



Künstliche Intelligenz für Adaptive, Responsive und Levelkonforme Interaktion im Fahrzeug der Zukunft

Abschlussbericht

Berichtszeitraum

01. Juli 2021 – 30. September 2024

Autoren

Lesley-Ann Mathis, Harald Widlroither (Fraunhofer IAO)
Frederik Diederichs, David Lerch, Manuel Martin (Fraunhofer IOSB)

Zuwendungsempfänger



Förderkennzeichen

19A21031F (Fraunhofer IAO)

19A21031F (Fraunhofer IOSB)

Vorhabensbezeichnung und Laufzeit

Künstliche Intelligenz für Adaptive, Responsive und Levelkonforme Interaktion im Fahrzeug der Zukunft (KARLI) vom 1. Juli 2021 bis 30. September 2024

Gefördert durch:



Finanziert von der
Europäischen Union
NextGenerationEU

Gefördert durch:



Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

Projekträger:



TÜVRheinland®

I Inhaltsverzeichnis

I	Inhaltsverzeichnis	II
II	Abbildungsverzeichnis	IV
III	Tabellenverzeichnis	VI
IV	Abkürzungsverzeichnis	VII
1	Kurzdarstellung	2
1.1	Aufgabenstellung	2
1.1.1	Fraunhofer IAO.....	2
1.1.2	Fraunhofer IOSB	4
1.2	Voraussetzungen	5
1.2.1	Fraunhofer IAO.....	5
1.2.2	Fraunhofer IOSB	6
1.3	Planung und Ablauf des Vorhabens.....	6
1.3.1	Fraunhofer IAO.....	6
1.3.2	Fraunhofer IOSB	7
1.4	Wissenschaftlicher und technischer Stand, an den angeknüpft wurde	7
1.5	Zusammenarbeit mit anderen Stellen	9
1.5.1	Fraunhofer IAO.....	9
1.5.2	Fraunhofer IOSB	9
2	Eingehende Darstellung der Ergebnisse	10
2.1	Verwendung der Zuwendung, erzielte Ergebnisse.....	10
2.1.1	Fraunhofer IAO.....	10
2.1.1.1	Arbeitspaket AP 100 (UAP 110, 120)	11
2.1.1.2	Arbeitspaket AP 200 (UAP 210, 220, 230)	13
2.1.1.2.1	Versuchsträger des Fraunhofer IAO	13
2.1.1.2.2	Datenerhebung im Realverkehr (LKV).....	14
2.1.1.3	Arbeitspaket AP 600 (UAP 610, 630, 640)	19
2.1.1.4	Arbeitspaket AP 700 (UAP 730).....	22
2.1.1.5	Arbeitspaket AP 800 (UAP 810, 820)	23
2.1.1.6	AP-Koordination, Bereitstellung von Verfahren.....	25
2.1.2	Fraunhofer IOSB	26
2.1.2.1	Arbeitspaket AP 100 (UAP 120).....	26
2.1.2.2	Arbeitspaket AP 200 (UAP 210, 220, 230)	27
2.1.2.3	Arbeitspaket AP 300 (UAP 312, 321)	29
2.1.2.4	Arbeitspaket AP 500 (UAP 510, 520)	33
2.1.2.5	Arbeitspaket AP 600 (UAP 620, 630, 640)	35

2.1.2.6	Arbeitspaket AP 700 (UAP 720, 730)	36
2.1.2.7	Arbeitspaket AP 800 (UAP 810, 820)	38
2.1.2.8	AP-Koordination, Bereitstellung von Verfahren	46
2.2	Zahlenmäßiger Nachweis	46
2.2.1	Fraunhofer IAO	46
2.2.2	Fraunhofer IOSB	46
2.3	Notwendigkeit und Angemessenheit der geleisteten Arbeit	46
2.3.1	Fraunhofer IAO	46
2.3.2	Fraunhofer IOSB	47
2.4	Voraussichtlicher Nutzen, Verwertbarkeit der Ergebnisse	47
2.4.1	Fraunhofer IAO	47
2.4.2	Fraunhofer IOSB	47
2.5	Fortschritt bei anderen Stellen	48
2.6	Erfolge und geplante Veröffentlichungen	49
VII	Literaturverzeichnis	50

II Abbildungsverzeichnis

Abbildung 1: KARLI-Nutzerrollen, in Anlehnung an die SAE-Level.....	12
Abbildung 2: White List - Darstellung von levelkonformen FFT aufgeteilt in die 5 KARLI Nutzerrollen (K-R1 - K-R5)	13
Abbildung 3: Studienablauf zur Erhebung von Trainingsdaten im Realverkehr unter der Berücksichtigung verschiedener Automatisierungsstufen und FFT.....	15
Abbildung 4: Studiensetup der LKV-Datenerhebung im Realverkehr mit dem WoZ Fahrzeug....	16
Abbildung 5: Hin- und Rückfahrt der Datenerhebung im Realverkehr (Stuttgart - Pforzheim), unterteilt in die verschiedenen Streckenabschnitte zur Simulation der KARLI Nutzerrollen.....	17
Abbildung 6: Exemplarische Darstellung der "Mission" während einer automatisierten Fahrt in KR-3.	18
Abbildung 7: Wizard-of-Oz Fahrzeug des Fraunhofer IAO, als Versuchsträger der Applikation LKV und MS	20
Abbildung 8: Hardwareübersicht des Wizard-of-Oz Fahrzeug.....	21
Abbildung 9: Demonstration des WoZ-Fahrzeugs auf der AbP.....	23
Abbildung 10: Demonstration des Projekts KARLI und den Einsatz des WoZ Fahrzeugs auf dem Fraunhofer IZS Campus in Stuttgart	24
Abbildung 11: Präsentation von KARLI Studienergebnissen auf der AHFE2023 in der Session „Vehicle Automation“	24
Abbildung 12: Präsentation von KARLI Studienergebnisse auf der CUI2024	25
Abbildung 13: Präsentation von KARLI Studienergebnisse auf der Mensch und Computer 2024	25
Abbildung 14: Personas die im Rahmen des KARLI Projekts entwickelt wurden	27
Abbildung 15: Gastvortrag von Vicomtech zum Thema Labelingtools und Semi-automatisches Labeling	28
Abbildung 16: Screenshot aus dem Vicomtech Labeling-Tool. Für jedes Frame im Video lassen sich Aktivitäten, Objekte und wichtige Punkte festlegen.	29
Abbildung 17: Aufzeichnungssystem des Fraunhofer IOSB integriert im Versuchsträger des Fraunhofer IAO.....	30
Abbildung 18: Innenraumkamera des Fraunhofer IOSB im Versuchsträger des Projektpartners Continental.....	30
Abbildung 19: Evaluation der Reaktionszeiten für ein statisches und ein generiertes User Interface.....	37
Abbildung 20: Evaluation der Fehlerquoten für ein statisches Interface (Simulation A) und ein generiertes Interface (Simulation B).....	37
Abbildung 21: Auswertung der Aktivitätenerkennung für „telefonieren“ vom Projektpartner Continental.....	38
Abbildung 22: Auswertung der Aktivitätenerkennung für „trinken“ vom Projektpartner Continental.....	38
Abbildung 23: Ergebnisverbreitung bei der ECCV 2022	39
Abbildung 24: Ergebnisverbreitung bei der CarHMI 2023.....	40
Abbildung 25: Ergebnisverbreitung bei der InCabin Konferenz 2024.....	40
Abbildung 26: Ergebnisverbreitung bei der CarHMI 2024.....	41
Abbildung 27: Ergebnisverbreitung bei der ITSC 2024	42
Abbildung 28: Ergebnisverbreitung bei der InCabin.Sensing 2024.....	42
Abbildung 29: Demonstration eines KI-Assistenten für den Fahrzeuginnenraum beim KARLI-Konsortialmeeting in Karlsruhe am Fraunhofer IOSB.	43

Abbildung 30: Screenshot des KI-Assistenten mit dem KARLI Avatar des Projektpartners Studiokurbos.....	43
Abbildung 31: Screenshot des KI-Assistenten für den Fahrzeuginnenraum. Der Assistent erkennt Motion Sickness induzierende Aktivitäten und gibt Hilfestellungen basierend auf der aktuellen Aktivität.....	44
Abbildung 32: Demonstrator des Motion Sickness Assistenten auf der KARLI Abschlusspräsentation.....	45
Abbildung 33: Screenshot der Demonstration des KI-Assistenten auf der KARLI Abschlusspräsentation. Der Assistent erkennt die Motion Sickness induzierende Aktivität „lesen“ und warnt den Fahrer vor dem Motion Sickness induzierenden „Kreisverkehr“.....	45

III Tabellenverzeichnis

Tabelle 1: Meilensteine und erzielte Ergebnisse während der Projektlaufzeit (Fraunhofer IAO)..	11
Tabelle 2: Messtechnik und Verwendungszweck für LKV-Datenerhebung und MS Studiendurchführung	22
Tabelle 3: Meilensteintabelle Fraunhofer IOSB.....	26

IV Abkürzungsverzeichnis

FFT	<i>Fahrfremde Tätigkeit</i>
IAO	<i>Institut für Arbeitswirtschaft und Organisation</i>
IOSB	<i>Institut für Optronik, Systemtechnik und Bildauswertung</i>
KI-I	<i>KI-Interaktion</i>
LKV	<i>Levelkonformes Verhalten</i>
MS	<i>Motion Sickness</i>
UAP	<i>Unterarbeitspaket</i>
WoZ	<i>Wizard-of-Oz</i>

1 Kurzdarstellung

Die Beiträge der Fraunhofer Gesellschaft zum Projekt KARLI lassen sich in zwei Themengebiete untergliedern. Diese beinhalten zum einen die Bereitstellung einer Versuchsumgebung zur realitätsnahen Datenerhebung mit dem Wizard-of-Oz Fahrzeug (WoZ) des Fraunhofer Instituts für Arbeitswirtschaft und Organisation (IAO), sowie die Entwicklung und Erprobung eines praxisnahen Kamera Setups zur Detektion von Fahrerzuständen, entwickelt vom Fraunhofer Institut für Optronik, Systemtechnik und Bildauswertung (IOSB).

