

Abschlussbericht

Teil I: Kurzbericht

DrAlve: KI-basierte Fahrerzustandserkennung im Kontext von Übergabesituationen des
(teil-) autonomen Fahrens

Zuwendungsempfänger: Rheinische Hochschule Köln gGmbH
(ehemals Rheinische Fachhochschule Köln gGmbH)
Fachbereich Wirtschaft und Recht
Schaevenstr. 1 a-b
50676 Köln

FKZ: 16SV8694

Berichtszeitraum 01.06.2021 – 31.05.2024

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Ursprüngliche Aufgabenstellungen sowie der wissenschaftliche und technische Stand, an den angeknüpft wurde

Zukunftstechnologien wie beispielsweise das (teil)autonome Fahren haben zum Ziel, unsere Sicherheit auf der Straße zu gewährleisten und zusätzlich den Komfort beim Fahren zu erhöhen. Im Rahmen von teilautonomen Szenarien (Stufen 2-4) muss eine sichere Übergabe der Fahrkontrolle vom Fahrzeug zum Fahrer und umgekehrt sichergestellt werden. So beinhaltet Stufe 2, die Ausführung von Lenk-, Beschleunigungs- und Bremsvorgängen durch Fahrerassistenzsysteme, unter Verwendung von Informationen über die Fahrumgebung und mit der Erwartung, dass der menschliche Fahrer alle verbleibenden Aspekte der dynamischen Fahraufgabe ausführt. Auf der Stufe 3 „Bedingte Automatisierung“ des Autonomielevels hingegen muss der Fahrer auf Anforderungen zum Eingreifen reagieren. Im Konsortialprojekt DrAIve wird ein KI-basiertes Fahrerzustandserkennungs-System entwickelt, um „Driver Readiness“ in solchen Übergabesituationen des autonomen Fahrens, z. B. bei einer Übergabe von automatisiertem zu manuellem Fahren, zu gewährleisten. Zu diesem Zweck wird der Fahrerzustand (z. B. Ablenkung, Wachheit, Schmerz) sowie die Fahrzeuginnenraum-Aktivität über ein audio-visuell gestütztes Fahrerassistenzsystem ermittelt. Detektiert wird der Fahrerzustand u.a. über Mimik- und Lidbewegungserkennung, Video-basierte-Herzratenenerkennung, Neurophonetische Stimmanalyse und Postur sowie Sitzposition. Zum Zweck der Modellentwicklung werden kognitive, emotionale und physische Parameter sowie Ablenkungsreize im Fahrzeuginnenraum in einer Labor- sowie in einer Testtrackstudie (im Realfahrzeug) erhoben und analysiert.

Die RH Köln ist im Rahmen des Projektes an der Entwicklung des Audio-Moduls (Stimmsignalanalyse) beteiligt. Hier ist das Ziel, den emotionalen Zustand des Fahrers anhand von akustischen Merkmalen wie z.B. Intonation und Frequenz zu bestimmen (Schuller et al., 2011; Batliner et al., 2011) und eine mögliche Ablenkung durch übermäßige Emotionalität zu detektieren. Das Fahrerzustandserkennungs-System soll nach Entwicklung hinsichtlich der Funktionalität, Nutzerakzeptanz und Nutzererfahrung evaluiert werden. An dem Projektvorhaben sind die Konsortialpartner CanControls GmbH und das Institut für experimentelle Psychophysiologie GmbH (IXP) beteiligt.

Ablauf des Vorhabens

Das Verbundprojekt DrAIve bestand aus acht Arbeitspaketen: AP1 Anforderungsmanagement; AP2 Corpus Engineering; AP3 Audio-Modul Fahrerzustandserkennung; AP4 Video-Modul Fahrerzustandserkennung; AP5 Kabinenaktivitäts-Erkennung; AP6 Prädiktion Fahrerzustand und Kabinenaktivität über Deep-Learning Modellierung; AP7 Anwendungs- und Gesamtsystemfunktionstest; AP 8 Dissemination. Die Rheinische Hochschule Köln (RH, ehemals Rheinische Fachhochschule Köln) war an AP1 Anforderungsmanagement, AP 2 Corpus Engineering, AP3 Audio-Modul Fahrerzustandserkennung, AP6 Prädiktion Fahrerzustand und Kabinenaktivität über Deep-Learning Modellierung, AP7 Anwendungs- und Gesamtsystemfunktionstest und AP 8 Dissemination beteiligt. Das Teilvorhaben der RH Köln befasste sich hauptsächlich mit der Entwicklung des Audio-Prototyps. Die Kernkompetenzen umfassten dabei alle Schritte von der Anforderungsanalyse der technischen und ELSI-Komponente, der Datenerhebung bis zur Datenvorbereitung und Datenbereinigung. Darüber hinaus war die RH an der Evaluation des Audio-Teilprototypen beteiligt.

Zwecks Modellentwicklung der Video- und Audio-Module, mussten Lerndaten im Rahmen des Projektes erhoben werden. Dies erfolgte in zwei Erhebungen: im stehenden

Fahrzeug sowie auf einer Teststrecke unter Realbedingungen. Aufgrund von aufwändigen Erhebungsprozessen und großer Teilnehmerzahl bei den beiden Datenerhebungen gab es vorläufig eine 4-monatige Verzögerung, die jedoch im Laufe der Datenvorbereitung sowie Modellierungsarbeiten kompensiert werden konnte. Nach Fertigstellung der Prototypen war die RH an einem Gesamtfunktionstest beteiligt, um Optimierungsbedarfe zu identifizieren. Zudem hat die RH Köln eine Evaluation des Audio-Teilprototyps mithilfe einer Online-Umfrage durchgeführt, um Bedenken, Vorlieben und Vorschläge der Nutzer:innen zu erfassen sowie prototypische Gestaltungsempfehlungen abzuleiten.

Die wesentlichen Ergebnisse sowie ggf. die Zusammenarbeit mit anderen Forschungseinrichtungen

Im Rahmen des Verbundprojektes wurde ein System zur Fahrerzustandserkennung entwickelt, welches in Echtzeit Informationen über emotionale, kognitive und gesundheitliche Faktoren liefert. Beispielsweise können dadurch Müdigkeit, Ablenkung, Schmerz oder Wut beim Fahrer detektiert werden. Das passiert durch Analyse von Audio- und Videodaten. Das Video-Modul sowie die Konzeption des finalen Demonstrators wurde von CanControls GmbH (Verbundkoordinator) durchgeführt. An das Audio-Modul, welches sich ausschließlich mit der Erfassung des emotionalen Fahrerzustandes beschäftigte, waren das Institut für experimentelle Psychophysiologie GmbH (IXP) und die RH Köln beteiligt. Die RH Köln war dabei hauptsächlich für die Datenverarbeitung und Datenbereinigung zuständig, während das IXP die Modellierung übernommen hat.

Der entwickelte Demonstrator zur Fahrerzustandserkennung verfügte über keine Benutzeroberfläche, über die Nutzer:innen mit dem System interagieren können. Aus diesem Grund erwies sich eine typische Evaluation der Nutzerakzeptanz durch Usability Testung als nicht möglich. Um die Usability und Nutzerakzeptanz zu testen, wurde mit den Projektpartnern entschieden, eine low-fidelity nicht funktionale Smartphone-Anwendung zu konzipieren, über die Nutzer:innen mit dem System interagieren können. Beispielsweise kann die Anwendung die Fahrerzustände der vergangenen Fahrten abbilden und Nutzer:innen können sich genauer durch verschiedenen Funktionen über eine bestimmte Strecke oder einen bestimmten Zustand informieren. Die RH Köln war dabei an einer ausführlichen Literaturrecherche von Heuristiken für die Datenvisualisierung beteiligt. Die prototypische Entwicklung der App wurde vom IXP übernommen. Schließlich wurde eine Usability-Evaluation von IXP durchgeführt, mit dem Ziel Nutzererfahrungen und Verbesserungsvorschläge abzuleiten.

Im Laufe des Forschungsprojektes wurde stets im engen Austausch mit den anderen Projektpartnern – CanControls GmbH und Institut für experimentelle Psychophysiologie GmbH gearbeitet. Der Informationsaustausch fand im Rahmen von zweiwöchentlichen Regelmeetings und jährlichen Statustreffen sowie zusätzlich bei Übergabesituationen und bei der gemeinsamen Erarbeitung von Inhalten statt.

