here is that although Machine Learning and Cloud Computing did not appear as topics, they are fundamental in IoHT. Many works mention the use of Machine Learning techniques (SIGNORELLI *et al.*, 2019; ELBASANI *et al.*, 2020) and cloud capabilities (AHMAD *et al.*, 2016; YACCHIREMA *et al.*, 2018).

For SRQ2, nine hot topics were found from the execution of the LDA algorithm. They are IoT Healthcare Services, Elderly Healthcare, Big Data, Sensors and Wearable, Security and Privacy, Network and Communication, Health Activity Monitoring, Health Prediction, and Well-being and Comfort. These clusters can be used as a taxonomy to organize the papers published in the IoHT area.

Figure 45 – Number of papers considering (A) the proposed clustering, (B) the research type, and (C) the contribution type



Source: author.

In the SRQ3, three aspects were used to understand what kind of research has been conducted: i) the research type was based on the taxonomy proposed by WIERINGA *et al.* (2006), ii) the contribution type using the BREIVOLD (2017) classification, and iii) the type of solution (monitoring, acting, or both).

The research type has six categories: solution proposal for papers that propose a solution but without a robust validation; validation research for studies that investigate a solution in controlled environments using experiments or simulations; evaluation research for papers that present a practical validation; conceptual proposal for studies that propose conceptual frameworks, ontologies or taxonomies; experience papers for those which report lessons learned from projects in practice; and, opinion papers for those that bring the author's opinion about a specific theme. In Figure 45.B, it is possible to observe the number of papers for each research type. Most works were classified as solution proposals (29), followed by evaluation research (21), validation research (17), conceptual proposal (9), experience paper (5), and two studies were classified as opinion papers. For 11 papers, this aspect was not clearly identified.

The contribution type also has six categories: method, when the contribution is a new method, approach, process, procedure, technique, strategy, or algorithm; model for architectures, conceptual models, frameworks, or system designs; tool, when the paper describes its main contribution as a tool or a system; formal study when the contribution is a theory or a formal analysis; experience when the contributions are lessons learned; and, others for those that not fit in any of these categories. Usually, this last category encompasses secondary studies and papers discussing challenges and opportunities. This mapping found 25 tools, 21 models, 13 methods, seven formal studies, and three experiences. Also, there is a high number of papers classified as others (25). These numbers can be observed in Figure 45.C.

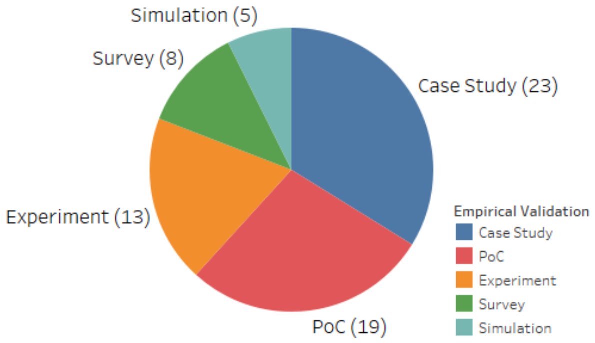
The third aspect is the type of solution. This aspect was used to classify the IoT solutions presented in the paper into only monitoring, or monitoring and acting, and it is important to understand the maturity of the services provided by these solutions (AL-FUQAHA *et al.*, 2015). In this way, 47 papers proposed only monitoring solutions, and six studies presented monitoring and acting solutions. For 41 studies, this aspect was not clearly identified. Thus, this result reinforces the discussion made by AL-FUQAHA *et al.* (2015) that there are many information aggregation services, and the IoT needs to evolve for more collaborative-aware services (which use the obtained data to react appropriately) and ubiquitous services (which provide collaborative-aware services anytime, for anyone, anywhere). Also, the low number of collaborative-aware and ubiquitous services can be seen as a gap in the Internet of Health Things area.

Thus, the SRQ3 answer is that most papers present solution proposals focused on monitoring health data and classified as tools or models. Also, only 22% were classified as evaluation research indicating that the percentage of solutions evaluated in practice is still low.

Regarding the empirical validation strategies, it was used the classification proposed by WOHLIN (2012), including usability evaluation, Proof-of-Concept (PoC), and simulation. These three last categories were included to expand the classification scheme. Thus, only 68 papers (72%) present a well-described empirical validation (see Figure 46). There are 23 case studies, 19 PoCs, 13 experiments, eight surveys, and five simulations. This result directly correlates with the research type, as case studies are used to evaluate a solution in practice. Also, considering that 28% of works did not present a strong validation and that 54,4% were evaluated in controlled environments with PoCs, experiments, and simulations, there is a large room for opportunities for partnerships with industry to conduct practical validations.

The answer for SRQ4 is that only 72% of studies describe an empirical validation, and the most used strategies are case studies (33,8%), PoCs (27,9%), controlled experiments (19,1%), surveys (11,8%), and simulations (7,4%). In addition, no usability evaluation was found.

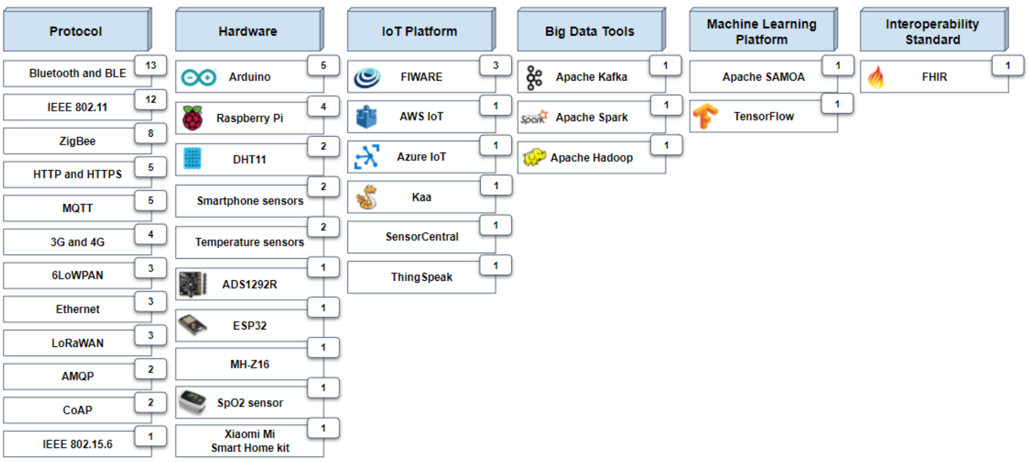
Figure 46 – Number of papers that present a well-described empirical validation



Source: author.

During the papers' full reading, the technologies used to build the IoT solutions were extracted. These technologies were then grouped into six categories: protocols, hardware, IoT platforms, big data tools, machine learning platforms, and interoperability standards. Figure 47 shows the technologies ordered by the number of mentions. However, due to the low number of recurrences, it is not possible to point out trends or gaps.

Figure 47 – Technologies mentioned in the solutions found in the mapping



Source: author.

Regarding SRQ5, many different technologies were found, but none with a significant occurrence. In the protocols category, the three most mentioned were Bluetooth and Bluetooth Low Energy (BLE), IEEE 802.11, and ZigBee. For the hardware, two well-known platforms have the highest citations: Arduino and Raspberry PI. For the IoT platform, FIWARE was used three times. Five tools were cited in big data and machine learning: Apache Kafka, Apache Spark, Apache Hadoop, Apache SAMOA, and TensorFlow. Finally, one paper was found mentioning the usage of FHIR as an interoperability standard.

For the user profile, the classification scheme has five categories (fetus, children, young, adult, and elderly) in order to cover all age groups. Furthermore, the label "everyone" was used for those papers in which a solution was proposed for all profiles. In this way, it was found 33 studies labeled as "everyone", 19 papers focused on the elderly, one study focused on children, and one for the fetus. Despite the high number of studies in which it was not possible to clearly identify the user profile (42%), a prominent interest in projects focused on older adults can be observed. This trend is related to the population aging discussed at the beginning of this work and the search for longevity with a higher quality of life.

The answer for SRQ6 is that considering the papers in which it was possible to identify a specific user profile, the most expressive interest is for the elderly. This trend was expected due to the pressure that aging populations are imposing on health systems.

Table 13 – Most mentioned health issues in the papers.