Das **Fraunhofer IAO** verfolgte dabei in KARLI, im Rahmen der Applikation Levelkonformes Verhalten (LKV), das Ziel möglichst variantenreiche und realitätsnahe Daten zu erheben. Diese wurden von den Projektpartnern Continental und Fraunhofer IOSB als Trainingsdaten für die KI-Entwicklung weiterverarbeitet. Um eine möglichst realitätsnahe Datenerhebung zu gewährleisten, wurde das Versuchsfahrzeug des Fraunhofer IAO (WoZ Fahrzeug) über die Projektlaufzeit hinweg umgebaut, mit der entsprechenden Hardware der Projektpartner aufgerüstet sowie der Datenerhebungsprozess definiert.

Die **Fraunhofer IOSB** hat ein neuartiges Stereokamerasystem zur 3D-Erfassung von Körperpositionen entwickelt. Dieses wurde in Fahrzeugen der kooperierenden Unternehmen Continental, Ford sowie Fraunhofer IAO erfolgreich implementiert und umfassend evaluiert. Die Verarbeitung der umfangreichen Videodaten erfolgt mittels der im Projekt beschafften GPU-Workstation. Das Fraunhofer IOSB erzielte Fortschritte bei der Modellentwicklung und optimierte die Aktivitätenerkennung mithilfe von semi-supervised learning. Zur Datenannotation wurde das Vicomtech-Tool für manuelles und automatisiertes Labeling eingesetzt. Für die videobasierte Innenraumanalyse wurden Vision-Language-Modelle verwendet.

Nachfolgend sind die Aufgabenstellungen sowie Ergebnisse nach Projektpartnern Fraunhofer IAO und IOSB dargestellt und beschrieben.

1.1 Aufgabenstellung

1.1.1 Fraunhofer IAO

Im Rahmen der Applikation LKV wurde im **AP100 Theoriebildung** zunächst eine Literaturrecherche durchgeführt, um den aktuellen Stand der Forschung bezüglich rechtlicher und empirischer Betrachtung von unterschiedlichen Levels der Automatisierung abzuleiten. Ziel hierbei war es, zusammen mit den anderen Projektpartnern der Applikation LKV, nutzerzentrierte Automatisierungsstufen abzuleiten. Anhand dieser sollte im weiteren Projektverlauf levelkonformes und nicht levelkonformes Verhalten definiert werden.

Darauf basierend wurden Erhebungsszenarien für die Datenerhebung mit dem WoZ Fahrzeug definiert, mit dem Ziel levelkonforme und nicht levelkonforme fahrfremde Tätigkeiten (FFT) während verschiedener Automatisierungsstufen abzudecken. In engem Austausch mit den anderen Projektpartnern der Applikation sollte so der Umfang der zu erhebenden Realdaten, sowie Anforderungen an die Realdaten mit dem WoZ Fahrzeug definiert werden.

Die aus der Literaturrecherche abgeleiteten Erhebungsszenarien sollten anschließend im Rahmen des **AP200 Erzeugung einer Datenbasis für KI-Algorithmen** umgesetzt werden. Das AP200 gliederte sich hierbei in mehrere Unterarbeitspakete (UAP), die die Definition des

Erhebungsdesigns, sowie den Prozess der Datenerhebung bis hin zum Labeling beschreiben. Die Aufgaben des Fraunhofer IAO bestanden hierbei in der Datenerhebung im Realverkehr mit dem WoZ Fahrzeug, welches im Rahmen des AP600 entsprechend umgebaut und angepasst wurde. Der Fokus der Datenerhebung lag auf der Erhebung eines möglichst variantenreichen und realitätsnahen Datensatz, der verschiedene Automatisierungsstufen und FFT beinhaltete. Aus diesem Grund wurde mit einer iterativen Datenerhebung geplant, welche häufige Messwiederholungen über einen längeren Zeitraum beinhaltete.

Die dafür notwendigen Arbeitsschritte sahen dabei wie folgt aus:

- Definition der Szenarien
- Abstimmung des Fahrzeugaufbaus und der Messtechnik im Rahmen des AP600
- Feinplanung und Wizard Training der Szenarien
- Probandenakquise
- Planung und Erhebung
- Vorversuche und Voranalyse der Daten

Die Datenerhebung wurde gemeinsam mit dem Fraunhofer IOSB und Continental abgestimmt, um die Anforderungen an die KI-Trainingsdaten zu definieren. Das WoZ Fahrzeug wurde außerdem für die Datenerhebung der Applikation Motion Sickness (MS) und KI-Interaktion (KI-I) gemeinsam mit dem IAT der Universität Stuttgart genutzt.

Eine weitere Aufgabe des Fraunhofer IAO lag darin, in regelmäßigen Abstimmungsrunden zum UAP230 Labeling und Datenaufbereitung, gemeinsam mit den anderen Projektpartnern eine Labelingstrategie und die dafür notwendigen Voraussetzungen an die Datenstruktur zu definieren. Hierbei sollten verschiedene Labelingtools getestet werden sowie ein Goldstandard für das Datenlabeling der Zielapplikation LKV festgelegt werden.

Um das WoZ Fahrzeug den Datenerhebungen entsprechend anzupassen, lagen die Aufgaben des Fraunhofer IAO in **AP600 Integration und Absicherung des KI-Gesamtsystems** darin eine Integrationsarchitektur festzulegen und damit eine einheitliche Datenerhebungs- und Synchronisierungsstruktur zu gewährleisten. Um die Simulation verschiedener Automatisierungsstufen sowie Transitionen während der Fahrt durchführen zu können, sollte zunächst mit dem Hersteller des WoZ Fahrzeugs (Paravan) die technischen Komponenten des Fahrzeugs umgebaut und angepasst werden. Das anschließende Architekturkonzept, bestehend aus integrierender Soft- und Hardware sollten gemeinsam mit einem Unterauftragnehmer umgesetzt werden. Für die Integration der Hardware umfassten die Aufgaben die Auswahl und Beschaffung der Messtechnikkomponenten, die kontinuierliche Qualitätssicherung der damit erhobenen Daten, die Anpassung der Messtechnik und Spezifikation an unterschiedliche Erhebungsanforderungen, sowie die Koordination mit anderen Partnern und anderen Arbeitspaketen (AP). Die Auswahl geeigneter Software bezog sich auf die Datensynchronisierung und Speicherung, sowie die Abstimmung der Schnittstellen zur Datenkommunikation zwischen verschiedenen Messtechnikkomponenten.

Die erzielten Ergebnisse, der Fahrzeugumbau des WoZ Versuchsfahrzeugs und die Datenerhebung der Zielapplikation LKV sollten auf der Halbzeit- und Abschlusspräsentation sichtbar gemacht werden. Im Rahmen des **AP700 Abschließende Evaluation und Demonstration** lagen die Aufgaben daher in der Planung und Umsetzung der Demonstrationen, mit dem Ziel die Datenerhebung und Weiterverwendung der erhobenen Daten erlebbar und verständlich zu machen. Daher sollten die verschiedenen Datenströme visuell sichtbar gemacht werden und die technischen Spezifikationen des WoZ Fahrzeugs sowie der Einsatz im Realverkehr dargestellt werden.

Die erzielten Ergebnisse sollten dabei nicht nur auf der Halbzeit- und Abschlusspräsentation gezeigt werden, sondern im Rahmen des **AP800 Projektmanagement und Ergebnisverbreitung** auch über das KARLI Projekt hinaus auf Konferenzen und Social-Media-Kanälen.

1.1.2 Fraunhofer IOSB

Im Rahmen des Projekts hatte das Fraunhofer IOSB eine zentrale Rolle in der Entwicklung und Umsetzung der videobasierten Innenraumanalyse für die Applikationen LKV und MS inne. Im Rahmen des AP 100 erfolgte eine Mitwirkung bei der Erstellung der Use Cases und Erhebungsszenarien für die Anwendungen LKV und MS. Das Fraunhofer IOSB fungierte hierbei als Beratungsinstanz hinsichtlich der Erfassbarkeit gewünschter Merkmale mit videobasierten Systemen. Gegenstand der Beratung war die Prüfung der technischen Umsetzbarkeit von Merkmalen der Innenraumerfassung, wie etwa Körperbewegungen oder Kopfpositionen, für die geplanten Anwendungen. Die frühzeitige Beratung sicherte die Ausrichtung der nachfolgenden technischen Entwicklungen an realistischen und durchführbaren Anforderungen.

Im Rahmen des AP 200 erfolgte die Erstellung und Bereitstellung eines umfassenden Datensatzes, welcher für die Entwicklung und das Testen der Verfahren zur videobasierten Innenraumanalyse für beide Anwendungen herangezogen wurde. In einem ersten Schritt erfolgte durch das Fraunhofer IOSB eine Sichtung der aufgezeichneten Daten aus den Applikationen LKV und MS. Dabei wurde eine Bewertung hinsichtlich deren Eignung für die beabsichtigten maschinellen Lernverfahren vorgenommen. Im Anschluss wurde die Datenbasis, welche aus Videomaterial und weiteren Quellen besteht, im Hinblick auf die Anforderungen der Innenraumanalyse analysiert.

Ein wesentlicher Bestandteil dieses Arbeitspakets war die Erstellung einer Labeling-Hierarchie, welche auf den zuvor definierten Anforderungen der beiden Anwendungen basierte. Dies ermöglichte eine konsistente und systematische Annotation der Daten, wodurch die Qualität der darauf aufbauenden maschinellen Lernmodelle maximiert werden konnte. Dabei erfolgte eine Abstimmung mit den Anforderungen der Innenraumanalyse, sodass sowohl die Erkennung von Motion Sickness als auch das levelkonforme Verhalten korrekt und vollständig abgebildet wurden. Das Fraunhofer IOSB führte das Labeling der Videodaten anhand von Goldstandards durch, welche aus anderen Quellen stammten, und ergänzte dies, wo erforderlich, durch manuelles oder semiautomatisches Labeling. Im Anschluss wurden die gelabelten Daten in Trainings-, Validierungs- und Testdatensätze für das überwachte Lernen (Supervised Learning) aufgeteilt. Des Weiteren wurden Daten ohne Label für das Verfahren des "Semi-supervised Learning" definiert, um die Anwendungsmöglichkeiten des Modells zu erweitern. Der so erstellte Datensatz wurde schließlich dokumentiert und den Projektpartnern zur Verfügung gestellt. Ein weiterer wesentlicher Output dieses Arbeitspakets war die Publikation der Daten in einer wissenschaftlichen Zeitschrift, wodurch die Verbreitung der Projektergebnisse in der wissenschaftlichen Gemeinschaft sichergestellt wurde.