Literaturverzeichnis

Ein gemeinsames Literaturverzeichnis für Teil I und Teil II ist auf Seite 17 von Teil II zu finden.

Abschlussbericht

Teil II - Eingehende Darstellung

DrAlve: KI-basierte Fahrerzustandserkennung im Kontext von Übergabesituationen des
(teil-) autonomen Fahrens

Zuwendungsempfänger: Rheinische Hochschule Köln gGmbH
(ehemals Rheinische Fachhochschule Köln gGmbH)
Fachbereich Wirtschaft und Recht
Schaevenstr. 1 a-b
50676 Köln

FKZ: 16SV8694

Berichtszeitraum 01.06.2021 – 31.05.2024

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Inhaltsverzeichnis

<i>Ausführlichere Darstellung der im Rahmen des Vorhabens durchgeführten Arbeiten, insbesondere im Vergleich zur ursprünglichen Vorhabensbeschreibung.....</i>	<i>3</i>
<i>Die wichtigsten Positionen des zahlenmäßigen Nachweises</i>	<i>15</i>
<i>Die Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten.....</i>	<i>15</i>
<i>Der voraussichtliche Nutzen, insbesondere die Verwertbarkeit der Ergebnisse – auch konkrete Planungen für die nähere Zukunft – im Sinne des fortgeschriebenen Verwertungsplans.....</i>	<i>16</i>
<i>Der während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen</i>	<i>16</i>
<i>Die erfolgten oder geplanten Veröffentlichungen des Ergebnisses nach Nr. 5 der NABF</i>	<i>16</i>
<i>Literaturverzeichnis</i>	<i>17</i>

Ausführlichere Darstellung der im Rahmen des Vorhabens durchgeführten Arbeiten, insbesondere im Vergleich zur ursprünglichen Vorhabensbeschreibung

API Anforderungsmanagement

AP 1.2. Anforderungsanalyse Daten u. Messsysteme

Die Rheinische Hochschule (RH) Köln hat sich im Rahmen dieses Arbeitspaketes schwerpunktmäßig um die Zusammenstellung der Anforderungen an die Aufnahme von Audiodaten gekümmert. Clip-On Mikrofone zeichnen sich durch einen gleichbleibenden Abstand zu Sprecher aus, jedoch haben jedoch eine schlechtere Sound-Qualität. Dadurch dass eine gute Audioqualität für das Forschungsvorhaben unerlässlich ist, wurde gemeinsam die Nutzung eines stationären Mikrofons (Rode VideoMic NTG) entschieden. Ein variierender Abstand zwischen Mikrophon und Sprecher sollte durch feste Installation während gleichbleibender Position der Probanden bei allen Erhebungen vermieden werden.

Weitere Anforderungen bezüglich des Datenformates wurden ebenso ausgearbeitet. Berücksichtigt wurden dabei Besonderheiten von Aufzeichnung bis hin zu der Speicherung der Daten. Die Aufzeichnungsqualität sollte bei mindestens 44.1 kHz und 24 Bit liegen und statt komprimierenden Dateiformaten (z.B. mp3, m4a) wurden .wav verwendet. Zum Vergleich der Daten wurden immer auch Hintergrundgeräusche sowie neutrale Sprachgeräusche aufgenommen. Weitere Anforderungen an das Audio-System sind auf Abbildung 1 zusammengefasst.




		
<ul style="list-style-type: none"> • Aufzeichnungsqualität sollte bei mindestens 44.1 kHz und 24 Bit liegen • Soundkarte muss für Aufzeichnungsfrequenz und -tiefe geeignet sein • Standardcomputer sind nicht in der Lage, mehrere Audiospuren gleichzeitig zu verarbeiten → ggf. wird ein Digitizer benötigt 	<ul style="list-style-type: none"> • Anwendung eines Pop-Filters von Vorteil • Die Verwendung von Bluetooth sollte vermieden werden, da bei der Transmission üblicherweise Daten verloren gehen • Statt komprimierenden Dateiformaten (z.B. mp3, m4a) sollte .wav verwendet werden 	<ul style="list-style-type: none"> • Wenn möglich, sollten der Sprecher mit einer anderen Person reden (Dialoge ohne Empfänger unterscheiden sich phonetisch von echten Unterhaltungen) • Zum Vergleich der Daten sollten immer auch Hintergrundgeräusche sowie neutrale Sprachgeräusche aufgenommen werden • Alles testen! Das Messsystem, die Anwendung und das finale Setup sollten vor der finalen Durchführung immer überprüft werden

Abb. 1. Anforderungen an das Audio-System

AP 1.4. Anforderungsanalyse ELSI-Aspekte

Es wurde eine ausgiebige Literaturrecherche bezüglich relevanter ELSI-Aspekte im Automotive Bereich sowie bezüglich AI-Systemen durchgeführt. Durch die Grundsätze des Konzeptes Privacy by Design, wird der Datenschutz bereits während der konzeptionellen und Entwicklungsphase berücksichtigt (Abb. 1). Dadurch sollen wichtige Datenschutzaspekte bereits bei der Technikgestaltung ins System mit einfließen. Der Begriff Privacy by Design beschreibt einen umfassenden Prozess, um Technologien zu entwerfen, die die Privatsphäre der Nutzer schützen. Durch den Grundsatz „Privacy by default“ müssten Autofahrer in der Lage sein, selbst entscheiden zu dürfen, welche Daten sie übermitteln wollen und welche nicht. Bei einem Umgang mit personenbezogenen Daten sind insbesondere die Komponenten der Authentisierung und Authentifizierung, Anonymisierung und Pseudonymisierung, Techniken der Datenvermeidung und Datensparsamkeit, Trennung von Daten und die Verschlüsselung

von Daten zu berücksichtigen (EU-Datenschutz-Grundverordnung). Zusätzlich wurden über den Anbieter adesso SE ausführliche Workshops zum Thema IT-Security und Datenschutz gebucht.

Zweckdefinition	<ul style="list-style-type: none"> • Zweck des Systems muss vor Beginn des Prozesses feststehen
Präventiver Datenschutz	<ul style="list-style-type: none"> • Alle Aktionen, die die Privatsphäre bedrohen könnten, werden durch präventive Maßnahmen verhindert • Aktionen werden verhindert, wenn sie nicht notwendig sind
Datenschutz als Standard	<ul style="list-style-type: none"> • Nur zwingend notwendige Daten werden erhoben • Verarbeitung nur für definierten Zweck und nach vorher klar definierter Vorgehensweise
Positives Gesamtergebnis	<ul style="list-style-type: none"> • Bei Datenschutzkonflikten: schrittweises Vorgehen, Kompromisse sind nur bei positivem Gesamtergebnis zulässig • Kriterien für positives Gesamtergebnis werden vorher definiert
Datenschutzziele	<ul style="list-style-type: none"> • Vertraulichkeit, Integrität, Verfügbarkeit, Transparenz, Nicht-Verknüpfbarkeit, Intervenierbarkeit • Einhaltung der Ziele ist zu beachten und dokumentierten
Schutz des vollen Lebenszyklus	<ul style="list-style-type: none"> • Alle Maßnahmen zur Einhaltung des Datenschutzes müssen von der Erhebung bis zur endgültigen Löschung der Daten angewendet werden
Individueller Datenschutz	<ul style="list-style-type: none"> • Es sollte möglich sein, das System zu konfigurieren und an individuelle Bedürfnisse anzupassen

Abb. 2. Die Kernkomponenten des Privacy by Design Prinzip

Mithilfe des ELSI-SAT Fragebogens (<https://www.elsi-sat.de/index.html>) wurden für das Vorhaben relevante ELSI-Aspekte aufgearbeitet und überprüft. Entlang von 7 Dimensionen (Allgemein, Datenschutz, Fürsorge, Schadensvermeidung, Autonomie, Gerechtigkeit, Transparenz) werden hier zentrale Fragestellungen adressiert. Im Rahmen der Ergebnisse (Abb. 2 bis 4) sind für das Projekt DrAlve in fast allen Dimensionen gute Werte zu beobachten. Bei der Dimension Fürsorge soll sichergestellt werden, dass unterschiedliche Subpopulationen bei der Modellbildung miteinbezogen werden, um möglicher Bias zu reduzieren.

