FH Vorarlberg
University of Applied Sciences

Bringing Context to Just-in-Time Adaptive Interventions (JITAIs) based on Stress Monitoring Using HRV

Master Thesis
Submitted in Fulfillment of the Degree

Master of Science in Engineering (MSc)

University of Applied Sciences Vorarlberg
Computer Science Master

Supervised by
Dr. Katrin Paldán

Submitted by
Matthias Lerch, BSc
Dornbirn, July 2022

Abstract

Bringing Context to Just-in-Time Adaptive Interventions (JITAIs) based on Stress Monitoring Using HRV

With the rise of people wearing smartwatches and the ever-lasting issue of stress, there has been an interest in detecting stress with wearables in real-time. This allows for interventions that take place exactly when stress occurs. However, many situations require all of our attention, making them unsuitable for any interventions. Additionally, many approaches currently do not factor in this aspect, running the risk of offering users undesirable interventions.

This thesis examines how contextual user information can be incorporated into a stress intervention system to reduce undesirable intervention timings. The system is split into detecting stress using heart rate variability (HRV) metrics obtained from a photoplethysmography (PPG) signal, and inferring user context from available sensor data. It is evaluated with a simulation-based approach using daily schedules of created personas and randomly sampled stressors during daily life.

The results obtained indicate the benefit of adding contextual user information to a stress intervention system. Depending on the busyness of the schedule, it can greatly decrease the number of received interventions. However, as these findings are attained without performing a user testing, it is unclear how they compare to results from real-world usage.

Kurzreferat

Hinzufügen von Kontext für just-in-time adaptive interventions (JITAIs) basierend auf Stressüberwachung mit HRV

Da immer mehr Menschen Smartwatches tragen und Stress ein ständiges Thema ist, besteht ein Interesse daran, Stress mit Wearables in Echtzeit zu erkennen. Dies ermöglicht Interventionen, die genau dann erfolgen, wenn Stress auftritt. Viele Situationen erfordern jedoch unsere ganze Aufmerksamkeit, sodass sie für Interventionen ungeeignet sind. Außerdem berücksichtigen viele Ansätze diesen Aspekt derzeit nicht, sodass die Gefahr besteht, dass den Nutzer*innen unerwünschte Interventionen angeboten werden.

In dieser Arbeit wird untersucht, wie kontextbezogene Informationen über Benutzer*innen in ein Stress-Interventionssystem integriert werden können, um unerwünschte Interventionszeitpunkte zu reduzieren. Das System ist aufgeteilt in die Erkennung von Stress mit Hilfe von Metriken der Herzfrequenzvariabilität, die aus einem PPG-Signal gewonnen werden, und die Ableitung des Benutzer*innenkontextes aus verfügbaren Sensordaten. Es wird mit einem simulationsbasierten Ansatz evaluiert, bei dem Tagespläne von erstellten Personas und zufällig ausgewählte Stressoren des täglichen Lebens verwendet werden.

Die erzielten Ergebnisse weisen auf den Nutzen der Hinzufügung von kontextbezogenen Informationen der Benutzer*innen für ein Stressinterventionssystem hin. Je nach Auslastung des Zeitplans kann dies die Anzahl der erhaltenen Interventionen stark verringern. Da diese Ergebnisse jedoch ohne reale Testpersonen erzielt wurden, ist unklar, wie sie sich mit den Ergebnissen aus der realen Nutzung vergleichen lassen.

Contents

List of Figures

List of Tables

List of abbreviations

ANS autonomic nervous system

DISE Daily Inventory of Stressful Events

ECG electrocardiogram

EMI ecological momentary intervention

HF high frequency

HRV heart rate variability

HR heart rate

JITAI just-in-time adaptive intervention

LF low frequency

NN normal-to-normal

PNS parasympathetic nervous system

PPG photoplethysmography

PSS Perceived Stress Scale

VASS Stress Visual Analogue Scale

PSD Power Spectral Density

PT persuasive technology

PTSD post-traumatic stress disorder

SNS sympathetic nervous system

ULF ultra-low frequency

VLF very-low frequency

WHO World Health Organisation

1. Introduction

Health is arguably one of the most important facets of life and is influenced by many different factors. It is important to note that health does not just mean being free of disease. The World Health Organisation (WHO) defines health as physical, mental, social, and spiritual well-being (Svalastog et al., 2017). Being healthy is not only an individual's desire, but also a societal one. Ignoring the personal motivations for being healthy, healthcare itself is by no means cheap. In the EU in 2019 alone, healthcare expenditures were 1.38 billion euros, with 41.4 million coming from Austria. Additionally, these numbers increased for all countries in the EU from 2012 to 2019, except for Greece ("Healthcare expenditure statistics", n.d.).

One important aspect of health, that affects everybody, is stress. Chronic stress has been associated with a higher risk and mortality of cardiovascular diseases (Fishta & Backé, 2015), it affects the ageing process of cells and is suggested to be a risk factor for specific cancers (Kruk et al., 2019), and is associated with Alzheimer's disease and diabetes (Aguiló et al., 2015). Additionally, chronic stress is related to many psychological diseases like schizophrenia (Aguiló et al., 2015), post-traumatic stress disorder (PTSD), burnout, atypical depression, and chronic fatigue disease (Marin et al., 2011).

As so many people are in possession of mobile phones nowadays, the interest for mHealth applications is very high. A mHealth application is used to "capture, analyze, process, and transmit health-based information from sensors and other biomedical systems" (Adibi, 2015, p.1). They are often used in conjunction with eHealth, which are healthcare practises that are assisted by electronic processes and communication systems (Adibi, 2015). According to Rowland et al. (2020), 26% of physicians were asked about mHealth applications by their patients and 56% have discussed them with patients.

Therefore, it is not astounding that a rather new trend has emerged: *Quantify yourself*. This movement is all about measuring indicators for their personal well-being with smart devices, like smartphones and smartwatches. These indicators include sleep duration and quality, step counters, heart rate and variability, and stress (Massa et al., 2017). With the help of a multitude of apps, users can nowadays also be prompted when certain conditions arise. One of these conditions is stress, that can be recognized with a PPG sensor of a smartwatch. However, stress is by far not bad in all situations, and not every situation allows for time and space for an intervention. Thus, determining whether an intervention is beneficial in a specific situation is of great importance, to avoid unfavourable intervention timings.

This leads to the research question of this thesis:
How can contextual information be integrated to reduce unfavourable timed stress-related interventions for just-in-time adaptive interventions (JITAIs)?

2. Heart rate variability (HRV) and Stress

This chapter explains what HRV is, what parameters can be computed from it, what stress is and its relationship with HRV.

2.1. HRV

Heart rate variability (HRV) describes the time change between heartbeats (also referred to as RR intervals), which is controlled by the autonomic nervous system (ANS). The ANS is divided into two parts, the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). The SNS is responsible for the "fight or flight" response and is active during stressful situations. The PNS is responsible for "rest and digest" and is active in relaxed scenarios. An increase in heart rate (HR) is generally related to higher activity in the SNS, while a decrease is related to the PNS. (Pham et al., 2021) In healthy individuals, the two parts work together and influence each other. Disturbances, especially with an increase in sympathetic activity, are associated with cardiovascular and some mental diseases. (Cygankiewicz & Zareba, 2013) Additionally, HRV can also be used as an indicator of acute stress, which is discussed in chapter 2.2. (Castaldo et al., 2015; Kim et al., 2018; Munla et al., 2015)

Once HRV is measured, e.g., with an electrocardiogram (ECG), there are multiple features that can be computed. They can be categorized into time-domain and frequency-domain methods, as well as nonlinear methods. Table 2.1 visualizes the different measures for time-domain and frequency-domain HRV analysis, which are the most common indices for HRV. (Pham et al., 2021)

2.1.1. Time-based analysis

Time-based HRV analysis is calculated by either working with the entire sequence of normal-to-normal (NN) intervals of the whole recording, or with

HRV indices		
Variable	Units	Description
Time-domain analysis		
SDNN	ms	Standard deviation of all normal-to-normal (NN) intervals
SDANN	ms	Standard deviation of average NN intervals in all 5 minute segments
SDNN Index	ms	Mean of the standard deviations of NN intervals in 5 min segments
RMSSD	ms	Root mean square of adjacent NN intervals
pnn50	%	Percentage difference between adjacent NN intervals greater than 50 ms
Frequency-domain analysis		
ULF	ms^2	Ultra-low frequency <0.003 Hz
VLF	ms^2	Very-low frequency <0.003-0.04 Hz
LF	ms^2	Low frequency power 0.04-0.15 Hz
HF	ms^2	High frequency power 0.15-0.4 Hz
LF/HF	Ratio	Ratio of low-high frequency power

Table 2.1.: Summary of HRV indices; adapted from (Cygankiewicz & Zareba, 2013, p.381; Pham et al., 2021, p.7)

consecutive NN intervals. Then, statistical metrics can be calculated, which lead to the different HRV indices. An example of this is the SDNN, which is simply the standard deviation of all NN intervals recorded. Here, the index is computed directly from the entire NN sequence, while other metrics like the RMSSD (root-mean-square of consecutive differences) are calculated from successive intervals.
Different HRV parameters all have different meanings and interpretations, the SDNN for instance is considered to represent the overall heart variability, whereas the RMSSD and pNN50 indicate parasympathetic heart modulation. Additionally, indices derived from sequential NN intervals mostly indicate short-term variations and can often be used with shorter recording times. However, there are no indices that represent sympathetic heart modulation in time-based analysis. (Cygankiewicz & Zareba, 2013; Pham et al., 2021)

2.1.2. Frequency-based analysis

Heart rate – and therefore also its variability – is determined by different systems in the body, each system using different frequencies. Thus, frequency-based analysis can help evaluate specific components better than time-based analysis. The individual signals can be grouped into four different frequency bands:

- ultra-low frequency (ULF): ≤ 0.003 Hz
- very-low frequency (VLF): $\leq 0.003 - 0.04$ Hz
- low frequency (LF): $\leq 0.04 - 0.15$ Hz
- high frequency (HF): $\leq 0.15 - 0.4$ Hz

Generally speaking, when it comes to the autonomic nervous system (ANS), HF is associated with activity of the parasympathetic nervous system (PNS), and LF with activity of the sympathetic nervous system (SNS). Additionally, the ratio between the two activities (LF/HF) is often used to compare the activity of the two systems with each other. (Pham et al., 2021)

However, one limitation of frequency-based analysis is the fact that different spectral analysis algorithm can output different values. This makes it difficult to compare values obtained from different studies and examinations. And even if the same algorithms are chosen, some of them still have parameters that can differ between studies. (Pham et al., 2021)

2.2. Stress

Stress is part of human nature and has peaked the interest of scientists for a long time. In 1950, it was defined as "the non-specific response of the body to any noxious stimulus" (Koolhaas et al., 2011, p.1292). Then, the concepts of *stressors* and *stress responses* were introduced. A stressor is a situation that threatens *homeostasis*, and a stress response is the action that tries to return to the state of homeostasis (Koolhaas et al., 2011). Homeostasis can be seen as the body's internal balance. Each bodily function has its own narrow range of set-points and tries not to deviate too far from these set-points. Homeostasis is only possible with observations and corrections. Control systems identify disruptions in homeostasis (like body temperature and blood sugar levels) and try to mitigate them. (Libretti & Puckett, 2022)

An addition to this theory is the adaptiveness of the stress response, which is described by the General Adaption Syndrome (GAS). In essence, organisms are usually able to adapt to stressors and return to homeostasis. However, when stressors are present for a longer period of time, this adaption might fail. This distinction of adaptiveness is also part of the principles behind *eustress* and *distress*. However, it can sometimes be difficult to distinguish the two of them, especially in studies where it could lead to some form of interpretation bias. Therefore, another approach is looking at the unpredictability and controllability of a stimulus. As the names suggest, controllability looks at how much influence a subject has on the situation, whereas predictability defines how well a scenario can be anticipated. When it comes to stressors, negative consequences, such as pathology, do not occur due to the stimulus' physical nature. Rather, it is the level of predictability and controllability that causes these outcomes. Specifically, in humans, it is the *perceived* control that is of importance here. (Koolhaas et al., 2011)

Additionally, stressors have another dimension besides predictability and controllability: intensity. This intensity can also be seen as the possible consequences for one's life, and ranges from almost non-existing to life-threatening. Once more, this intensity is graded by the individual's perception and is very likely to add to the severity of the outcome. For example, traumatic events cause post-traumatic stress disorder (PTSD) in around 20-30% of humans, although the predictability and controllability of these situations might objectively be very similar. (Koolhaas et al., 2011)

Another principle is the differentiation between *regulatory range* and *adaptive capacity*. Regulatory range describes the theoretically possible range of situations that a healthy organism can deal with. The adaptive capacity refers to the range of responses in behaviour and physiology to cope with the environment. This principle can be clarified by looking at an animal that is living in an area that is getting colder and has food shortage. The regulatory range would include a range of temperatures where the animal can live without freezing to death or dying of heat. The adaptive capacity consists of behaviours like leaving the environment in search of a place with more food options and warmer temperatures, food hoarding, and building nests. Additionally, depending on the species, the adaptive capacity could also include physiological responses, such as lowering the set-point of body temperature.