Im Rahmen des Arbeitspakets 300 erfolgte die Umsetzung der Anforderungen der Anwendung LKV an die videobasierte Innenraumanalyse. Die entwickelten Hardware- und Algorithmenkonzepte basierten auf der Aggregation der Anforderungen für die Erfassung des levelkonformen Verhaltens. Zu diesem Zweck konzipierte das Fraunhofer IOSB ein Kamerasystem, welches den spezifischen Anforderungen gerecht wurde. Diesbezüglich sind insbesondere die Erkennung von Körperpose, Kopforientierung und Nebentätigkeiten zu nennen.

In der Folge entwickelte das Fraunhofer IOSB ein echtzeitfähiges Verarbeitungskonzept, welches verschiedene maschinelle Lernverfahren nutzte. In einem ersten Schritt wurden überwachte Lernverfahren implementiert, um die Daten auszuwerten und die Verhaltensweisen der Fahrzeuginsassen zu analysieren. Die erste Entwicklungsstufe diente als Grundlage zur korrekten Erfassung des Verhaltens der Insassen in Echtzeit. In einer zweiten Phase wurde untersucht,

inwiefern semi-supervised und unsupervised Learning-Methoden zur Verbesserung der Algorithmen eingesetzt werden können. Ziel war die weitere Optimierung der Modelle durch eine effizientere Nutzung ungelabelter Daten.

Im Rahmen der Kooperation mit der Entwicklung für die Anwendung MS wurde das Ziel verfolgt, Synergien zwischen beiden Anwendungsbereichen zu nutzen.

Gleichzeitig wurden die Anforderungen der Anwendung MS im Rahmen des AP 500 durch das Fraunhofer IOSB umgesetzt. Des Weiteren wurden die im Rahmen der videobasierten Innenraumanalyse entwickelten Konzepte in Hardware und Algorithmen überführt. Analog zum Vorgehen im AP 300 wurden auch hier überwachte Lernverfahren implementiert, um die erforderlichen Daten zu verarbeiten. Im zweiten Schritt wurden Methoden des semi-supervised und unsupervised Learnings eingesetzt, um die Effizienz der Modelle weiter zu steigern. Das Ziel dieser Anwendung bestand in der Entwicklung eines Systems, welches in Echtzeit Signale erkennt, die auf eine beginnende Motion Sickness hindeuten, wie beispielsweise unruhiges Verhalten oder Kopfdrehungen. Auch hier wurden Synergien mit der Entwicklung für das levelkonforme Verhalten genutzt, um die Prozesse effizient zu gestalten.

Das AP 600 hatte die Entwicklung der Systemarchitektur für das WoZ-Fahrzeug des Fraunhofer IAO zum Gegenstand, welches für die Datenerfassung zum Einsatz kam. In enger Zusammenarbeit mit dem Fraunhofer IAO entwickelte das Fraunhofer IOSB Synchronisations- und Aufzeichnungskonzepte für die videobasierte Innenraumanalyse. Das Aufzeichnungssystem wurde in gemeinsamer Konzeption und Implementierung erstellt. Im Anschluss integrierte das Fraunhofer IOSB die entwickelten Systeme zur Innenraumanalyse in die Versuchsträger und kalibrierte die Kameras für die jeweiligen Fahrzeugsysteme.

Im abschließenden AP 700 erfolgte eine Evaluierung der entwickelten Technologie sowie eine Anpassung für die Demonstration in den Versuchsträgern. Die Algorithmen zur Aktivitätenerkennung und zur Vermeidung von Motion Sickness wurden auf Basis der Testdaten validiert und dokumentiert, um eine zuverlässige und praxistaugliche Funktionsweise der Systeme sicherzustellen. Die Verfeinerung und Optimierung der entwickelten Modelle diente der Vorbereitung für deren Einsatz in zukünftigen Anwendungen.

Das Fraunhofer IOSB übernahm in KARLI auch die wissenschaftliche und technische Koordination des Konsortiums.

1.2 Voraussetzungen

1.2.1 Fraunhofer IAO

Das Fraunhofer IAO aus Stuttgart wurde 1981 gegründet und steht mit rund 600 Wissenschaftlerinnen und Wissenschaftlern für angewandte Forschung mit dem Menschen im Mittelpunkt von Technologie und Organisation. In großen Initiativen wie dem Cyber-Valley und dem KI-Fortschrittszentrum treibt das IAO und die Arbeitsgruppe Mensch-Technik Interaktion viele Themen rund um Human-Centered AI. Das Team Ergonomics & Vehicle Interaction befasst sich seit 30 Jahren mit nutzerzentrierten Innovationen im Fahrzeug als Auftragsforschung für OEM und Tier1-2 Supplier sowie in zahlreichen Konsortialprojekten. Schwerpunkte sind Fahrerzustandserfassung in Fahrsimulatoren und Realfahrversuchen sowie Durchführung von menschenbezogener Datenerhebungen für KI-Anwendungen. Das Fraunhofer IAO bringt ein Wizard-of-Oz-Fahrzeug mit Paravan Steer-by-Wire Technologie ins Projekt ein, welches es erlaubt verschiedene Level des automatisierten Fahrens, sowie Übergaben zwischen automatisierten und manuellen Fahren durchzuführen. Des Weiteren kann es für die Datenerhebung zur Fahrerzustandsbewertung beliebig mit Sensorik ausgerüstet werden. Des Weiteren bringt das Institut einen Fahrsimulator

auf Basis eine Porsche Macan mit der Simulationssoftware Silab in das Projekt ein, mit dem bereits langjährige Erfahrungen in der Fahrerzustandsbewertung und HMI-Evaluation bestehen.

1.2.2 Fraunhofer IOSB

Das Fraunhofer IOSB aus Karlsruhe beschäftigt rund 700 Wissenschaftlerinnen und Wissenschaftler. Es entstand 2010 durch Fusion von Vorgängerinstituten. Das Institut deckt die komplette Prozess- und Verwertungskette von der Entwicklung neuartiger Sichtsysteme, dem optimalen Einsatz von Sensoren und der Verwertung der entstehenden Datenströme ab. Auf dieser Basis entwickeln die Wissenschaftlerinnen und Wissenschaftlern des Instituts Systeme, die Menschen effizient unterstützen, fundierte Entscheidungen treffen, Prozesse optimieren und autonome Systeme intelligent steuern. Im Jahr 2019 generierte das Institut hierdurch €64 Mio. Umsatz.

In der Abteilung „Human-AI Interaction“ (HAI) werden innovative Interaktionsmethoden und Assistenzsysteme entwickelt, die Menschen bei ihren Aufgaben unterstützen. Mit der Entwicklung einer kamerabasierten Wahrnehmung und angepassten Nutzerschnittstellen liegt der Fokus insbesondere auf der Erfassung von Personen und der Auswertung ihrer Aktivitäten für die multimodale MMI in intelligenten und proaktiven Umgebungen. Im BMBF-Projekt InCarIn und PAKoS konnte das IAD die Körperpose und Interaktionen aller Insassen erfassen und an personalisierte Verhaltensweisen, inkl. Fahrtübernahmen anpassen. In KARLI werden diese Kompetenzen eingebracht und weiter ausgebaut. Hierfür kommen bei der internen Verfahrensentwicklung zusätzlich zu den Versuchsträgern im Projekt auch die Simulatoren des Fraunhofer IOSB und die Rechenkapazitäten für das Training komplexer maschineller Lernverfahren zum Einsatz. Zusammen mit dem Fraunhofer IAO baut das Fraunhofer IOSB das Wizard-of-Oz-Fahrzeug für KARLI auf, steuert die Fahrzeugmesstechnik und unterstützt die Datenerhebung im Realverkehr und auf Teststrecken. Kompatibel zum Fraunhofer IAO Fahrsimulator steht auch am IOSB ein Fahrsimulator auf SILAB Basis für das Projekt zur Verfügung.

1.3 Planung und Ablauf des Vorhabens

1.3.1 Fraunhofer IAO

In der Applikation LKV fanden alle 4 Wochen Regelmeetings statt, um den aktuellen Stand mit den jeweiligen Projektpartnern auszutauschen und Zwischenergebnisse zu präsentieren. Zur Abstimmung der Datenerhebung im Realverkehr mit dem WoZ Fahrzeug fanden außerdem weitere Abstimmungen mit den Projektpartnern Continental und Fraunhofer IOSB statt. Hierbei sollte zum einen der Einbau der Head-Eye Tracking Kamera (Continental), sowie der Innenraumkamera zur Aktivitätenerkennung (Fraunhofer IOSB) und die dafür notwendigen Schnittstellen besprochen und definiert werden. Des Weiteren wurde auch das Studiendesign der Datenerhebung, sowie die Anforderungen an die Daten besprochen, sodass diese anschließend als Trainingsdatensätze verwendet werden konnten.

Im Rahmen des UAP230 fanden Regelmeetings statt, um den Fortschritt zum Labeling und Datenerhebung mit den jeweiligen Projektpartnern zu besprechen. Diese Meetings wurden dazu genutzt, ein einheitliches Labelingformat, sowie Schnittstellen zwischen Sensorik und Middleware in den verschiedenen Versuchsfahrzeugen für die Datenerhebungen zu definieren.

Der Fahrzeugumbau des WoZ Fahrzeugs hinsichtlich verbesserter Steuerung vom vorderen, sowie hinteren Bereich des Fahrzeugs und der Durchführung von Transitionen während der Fahrt, wurde im Rahmen eines Unterauftrags mit dem Hersteller Paravan geprüft und durchgeführt. Für die Integration von Messtechnik, sowie weiterer Soft- und Hardware wurde ein Unterauftrag ausgeschrieben, da diese technischen Kompetenzen nur bis zu einem gewissen Grad am IAO vorlagen. Der Unterauftrag wurde an die Firma AIOCAS Consulting & Solutions vergeben. Im Verlauf des AP600 konnten somit alle Anforderungen an den Fahrzeugumbau, sowie die Integration der Messtechnik und einer gemeinsamen Middleware umgesetzt werden.

