Datenschutzfreundliche Vermessung und Auswertung persönlicher Daten mit Wearables zur Erhöhung der Arbeitssicherheit

Ergebnisbericht D6.1

Evaluationsbericht

Version	1.0
Datum	29.11.2024
Verfasser	Bianca Steffes (UdS) Reinhard Schwarz (IESE) Marcus Schröder, Jule Marie Hucke (neusta) Esteban Bayro-Kaiser (WearHealth)

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter den Förderkennzeichen 16KIS1511K, 16KIS1512, 16KIS1514 und 16KIS1665 gefördert.

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Ansprechperson

Marcus-Sebastian Schröder
neusta mobile solutions GmbH
Konsul-Smidt-Str. 24
28217 Bremen

E-Mail: m.schroeder@neusta.de

Inhaltsverzeichnis

Liste der Abkürzungen

AES	Advanced Encryption Standard
AWS	Amazon Web Services
DSGVO	Datenschutzgrundverordnung
EU	Europäische Union
HRV	Herzratenvariabilität
KI	Künstliche Intelligenz
SOC	System and Organization Controls
TLS	Transport Layer Security
TOM	Technisch-Organisatorische Maßnahme

1 Einleitung

Im Rahmen des WearPrivate-Projekts wurden Konzepte entwickelt, um eine datenschutzverträgliche Erfassung von Vital- und Kontextdaten mittels Wearables zum Zwecke des Arbeits- und Gesundheitsschutzes zu ermöglichen. Aus diesen Konzepten sind eine Reihe von Lösungsbausteinen hervorgegangen, deren Wirksamkeit und Praktikabilität die Projektpartner im weiteren Verlauf des Projekts evaluiert haben. Die Analysen basieren zum Teil auf praktischen Erprobungen, die mit dem WearPrivate-Demonstrator durchgeführt wurden. Andere Betrachtungen stützten sich auf anonyme Rohdaten, die von anderen Forschungsprojekten und vom Verbundpartner WearHealth bereitgestellt wurden. Ergänzend dazu führten die Projektpartner zu Beginn und zum Ende des Projekts Befragungen durch, um die Bedürfnisse, Erwartungen und Vorbehalte potenzieller Nutzer solcher Wearable-basierten Messprogramme zu ergründen und zu beurteilen, inwieweit die Lösungsbausteine zu einer vorteilhafteren Datenschutzeinschätzung von Wearable-Messprogrammen beitragen.

Dieser Bericht beschreibt die Evaluationsschritte, die im Projekt durchgeführt wurden, und deren Befunde.

2 Evaluation der Anonymisierungskonzepte

Im Folgenden soll auf die verschiedenen Konzepte eingegangen werden, mit denen die Anonymität im Kontext von WearPrivate erreicht oder verbessert werden soll.

2.1 Schutz der Profildaten mithilfe von k-Anonymität

Die vom Analysedienst verarbeiteten Profildaten Geschlecht, Geburtsjahr, Größe und Gewicht sollen durch k-Anonymität geschützt werden. Dazu sollen die Nutzer in Äquivalenzklassen eingeteilt werden, deren Profildaten dann verfremdet werden sollen, sodass alle Mitglieder einer Äquivalenzklasse nicht mehr voneinander zu unterscheiden sind.

Werden die Daten jedoch verfremdet, kann es zu einer Verschlechterung der Befundqualität des Analysedienstes kommen. Daher werden nun zum einen die Auswirkungen des Verfremdens der einzelnen Profildaten auf die Analyseergebnisse betrachtet und zum anderen anschließend daran Empfehlungen abgeleitet, welche möglichen Varianten des Verfremdens aus datenschutzrechtlicher Sicht zu empfehlen sind.

2.1.1 Auswirkungen der Profildatenverfremdung auf die Qualität des Analysedienstes

Das demografische Profil eines Nutzers wird durch sein Alter, seine Größe, sein Gewicht und sein Geschlecht definiert. Um die Auswirkungen von Profilrauschen (Veränderungen in den Profildaten) auf die Qualität der Belastungsanalyse zu bewerten, wurden Variationen in Alter, Größe und Gewicht in die demografischen Profile eingeführt. Anschließend wurden die auf den modifizierten Profilen basierenden Vorhersagen mit denen der Ausgangsprofile verglichen.

Zur Berücksichtigung natürlicher Variationen in demografischen Profilen wurden 468 Ausgangsprofile ausgewählt. Dazu wurden die insgesamt verfügbaren Profile in spezifische Gruppen unterteilt: 6

Gruppen für das Alter, 8 Gruppen für die Größe und 10 Gruppen für das Gewicht. Zudem wurden Kombinationen basierend auf diesen Gruppierungen generiert. Ursprünglich ergaben sich so 480 mögliche Ausgangsprofile; jedoch wurden Kombinationen ausgeschlossen, die eine große Körpergröße mit geringem Gewicht kombinierten, was zu einer endgültigen Anzahl von 468 Ausgangsprofilen führte.

Um den Effekt von Profilverfremdung zu analysieren, wurden Variationen in die Ausgangsprofile eingeführt, indem Deltas auf die Parameter Alter, Größe und Gewicht angewendet wurden. Insgesamt wurden 9 Variationen pro Parameter untersucht, wobei Modifikationen gleichzeitig auf alle drei Parameter angewendet wurden. Die spezifischen Variationen waren wie folgt:

- Alter: ± 0, 2, 5, 10, 15 Jahre
- Größe: ± 0, 2, 5, 10, 15 cm
- Gewicht: ± 0, 5, 10, 15, 20 kg

Beispielsweise könnte ein Ausgangsprofil mit {Alter = 35 Jahren, Größe = 170 cm, Gewicht = 75 kg} durch Anwendung von Alter_delta = 0, Größe_delta = +5 und Gewicht_delta = –2 modifiziert werden. Das resultierende Profil wäre dann {Alter = 35 Jahre, Größe = 175 cm, Gewicht = 73 kg}. Durch Anwendung aller möglichen Kombinationen der Deltas auf die drei Parameter könnte jedes Ausgangsprofil insgesamt 739 Variationsprofile ergeben.

Anschließend wurden die Belastungswerte unter Verwendung der modifizierten Profile berechnet und mit den Vorhersagen verglichen, die aus den Ausgangsprofilen generiert wurden. Für jede Profilvariation (Alter_delta, Größe_delta, Gewicht_delta) wurden deskriptive Metriken berechnet, einschließlich prozentualer Fehler und statistischer Kennzahlen, um die Auswirkungen der Variationen zu bewerten.

Basierend auf den berechneten Metriken wurden Profilvariationen identifiziert, die zu Änderungen von weniger als 15 Prozent in den Vorhersagen für jeden Nutzer führten, mit einer medianen Änderung von weniger als 5 Prozent. Diese Variationen wurden als solche betrachtet, die die Ergebnisse nicht signifikant beeinflussen.

Daher werden die in Tabelle 1 angegebenen Profilvariationen empfohlen. Bei Überschreitung der empfohlenen Variationsbreite sinkt die Analysequalität merklich. Bei Anwendung der Variationen wiesen 90 Prozent der Nutzer weniger als 10 Prozent Änderungen in ihren Vorhersagen auf.

Es ist wichtig zu beachten, dass diese Variationen eine Anonymisierung basierend auf WearHealth-Analysen bieten und nicht als allgemeines oder übertragbares Ergebnis für die demografische Anonymisierung in anderen Diensten interpretiert werden sollten.

Tabelle 1 Ermittelte Spielräume für Profildatenverfremdung im WearPrivate-Demonstrator: Werden die Profildaten maximal um die angegebenen Beträge variiert, ändert sich die Bewertung der Wearable-Messdaten durch den Analysedienst bei 90 Prozent der Betroffenen um weniger als 10 Prozent.

Variation des Alters	Variation des Gewichts	Variation der Körpergröße
5	0	2
2	0	2
0	5	-10
		-15

Variation des Alters	Variation des Gewichts	Variation der Körpergröße
	0	10
		5
		2
		0
		-2
		-5
		-10
	-5	15
		10
-5	0	-2

2.1.2 Datenschutzrechtliche Empfehlung zur Wahl der Datenverfremdung

Nun soll bestimmt werden, welche Varianten der Datenverfremdung aus datenschutzrechtlicher Sicht zu empfehlen sind. Ziel des Schutzes durch Verfremden ist es, dass eine unbeteiligte dritte Person D, der die verfremdeten Profildaten von einer Person P in die Hände fallen, diese Daten nicht auf P zurückführen kann. Konkret soll verhindert oder erschwert werden, dass D anhand der Liste aller verfremdeten Profildaten eindeutig bestimmen kann, welche Daten zu P gehören. Es wird folglich vorausgesetzt, dass P den Analysedienst nutzt. Für die verschiedenen Profildaten gehen wir von folgendem Wissen von D aus, welches darauf basiert, dass D zumindest ein Foto von P vorliegt:

- **Geschlecht:** Wie in Abschnitt 2.1.1 beschrieben, können nur die Parameter Gewicht, Größe und Geburtsjahr angepasst werden. Daher wird davon ausgegangen, dass D das biologische Geschlecht von P kennt. Es kann davon ausgegangen werden, dass D mit hoher Wahrscheinlichkeit das biologische Geschlecht aus dem äußeren Erscheinungsbild von P ableiten kann.
- **Geburtsjahr:** Für die Nutzergruppe der Beschäftigten ist mit einer möglichen Alterspanne von etwa 15 bis 65 Jahren zu rechnen; der konkrete Datensatz von WearHealth umfasst jedoch nur Personen in der Altersspanne von 25 bis 53 Jahren. Einer Studie [1] zufolge können Menschen das Alter anderer Personen im Durchschnitt etwa auf 8 Jahre genau schätzen. Daher wird davon ausgegangen, dass D Datensätze als potenziell relevant identifizieren kann, die sich im Geburtsjahr nicht mehr als 8 Jahre vom Geburtsjahr von P unterscheiden.
- **Gewicht:** Die mögliche Spanne im Bereich des Gewichts liegt in den Altersklassen der männlichen Beschäftigten etwa zwischen 65 und 110 kg [3], im Datensatz von WearHealth sind Werte von 60 bis 110 kg vertreten. Untersuchungen [4] zufolge können Dritte das Gewicht einer Person durchschnittlich mit einer Abweichung von etwa 13 bis 15 kg schätzen. Folglich gehen wir davon aus, dass D ebenfalls Datensätze als potenziell relevant identifizieren kann, die sich nicht mehr als 15 kg von Ps Gewicht unterscheiden.
- **Größe:** In der Altersklasse aller Beschäftigten sind Größenwerte zwischen etwa 167 und 190 cm üblich [3], im Datensatz von WearHealth sind Personen mit einer Größe von 165 bis 198 cm vertreten. Mithilfe eines Fotos ist es Menschen einer Studie [4] zufolge durchschnittlich möglich, die Größe von Menschen mit einer Abweichung von etwa 2 bis 6 cm zu bestimmen.

Begegnen sich D und P sogar, so könnte D die Größe von P unter Umständen noch exakter bestimmen (etwa im Vergleich zur eigenen Körpergröße). Daher wird davon ausgegangen, dass D Datensätze bestimmen kann, die nicht mehr als 2 cm von der Größe von P abweichen.

Um die Widererkennung von P zu erschweren, können zwei verschiedene Ansätze gewählt werden.