However, the adaptive capacity can be lowered by environmental factors, such as droughts leading to less food availability, which results in less fat tissue. This leads to a reduced ability to cope with the cold, although the temperature still lies within the normal regulatory range. Ultimately, temperature as a stimulus might now be seen as a stressor and a stress response might occur.
Following this principle, a stressor can either influence the regulatory range or the adaptive capacity. A reduction in adaptive capacity like in the example above entails that a stimulus, which was formerly not identified as a stressor, is now perceived as one. Additionally, alteration in the regulatory range implies that situations previously viewed as stressors are now controllable and predictable and no longer count as stressors. (Koolhaas et al., 2011)

2.2.1. Measuring stress

Stress can be measured by looking at different groups of variables. The two most important physiological groups are electrophysiological, and biochemical variables. Additionally, stress can be measured with psychometric tests.
Electrophysiological parameters include blood pressure, skin conductance response, brainwaves, skin temperature, blood volume pulse, heart rate, and HRV. Biochemical parameters include levels of cortisol, α-amylase, copeptin and prolactin. Some of these can be measured with saliva samples, such as cortisol and α-amylase, whilst others require blood samples (Massa et al., 2017).

Psychometric tests include the Perceived Stress Scale (PSS), and Stress Visual Analogue Scale (VASS). The PSS assesses the general stress level in a subject's everyday life. VASS measures how stressful a subject perceives a specifically given situation (Massa et al., 2017).

2.2.2. Stress-related smartwatch apps

When searching for "stress", "stress hrv", and "hrv" in Google Play Store (filtered to show only apps that support smartwatches), the following types of apps are shown: puzzle games for stress reduction, and several health/fitness and lifestyle apps and trackers. In the latter category, the most popular applications are *"Calm – Meditate, Sleep, Relax"*, *"Stila | Stress Tracking and Monitoring"*, *"Cardiogram: Heart Rate Monitor"*, and *"Cardiograph – Heart Rate Meter"*. *Calm* is an app for sleep and meditation and does not detect stress. *Cardiograph* has one functionality – it displays the current heart rate when the app is active. *Cardiogram* also measures heart rate, with the added functionality of tracking it while running in the background and creating charts. However, no stress or HRV is measured. *Stila* is still in early access and currently supports WearOS smartwatches and FitBit PurePulse wristbands. It is part of a research project at the Institute for Informatics at Ludwig-Maximilian University of Munich with the goal of providing "students and professionals with personalized recommendations aiming at improving their learning performances" ("Stila - About", n.d.). It also uses HRV to measure the stress level, but does currently not offer any interventions when stress is detected. However, in the app itself is a segment of an assistant, which currently has no functionality, but will be available in the upcoming release. This assistant might also be part of an intervention system, but no additional information is available yet.

2.3. Stress evaluation with HRV

When it comes to evaluating stress responses with HRV, there are several factors that need to be considered. HRV can be influenced by several different components, like physiological health and psychological disorders. These factors should be taken into consideration when evaluating HRV clinically. Additionally, many lifestyle habits can also influence HRV like drinking, smoking, physical activity, and medications. However, most studies found that stress lead to reduction of parasympathetic activity and an activation of the sympathetic nervous system, which is represented by a decrease of HF and an increase in LF. This change can also be tracked by examining the LF/HF ratio when frequency-based analysis is performed (Kim et al., 2018). Additionally, when it comes to time-based analysis, the measures RR, RMSSD are decreased when being stressed (Castaldo et al., 2015). The concrete changes in HRV parameters during stress are visualized in table 2.2, which also contains the mean values during rest.

However, one limitation of stress response assessment with HRV is the in-

Differences in HRV during stress				
Parameter	Mean* (Rest)	SD* (Rest)	MD	CI95%
Time-based				
RR	845.00	159.29	−142.20	(-168.9; -115.47)
RMSSD	47.92	27.35	−12.03	(-16.78; -7.28)
pNN50	33.46	19.82	−7.98	(-14.52; -1.45)
Frequency-based				
LF	829.14	556.83	156.1	(157.6; 469.8)
HF	1971.64	1841.88	−359.7	(-559.20;-160.25)
LF/HF	1.02	1.11	0.6	(0.14; 1.08)

MD: mean difference during stress | CI95%: 95% confidence interval
*Weighted average calculated from (Castaldo et al., 2015, p. 374)

Table 2.2.: Change in HRV parameters during stress adapted from (Castaldo et al., 2015, p.376)

ability to differentiate between different forms of stress (Oksman et al., 2016). Depending on the stress model used, this could be eustress and distress or the difference in controllability and predictability. Nevertheless, HRV represents a non-invasive and accurate way of assessing stress (Kim et al., 2018).
Traditionally, the HRV measures are calculated from ECG signals, and is still considered as the gold standard in clinical settings. However, with the introduction of newer technology and wearables, new possibilities have arisen. Ranging from ECG chest straps, to optical PPG sensors that are built into smartwatches, the availability of these sensors is very high. When it comes to how accurately these measures can be obtained from PPG sensors, it depends on the setting and the measures themselves. In one study, Jeyhani et al. (2015) compared time-based analysis HRV measures obtained from ECG sensors with the same measures calculated from PPG sensors. Besides pNN50, which had a relative error of 29.89%, all other parameters (SDNN, RMSSD, SD1, and SD2) registered an error of less an 6%. Other studies also report similar observations. However, this assumption should only be made with resting healthy subjects that do not take part in physical activity, such as exercise (Lin et al., 2014; Pinheiro et al., 2016). Another factor that can influence the accuracy of PPG derived parameters is breathing. One study compared HRV metrics calculated from ECG and PPG data during rest and during breathing tasks, such as taking deep and shallow breaths, both rapidly and slowly. Although

reporting similarities in rest, the parameters obtained during breathing tasks demonstrate a significant difference (Jan et al., 2019).

3. State-of-the-art

The following chapter describes the current state of the art and introduces related work as the foundation for this thesis.

3.1. Digital signal processing

Whenever dealing with digital signals, in this case PPG data, digital signal processing techniques can help to reduce noise and extract underlying patterns. This section gives a short introduction to the relevant methods that are later used in the thesis.

3.1.1. Filters

Generally speaking, a filter allows some things to pass, while blocking others. In digital signal processing frequencies are filtered, and there are four major types of digital filters: low-pass, high-pass, band-pass, and band-stop filters. As the name suggests, low-pass filters allow low frequencies to pass and block high frequencies. The opposite is true for high-pass filters. Band-pass filters define a passband – an interval of frequencies that it allows – while blocking the rest. The opposite to that is a band-stop filter, that blocks the defined interval of frequencies and allows all others (Lyons & Fugal, 2014).
Figure 3.1 illustrates the four different basic types of digital filters. In addition to the type of filter and a cut-off frequency (where the transition between passband and stopband is), each of them also has an *order*. The lower the order, the flatter and smoother the cut-off response of the filter. Similarly, with a higher order, the cut-off response will be similar to a "brick-wall" response, the most extreme case that the visualized filters in figure 3.1 display. (Jones et al., 2020a)

3.1.2. Change point detection

In digital signal processing, it is often a necessary task to detect and identify data change points. Generally speaking, a change point is a shift in the underlying model of the data. Besides digital signal processing, it is also often

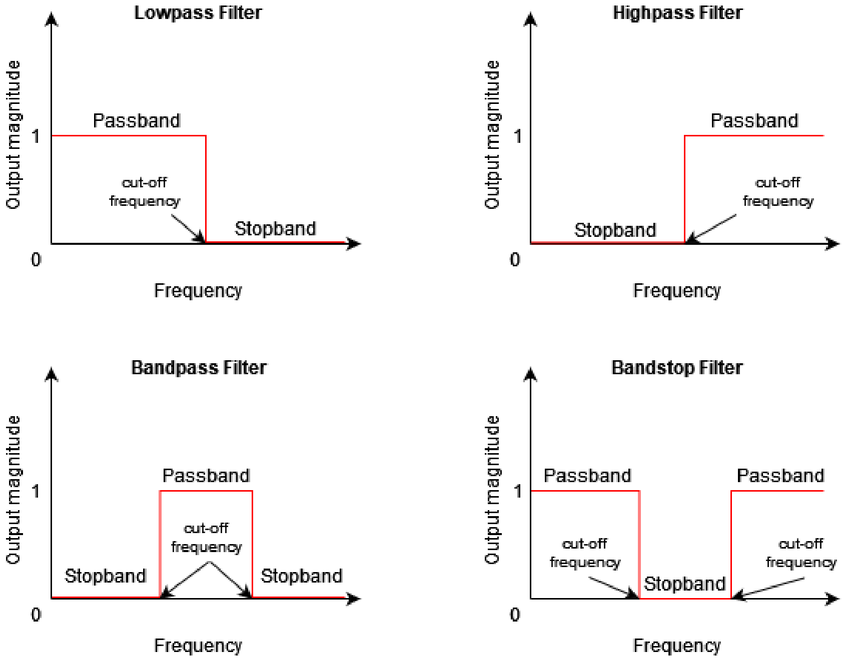

Figure 3.1.: The four basic types of digital filters modified from (Jones et al., 2020b)

used in finance, climatology, bio-informatics, and speech processing. There are two distinctive types of change point detection: online, and offline techniques. The first tries to identify changes in real time and is often called *anomaly or event detection*, whereas the second looks at the data only once all samples are present. It is also referred to as *signal segmentation*. Additionally, change point detection techniques differ whether the number of change points is known beforehand or not. They can be categorized by:

- **Cost function**: how "homogeneous" (no change points) or "heterogeneous" (at least one change point) a segment of the signal is (Truong et al., 2020)
- **Search method**: the algorithm how this optimization problem is solved (such as window slicing, binary segmentation, and bottom-up segmentation) (Truong et al., 2020)
- **Constraint**: in cases where the number of change points are not known, a penalty is added to the optimization problem. However, if this penalty is too big, not enough change points will be detected. Similarly, if the penalty is too small, too many are found, and even noise might be included in the result (Truong et al., 2020)

3.2. Persuasive Technologies

Persuasive Technologies (PTs) focus on trying to motivate people to adopt and reinforce beneficial behaviours and attitudes, and to mitigate and avoid harmful ones. This process is always done voluntarily, without coercion or misinformation, otherwise the system is no longer considered as PT. Generally, there are two use cases for these systems, preventive health promotion and disease management. The first one focuses on a healthy lifestyle and/or to prevent diseases in the first place. Examples of this are: increasing physical activity, making diets more healthy, and decreasing alcohol consumption or smoking.
The second type of PTs (disease management) focuses on reminding people to go along with their treatment or medication plans, and help them to live with their health conditions. (Orji & Moffatt, 2018)

One major advantage that PTs have over types of media that use persuasion is the adaptability and interactivity with users. Persuasive technology can be adjusted based on (user) data that can help tailor the system to specific needs and situations. A system using PTs that focuses on more physical activity by encouraging walks could find out when a user is more likely or even unwilling to take a stroll, the preferred length of a walk, or suggest a place near the user's location. (Fogg, 2003)

We humans can be very good persuaders. We can sense emotions and the overall mood of other people, understand the social aspects of a situation, and know when to use what forms of persuasions. However, there are some aspects in which PT systems have advantage over human persuaders: (IJsselsteijn et al., 2006)

- Depending on the field used, the anonymity that the users perceive can help PT systems be more effective and help people make changes in their lives more easily. This is especially true in sensitive fields such as psychological issues, sexuality, and substance abuse.

- PTs have more forms of presentation at their disposal than human persuaders. They can use different video, audio, and text formats to convey their messages, combine several modes of presentation, and match the type of presentation used with the user's preferences.

- As these systems often have huge amounts of data to work with, they can offer information that a single human persuader could never know. They can use this data, both past data from the individual user to create trends, goals and statistics, and data from other users to create competitions and recommendations. (Fogg, 2003)

- The replicability and scalability of PT systems allows these persuasive systems to reach a significant number of people all at once.
- The size of PT systems often allows them access to places that are inaccessible for humans (such as clothing or wearables), and are often granted access to places where human persuaders would not be allowed (e.g., bathroom or bedroom). (IJsselsteijn et al., 2006)

Persuasive technology systems can play different roles depending on the use cases. These roles are described by the functional triad, a conceptional framework that describe how these technologies can be classified as: tools, media, and social actors.
Tool-PTs focus on making certain behaviour easier to accomplish for users, some of which they would not be able to accomplish without the use of technology.
Media-PTs focus on conveying messages through different forms of media, which can be further categorized into symbolic and sensory media. Symbolic media use texts, charts, and graphics, while sensory media operate with video and audio.
Social actor-PTs try to create an image of the system being alive. When achieved, users interact with these systems almost as if it were a living being. The idea is that the perception of the system being alive can apply social influence, such as peer pressure. (Fogg, 2003)

3.3. Context

This section introduces what user context is, which characteristics it has, and concepts to infer higher level context.