Über den Projektverlauf waren vier große Meilensteine geplant und entsprechend der Projektdauer wie folgt festgelegt:

- MS1: Monat 9
- MS2: Monat 18, Halbzeitpräsentation (HzP)
- MS3: Monat 27
- MS4: Monat 36, Abschlusspräsentation (AbP)

Die Zielerreichung und Abweichungen sind in Kapitel 2.1.1 dargestellt (Tabelle 1).

1.3.2 Fraunhofer IOSB

Das Projekt verfolgt einen modularen und agilen Projektansatz mit iterativer Nutzereinbindung und kontinuierlicher Verbesserung der KI-Modelle anhand laufend erhobener Daten.

Die Implementierung des Datenaufzeichnungssystems für das WoZ-Fahrzeug des Fraunhofer IAO (MS1) markierte einen entscheidenden ersten Meilenstein im Projektverlauf. Darauf aufbauend wurde das Labelingkonzept für die videobasierte Innenraumanalyse erfolgreich implementiert, und erste Daten konnten annotiert werden (MS2). Parallel dazu erfolgte die Entwicklung und erste Demonstration der Verfahren, wodurch deren technische Machbarkeit validiert wurde. Aufgrund verzögerter Verfügbarkeit der Karli-Daten wurden diese Verfahren zunächst auf öffentlichen Datensätzen entwickelt, um den Forschungsfortschritt sicherzustellen.

Im weiteren Verlauf wurden die Algorithmen durch den Einsatz von Semi-supervised und Unsupervised Learning weiter optimiert (MS3), wodurch sowohl die Robustheit als auch die Effizienz der Verfahren signifikant gesteigert werden konnten. Schließlich erfolgte die Integration der finalisierten Verfahren in die Versuchsträger, gefolgt von einer erfolgreichen Demonstration (MS4). Aufgrund einer dreimonatigen Projektverlängerung wurde dieser Meilenstein im Projektmonat 39 erreicht. Durch diese iterative Vorgehensweise konnte eine sukzessive Verbesserung der entwickelten Systeme gewährleistet und deren Anwendbarkeit in realen Szenarien fundiert validiert werden.

Die Zielerreichung und Abweichungen sind im Kapitel 2.1.2 dargestellt.

1.4 Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Die Entwicklung von SAE-Level-3-Systemen hin zu SAE-Level-4- oder sogar SAE-Level-5-Systemen ermöglicht es dem Fahrer, die Fahrzeugsteuerung für vordefinierte Strecken oder sogar während der gesamten Fahrt an das System zu übertragen. Dadurch kann der Nutzer mehr Zeit für fahrfremde Tätigkeiten wie Arbeiten, Lesen oder sogar Schlafen während der Fahrt aufwenden (Shi et al., 2020). Je nach Automatisierungsgrad des Fahrzeugs muss die Aufmerksamkeit des Nutzers dabei mehr oder weniger auf die Straße gerichtet sein. Aufgrund intransparenter Spezifikationen levelkonformen und nicht levelkonformen Verhaltens in einem bestimmten

Automatisierungsgrad ist es für den Nutzer oft schwierig, sich richtig zu verhalten, insbesondere bei teilautomatisierten Fahrten. Der Zustand des Fahrers ist besonders vor, während und nach Übergaben relevant (Diederichs et al., 2020, Erikson et al., 2017, Lu & de Winter, 2015). Um ein dem Automatisierungsgrad entsprechendes Verhalten zu bewerten, muss der aktuelle Zustand des Fahrers klassifiziert und mit den Anforderungen des aktuellen Automatisierungsgrads verglichen werden (Diederichs et al., 2022). Dies bedarf geeigneter KI-Applikationen, welche in der Lage sind, die aktuelle Tätigkeit des Fahrers zu klassifizieren und mit dem aktuellen Automatisierungslevel zu vergleichen.

In der Vergangenheit wurden die meisten Daten für die Entwicklung solcher KI-basierter Systeme in Fahrsimulatorstudien erhoben, die eine gut kontrollierte Studienumgebung darstellen. Im Vergleich dazu haben Fahrstudien im Realverkehr den Vorteil, dass sie das Fahrverhalten im realen Verkehr darstellen und dabei auch andere Verkehrsteilnehmer und Fußgänger berücksichtigen. Dies ermöglicht die Erhebung realistischerer Daten zum Verhalten der Nutzer. Darüber hinaus bietet ein mit Sensorik ausgestattetes Fahrzeug im Realverkehr den Vorteil, bessere Einblicke in das Verhalten der Nutzer in Echtzeit zu erlangen und stellt daher eine gute Ergänzung zu den Standardmethoden der Datenerhebung mit Fragebögen vor und nach der Fahrt dar (Carsten et al., 2013). Insbesondere die gleichzeitige Verwendung mehrerer Sensoren im Auto ermöglicht eine multimodale Datenerfassung von z. B. Videodaten, physiologischen Daten und fahrzeugspezifischen Daten (Li et al., 2013) und liefert somit verschiedene naturalistische Daten für das Training von KI-basierten Systemen.

Eine große Herausforderung für die Entwicklung und Bewertung autonomer Fahrfunktionen ist jedoch die geringe Verfügbarkeit hochautomatisierter Fahrzeuge für die Forschung (Habibovic et al. 2016). Eine bewährte Technik zur Konzeption und Bewertung automatisierter Systeme ist der Wizard-of-Oz-Ansatz, bei dem ein Mensch („Wizard“) die Aufgaben des KI-basierten Systems ausführt, wodurch der Benutzer den Eindruck erhält mit einem automatisierten System zu interagieren (Fraser & Gilbert, 1991). Diese Technik stammt aus dem Bereich des Voice-Interface-Designs, wird aber auch im Automobilsektor eingesetzt, um HMI-Funktionen sowie das Benutzerverhalten während teil- oder vollautomatisierter Fahrten und Übergänge zu simulieren und zu bewerten (Kiss et al., 2008). Die Umsetzung des WoZ-Ansatzes während des realen Fahrens ermöglicht somit die Datenerfassung und die Untersuchung eines natürlicheren Verhaltens des Nutzers während der Fahrt.

Um hochautomatisierte Fahrzeuge zu ermöglichen, ist es notwendig die Levelkonformität festzustellen. Dafür ist es notwendig, den Fahrzeuginnenraum zu erfassen. Fahrfremde Tätigkeiten können dabei mit verschiedenen bildbasierten Verfahren erkannt werden.

Skelettbasierte Techniken zur Erkennung menschlicher Handlungen gelten als robuste Alternative zu bildbasierten Ansätzen, die oft durch variierende Bildgebungsbedingungen und Datenschutzprobleme beeinträchtigt werden. Dabei verwenden sie Gelenkdaten, um menschliche Aktivitäten informativ und kompakt darzustellen. Die Verwendung von Deep-Learning-Modellen wie CNNs, RNNs, GCNs und Transformern hat diesen Ansatz deutlich vorangebracht: CNNs extrahieren räumliche Merkmale, RNNs behandeln zeitliche Sequenzen, und GCNs modellieren topologische sowie räumlich-zeitliche Graphenstrukturen in den Skelettdaten (Du et al., 2015; Shi et al., 2019; Martin et al., 2020; Zhou et al., 2023). Transformermodelle nutzen Selbstaufmerksamkeit, um Abhängigkeiten zwischen den Gelenken zu erfassen und so globales Repräsentationslernen zu ermöglichen (Cheng et al., 2021; Kim et al., 2022; Su et al., 2020).

Für die unüberwachte Erkennung menschlicher Aktionen existieren spezifische Ansätze wie LongT-GAN (Zheng et al., 2018) und PCRP (Xu et al., 2023), die Methoden wie GRU-basierte Sequenzrekonstruktion und lernbare Klassenprototypen zur Verbesserung der Modellleistung

nutzen. Weitere Modelle wie Predict & Cluster (Su et al., 2020) und AS-CAL (Rao et al., 2021) setzen kontrastives Lernen oder kombinieren GRU-Architekturen, um langfristige Abhängigkeiten und präzise Klassifikationen zu ermöglichen. In der selbstüberwachten Erkennung setzen Methoden wie Autoencoder, einschließlich SkeletonMAE (Wu et al., 2022), auf Maskierungstechniken zur Rekonstruktion von Skelettsequenzen und erfassen gleichzeitig subtile Bewegungsmuster. Der PSTL-Ansatz (Zhou et al., 2023) findet erstmals Anwendung in der skelettbasierten Fahreraktivitätserkennung und stützt sich auf die räumliche und zeitliche Dynamik der Gelenke.

Aktuelle Entwicklungen in Transformer-Architekturen zeigen vielversprechende Ergebnisse, die RNNs häufig übertreffen (Song et al., 2017; Qiu et al., 2022). Neuere Modelle, darunter unser Vanilla-Transformer-Modell mit Cross-Attention, nutzen das gesamte Wissen über die Sequenz, um sowohl die Aktionserkennung als auch die Datenproduktion zu verbessern und eine verlässliche Brücke zur selbstüberwachten Datengenerierung zu schlagen.

1.5 Zusammenarbeit mit anderen Stellen

1.5.1 Fraunhofer IAO

Das Fraunhofer IAO arbeitete während der Projektlaufzeit eng mit den Projektpartnern Continental und Fraunhofer IOSB zusammen, um den Einbau der Sensorik in das WoZ Fahrzeug zu koordinieren, sowie die Anforderungen an die zu erhebenden Daten für die Applikation LKV zu definieren. Im Rahmen der Applikation Motion Sickness kam das WoZ Fahrzeug als Versuchsträger in den Studien der Universität Stuttgart IAT zum Einsatz. Auch hierfür fanden Absprachen zur Planung der Studien, sowie den Voraussetzungen an das WoZ Fahrzeug statt. Des Weiteren wurde der Fahrsimulator für Studien innerhalb der Applikation KI-I zur Verfügung gestellt. Im Rahmen der Unteraufträge für den Fahrzeugumbau und die Konzipierung und Umsetzung der Systemarchitektur wurde mit der Paravan GmbH und AIOCAS Consulting & Solutions zusammengearbeitet.