Variante A: Zum einen kann der Datensatz von P derart verändert werden, dass seine Profildaten nicht mehr in dem Bereich liegen, den ein Angreifer durch Schätzen als mögliche Profildaten von P identifizieren würde. Kann die Körpergröße etwa bis auf 2 cm genau geschätzt werden, könnte die Größe von P in den Daten um mehr als 2 cm verändert werden, um so als nicht mehr relevant eingestuft zu werden. Dieses Vorgehen bietet jedoch keine absolute Sicherheit: Handelt es sich bei P um eine Person, deren Profildaten gravierend von den restlichen Personen in der Gesamtgruppe abweichen (bspw. ein Schulpraktikant in einer Gruppe Bauarbeiter mittleren Alters), so könnte ein Angreifer dennoch recht einfach die Daten von P erraten, wenn er bloß den Datensatz auswählt, der P am ähnlichsten zu sein scheint.

Variante B: Zum anderen ist es möglich, die Daten aller Gruppenmitglieder derart anzupassen, dass es stets mehrere Personen gibt, deren Daten identisch (k-Anonymität) oder zumindest so ähnlich sind, dass sie allein durch Schätzen eines Angreifers nicht eindeutig P zugeordnet werden können. Für einen Angreifer würde es also stets mehrere Personen geben, die P am ähnlichsten erscheinen, auch wenn die ursprünglichen Daten von P erheblich von der Gesamtgruppe abweichen. Dies birgt jedoch einen Nachteil in sich: Sowohl die Profildaten von P als auch die Profildaten der weiteren Personen, die P ähneln sollen, müssen so sehr angepasst werden, dass sie allein durch Schätzen nicht unterscheidbar sind. Liegen die Profildaten der jeweiligen Personen weit auseinander, müssen sie äußerst stark verändert werden, was eine gravierende Verschlechterung der Analysequalität des Dienstes mit sich bringen kann. Liegen die Daten der Gesamtgruppe jedoch alle recht nah aneinander, müssen sie nur leicht angepasst werden und die Analysequalität wird nur wenig beeinträchtigt.

Bezugnehmend auf die Analysen aus Abschnitt 2.1.1 soll nun bestimmt werden, wie stark die einzelnen Daten verfremdet werden können. Nach Möglichkeit sollen dabei mehrere Datenkategorien zugleich verändert werden, daher fallen alle Verfremdungsvarianten weg, die nur eine Datenkategorie ändern. Bezüglich der übrigen Varianten sollte eine Abwägung durchgeführt werden: Für die Umsetzung von Variante A sollten die verfremdeten Werte möglichst nah am Rand der Spanne sein, die D etwa schätzen würde oder diese sogar überschreiten. Für Variante B kann einerseits betrachtet werden, wie viele andere Personen Werte in diesem Bereich haben, andererseits kann auch die Spanne an insgesamt möglichen Werten betrachtet werden. So liegen etwa die möglichen Alterswerte in einer Spanne von 28 Jahren. Wenn das Verfremden Werte in einem großen Bereich dieser Spanne ermöglicht, steigt auch die Wahrscheinlichkeit, mit den Daten anderer Personen zusammenzufallen.

Aus den Analysen aus Abschnitt 2.1.1 ergeben sich mehrere Möglichkeiten der Veränderung der Profildaten, welche eine Umsetzung der beiden Varianten ermöglichen und etwa dieselbe Fehlerrate aufweisen. Besonders datenschutzfreundlich wäre es daher, wenn alle zutreffenden Varianten in der Erstellung der Klassen ermöglicht werden. Denn wenn es nur eine exakte Variante zur Veränderung der Daten gäbe, könnte D, wenn er Kenntnis über diese Variante erhält, relativ einfach die veränderten Daten auf die Originaldaten zurückführen. Daher ist es empfehlenswert, alle datenschutzfreundlichen Varianten zu ermöglichen und somit für D ein einfaches Zurückführen der Daten zu verhindern. Wichtig

ist dabei auch, dass Datenverfremdungen im gesamten Wertebereich möglich sein sollen und nicht nur bis auf die Grenzen des Wertebereichs.

2.2 Schutz der Messdaten mithilfe von Differential Privacy

Die Messdaten können ebenfalls zur (Re-)Identifizierung von Personen genutzt werden. Die Bedrohung besteht dabei durch einen Angreifer, der Vergleichsdaten seines Opfers zur Verfügung hat. Dabei könnte es sich etwa um den Manager M eines Arbeitsteams handeln, der vermutet, dass eine bestimmte Person P in jedem Gruppenbericht die Durchschnittswerte der Gruppe durch ihre schlechten Leistungen herabsetzt. Erhält M nun Zugang zu den Messdaten der schlechtesten Person S seiner Gruppe (etwa durch ein Datenleck oder Kollaboration mit einem Analysedienstmitarbeiter) und verschafft sich Zugang zu weiteren Beschleunigungsdaten von P (etwa über öffentlich verfügbar gemachte Daten in einer Fitness-App), so könnte M herausfinden, ob S und P dieselbe Person sind.

Um dies zu verhindern, sollen die vom Analysedienst verarbeiteten Messdaten der Beschleunigung und Herzratenvariabilität durch Differential Privacy geschützt werden. Die Anwendung von Differential Privacy führt jedoch zu Veränderungen der Rohdaten, die eine Verschlechterung des Analysequalität bewirken können. Um die Auswirkung dieser Datenveränderung auf für den betrieblichen Arbeitsschutz relevanten Handlungen einerseits und die Identifizierung von Personen andererseits zu untersuchen, wird im Folgenden sowohl für die Beschleunigungsdaten als auch die Herzraten-variabilitätsdaten jeweils ein Anwendungsfall für beide Aspekte betrachtet und daraus eine konsolidierte Empfehlung für die datenschutzfreundliche Nutzung im Projekt abgeleitet.

2.2.1 Datentypen

Beschleunigungsdaten beinhalten die Beschleunigung eines Wearables von einem Zeitpunkt zum nächsten. Sie bestehen daher aus den Beschleunigungen in x-, y- und z-Richtung seit dem letzten Messzeitpunkt. Aus diesen Daten lassen sich dann sowohl die zurückgelegte Strecke des Wearables im Messzeitraum als auch konkrete Bewegungen ableiten.

Die Herzratenvariabilität (HRV) beschreibt die Variation zwischen aufeinander folgenden Herzschlägen [7]. HRV-Analysen arbeiten mit der Information, wie sehr sich die Zeiten von einem Herzschlag zum nächsten (RR-Intervall oder auch NN-Intervall) über eine gewisse Zeitspanne unterscheiden. Da es keinen konkreten Messwert für die Herzratenvariabilität als solches gibt, wird sie anhand der Zeiten zwischen den Herzschlägen erhoben und später durch ein passendes Maß ausgewertet. Die Auswertung kann mithilfe verschiedener Methoden stattfinden, die bekanntesten Gruppen von Methoden sind dabei *Time-Domain*-Methoden und *Frequency-Domain*-Methoden. Oft genutzte Maße sind etwa die Standardabweichung aller NN-Intervalle (SDNN) oder die Standardabweichung der Differenzen zwischen benachbarten NN-Intervallen (SDSD).

2.2.2 Anwendungsfall 1: Identifizierung von Stürzen

Ein sinnvoller Anwendungsfall für die Nutzung von Wearable-Daten im Arbeitskontext ist die Sturzerkennung: Findet eine schnelle Bewegung des Körpers nach unten statt und keine anschließende Bewegung mehr nach oben (Bewegungen auf der z-Achse), so kann es zu einem Sturz gekommen sein. Das Erkennen von Stürzen im betrieblichen Kontext kann etwa vorteilhaft sein, wenn Personen in gefährlichen Bereichen oder isoliert von anderen Personen arbeiten. Werden Stürze erkannt, können andere Personen auf die mögliche Notlage hingewiesen werden und Betroffenen zur Hilfe kommen.

Auch auf die Herzratenvariabilität kann dies übertragen werden: Studien (beispielsweise von Melilo et al. [8]) haben belegt, dass HRV-Daten für die Erkennung von Stürzen genutzt werden können. Daher wird im Folgenden der Anwendungsfall »Sturzerkennung« als ein beispielhafter Anwendungsfall für den betrieblichen Kontext betrachtet.

Die Beschleunigungsdaten für diesen Anwendungsfall wurden erstmals in den Arbeiten von Nho et al. [5] vorgestellt. Sie analysieren Beschleunigungsdaten für verschiedene Sturzszenarien von 21 Personen, die durch ein Wearable am linken Handgelenk gemessen wurden. Die Daten wurden mit einer Frequenz von 50 Hertz erhoben. Der Datensatz beinhalten ebenfalls die dazu gehörenden Herzratenvariabilitätsdaten, welche mit einer Frequenz von 50 Hertz erhoben wurden.

2.2.3 Anwendungsfall 2: Identifizierung von Personen

Um die Identifizierung von Personen über Wearable-Daten zu überprüfen, soll ein anderes Szenario gewählt werden.[1] Da in der Nutzung von Wearables zur Aufzeichnung von Wearable-Daten nicht nur arbeitsrelevante Handlungen aufgezeichnet werden, sondern auch alltägliche Situationen, sollen diese nun auf ihre Identifizierbarkeit untersucht werden.

Dazu wurde für die Beschleunigungsdaten der Anwendungsfall »Händewaschen« ausgewählt, da davon auszugehen ist, dass diese Aktivität einerseits sehr wahrscheinlich während der Arbeitszeit auftreten wird und andererseits die Arbeitnehmer sich wahrscheinlich nicht dazu entscheiden werden, stets die Datenübertragung zu unterbrechen, wenn sie ihre Hände waschen möchten. Diese Daten liegen im Datensatz von Climent-Pérez et al. [2] vor und wurden über ein Wearable am Handgelenk von 52 Studienteilnehmern mit einer Frequenz von 32 Hertz gemessen.

Für die Herzratenvariabilität wurden zwei Datensätze genutzt, welche Daten in Ruhe- sowie auch in unterschiedlichen Stressphasen umfassen. Der erste Datensatz von Irurzun et al. [9] umfasst 147 Personen im Alter von einem Monat bis 55 Jahren mit einer Sampling Rate von 128 Hz. Da diese Alterspanne nicht dem Anwendungsfall entspricht, wurden für die folgende Analyse nur die Personen zwischen 18 und 55 Jahren berücksichtigt (33 Personen). Im zweiten Datensatz von Koldijk et al. [10] sind 25 Personen vertreten, deren Daten mit einer Frequenz von 2048 Hz erfasst wurden.

2.2.4 Analyse

Ziel der folgenden Untersuchungen ist es, herauszufinden, ob sich Differential Privacy für die Beschleunigungs- und Herzratenvariabilitätsdaten im Projektkontext eignet. Konkret beinhaltet das die Überprüfung, ob Differential Privacy derart parametrisiert werden kann, dass einerseits die Privatsphäre der Nutzer geschützt und andererseits eine akzeptable Analysequalität für arbeitsrelevante Dienste ermöglicht werden kann. Die beiden Probleme der Sturzerkennung und der Identifizierung werden dabei jeweils als binäre Entscheidungen definiert. In der Sturzerkennung soll bei einem gegebenen Datenpunkt entschieden werden, ob es sich um einen Sturz handelt, bei der Identifizierung soll zu einem gegebenen Datenpunkt entschieden werden, ob es sich um die Daten einer im Vorhinein zufällig festgelegten Person handelt oder nicht.