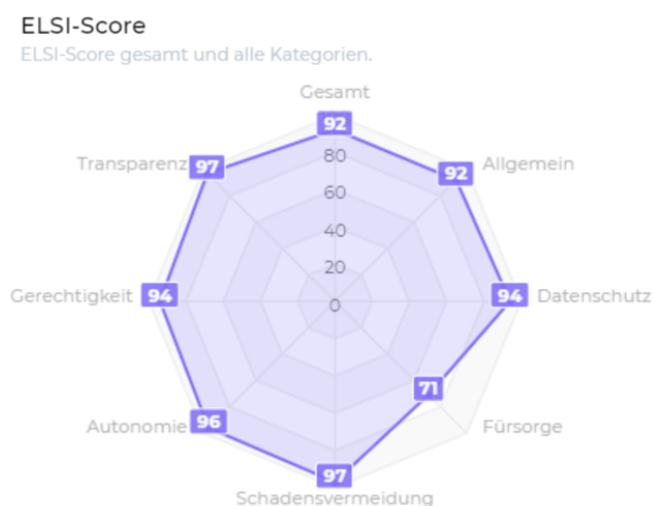


Abb. 3. Zusammenfassende ELSI-Bewertung des DrAlve-Projekts entlang der Hauptdimensionen (Allgemein, Datenschutz, Fürsorge, Schadensvermeidung, Autonomie, Gerechtigkeit, Transparenz) des ELSI-SAT Tools

Risiko und Adressierung

Die Risiko- und Adressierungs-Werte für jede Kategorie.



Heatmap

Risiko, Adressierung und ELSI-Score als Ampel für alle Kategorien.

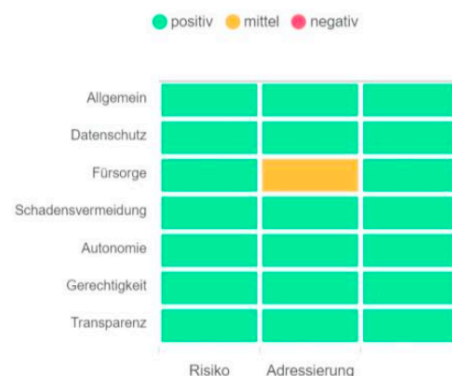


Abb. 4. Klassifikation der ELSI-SAT der Hauptdimensionen in Bezug auf das DRAIVE Projekt. Bestimmung von Risikohöhe und Risikoadressierung.

AP 1.5. Spezifikation von Test und Validierungsszenarien

Die RFH hat das IXP in der Erstellung von möglichen Anwendungs- und Testszenarien mit dem Fokus der Generierung von Audioevents unterstützt. Hauptziel des Audio-Modules ist die Erkennung verschiedener emotionaler Zustände im Straßenverkehr durch Stimmparameter. Gemeinsam wurden die wichtigsten Emotionen herausgearbeitet, die für den Kontext Autofahren relevant sind: Wut, Trauer und Überraschung. Ein weiteres systematisches Review konnte zeigen, dass die genannten Emotionen eine negative Wirkung auf das manuelle Fahrverhalten aufweisen (Braun et al., 2021). Beispielsweise könnten Du et al. (2019) zeigen, dass Wut zu einem aggressiven und riskanten Fahren und zudem zu einer geringeren Übernahmefähigkeit (Take-over readiness) geführt hat. Braun und Kollegen (2021) weisen zudem darauf hin, dass Trauer sich ebenso negativ auf das Fahrverhalten auswirkt sowie dass Überraschung zu einer höheren kognitiven Belastbarkeit führen kann. Emotionsmodellierung und -erkennung über Sprache, z.B. Tonhöhe, Lautstärke, ist besonders geeignet, um den emotionalen Zustand einer Person zu bestimmen und gewinnt in der Automobilindustrie zunehmend an Interesse (Zepf et al., 2020).

Zwecks Emotionsinduktion wurden entweder vorgegebene Szenarien oder die Erinnerung eines eigenen Szenarios mit einem selbstkonstruierten Satz benutzt. Für die jeweilige Emotion wurden Szenarien ausgearbeitet, die den Probanden vorgelesen werden sollten. Die Teilnehmer sollen sich dann in die Szenarien möglichst gut hineinversetzen und die Emotion spüren. Nach dieser Emotionsinduktion wurden die Teilnehmer gebeten, die jeweilige Emotion akustisch darzustellen. Dies erfolgte entweder durch einen Freitext (Probanden dürfen selber der Inhalt das Gesagte auswählen) oder mithilfe von bestimmten phonetischen Sätzen, basierend auf der deutschen akustischen Emo-DB Datenbank (Burkhardt et al., 2005). Diese ausgewählten phonetischen Sätze beinhalten möglichst viele Vokale der deutschen Sprache, was für die Modellbildung von wesentlichem Vorteil ist. Allerdings reduziert das die ökologische Validität der Szenarien, da die Inhalte der Sätze mit der zu darstellenden Emotionen nicht kompatibel sind (z.B. „Der Lappen liegt auf dem Eisschrank“). Probanden wurden aus dem Grund gebeten, der Inhalt der Sätze möglichst zu ignorieren. Im Anschluss sollten die Teilnehmer auf einer Skala von 1 bis 10 angeben, wie gut sie sich in die jeweilige Emotion hineinversetzen konnten. Das ermöglichte unter anderem einzelne Aufgaben und Emotionen, wo die Emotionsinduktion nicht gelungen ist, aus der Modellbildung zu entfernen und dadurch die Qualität des Modells zu steigern.

AP2 Corpus Engineering

AP 2.1. Datenrecording in Labor Studie 1

Im Arbeitspaket 2.1. Datenrecording in Labor (Studie 1), geleitet vom IXP, war die RFH unterstützend und beratend beteiligt. Dabei ist eine kontinuierliche Anpassung der Testszenarien und Bedingungen im Laufe der Datenerhebung zu nennen sowie die Qualitätssicherung der Audio-Dateien. Wissenschaft-fundierte Methoden zur Generierung von emotionalen Audioaufnahmen und deren akustischen Ausdrucksweise wurden auf Basis der deutschen akustischen Emo-DB Datenbank abgeleitet und im Versuchsprotokoll implementiert. Die Qualität der Audioaufnahmen wurde im Laufe der Datenerhebung kontinuierlich überwacht und Datensätze mit Störgeräuschen, schlechter Position des Mikrofons oder weiteren technischen Störungen entsprechend agiert.

AP 2.2. Datenrecording in Testtrack/Realfahrzeug Studie 2

Im Rahmen des Datenrecordings im Testtrack/Realfahrzeug (Studie 2), geleitet vom IXP, lag der Fokus vor allem auf der kontinuierlichen Qualitätssicherung der Audio-Dateien, auf der technischen Unterstützung bezüglich der Audio-Bestandteile sowie auf der Entwicklung der Test-Szenarien und -aufgaben zum Zweck der Erfassung des emotionalen Zustands des Fahrers. Wissenschaftlich fundierte Methoden zur Generierung von emotionalen Audioaufnahmen und ihrer akustischen Ausdrucksweise wurden, wie bei der Datenerhebung im stehenden Fahrzeug (Studie 1) auf Basis der deutschen akustischen Emo-DB Datenbank abgeleitet und im Versuchsprotokoll implementiert. Die Qualität der Audioaufnahmen wurde ebenso kontinuierlich überwacht.

AP 2.3. Datenkorpora Qualitätsprüfung, Bereinigung und Fusionierung

Nach dem abgeschlossenen Datenrecording in Labor sowie in Realfahrzeug legte die RFH den Schwerpunkt auf die Aufbereitung und Qualitätssicherung des Audio-Datenkorporpus. Dazu wurde das Open Source Tool Label Studio eingesetzt, um die Audiodateien systematisch nach Emotionen und möglichen Störgeräuschen zu labeln. Label Studio ermöglichte es uns, präzise Annotationen durchzuführen, indem es eine benutzerfreundliche Oberfläche bereitstellte, die die manuelle Kennzeichnung und Kategorisierung der auditiven Daten erleichterte.