1.5.2 Fraunhofer IOSB

Während der Projektlaufzeit arbeitete das Fraunhofer IOSB in enger Kooperation mit den Projektpartnern Continental, Ford und dem Fraunhofer IAO zusammen. Ziel dieser Zusammenarbeit war die Koordination des Einbaus der Sensorik in die Versuchsträger sowie die effiziente Verwaltung der erhobenen Daten. Ein weiterer wesentlicher Kooperationsschwerpunkt lag auf der Integration des KARLI-Avatars, der im Rahmen des Projekts von Studiokurbos entwickelt wurde, um ein interaktives und adaptives Interface zu kreieren. Zur Evaluierung des Interfaces wurde eine umfassende Studie durchgeführt, welche entscheidende Einblicke in die Nutzerakzeptanz und -interaktion lieferte. Darüber hinaus ermöglichte das von Vicomtech entwickelte Labeling-Tool eine effiziente Annotation und Weiterverarbeitung der gesammelten Sensordaten und stellte eine wichtige Grundlage für die Weiterentwicklung der KI-Modelle zur Erkennung und Interpretation menschlicher Aktionen und Intentionen dar.

2 Eingehende Darstellung der Ergebnisse

Die erzielten Projektergebnisse werden nachfolgend für das Fraunhofer IAO und IOSB separat nach Arbeitspaketstruktur dargestellt.

2.1 Verwendung der Zuwendung, erzielte Ergebnisse

Das **Fraunhofer IAO** hat sich in KARLI zum Ziel gesetzt, Fahrerzustandserkennung möglichst realitätsnah umzusetzen und zu erproben. Dafür wurde ein industrietaugliches Labeling- und Datenerhebungskonzept entwickelt, welches den Ablauf beginnend bei der Versuchsplanung, über die Datenerhebung bis hin zur Datenaufbereitung und dem Labeling der Daten für den Einsatz in KI-Algorithmen darstellt. Umgesetzt wurde dies durch den Einsatz des Wizard-of-Oz Fahrzeugs zur Datenerhebung aus verschiedenen Sensorquellen, sowie der Erweiterung der bestehenden Sensorik mit Innenraumkameras mehrerer Projektpartner.

Innerhalb der Projektlaufzeit fanden umfangreiche Anpassungen am Wizard-of-Oz Fahrzeug statt, welche auf Grund von Verzögerungen in der Hardwarebeschaffung mehr Zeit benötigten als ursprünglich angenommen. Dies umfasste zum einen die Konzeption und Umsetzung eines Synchronisierungskonzeptes, sowie die Integration von Hardware in das bestehende System. Des Weiteren beanspruchte die Vorbereitung der Realfahrtstudie im Rahmen der Applikation LKV mehr Zeit als ursprünglich geplant. Grund hierfür war das Training von Wizard und Sicherheitsfahrer, welches auf Grund des komplexen Studiensetups und dem Umgang mit dem WoZ Fahrzeug, in mehreren Schritten und mit verschiedenen Trainingseinheiten erfolgte. Auf Grund der Verzögerungen im Fahrzeugumbau und in der Beschaffung der Messtechnik am IAO sowie bei anderen Projektpartnern, was sich ebenfalls auf die geplanten Datenerhebungen auswirkte, verlängerte sich die Projektlaufzeit um 3 Monate (Tabelle 1).

Das **Fraunhofer IOSB** hat sich in KARLI zum Ziel gesetzt, ein kamerabasiertes Advanced Occupant Monitoring System mit seriennaher Hardware aufzubauen, die Hardware in mindestens 3 unterschiedliche Realfahrzeuge so zu integrieren, dass generische Forschungsdaten damit erhoben werden können, diese Daten im realen Fahrbetrieb bei Projektpartnern zu erheben und einen Realfahrdatensatz mit hochwertigen Labels zu erstellen. Auf dieser Grundlage entwickelt das Fraunhofer IOSB ein Advanced Occupant Monitoring System für die KI-Fusion mit anderen Sensoren der Partner.

2.1.1 Fraunhofer IAO

Meilenstein	Erzielte Ergebnisse	Ggf. angepasste Ziele oder Meilensteine	Referenz
MS1 (M9): LKV in messbare Indikatoren operationalisiert.	Levelkonforme und nicht levelkonforme FFT wurden anhand einer Black- und White-List definiert.	--	Kapitel 2.1.1.1
MS2 (M18): Fraunhofer WoZ-Fahrzeug zeigt Datenerhebung für LKV auf HzP.	Integrierte Sensorik und Konzept der LKV Datenerhebung wurden demonstriert.	Verzögerungen im Fahrzeugumbau, daher vollständige Integration der Sensorik und Middleware erst zu M21 fertiggestellt.	Kapitel 2.1.1.3 2.1.1.4

MS3 (M27): Datenerhebung abgeschlossen	Datenerhebung mit der Universität Stuttgart IAT innerhalb der Applikation Motion Sickness durchgeführt. Datenerhebung innerhalb der Applikation LKV im Realverkehr durchgeführt.	Folgestudie innerhalb der Applikation MS in M32 durchgeführt, um Ergebnisse zu validieren und Stichprobe zu erweitern.	Kapitel 2.1.1.2.2
MS4 (M36): Lessons Learned aus WoZ Datenerhebung	Lessons Learned wurden tabellarisch definiert und können für nachfolgende Studien angewandt werden.	Auf Grund der Projektverlängerung um 3 Monate in M39 fertiggestellt.	Kapitel 2.1.1.2.2

Tabelle 1: Meilensteine und erzielte Ergebnisse während der Projektlaufzeit (Fraunhofer IAO)

2.1.1.1 Arbeitspaket AP 100 (UAP 110, 120)

Die Arbeiten des AP100 umfassten eine umfassende Recherche zu den rechtlichen und empirischen Aspekten der unterschiedlichen Automatisierungslevel und diente vor allem dazu einen Überblick über die unterschiedlichen Definitionen der Automatisierung zu erlangen. Die Recherche befasste sich hierbei mit den Definitionen der SAE-Level, der Euro NCAP, ADAC und BAST. Die Rechercheergebnisse zeigten, dass der Fokus der bisherigen Definitionen der Automatisierungslevel stark auf die technischen Komponenten des automatisierten Fahrens gerichtet ist und die Aufgaben des Nutzers bzw. die durch die Automatisierung entstehende Möglichkeit zur Ausführung fahrfremder Tätigkeiten (FFT) wenig betrachtet werden. Da in KARLI der Nutzer im Vordergrund steht, wurden auf Grund dieser Ergebnisse in Zusammenarbeit mit den anderen Projektpartnern der Applikation LKV, unter der Leitung von Continental, KARLI Nutzerrollen definiert (Abbildung 1). Diese Nutzerrollen beschreiben die Anforderungen an den Fahrer in verschiedenen Automatisierungsstufen und stellen dar, welcher Nutzerzustand und welche FFT in der jeweiligen Automatisierungsstufe zugelassen sind.






Nutzerrolle	Kürzel	Beschreibung	
Aktiver Fahrer	K-R1	Verantwortlicher, aktiver Fahrzeugführer, möglicherweise unterstützt durch Assistenzsysteme	
Überwachender Fahrer	K-R2	Verantwortlicher Fahrzeugführer, unterstützt durch Assistenzsysteme/Automatisierungssysteme, die die Fahraufgabe unter dauerhafter Überwachung des Fahrers übernehmen	
Bereitschaftsfahrer	K-R3	Verantwortlicher Fahrzeugführer, unterstützt durch Automatisierungssysteme, die die Fahraufgabe vollständig übernehmen, jedoch jederzeit die Übernahme durch den Nutzer anfordern können und diesen als Rückfallebene nutzen	
Zeitweiser Passagier	K-R4	Auf definierten Streckenabschnitten beteiligt sich der Nutzer in keiner Weise an der Fahrzeugführung, die Fahraufgabe wird vollständig dem Automatisierungssystem übergeben. Dieses ist in der Lage, alle Situationen, ohne das Zutun des Nutzers zu bewältigen. Außerhalb dieser Streckenabschnitte wird ein Teil der Fahraufgabe an den Nutzer übertragen	
Dauerhafter Passagier	K-R5	Der Nutzer beteiligt sich während der gesamten Fahrt in keiner Weise an der Fahrzeugführung, die Fahraufgabe wird vollständig von dem Automatisierungssystem übernommen.	

Abbildung 1: KARLI-Nutzerrollen, in Anlehnung an die SAE-Level

Basierend auf der Recherche sowie der definierten KARLI Nutzerrollen wurde eine Black-White List für Aktivitäten und Fahrzustände definiert, welche sowohl levelkonforme als auch nicht levelkonforme FFT gegenüberstellt. Die hierfür herangezogenen FFT und Fahrzustände wurden in Absprache mit den anderen Projektpartnern der Applikation LKV vorab definiert, um sicherzustellen, dass diese auch in den folgenden Datenerhebungen und Studien untersucht werden und für die KI-Entwicklung relevant sind. So sind FFT wie „Trinken + 1 Hand am Lenkrad“ und „Telefonieren über die Freisprechanlage“ in allen KARLI Nutzerrollen erlaubt, da auch in den niedrigen Levels die Fahrzeugführung bzw. eine Übernahme der Fahrzeugführung gewährleistet werden kann und der Blick weiterhin auf die Straße gerichtet ist. Im Gegensatz dazu sind visuelle Ablenkungen von mehr als 2 Sekunden erst ab K-R3 erlaubt, wenn die reine Fahraufgabe nicht mehr vollständig beim Nutzer liegt. Dies ermöglicht des Weiteren FFT wie „Konsumieren von Inhalten auf dem Smartphone“ oder „Augen schließen“ (Abbildung 2).