Group	Health Issue	Papers
Diseases	Sleep apnea syndrome (SAS) (HAOYU *et al.*, 2019; YACCHIREMA *et al.*, 2018)	2
	Stress (KOLDIJK *et al.*, 2016; PRIYADARSHINI *et al.*, 2018)	2
	Cancer (ONASANYA; ELSHAKANKIRI, 2019; SIGNORELLI *et al.*, 2019)	2
	Cardiovascular and cardio-respiratory diseases (GATOUILLAT *et al.*, 2018; ALBAHRI *et al.*, 2019)	2
	Chronic diseases (FILHO; AQUINO G.S., 2017; BELESIOTI *et al.*, 2018)	2
	Hypertension health issues (WANG *et al.*, 2014; PRIYADARSHINI *et al.*, 2018)	2
	Diabetes (PRIYADARSHINI *et al.*, 2018)	1
Health Monitoring	Continuous and real-time monitoring of vital sign (DISI *et al.*, 2018; QURESHI; KRISHNAN, 2018; CHONG *et al.*, 2019)	3
	Monitor air and water quality (in and outdoor) (LJUBOJEVIC *et al.*, 2016; SHAFI *et al.*, 2018; LIU *et al.*, 2018)	3
	Fitness tracking (still limited when compared to vital body parameters in a clinical context) (ATHAVALE; KRISHNAN, 2017; CARBONARO *et al.*, 2018)	2
	Monitor patient behaviors	1
	Monitor patients who live further away from the city	1
Elderly Healthcare	Cognitive decline and dementia (MILOVICH; BURLESON, 2017; KANG; KANG, 2017; ENSHAEIFAR *et al.*, 2018)	3
	Falls (RAFFERTY *et al.*, 2019; HUYNH *et al.*, 2020)	2
	Self-independence of elderly people (MARQUES; PITARMA, 2019; BELEN *et al.*, 2019)	2
	Parkinson's disease (TEWELL *et al.*, 2019)	1
	Reduced physical ability in elderly (GKOUSKOS; BURGOS, 2017)	1
Illness detection	Early illness detection (MEKKI *et al.*, 2017; JAGADEESWARI *et al.*, 2018; QUINN *et al.*, 2019)	3
	Detect emotional state (MANO *et al.*, 2019)	1
	Detect fetal bio-signals for the early detection of embryonic developmental impairments (BALAKRISHNA *et al.*, 2019)	1
	Disease prevention and personalized wellness management (MCRAE *et al.*, 2016)	1
Ethics	Ethical responsibility over health data (MITTELSTADT, 2017; LAURIE, 2019)	2
Health at Work	Provide an environment that brings better well-being to employees (GOMEZ-CARMONA *et al.*, 2018; NABUCO *et al.*, 2019)	2

Source: author.

SRQ7 guided the investigation to understand what kinds of health issues have been addressed in the QoL-related IoHT literature. It was found 33 health issues, and they were clustered into six groups: diseases, health monitoring, elderly healthcare, illness detection, ethics, and health at work. Table 13 presents all these issues and their occurrences.

Regarding diseases, studies about sleep apnea syndrome, stress, cancer, cardiovascular and cardio-respiratory diseases, chronic diseases, hypertension, and diabetes were found. In the health monitoring category, the most recurrent issue is how to provide continuous and real-time monitoring of vital signs. For elderly healthcare, dementia and falls were the most mentioned. Finally, it was also found studies focused on early illness detection, ethical responsibility over health data, and how to provide a better environment to employees.

Thus, it is possible to answer the SRQ7 stating that many health issues can take advantage of the Internet of Things. Considering the data extracted in this work, 33 health issues were found, and they were grouped into six categories: diseases, health monitoring, elderly healthcare, illness detection, ethics, and health at work. It is also worth mentioning that soon, there should be a growing interest in the early detection of health problems based on consolidating monitoring systems and applying Machine Learning techniques.

Table 14 – Papers that correlate their proposals with the WHO QoL domains and facets.

Domain	Facet	Papers	Nº
Physical	Activities of daily living	(LJUBOJEVIC *et al.*, 2016; GATOUILLAT *et al.*, 2018; REN *et al.*, 2020; HUYNH *et al.*, 2020)	4
	Energy and fatigue	(GATOUILLAT *et al.*, 2018; DOBRE *et al.*, 2019)	2
	Mobility	(DOBRE *et al.*, 2019; HUYNH *et al.*, 2020)	2
	Sleep and rest	(YACCHIREMA *et al.*, 2018)	1
	Pain and discomfort	(VICINI *et al.*, 2012)	1
	Dependence on medicinal substances and medical aids	(VICINI *et al.*, 2012)	1
	Work capacity	-	0
Psychological	Thinking, learning, memory and concentration	(MILOVICH; BURLESON, 2017; DOBRE *et al.*, 2019)	2
	Self-esteem	(VICINI *et al.*, 2012)	1
	Positive feelings	(MANO *et al.*, 2019)	1
	Negative feelings	(MANO *et al.*, 2019)	1
	Bodily image and appearance	-	0
	Religion/Spirituality/Personal beliefs	-	0
Social relationships	Personal relationships	(VICINI *et al.*, 2012)	1
	Social support	(MILOVICH; BURLESON, 2017)	1
	Sexual activity	-	0
Environment	Physical environment (pollution/noise/traffic/climate)	(LJUBOJEVIC *et al.*, 2016; LIU *et al.*, 2018)	2
	Home environment	(REN *et al.*, 2020)	1
	Health and social care: accessibility and quality	(NABUCO *et al.*, 2019)	1
	Financial resources	-	0
	Opportunities for acquiring new information and skills	-	0
	Participation in and opportunities for recreation/leisure	-	0
	Freedom, physical safety and security	-	0
	Transport	-	0

Source: author.

Finally, the last secondary research question (SRQ8) is about how the QoL domains and facets have been investigated. Here, it was considered the WHO's QoL definition (WHOQoL Group, 1994), and the domains and facets defined by the WHOQOL-BREF questionnaire (SKEVINGTON *et al.*, 2004).

In WHOQOL-BREF, there are four domains (physical, psychological, social, and environmental) and 24 facets. The results showed that despite the large number of studies that use the term "Quality of Life", few studies (11,7%) correlate this term to the definition proposed by WHO. In addition, although there are proposals for individualized monitoring of aspects related to QoL, no initiatives have been found for comprehensive and holistic Quality of Life monitoring using intelligent objects in IoT environments. Table 14 brings the results found for this aspect.

The answer for SRQ8 is that many researchers have investigated strategies to improve people's QoL, but these strategies generally focus on a specific QoL aspect or health issues. Also, few papers correlated their results with reference models for measuring the QoL, such as the WHOQL-BREF. The domain with the most mention was the physical, followed by the psychological, environmental, and social relationships.

Together, the SRQ1 to SRQ8 answers provide a comprehensive overview of the context of the papers published in the QoL-related IoHT literature. Thus, the answer for RQ1 can be written by summarizing their discussions.

RQ1 What is the context of the papers published in the QoL-related IoHT literature?

Summarized Answer:

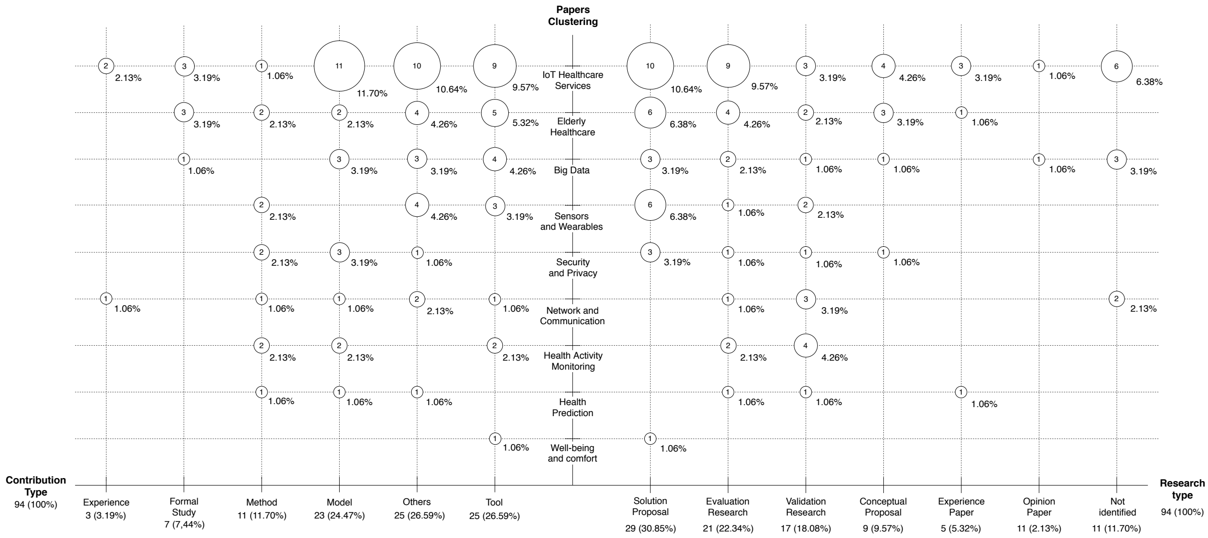
- **Spatiotemporal distribution**: in the last years, the number of studies has grown with a strong interest in the USA, United Kingdom, Italy, and India. Also, 55% and 27% of the papers were published in journals and conferences, respectively. However, no significant concentration of papers was found in specific venues.
- **Hot topics**: nine hot topics were found: IoT Healthcare Services (38.3%), Elderly Healthcare (17%), Big Data (11.7%), Sensors and Wearable (9.6%), Security and Privacy (6.4%), Network and Communication (6.4%), Health Activity Monitoring

(6.4%), Heath Prediction (3.2%), and Well-being and Comfort (1%). These clusters can be used as a taxonomy for the IoHT studies.