[1] Eine Betrachtung der Sturzdaten für die Identifizierung erscheint als nicht zielführend: Die reine Information, dass eine Person gefallen ist, kann bereits identifizierend sein, wenn nur ein einziger Arbeitnehmer gefallen ist. Diese Art der Identifizierbarkeit kann für eine sinnvolle Nutzung im Arbeitskontext jedoch nicht entfernt werden.

Zur Anwendung von Differential Privacy wird die Diffprivlib von IBM [6] verwendet. Die beiden zu belegenden Parameter sind ε und *Sensitivity*. Die Sensitivity gibt an, wie stark sich die Daten von einem Datenpunkt zum Nächsten maximal verändern können. Theoretisch könnte dies eine Änderung vom Minimum bis zum Maximum des jeweiligen Wertebereiches sein. Die Wertebereiche sind jedoch nicht eindeutig bestimmbar und eine Änderung von Minimum auf Maximum des Wertebereichs von einem Zeitintervall zum nächsten ist nicht in allen Daten (bspw. z-Achse der Beschleunigung) realistisch. Für die Analysen wurden daher die maximalen Änderungsdaten der vorhandenen Daten bestimmt und von diesen das 0.95-Quantil als Sensitivity genutzt, um etwaige Ausreißer und Messfehler zu umgehen. Der Parameter ε ermöglicht die Einstellung, wie viele Informationen offengelegt werden und wird auch als *Privacy-Budget* bezeichnet. Ein kleines ε bietet dabei hohen Schutz und legt wenige Daten offen, während ein großes ε wenig schützt und mehr Daten offenlegt.

Die Wahl eines konkreten ε ist vom Anwendungsfall abhängig. Um ein passendes ε zu bestimmen, sollen im Folgenden verschiedene Werte angewandt und eine Abwägung durchgeführt werden, was ein passendes ε für den hier beschriebenen Anwendungsfall sein könnte.

Für die letztendliche Identifizierung und Sturzerkennung wurden verschiedene Algorithmen getestet. Dazu gehörten unter anderem verschiedene lineare Modelle (bspw. Regression), aber auch Ensemble-Modelle (bspw. Decision Trees) und ein One Class Classifier (Isolation Forest). Dazu wurden die verschiedenen Modelle für jede Parameterkonstellation einem Training unterzogen und anhand eines separaten Testdatensatzes der beste Algorithmus ausgewählt. Die Ergebnisse der Analysen mit den verschiedenen ε-Werten sollen dabei stets mit den Ergebnissen ohne etwaige Änderungen an den Daten verglichen werden.

2.2.5 Auswertung

Für die Auswertung der Qualität der verschiedenen Modelle und Algorithmen wurde der F1-Score genutzt. Dieser stellt das harmonische Mittel aus Precision (Anteil der Vorhersagen der relevanten Klasse, die korrekt waren) und Recall (Anteil der korrekten Vorhersagen der relevanten Klasse an allen Vorkommnissen der relevanten Klasse) dar. Ein hoher F1-Score (nahe an 1) zeigt eine hohe Genauigkeit, während ein niedriger Wert (nahe an 0) eine schlechte Genauigkeit anzeigt. Für die Sturzerkennung sind daher hohe F1-Werte wichtig, während ein Schutz gegen Identifizierung besonders niedrige F1-Werte erfordert, um die Privatheit der Personen nicht zu beeinträchtigen.

Betrachten wir nun die Ergebnisse der Analysen für die Beschleunigungsdaten (Abbildung 1). Die orange Linie zeigt den F1-Score, wenn die Daten ungeschützt für die jeweilige Analyse genutzt werden, und die blaue Kurve zeigt den F1-Score für die verschiedenen ε-Werte. Es zeigt sich, dass für beide Datentypen auf ungeschützten Daten gute Ergebnisse erzielt werden können, wobei die Identifizierung etwas bessere Werte aufweist als die Sturzerkennung. Besonders niedrige F1-Scores in der Identifizierung finden sich bei ε-Werten kleiner oder gleich 0,1 (10^{-1}), besonders hohe F1-Scores der Sturzerkennung nahe an den Ergebnissen der ungeschützten Daten sind jedoch erst ab ε-Werten größer 100 (10^{2}) zu finden. Diese beiden Wertebereiche weisen keine Überlappung auf, so dass eine sinnvolle Nutzung von Differential Privacy bei gleichzeitigem Erhalt der Nutzbarkeit schwierig ist.

Die Ergebnisse der Analysen mit den Herzratenvariabilitätsdaten (Abbildung 2) lassen ähnliche Schlüsse zu. Auch hier zeigt die orange Linie die F1-Scores ohne jeglichen Schutz und die blaue Kurve die Ergebnisse mit Differential Privacy bei der Nutzung verschiedener ε-Werte. Die F1-Scores der Identifizierung zeigen ähnlich wie bei den Beschleunigungsdaten, dass ein ε-Wert kleiner oder gleich

0,1 (10^{-1}) vor einer Identifizierung schützen kann. Die Ergebnisse der Sturzerkennung sind dagegen nicht so aussagekräftig, da auch die ungeschützten Daten keine hohe Qualität in der Sturzerkennung ermöglichen. Die Anwendung von Differential Privacy verringert den F1-Scores nur geringfügig, wobei sich die Kurve ab einem Wert von 10 (10^{1}) den Ergebnissen der ungeschützten Analyse anzugleicht.

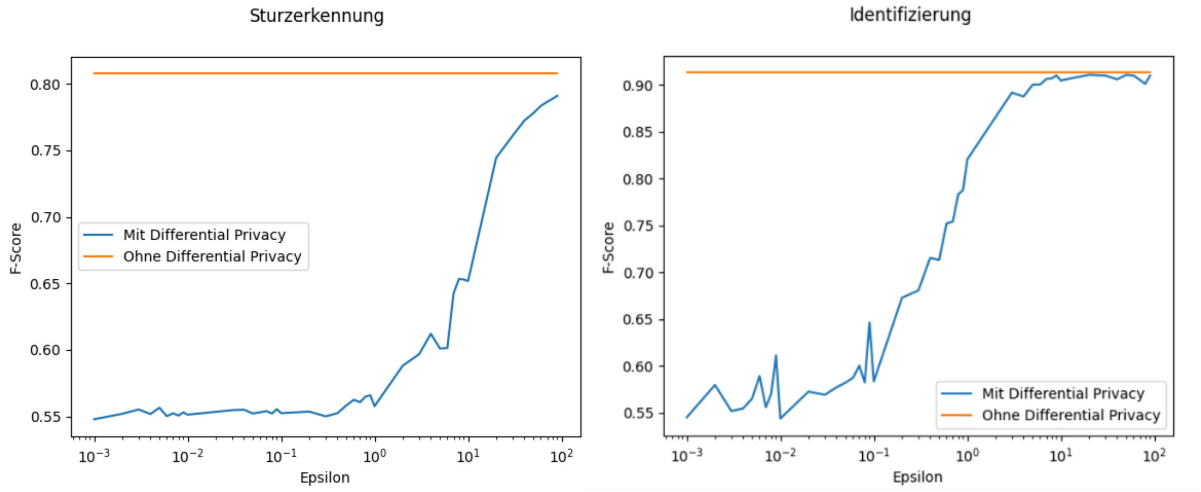

Abbildung 1 Ergebnisse der Analysen für Beschleunigungsdaten (links Sturzerkennung, rechts Identifizierung)

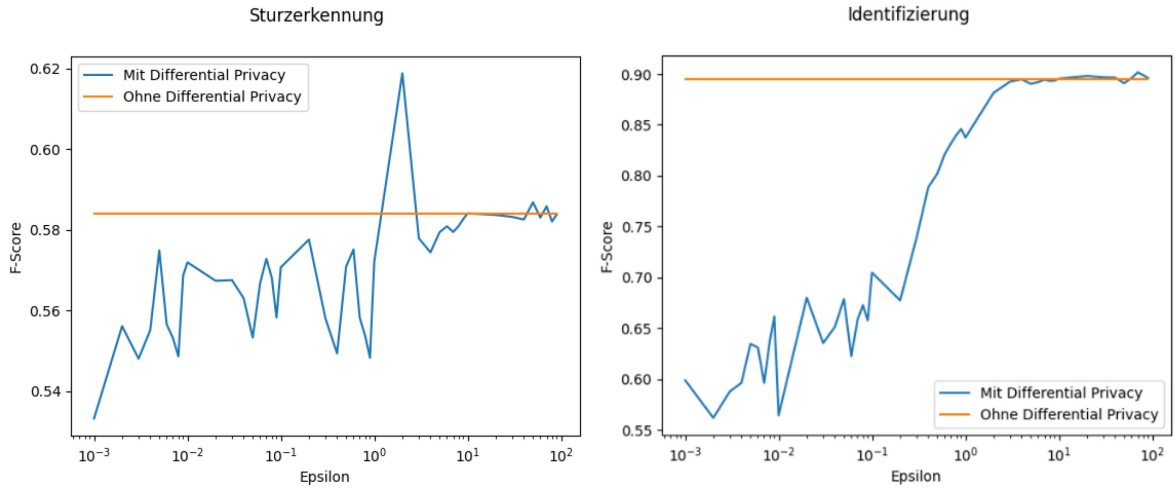

Abbildung 2 Ergebnisse der Analysen für Herzratenvariabilitätsdaten (links Sturzerkennung, rechts Identifizierung)

Legt man die Befunde gemäß Abbildung 2 als Maßstab zugrunde und betrachtet die nachfolgend beschriebenen Evaluationsergebnisse, so zeigt sich, dass die Klassifizierung der Belastung von Wearable-Trägern sehr sensibel auf eine Verfälschung der Herzratenvariabilitätsdaten reagiert. In unseren Experimenten haben wir zum Beispiel die Klassifizierung der unveränderten Rohdaten mit ihrer Klassifizierung nach einer Verfremdung der Herzratendaten verglichen.

Abbildung 3 zeigt ein Beispiel einer solchen Analyse, bei der jeweils die ursprünglich zugewiesene Klasse der neu ermittelten Klasse nach Datenverfremdung gegenübergestellt ist. Im Diagramm kann man für jede Klasse die Anzahl der korrekten positiven und negativen Klassenzuordnungen sowie die

Anzahl der falsch-positiven und falsch-negativen Zuordnungen ablesen. Daraus lässt sich für jede Klasse die Precision und der Recall der Klassenzuordnung berechnen.