Komplementär dazu beschreibt die Black-List, welche FFT und Fahrzustände in den jeweiligen KARLI Nutzerrollen nicht erlaubt sind. Black- und White-List dienen im weiteren Projektverlauf

dazu, die während der Realfahrtstudien erhobenen Daten entsprechend zu labeln und somit einen Trainingsdatensatz für die Klassifizierung von levelkonformen und nicht levelkonformen FFT zu definieren.

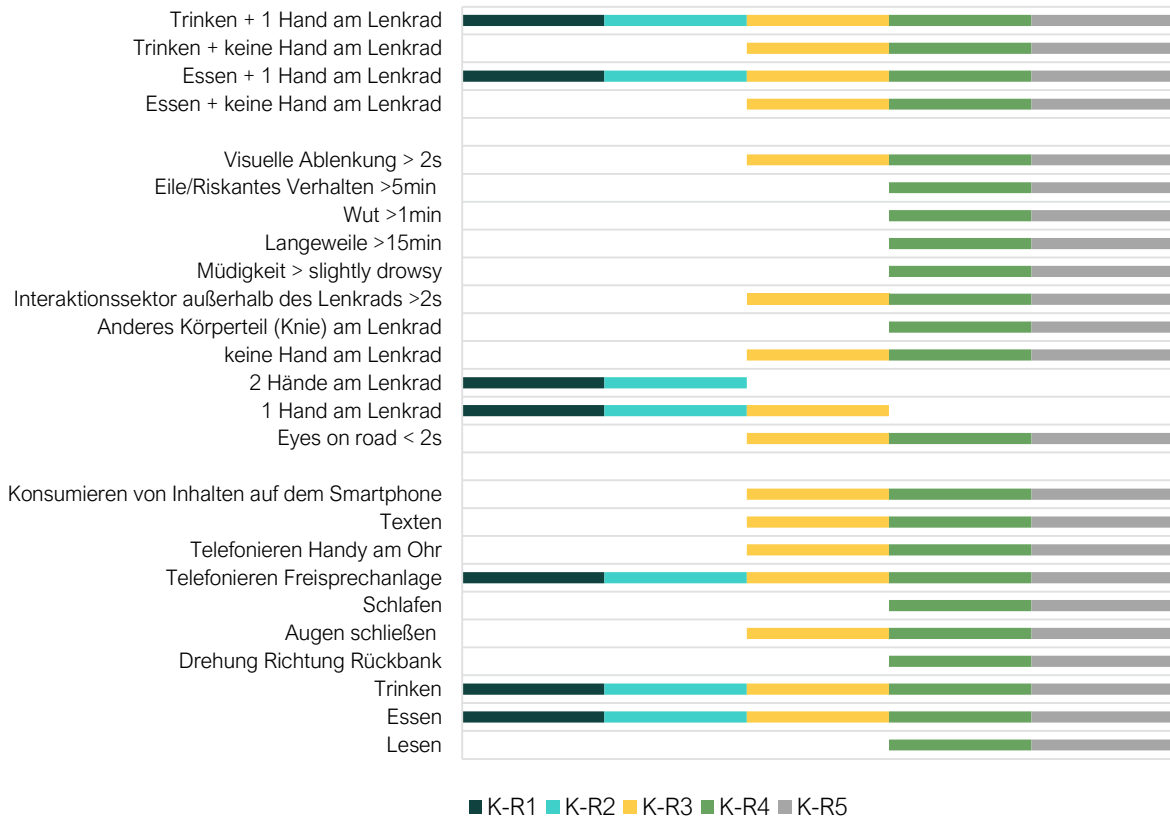


Abbildung 2: White List - Darstellung von levelkonformen FFT aufgeteilt in die 5 KARLI Nutzerrollen (K-R1 - K-R5)

2.1.1.2 Arbeitspaket AP 200 (UAP 210, 220, 230)

2.1.1.2.1 Versuchsträger des Fraunhofer IAO

Nachfolgend wird das Wizard-of-Oz (WoZ) Fahrzeug als Versuchsträger des Fraunhofer IAO kurz erläutert, was zum besseren Verständnis der in Kapitel 2.1.1.2.2 beschriebenen Datenerhebung im Realverkehr dient. Nähere Details zum technischen Aufbau und Umbau des Fahrzeugs können Kapitel 2.1.1.3 entnommen werden.

Das WoZ Fahrzeug dient dazu, verschiedene Automatisierungsstufen sowie Übergaben von automatisierter zu manueller Fahrt (und vice versa) zu simulieren. Dafür ist das Fahrzeug mit drei Sitzplätzen (vorne rechts, vorne links, hinten) ausgestattet. Jeder Platz ist mit einem eigenen Lenkrad und Pedalerie ausgestattet, welche es ermöglichen, das Fahrzeug von allen 3 Plätzen zu steuern. Daraus ergibt sich die folgende Rollenverteilung für die Studiendurchführung: Während der Simulation der automatisierten Fahrt steuert der Wizard auf dem hinteren Platz das Fahrzeug. Der Nutzer sitzt, wie gewohnt, vorne links, ein Sicherheitsfahrer auf dem vorderen rechten Platz überwacht das Fahrgeschehen dauerhaft und greift in Notfallsituationen ein, indem er die Kontrolle

über das Fahrzeug übernimmt. Um zu vermeiden, dass der Nutzer des Fahrzeugs sieht, dass dieses von einem menschlichen Wizard gesteuert wird, sind vorderer und hinterer Bereich durch eine verspiegelte Scheibe voneinander getrennt (Abbildung 4). Je nach Anforderung an die Datenerhebung kann das Fahrzeug mit verschiedener Sensorik ausgestattet werden, welche in den bestehenden Synchronisierungsprozess integriert werden kann.

Um das Fahrzeug sowohl als Wizard oder auch als Sicherheitsfahrer steuern zu können, bedarf es einem Sicherheitstraining, welches im Rahmen des AP200 konzipiert wurde. Dieses Training ist notwendig, um Mitarbeitende darin zu schulen, das Fahrzeug von vorne rechts als auch von hinten zu fahren, sowie Kenntnisse im technischen Aufbau des Fahrzeugs und der integrierten Messtechnik zu vermitteln. Jeder Mitarbeitende muss dabei das mehrstufige Fahrtraining, sowie die Sicherheitsunterweisung absolviert haben, um die Rolle des Sicherheitsfahrers und/oder Wizards zu übernehmen. Das Training setzt sich dabei aus den folgenden Stufen zusammen:

- 1) Theorie
 - Theoretische Grundlagen des Fahrzeugs, dessen Ausstattung und technische Funktionen werden zusammen mit einem erfahrenen Mitarbeiter durchgesprochen
 - Integrierte Messtechnik, sowie dazugehörige Software und Datensynchronisierungsprozess werden zusammen mit einem erfahrenen Mitarbeiter durchgesprochen
- 2) Fahrzeug kennenlernen (im Stand)
 - Aufbau und Funktionen des WoZ werden durch einen bereits erfahrenen Mitarbeiter direkt am Fahrzeug gezeigt
 - Stromversorgung und Messtechnik werden gestartet und demonstriert
- 3) Fahrübungen mit Nachweis der Durchführung (Sicherheitsfahrer)
 - Fahren auf Privatgelände ohne andere Verkehrsteilnehmer mit Fokus auf Übersteuerung der anderen Fahrer (vorne links und hinten) und Übergabeszenarien
 - Fahrübung bei unübersichtlichen Gegebenheiten mit anderen Verkehrsteilnehmern, Fußgängern sowie Parken und Rangieren
 - Fahren im Realverkehr
- 4) Fahrübung mit Nachweis der Durchführung (Wizard)
 - Fahren auf Privatgelände ohne andere Verkehrsteilnehmer mit Fokus auf Starten und Kalibrierung des Systems, routiniertes Fahren von hinten (Sichtverhältnisse, Pedalerie), Übergabeszenarien
 - Fahrübung bei unübersichtlichen Gegebenheiten mit anderen Verkehrsteilnehmern, Fußgängern
 - Fahren im Realverkehr (wurde auf Grund der schwierigen Sichtverhältnisse bisher nicht durchgeführt)

Die Fahrübungen für den Wizard umfassen dabei auf Grund der eingeschränkten Sichtverhältnisse mehr Iterationen als die Fahrübungen für den Sicherheitsfahrer.

2.1.1.2.2 Datenerhebung im Realverkehr (LKV)

Ziel des AP200 war es, möglichst reale und variantenreiche Trainingsdatensätze zu erheben, welche für die Entwicklung der KI-Applikationen zur Nutzerzustandserkennung und Aktivitätenerkennung der Projektpartner Continental und Fraunhofer IOSB genutzt werden können. Hierfür wurde basierend auf den Ergebnissen und Definitionen aus AP100 zunächst eine

Realfahrtstudie mit dem Wizard-of-Oz Fahrzeug konzipiert (UAP210), welche es erlaubt, kontinuierlich Daten zum Nutzerzustand sowie levelkonformen und nicht levelkonformen FFT in verschiedenen Automatisierungsstufen zu erheben (UAP220). Im Rahmen von Regelmeetings wurden außerdem die Anforderungen an das Datenformat und das Labeling definiert (UAP230).

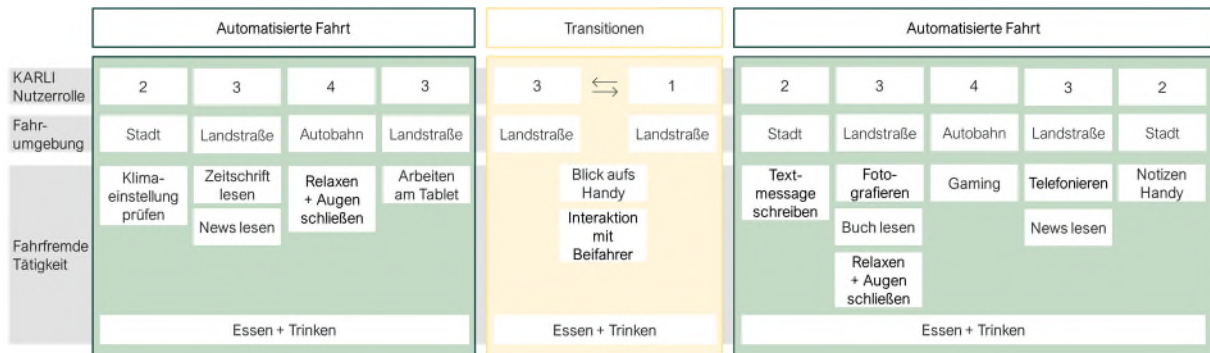


Abbildung 3: Studienablauf zur Erhebung von Trainingsdaten im Realverkehr unter der Berücksichtigung verschiedener Automatisierungsstufen und FFT.