– **Kind of research**: many papers classified as solution proposals have a monitoring tool as a contribution. Also, 78% of the works were not evaluated in practical scenarios.
– **Empirical validation strategies**: the distribution of strategies was 33,8% of case studies, 27,9% of proofs-of-concept, 19,1% of experiments, 11,8% of surveys, 7,4% of simulations, and was not found usability evaluations. This result reinforces the need to reinforce partnerships with the industry to conduct practical validations.
– **Technologies**: it was found many different technologies and they were classified into six categories, but none with a significant recurrence. The most mentioned were Bluetooth and BLE as protocols, Arduino in the hardware category, and FIWARE as an IoT platform.
– **User profile**: the major interest is for the elderly profile. This result is a consolidated trend due to population aging and its pressure on the health system. However, with the development of solutions for the early detection of health problems, this trend can gradually change, providing greater attention to young people and adults.
– **Health issues**: it was found 33 health issues into six categories: diseases, health monitoring, elderly healthcare, illness detection, ethics, and health at work. With the consolidation of monitoring systems, a growing interest in applying Machine Learning to detect early diseases should be noted.
– **QoL domains and facets**: many researchers have investigated strategies to improve people's QoL, but these strategies are generally focused on a specific QoL aspect or health issues. Moreover, it was not found studies focused on providing a semantic correlation for the QoL data and able to use smart objects in seamless QoL monitoring.

Finally, Figure 48 presents a classic map visualization correlating the paper clustering, the contribution type, and the research type. This graphic brings an overview of the papers' concentration (the upper area near the center of the image), and little-explored subjects (the lower left and right areas). It can be better analyzed using its expanded version in Appendix B or downloading the full version on GitHub (github.com/great-ufc/healful-thesis/).

Figure 48 – Systematic mapping correlating the paper clustering, the contribution type, and the research type. To zoom this visualization, access the GitHub repository with the link: github.com/great-ufc/healful-thesis/



Source: author.

IoHT to monitor and improve the people's Quality of Life

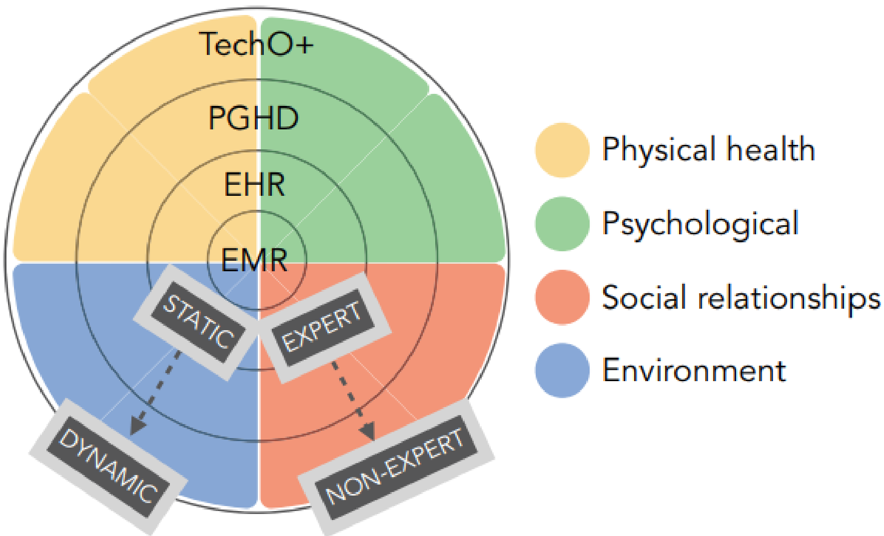
It is a consensus that health and quality of life are closely related (ESTRADA-GALINANES; WAC, 2018), and there is also a convergence between technology and health researchers that technology (when applied correctly) can improve people's quality of life (FILHO; JUNIOR, 2017; VALLEE *et al.*, 2016). In addition, the World Health Organization (WHO) indicates that measuring QoL can provide valuable data for medical practice (ESTRADA-GALINANES; WAC, 2018). However, there is no significant interest in technological strategies to measure this QoL gain. Currently, the most known strategies to measure QoL are based on questionnaires, which are tiring, hardly engage the user, and still suffer from the bias of the respondent (SANCHEZ *et al.*, 2015; DOBRE *et al.*, 2019). In this way, seamless and unnoticeable (ubiquitous) IoHT-based monitoring can be more helpful in detecting variations in the users' Quality of Life, promoting early interventions. This idea seems simple, but it involves the challenge of dealing with several smart objects in indoor and outdoor environments to map the various domains and facets that compose QoL.

Considering this context, the RQ3 was proposed from the hypothesis that many works propose IoHT solutions to improve people's QoL. However, only a few studies are concerned with holistically analyzing this indicator. Only eleven (11) papers explicitly mentioned some QoL domain or facet in the mapping result. As previously mentioned in Section B, there are still unexplored facets, such as work capacity, bodily image, and transport.

The study published by DOBRE *et al.* (2019) is probably the most correlated with the idea presented in this section. The authors present a broad discussion about QoL for the elderly and its relation to health. In addition, it was identified a set of instruments to measure QoL, such as EQ-SD-3L, SF-36, WHOQOL-BREF, WHOQOL-OLD, and many others. They conclude by proposing an architecture for non-intrusive monitoring of older adults. The main drawback of this proposal is that the monitoring module still uses questionnaires to acquire data, and it was not presented any strategy focused on the semantic structure of the QoL domain or facet and how the IoT smart devices can produce data to infer the QoL measure.

The other works address specific points, for example, QoL of hospitalized children (VICINI *et al.*, 2012), plant wall management to improve indoor comfort (LIU *et al.*, 2018), sleep monitoring for apnea treatment (YACCHIREMA *et al.*, 2018), recognition of emotions from face and heart rate analysis (MANO *et al.*, 2019), a wearable to monitor the cardio-respiratory functions (GATOUILLAT *et al.*, 2018), a cloud-based system for fall detection (HUYNH *et al.*, 2020), daily activity monitoring using a cyber-physical system (NABUCO *et al.*, 2019), the usage of social media to reduce the risk of cognitive decline (MILOVICH; BURLESON, 2017), indoor air quality management (LJUBOJEVIC *et al.*, 2016), and the liquid level sensing for smart homes (REN *et al.*, 2020).

Figure 49 – The wheel of QoL data



Source: ESTRADA-GALINANES; WAC (2018).

Finally, the position paper of ESTRADA-GALINANES; WAC (2018), despite not focusing on a specific QoL domain or facet, presented the requirements and design choices for an Open Health Archive (OHA) to avoid the data silos issue. The idea is to facilitate access to personal health data to encourage the development of new QoL-based studies. In addition, the authors presented an interesting organization for QoL data. The QoL wheel – shown by Figure 49 – brings the QoL domains (physical health, psychological, social relationships, and environment) and the models for obtaining data. The closer to the center, the more data are collected by experts and are generally static. In the center, there are electronic medical records (EMR) with data collected directly by doctors. Then, electronic health records (EHR) with more comprehensive data on the patient. The last two levels are the patient-generated health data (PGHD), which includes data from family members and obtained by caregivers, and the technology-reported outcome plus (TechO+), with data obtained from sensors and other IoT devices. This organization suggests a path for the evolution of the QoL-based IoHT systems.

RQ3 What is the evidence that IoHT can monitor and improve people's Quality of Life?
Summarized Answer: there is an agreement among researchers that the IoT can improve people's QoL. However, most studies did not seek to measure this gain. In addition, few studies proposed a QoL automated monitoring approach. Finally, there is a lack of studies that holistically consider QoL, providing models for the semantic organization of the data that compose it and using artificial intelligence to estimate its value.