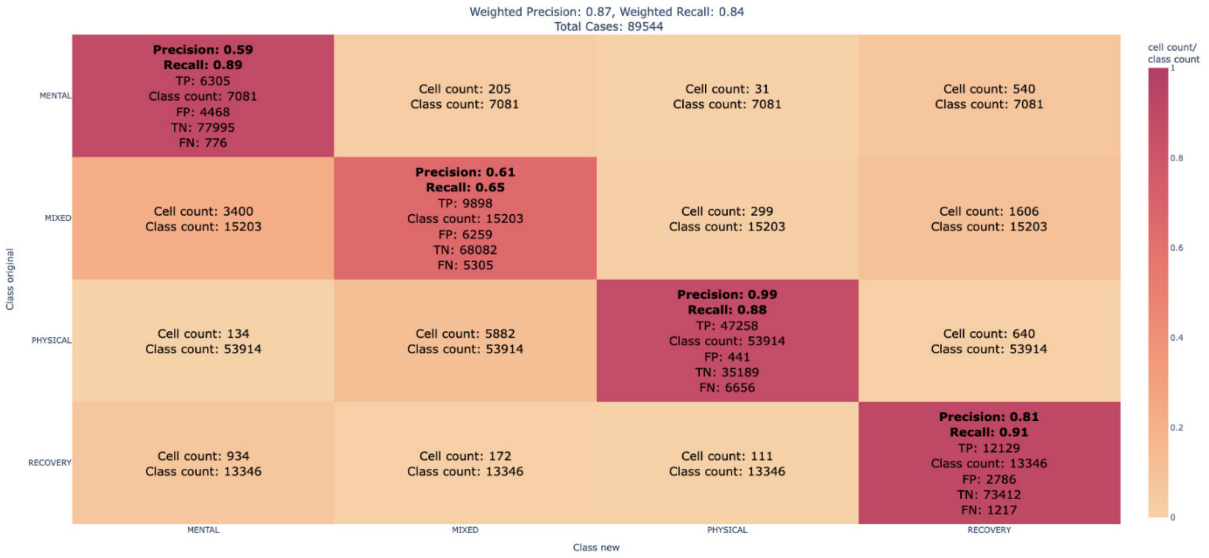

Abbildung 3 Beispiel für die Bewertung des Effekts einer Datenverfremdung der gemessenen Herzratendaten auf die Klassifizierung der persönlichen Workload von Wearable-Trägern. Die Auswertung zeigt, wie die unveränderten Vitaldatenproben (89544 Fälle) ursprünglich klassifiziert wurden (vertikale Achse), und wie deren Klassifizierung nach Anwendung von Differential-Privacy-Techniken geändert hat (horizontale Achse). Die Verfremdung beeinträchtigt den Recall und die Precision der Klassifizierung.

Solche Messungen wurden im Projekt für verschiedene Verfremdungsstärken – charakterisiert durch die Wahl von ε – vorgenommen. Abbildung 4 zeigt die Befunde für die Klassifizierungsaufgabe gemäß Abbildung 3. Schon bei milder Datenverfremdung mit ($10 \leq \varepsilon \leq 1000$) sinken Precision und Recall der für unverfälschte Rohdaten ermittelten Workload-Klassen merklich; ab $\varepsilon \leq 0{,}9$ nehmen die Klassifizierungsfehler rapide zu.

Auch die Wirkung der Verfremdung auf die Bewertung einer Belastungsstärke als LOW, MODERATE oder HIGH wurde für unsere Demonstratoranwendung ermittelt. Abbildung 5 zeigt das Ergebnis unserer Messungen. In diesem Fall ist eine Verfremdung mit einem $\varepsilon \geq 0{,}8$ noch vertretbar. Danach sinkt die Analysequalität merklich.

Für unsere Demonstratoranwendung ergibt sich somit, dass für die Verfremdung der Herzraten-variabilitätsdaten ein $\varepsilon \geq 0{,}9$ gewählt werden sollte, um die Analysequalität in Bezug auf die Belastungsklassifizierung nicht zu sehr zu beeinträchtigen. Die Beurteilung der Belastungsstärke reagiert nur geringfügig besser auf eine ε-Verfremdung der Daten. Der Spielraum für Differential-Privacy, der sich daraus ergibt, bietet laut dem rechten Diagramm in Abbildung 2 jedoch nur einen geringen zusätzlichen Schutz vor einer Identifizierung der Probanden. Bei diesem Befund ist allerdings zu berücksichtigen, dass gerade die Herzratenvariabilität sehr empfindlich von den gemessen RR-Intervallen abhängt, so dass für eine HRV-Bestimmung zu erwarten ist, dass überlagerte Rauschen schnell zu Fehleinschätzungen führt.

Zudem ist zu berücksichtigen, dass mit ε zwar die Stärke der Datenverfremdung charakterisiert wird, nicht aber das genaue Verfremdungsverfahren. Das von uns verwendete Verfahren ist im Ergebnisbericht D3.3 genauer beschrieben. Es zeigt sich, dass nach einer Verfremdung zur

Konsistenzwahrung verschiedene kleinere Anpassungen an den Daten vorgenommen werden müssen, etwa, damit der übermittelte Puls-Wert noch zu den gemessenen Herzraten-Intervallen passt. Für diese Konsistenzmaßnahmen gibt es verschiedene Möglichkeiten, so etwa das Weglassen bzw. Einfügen von Messwerten oder das Stauchen beziehungsweise Strecken der zuvor verfremdeten Intervalle. Daher ist es denkbar, dass mit einem anderen als dem von uns implementierten Verfremdungsverfahren die Analysequalität besser gewahrt werden kann, was einen stärkeren Schutz gegen Identifizierung von Probanden ermöglichen würde.

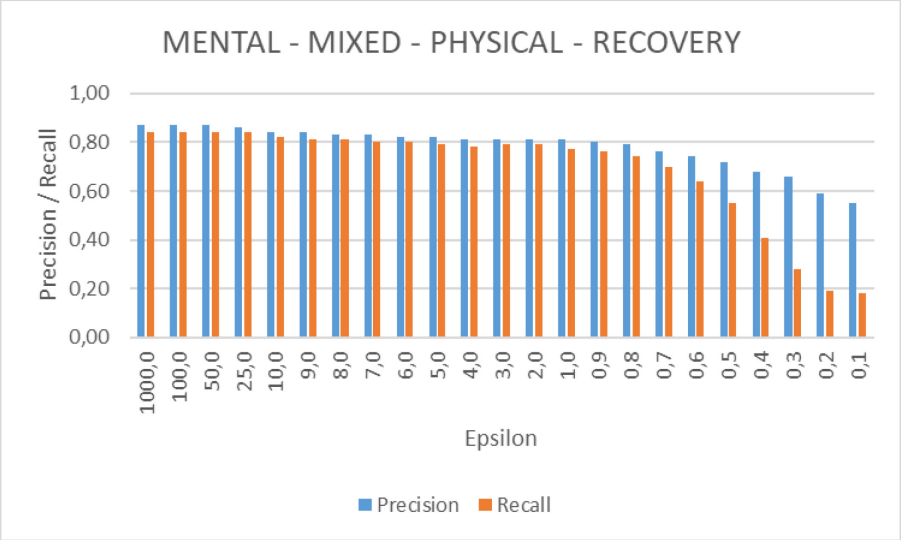

Abbildung 4 Bewertung Effekts einer Datenverfremdung der gemessenen Herzratendaten auf die Klassifizierung der persönlichen Workload von Wearable-Trägern nach Belastungsart. Die Darstellung zeigt die erzielte gewichtete Precision und den erzielten gewichteten Recall der Klassifizierung gemäß der Klasseneinteilung aus Abbildung 3 für verschiedene Werte von ε im Bereich zwischen 0,1 und 1000. Für starke Datenverfremdung ($\varepsilon \leq 0{,}9$) sinkt die Analysequalität deutlich; insbesondere der Recall bricht ein.

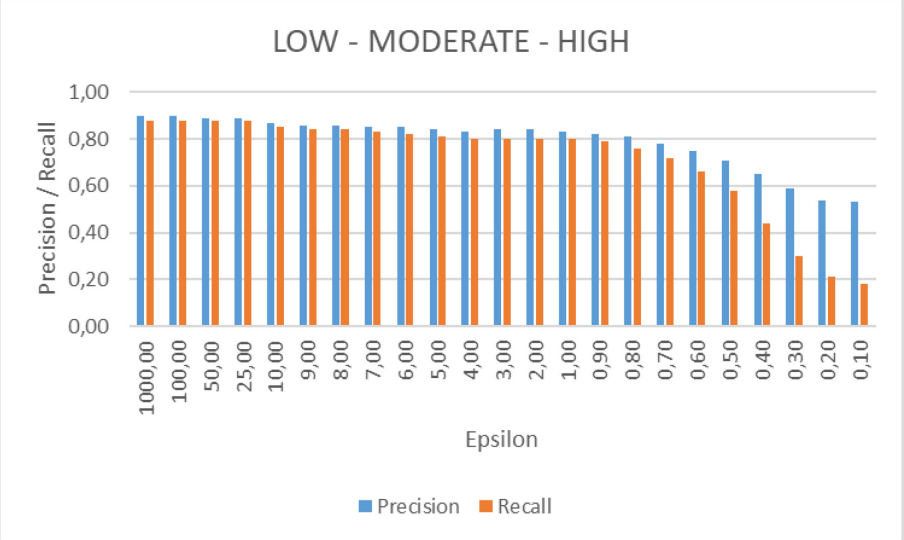

Abbildung 5 Bewertung Effekts einer Datenverfremdung der gemessenen Herzratenvariabilitätsdaten auf die Bewertung der persönlichen Belastungsstärke als LOW, MODERARTE oder HIGH. Für starke Datenverfremdung ($\varepsilon \leq 0{,}8$) sinkt die Analysequalität deutlich; wie in Abbildung 4 bricht der Recall rapide ein.

Die von uns ermittelten Spielräume für die Anwendung von Differential Privacy als Anonymisierungsmaßnahme bei HRV-Messungen sind darüber hinaus auch insoweit eine eher konservative Schätzung, als der Anonymisierungseffekt gemäß Abbildung 2 auf Auswertungen über einer geringen Zahl von Probanden beruht: Laut Abschnitt 2.2.3 wurden nur Daten von 58 Personen zugrunde gelegt. Es ist naturgemäß leichter, ein HRV-Profil in einer kleinen Menge von Kandidaten wiederzufinden als in einer sehr großen Grundgesamtheit. Wenn also ein Analysedienstleister HRV-Messungen von sehr vielen Kunden verarbeitet, wird es einem Angreifer schwerer als in unseren Experimenten fallen, einen Vergleichsdatensatz eindeutig einem bestimmten Probanden zuzuordnen. In diesem Fall ist bei einer Datenverfremdung mit $\varepsilon \approx 1$ voraussichtlich mit einem deutlich besseren Anonymisierungseffekt zu rechnen.

Die Analysekomponente des WearPrivate-Demonstrators basiert auf künstlicher Intelligenz (KI). Um den Analysedienst robuster gegenüber Datenverfremdungen zu machen, könnte man die KI gezielt mit verfremdeten Daten trainieren. In unseren Experimenten zeigte sich nämlich, dass die Verfremdung durch Differential-Privacy-Techniken die Herzraten-Variabilität tendenziell erhöht, was von dem Analyseverfahren meist als »geringere Belastung« interpretiert wird. Das bedeutet, dass der Belastungsindex bei verfremdeten Daten systematisch etwas zu gering eingeschätzt wird, wie die Beispielmessung in Abbildung 6 verdeutlicht. Teilt man der Analyse also mit, in welchem Maße die Messdaten verfremdet sind, so kann das KI-Verfahren die dadurch verursachte Fehlertendenz bei geeignetem Training zumindest teilweise kompensieren. Im Projekt fehlte leider die Zeit, diese Idee weiterzuverfolgen. Die vorläufigen Messungen weisen jedoch klar darauf hin, dass hier noch Verbesserungspotenziale erschlossen werden können.