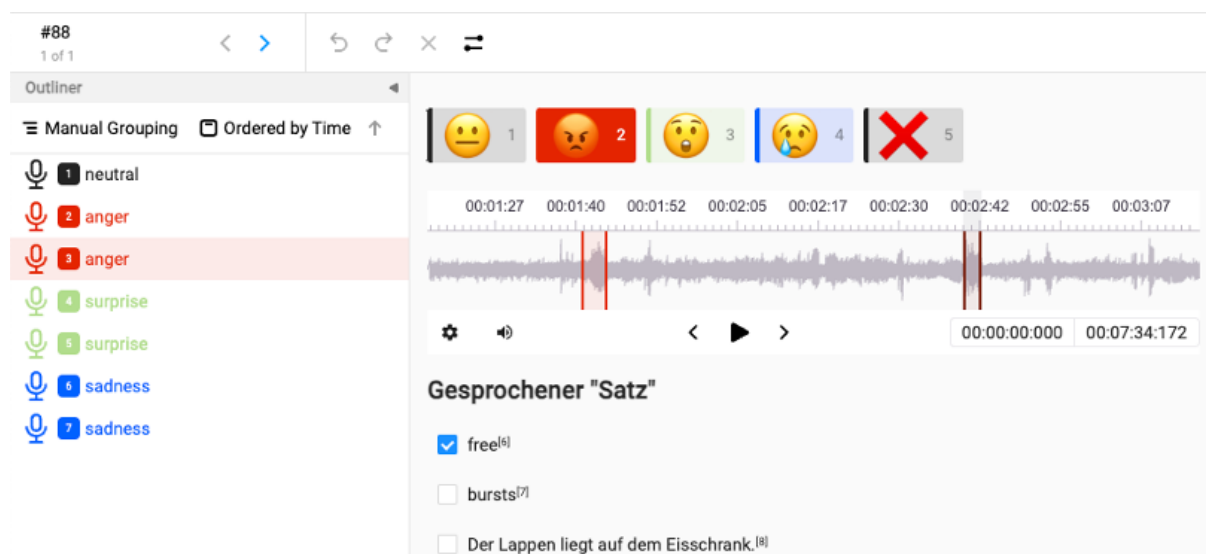


Abb. 5. Labeling Tool

AP 2.3.1 Audio-Datenkorpora Qualitätsprüfung, Bereinigung.

Das Labeln der Audiodaten war äußerst arbeitsintensiv und erforderte viel Zeit und Sorgfalt von unserem Team, da jede einzelne Datei detailliert analysiert und entsprechend markiert werden musste. Durch diese aufwendige händische Labelarbeit konnten wir die Emotionen in den Aufnahmen zuverlässig identifizieren und gleichzeitig unerwünschte Geräusche, wie z.B. das Geräusch des Blinkers, markieren, die die Analyse beeinflussen könnten.

Anschließend konnten Anhand der vergebenen Labels die Audiodateien in kleinere Segmente geschnitten werden, die nur noch den entsprechenden Satz enthielten. Dabei wurden Störgeräusche und das Hintergrundrauschen möglichst entfernt, sodass die Aufnahmen sich aus dem stehenden und aus dem fahrenden Fahrzeug sich nicht mehr systematisch unterscheiden. Ebenfalls wurden die Audioaufnahmen normalisiert, damit der Abstand zwischen Mikrofon und Mund nicht zu unterschiedliche Lautstärke führt, da dies einen Bias in die Daten bringen könnte. Die Sätze, bei denen während des Labelings unbehebbar Probleme aufgefallen sind, wie z.B. starkes Übersteuern oder aber auch nicht natürlich klingende Sätze wurden entfernt.

AP 2.3.2 Audio-Datenkorpora Fusionierung.

Für die Fusionierung der zwei Korpora wurden Metadaten in Form von CSV-Dateien erstellt, die für jede Audiodatei Informationen über Dateinamen, MD5_Hash, Korpusname, Emotion, ID der Sprechenden Person, Geschlecht, Alter, Art der Induktion, Art des Satzes (frei oder vorgegeben), ID des Satzes, Sprache und nicht herausgefilterte kleinere Probleme. Diese Metadaten erlauben es bei der Entwicklung des Machine Learning Algorithmus jeden Datensatz gleich zu behandeln und direkt jede Information aus einer Quelle zu haben und so schnell möglichen Bias finden zu können.

AP3 Audio-Modul Fahrerzustandserkennung

AP 3.1. Anpassung der Vorverarbeitung und Audio-Merkmalsextraktion für den Kontext Fahrerzustandserkennung

In Absprache mit IXP haben wir für alle erhobenen und externen Datensätze die OpenSmile (Eyben et al., 2010) Features extrahiert und ihnen diese zur Verfügung gestellt. Zusätzlich haben wir an der Entwicklung einer ressourcensparenden Echtzeit-Version gearbeitet. Diese Effizienz ist entscheidend, damit die Feature-Extraktion auch auf weniger leistungsstarken Systemen zuverlässig läuft und in Echtzeit erfolgen, sodass das Auto unmittelbar Feedback geben kann. Dieses Vorhaben haben wir jedoch eingestellt, nachdem das IXP auf ein Transformermodell mit integriertem Encoder umgestiegen ist.

AP 3.2. Aufbereitung Fahrerzustandserkennung relevanter Zusatz-Datencorpora

Die im Projekt erhobenen Datensätze wurden durch frei zugängliche Daten ergänzt, um die Analyse der Audiodateien zu erweitern und zu vertiefen. Hierbei kamen die Korpora Thorsten Emotional (Müller & Kreutz, 2021), Montreal Affective Voices (Belin et al., 2007), Emo-DB (Burkhardt et al., 2005) und das YouTube-8M Segments Dataset (Abu-El-Haija 2016) zum Einsatz, da diese die geforderten Emotionen bzw. Gesundheitsmerkmale enthalten. Die Integration dieser Korpora ist wichtig, um eine verbesserte Generalisierbarkeit der Ergebnisse und eine robustere Validierung der entwickelten Modelle zu erlangen, da sie eine Vielfalt an Stimmen, Sprachen und Emotionen beinhalten.

Für die Aufbereitung der jeweiligen Korpora wurden die gleichen Metadaten wie in AP 2.3.2 erstellt. Dies war hier besonders wichtig, da sich so das Entwicklungsteam nicht die Arbeit machen muss jedes Paper und vorhandene Dokumentation durchzugehen, um so an die relevanten Informationen zu kommen. Durch diese Metadateien ist es möglich die externen Datensätze wie die in diesem Projekt erhobenen Datensätze zu behandeln.

AP 3.4 Prüfung der Umsetzung der Nutzerakzeptanz u. ELSI-Aspekte des Teilprototypen

Dadurch, dass der Audio-Teilprototyp über keine Benutzeroberfläche verfügt und Nutzer:innen nicht direkt mit dem System interagieren können, erwies sich eine typische Evaluation vor Ort als nicht sinnvoll. Stattdessen wurde entschieden, die Forschungsfragen in APs 3.4. und 3.5. mittels einer Online-Umfrage zu untersuchen. In der Online-Umfrage werden realitätsnahe Anwendungsfälle im Fahrzeug beschrieben und Nutzer:innen werden gebeten, den Teil-Prototyp anhand dieser Anwendungsfälle zu evaluieren. Zudem können hiermit gleichzeitig verschiedene akustischen Feedback-Optionen des Systems evaluiert werden, um Gestaltungsempfehlungen zur prototypischen Realisierung auszuarbeiten.

Es erfolgte auch eine Prüfung der zuvor ausgearbeiteten ELSI-Aspekte, genannt in AP 1.4. Der Audio-Prototyp war von dem Projektpartner IXP zu entwickeln und stellt eine Teilkomponente des Gesamtsystems dar. Hierbei wurden ausschließlich Audio-Daten zur Analyse des emotionalen Zustandes verwendet. Geplant ist hierbei zukünftig die gesammelten Daten in Echtzeit zu analysieren und nicht zu speichern, sodass keine personenbezogenen Daten gesammelt werden. Dies kann allerdings in folgenden Entwicklungsschritten umgesetzt werden. Dadurch dass das Teilprototyp nur eine Art von Daten verwendet, entfällt hier die Komponente der Konfigurierbarkeit (Individueller Datenschutz). Nutzer:innen sind in der Lage das System an- und auszuschalten. Die Konfigurierbarkeit bei multiplen Datenflüssen (z.B. Audio und Video) soll bei dem Gesamt-Demonstrator geprüft werden. Die weiteren Datenschutz-Kriterien von AP1.4. sind durch den Audio-Teilprototyp erfüllt. Die Nutzerakzeptanz wurde im Rahmen der Online-Umfrage unter AP 3.5. evaluiert.