Final Remarks

This study was conducted to summarize IoHT literature focused on Quality of Life. The work context involves the need for solutions to support the healthcare system to provide better living conditions while optimizing resources. Its motivation relies on the valuable knowledge spread in the academic literature, which can support researchers and professionals in deciding which areas should be investigated. Before this study, several surveys and systematic reviews were found, but none of them had the same objective and updated data.

In this way, it was performed a systematic literature mapping following the well-known guidelines proposed by KITCHENHAM *et al.* (2015) and GAROUSI *et al.* (2019), and adapting the process defined by PETERSEN *et al.* (2008). Two of the differentials of this adaptation were the inclusion of a topic modeling activity for clustering the papers and building a robust classification scheme. Moreover, this kind of empirical methodology helps answer questions about the challenges and opportunities in a specific research area. Thus, three research questions were defined:

RQ1 What is the context of the papers published in the QoL-related IoHT literature?

The rationale for this research question is the strong need to get the studies' context to understand the IoHT area's behavior. Eleven (11) aspects were chosen to characterize this context, including open fields, fields proposed by researchers, and fields based on taxonomies already published. The answer to this question is comprehensive, but it is possible to highlight the growing interest in IoHT studies, primarily aimed at the elderly. In general, there is still room for more partnerships with the industry to carry out validations in practice. Finally, several solutions for monitoring and diagnosing diseases have been proposed, but it is expected to increase the number of solutions that use machine learning for an increasingly early diagnosis.

RQ2 What are the challenges and opportunities related to IoHT for Quality of Life?

This question is focused on understanding the challenges to point out research paths and opportunities. Thus, passages were extracted and iteratively categorized to provide more intelligible information about these challenges. The eight groups ordered by their recurrence were General IoT Challenges, Security and Privacy, Data Science, Network and Communication, Sensors and Wearable, Software Engineering, Human-Computer Interaction, and Cloud Computing. Regarding these challenges, there is a desire for increasingly personalized and intelligent services that provide continuous, fast, secure, and effective health data monitoring and processing.

RQ3 What is the evidence that IoHT can monitor and improve people's Quality of Life?

The last research question was defined considering the hypothesis that many studies discuss that their proposals improve the users' quality of life, but few effectively measure this gain. Also, most measurement methods are based on questionnaires, which makes it challenging to engage the users. Thus, this question can validate the beginning of a more in-depth investigation toward a semantic structure of the QoL domains and facets and an intelligent model for capturing and inferring this metric using the data collected by smart objects. In fact, the results confirmed the hypothesis. Unfortunately, only some studies seek a holistic approach to dealing with QoL, and the works closer to this proposal still use questionnaires to collect data.

Thus, regarding the trends, it was identified the development of IoHT focused on elderly healthcare, a large number of tools and models aimed at monitoring health data, the usage of big data for personalized systems, and the extensive adoption of mobile devices as wearables.

Moreover, concerning the gaps, there is a need for collaborative-aware and ubiquitous services, that is, services that use intelligence to anticipate events and to act in the environment to improve the living conditions, partnerships with the industry to conduct validations in practice, methods and protocols capable of guaranteeing data security and privacy even on restricted devices, network designs to ensure low latency and high reliability, techniques for validation and verification of fitness tracking apps, approaches to the development of intelligent systems with the proximity of the domain experts, as well as studies focused on user experience when using this type of system.

Selected Studies

This section brings two tables (Tables 15, 16 and 17) with the 94 studies selected in this systematic mapping sorted by year, and Table 18 with the description of the tools. Also, the public dataset created from the data extraction conducted in this work is public and can be accessed by the following GitHub repository: https://github.com/great-ufc/healful-thesis.

Table 15 – Primary studies selected (1 to 30)

ID	Citation	Research Type	Contribution	User Profile	Empirical Validation	Our Clustering
1	BROWN *et al.* (2004)	Evaluation Research	Model	Conceptual Model	Case Study	Elderly Healthcare
2	CALHOUN *et al.* (2012)	Not identified	Tool	Platform	Survey	Networks
3	VICINI *et al.* (2012)	Experience Paper	Method	Strategy	Case Study	IoT Healthcare Services
4	BIDMESHKI; JAFARI (2013)	Evaluation Research	Model	Architecture	Experiment	IoT Healthcare Services
5	WANG *et al.* (2014)	Validation Research	Method	Approach	Experiment	Health Prediction
6	AHMED *et al.* (2015)	Solution Proposal	Formal Study	Standards	PoC	IoT Healthcare Services
7	RAHMANI *et al.* (2015)	Evaluation Research	Tool	Platform	Case Study	Big Data
8	LJUBOJEVIC *et al.* (2016)	Solution Proposal	Method	Monitoring Method	Experiment	Wearables and Sensors
9	KOLDIJK *et al.* (2016)	Solution Proposal	Model	Framework	PoC	IoT Healthcare Services
10	BANDODKAR *et al.* (2016)	Validation Research	Method	Strategy	Survey	Wearables and Sensors
11	KARAMITSIOS *et al.* (2016)	Validation Research	Others	Recommendations	Simulation	Networks
12	MATTSSON *et al.* (2016)	Evaluation Research	Others	Recommendations	Case Study	Wearables and Sensors
13	MCRAE *et al.* (2016)	Solution Proposal	Tool	Sensors and Wearable	PoC	Wearables and Sensors
14	AHMAD *et al.* (2016)	Validation Research	Model	Framework	Experiment	IoT Healthcare Services
15	SRINIVASAN *et al.* (2016)	Evaluation Research	Model	Framework	Case Study	Big Data
16	ATHAVALE; KRISHNAN (2017)	Solution Proposal	Others	Recommendations	Survey	Wearables and Sensors
17	ALTHOFF (2017)	Experience Paper	Others	Recommendations	Not identified	Health Prediction
18	ARULANANTHAN; HANIFA (2017)	Validation Research	Others	Survey	Survey	IoT Healthcare Services
19	AZIMI *et al.* (2017)	Conceptual Proposal	Others	Review	Not identified	Elderly Healthcare
20	KANG; KANG (2017)	Solution Proposal	Formal Study	User Experience	PoC	Elderly Healthcare
21	FILHO; AQUINO G.S. (2017)	Solution Proposal	Model	Architecture	PoC	IoT Healthcare Services
22	MEKKI *et al.* (2017)	Solution Proposal	Model	Architecture	PoC	Security
23	FILHO; JUNIOR (2017)	Not identified	Others	Review	Not identified	IoT Healthcare Services
24	VALLEE *et al.* (2016)	Solution Proposal	Others	Recommendations	Not identified	IoT Healthcare Services
25	GKOUSKOS; BURGOS (2017)	Conceptual Proposal	Others	Recommendations	Not identified	Elderly Healthcare
26	MITTELSTADT (2017)	Conceptual Proposal	Others	Recommendations	Not identified	Security
27	LATIF *et al.* (2017)	Evaluation Research	Experience	Lessons Learned	Case Study	Networks
28	MILOVICH; BURLESON (2017)	Validation Research	Method	Strategy	Experiment	Elderly Healthcare
29	PAWAR; GHUMBRE (2016)	Solution Proposal	Method	Security Algorithm	Survey	Security
30	SHETH *et al.* (2017)	Not identified	Others	Recommendations	Not identified	Big Data

Source: author.