Die in UAP210 konzipierte Realfahrtstudie mit dem WoZ Fahrzeug, beinhaltete zum einen Streckenabschnitten automatisierter Fahrt mit dem Fokus auf K-R3 und K-R4 auf Landstraßen, sowie Übergabeszenarien zwischen manuellem und automatisiertem Fahren. Ziel der Datenerhebung war es, sowohl levelkonforme als auch nicht levelkonforme FFT zu erfassen und dies in möglichst realer Umgebung. Die FFT umfassten hierbei z.B. das Lesen einer Zeitschrift, Telefonieren oder Arbeiten am Tablet (Abbildung 3).

Um möglichst variantenreiche Datensätze zu erheben, konnten die Teilnehmenden bis zu 3-mal an der Studie teilnehmen. Somit war es möglich verschiedene Tageszeiten, Wetterbedingungen und Verkehrsverhältnisse in den Datensätzen zu berücksichtigen. Vor Beginn der Datenerhebung wurde das Konzept, sowie die definierten FFT mit den Projektpartnern Continental und Fraunhofer IOSB abgestimmt.

Die zu fahrende Strecke, von Stuttgart nach Pforzheim und zurück, umfasste insgesamt 83km mit einer Fahrtdauer von 90-120 Minuten und beinhaltete zum größten Teil Landstraßenabschnitte. Die gesamte Strecke wurde vorab in verschiedene Streckenabschnitte eingeteilt, in denen unterschiedliche KARLI Nutzerrollen (K-R1 bis K-R4) simuliert wurden (Abbildung 5). Hierbei lag der Fokus im Realverkehr auf K-R3 und K-R4. Um die Fahrt für die Teilnehmenden möglichst realistisch darzustellen, beinhaltete sie außerdem zu Beginn und Ende kurze Abschnitte in KR-2 im Stadtverkehr. Aus Sicherheitsaspekten wurde die Simulation der automatisierten Fahrt in dieser Datenerhebung im Realverkehr den Teilnehmenden offen gelegt. Die zuvor absolvierten Trainingseinheiten von Wizard und Sicherheitsfahrer ergaben, dass es auf Grund der eingeschränkten Sicht mehr Routine in der Rolle des Wizard verlangt, um das Fahrzeug sicher im realen Straßenverkehr, sowie in unübersichtlichen Verkehrssituationen steuern zu können. Aus diesem Grund wurde das Fahrzeug im Realverkehr von vorne rechts gesteuert (Abbildung 4).



Abbildung 4: Studiensetup der LKV-Datenerhebung im Realverkehr mit dem WoZ Fahrzeug

Des Weiteren wurden Übergaben zwischen automatisierter und manueller Fahrt auf Privatgelände auf einem Verkehrsübungsplatz in Pforzheim durchgeführt. Hierfür wurde die Fahraufgabe von den Teilnehmenden an den Wizard im hinteren Teil des Fahrzeugs übergeben. Dies erlaubte eine realistischere Datenaufzeichnung während der Übergabesituationen. Die Datenerhebung der FFT während der Übergabesituationen, sowie der anschließenden manuellen Fahrt in KR-1 fand somit nicht im Realverkehr statt. Da das Ziel der Datenerhebung in der Generierung von Trainingsdatensätzen zur Klassifizierung von FFT lag und nicht in der Untersuchung der automatisierten Fahrt selbst, war dieses Studiendesign für die Ergebnisse angemessen und vorab mit den Projektpartnern abgesprochen.

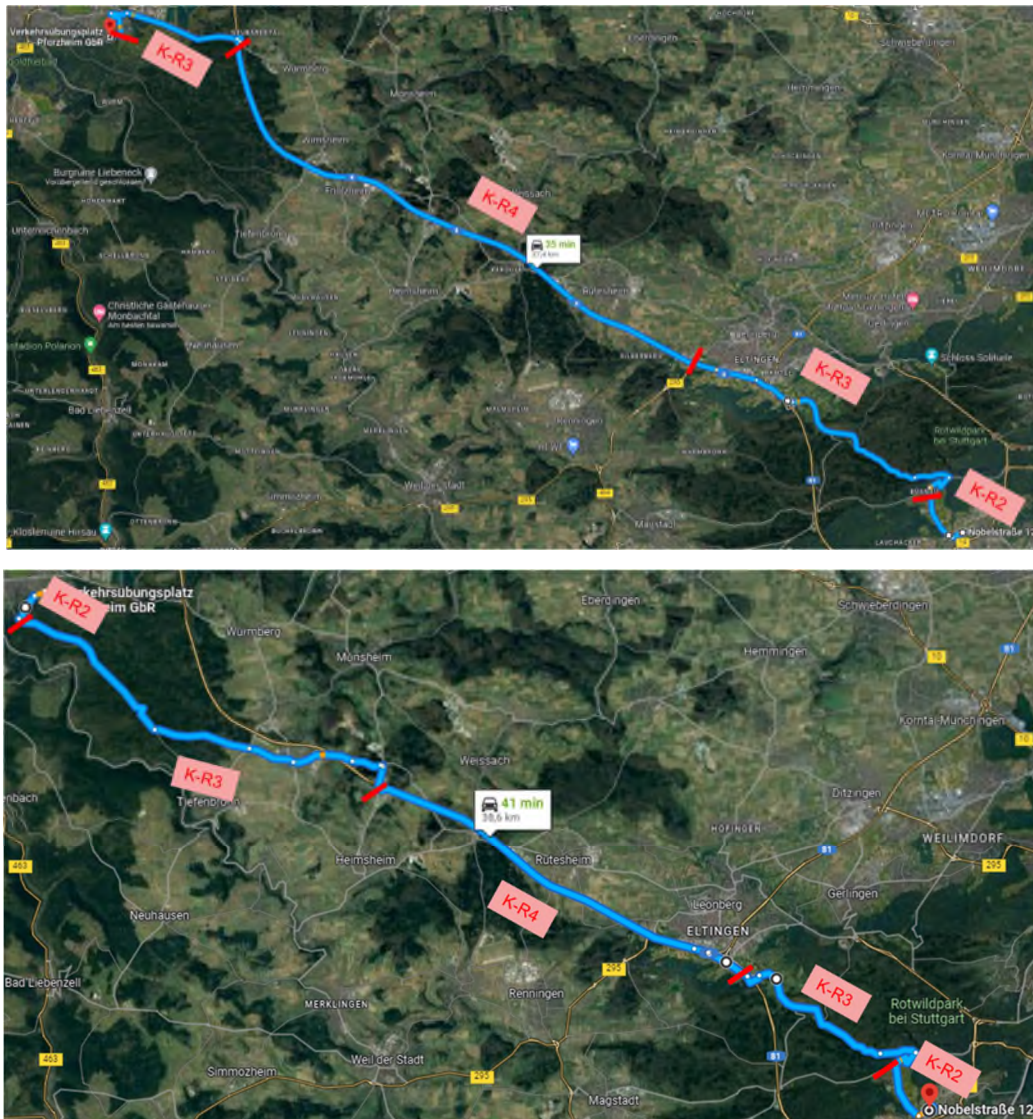


Abbildung 5: Hin- und Rückfahrt der Datenerhebung im Realverkehr (Stuttgart - Pforzheim), unterteilt in die verschiedenen Streckenabschnitte zur Simulation der KARLI Nutzerrollen.

Die aktuelle KARLI Nutzerrolle wurde den Teilnehmenden, zusammen mit den zu absolvierenden „Missionen“, auf einem Tablet in der Mittelkonsole des Fahrzeugs angezeigt (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Die Missionen umfassten die Tätigkeitsbereiche *Klimaeinstellung anpassen, Zeitschrift und Buch lesen, Handy- und Tablet Nutzung, Relaxen, Interaktion mit dem Beifahrer sowie Gaming*. Eine Mission bestand dabei aus einer oder mehreren levelkonformen und nicht levelkonformen FFT. Die FFT „Essen und Trinken“ war die gesamte Fahrt über möglich. Um möglichst natürliche Daten zu erhalten, wurde den Teilnehmenden die Reihenfolge und Dauer der Ausführung dieser Nebentätigkeiten freigestellt. Zu fest definierten Streckenabschnitten wurde die nächste Mission durch den Versuchsleiter eingeblendet.

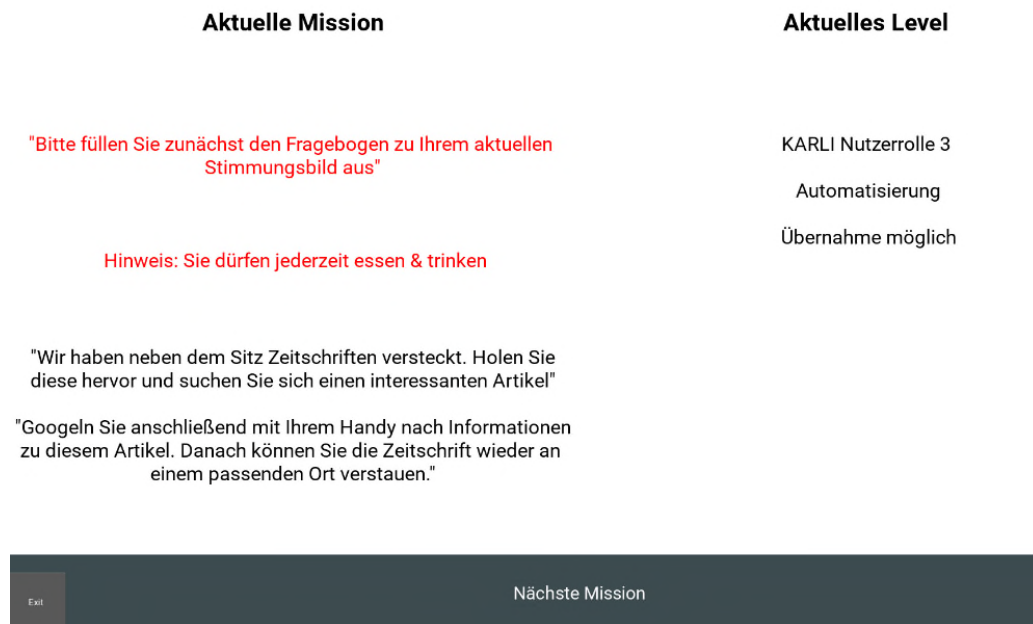


Abbildung 6: Exemplarische Darstellung der "Mission" während einer automatisierten Fahrt in KR-3.