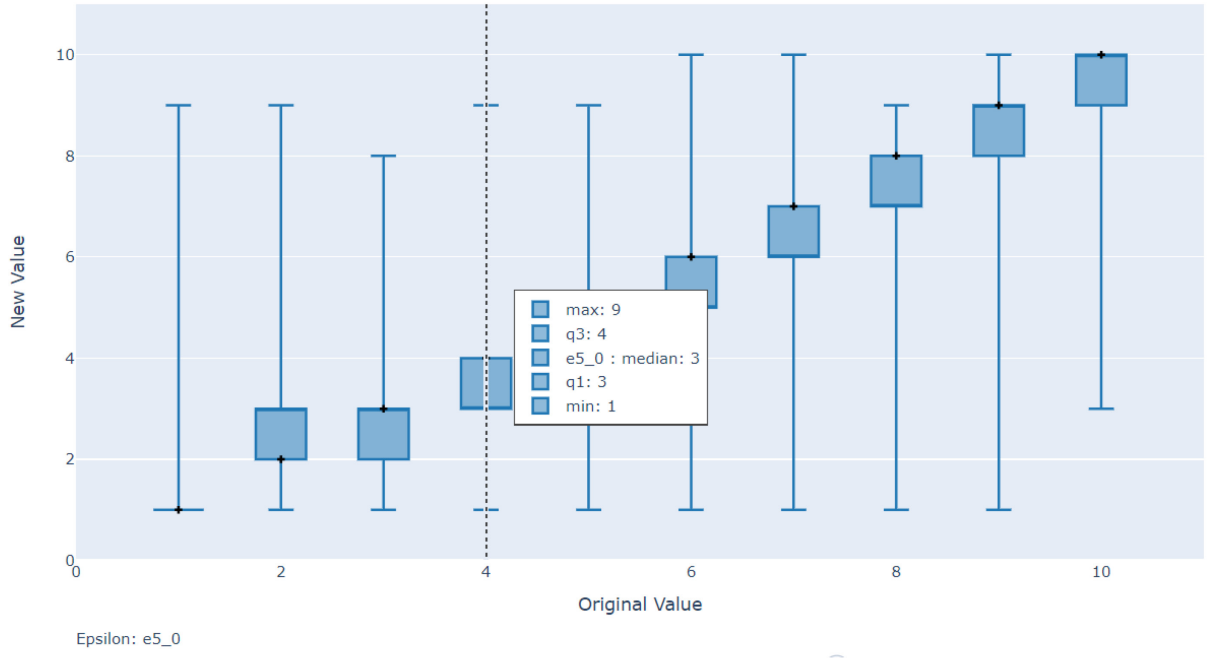

Abbildung 6 Beispiel für die systematische Verschiebung des ermittelten Belastungsindex [1 … 10] bei Verfremdung der Daten (Im Beispiel: $\varepsilon = 5$). Ähnliche Verschiebungen wurden im Projekt für unterschiedliche Werte von ε ermittelt.

3 Evaluation des Interaktionskonzepts

Die Interaktion mit einer Wearable-basierten Gesundheits-App soll möglichst reibungslos und verständlich möglich sein. Im Alltag darf sie die App-Nutzenden nicht überfordern, nicht zu viel Aufmerksamkeit im Tagesablauf beanspruchen und nicht zu viel Zeit und Mühe kosten. Gerade im beruflichen Umfeld darf die Interaktion mit dem System nicht selbst zum Stressfaktor oder gar zu einer Gefahrenquelle mutieren.

Darüber hinaus soll das Interaktionskonzept aber auch einladend und ansprechend sein. Im Idealfall animiert es zur Nutzung der Gesundheits-App und belohnt die Wearable-Träger in irgendeiner Form für ihr Tun.

Die Nutzerschnittstelle und die damit realisierten Interaktionsmöglichkeiten mit dem System sind ein entscheidender Baustein, um dem potenziellen Anwenderkreis die Nutzung des Systems schmackhaft zu machen und tatsächliche Nutzerinnen und Nutzer längerfristig zum Mitmachen zu bewegen. Im Folgenden beschreiben wir unsere Ansätze, mit denen wir das Interaktionskonzept unseres Demonstrators evaluiert haben.

3.1 Studienüberblick

Im Rahmen des Projekts WearPrivate wurde ein Demonstrator entwickelt, der den Stress am Arbeitsplatz erfassen und den Nutzern eine Hilfestellung zur Stressbewältigung geben soll. Teil dieses Demonstrators ist die *HealthyWork*-App. Um die Usability dieser App als Teil des Gesamtsystems zu untersuchen, wurde die im Folgenden beschriebene Studie durchgeführt.

Die Hauptziele dieser Studie bestanden darin, die Bedienbarkeit der App, die Genauigkeit der Stressmessungen sowie die Nutzerzufriedenheit zu analysieren. Zusätzlich wurden Interviews durchgeführt, um tiefere Einblicke in die Nutzererfahrungen zu gewinnen. Diese Ergebnisse wurden mit den Daten aus den Fragebögen und Tabellen korreliert, um zu einer Einschätzung der Bedienbarkeit der App zu gelangen sowie um Empfehlungen für deren Weiterentwicklung zu gewinnen.

Im Vorfeld der eigentlichen Studie wurde bereits eine separate Umfrage mit unterschiedlichen Teilnehmern durchgeführt. Diese Umfrage untersuchte die allgemeine Einstellung gegenüber Wearable-Technologien, jedoch ohne die Nutzung eines spezifischen Demonstrators. Die Teilnehmer dieser Vorbefragung hatten somit keine Gelegenheit, die Technologie selbst zu testen, sondern äußerten ihre Meinungen basierend auf theoretischen Vorstellungen und bisherigen Erfahrungen mit anderen Produkten. Diese Ergebnisse lieferten wertvolle Vorabinformationen, die in der Nutzerstudie, in der die App direkt getestet wurde, weiter untersucht wurden.

3.2 Teilnehmer und Methodik

Die Experteninterviews wurden mit vier männlichen Teilnehmern im Alter von 20 bis 50 Jahren aus verschiedenen Branchen, darunter Sicherheit, Pflege und Industrie, durchgeführt. Die Teilnehmer bekamen den Demonstrator im Rahmen der Studie zunächst als Klick-Dummy präsentiert. Dabei wurde ihnen die grundlegende Funktionsweise der App anhand eines Prototyps veranschaulicht, der wesentliche Features und Interaktionen zeigte, jedoch noch keine vollständige Funktionalität oder

Echtzeit-Stressmessungen bot. Nach dieser Demonstration beantworteten die Probanden Fragen zur Benutzerfreundlichkeit und zum Datenschutz. Diese Fragen wurden später erneut aufgegriffen, um Veränderungen in den Einschätzungen nach der Demonstration zu erfassen und zu evaluieren, wie der Prototyp das Vertrauen der Probanden in die Technologie beeinflusst hat.

Die Entscheidung für qualitative Interviews als Methode basierte darauf, tiefere Einblicke in die persönlichen Meinungen, Bedenken und Akzeptanzkriterien der Teilnehmer zu gewinnen, was durch standardisierte Fragebögen allein nur eingeschränkt möglich gewesen wäre. Durch die direkte Interaktion mit dem Demonstrator konnten die Teilnehmer spezifische Aspekte der App nachvollziehen und im Gespräch detailliert erläutern, welche Funktionen für sie besonders wichtig oder problematisch waren. Diese Methodik fördert ein umfassenderes Verständnis der individuellen Nutzerbedürfnisse und erlaubt es, Optimierungsmöglichkeiten gezielt zu identifizieren.

In der angeschlossenen Nutzerstudie wurde die tatsächliche Funktionsweise des Demonstrators von einem Mitarbeiter (Haupttätigkeit am Schreibtisch, Altersgruppe 40-45) der NMS unter Bedingungen untersucht, die den realen Einsatz im Alltag widerspiegeln. Indem der Proband die App zur Stressmessung in praxisnahen Szenarien nutzte, konnten Rückmeldungen zur Bedienbarkeit, Messgenauigkeit und Effizienz der App gesammelt werden. Die gewählten Aufgaben zielten darauf ab, die App auf ihre Alltagstauglichkeit zu prüfen und zugleich herauszufinden, ob und wie sie den Nutzern hilft, ihr Stresslevel zu erkennen und zu managen.

3.3 Ergebnisse aus der Studie

In der Befragung bewertete der Proband die nachfolgenden Aspekte jeweils auf einer Skala von 1 (stimme überhaupt nicht zu) bis 5 (stimme vollkommen zu).

3.3.1 Benutzerfreundlichkeit und Effizienz

- Bewertung der Aussage: *"Das System war einfach zu bedienen"* **5/5**
- Bewertung der Aussage: *"Ich konnte die Aufgaben effizient und ohne größere Schwierigkeiten bewältigen"* **5/5**

Diese Höchstbewertungen unterstreichen, dass die App intuitiv und leicht zu bedienen ist. Der Proband bewertete sowohl die Benutzerfreundlichkeit als auch die Effizienz als hervorragend. Das deutet darauf hin, dass das Interface der App gut gestaltet ist und Nutzer ohne Schwierigkeiten durch die Aufgabenführung navigieren können.

3.3.2 Genauigkeit der Stressmessungen

- Bewertung der Aussage: *"Die App hat meine Stresslevels präzise erfasst"* **3/5**

Hier gab es gemischte Rückmeldungen. Der Proband äußerte mitunter Zweifel an der Genauigkeit. Dies könnte auf Schwächen bei der technischen Datenerhebung oder aber auch bei der KI-gestützten Auswertung hinweisen. Es ist empfehlenswert, die Ursachen hierfür in zukünftigen Studien genauer zu untersuchen.

3.3.3 Verständlichkeit der Benutzeroberfläche

- Bewertung der Aussage: *"Die Benutzeroberfläche der App war klar und verständlich"* **4/5**

Die meisten Studienteilnehmer (Interviewpartner und Praxisproband) fanden die Benutzeroberfläche klar und einfach zu bedienen, was den positiven Gesamteindruck der Benutzerfreundlichkeit unterstreicht. Einige Teilnehmer merkten jedoch an, dass bestimmte Funktionen klarer dargestellt werden könnten.

3.3.4 Anwendung im Alltag

- Bewertung der Aussage: *"Ich würde das System auch in meinem Alltag zur Stressbewertung verwenden"* **3/5**

Diese mittlere Bewertung deutet darauf hin, dass die Nutzer die App als nützlich, aber nicht als unverzichtbar betrachten. Mit weiteren Nutzerstudien könnte herausgearbeitet werden, mit welchen Änderungen oder Erweiterungen die App attraktiver für die alltägliche Benutzung werden würde.

3.3.5 Erkennung und Verständnis des Stresses

- Bewertung der Aussage: *"Das System half mir, meinen Stress besser zu erkennen und zu verstehen"* **2/5**

Die niedrige Bewertung bei dieser Aussage zeigt, dass die App dem Nutzer nicht ausreichend half, seinen Stress klarer zu verstehen. Zur Weiterentwicklung des Systems sollte zunächst genauer untersucht werden, warum die Nutzerzufriedenheit mit diesem Punkt so gering ist.

3.3.6 Allgemeine Zufriedenheit

- Bewertung der Aussage: *"Insgesamt war ich zufrieden mit meiner Erfahrung mit dem System"* **3/5**

Die Zufriedenheit lag im mittleren Bereich. Diese neutrale Haltung spiegelt wider, dass die App zwar nicht enttäuschte, aber auch nicht die Erwartungen übertraf. Die Bewertung weist darauf hin, dass vor allem die Genauigkeit der Stressmessung und die Nützlichkeit im Alltag weiter optimiert werden sollten.