3.3.1. Definition of context

Before discussing how contextual information of users can be integrated into JITAIs, it is important to define what user context even is. Unfortunately, there does not seem to be one general, universally accepted definition and its meaning is based on the interpretation of researchers across different domains (Pradeep & Krishnamoorthy, 2019). For example, in the study conducted by Oinas-Kukkonen and Harjumaa (2008) the authors differentiate between a *use context* and a *user context*. Here, use context refers to what information is

important for a user in a given situation, whereas the user context focuses on the bigger picture, like motivation, needs, and interest of the user.
Other researchers do not distinguish between a use and user context. Additionally, depending on the field of their research, their definition of user context can vary quite a bit and is often very specific. In those cases, user context sometimes only includes location, time, people, and places nearby and the changes of them, without addressing psychological aspects such as emotional well-being (Shin et al., 2009).

A more broad definition of context is given by Almazan: *"Context consists of one or more relationships an information item has to another information item. An information item can be any entity, either physical (such as a person, a computer, an object), virtual (such as a computer service, a message), or a concept (location, time, and so on). A relationship describes a predicate connecting two or more information items, which may change at any time for any reason"* (Pradeep & Krishnamoorthy, 2019, p. 47). This definition is suitable for this thesis, and is kept in mind when mentioning (user) context from now on.

3.3.2. Characteristics of context

With context being now defined, some important characteristics and properties of context can now be examined. It is important to note that depending on the definition of context, some characteristics might vary. However, for the most part, these properties can also be found with other context definitions.

- **Context can be temporally static or dynamic**
  Depending on the kind of information, its relevancy can be either very short or very long (sometimes even indefinitely). A user's date of birth, for example, never changes and his or her relationships likely stays the same for months or years. However, physiological data for instance changes very fast and its relevancy is therefore short-lived. (Henricksen et al., 2002)

- **Context is imperfect**
  Contextual information can be incorrect, outdated, inconsistent, or incomplete. Errors in sensors, time-delays, inconsistencies between sensors and the inability to cover all aspects of the context will lead to imperfect context. This becomes especially true if some contextual information is derived from other sensors that provide false information (Henricksen et al., 2002).

- **Context arises from multiple sources**
  Context can often only be derived from multiple sources. These sources on their own lack the ability to paint the bigger picture on their own (Pradeep & Krishnamoorthy, 2019).

- **Sources for context information**
  There are three major types of sources for contextual information: sensors, human input and inference. Generally speaking, sensor information is mostly short-lived and changes frequently, whereas data obtained from human input is mostly longer relevant than sensor-based data. The relevancy of inferred data is dependent upon on what kind of information it was derived from (Henricksen & Indulska, 2004).

Multiple pieces of context information cannot only vary from how they are obtained and their time of relevancy, but also on their quality. One idea is to rate their quality with the following five metrics:
(i) accuracy – the precision of the obtained data,
(ii) confidence – how likely it is that the information is correct,
(iii) freshness – how recent the information is,
(iv) resolution – the granularity of the data,
(iv) credibility – how reliable the information is.
(Pradeep & Krishnamoorthy, 2019)

3.3.3. Context reasoning

With context and its characteristics now being defined, and the available sensors and pieces of context information introduced, the next step is deriving concrete context from these data points. This process can also be referred to as *context reasoning*. One approach is a three-step process: **(i)** data pre-processing – handling null values, normalizing data, removing outliers, **(ii)** combining sensor data – aggregating data points from multiple sources, and **(iii)** inference – inferring a higher level context from this lower level context. How this process is performed is highly dependant upon the technologies and algorithms used in the approach (Pradeep & Krishnamoorthy, 2019). Some of the most common techniques are as follows:

- **Rule-based:** Rules are defined by a human expert, and are IF-THEN statements that link certain conditions to actions. They are easy to understand and create, but lack validation and are more prone to errors than other methods (Pradeep & Krishnamoorthy, 2019; Abu-Nasser & Abu Naser, 2018).

- **Supervised learning**: This method uses a significant amount of labelled data – where each training example has a label of the corresponding ground-truth – to create predictive models. The most common techniques are Bayesian Networks, Artificial Neural Networks, Support Vector Machines, and Decision Trees (Zhou, 2018; Pradeep & Krishnamoorthy, 2019).

- **Unsupervised learning**: Unsupervised learning techniques use unlabelled data to find hidden structures of the provided data. The most common techniques are k-Nearest Neighbour, Clustering, and Kohonen Self Organization Map (Pradeep & Krishnamoorthy, 2019)

- **Probabilistic logic:** Here, numerical values all have associated probabilities that can be used to handle uncertainty by combining multiple sources. The most common techniques are Hidden Markov Models and the Dempster Schafer Theory of Evidence (Pradeep & Krishnamoorthy, 2019).

3.4. Just-in-time adaptive interventions (JITAIs)

Just-in-time adaptive interventions (JITAIs) are interventions "aiming to provide the right type/amount of support, at the right time, by adapting to an individual's changing internal and contextual state" (Nahum-Shani et al., 2017, p. 446). As the name suggests, JITAIs try to time interventions so that they are neither too early nor too late. Timing can either refer to clock time or can be event-based depending on the use cases. A time-bound JITAI could help comply with the medication plans in the mornings and evenings. An example of an event-based system would be when an alcoholic in the proximity of a favourite bar or liquor shop (or any other high-risk location). As these events often happen irregularly, it would be hard to predict them on clock time alone. To try to find a suitable timing and type of intervention, JITAIs look for three key points. (Nahum-Shani et al., 2017)

- Check if the person is in a state that needs support

- Check what type and amount of support is necessary for this current state

- Check if offering this support could potentially lead to the desired behaviour.

To identify states that need support, *states of vulnerability* are introduced. They describe situations, where people are more prone to make decisions that have adverse health consequences for them. The problem is that they can happen multiple times in short time windows, and designers might be tempted to offer interventions very often as a counter measure. However, this can backfire, because the users might be overwhelmed by this.
Also, depending on the situation and the context, the person's attention might be needed somewhere else entirely. These two factors introduce new challenges for intervention, as badly timed ones can lead to intervention fatigue, where the person is emotionally or cognitively burnout. Therefore, interventions should only be offered when the individual can and is willing to "receive, process, and utilize just-in-time support" (Nahum-Shani et al., 2017, p.450).

3.4.1. Components

To fulfil the unique properties of JITAIs, there exists a framework describing four major components suggested for these interventions: *decision points*, *intervention options*, *tailoring variables* and *decision rules*.

Decision points define at which points in time the decision for an intervention should occur. It depends on how often a change in a tailoring variable can realistically occur. In a JITAI that focuses on medication plan adherence, it does not make sense to set a decision point every minute, if the user should take the medication at a certain time. However, when these tailoring variables are expected to change quickly, decision points could occur as often as every minute. Also, the day of the week could also play an important role, such as when dealing with work-related stress. Then, decision points on the weekend are not meaningful (Nahum-Shani et al., 2017).

Intervention options are the strategies that could potentially be offered at a decision point. Depending on the application, they can vary drastically, from offering feedback or advice, the intensity of support, to the way the support is offered. Many JITAIs follow a framework for their interventions, a concept called ecological momentary interventions (EMIs). These EMIs are delivered to people in their daily lives – in the real world and in real time – and should therefore be quick and short. (Nahum-Shani et al., 2017; Heron & Smyth, 2010)

Tailoring variables are pieces of information about the subject that are used to determine under which conditions an intervention should take place, and which of the intervention options should be applied. An example could be the distance to a high-risk location, such as a bar or liquor shop, for alcohol-consumption JITAIs. These tailoring variables can either be active, where the user has some form of interaction (e.g., self-reporting), or passive ones that need no interactions (e.g., tracking the GPS location of the person).

Decision rules combine intervention options with tailoring variables. They include thresholds that define which intervention option should be offered with a given tailoring variable. In the example with the distance to a high-risk location (the tailoring variable) the decision rule could look like this: (Nahum-Shani et al., 2017, p. 451-452)

```
If distance ≤ threshold, then
     Provide intervention i
Else (distance > threshold),
     Provide nothing
```

3.4.2. Design

These four components should be chosen with a clear goal in mind. JITAIs differentiate between proximal and distal goals. The distal ones can be seen as the long-term goals that the interventions try to achieve, such as reducing alcohol consumption or spending more time being physically active. The proximal outcomes are short-term goals that should ideally be measurable shortly after the intervention. An example would be a step counter measuring if the subject is taking a walk during lunch breaks. The idea is that if proximal goals are regularly and consistently reached, then it is only a matter of time until the distal goals are achieved as well. The following figure 3.2 visualizes the interactions between the four JITAI components and the proximal and distal outcomes (Nahum-Shani et al., 2017).

Additionally, these two types of goals also have an impact on the compliance and quitting of JITAIs. Depending on the type of distal outcome, the minimum amount of engagement time with the JITAI system can range from short-term to long-term. Especially with long-term JITAIs, it can be difficult to keep people from lowering their use or discarding the whole system entirely. Generally, this intervention fatigue can be characterized into three groups. Firstly, *cognitive overload*, which is when people come across tasks that are too mentally challenging. Secondly, *habituation*, which describes the phenomenon where an excessive number of interventions leads to a reduction in a user's response to the intervention. And lastly, *negative emotions* (boredom, anger, disappointment, etc.) towards the JITAI application in general.

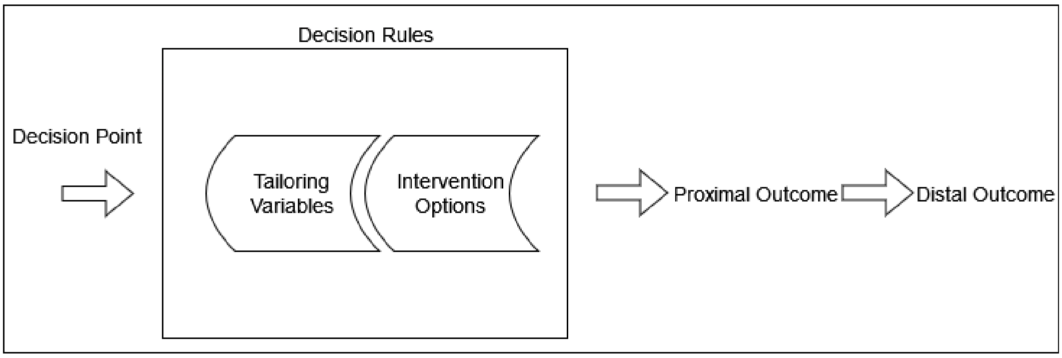

Figure 3.2.: Concept for JITAI components
modified from (Nahum-Shani et al., 2017, p.451)

4. Hypotheses

H0: There is no difference between stress-related JITAIs, which do and do not incorporate contextual user information, when it comes to reducing unfavourable intervention timings.

H1: Incorporating contextual user information does help to reduce unfavourable intervention timings for stress-related JITAIs.

5. Design

This chapter gives an overview of the design of the proposed intervention application and discusses the choices made during this process.

5.1. The JITAI components

As a general guideline, the intervention application described in this paper follows the framework of just-in-time adaptive interventions (JITAIs), discussed in section 3.4. This section includes the specific components with their respective parameters chosen for this use-case.

5.1.1. Decision points

As explained, decision points describe how often or when an intervention could be deemed meaningful. This depends on the application and is very contingent upon how often the tailoring variables change. For this application, the two most important times to avoid interventions are both during sleep and in the period directly after an offered intervention. During the day, decision points are chosen in a pre-defined time interval. However, the length of this interval is of great importance. On the one hand, more frequent decision points will increase the probability of detecting moments of stress. On the other hand, constant checks via the smartwatch's PPG sensor will drain its battery rather quickly. Depending on how the smartwatch is used otherwise, and how often the user is willing to charge the device, this number might be changed accordingly.
Therefore, the following rules will be applied for this system:

- Decision points only occur from when the application is launched until 23:00. This is to preserve battery life and to prevent intervention prompts during sleep.

- After an intervention took place, wait for one hour until which decision points will continue normally. This is primarily to prevent intervention fatigue.

- During the day, a decision point is due every 15 minutes. This is a compromise between battery duration and detecting as many stressful situations as possible.

5.1.2. Intervention options

Intervention options define what the interventions are. As a guideline, the framework of EMIs was used. One key feature of EMIs is that they are provided during the user's everyday lives – in real time and in the real world. Ideally, they incorporate personalized feedback and contextual information and are most often rather short. In a clinical study, in which EMIs are most often used, psychologists and other experts work together to create these tailored messages (Heron & Smyth, 2010). For this thesis, a general purpose intervention will be used, as creating tailored interventions is not the focus of it and would require expert knowledge in psychology. More precisely, a prompt is shown where the user is asked to use the five senses to label objects in the vicinity, which is a mindfulness exercise taught in cognitive behavioural therapy. (Howe et al., 2022)

5.1.3. Tailoring variables

Tailoring variables are information used to determine if an intervention is going to be performed, and can also be used to select which intervention option to offer. For this system, the most important variable is data from the smartwatch's PPG sensor. More specifically, the RR intervals that are used to calculate all time-domain and frequency-domain HRV parameters. This is the primary source of data.
The secondary source is the combination of different sensor and device information that is used to infer the context of the user. While this information is not used to detect stress itself, it is needed to differentiate between suitable and unsuitable user contexts.