AP 3.5 Evaluation Teilprototyp u. Ausarbeitung von Gestaltungsempfehlungen zur prototypischen Realisierung

Der Audio-Teilprototyp verfügt über keine Benutzeroberfläche und Nutzer:innen können somit nicht direkt mit dem System interagieren. Dadurch erwies sich eine typische Evaluation vor Ort als nicht sinnvoll. Stattdessen wurde entschieden, die Forschungsfragen in APs 3.4. und 3.5. mittels einer Online-Umfrage zu untersuchen. In der Online-Umfrage wurden realitätsnahe Anwendungsfälle im Fahrzeug beschrieben und Nutzer:innen wurden gebeten, den Teil-Prototyp anhand dieser Anwendungsfälle zu evaluieren. Zudem wurden hiermit gleichzeitig verschiedene akustische Feedback-Optionen des Systems (Warnton, Melodie, menschliche Stimmen) evaluiert, um Gestaltungsempfehlungen zur prototypischen Realisierung auszuarbeiten.

Es wurden mehrere Nutzerszenarien ausgearbeitet. Bezüglich des Systems gab es zwei Varianten: a) eine Übernahme der Fahrkontrolle nach manueller Fahrt und b) ein Take-Over Request (Übernahmeanfrage) an den Fahrer während autonomer Fahrt aufgrund von einer systemkritischen Situation z.B. Baustelle mit Personen auf der Fahrbahn. Der Fahrer befindet sich bei den oberen Szenarien in einem unsicheren emotionalen Zustand, bedingt durch a) Wut oder b) Angst/Stress. Es wurden bei jedem Szenario mehrere Feedback-Optionen des Systems untersucht: Warnton, Melodie, weibliche und männliche Stimme. Darüber hinaus wurde bei einem Warnsignal durch Voice auch die Wirkung des Inhalts des Gesagten untersucht.

Beim Studiendesign handelte es sich um ein Between-Subjects-Design. Zwei verschiedene Umfrageteile mit verschiedenen Szenarien wurden erstellt, um mögliche Effekte der Emotionen auszuschließen. Jede Emotion wurde jeweils mit einem der Systemszenarien kombiniert. Ein Beispielszenario ist auf Abb.6 zu finden.

Um zu prüfen, ob die Teilnehmer sich in die Szenarien hineinversetzen konnten, wurde neben einer subjektiven Eigeneinschätzung auf einer Skala von 1 bis 5 auch die eigene Fähigkeit zur Perspektivenübernahme mittels dem Saarbrücken Persönlichkeitsfragebogen SPF(IRI) (Paulus, 2009; Skalen Perspektivenwechsel und Fantasie) erhoben. Personen, die unten den von den Autoren angegebenen Durchschnittswerte lagen, wurden aus der Analyse ausgeschlossen.

Szenario 1

Stellen Sie sich vor, Sie fahren auf der Autobahn. Sie sind gerade nach einem langen, harten Arbeitstag auf dem Heimweg und freuen sich auf den wohlverdienten Feierabend. Plötzlich klingelt Ihr Telefon. Sie nehmen den Anruf über die Freisprechanlage des Fahrzeugs entgegen. Es ist Ihr Chef. Er ist sehr ungehalten und verlangt, dass Sie sich sofort mit einem Ihrer Kunden in Verbindung setzen, um ein Verständnisproblem zu klären. Sie sind sehr wütend, weil Sie zum wiederholten Mal von Ihrem Chef nach der Arbeitszeit kontaktiert werden. Das Problem des Kunden erscheint Ihnen auch nicht als so dringend, dass es nicht bis morgen warten könnte. Sie regen sich sehr über die ganze Situation auf. Was für ein wirklich mieser Tag. Es ertönt ein Warnsignal, das Sie darauf hinweist, dass Ihre Verfassung eine sichere Führung des Fahrzeugs einschränkt. Der Sprachassistent bietet Ihnen an, das autonome Fahren zu aktivieren. Nachdem Sie zustimmen, übernimmt die autonome Fahrassistenz die Kontrolle über das Fahrzeug.

Abb. 6. Beispielhaftes Szenario zur Übernahme der Fahrkontrolle durch das System; Fahrerzustand: Wut.

Insgesamt 26 Personen nahmen an der Studie teil. Davon wurden 4 Personen aufgrund von unterdurchschnittliche Perspektivenübernahme-Werte ausgeschlossen. Die finale Stichprobe bestand daher aus 22 Personen (Durchschnittsalter 34,7 Jahre; 14 weiblich; 8 männlich). Die durchschnittliche Fahrerfahrung beträgt 17 Jahre und die Mehrheit der Teilnehmer benutzen keine Fahrerassistenzsysteme ($N=14$). Es wurden keine signifikanten Unterschiede zwischen den beiden Umfrageteilen gefunden.

Aufgrund der kleinen Stichprobe wurde nur eine explorative Analyse durchgeführt. Anhand dieser stellte sich ein möglicher Geschlechtereffekt heraus, wobei Frauen tendenziell die Sicherheit (Männer $M= 2.81$, Frauen $M=3.63$) und Unterstützung (Männer $M= 2.52$, Frauen $M=3.83$) des Systems sowie die Bereitschaft dieses im Alltag zu nutzen (Männer $M= 2.75$, Frauen $M=3.67$) höher bewertet haben als Männer. Komfort wurde lediglich nur im Szenario zur Übernahme der Fahrkontrolle nach manueller Fahrt besser bewertet (Männer $M= 3.29$, Frauen $M=4.5$). Signifikant ist davon jedoch kein Unterschied.

Bezüglich der Warnsignale wurde die Melodie über alle Szenarien hinweg am schlechtesten bewertet ($M = 1.59$). Ein Warnsignal via Voice war tendenziell beliebter als ein Warnton:

Warnton ($M = 3.13$), Stimme weiblich ($M = 3.62$), Stimme männlich ($M = 3.72$). Generell ist zwischen den Szenarien kein Unterschied zu sehen.

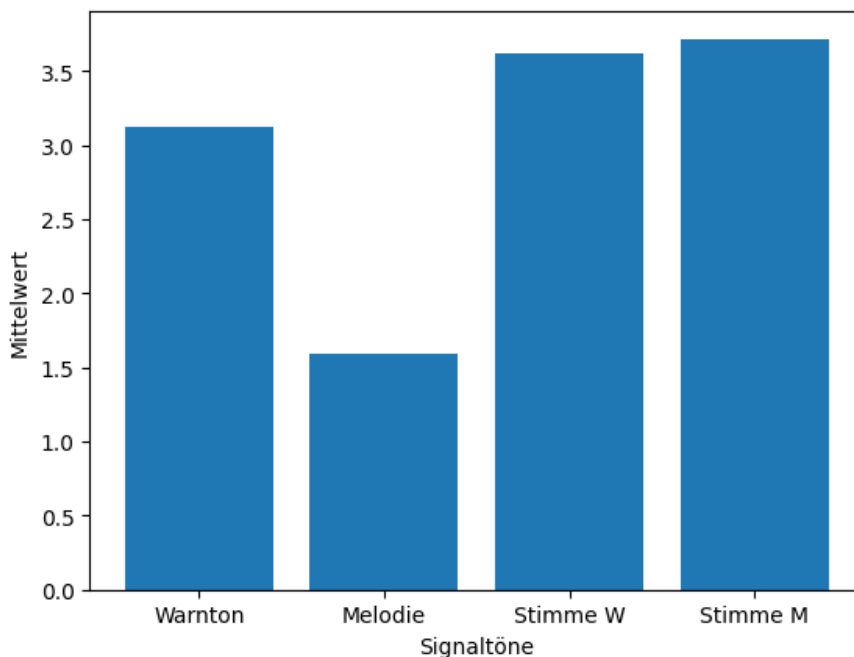


Abb. 7. Ergebnisse der Eignung verschiedener Warnsignaltöne

Eine Kommunikation mit dem Fahrzeug durch einen Sprachassistenten wurde unter den Teilnehmern am besten bewertet. Es wurden zusätzlich zwei verschiedene Varianten, um den Warnton mittels Sprache darzustellen, untersucht:

- Variante A): Sprachnachricht, die über den Systemeingriff aufgrund einer unsicheren Situation informiert (z.B. „*Achtung, die Situation ist nicht sicher. Ich leite einen Nothalt ein/übernehme die Steuerung des Fahrzeugs*“)
- Variante B): eine Sprachnachricht, die zusätzlich Information zum aktuellen Fahrerzustand gibt („*Achtung, sie sind stark gestresst/aufgebracht. Ich leite einen Nothalt ein/übernehme die Steuerung des Fahrzeugs*“)

Im Rahmen des Szenarios zu einer Systemübernahme vom manuellen zum autonomen Fahren waren die meisten Teilnehmern nach der *Variante A*) überrascht und unsicher aufgrund von mangelnden Informationen und gaben an Stress/Nervosität oder ein Gefühl des Kontrollverlusts zu verspüren ($N=14$). Andere Teilnehmer fanden allerdings die Information hilfreich und fühlten sich dadurch unterstützt und sicher ($N=6$). Die *Variante B*) löste unter den meisten Teilnehmern negative Gefühle wie mehr Stress, Nervosität, Bevormundung ($N=13$; z.B. „*Vermittelt ein schlechtes Gefühl und Irritation da einem Zurechnungsfähigkeit abgesprochen wird*“) aus. Andere hingegen fanden die Nachricht informativ und fühlten sich entlastet und verstanden ($N=6$).