Table 16 – Primary studies selected (31 to 88)

ID	Citation	Research Type	Contribution	User Profile	Empirical Validation	Our Clustering
31	QURESHI; KRISHNAN (2018)	Solution Proposal	Others	Review	Not identified	Wearables and Sensors
32	ENSHAEIFAR *et al.* (2018)	Experience Paper	Tool	Monitoring System	Case Study	Elderly Healthcare
33	ALKHATIB *et al.* (2018)	Solution Proposal	Formal Study	Privacy	PoC	Elderly Healthcare
34	GOMEZ-CARMONA *et al.* (2018)	Solution Proposal	Formal Study	User Experience	PoC	IoT Healthcare Services
35	JAGADEESWARI *et al.* (2018)	Not identified	Others	Recommendations	Survey	Big Data
36	GATOUILLAT *et al.* (2018)	Validation Research	Tool	Sensors and Wearable	Experiment	Wearables and Sensors
37	BELESIOTI *et al.* (2018)	Evaluation Research	Tool	Smart Healthcare System	Case Study	IoT Healthcare Services
38	NEWCOMBE *et al.* (2017)	Solution Proposal	Others	Survey	Not identified	Elderly Healthcare
39	REDA *et al.* (2018)	Solution Proposal	Tool	Monitoring System	PoC	Big Data
40	SALAMA *et al.* (2018)	Solution Proposal	Model	Framework	Not identified	Security
41	DAUWED *et al.* (2018a)	Not identified	Others	Survey	Survey	IoT Healthcare Services
42	SILVA; JUNIOR (2018)	Not identified	Others	Review	Not identified	Networks
43	PRIYADARSHINI *et al.* (2018)	Evaluation Research	Model	Conceptual Model	Case Study	Health Prediction
44	SHAFI *et al.* (2018)	Evaluation Research	Tool	Monitoring System	Case Study	IoT Healthcare Services
45	CARBONARO *et al.* (2018)	Conceptual Proposal	Tool	Monitoring System	Not identified	Big Data
46	ABDELNAPI *et al.* (2018)	Not identified	Others	Survey	Not identified	IoT Healthcare Services
47	DAUWED *et al.* (2018b)	Conceptual Proposal	Others	Recommendations	Not identified	IoT Healthcare Services
48	YACCHIREMA *et al.* (2018)	Validation Research	Model	Architecture	Experiment	Health Activity Monitoring
49	DISI *et al.* (2018)	Evaluation Research	Model	Architecture	Case Study	IoT Healthcare Services
50	YAO *et al.* (2018)	Validation Research	Tool	Development Support	Experiment	Health Activity Monitoring
51	LIU *et al.* (2018)	Solution Proposal	Tool	Monitoring System	PoC	Well-being and comfort
52	RAMU (2018)	Evaluation Research	Model	Framework	Case Study	Security
53	MARTíNEZ-CARO *et al.* (2018)	Conceptual Proposal	Model	Conceptual Model	Not identified	IoT Healthcare Services
54	RODRIGUES *et al.* (2018)	Solution Proposal	Others	Survey	Survey	IoT Healthcare Services
55	LAURIE (2019)	Not identified	Formal Study	Ethics	PoC	Big Data
56	ONASANYA; ELSHAKANKIRI (2019)	Evaluation Research	Tool	Smart Healthcare System	Experiment	IoT Healthcare Services
57	ALBAHRI *et al.* (2019)	Evaluation Research	Model	Framework	Case Study	IoT Healthcare Services
58	KHAREL *et al.* (2019)	Evaluation Research	Tool	Monitoring System	Case Study	IoT Healthcare Services
59	DONATI *et al.* (2019)	Evaluation Research	Tool	Monitoring System	Case Study	IoT Healthcare Services
60	HAOYU *et al.* (2019)	Validation Research	Method	Monitoring Method	Experiment	Health Activity Monitoring
61	NABUCO *et al.* (2019)	Solution Proposal	Model	Conceptual Model	Survey	IoT Healthcare Services
62	KHODKARI *et al.* (2018)	Solution Proposal	Model	Ontology	Not identified	IoT Healthcare Services
63	BAJENARU; CUSTURA (2019)	Solution Proposal	Tool	Platform	PoC	Elderly Healthcare
64	RAFFERTY *et al.* (2019)	Validation Research	Tool	Monitoring System	Simulation	Elderly Healthcare
65	WALLACE *et al.* (2019)	Not identified	Others	Recommendations	Not identified	IoT Healthcare Services
66	BALAKRISHNA *et al.* (2019)	Solution Proposal	Others	Recommendations	Not identified	Wearables and Sensors
67	QUINN *et al.* (2019)	Evaluation Research	Method	Approach	Case Study	Health Activity Monitoring
68	POPENTIU-VLăDICESCU; AL-BEANU (2019)	Validation Research	Others	Recommendations	PoC	Big Data
69	CHUI *et al.* (2019)	Solution Proposal	Tool	Monitoring System	Not identified	Big Data
70	MANO *et al.* (2019)	Validation Research	Model	Architecture	Simulation	Health Activity Monitoring
71	BELEN *et al.* (2019)	Solution Proposal	Formal Study	Assistive Technology	PoC	Elderly Healthcare
72	ESTRADA-GALINANES; WAC (2018)	Opinion Paper	Model	Conceptual Model	Not identified	Big Data
73	TEWELL *et al.* (2019)	Evaluation Research	Tool	Sensors and Wearable	Case Study	Health Activity Monitoring
74	FAHEEM *et al.* (2019)	Validation Research	Method	Routing Protocol	Simulation	Networks
75	DOBRE *et al.* (2019)	Solution Proposal	Tool	Monitoring System	PoC	Elderly Healthcare
76	ONASANYA *et al.* (2019)	Experience Paper	Tool	Smart Healthcare System	Case Study	IoT Healthcare Services
77	CHONG *et al.* (2019)	Solution Proposal	Tool	Sensors and Wearable	Not identified	Wearables and Sensors
78	SIGNORELLI *et al.* (2019)	Solution Proposal	Formal Study	Standards	PoC	IoT Healthcare Services
79	COBAN *et al.* (2018)	Solution Proposal	Model	Architecture	PoC	Big Data
80	MARQUES; PITARMA (2019)	Evaluation Research	Tool	Monitoring System	Case Study	Elderly Healthcare
81	ELMISERY *et al.* (2019)	Validation Research	Method	Approach	Experiment	Security
82	GOPAL *et al.* (2019)	Opinion Paper	Experience	Lessons Learned	Not identified	IoT Healthcare Services
83	CELESTI *et al.* (2019)	Experience Paper	Experience	Lessons Learned	Case Study	IoT Healthcare Services
84	ASIF-UR-RAHMAN *et al.* (2019)	Validation Research	Model	Framework	Simulation	Networks
85	LEE *et al.* (2019b)	Conceptual Proposal	Model	Framework	Experiment	IoT Healthcare Services
86	TUN *et al.* (2021)	Conceptual Proposal	Others	Review	Not identified	Elderly Healthcare
87	AL-TURJMAN *et al.* (2020)	Not identified	Others	Review	Not identified	IoT Healthcare Services
88	MEA *et al.* (2020)	Solution Proposal	Tool	Platform	PoC	IoT Healthcare Services

Source: author.

Table 17 – Primary studies selected (89 to 94)

ID	Citation	Research Type	Contribution	User Profile	Empirical Validation	Our Clustering
89	ELBASANI *et al.* (2020)	Evaluation Research	Method	Approach	Case Study	Elderly Healthcare
90	KOTHA (2020)	Not identified	Others	Recommendations	Not identified	IoT Healthcare Services
91	PAZIENZA *et al.* (2020)	Evaluation Research	Tool	Platform	Case Study	IoT Healthcare Services
92	VIJAYAKUMAR; BHUVANESWARI (2020)	Conceptual Proposal	Model	Framework	PoC	IoT Healthcare Services
93	HUYNH *et al.* (2020)	Evaluation Research	Model	Architecture	Case Study	Elderly Healthcare
94	REN *et al.* (2020)	Validation Research	Tool	Monitoring System	Experiment	IoT Healthcare Services

Source: author.

Table 18 – Description of the tools found.