5.1.4. Decision rules

As mentioned, decision rules combine tailoring variables with intervention options. Often times, machine learning models or knowledge-based systems are used for this task. Because of its simplicity and speed of implementation (and also because of the lack of pre-existing datasets), a rule-based system was chosen for this thesis. The thresholds for the HRV metrics are taken from the literature and are discussed in the next chapter.

With all these components in mind, the general idea of the application is visualized in figure 5.1 and works like this: When a decision point is reached, stress is classified with the HRV parameters. If stress occurred, contextual information is retrieved and inferred. Lastly, if the user context allows for an intervention, the intervention is displayed. Then there is a wait duration of one hour, until the decision points continue normally. If any of the other checks are false, then the decision points also continue normally.

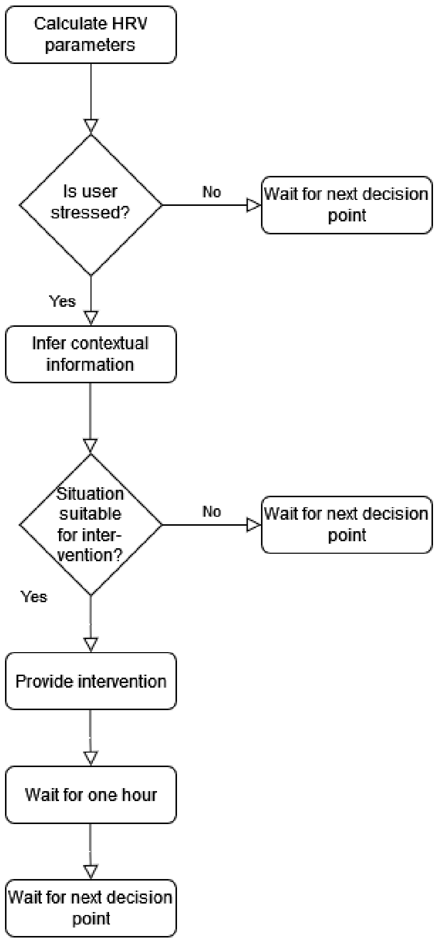

Figure 5.1.: A flow chart visualizing the decision rules

5.2. Context

This section discusses the available sources for context information, which were chosen for this thesis, and how they are inferred.

5.2.1. Available context information

This thesis discusses how to bring context to stress-monitoring JITAIs with evaluating HRV. Specifically, this process is done with the help of a smartwatch, which delivers the PPG data for the HRV analysis, and a smartphone which functions as the companion device. Therefore, both the smartwatch and the smartphone can be used as devices for context information. The typical sensors available in a smartphone nowadays are visualized in figure 5.2.

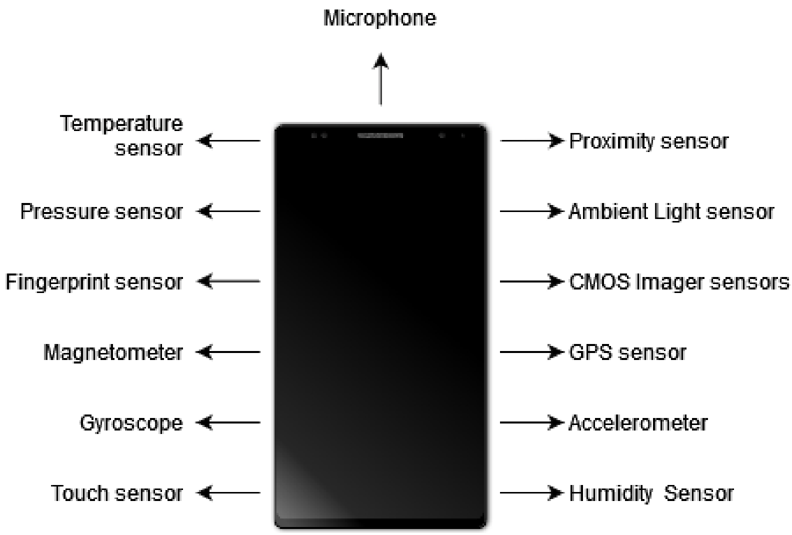

Figure 5.2.: Typical sensors present in smartphones
modified from (Majumder & Deen, 2019, p.5)
with ("Android black smartphone | Free SVG", n.d.)

Additionally, other contextual information can be used from smartphones besides hardware sensors. The day of the week (or the distinction between work week and weekend), the time of day, and calendar events can all be used to derive the context of the user (Santos et al., 2010). Furthermore, how the smartphone is being used can also be used as contextual information. What kinds of apps are used for how long, web activity of users, and even with whom they talk to or text (and the contents of their conversations) can all be indicators for the user's context. Also, even when location services are turned off, available Wi-Fi and Bluetooth devices can give hints about the user's whereabouts (Sarker, 2019).

Besides the smartphone, the sensors of the smartwatch can also be used. As this thesis uses a Compal Fossil Sport (model FTW6024 – released in 2019), the list of available sensors were retrieved and printed via Android's SensorManager class. The retrieved list of accessible sensors is as follows:

- Wrist tilt gesture
- Step detection
- Step counter
- Stationary detection
- Significant motion
- Game rotation vector
- Geomagnetic rotation vector
- Pressure
- PPG
- Motion detection
- Magnetic field
- Linear acceleration
- Uncalibrated accelerometer
- Heart rate
- Off-body detect
- Gravity
- Gyroscope
- Ambient light sensor

However, it is important to note that some of these listed sensors might not all be hardware sensors. As stated on the Android developers website ("Motion sensors | Android Developers", 2022), some sensors like linear accelerometer can either be hardware-based or software-based, depending on the model of the device. However, based on the API that is provided by Wear OS, sensors cannot be differentiated between hardware- and software-based.

5.2.2. Chosen sensors

To infer context later, the following sensors (or pieces of information from apps) were used:

- Calendar (phone)
- PPG (watch)
- GPS location (phone)
- List of Wi-Fi networks (phone)

- Gyroscope (watch)

Generally speaking, if both the smartphone and the smartwatch have the same type of sensor (like GPS), then the phone's sensor will be used. This has the effect of prolonged battery life and removes the need to send data to the smartphone that it already has, as the smartphone will handle all the logic to determine the context. The only exception to this rule is the gyroscope. The data comes from the smartwatch and is used to detect abrupt hand movements (e.g., during sport) that would lead to measurement errors of the PPG sensor.

With these sensors, the following contextual information is inferred: the location of the user via GPS, differentiation between work and home via the names of available Wi-Fi networks if location services are disabled, the availability during work time, and the detection of physical activities. Also, users can manually blacklist locations if they wish to remain undisturbed, such as when visiting a place of worship or going to a healthcare facility.

6. Methods

This chapter describes the general flow of the apps, their communication with the server, how the PPG signal was preprocessed and analysed, and how context is inferred.

6.1. General procedure

The system developed consists of three different parts: the Wear OS (Google's operating system for smartwatches) app running on the smartwatch, the Android app running on the smartphone, and a Python application running on a server. Once a decision point is reached (and under the assumption that the user is stressed and the user's context allows for an intervention), the communication between the three applications is visualized in figure 6.1.

The smartphone app acts as the main component and communicates with the smartwatch via the Wearable Data Layer API. The API comes with a MessageClient and a DataClient. The former is used for requests and remote procedure calls, while the latter synchronizes data between the devices ("Send and sync data on Wear OS | Android Developers", 2022). When a decision point occurs, it sends a message to the Wear OS app on the smartwatch. Then, a listener is registered on the smartwatch that captures incoming PPG data with around 100 samples per second for four minutes. The data is then transferred via the DataClient to the smartphone app, that sends an HTTP request to the python server. The python server analyses the signal – discussed in section 6.2 – and then responds with the calculated RMSSD value.

The smartphone compares the calculated RMSSD value with the RMSSD confidence intervals for stress and rest. The values used for the calculation were discussed in section 2.3 and visualized in table 2.2. The confidence levels for the intervals are 99.95%. The reason behind this number is to maximize the size of the two intervals, while still avoiding any overlaps between the two. The calculated value is discarded when it does not lie in one of the two intervals, as it is most likely due to a measurement error. When the value indicates rest, then no further action is performed and the smartphone apps waits until the

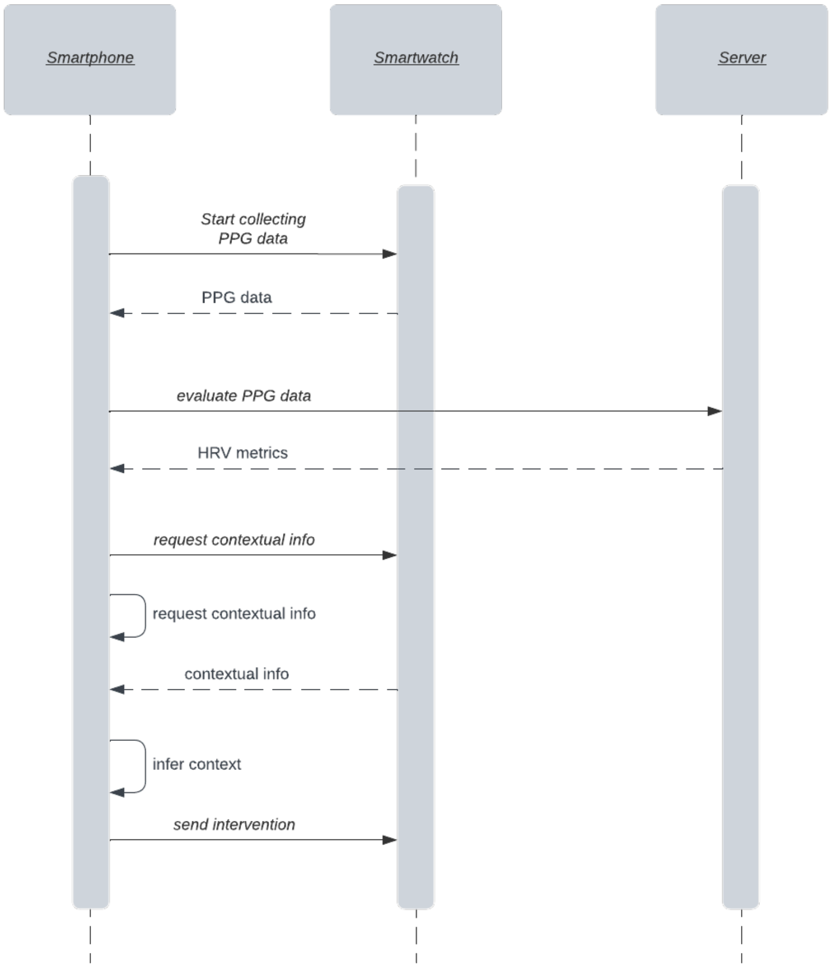

Figure 6.1.: Communication between the apps and the server

next decision point.

When the RMSSD value hints at stress, the subcomponents for context inference are started. Currently, there are four different in use: CalendarContext, DnDContext, LocationContext, and ActivityContext. Each of them decides, with provided data, whether this current context is suitable for an intervention or not – which is discussed in section 6.3 in greater detail. Only when all components determine that the situation is convenient will an intervention start and a message is sent to the smartwatch. As mentioned before, to maximize the benefit of the JITAI application, the interventions themselves should be modelled and chosen with care and expertise. However, as this thesis focuses on technical implementation and a general proof-of-concept, only the discussed exercise from cognitive behavioural therapy is offered. Finally, the calculated RMSSD value, the inferred context, and the intervention decision are saved for analytic purposes.

6.2. PPG signal analysis

As mentioned before, the raw PPG signal is analysed by a Python server. The signal is sent in a JSON format, consisting of the timestamps and the recorded values. For the HRV analysis, the python library HeartPy (Van Gent et al., 2019; van Gent et al., 2019) was chosen.

HeartPy offers functions to preprocess the PPG signal, such as clipping detection and interpolation, digital filters (discussed in 3.1.1), and peak enhancement. The filtered signal is then passed to a peak detection algorithm, that either accepts and marks or rejects the peaks of the signal. In the case of a PPG signal, these peaks are called diastolic peaks and are used for time-domain and frequency-domain analysis. For time-domain analysis, the time between these peaks (also called R-R interval) is used for the metrics, such as the root-mean-square of successive differences (RMSSD).
For frequency-domain analysis, the signal is first interpolated to evenly space out the data, then the Power Spectral Density (PSD) is calculated, e.g., with a fast Fourier transform. The PSD is then used to integrate the required frequency spectrums, such as 0.05-0.15Hz for low frequency (LF).