3.4 Ergebnisse der Interviews

Die Interviews bieten tiefere Einblicke in die individuellen Erfahrungen und Bedenken der Nutzer im Umgang mit der App und ihrer Funktionsweise. Befragt wurden vier Personen mit unterschiedlichem beruflichen Hintergrund.

3.4.1 Vertrauen in die Lösung und Datenschutz (Interview A und D)

Proband A (Sicherheitsexperte) und Proband D (Pflegefachkraft) betonten beide die Bedeutung von Datenschutz und Kontrolle über die persönlichen Daten. Beide gaben hohe Bewertungen für die Möglichkeit, selbst zu entscheiden, welche Daten erfasst und weitergegeben werden dürfen

Proband C (Industriearbeiter) teilte ebenfalls diese Bedenken, war jedoch grundsätzlich offen für die Nutzung der App, insbesondere durch den anonymisierten Anmeldeprozess und die Möglichkeit, die Datenerfassung zu steuern.

Diese Aussagen lassen darauf schließen, dass Datenschutz eine entscheidende Rolle bei der Akzeptanz solcher Technologien spielt. Auch in der Tabellenauswertung spiegeln sich diese Aspekte wider, da die Nutzer vor allem aufgrund der Anonymität und der Kontrolle über ihre Daten Vertrauen schöpften.

3.4.2 Genauigkeit und Funktionalität der Messungen (Interview B und C)

Proband B (Health & Safety Manager) äußerte Skepsis gegenüber der Genauigkeit der Wearable-Technologie, insbesondere in Bezug auf seine spezifische Arbeitsumgebung. Er gab an, dass er die Technologie zwar als nützlich ansieht, aber Zweifel an der praktischen Umsetzung in seinem maritimen Arbeitsumfeld hat.

Proband C betonte, dass die Verrauschung der Daten (zum Schutz der Privatsphäre) positiv bewertet wurde, jedoch auch zur Unsicherheit über die Genauigkeit der Analysen führte.

Diese Bedenken zur Genauigkeit spiegeln sich in der 3/5-Bewertung der Frage zur Stress-messgenauigkeit in den Fragebögen wider. Es wird deutlich, dass die App hier noch Optimierungs-potenzial aufweist, um das Vertrauen in die erfassten Daten zu steigern.

3.4.3 Anwendungspotenzial und Nutzerfreundlichkeit (Interview A und C)

Proband A und C betonten, dass die App gut gestaltet und einfach zu bedienen sei. Sie empfanden die Transparenz bei der Datenverarbeitung als positiv und vertrauten der Technologie mehr durch die Möglichkeit, den Gruppenbericht selbst einzusehen, den ihre Vorgesetzten erhalten.

Diese Einschätzungen korrelieren mit den hohen Bewertungen zur Benutzerfreundlichkeit in den Tabellen (5/5). Die einfache Bedienung der App wurde von den meisten Teilnehmern als eine ihrer Stärken hervorgehoben.

3.5 Zusammenfassung und Empfehlungen

Die Usability-Studie – ergänzt durch qualitative Interviews – zeigt, dass die App eine solide Grundlage hinsichtlich Benutzerfreundlichkeit und Effizienz bietet. Gleichzeitig wurde deutlich, dass die Genauigkeit der Stressmessungen sowie der Mehrwert der App für den Alltag weiter verbessert werden müssen. Der Datenschutz und die Kontrolle über die eigenen Daten erwiesen sich als zentrale Faktoren, die das Vertrauen der Nutzer maßgeblich beeinflussten.

Besonders auffällig war, dass der Klickprototyp von den Teilnehmern etwas besser bewertet wurde als die voll funktionsfähige Version der App. Die reduzierte Interaktion und klar strukturierte Darstellung im Klick-Dummy hatten eine positive Wirkung, da sie es den Nutzern erlaubten, sich direkt auf die Benutzeroberfläche, die Navigation und vor allem auf das Datenschutzkonzept zu konzentrieren. Der Klick-Dummy ermöglichte den Teilnehmern eine ungestörte, visuelle Erfahrung ohne Ablenkung durch technische Details oder Echtzeitmessungen. Dieser Fokus auf die Interaktion mit den Datenschutz-funktionen, wie die individuelle Kontrolle und die Transparenz bei der Datenerfassung, deutet darauf hin, dass der Datenschutz ein entscheidender Erfolgsfaktor für die Nutzerakzeptanz ist.

Ein weiterer Grund für die bessere Bewertung des Klick-Dummies könnte daran liegen, dass die bewertenden Teilnehmer in diesem keine Erfahrungen mit Stresshinweisen machen konnten, die zu einem mehr oder weniger nachvollziehbaren Zeitpunkt eintrafen, wie es mit der App geschah. Es ist vorstellbar, dass dieser technische Aspekt in der Vorstellung der Teilnehmer beim Klick-Dummy einfach „perfekt“ funktionierte. In der praktischen Nutzung waren insbesondere der Messwert der mentalen Belastung sowie die zugehörigen Warnungen nicht immer intuitiv zu den aktuell durchgeführten Aktivitäten zuzuordnen.

Zusammengefasst lässt sich erkennen, dass die für die Anwender am deutlichsten sichtbare Komponente des *WearPrivate*-Gesamtsystems – die *HealthyWork*-App – durch ihre als ansprechend und gut bedienbar bewertete Oberfläche eine wichtige Rolle darin spielt, den Nutzern einen selbstbestimmten Umgang mit ihren Daten zu ermöglichen. Gleichzeitig zeigen die Bewertungen der Genauigkeit der Messungen sowie der Erkenntnis und dem Verständnis des Stresses, dass das *WearPrivate*-System in diesen Bereichen noch verbessert werden kann.

Es ist empfehlenswert, zu diesen Themen weitergehende Untersuchungen durchzuführen, um die Gründe für die hier beobachteten Probleme genauer zu identifizieren, damit diese anschließend behoben werden können. Beispielsweise ist aktuell unklar, ob die empfundene Ungenauigkeit in der Belastungsmessung eine tatsächliche Ungenauigkeit ist, die durch technische Maßnahmen verbessert werden könnte, oder nur eine empfundene Ungenauigkeit. Insbesondere mentale Belastungen könnten für die betroffenen Personen in dem Moment des Geschehens gar nicht als solche wahrgenommen werden. Die genauere Analyse solcher Situationen in weiteren Studien sowie die differenziertere Formulierung der Warntexte könnte von der Hinzunahme psychologischer und medizinischer Kompetenzen profitieren.

4 Evaluation des Demonstrators

4.1 Evaluationsgegenstand

In dieser Evaluation soll der entwickelte Demonstrator im Hinblick auf seine Konformität mit den im Ergebnisbericht D3.1 aufgestellten Technisch-Organisatorischen Maßnahmen (TOMs) evaluiert werden. Die App und der Analysedienst werden in diesem Kapitel separat evaluiert, da beide auch eigene TOMs definiert haben.

4.2 Notwendigkeit der Evaluation

Die vorhergehenden Evaluationen haben den konzeptionellen und theoretischen Unterbau des Demonstrators untersucht. Die auf Grundlage von Art. 32 Abs. 1 DSGVO aufgestellten TOMs sollen sicherstellen, dass ein angemessenes Schutzniveau für Nutzer des Systems gewährleistet wird. Hieraus ergibt sich die Notwendigkeit, die tatsächliche Implementierung auf Konformität mit den zuvor genannten Maßnahmen zu überprüfen.

4.3 Evaluationskriterien

Für die Evaluation des Demonstrators konzentrieren wir uns auf die Vollständigkeit der Implementierungen mit den zuvor aufgestellten Technisch-Organisatorischen Maßnahmen. Bereits die Vernachlässigung einer einzelnen Maßnahme kann zu einem Verlust der Anonymität von Nutzern führen. Daher soll jede im Anforderungsdokument D3.1 definierte Maßnahme zu dieser Evaluation herangezogen werden.

Für ein Produktivsystem wäre auch die Korrektheit der Umsetzung der Maßnahmen ein essenzielles Kriterium. So bestünde zum Beispiel aufgrund von Implementierungsfehlern oder Zugrundelegung eines veralteten Standes der Technik trotz der Umsetzung einer TOM die Gefahr eines Verlusts der

Anonymität von Nutzern. Im Rahmen dieses Projektes wollen wir jedoch nur die grundsätzliche Gangbarkeit unserer Ansätze und die Zufriedenheit der Anwender mit den vorgeschlagenen Lösungsbausteinen nachweisen. Selbst wenn also noch unentdeckte Implementierungsschwächen vorhanden sein sollten, beeinträchtigt dies nicht die Funktion des Demonstrators für das Projekt, da der Demonstrator nie im Produktivbetrieb eingesetzt werden soll. Darüber hinaus müsste für eine Überprüfung auf Korrektheit eine unabhängige Partei die entsprechenden Prüfungen vornehmen, was im Rahmen des Projektes nicht möglich war.

4.4 Analyse

4.4.1 Bewertung der App

Wir führen die Analyse der App entlang der im Ergebnisbericht D3.1 (Sicherheitsarchitektur und Datenschutzkonzept) in Abschnitt 2 (Technisch-Organisatorische Maßnahmen für die App) genannten Anforderungen durch.

4.4.1.1 Verschlüsselte Speicherung

In der finalen Version des Demonstrators werden nur wenige Daten tatsächlich in der App persistiert. Das Paar Benutzername/Passwort wird zur Erleichterung des Logins in der Keychain des Betriebssystems abgelegt. Das trifft auch auf die eventuell bezogenen Login- und Refresh-Token zu. Durch Sandboxing und Verschlüsselung[2] betrachten wir die dort hinterlegten Werte als hinreichend geschützt. Mit iOS 18 wurde zudem auch die Möglichkeit eingeführt, das Starten einer beliebigen App von einer erfolgreichen biometrischen Identifizierung am Gerät abhängig zu machen, was daher vom Nutzer auch für die WearPrivate-App festgelegt werden kann.

Weiterhin werden in der finalen Version die Beschleunigungsdaten des Wearables sowie die Einstellungen für den MyData-Dienst in einer CoreData-Datenbank mit dem Protection Level *completeUnlessOpen*[3] zwischengespeichert.

4.4.1.2 Auditing von Frameworks

Nicht zuletzt, um diesen Aufgabenteil klein zu halten, waren wir bei der Entwicklung bestrebt, die im Projekt verwendeten Abhängigkeiten zu minimieren. Das hat abschließend betrachtet nicht funktioniert: Allein das SDK für den Polar-Pulsgurt hat als eine Abhängigkeit das RxSwift-Framework, das in über 11 MB mehr als 1300 Dateien unterbringt. Da noch weitere Dependencies vorhanden sind, die ihrerseits größere Abhängigkeiten mitbringen, ließ sich das Auditing nicht im Rahmen des Forschungsvorhabens durchführen.

4.4.1.3 Prüfen auf Jailbreaks

Wir haben das beschriebene Package *IOSSecuritySuite* in der beim Verfassen dieses Textes aktuellen Version 1.9.10 vom 14. Oktober 2023 in die App integriert. Die Integration ist dergestalt, dass wir beim Start der App auf das Vorhandensein eines Jailbreaks prüfen und den Start der Anwendung verhindern, falls ein Jailbreak erkannt wird. In der aktuellen Version geschieht das ohne weitere Nutzer-

[2] https://support.apple.com/de-de/guide/security/secb0694df1a/web

[3] https://developer.apple.com/documentation/foundation/fileprotectiontype/1617188-completeunlessopen

benachrichtigung. Als Weiterentwicklung wäre denkbar, dem Nutzer eine genauere Rückmeldung darüber zu liefern, warum die App nicht nutzbar ist.