ID	Citation	Name	Summary
2	CALHOUN *et al.* (2012)	Not defined	A new Body sensor networks hardware platform that integrates novel circuit designs and cutting-edge technologies to reduce the cost of communication and computation by several orders of magnitude.
7	RAHMANI *et al.* (2015)	UT-GATE	Provides efficient local services for health monitoring applications such as local repository, compression, signal processing, data standardization, WebSocket server, protocol translation and tunneling, firewall, and data mining and notification.
13	MCRAE *et al.* (2016)	p-BNC	A "platform to digitize biology" in which small quantities of patient sample generate an immunofluorescent signal on agarose bead sensors that is optically extracted and converted to antigen concentrations. The platform comprises disposable microfluidic cartridges, a portable analyzer, automated data analysis software, and intuitive mobile health interfaces.
32	ENSHAEIFAR *et al.* (2018)	TIHM	Combine environmental and physiological data to learn and discover changes in patient's health and well-being.
36	GATOUILLAT *et al.* (2018)	REC Heart	A wearable cardiorespiratory sensor that can be easily integrated into a wider-scale healthcare framework.
37	BELESIOTI *et al.* (2018)	VICINITY	Aims to redesign modern healthcare services with promising technological, economic, and social prospects. The proposed services can have positive results like improved everyday life with increased quality.
39	REDA *et al.* (2018)	Not defined	An eHealth ontology-based Semantic Web system that transforms low-level data obtained from IoT fitness and wellness devices into enriched information model encoded using Resource Description Framework and Web Ontology Language.
44	SHAFI *et al.* (2018)	IoT-based water quality monitoring system	An IoT-based solution to monitor the water quality in real-time. The system provides remote water quality assessment monitoring and water flow control via a mobile app.
45	CARBONARO *et al.* (2018)	IoT Fitness Ontology (IFO)	An ontology-based system for the eHealth domain. It provides semantic interoperability among heterogeneous IoT fitness and wellness devices and facilitates data integration and sharing.
50	YAO *et al.* (2018)	WITS	An end-to-end solution to facilitate the development of smart home applications.
51	LIU *et al.* (2018)	Active Plant Wall	A remote monitoring and management solution specific to a plant wall system based on the Azure public cloud platform and is aimed at contributing to indoor climate monitoring and control in public or private buildings.
56	ONASANYA; ELSHAKANKIRI (2019)	Not defined	An IoT-enabled medical system for enhanced treatment, diagnosis, detection, and monitoring of cancer patients based on cancer care services and business analytics/cloud services, where the business analytics/cloud services constitute enablers for actionable insights; decision making; data transmission; and reporting.
58	KHAREL *et al.* (2019)	FC-based smart health monitoring	System can be promising for changing the clinic-centric health system to a smart patient-centric health system and for providing seamless health services to all.
59	DONATI *et al.* (2019)	EasyCare	A telemedicine platform for data acquisition, distribution, processing, presentation, and storage, aimed to remotely monitor chronic patients' clinical status.
63	BAJENARU; CUSTURA (2019)	MONISAN	A platform that focuses on remote monitoring of physical and environmental parameters and general assessment of public health.
64	RAFFERTY *et al.* (2019)	Thermal Vision for Fall Detection	An ensemble of thermal vision-based, big data facilitated solutions which aim fall detection.
69	CHUI *et al.* (2019)	Not defined	A big data and IoT-based patient behavior monitoring system
73	TEWELL *et al.* (2019)	SCAMPI sensor toolkit	A toolkit comprised of off-the-shelf, affordable sensors to allow persons with dementia and Parkinson's disease to monitor meaningful activities as well as activities of daily living in order to self-manage their life and well-being.
75	DOBRE *et al.* (2019)	vINCI	System aims to provide monitoring and non-intrusive care to the elderly in order to improve their Quality of life (QoL) and well-being by offering modern and efficient solutions to solve their problems.
76	ONASANYA *et al.* (2019)	Smart Saskatchewan Healthcare	System based on IoT technology in the context of four services, namely: business analytics and cloud services, cancer care services, emergency services, and operational services.
77	CHONG *et al.* (2019)	eMeD	Provide an autonomous wearable device that can be used in Internet of Medical Things (IoMT) application. / Introduced as a self-sustainable wearable device that sense, process, and transmit vital sign data via ZigBee.
80	MARQUES; PITARMA (2019)	AirPlus	A real-time indoor environmental quality monitoring system
88	MEA *et al.* (2020)	PollicIoT	A complete system for outdoor patient localization from the hardware to the management platform.
91	PAZIENZA *et al.* (2020)	eLifeCare	Which is enhanced by the In-Edge computation of AI-based techniques to perform reliable decision-making activities in a high complexity scenario such as the healthcare domiciliary hospitalization in an AAL fashion.
94	REN *et al.* (2020)	LiquidSense	A liquid level sensing system that is low-cost, high accuracy, widely applicable to different daily liquids and containers, and can be easily integrated with existing smart home networks.

Source: author.

APPENDIX C – PREDICTORS (FEATURES) DETAILING

Tables 19 and 20 detail the features used to train the Machine Learning models during the case study. With this detailing, it is possible to understand what data was collected and how they are represented.

Table 19 – Features used in the case study (part I).

#	Feature	Short description	Data type
1	height	Participant height	Float
2	weight	Participant weight	Float
3	steps	Daily steps	Integer
4	calories	Daily calories	Float
5	lightsleep	Light sleep time	Integer
6	deepsleep	Deep sleep time	Integer
7	remsleep	REM sleep time	Integer
8	awakesleep	Awake time during last sleep session	Integer
9	incomingcalls	Number of incoming calls	Integer
10	rejectedcalls	Number of rejected calls	Integer
11	blockedcalls	Number of blocked calls	Integer
12	missedcalls	Number of missed calls	Integer
13	outgoingcalls	Number of outgoing calls	Integer
14	incomingcallsaverageduration	Average duration of incoming calls	Float
15	outgoingcallsaverageduration	Average duration of outgoing calls	Float
16	differentlocations	Number of locations visited (50 meters apart)	Integer
17	differentwifi	Daily number of WiFi networks connected	Integer
18	whatsappinvoice	Number of WhatsApp incoming voice call	Integer
19	whatsappinvideo	Number of WhatsApp incoming video call	Integer
20	whatsappoutvoice	Number of WhatsApp outgoing voice call	Integer
21	whatsappoutvideo	Number of WhatsApp outgoing video call	Integer
22	whatsappnotification	Number of WhatsApp notifications	Integer
23	mean_nni	The mean of RR-intervals	Float
24	sdnn	The standard deviation of the time interval between successive normal heart beats (i.e. the RR-intervals)	Float
25	sdsd	The standard deviation of differences between adjacent RR-intervals	Float
26	nni_50	Number of interval differences of successive RR-intervals greater than 50 ms.	Float
27	pnni_50	The proportion derived by dividing nni_50 (The number of interval differences of successive RR-intervals greater than 50 ms) by the total number of RR-intervals.	Integer
28	nni_20	Number of interval differences of successive RR-intervals greater than 20 ms.	Float
29	pnni_20	The proportion derived by dividing nni_20 (The number of interval differences of successive RR-intervals greater than 20 ms) by the total number of RR-intervals.	Integer
30	rmssd	The square root of the mean of the sum of the squares of differences between adjacent NN-intervals. Reflects high frequency (fast or parasympathetic) influences on hrV	Float
31	median_nni	Median Absolute values of the successive differences between the RR-intervals.	Float
32	range_nni	Difference between the maximum and minimum nn_interval.	Integer
33	cvsd	Coefficient of variation of successive differences equal to the rmssd divided by mean_nni.	Float
34	cvnni	Coefficient of variation equal to the ratio of sdnn divided by mean_nni	Float
35	mean_hr	Average of Heart Rate	Float
36	max_hr	Maximum of Heart Rate	Integer
37	min_hr	Minimum of Heart Rate	Integer
38	std_hr	Standard Deviation of Heart Rate	Float
39	specificage	Age	Integer
40	group	Study Group	Categorical (Initial Set, UFPI, UFC)
41	gender	Participant gender	Categorical (1 for male and 0 for female)
42	income	Participant income	Categorical (1 for 0 to 1 considering minimum wage; 2 for 2 to 4; 3 for 5 to 7; 4 for 8 to 10; and 5 for more than 10)

Source: author.

Table 20 – Features used in the case study (part II).

#	Feature	Short description	Data type
43	children	Number of children	Categorical (1 for none; 2 for 1 or 2; 3 for 3 or 4; 4 for more than 4)
44	edulevel	Participant educational level	Categorical (1 for primary level; 2 for secondary level; 3 for university level; and 4 for postgraduate level)
45	familyarr	Number of people living with the participant	Categorical (1 for lives alone; 2 for lives with more 1 or 2; 3 for lives with more 3 or 4; 4 for lives with 5 or more)
46	residence	Number of WhatsApp outgoing voice call	Categorical (1 for urban area and 0 for rural area)
47	social	Daily time spent on apps from category: social	Integer
48	lifestyle	Daily time spent on apps from category: lifestyle	Integer
49	communication	Daily time spent on apps from category: communication	Integer
50	entertainment	Daily time spent on apps from category: entertainment	Integer
51	music	Daily time spent on apps from category: music	Integer
52	photography	Daily time spent on apps from category: photography	Integer
53	finance	Daily time spent on apps from category: finance	Integer
54	video	Daily time spent on apps from category: video	Integer
55	health	Daily time spent on apps from category: health	Integer
56	productivity	Daily time spent on apps from category: productivity	Integer
57	business	Daily time spent on apps from category: business	Integer
58	maps	Daily time spent on apps from category: maps	Integer
59	shopping	Daily time spent on apps from category: shopping	Integer
60	travel	Daily time spent on apps from category: travel	Integer
61	food	Daily time spent on apps from category: food	Integer
62	vehicles	Daily time spent on apps from category: vehicles	Integer
63	sports	Daily time spent on apps from category: sports	Integer
64	news	Daily time spent on apps from category: news	Integer
65	games	Daily time spent on apps from category: games	Integer
66	art	Daily time spent on apps from category: art	Integer
67	education	Daily time spent on apps from category: education	Integer
68	events	Daily time spent on apps from category: events	Integer
69	educational	Daily time spent on apps from category: educational	Integer
70	weather	Daily time spent on apps from category: weather	Integer
71	books	Daily time spent on apps from category: books	Integer
72	beauty	Daily time spent on apps from category: beauty	Integer
73	house	Daily time spent on apps from category: house	Integer
74	dating	Daily time spent on apps from category: dating	Integer
75	walking	Daily time spent on apps from category: walking	Integer
76	other	Daily time spent on apps from category: other	Integer
77	invehicle	Daily time performing the activity: invehicle	Integer
78	running	Daily time performing the activity: running	Integer
79	biking	Daily time performing the activity: biking	Integer
80	running(treadmill)	Daily time performing the activity: running(treadmill)	Integer
81	strengthtraining	Daily time performing the activity: strengthtraining	Integer
82	profession_fulltimeworker	True for full time worker (participant)	Integer (0 for false and 1 for true)
83	profession_parttimeworker	True for part-time worker (participant)	Integer (0 for false and 1 for true)
84	profession_selfemployed	True for self employed worker (participant)	Integer (0 for false and 1 for true)
85	profession_student	True for students (participant)	Integer (0 for false and 1 for true)
86	maritalstatus_married	True for participants married	Integer (0 for false and 1 for true)
87	maritalstatus_single	True for participants single	Integer (0 for false and 1 for true)
88	phy_ref_score or psy_ref_score	Reference score for physical or psychological domain depending on the selected dataset	Float