However, after some testing with self-recorded PPG data from the smartwatch, the provided methods for signal preprocessing did not suffice for a satisfactory result of the peak detection algorithm. While working as intended when remaining still, abrupt hand movements cause errors in the calculated RMSSD value. Thus, before passing the raw data to a band-pass filter, an off-line change point detection algorithm (discussed in 3.1.2) is used to split the data into segments. A sliding window algorithm from the Python package ruptures (Truong et al., 2020) was chosen, and the result is visualized in figure 6.2.

The next step is to pass the segments that were obtained from the step detection with a minimum length of 20 seconds to a bandpass filter that removes frequencies smaller than 0.7 and greater and 3.5Hz. This is equivalent to 42 and 210 beats per minute and is done to filter unnaturally low or high frequencies. The filtered signal is then passed to the peak detection algorithm. The signal before and after band-pass filtering is visualized in figure 6.3 and the output of the peak detection algorithm in figure 6.4. The peaks are then used to calculate the time-domain metrics. Because the original signal was first split into segments, and each segment has its own HRV metrics, the weighted average HRV metrics across all segments are then calculated. The weights themselves are the lengths of the specific intervals. The weighted average HRV metrics, which currently only includes the RMSSD, is then returned by the server to the smartphone.

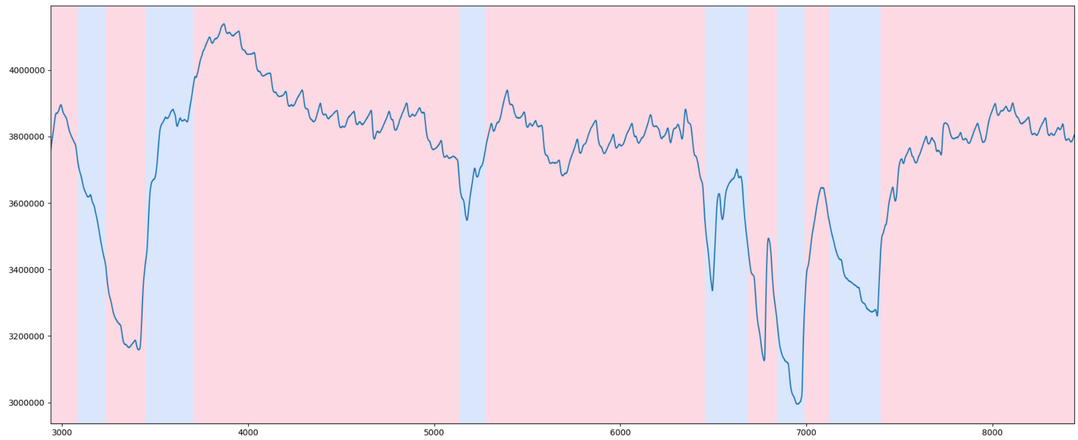

Figure 6.2.: Results of the change point detection algorithm

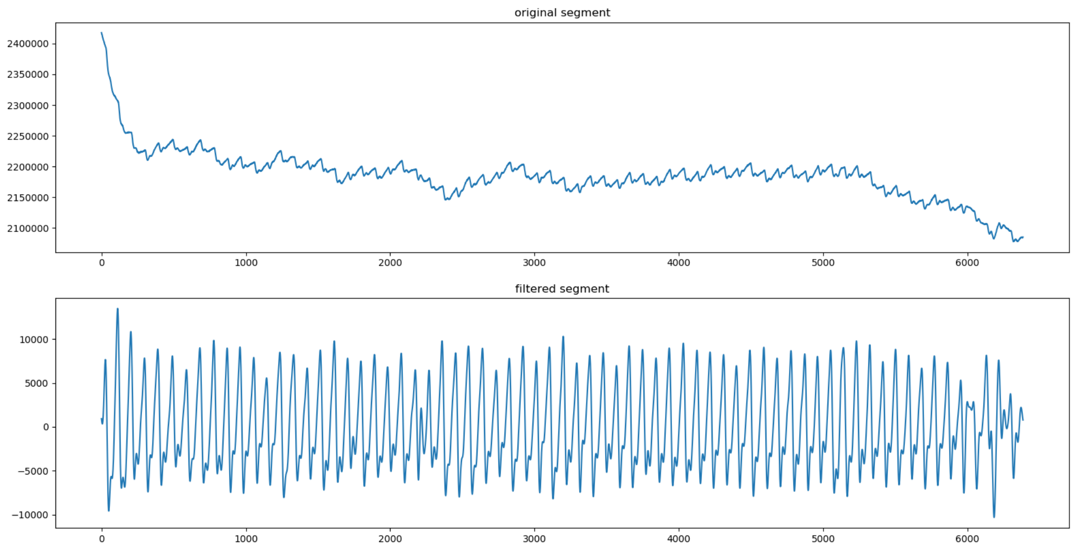

Figure 6.3.: Segment before and after band-pass filtering

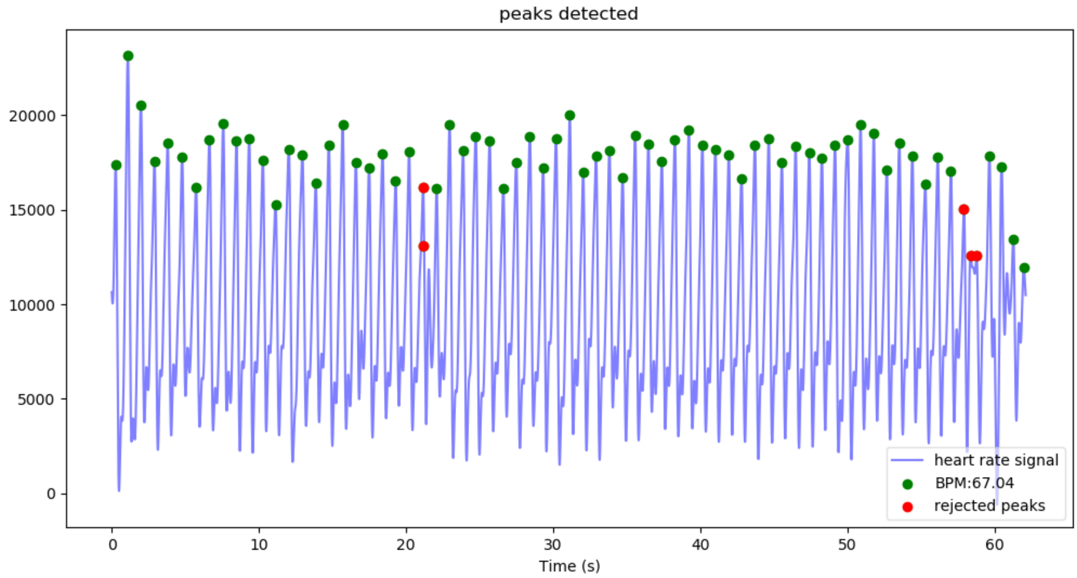

Figure 6.4.: Peak detection with HeartPy

6.3. Context inference

After detecting stress with the HRV metrics, contextual information is retrieved and analysed. The system was designed modularly, so that additional aspects of user context can be added at any given time. Currently, there are four different topics being evaluated.

Calendar: all calendars that are linked with the user's device are searched for events that overlap with the recorded time of the stressful situation. Events that are automatically added (such as start of a calendar week or public holidays) are ignored. The same is true for any birthday entries.

Location: for this module, the user can specify if a certain location is inconvenient for an intervention. For this, the GPS location (longitude and latitude) with a maximum distance to this location can be defined. Additionally, a list of the available Wi-Fi networks is also saved, which allows for the detection of a location even if location services are turned off.

Do not disturb: this simply checks if the smartphone or the smartwatch are currently in do-not-disturb mode or not. Additionally, the user can specify times when interventions are not desired in general.

Physical activity: the gyroscope and accelerometer of the smartwatch are used to capture and evaluate data related to the position and movement of the arm where the smartwatch is worn. This data is used to detect if the user is probably engaged in any physical activities, such as sport. This step might not be that important for the user, but is especially important to ensure the proper performance of the systems as a whole. This is because the readings of the PPG are more prone to measurement errors during physical activities, therefore limiting the possibility of accurately assessing the user's HRV parameters.

7. Evaluation and Results

This chapter discusses the methods used to evaluate the system and presents the obtained results from said methods.

7.1. Evaluation

To evaluate the system, a simulation-based approach was chosen. Different personas – discussed in subsection 7.1.1 – were created, each of them with a unique daily routine. The idea of this approach is that even though not all contextual information is temporal in nature, the fact whether a context is suitable for an intervention always has a time-component associated with it. For example, if a user is doing sports or is at a location where no intervention is desired, the contextual information are sensor values of the gyroscope and the longitude and latitude of the GPS position. However, all of these situations take place at a certain time during the day. Therefore, for the sake of this type of evaluation, any context can be broken down in time periods when they take place, and when simply no intervention is desired.

7.1.1. Personas and daily routines

Before creating the timetables for the evaluation, three personas were created. The idea behind a persona is to make the concept of the user more realistic. It includes a fictional name, picture, and background, and other information about the person, such as goals, behaviours, and motives. They are used in order to avoid vague concepts like “the user” and “user-friendly” that can be hard to grasp. (Blomkvist, 2002)

Three personas were created with a template of a persona canvas (Weidmann, 2018), which are included in appendix B: a 24-year-old physics student (Filippo), a 32-year-old human resource manager (Mujika), and a 42-year-old financial analyst (Martha). The average times (in hours per day), when these three people are not available or do not desire intervention are as follows:

- Filippo: total of 5.55h per day of no interventions desired
  - Calendar 2.8h → lectures and seminars
  - Do not disturb (DnD) 1.5h → submissions, uni projects and learning
  - Location 0.75h → flat of his girlfriend
  - Sport 0.5h → various types of sports across the seasons
- Mujika: total of 7h per day
  - Calendar 4.5h → work
  - Do not disturb (DnD) 1.5h → spending time with her son and husband
  - Location 0.3h → homes of other family members
  - Sport 0.70h → running
- Martha: total of 8.7h per day
  - Calendar 7h → work related meetings
  - Do not disturb (DnD) 0.75h → wants to be undisturbed while driving
  - Location 0.75h → various restaurants and church
  - Sport 0.2h → some tennis and miniature golf on weekends

As not every day is the same, these average values are taken as a guideline value and each day has a random deviation of ±5% (evenly distributed). Additionally, the average values are changed during the weekend. As a rule of thumb, the time of the calendar context is reduced by 80% and the rest of the contexts are all increased by 35%. This is done to reflect work- and study-free time, while also factoring in spending more time doing sports, meeting with friends and family and recuperating from the work week.

7.1.2. Stress simulation

In order to realistically model stressors in daily life, the average number of daily stressors is of utmost importance. The University of California, San Francisco, and the National Institute of Aging offer *The Stress Measurement Network*, which is a toolbox describing techniques for capturing different types of stress (Crosswell & Lockwood, 2020). For daily stress, the Daily Inventory of Stressful Events (DISE) is recommended, which is a semi-structured interview. (D. Almeida, 2018)

Studies utilizing this technique found that on average, participants reported at least one stressor on around 37% of days, and at least two stressors on around 12.5% of days. The average number of experienced stressors per day is $M = 0.51$ and $M = 0.65$ with a standard deviation of $SD = 0.74$ and $SD = 0.86$, respectively. (Koffer et al., 2016; Bellingtier et al., 2017; D. M. Almeida et al., 2002)

With these numbers in mind, the concrete simulation of stressors is performed as follows: a typical day is assumed to have a duration of 14 hours (from 08:00h until 22:00h), each day with associated probabilities of having N stressors. The chances of a day having zero or one stressor (0.64 and 0.21) have been calculated from Bellingtier et al. (2017), with the remaining probability of 0.15 distributed between two and three stressors (with probabilities of 0.1 and 0.05). If a sampled day has at least one stressor, then random times between 08:00h and 22:00h will be sampled and assigned to the stressor.
When the stress generation is performed with the above-mentioned probabilities, and run for 50 million days, the average stressors per day is $M = 0.56$ with a standard deviation of $SD = 0.86$. Both of these statistical parameters are close enough to the ones found in the papers mentioned, thus making this stress simulation model a suitable approximation.

7.2. Results

The length of the simulation was chosen to be 365 days and was run with the mentioned parameters. For each day, the stressors with corresponding timestamps for that particular day were sampled. It is important to note that while the schedules of the three personas differ, the time when stressors arise are equal for all. Figure 7.1 shows how often stressors occurred (with absolute and relative values). The vast majority of days do not have any stressors (65.21%), followed by one (22.19%), two (8.77%), and three stressors per day (3.83%).
After that, the schedules of each persona were examined and checked which of the interventions that were caused by the stressors are desirable or not. If an overlap between their schedule and the time of the stressor was found, an intervention was deemed to be undesirable. This information – how many interventions are blocked per day – was combined with how many stressors occur each day to create figure 7.2. It shows the ratio of how many interventions are blocked on days with the given number of stressors.
Filippo blocks an average of $M = 0.382$ with a standard deviation of $SD = 0.093$ of interventions in any given day. The values for Mujika and Martha are $M = 0.479, SD = 0.028$, and $M = 0.498, SD = 0.083$ respectively.