Bei dem Szenario über eine Übernahmeanfrage vom autonomen zum manuellen Fahren löste die *Variante A*) unter den meisten Teilnehmer Panik und Stress ($N=17$), insbesondere aufgrund von mangelnden Informationen (z.B. „*Was mache ich gerade verkehrt?*“). Vier Probanden waren erschrocken, jedoch dankbar für die Unterstützung. Allerdings führte auch die *Variante B*) meist zu einen erhöhten Stress-/Emotionslevel ($N=18$) und die meisten

Teilnehmer fühlen sich beobachtet und bevormundet („*Ich wäre genervt, denn dass ich gestresst bin, weiß ich ja in der Situation selbst, das muss mit niemand mehr sagen*“). Andere drei Probanden fanden hingegen, dass die Aussage eine stärkere und kritischere Wirkung auf sie hatte. Bei diesem Szenario hatten zudem fünf Probanden Anmerkungen und Sorgen bezüglich des Nothaltes und wünschten sich mehr Entscheidungsfreiheit z.B. den Nothalt selbst genehmigen nach Systemanfrage.

Zusammenfassend deuten die Ergebnisse darauf hin, dass Frauen insgesamt das System besser bewerten und eher nutzen würden als Männer. Als bevorzugter Kommunikationsweg erwies sich ein Warnsignal mittels Voice (Sprachassistenten). Eine Melodie ist hingegen am wenigsten geeignet. Der Inhalt des Gesagten bei einem Warnsignal via Voice stellt ebenso einen wichtigen Faktor dar - wie die Teilnehmer das Warnsignal wahrnehmen und wie sich dieses auf ihren emotionalen Zustand auswirkt. Ein Warnsignal mit möglichst geringer Informationsmenge war bei den meisten Teilnehmern mit Panik, Stress, Überraschung und Unsicherheit verbunden. Nach einem Warnsignal mit mehr Informationen über den eigenen Zustand fühlten sich die meisten Teilnehmern allerdings bevormundet und beobachtet, was zu einem erhöhten Stress/Emotionslevel führte. Diese explorativen Ergebnisse sollten jedoch aufgrund der kleinen Stichprobe mit Vorsicht betrachtet werden. Weitere Forschung ist hierbei nötig, um den optimalen Informationsgrad sowie sprachliche Formulierungen abzuleiten, die zu möglichst wenig Frustration und negativen Emotionen führen. Auch weitere Möglichkeiten zur Personalisierung der Warnsignale könnten evaluiert werden. Allerdings soll dabei auch untersucht werden, inwiefern auch nach einer solchen Personalisierung die Fahrende die aktuelle Situation und den Systemeingriff verstehen.

AP 6 Prädiktion Fahrerzustand und Kabinenaktivität über Deep-Learning Modellierung

AP 6.3. Integration ganzheitlicher Prädiktionsmodule

In der Praxis des maschinellen Lernens gibt es verschiedene Methoden, um Modelle miteinander zu kombinieren und dadurch ihre Leistungsfähigkeit zu erhöhen. Zu den gängigen Techniken gehören unter anderem Bagging (Bootstrap Aggregating), Boosting und Stacking. Bagging verbessert die Genauigkeit eines Modells, indem es mehrere Versionen des ursprünglichen Modells auf verschiedenen Stichproben des Trainingsdatensatzes trainiert und deren Ergebnisse mittelt. Boosting erhöht die Modellleistung, indem es sequenziell mehrere Modelle trainiert, wobei jedes neue Modell die Fehler des vorhergehenden Modells korrigiert. Stacking kombiniert die Vorhersagen mehrerer Basismodelle, indem es einen Metamodel-Ansatz nutzt, der auf diesen Vorhersagen trainiert wird.

Für die Emotionserkennung kann es sinnvoll sein, das Audio- und das Videomodell durch Stacking zu kombinieren, um die Stärken beider Modalitäten auszunutzen und eine robustere Klassifikation zu erzielen. Bei der Erkennung der Fahreraufmerksamkeit (Driver Readiness) ist es jedoch am sinnvollsten, ein regelbasiertes System einzusetzen, das aus den verschiedenen Ergebnissen der Modelle einen Score berechnet. Ein solches System kann spezifische Regeln und Gewichtungen anwenden, um die unterschiedlichen Signale zu integrieren und eine zuverlässige Einschätzung der Fahreraufmerksamkeit zu gewährleisten. Deshalb wurde ein erster Prototyp eines solchen Systems in den Prototypen integriert.

AP 6.5. Evaluation Prädiktionsmodule u. Ausarbeitung weiterer Entwicklungsroadmaps

Für zukünftige Entwicklungsschritte könnte es sinnvoll sein, erneut Daten zu erheben und diese direkt auf die Erkennung der Driver Readiness zu operationalisieren. Dabei könnten die einzelnen Modelle zur Emotionserkennung und anderen relevanten Aspekten als Input für ein neues, übergeordnetes Modell dienen. Durch den Einsatz von Stacking könnten wir die Vorhersagen der einzelnen Modelle kombinieren und ein Metamodell entwickeln, das auf den neuen, speziell erhobenen Daten trainiert wird. Dies würde die Genauigkeit und Zuverlässigkeit der Erkennung der Fahreraufmerksamkeit weiter erhöhen und die Leistungsfähigkeit des Systems optimieren.

AP 7 Anwendungs- und Gesamtsystemfunktionstest

AP 7.1 Finaler Anwendungs- und Gesamtfunktionstest

Der finale Gesamtfunktionstest wurde Anfang 2024 in den Räumlichkeiten der RH Köln durchgeführt. Dabei wurde das System in einem Fahrsimulator unter Laborbedingungen mit Nutzer:innen evaluiert, um Fehler oder Optimierungsbedarfe zu identifizieren. Für die Fahrsimulation sollten die Teilnehmenden in dem Spiel Assetto Corsa mehrere Runden in einem Audi S1 über die Strecke Silverstone fahren.

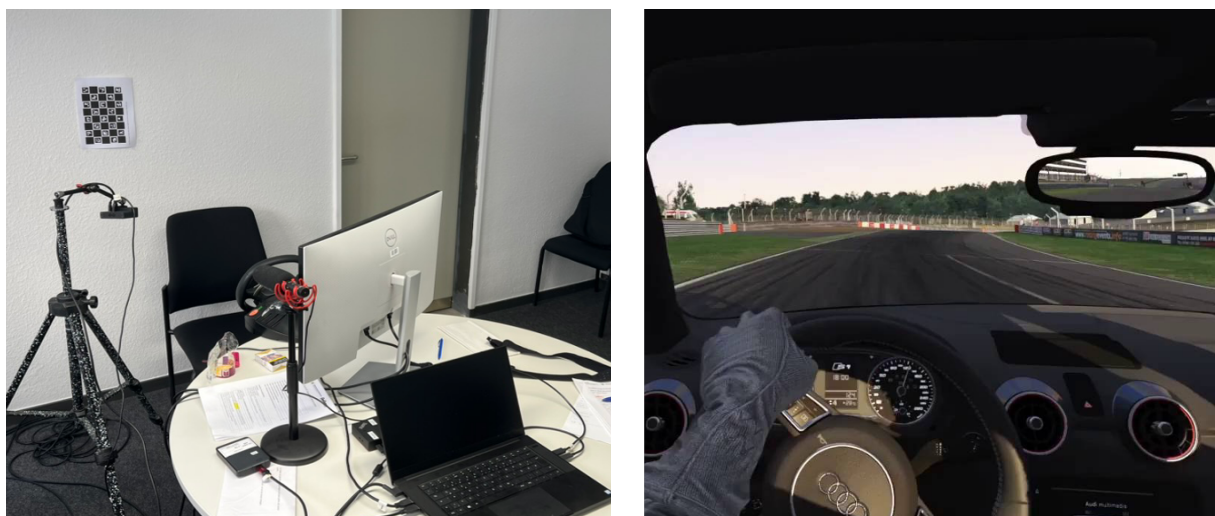


Abb. 8. Gesamtfunktionstest Setup (links) und Fahraufgabe (rechts) aus Sicht der Studienteilnehmer

Insgesamt 24 Teilnehmer (Durchschnittsalter: 28 Jahre; 13 weiblich, 8 männlich) nahmen an der Studie teil. Die Stichprobe bestand hauptsächlich aus Studierende ($N=17$) oder Auszubildende ($N=2$), dennoch nahmen auch Berufstätige ($N=4$) sowie Freiberufler ($N=1$) teil. Nach Begrüßung und kurzer organisatorischen Phase (Einwilligungserklärungen aushändigen) haben die Probanden insgesamt 8 Aufgabenblöcke bearbeitet (Abb. 9).