Source: author.

APPENDIX D – TERMS RELATED TO THE HEALFUL PLATFORM

This appendix provides a list of terms related to the Healful platform.

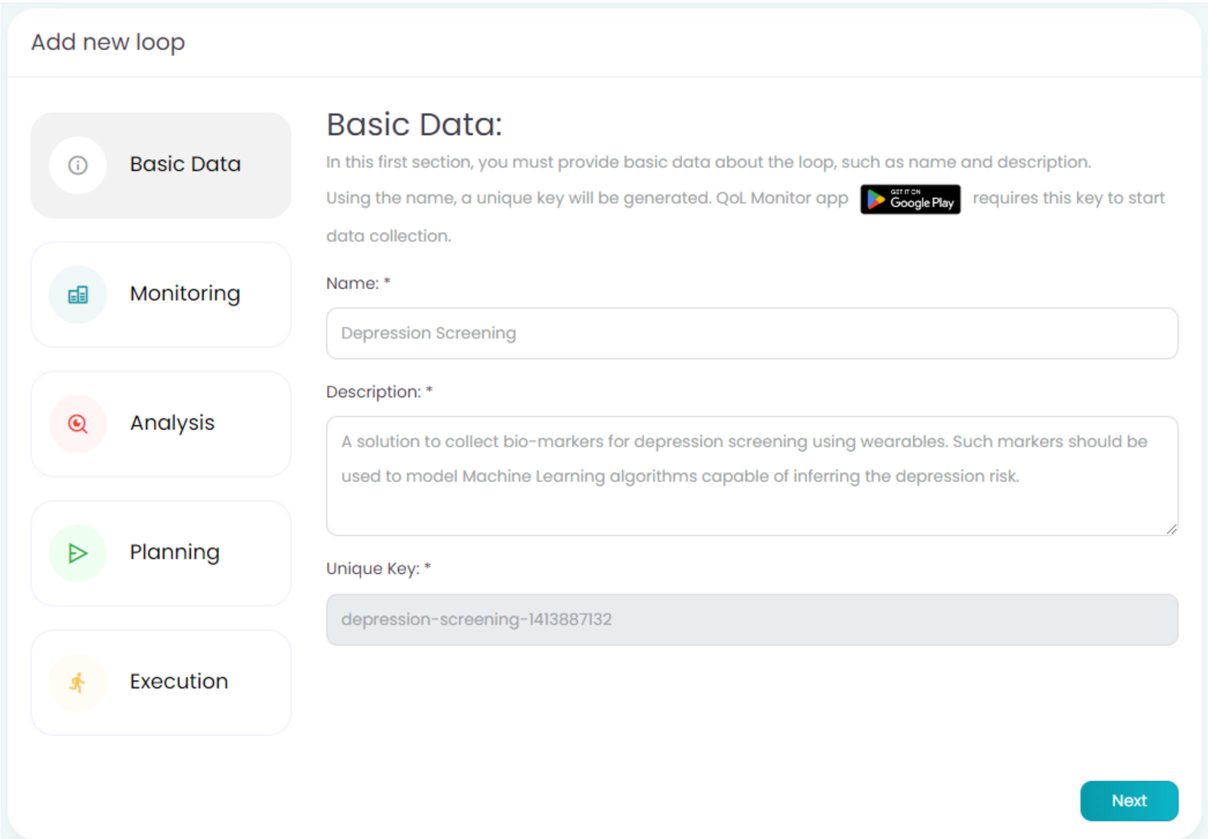
- **Healful platform**: name of the platform presented in this thesis to support the development of solutions to monitor users' Quality of Life. This name was defined from the acronym HEALth THings platForm for qUality of Life.
  Link: https://healful.life.
- **QoL Monitor**: mobile app developed to aid IoHT data collection. In addition, it integrates with the Google Fit platform to deal with the challenge of the high heterogeneity of wearable devices.
  Link: https://www.qol-monitor.com
- **Athena**: web system developed to assist the construction of computational intelligence systems. Athena abstracts algorithms into modules that can be interconnected to process data and model intelligent algorithms.
  Link: https://athenasystem.com.br
- **Google Fit**: it is a health tracking platform developed by Google for the Android operating system, Wear OS, and Apple iOS.
  Link: https://www.google.com/fit
- **MAPE-K loop**: adaptation loop proposed by IBM, which has four steps: monitor, analyze, plan, and execute. All steps are supported by the knowledge that is stored over the iterations (IBM, 2005).

APPENDIX E – HEALFUL RUNNING EXAMPLE

This appendix presents a running example of the Healful platform based on a scenario inspired by the studies of RYKOV *et al.* (2021) and SAEB *et al.* (2017). This scenario considers the need to collect bio-markers for depression screening using wearables. Such markers should be used to model Machine Learning algorithms capable of inferring the depression risk. In addition, it should be possible to define interventions to be applied when the risk is high.

According to WHO, depression is one of the main factors of disability worldwide (RODRIGUES *et al.*, 2022) with a severe negative impact on the global economy of US$16.3 trillion between 2011 and 2030 (WHO, 2021). Solutions like those mentioned earlier are valuable and have aroused a high interest (ZHOU *et al.*, 2022). However, developing IoHT solutions that involve data collection, data analysis, building intelligent models, and planning interventions based on context monitoring is a complex task (OLIVEIRA *et al.*, 2022a). Thus, the Healful platform seeks to attenuate the difficulties of each phase in addition to supporting QoL-based IoHT solutions.

Figure 50 – Step-by-step wizard to create a new adaption loop in the Healthful platform