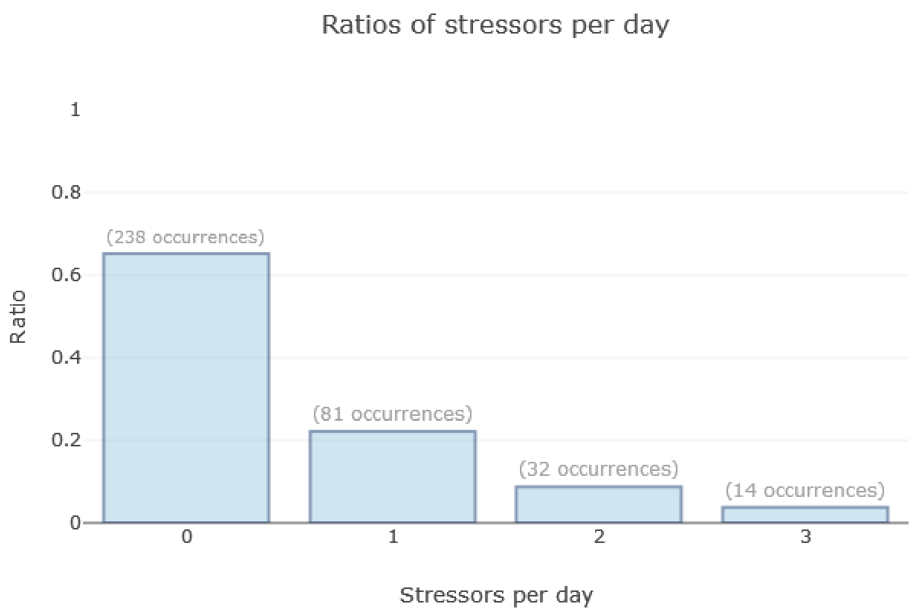

Figure 7.1.: Distribution of stressors per day

Although Martha normally blocks more interventions (around 55%) on most days compared to Mujika and Filippo, the opposite is true for days that have three stressors. Here, Martha prevents around 38%, while Filippo and Mujika intercept approximately 50% and 45%. This is discussed in the next chapter.

Due to the nature of the chosen evaluation, the stated hypotheses (chapter 4) cannot be empirically verified or falsified. However, the observed results show that a number of interventions are blocked due to undesirable user context. This indicates the interception of numerous interventions in a real-world setting, resulting in less unwelcome interventions offered to the user. Therefore, the results hint at the usefulness of incorporating user context in a stress-related JITAI system and at a potential acceptance of the H1 hypothesis.

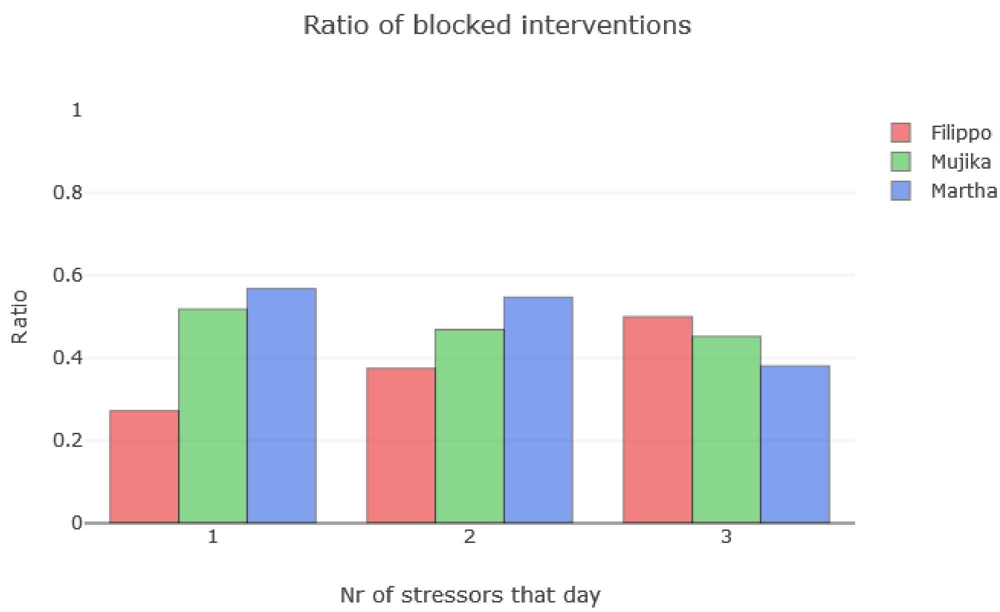

Figure 7.2.: How many interventions are blocked due to unsuitable context

8. Discussion

This chapter discusses the evaluation of this thesis and reflects upon the results that were obtained.

8.1. Evaluation

Ideally, the evaluation of the system would have been in the form of a long-term, large-scale user testing, combining objective data (such as HRV trends, contextual information, and data about the app usage) with subjective data (perceived stress levels and satisfaction with the system). However, due to encountered difficulties, the chances of this approach being successful were not sufficient.
Firstly, there was a lack of usable hardware that could have been given out to people for the user tests. Ergo, this would cause a need for a longer time period just for the user tests, as the number of tests that could be run in parallel would be quite small. Even if a user was in possession of a smartwatch, the use of their own device would not be recommendable, as differences in hardware (especially the PPG sensor) can cause alterations in the observed results.
Secondly, there is still the issue of measurement errors caused by hand movements. Without a highly reliable method to prevent this, there are chances of the HRV metrics being faulty as a consequence. This would then lead to either false positives or false negatives for intervention timings, decreasing the benefit for the user.

However, the chosen method of stress and schedule simulation is by far not perfect. While the parameters of stress simulation are backed by the discussed papers, the question remains if this average number of stressors is approximate to the number of stressors that would have been obtained during a real user test. The Daily Inventory of Stressful Events (DISE) examines the stressful events via asking the user for their experiences. Therefore, this method captures *perceived* stress. Whether less, equal, or more stressors are captured with HRV metrics is debatable. However, it could be argued, that unmemorable events during the day that still cause short-term stress (such as almost missing a connecting bus or being cut off in traffic) would not show up on DISE at the end of the day. Nevertheless, these situations would still be recognized by analysing HRV.

Another limitation of this thesis was the hardware used.
Firstly, a compromise between battery lifetime and accuracy/runtime of the application was necessary. Running the Wear OS app continuously on the smartwatch (Compal Fossil Sport – model FTW6024) drains the battery completely in around five hours, even without any additional use of the device. Therefore, there is the decision whether to limit the time the app is active and risk missing stressful situations, or increase the PPG sensor's delay and therefore decrease the quality of the data collected.
Secondly, with an mHealth application in mind, it is not feasible to ask the user to recharge their smartwatches at least once per day, so that just this application is able to run in the background continuously. Further research could be done to test the quality of HRV metrics obtained from less accurate PPG data, and to examine how much battery life per day users are willing to spend on this application.

8.2. Results

As stated before, the results obtained cannot be used to empirically verify or falsify the hypothesis, but indicate the usefulness of adding context information to JITAIs. The results themselves were quite expectable, because the personas were created in a way that reflects different lifestyles with different levels of busyness in different schedules. Martha was most busy with an average of 8.7h per day, followed by Mujika with 7h, and Filippo with 5.55h per day. Naturally, Martha would intercept most interventions per day on average, as she is the most busy persona. This is reflected in the chart discussed before, that visualizes the ratio of blocked interventions on days with one, two, or three stressors. However, while she does prevent more interventions than the other two personas on days with one or two stressors, the opposite is true on days with tree stressors. One explanation to that is how often three stressors per day occur in the simulation. With a probability of 5%, only 14 out of 365 simulated days have three stressors. Therefore, the observed difference could potentially be due to the rarity of these days.

With this theory and the law of large numbers in mind, another simulation was performed with one million days to show the difference. The law of large numbers state that the higher the sample size, the closer the observed value is to the expected value (Bobbitt, 2020). With a length of one million days, two observations can be made. Firstly, the ratio of days that have N stressors more closely resembles the initial distribution used for the sampling. Secondly, as the number of days having three stressors is sufficiently large enough (now 49 993 instead of 14 occurrences), the difference between days having three stressors per day and any other given day is now insignificantly low. For Martha, the maximum difference between the different types of days is now 0.27%, compared to 18.69% observed before. The same can be said for the other two personas as well. In the simulation with one million days, the difference between the types of days become irrelevantly minor for them as well. This suggests that intercepting interventions is similarly efficient, regardless of how many stressors occur per day, and that the obtained results were due to an insufficiently large sample size.

9. Conclusion

This thesis examined the concept of adding contextual user information for stress interventions using the framework of just-in-time adaptive interventions (JITAIs). This was done by introducing HRV and how it can be analysed, defining stress and ways to detect it, and giving an overview of the current state-of-the-art methods in digital signal processing, JITAIs, persuasive technologies, and user context.
Next, it was hypothesized that including contextual information from users decreases the amount of undesirable intervention timings. The following segment discusses the conceptional structure of the application, that closely follows the concept and components of JITAIs. Additionally, the way how user context is captured and which sensors are to be used is also proposed. This was followed by the details of the implementation, how the PPG signal is retrieved and analysed, and which data sources are used to infer the user context. After that, the way how the results were evaluated was explained, and the obtained results listed. The last part of this thesis discussed the results and examined the limitations of it.

The results of the simulation-based approach indicate the benefit of including user context in stress-related JITAIs. However, since the evaluation did not include a user testing, the extent of this benefit can currently only be surmised. The next step would be to recruit people willing to participate in a mid- to long-term user test to obtain and evaluate realistic data. Additionally, the system itself could also be improved to increase user benefits. Concretely speaking, enhancing the reliability of the obtained HRV metrics (especially during movement) and looking into personalized context inference, such as by using machine learning methods.

Bibliography

Abu-Nasser, B. S., & Abu Naser, S. S. (2018). *Rule-Based System for Watermelon Diseases and Treatment* (SSRN Scholarly Paper No. 3242597). Social Science Research Network. Rochester, NY. Retrieved April 11, 2022, from https://papers.ssrn.com/abstract=3242597

Adibi, S. (2015). Introduction [Series Title: Springer Series in Bio-/Neuroinformatics]. In S. Adibi (Ed.), *Mobile Health* (pp. 1–7). Springer International Publishing. https://doi.org/10.1007/978-3-319-12817-7_1

Aguiló, J., Ferrer-Salvans, P., García-Rozo, A., Armario, A., Corbí, Á., Cambra, F. J., Bailón, R., González-Marcos, A., Caja, G., Aguiló, S., et al. (2015). Project ES3: Attempting to quantify and measure the level of stress. *Rev Neurol*, *61*(9), 405–415.

Akemi, C. D. D. E. T. 1., Urayasu-shi, Chiba, & Japan. (2019). The effects of workplace norms on women's work behaviour in Japan. Retrieved June 15, 2022, from https://researchoutreach.org/articles/effects-workplace-norms-womens-work-behaviour-japan/

Almeida, D. (2018). Daily Stressors. Retrieved June 14, 2022, from https://www.stressmeasurement.org/daily-stressors

Almeida, D. M., Wethington, E., & Kessler, R. C. (2002). The Daily Inventory of Stressful Events: An Interview-Based Approach for Measuring Daily Stressors. *Assessment*, *9*(1), 41–55. https://doi.org/10.1177/1073191102091006

Android black smartphone | Free SVG. (n.d.). Retrieved June 16, 2022, from https://freesvg.org/1549453830

Bellingtier, J. A., Neupert, S. D., & Kotter-Grühn, D. (2017). The Combined Effects of Daily Stressors and Major Life Events on Daily Subjective Ages. *The Journals of Gerontology: Series B*, *72*(4), 613–621. https://doi.org/10.1093/geronb/gbv101

Blomkvist, S. (2002). Persona - an overview (Extract from the paper "The User as a personality. Using Personas as a tool for design". Position paper for the course workshop "Theoretical perspectives in Human-Computer Interaction" at IPLab, KTH, September 3, 2002).