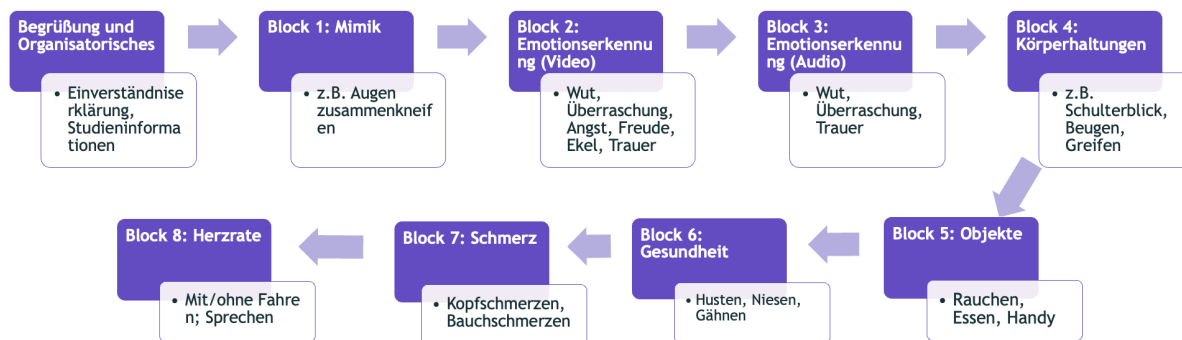


Abb. 9. Übersicht Studienablauf

Eine genaue Übersicht der einzelnen Bestandteile der Aufgaben ist unten zu finden. Die Teilnehmer wurden gebeten die jeweilige Aufgabe möglichst realistisch auszuführen. Der Gesamtdauer der Studie beträgt ca. 1 Stunde pro Person.

Block 1 Mimik:

- Augen weit geöffnet
- Augen zusammenkneifen
- Augenbrauen hochziehen/Stirn horizontal runzeln
- Augenbrauen zusammenziehen/Stirn senkrecht runzeln
- Nase rümpfen
- Weit geöffneter Mund
- Zähne zeigen (Schmerz)
- Lippe beißen/saugen
- Lächeln
- Kinnlade runter (Überraschung)
- Oberlippe; Kinn hoch

Block 2 Emotionserkennung Audio

- Neutraler Ausdruck ohne Bewegung
- Neutraler Ausdruck mit Bewegung
- Wut
- Überraschung
- Angst
- Freude
- Ekel
- Trauer

Block 3: Emotionserkennung Audio:

- Neutral 1
- Neutral 2
- Wut 1 fiktiv und Wut-Assessment (Skala 1 bis 10)
- Wut 2 Erinnerung und Wut-Assessment (Skala 1 bis 10)
- Überraschung 1 und Überraschungs-Assessment (Skala 1 bis 10)
- Überraschung 2 und Überraschungs-Assessment (Skala 1 bis 10)
- Trauer 1 fiktiv und Trauer-Assessment (Skala 1 bis 10)
- Trauer 2 Erinnerung und Trauer-Assessment (Skala 1 bis 10)

Block 4: Körperhaltungen:

- Schulterblick

- Sitzen
- Arme hinter den Kopf
- Arme vor die Brust
- Beugen, drehen, greifen

Block 5: Objektinteraktion

- Zigarette anzünden, aus dem Fenster aschen
- Snack essen
- Handy: telefonieren, Textnachricht schreiben

Block 6: Gesundheit

- Niesen mit und ohne Hand vor dem Gesicht
- Husten mit und ohne Hand vor dem Gesicht
- Gähnen mit und ohne Hand vor dem Gesicht

Block 7: Schmerzen

- Kopfschmerzen (plötzlich und allmählich einsetzende)
- Bauchschmerzen (plötzlich und allmählich einsetzende)

Block 8: Bildbasierte Erkennung der Herzrate

- Ruhig sitzen ohne fahren
- Ruhig sitzen und Autofahren
- Sprechen und Autofahren

Die Audiodaten wurden dann erneut, wie in AP 2.3 annotiert und verarbeitet. Die weiterführende Auswertung der erhobenen Daten wurde von den Projektpartnern durchgeführt, da sie an der Entwicklung der einzelnen Systembestandteile (Video- und Audio-Modelle) beteiligt sind.

AP 7.2 Usability und Nutzerakzeptanz-Studie

Der zu entwickelnde Prototyp zur Fahrerzustandserkennung verfügt über keine Benutzeroberfläche, über die Nutzer:innen mit dem System interagieren können. Aus diesem Grund erwies sich eine typische Evaluation der Nutzerakzeptanz durch Usability Testung als nicht möglich. Um die Usability und Nutzerakzeptanz zu testen, wie im AP 7.3. vorgesehen, wurde mit den Projektpartnern entschieden, eine low-fidelity nicht funktionale Smartphone-Anwendung zu konzipieren, über die Nutzer:innen mit dem System interagieren können. Beispielsweise soll die Anwendung die Fahrerzustände der vergangenen Fahrten abbilden können und Nutzer:innen können durch verschiedenen Funktionen sich genauer über eine bestimmte Strecke oder einen bestimmten Zustand weiter informieren.

Die RH Köln hat sich an einer ausführlichen Literaturrecherche beteiligt. Diese beinhaltet unter anderem die Recherche von Heuristiken für die Datenvisualisierung, z. B. Zielerfüllung, Visualisierung biometrischer Daten, Zeitreihendaten von mehrere Datenströme (Cuttone et al., 2014) sowie mögliche Visualisierungen, die selbst generierte Gesundheitsdaten darstellen (Kim, 2022). Zudem wurden Beispiele für verschiedene Visualisierungen des Fahrerzustands und ihre Anpassung an Alter, Geschlecht und Persönlichkeit (Braun et al., 2019) sowie ein Beispiel für Datenvisualisierung, die die kardiovaskuläre Psychophysiologie mit interaktiver Karte verbindet, die mit Standort, Fotos, Geschwindigkeit und Verkehrsdichte verknüpft ist (Fairclough & Dobbins, 2020), bereitgestellt. Die Entwicklung sowie die Usability-Testung der Anwendung wurde von IXP übernommen.

AP 8 Dissemination

AP 8.1 Öffentlichkeitswirksame Presse- und Medienarbeit

Das Verbundprojekt wurde im Rahmen des BMBF-Forschungstours „Ich, Zukunft und Künstliche Intelligenz“ am 08.11.2022 online vorgestellt. Des Weiteren ist das Projekt DrAlve auf der [RH-Webseite](#) präsent.

AP 8.2 Wissenschaftliche Publikationen

- Achtermann, K., Weber, F., Storch, D., Krajewski, J., Canzler, U., Karos, L., Wong, M., Schnieder, S. (2022). KI-basierte Fahrerzustandserkennung im Kontext von Übergabesituationen des teil autonomen Fahrens: Entwicklung eines Audio-Video-basierten Fahrerassistenzsystems zur Detektion sicherheitskritischer „Driver Readiness“. In Bosau, C., Kiell, G., Pieczykolan, A. (Hrsg.) *Abstracts zur 25. Fachtagung der Gesellschaft für angewandte Wirtschaftspsychologie (GWPs)*. S. 64. ISBN 978-3-95853-811-5
- Projektvorstellung im Rahmen der Mittelstandskonferenz „KMU stärken durch digitale Innovationen“ am 21 und 22 November 2023 in Berlin
- Projektvorstellung bei dem „Tag der Forschung“ an der Rheinischen Hochschule Köln am 06.09.2023. Vortrag mit dem Titel *“Fokus auf den Fahrer: Entwicklung und Evaluation eines audiovisuellen Fahrerzustandserkennungssystem“*

AP 8.3 Wissenschaftliche Nachwuchsförderung

Projektinformation und -ergebnisse sind in die Lehre der RH eingeflossen. Eine Förderung des wissenschaftlichen Nachwuchses wurde durch Projektmitarbeit in Rahmen von studentischen Praktika und Hausarbeiten gewährleistet. Im Projektzeitraum konnten keine geeigneten Kandidaten für eine Abschlussarbeit rekrutiert werden.