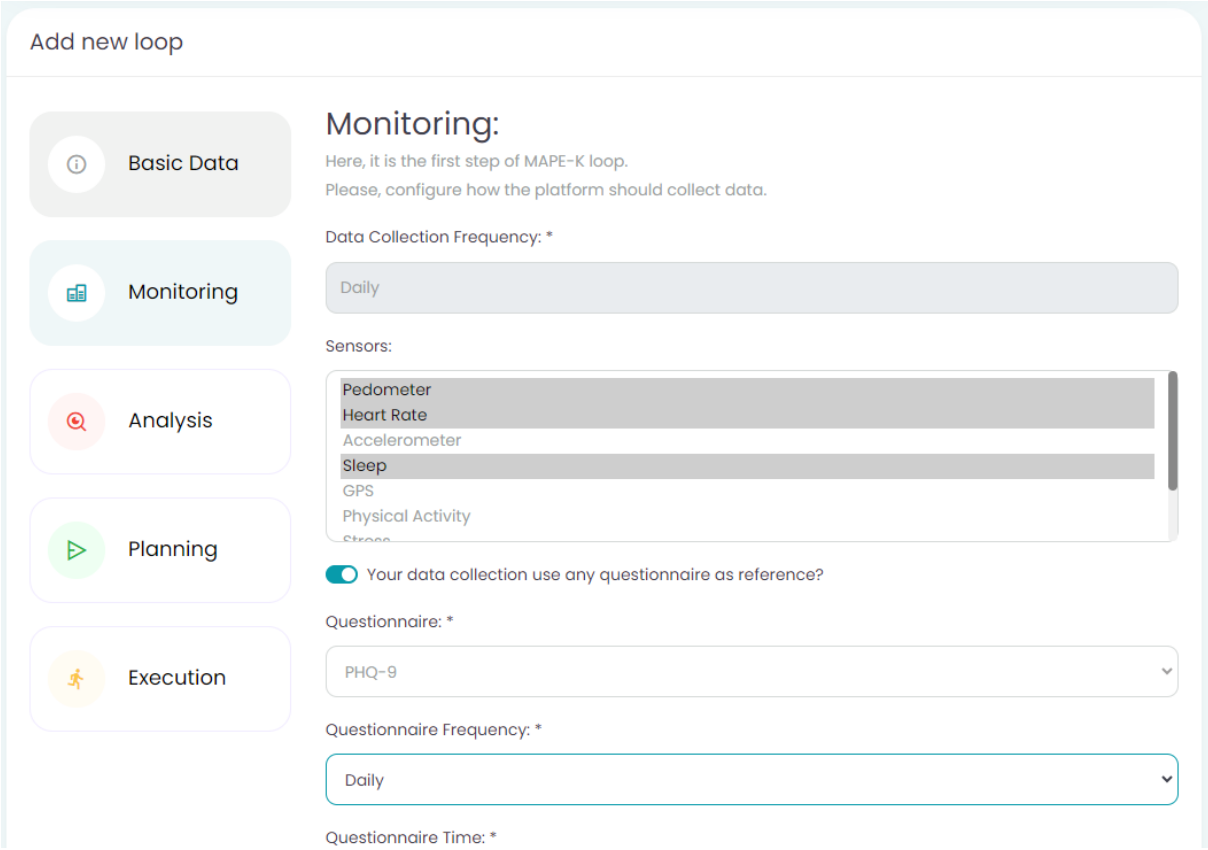


Source: author.

When using the Healful platform, the user can create a new adaptation loop. This loop is the platform primary abstraction, and its attributes are defined through a step-by-step wizard. The first step (Figure 50) requires a name, a description, and a unique key to link collected data. In the second step (Figure 51), the user must inform the frequency of data collection[1], the sensors and, if the collection uses a questionnaire as a reference, it must also inform the name and frequency of the questionnaire. In the current version of the platform, the questionnaires are registered in an electronic spreadsheet, and inserting new ones requires manual entry of the questions and possible answers. However, such management can be included within the platform in future releases.

As proposed by RYKOV *et al.* (2021), it is possible to identify the depression risk based on the daily steps, heart rate, and time spent in each sleep stage. Therefore, as presented in Figure 51, the respective sensors (pedometer, heart rate sensor, and sleep sensor) were selected to collect these data. Also, as the baseline to train the Machine Learning models, the authors recommend the PHQ-9 questionnaire (KROENKE *et al.*, 2001).

Figure 51 – Setting up monitoring for depression risk monitoring

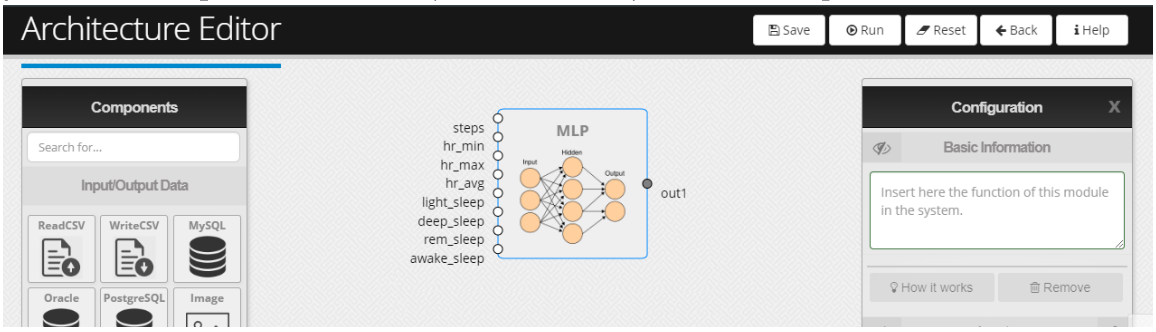


Source: author.

[1] Currently, the platform only allows daily data collection.

In the third step, the Athena system responsible for processing the data and building the intelligent model must be selected. This system must be created in the Athena tool using the built-in modules. Figure 52 presents a module arrangement that uses the raw data to train a Multilayer Perceptron (MLP) to classify depression risk.

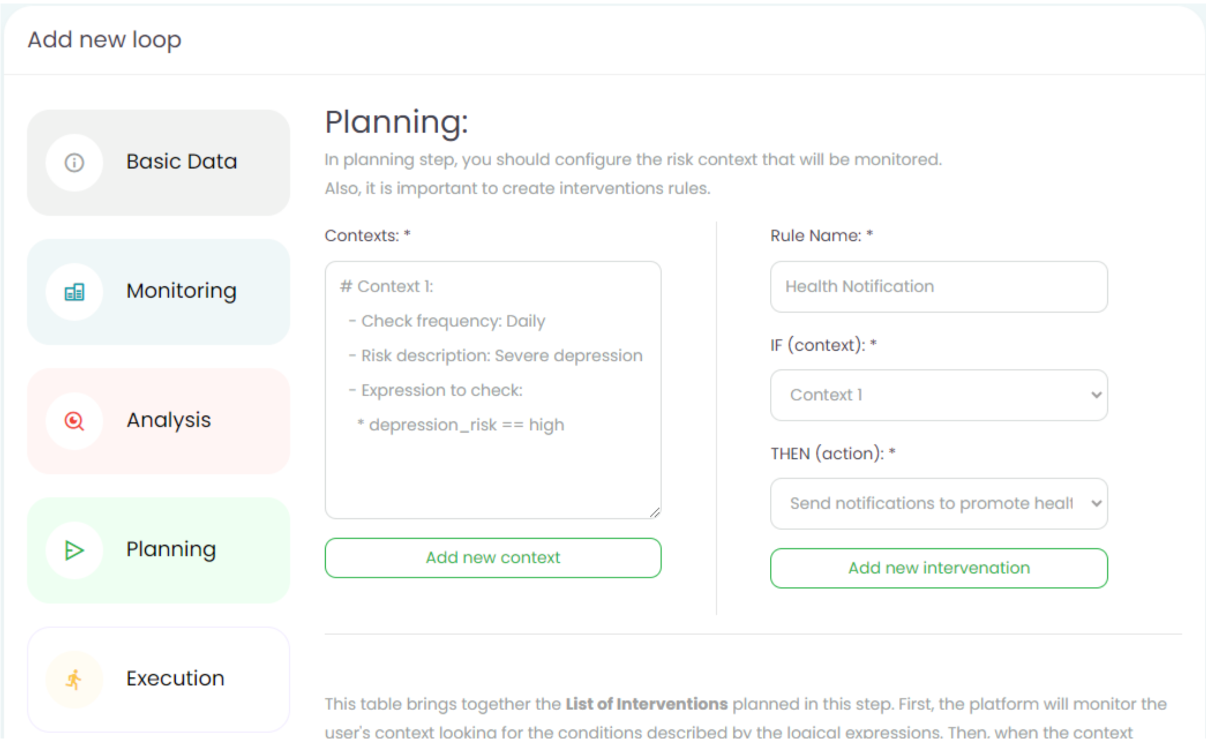
Figure 52 – Example of an Athena system to classify the risk of depression



Source: author.

It is worth mentioning that the current version of Healful requires the intelligent model to be previously trained in Athena. However, in future work, it is possible to create a structure for online training in Athena using the data collected by Healful.

Figure 53 – Screen to specify the risk contexts and healthcare interventions



Source: author.

In the fourth step (Figure 53), healthcare specialists can plan health interventions.

Thus, it is possible to define the risk contexts and which actions should be performed in each context. For this task, the platform has a grammar of mathematical and logical expressions, in which it is possible to specify variables to be monitored and the period. In addition, the healthcare specialist can inform the risks associated with each context. For example, in Figure 53, there is a context in which the depression-risk variable (achieved using the MLP model) is high. If this scenario persists, the patient can evolve into severe depression. In Figure 53, it is also possible to observe the actions (healthcare interventions) defined for the previously described context. In this case, periodic notifications will be sent to users.

Finally, the user should review all the definitions in the fifth step and confirm the creation of the loop. However, users must install the QoL Monitor app and provide the unique study key to start data collection. Thus, the data acquired from the selected sensors will be sent to the Healful database, and the platform will execute the following stages. In order to bring more details about Healful usage, there is a video presenting the whole process discussed in this subsection (link: github.com/great-ufc/healful-thesis) and the appendix D bring a glossary of terms related to the platform.