4.4.1.4 Button für Unterbrechung implementieren

Als eindeutiger Hinweis auf die Datenerfassung in der App wurde ein Banner auf oberster Ebene in die App eingebaut, das bei aktivierter Datenerfassung vom Wearable einen entsprechenden Hinweistext auf rotem Hintergrund anzeigt.

Tatsächlich geht die Implementierung der App inzwischen über das ursprüngliche Konzept dieser TOM hinaus: Durch die Integration von MYDATA kann eine Nutzerin in den Einstellungen Zeiträume definieren, innerhalb derer die Erfassung von Daten erlaubt ist. Bevor die Daten tatsächlich an den Analysedienst gesendet werden, findet eine Überprüfung jedes Datums durch MYDATA statt, in der auch sichergestellt wird, dass dieses in einen erlaubten Zeitraum fällt. Ist dies nicht der Fall, wird der Datensatz verworfen.

4.4.1.5 Initial deaktivierte Push-Mitteilungen

Aus Zeitgründen bei der Konzeption und Entwicklung war es nicht mehr möglich, einen separaten Screen für die Kontextualisierung der Frage nach der Erlaubnis für Push-Mitteilungen im Onboarding des Demonstrators hinzuzufügen. Daher erscheint diese aktuell direkt beim ersten App-Start.

4.4.1.6 Keine Gesundheitsdaten in Push-Mitteilungen

Der Demonstrator zeigt keine sensitiven Daten in seinen Push-Mitteilungen an.

4.4.1.7 Prüfung des Vertriebswegs

Der Demonstrator ist gegenwärtig nicht für Endnutzer verfügbar. Die TestFlight-Infrastruktur, die zur projektinternen Anwendung zum Verteilen von Entwicklungsständen und für die Durchführung von Tests genutzt wird, ist Teil der AppStore-Infrastruktur. Auch diese kann genutzt werden, ohne das Gerät einem Jailbreak auszusetzen. Das wesentliche Kriterium der geforderten TOM ist damit erfüllt.

Die prognostizierte Öffnung des App-Stores in der EU wurde tatsächlich vollzogen. In der Praxis ist die Veröffentlichung einer App über diesen Weg aktuell noch technisch aufwendig, so dass versucht werden sollte, den AppStore weiterhin als Vertriebsweg zu wählen. Erst nach einer eventuellen Ablehnung scheint sich die Beschäftigung mit anderen Stores zu lohnen.

4.4.1.8 Datensicherung

Das Dokumentenverzeichnis der WearPrivate-App ist mit dem für die Steuerung von Backup-Eigenschaften vorgesehenen Dateiattribut *NSURLIsExcludedFromBackupKey* versehen. Hierdurch werden sämtliche Dateien der App von lokalen und Cloudbackups ausgenommen.

4.4.1.9 Periodisches Löschen personenbezogener Daten

Diese Anforderung konnte im Rahmen des Demonstrators umgesetzt werden. Die vom Wearable erhobenen Gesundheitsdaten werden nur bis zu einer erfolgreichen Übermittlung an den Analysedienst in der App vorgehalten und danach verworfen.

4.4.1.10 State-of-the-Art Privacy-UI

Die Gestaltung des UI und des Interaktionskonzepts legte einen starken Fokus darauf, klar dazustellen, wann und welche Daten geteilt werden. Kapitel 3 beschäftigt sich ausführlich mit der Analyse unseres Ansatzes.

4.4.1.11 Pushnachricht an Nutzer (als Umgang mit Ausnahmesituationen)

Ein Benachrichtigungsmechanismus für Ausnahmesituationen konnte innerhalb der Projektlaufzeit nicht mehr im Demonstrator implementiert werden. Die Realisierung stellt jedoch kein grundsätzliches Problem dar und wäre grundsätzlich leicht im Einklang mit unseren Anforderungen gemäß Ergebnisbericht D1.1 [11] realisierbar.

4.4.2 Bewertung des Analysedienstes

Der Analyseservice wurde vollständig gemäß den Anforderungen des Ergebnisberichts D3.1 (Sicherheitsarchitektur und Datenschutzkonzept), Abschnitt 3 (Technisch-Organisatorische Maßnahmen für den Analyseservice) implementiert, um die Einhaltung der Datenschutz-Grundverordnung (DSGVO) sicherzustellen. Durch die Nutzung von AWS als Cloud-Service-Anbieter, der mit umfangreichen Zertifizierungen und Sicherheitsstandards (z. B. ISO 27001 für Informationssicherheitsmanagement, ISO 27017 für Cloud-Sicherheit, ISO 27018 für Datenschutz in der Cloud und SOC 2/3 für Service-Organisationskontrollen) eine DSGVO-konforme Infrastruktur bietet, wurden alle erforderlichen Maßnahmen umgesetzt. Diese Maßnahmen gewährleisten die Integrität, Vertraulichkeit und Verfügbarkeit der verarbeiteten Daten und minimieren Datenschutzrisiken durch klare Prozesse und sichere Systemintegration.

4.4.2.1 Integrität und Vertraulichkeit

Folgende Maßnahmen wurden zur Gewährleistung der Integrität und Vertraulichkeit ergriffen:

- Zugangskontrolle: Der Zugriff auf personenbezogene Daten ist durch ein rollenbasiertes Berechtigungssystem, Passwortschutz und Zwei-Faktor-Authentifizierung ausschließlich autorisierten Personen möglich. Zugangsdaten werden ausschließlich in einem SOC2-zertifizierten Passwort-Manager sicher gespeichert.
- Datenverschlüsselung: Alle Datenübertragungen erfolgen verschlüsselt mittels TLS 1.2/1.3, und gespeicherte Daten werden durch AES-256-Verschlüsselung geschützt. AWS als Datenhost erfüllt sämtliche DSGVO-Standards und Sicherheitsanforderungen.
- Pseudonymisierung und Datenminimierung: Die Verarbeitung erfolgt in pseudonymisierter Form, wodurch Rückschlüsse auf Einzelpersonen ausgeschlossen werden. Nur die für den Dienst notwendige minimale Datenmenge wird erfasst und verarbeitet.
- Trennungskontrolle: Kundendaten sind mandantenfähig organisiert und werden durch eindeutige Mandanten-IDs getrennt. Dies stellt sicher, dass Kunden ausschließlich Zugriff auf ihre eigenen Daten haben.

4.4.2.2 Interaktion und Integration

Folgende Maßnahmen gewährleisten die sichere Interaktion und Integration der Systemkomponenten:

- Sichere Integration: Die Integration des Analyseservices mit externen Systemen erfolgt über dokumentierte und verschlüsselte Anwendungsprogrammierschnittstellen. Alle Datenübertragungen werden protokolliert, um eine vollständige Nachvollziehbarkeit zu gewährleisten.
- Hosting und Infrastruktur: Die gesamte Infrastruktur wird über AWS bereitgestellt und auf Servern innerhalb der EU (z. B. Frankfurt am Main) gehostet. AWS garantiert durch das Data Processing Addendum, dass Daten nur in der festgelegten Serverregion verarbeitet werden. Die AWS-Rechenzentren verfügen über umfassende Sicherheitsmaßnahmen wie Zutrittskontrollen, redundante Stromversorgung und Notfallsysteme.
- Systemübergreifende Sicherheit: Alle integrierten Systeme wurden vor ihrer Einbindung auf Einhaltung strenger Sicherheits- und Datenschutzanforderungen geprüft.

4.4.2.3 Betrieb und Nutzung

Folgende Maßnahmen dienen dem nutzerfreundlichen Datenschutz beim Betrieb des Systems:

- Benutzerfreundliche Anwendungen: Die Mobile App und das Dashboard sind intuitiv gestaltet und standardmäßig datenschutzfreundlich konfiguriert. Nutzer und Administratoren wurden in den sicheren Umgang mit den Anwendungen geschult.
- Protokollierung und Monitoring: Systemaktivitäten – einschließlich Zugriffe, Änderungen und Datenübertragungen – werden automatisch protokolliert. Protokolldaten werden regelmäßig überprüft, um mögliche Sicherheitsvorfälle frühzeitig zu erkennen.
- Backup- und Wiederherstellungsverfahren: Ein automatisiertes Backup-System erstellt regelmäßige Snapshots der Datenbanken. Backups werden in redundanten AWS-Zonen gespeichert und gewährleisten die Wiederherstellung im Falle eines Ausfalls.

4.4.2.4 Privacy by Default

Das System realisierte folgende Prinzipien:

- Datensparsame Voreinstellungen: Der Analyseservice ist so konfiguriert, dass standardmäßig nur minimal erforderliche Daten erfasst und verarbeitet werden.
- Anonymisierung von Daten: Vor der Analyse werden personenbezogene Daten anonymisiert, um Datenschutzrisiken zu minimieren.
- Einhaltung von Datenschutzprinzipien: Alle Prozesse und Systeme wurden so gestaltet, dass sie den Prinzipien der Datenminimierung und Zweckbindung entsprechen. Sie werden regelmäßig auf Aktualität und Wirksamkeit überprüft.

4.4.2.5 Umgang mit Ausnahmesituationen

Für den Umgang mit Datenschutzzwischenfällen sind folgende Vorkehrungen getroffen:

- Vorfallmanagement: Ein Incident-Response-Plan wurde implementiert, der Sicherheitsvorfälle effektiv bearbeitet. Betroffene Personen und zuständige Behörden werden im Falle einer Datenschutzverletzung unverzüglich informiert.

- Schulungen zu Notfallszenarien: Mitarbeiter werden regelmäßig geschult, um sicherzustellen, dass sie in Ausnahmefällen angemessen reagieren können.
- Zentrale Ansprechpartner: Ein Datenschutzbeauftragter oder zentraler Ansprechpartner ist jederzeit verfügbar, um bei Vorfällen oder Fragen zum Datenschutz Unterstützung zu leisten.

4.4.3 Bewertung des Demonstrators

Abschließend wird im Folgenden wird betrachtet, wie der Demonstrator als Gesamtsystem die im Ergebnisbericht D3.1 erhobenen Anforderungen an Schutzmaßnahmen, die über TOMs hinausgehen, erfüllt.

4.4.3.1 Anforderungen 12 – 14: Informationspflicht

- *Vor der Einführung einer Wearable-basierten Vitaldatenerfassung zum Zwecke des Arbeits- und Gesundheitsschutzes muss der Arbeitgeber die Belegschaft umfassend über die geplante Maßnahme und deren Zwecke, die eingesetzten Mittel und ihre Wirkungsweise, die dabei erhobenen Daten und deren Verwendung sowie die Empfänger der aus dem Messprogramm abgeleiteten Informationen informieren.*
- *Basierend auf den vermittelten Informationen gemäß der voranstehenden Anforderung ist von den Mitarbeitenden eine informierte, freiwillige, ausdrückliche und ggf. schriftliche oder elektronische Einwilligung zur Teilnahme an der Vitaldatenerfassung mittels Wearables und deren Auswertung zum Zwecke des Gesundheits- und Arbeitsschutzes einzuholen. Diese Einwilligung ist zu dokumentieren.*
- *Alle Informationen, die der Arbeitgeber gemäß der voranstehenden Anforderung den Betroffenen vermittelt hat, müssen von der mobilen WearPrivate-App vom Nutzer bei Bedarf jederzeit abrufbar sein.*

Im Demonstrator wurden keine Methoden implementiert, um eine Informationskampagne eines Arbeitsgebers zu unterstützen. Es handelt sich um eine organisatorische Maßnahme, die besser in Form einer Mitarbeiterversammlung durchgeführt wird. Eine Bereitstellung der Informationsmaterialien im Demonstrator wurde aus Aufwandsgründen nicht realisiert.