Die wichtigsten Positionen des zahlenmäßigen Nachweises

Für die wichtigsten Positionen des zahlenmäßigen Nachweises siehe Tabelle 4. Für weitere Auskünfte bezüglich der Verwendung der Zuwendung möchten wir an der Stelle an den Verwendungsnachweis verweisen.

Tabelle 1: Übersicht Kostenart und Ausgaben

Kostenart	Ausgaben
Personalmittel	190.001,76 €
Vergabe von Aufträgen	14.280,00 €
Sonstige Verwaltungsausgaben	4.287,46 €
Dienstreisen	308,80 €
Gesamt	208.878,02 €

Die Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Jedes Jahr gibt es viele Verkehrsunfällen. Häufig ist auch ein menschlicher Fehler wie z.B. Müdigkeit, Alkoholkonsum oder rücksichtsloses Verhalten als Ursache zu nennen. KI kann

die Sicherheit auf den Straßen erhöhen, indem der Fahrerzustand überwacht wird. Bei kritischen Situationen kann somit rasch reagiert werden. Auch Fahrer:innen können durch das System Unterstützung bekommen, indem Fahraufgaben von dem System übernommen werden können. Mit fortschreitender Entwicklung könnten KI-gesteuerte Fahrzeuge die Unfallrate reduzieren und die Verkehrssicherheit erheblich steigern. Bei der Mobilität kann das Verkehrsmanagement und Verkehrsflüsse optimiert, (teil)autonomes Fahren sicherer gemacht und die Effizienz im öffentlichen Nahverkehr verbessert werden. Aufgrund des Umgangs mit sensiblen und personenbezogenen Daten, ist es notwendig, sicherzustellen, dass die Entwicklung von KI stets von ethischen und sozialen Aspekten begleitet wird. Der verantwortungsvolle Einsatz von KI und die Berücksichtigung von Datenschutz und Privatsphäre sind unerlässlich, um das volle Potenzial der Technologie zu nutzen und gleichzeitig potenzielle Risiken zu minimieren.

Der voraussichtliche Nutzen, insbesondere die Verwertbarkeit der Ergebnisse – auch konkrete Planungen für die nähere Zukunft – im Sinne des fortgeschriebenen Verwertungsplans

Die Projektergebnisse haben für zukünftige Forschung einen hohen Nutzen. Durch die gesammelten Nutzeranforderungen können zukünftige Projekte schneller Entscheidungen zu den Besonderheiten des Systems, beispielsweise die technischen Komponenten, treffen. Die entstandene Datenbank enthält anonymisierte technische Daten, die für die Weiterentwicklung des Systems benutzt werden können oder auf Nachfrage von anderen Forschern benutzt werden könnten. Die Evaluation beinhaltet mögliche Vorlieben und Bedenken der Nutzer:innen, Gestaltungsempfehlungen, Feedback-Optionen sowie Verbesserungsvorschläge. Diese könnten in zukünftigen Projekten umgesetzt werden und somit die Nutzererfahrung verbessern.

Der während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Es wurden keine Ergebnisse Dritter bekannt, die für das vorliegende Projekt relevant sind.

Die erfolgten oder geplanten Veröffentlichungen des Ergebnisses nach Nr. 5 der NABF

Achtermann, K., Weber, F., Storch, D., Krajewski, J., Canzler, U., Karos, L., Wong, M., Schnieder, S. (2022). KI-basierte Fahrerzustandserkennung im Kontext von Übergabesituationen des teil autonomen Fahrens: Entwicklung eines Audio-Video-basierten Fahrerassistenzsystems zur Detektion sicherheitskritischer „Driver Readiness“. In Bosau, C., Kiell, G., Pieczykolan, A. (Hrsg.) *Abstracts zur 25. Fachtagung der Gesellschaft für angewandte Wirtschaftspsychologie (GWPs)*. S. 64. ISBN 978-3-95853-811-5

Literaturverzeichnis

- Abu-El-Haija, S., Kothari, N., Lee, J., Natsev, P., Toderici, G., Varadarajan, B., & Vijayanarasimhan, S. (2016). *YouTube-8M: A Large-Scale Video Classification Benchmark*.
- Batliner, A., Schuller, B., Seppi, D., Steidl, S., Devillers, L., Vidrascu, L., Vogt, T., Aharonson, V., & Amir, N. (2011). The Automatic Recognition of Emotions in Speech. In *Cognitive Technologies* (pp. 71–99). https://doi.org/10.1007/978-3-642-15184-2_6
- Belin, P., Fillion-Bilodeau, S. & Gosselin, F. The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods* 40, 531–539 (2008). <https://doi.org/10.3758/BRM.40.2.531>
- Braun, M., Chadowitz, R., & Alt, F. (2019). *User Experience of Driver State Visualizations: A Look at Demographics and Personalities* (pp. 158–176). https://doi.org/10.1007/978-3-030-29390-1_9
- Braun, M., Weber, F., & Alt, F. (2021). Affective Automotive User Interfaces – Reviewing the State of Driver Affect Research & Emotion Regulation in the Car. *ACM Computing Surveys*, 54, 137. <https://doi.org/10.1145/3460938>
- Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., & Weiss, B. (2005). A database of German emotional speech. *Interspeech 2005*. <https://doi.org/10.21437/interspeech.2005-446>
- Cuttone, A., Petersen, M. K., & Larsen, J. E. (2014). Four Data Visualization Heuristics to Facilitate Reflection in Personal Informatics. In C. Stephanidis & M. Antona (Eds.), *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice* (pp. 541–552). Springer International Publishing.
- Du, N., Ayoub, J., Zhou, F., Pradhan, A., Robert, L., Tilbury, D., Pulver, E., & Yang, X. J. (2019). Examining the impacts of drivers’ emotions on takeover readiness and performance in highly

- automated driving. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 63, 2076–2077. <https://doi.org/10.1177/1071181319631391>
- Eyben, F., Wöllmer, M., & Schuller, B. (2010). Opensmile. *Proceedings of the 18th ACM International Conference on Multimedia*. <https://doi.org/10.1145/1873951.1874246>
- Fairclough, S. H., & Dobbins, C. (2020). Personal informatics and negative emotions during commuter driving: Effects of data visualization on cardiovascular reactivity & mood. *International Journal of Human-Computer Studies*, 144, 102499. <https://doi.org/10.1016/j.ijhcs.2020.102499>
- Kim, S.-H. (2022). A Systematic Review on Visualizations for Self-Generated Health Data for Daily Activities. *International Journal of Environmental Research and Public Health*, 19(18). <https://doi.org/10.3390/ijerph191811166>
- Müller, T., & Kreutz, D. (2021). *Thorsten-Voice Dataset 2021.06 emotional*(Version 2.0) [Data set]. doi:10.5281/zenodo.5525023
- Paulus, C. (2009). Der Saarbrücker Persönlichkeitsfragebogen SPF(IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index. [Http://www.Uni-Saarland.de/Fak5/Ezw/Personal/Paulus/Empathy/SPF_Artikel.Pdf](http://www.Uni-Saarland.de/Fak5/Ezw/Personal/Paulus/Empathy/SPF_Artikel.Pdf).
- Schuller, B., Batliner, A., Steidl, S., & Seppi, D. (2011). Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. *Sensing Emotion and Affect - Facing Realism in Speech Processing*, 53(9), 1062–1087. <https://doi.org/10.1016/j.specom.2011.01.011>
- Zepf, S., Hernandez, J., Schmitt, A., Minker, W., & Picard, R. (2020). Driver Emotion Recognition for Intelligent Vehicles: A Survey. *ACM Computing Surveys*, 53, 1–30. <https://doi.org/10.1145/3388790>