Bobbitt, Z. (2020). Law of Large Numbers: Definition + Examples. Retrieved June 29, 2022, from https://www.statology.org/law-of-large-numbers/

Business woman clapping isolated over a white background | Freestock photos. (n.d.). Retrieved June 15, 2022, from https://www.freestock.com/free-photos/business-woman-clapping-isolated-white-background-11417599

Castaldo, R., Melillo, P., Bracale, U., Caserta, M., Triassi, M., & Pecchia, L. (2015). Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis. *Biomedical Signal Processing and Control*, *18*, 370–377. https://doi.org/10.1016/j.bspc.2015.02.012

Crosswell, A. D., & Lockwood, K. G. (2020). Best practices for stress measurement: How to measure psychological stress in health research. *Health Psychology Open*, *7*(2), 205510292093307. https://doi.org/10.1177/2055102920933072

Cygankiewicz, I., & Zareba, W. (2013). Heart rate variability. *Handbook of Clinical Neurology* (pp. 379–393). Elsevier. https://doi.org/10.1016/B978-0-444-53491-0.00031-6

Fishta, A., & Backé, E.-M. (2015). Psychosocial stress at work and cardiovascular diseases: An overview of systematic reviews. *International Archives of Occupational and Environmental Health*, *88*(8), 997–1014. https://doi.org/10.1007/s00420-015-1019-0

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann Publishers.

Free Images : Man, boy, photo, male, portrait, young, professional, profession, shirt, face, eyes, free, head, the student, a person, general practitioner, id card, the id card 3168x4752 - - 795555 - Free stock photos - PxHere. (2017). Retrieved June 15, 2022, from https://pxhere.com/en/photo/795555

Healthcare expenditure statistics. (n.d.). Retrieved April 11, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Healthcare_expenditure_statistics

Henricksen, K., & Indulska, J. (2004). Modelling and using imperfect context information. *IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second*, 33–37. https://doi.org/10.1109/PERCOMW.2004.1276901

Henricksen, K., Indulska, J., & Rakotonirainy, A. (2002). Modeling Context Information in Pervasive Computing Systems [Series Title: Lecture Notes in Computer Science]. In G. Goos, J. Hartmanis, J. van Leeuwen, F. Mattern, & M. Naghshineh (Eds.), *Pervasive Computing* (pp. 167–180). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45866-2_14

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour

treatments. *British Journal of Health Psychology*, *15*(1), 1–39. https://doi.org/10.1348/135910709X466063
Howe, E., Suh, J., Bin Morshed, M., McDuff, D., Rowan, K., Hernandez, J., Abdin, M. I., Ramos, G., Tran, T., & Czerwinski, M. P. (2022). Design of Digital Workplace Stress-Reduction Intervention Systems: Effects of Intervention Type and Timing. *CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3491102.3502027
IJsselsteijn, W., de Kort, Y., Midden, C., Eggen, B., & van den Hoven, E. (2006). Persuasive Technology for Human Well-Being: Setting the Scene [Series Title: Lecture Notes in Computer Science]. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, W. A. IJsselsteijn, Y. A. W. de Kort, C. Midden, B. Eggen, & E. van den Hoven (Eds.), *Persuasive Technology* (pp. 1–5). Springer Berlin Heidelberg. https://doi.org/10.1007/11755494_1
Jan, H.-Y., Chen, M.-F., Fu, T.-C., Lin, W.-C., Tsai, C.-L., & Lin, K.-P. (2019). Evaluation of Coherence Between ECG and PPG Derived Parameters on Heart Rate Variability and Respiration in Healthy Volunteers With/Without Controlled Breathing. *Journal of Medical and Biological Engineering*, *39*(5), 783–795. https://doi.org/10.1007/s40846-019-00468-9
Jeyhani, V., Mahdiani, S., Peltokangas, M., & Vehkaoja, A. (2015). Comparison of HRV parameters derived from photoplethysmography and electrocardiography signals. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 5952–5955. https://doi.org/10.1109/EMBC.2015.7319747
Jones, A., Clarke, B., & Picton, P. (2020a). Electronic applications. Retrieved June 9, 2022, from https://www.open.edu/openlearn/science-maths-technology/electronic-applications/content-section-2.7
Jones, A., Clarke, B., & Picton, P. (2020b). Electronic applications. Retrieved June 9, 2022, from https://www.open.edu/openlearn/science-maths-technology/electronic-applications/content-section-2.2
Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., & Koo, B.-H. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature [Number: 3]. *Psychiatry Investigation*, *15*(3), 235–245. https://doi.org/10.30773/pi.2017.08.17
Koffer, R. E., Ram, N., Conroy, D. E., Pincus, A. L., & Almeida, D. M. (2016). Stressor diversity: Introduction and empirical integration into the daily

stress model. *Psychology and Aging*, *31*(4), 301–320. https://doi.org/10.1037/pag0000095
Koolhaas, J. M., Bartolomucci, A., Buwalda, B., Boer, S. F. d., Flügge, G., Korte, S. M., Meerlo, P., Murison, R., Olivier, B., Palanza, P., Richter-Levin, G., Sgoifo, A., Steimer, T., Stiedl, O., Dijk, G. v., Wöhr, M., & Fuchs, E. (2011). Stress revisited: A critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews*, *35*(5), 1291–1301. https://doi.org/https://doi.org/10.1016/j.neubiorev.2011.02.003
Kruk, J., Aboul-Enein, B. H., Bernstein, J., & Gronostaj, M. (2019). Psychological Stress and Cellular Aging in Cancer: A Meta-Analysis. *Oxidative Medicine and Cellular Longevity*, *2019*, 1–23. https://doi.org/10.1155/2019/1270397
Libretti, S., & Puckett, Y. (2022). Physiology, Homeostasis. *StatPearls*. StatPearls Publishing. Retrieved March 28, 2022, from http://www.ncbi.nlm.nih.gov/books/NBK559138/
Lin, W.-H., Wu, D., Li, C., Zhang, H., & Zhang, Y.-T. (2014). Comparison of Heart Rate Variability from PPG with That from ECG [Series Title: IFMBE Proceedings]. In Y.-T. Zhang (Ed.), *The International Conference on Health Informatics* (pp. 213–215). Springer International Publishing. https://doi.org/10.1007/978-3-319-03005-0_54
Lyons, R. G., & Fugal, D. L. (2014). *Essential Guide to Digital Signal Processing, The* [OCLC: 1224592096]. Pearson. https://learning.oreilly.com/library/view/-/9780133812220/
Majumder, S., & Deen, M. J. (2019). Smartphone Sensors for Health Monitoring and Diagnosis. *Sensors*, *19*(9), 2164. https://doi.org/10.3390/s19092164
Marin, M.-F., Lord, C., Andrews, J., Juster, R.-P., Sindi, S., Arsenault-Lapierre, G., Fiocco, A. J., & Lupien, S. J. (2011). Chronic stress, cognitive functioning and mental health. *Neurobiology of Learning and Memory*, *96*(4), 583–595. https://doi.org/10.1016/j.nlm.2011.02.016
Massa, P., Mazzali, A., Zampini, J., & Zancanaro, M. (2017). Quantify Yourself: Are Older Adults Ready? [Series Title: Lecture Notes in Electrical Engineering]. In F. Cavallo, V. Marletta, A. Monteriù, & P. Siciliano (Eds.), *Ambient Assisted Living* (pp. 377–389). Springer International Publishing. https://doi.org/10.1007/978-3-319-54283-6_27
Motion sensors | Android Developers. (2022). Retrieved April 8, 2022, from https://developer.android.com/guide/topics/sensors/sensors_motion
Munla, N., Khalil, M., Shahin, A., & Mourad, A. (2015). Driver stress level detection using HRV analysis. *2015 International Conference on Advances in Biomedical Engineering (ICABME)*, 61–64. https://doi.org/10.1109/ICABME.2015.7323251

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2017). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support [Number: 6]. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Oinas-Kukkonen, H., & Harjumaa, M. (2008). A Systematic Framework for Designing and Evaluating Persuasive Systems [ISSN: 0302-9743, 1611-3349 Series Title: Lecture Notes in Computer Science]. In H. Oinas-Kukkonen, P. Hasle, M. Harjumaa, K. Segerståhl, & P. Øhrstrøm (Eds.), *Persuasive Technology* (pp. 164–176). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68504-3_15

Oksman, V., Ermes, M., & Tikkamäki, K. (2016). Eustress - findings concerning the indication and interpretation of positive stress among entrepreneurs - a case study. *The Business and Management Review*, *7*(3), 342–347.

Orji, R., & Moffatt, K. (2018). Persuasive technology for health and wellness: State-of-the-art and emerging trends [Number: 1]. *Health Informatics Journal*, *24*(1), 66–91. https://doi.org/10.1177/1460458216650979

Pham, T., Lau, Z. J., Chen, S. H. A., & Makowski, D. (2021). Heart Rate Variability in Psychology: A Review of HRV Indices and an Analysis Tutorial. *Sensors*, *21*(12), 3998. https://doi.org/10.3390/s21123998

Pinheiro, N., Couceiro, R., Henriques, J., Muehlsteff, J., Quintal, I., Goncalves, L., & Carvalho, P. (2016). Can PPG be used for HRV analysis? *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2945–2949. https://doi.org/10.1109/EMBC.2016.7591347

Pradeep, P., & Krishnamoorthy, S. (2019). The MOM of context-aware systems: A survey. *Computer Communications*, *137*, 44–69. https://doi.org/10.1016/j.comcom.2019.02.002

Rowland, S. P., Fitzgerald, J. E., Holme, T., Powell, J., & McGregor, A. (2020). What is the clinical value of mHealth for patients? *npj Digital Medicine*, *3*(1), 4. https://doi.org/10.1038/s41746-019-0206-x

Santos, A. C., Cardoso, J. M., Ferreira, D. R., Diniz, P. C., & Chaínho, P. (2010). Providing user context for mobile and social networking applications. *Pervasive and Mobile Computing*, *6*(3), 324–341. https://doi.org/10.1016/j.pmcj.2010.01.001

Sarker, I. H. (2019). Context-aware rule learning from smartphone data: Survey, challenges and future directions. *Journal of Big Data*, *6*(1), 95. https://doi.org/10.1186/s40537-019-0258-4

Send and sync data on Wear OS | Android Developers. (2022). Retrieved June 8, 2022, from https://developer.android.com/training/wearables/data/data-layer

Shin, D., Lee, J.-w., Yeon, J., & Lee, S.-g. (2009). Context-Aware Recommendation by Aggregating User Context. *2009 IEEE Conference on Commerce and Enterprise Computing*, 423–430. https://doi.org/10.1109/CEC.2009.38

Stila - About. (n.d.). Retrieved June 7, 2022, from http://stila.pms.ifi.lmu.de/

Svalastog, A. L., Donev, D., Jahren Kristoffersen, N., & Gajović, S. (2017). Concepts and definitions of health and health-related values in the knowledge landscapes of the digital society. *Croatian Medical Journal*, *58*(6), 431–435. https://doi.org/10.3325/cmj.2017.58.431

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, *167*, 107299. https://doi.org/10.1016/j.sigpro.2019.107299

Van Gent, P., Farah, H., Van Nes, N., & Van Arem, B. (2019). Analysing Noisy Driver Physiology Real-Time Using Off-the-Shelf Sensors: Heart Rate Analysis Software from the Taking the Fast Lane Project. *Journal of Open Research Software*, *7*(1), 32. https://doi.org/10.5334/jors.241

van Gent, P., Farah, H., van Nes, N., & van Arem, B. (2019). HeartPy: A novel heart rate algorithm for the analysis of noisy signals. *Transportation Research Part F: Traffic Psychology and Behaviour*, *66*, 368–378. https://doi.org/10.1016/j.trf.2019.09.015

Weidmann, K.-H. (2018). Persona Canvas, supplementary unpublished document from lecture "User Centered Software Development" during 4th semester of "Computer Science Bachelor" at Vorarlberg University of Applied Sciences.

Zhou, Z.-H. (2018). A brief introduction to weakly supervised learning. *National Science Review*, *5*(1), 44–53. https://doi.org/10.1093/nsr/nwx106

A. Healthcare expenditures in the EU

Current healthcare expenditure, 2019

	Million EUR	EUR per inhabitant	PPS per inhabitant	% of GDP
EU (¹)	*1 386 255*	*3 102*	*3207.51*	*9.9*
Belgium	50 759	4 418	3 901	10.7
Bulgaria	4 364	626	1 317	7.1
Czechia	17 546	1 644	2 443	7.8
Denmark	31 137	5 355	3 915	10.0
Germany	403 444	4 855	4 659	11.7
Estonia	1 892	1 426	1 792	6.7
Ireland	23 782	4 820	3 633	6.7
Greece	14 376	1 341	1 657	7.8
Spain	113 674	2 412	2 573	9.1
France	269 541	4 008	3 770	11.1
Croatia	3 785	931	1 440	7.0
Italy	155 249	2 599	2 611	8.7
Cyprus	1 562	1 771	1 946	7.0
Latvia	2 001	1 046	1 457	6.6
Lithuania	3 420	1 224	1 949	7.0
Luxembourg	3 411	5 502	3 870	5.4
Hungary	9 277	949	1 551	6.4
Malta (²)	1 110	2 290	2 754	9.0
Netherlands	82 365	4 749	4 102	10.2
Austria	41 483	4 672	4 078	10.4
Poland	34 400	906	1 636	6.5
Portugal	20 392	1 983	2 393	9.5
Romania	12 810	661	1 354	5.7
Slovenia	4 125	1 975	2 361	8.5
Slovakia	6 534	1 198	1 565	7.0
Finland	21 992	3 983	3 258	9.2
Sweden	51 824	5 042	3 968	10.9
Iceland	1 900	5 270	3 245	8.6
Liechtenstein	333	8 626	:	5.6
Norway	38 113	7 127	4 821	10.5
Switzerland	73 787	8 605	5 102	11.3
Bosnia and Herzegovina	:	:	:	9.1

(¹) 2019 EU calculated with 2018 Malta data
(²) 2018 data.
Source: Eurostat (online data codes: hlth_sha11_hf, demo_gind and nama_10_gdp)

eurostat

Figure A.1.: ("Healthcare expenditure statistics", n.d.)