4.4.3.2 Anforderung 15: Widerrufbarkeit der Einwilligung

- *Die Mobil-App für die Teilnahme an einem Vitaldaten-Messprogramm zum Zwecke des Arbeits- und Gesundheitsschutzes muss eine Funktion bereitstellen, mit der ein Nutzer seine Einwilligung zur Teilnahme an dem Messprogramm jederzeit widerrufen kann.*

Die App enthält keine Funktion, um die Einwilligung zur Teilnahme am Messprogramm zu widerrufen. Das Merkmal wurde im Demonstrator aus Aufwandsgründen nicht umgesetzt.

4.4.3.3 Anforderung 16: Anonyme Teilnahme am Messprogramm

- *Arbeitnehmer benötigen für die Teilnahme am Messprogramm nur einen anonymen Freischaltcode ohne personenidentifizierende Merkmale. Eine Erfassung von Namen, Adressen, E-Mail-Adressen, Telefonnummern oder anderen personenbezogenen Angaben über die Teilnehmer beim Dienstleister erfolgt nicht. Die teilnehmenden Individuen am Messprogramm*

werden vom Dienstleister nur unter einer zufällig vergebenen ID-Nummer geführt; ihre wahre Identität bleibt dem Dienstleister verborgen.

Der Mechanismus für eine anonyme Teilnahme am Messprogramm wurde im Demonstrator umgesetzt.

4.4.3.4 Anforderung 17: Rollentrennung zwischen Analysedienst und Vertrieb

- *Die Gesamtdienstleistung »Belastungsmessprogramm für Arbeitnehmergruppen« soll in zwei unabhängige Teildienste unterteilt werden: (1) einen Analysedienst, der die Rohdaten anonymer Teilnehmer zu Gruppenberichten für anonyme Gruppen verdichtet; (2) einem Vertrieb, der die Geschäftsabwicklung mit dem Arbeitgeber übernimmt (z. B. Vertragskonditionen, Bereitstellung anonymer Teilnahmetickets, Ticketverifikation und Entwertung, Lieferung der Ergebnisse, Inkasso, Kundenbetreuung).*

Die organisatorische Anforderung, die Dienstleistung der Belastungsmessung in zwei unabhängigen Betreibern für jeweils Analyse und Vertrieb zu unterteilen, ließ sich im Projektkontext für den Demonstrator nicht umsetzen.

4.4.3.5 Anforderung 18: Getrennte Rolle für Gruppenaggregation

- *[optional] Nutzer sollen gegenüber dem Analysedienst nicht ihre Gruppen-ID offenlegen. Stattdessen aggregiert der Analysedienst die Daten eines Individuums zu Individualreports. Diese gibt er an eine gesonderte Instanz weiter, den Gruppenintegrator. Der Gruppenintegrator aggregiert die erhaltenen Individualreports zu Gruppenreports, die er dann an den Vertrieb weiterleitet. Zu diesem Zweck teilt jeder individuelle Nutzer dem Gruppenaggregator (unter Umgehung des Analysedienstes) anonym seine ID und seine GID mit.*

Eine optionale, noch weitergehende Entkopplung zwischen Individualdaten und Gruppenreport durch einen gesonderten Gruppenaggregator (neben Analysedienst und Vertrieb) wurde nicht implementiert.

4.4.3.6 Anforderung 19 – 20: Abwägung zwischen Privatsphäreschutz und Analysegenauigkeit

- *Den Nutzern sollen drei verschiedene Varianten für die Übertragung ihrer Daten zu Verfügung stehen:*
  - *Das **Versenden der unveränderten Rohdaten**. Dies bietet die höchste Datenqualität und damit zusammenhängend auch die höchste Genauigkeit des Analyse-Ergebnisses. In diesem Fall werden jedoch keine Schutzmaßnahmen für die Daten getroffen, um die identifizierten Bedrohungen zu mildern.*
  - *Das **Versenden von leicht geschützten Daten**. Durch Anonymisierungs- und Aggregationsmethoden sollen die versendeten Daten leicht verändert werden, um die Bedrohungen der Nutzer zu verringern. Dadurch wird die Datenqualität ein wenig verschlechtert, was zu Beeinträchtigungen der Genauigkeit der Analysen führen kann.*
  - *Das **Versenden von stark geschützten Daten**. Durch Anonymisierungs- und Aggregationsmethoden sollen die versendeten Daten derart verändert werden, dass sie die identifizierten Bedrohungen der Nutzer stark verringern. Dadurch wird die*

Datenqualität stärker als in Variante 2 beeinträchtigt, was zu einer weiteren Verschlechterung der Analysegenauigkeit führen kann.

- *Der Arbeitnehmer soll eine geschützte Verarbeitung seiner Daten wählen können, auch wenn diese einen Verlust der Qualität des Dienstes mit sich bringt.*

Die beschriebenen Varianten zur Auswahl eines Schutzlevels werden dem Arbeitnehmer vom System bereitgestellt.

4.4.3.7 Anforderung 21 – 22: Anonymisierung der Profildaten

- *Charakteristische Profildaten des Nutzers sollen vor einer Übermittlung an den Analysedienst verfremdet werden, um eine Identifizierung des Nutzers anhand der Profildaten zu erschweren.*
- *Die Daten der Nutzer auf dem Smartphone sollen so verfremdet werden, dass mindestens drei Nutzer jeweils dieselbe Kombination aus Geburtsjahr, Geschlecht, Größe und Gewicht haben.*

Im Demonstrator findet keine Verfremdung der Profildaten vor der Übermittlung an den Analysedienst statt.

4.4.3.8 Anforderungen 23 – 24: Verfremdung der Messdaten

- *Die gemessenen Rohdaten (Vital- und Kontextdaten) des Nutzers sollen vor einer Übermittlung an den Analysedienst verfremdet werden, um eine Identifizierung des Nutzers anhand der Rohdaten zu erschweren.*
- *Bewegungsdaten sollen mit Differential Privacy geschützt werden, wobei für den Parameter ε ein Wert unter 4 gewählt werden sollte.*

Je nach gewähltem Schutzlevel (s.o.) wird eine Verfremdung der Daten vor der Übermittlung an den Analysedienst durchgeführt. Der Mechanismus zur Verfremdung der Daten wurde im Demonstrator beispielhaft für Herzratendaten implementiert. Zudem haben wir haben den Einfluss verschiedener ε-Werte auf die Analysequalität untersucht. Unsere Untersuchungen haben gezeigt, dass eine Verfremdung der Herzratendaten mit $\varepsilon \approx 1$ mit einem noch vertretbaren Qualitätsverlust einhergeht und daher im Anwendungsfall möglich ist. Nähere Details dazu beschreibt Abschnitt 2.2.

4.5 Werturteil

In der App-Komponente des Demonstrators konnten die meisten der gesetzten technisch-organisatorischen Maßnahmen erfüllt werden. Bei den Maßnahmen, die nicht erfüllt werden konnten, war dies ausschließlich durch zeitliche Einschränkungen begründet. Bei keinem dieser Punkte zeichnet sich die technische Nichtmachbarkeit ab. Die für die App getroffenen technischen Maßgaben erscheinen uns daher als valide.

Der Analyseservice des Demonstrators erfüllt alle relevanten DSGVO-Anforderungen durch die erfolgreiche Implementierung der beschriebenen Maßnahmen. Durch die Nutzung der AWS-Infrastruktur und die Umsetzung von Datenschutzprinzipien wird ein Höchstmaß an Sicherheit, Vertraulichkeit und Integrität für die verarbeiteten Daten gewährleistet. Durch die Integration der Datennutzungskontrolle durch MYDATA bietet das System den Nutzern sowohl Selbstbestimmung (welche Daten dürfen für welche Zwecke verarbeitet werden) als auch Transparenz (wann wurden welche Daten zu welchen Zwecken verarbeitet) hinsichtlich der Verwendung ihrer Daten.

Quellenverzeichnis

[1] Clifford Colin W. G., Watson Tamara L. and White David (2018): Two sources of bias explain errors in facial age estimation. *R. Soc. open sci.* 5180841180841. http://doi.org/10.1098/rsos.180841

[2] Pau Climent-Pérez, Ángela M. Muñoz-Antón, Angelica Poli, Susanna Spinsante, Francisco Florez-Revuelta (2022): Dataset of acceleration signals recorded while performing activities of daily living. *Data in Brief, Volume 41, 2022,* ISSN 2352-3409. https://doi.org/10.1016/j.dib.2022.107896

[3] Fischer, B., Sedlmeier, A.M., Hartwig, S. *et al.* (2020): Anthropometrische Messungen in der NAKO Gesundheitsstudie – mehr als nur Größe und Gewicht. *Bundesgesundheitsblatt 63,* pp. 290–300. https://doi.org/10.1007/s00103-020-03096-w

[4] Martynov, K., Garimella, K. & West (2020): R. Human biases in body measurement estimation. *EPJ Data Sci. 9, 31.* https://doi.org/10.1140/epjds/s13688-020-00250-x

[5] Y. -H. Nho, J. G. Lim and D. -S. Kwon (2020): Cluster-Analysis-Based User-Adaptive Fall Detection Using Fusion of Heart Rate Sensor and Accelerometer in a Wearable Device. *IEEE Access*, vol. 8, pp. 40389-40401. https://doi.org/10.1109/ACCESS.2020.2969453

[6] Holohan, Naoise; Braghin, Stefano; Mac Aonghusa, Pol and Levacher, Killian (2019): Diffprivlib: the IBM differential privacy library. https://arxiv.org/abs/1907.02444

[7] European Society of Cardiology et al. (1996): Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal (1996) 17,* pp. 354–381. https://www.escardio.org/static-file/Escardio/Guidelines/Scientific-Statements/guidelines-Heart-Rate-Variability-FT-1996.pdf

[8] P. Melillo, R. Castaldo, G. Sannino, A. Orrico, G. de Pietro and L. Pecchia (2015): Wearable technology and ECG processing for fall risk assessment, prevention and detection. In: Proceedings of the *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 2015, pp. 7740–7743. https://doi.org/10.1109/EMBC.2015.7320186

[9] Irurzun, I. M., Garavaglia, L., Defeo, M. M., & Thomas Mailland, J. (2021): RR interval time series from healthy subjects (version 1.0.0). *PhysioNet.* https://doi.org/10.13026/51yd-d219.

[10] Saskia Koldijk, Maya Sappelli, Suzan Verberne, Mark A. Neerincx, and Wessel Kraaij (2014): The SWELL Knowledge Work Dataset for Stress and User Modeling Research. In: Proceedings of the *16th International Conference on Multimodal Interaction (ICMI '14).* Association for Computing Machinery, New York, NY, USA, 291–298. https://doi.org/10.1145/2663204.2663257.

[11] Svenja Polst, Philipp Neuschwander, Reinhard Schwarz, Bianca Steffes, Simone Salemi (2024): Anforderungsdokument. Ergebnisbericht D1.1, Projekt WearPrivate, Kaiserslautern – Saarbrücken