Current healthcare expenditure, 2012-2019

	2012	2013	2014	2015	2016	2017	2018	2019	Overall change 2012-2019
	(EUR million)								(%)
EU (1)	:	:	1 178 168	1 213 418	1 247 799	1 291 767	1 334 056	*1 386 255*	18
Belgium	40 573	41 493	42 695	43 465	46 334	47 990	49 557	50 759	25
Bulgaria	3 186	3 004	3 306	3 386	3 637	3 936	4 121	4 364	37
Czechia (2)	:	12 314	11 989	12 202	12 610	13 864	15 872	17 546	42
Denmark	26 072	26 313	27 033	27 922	28 720	29 598	30 450	31 137	19
Germany	297 955	309 191	322 649	338 439	352 208	369 321	384 322	403 444	35
Estonia	1 045	1 138	1 227	1 319	1 410	1 573	1 734	1 892	81
Ireland	18 653	18 495	18 607	19 230	20 107	21 219	22 432	23 782	27
Greece	16 811	15 028	14 024	14 210	14 498	14 355	14 304	14 376	-14
Spain	94 409	92 512	93 817	98 343	99 708	104 068	108 266	113 674	20
France	235 981	241 706	248 178	251 667	256 291	260 341	264 355	269 541	14
Croatia (2)	:	2 855	2 908	3 028	3 184	3 328	3 561	3 785	33
Italy	142 676	141 526	144 317	146 613	147 963	150 697	153 685	155 249	9
Cyprus	1 274	1 250	1 212	1 208	1 254	1 328	1 451	1 562	23
Latvia (2)	:	1 233	1 291	1 389	1 556	1 610	1 804	2 001	62
Lithuania	2 097	2 147	2 266	2 424	2 581	2 733	2 972	3 420	63
Luxembourg	2 463	2 570	2 709	2 751	2 850	2 987	3 174	3 411	38
Hungary	7 429	7 396	7 488	7 731	8 124	8 573	8 899	9 277	25
Malta (4)	:	:	795	889	945	1 042	1 110	:	40
Netherlands	68 816	69 901	70 964	71 236	72 918	74 614	77 645	82 365	20
Austria	32 500	33 317	34 541	35 692	37 021	38 355	39 760	41 483	28
Poland (2)	:	25 166	25 681	27 280	27 756	30 664	31 502	34 400	37
Portugal	16 247	16 035	16 168	16 743	17 520	18 235	19 313	20 392	26
Romania	6 282	7 467	7 568	7 923	8 509	9 672	11 371	12 810	104
Slovenia (3)	:	:	3 200	3 309	3 429	3 520	3 797	4 125	29
Slovakia	5 550	5 583	5 256	5 418	5 666	5 721	5 991	6 534	18
Finland	19 271	20 034	20 237	20 389	20 399	20 654	21 111	21 992	14
Sweden	46 165	48 183	48 043	49 212	50 601	51 771	51 497	51 824	12
Iceland	939	992	1 109	1 275	1 523	1 821	1 872	1 900	102
Liechtenstein (2)	:	277	283	325	330	326	314	333	20
Norway	34 806	35 130	35 132	35 220	35 319	36 448	37 118	38 113	9
Switzerland	55 183	56 143	58 809	69 655	71 047	71 641	69 473	73 787	34
Bosnia and Herzegovina (2)	:	1 292	1 325	1 367	1 414	1 434	1 520		18

(1) 2019 EU calculated with 2018 Malta data

(2) Alternate year 2013 replaces 2012 in 'overall change 2012-2019'

(3) Alternate year 2014 replaces 2012 in 'overall change 2012-2019'

(4) 2019 data not available

Source: Eurostat (online data code: hlth_sha11_hf)

eurostat

Figure A.2.: ("Healthcare expenditure statistics", n.d.)

B. Personas

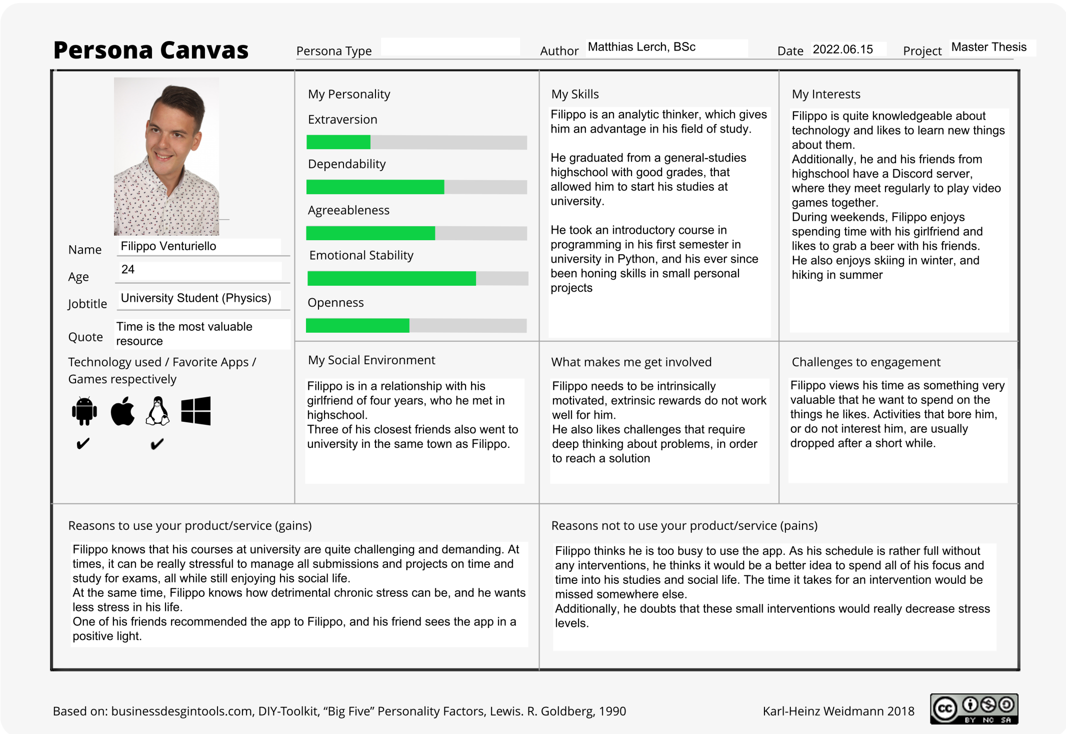

Figure B.1.: Persona 1
Template from (Weidmann, 2018),
picture from (“Free Images”, 2017)

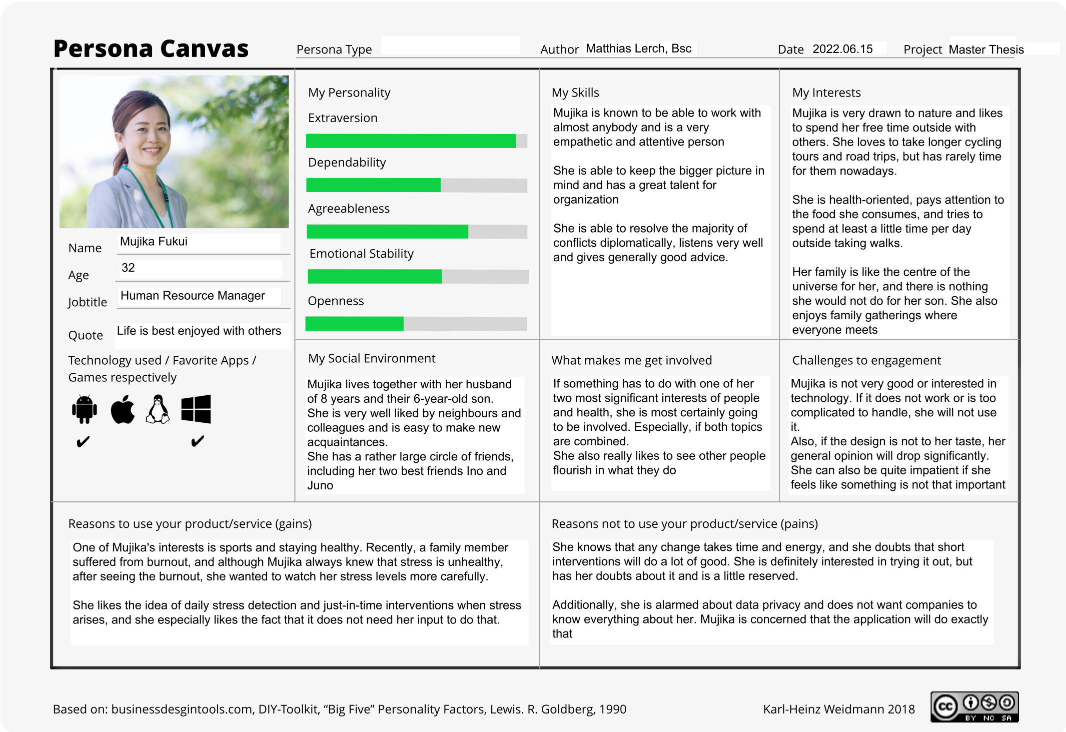

Figure B.2.: Persona 2
Template from (Weidmann, 2018),
picture from (Akemi et al., 2019)

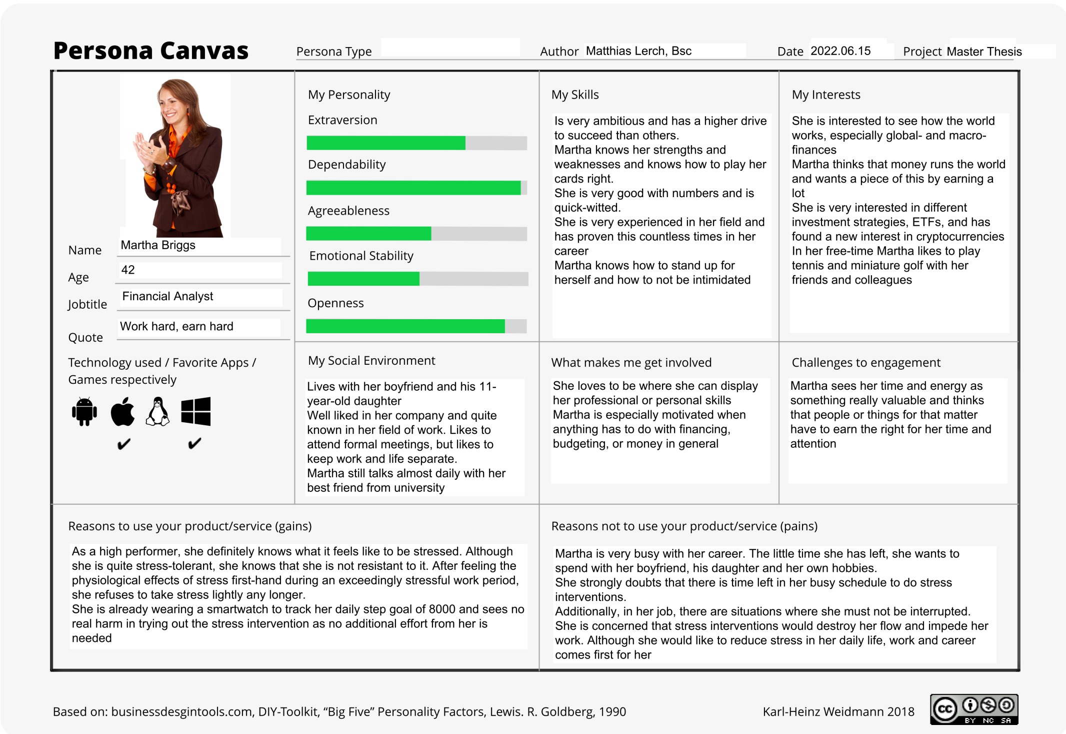

Figure B.3.: Persona 3
Template from (Weidmann, 2018),
picture from ("business woman clapping isolated over a white background | Freestock photos", n.d.)

Statutory Declaration

I declare that I have developed and written the enclosed work completely by myself, and have not used sources or means without declaration in the text. Any thoughts from others or literal quotations are clearly marked. This Master Thesis was not used in the same or in a similar version to achieve an academic degree nor has it been published elsewhere.

Dornbirn, the 7^{th} of July 2022 Matthias Lerch, BSc