Des Weiteren fand im Rahmen der LKV-Datenerhebung eine Studie der Universität Stuttgart IAT und zur Untersuchung von proaktiven Ansprachen bei Nebentätigkeiten statt, welche als Input für die Applikation KI-Interaktion (AP 400) diente. Hierbei sollte die Akzeptanz proaktiver Ansprachen bei verschiedenen FFT während der automatisierten Fahrt untersucht werden. Die Ansprachen wurden unter Nutzung eines Wizard-of-Oz Softwaretools ausgelöst. Die Ansprachen konnten dabei zu Beginn einer FFT, während der Ausführung der FFT oder nach Beenden der FFT auftreten. Die Aufgabe der Teilnehmenden war es, der Ansprache entweder zuzustimmen oder diese abzulehnen, je nachdem ob sie den Zeitpunkt als passend empfanden oder nicht. Der genaue Ablauf der Datenerhebung sowie die Ergebnisauswertung ist im partnerspezifischen Bericht der Universität Stuttgart IAT beschrieben.

Für die Datenerhebung wurde das Wizard-of-Oz Fahrzeug mit unten beschriebener Messtechnik ausgestattet, sowie Anpassungen im Synchronisierungsprozess und der Datenspeicherung vorgenommen. Dabei wurde sich an Herausforderungen und Lessons Learned der vorangegangenen Motion Sickness Studie in Zusammenarbeit mit der Universität Stuttgart IAT orientiert. Des Weiteren wurde die integrierte Messtechnik der Projektpartner Continental und Fraunhofer IOSB nochmals ausführlich getestet und Updates durchgeführt. Beide Kamerasysteme wurden in das aktualisierte Synchronisierungskonzept integriert und während mehrerer Probefahrten getestet. Des Weiteren wurde die bereits integrierte Sensorik des Fraunhofer IAO sowie die Abspeicherung und Synchronisierung verschiedener Label getestet. Während der Realfahrten wurden demnach die folgenden Daten aufgezeichnet (Abbildung 8):

- Driver Monitoring und Aktivitätenerkennung (Fraunhofer IOSB)
- Head-Eye-Tracking (Continental)
- Eye-Tracking zum Abgleich und als Backup für fehlende Daten (Fraunhofer IAO)
- Subjektive Daten als Label für das Wohlbefinden das aktuelle Stimmungsbild während der Fahrt (Fraunhofer IAO)

- KARLI Nutzerrollen und durchgeführte FFT pro Mission als Label für Videodaten (Fraunhofer IAO)
- Sprachdaten zur Bewertung von Ansprachezeitpunkten während verschiedener KARLI Nutzerrollen (Universität Stuttgart IAT)

Die aus den Realfahrten resultierende Datenmenge beinhaltet insgesamt ca. 29 Stunden an Videoaufzeichnung, sowie insgesamt ca. 1600 km zurückgelegte Fahrtstrecke. An der Datenerhebung nahmen 8 Personen im Alter von 20-65 Jahren teil, 5 davon mit einer Messwiederholung von 3-Mal. Die daraus resultierenden 19 durchgeführten Fahrten deckten sowohl verschiedene Licht- und Wetterverhältnisse als auch Verkehrssituationen (z.B. Stau) ab. Des Weiteren konnten durch die Messwiederholungen Gewohnheiten bei wiederholter Ausführung FFT im Verhalten der Teilnehmenden festgestellt werden. So wurden FFT in der dritten Messwiederholung freier und spontaner ausgeführt als noch während der ersten Fahrt. Dies ist vor allem für das Training der KI-Applikationen von großem Nutzen und zeigt, dass im Kontext der Datenerhebung im Realverkehr ein iteratives Studiendesign über einen längeren Zeitraum hinweg nicht nur von Vorteil für die Datenmenge, sondern auch für die Datenqualität ist.

Die anschließende Datenaufbereitung beinhaltete unter anderem den Export im gewünschten Datenformat und die Bereinigung der Daten, sowie einen Abgleich der gespeicherten Zeitstempel zur Synchronisierung verschiedener Datenquellen. Des Weiteren wurde die Datenqualität anhand der Kriterien „fehlerhafter bzw. unvollständiger Datensatz“ und „fehlender Zeitstempel“ bestimmt. Der Datenaustausch mit den entsprechenden Projektpartnern fand über die von Invensity bereitgestellte aws Cloud statt.

Basierend auf der Datenerhebung im Rahmen der Applikation LKV sowie der Studiendurchführung im Rahmen der Applikation MS, bestand ein weiterer Arbeitsschritt in der Definition der Lessons Learned zur Studiendurchführung mit dem Wizard-of-Oz Fahrzeug. Da das Wizard-of-Oz Fahrzeug in beide Applikationen eingebunden war, wurden diese gemeinsam definiert und tabellarisch festgehalten. Die Lessons Learned gliedern sich dabei in folgende Kategorien:

- Planung von Studien (Realfahrt & Teststrecke)
- Vorbereitung des WoZ Fahrzeugs
- Durchführung von Studien (Datenerhebung, Datensynchronisierung)
- Umgang mit Probanden (Instruktion zum Fahrzeug, Aufklärung der Illusion)
- Datenspeicherung und Export
- Datenaustausch

Die definierten Lessons Learned können für zukünftige Projekte mit dem WoZ Fahrzeug herangezogen werden, um somit die Studienplanung und Vorbereitung effektiv zu gestalten und Anforderungen an die Datenerhebung sowie an das Fahrzeug frühzeitig zu definieren.

2.1.1.3 Arbeitspaket AP 600 (UAP 610, 630, 640)

Ziel des AP600 war die Entwicklung einer System- und Software Architektur für das Wizard-of-Oz Fahrzeug, sowie der technische Umbau des Fahrzeugs zur Simulation verschiedener Automatisierungsstufen und zur Durchführung von Transitionen während der Fahrt (UAP610).

Für die anschließende Datenerhebung in LKV, sowie dem Einsatz des Fahrzeugs in der Applikation MS, musste die entsprechende Messtechnik beschafft und integriert werden (UAP630), sowie eine Kommunikation zwischen den verschiedenen Hardwarekomponenten über eine geeignete Middleware sichergestellt werden (UAP640).



Abbildung 7: Wizard-of-Oz Fahrzeug des Fraunhofer IAO, als Versuchsträger der Applikation LKV und MS

Um die im Rahmen des Projekts definierten KARLI Nutzerrollen während Realfahrtstudien oder Studien auf der Teststrecke simulieren zu können, sowie Übergaben von manuellem zu automatisiertem Fahren (oder vice versa) durchzuführen, waren technische Anpassungen des Fahrzeugaufbaus notwendig. Dieser wurde im Rahmen eines Unterauftrags mit der Firma Paravan durchgeführt. Hierbei wurden die Lenkräder der beiden vorderen Sitze durch eine mechanische Kette und das Lenkrad auf dem hinteren Fahrplatz mittels Steer-by-Wire Technologie mit den vorderen Lenkrädern gekoppelt. Diese Technologie ermöglicht eine elektronische Kraftübertragung des hinteren Lenkrads auf das Originalenkrad (vorne rechts) mit Hilfe einer Magnetkupplung (Paravan, 2022).

Die Übergabe der Fahrzeugkontrolle erfolgt über einen Release Button, welcher vorne mittig im Fahrzeug verbaut ist. Durch ein Drücken des Buttons erfolgt die Deaktivierung der Magnetkupplung, die Fahrzeugkontrolle liegt beim Sicherheitsfahrer und Studienteilnehmer. Durch ein Ziehen des Buttons erfolgt die Aktivierung der Magnetkupplung und der Wizard kann das Fahrzeug vom hinteren Fahrzeugplatz steuern. Aktivierung und Deaktivierung erfolgt durch den Sicherheitsfahrer und erlaubt eine einfache Übergabe der Fahraufgabe während der Fahrt. Der angepasste Fahrzeugaufbau wurde vom TÜV abgenommen und die Straßenzulassung erteilt, sodass eine Fahrzeugführung durch den Sicherheitsfahrer und den Wizard im Straßenverkehr möglich sind.

Um die integrierte Sensorik sowie die Car PCs zur Datenaufzeichnung und Speicherung mit Energie zu versorgen, ist das Fahrzeug mit zwei Lithium-Ionen-Batterien ausgestattet. Über einen Inverter wird die Gleichspannung in Wechselspannung umgewandelt. Während der Projektlaufzeit wurde das Fahrzeug außerdem mit einem Klimagerät im hinteren Bereich des Fahrzeugs aufgerüstet, um eine Überhitzung der Rechner im Sommer zu vermeiden, sowie den Fahrerplatz angenehmer zu gestalten und somit die Konzentration des Wizards zu gewährleisten.

Im Fahrzeug dient ein internes Netzwerk für die Kommunikation zwischen der Messtechnik und den verschiedenen Rechnern. Dieses Netzwerk wird über einen Router aufgebaut und ermöglicht die Kommunikation über TCP und UDP.

Um die Anforderungen an die Datenerhebung in LKV, sowie an die Studiendurchführung der Applikation MS nachkommen zu können, wurde das Fahrzeug im Projektverlauf mit verschiedener Sensorik ausgestattet. Dabei handelte es sich zum einen um Sensorik des Fraunhofer IAO, als auch um externe Sensorik und die dafür notwendige Hardware der Projektpartner Continental und

Fraunhofer IOSB. Abbildung 8 gibt einen Überblick über die Rollenverteilung (Sicherheitsfahrer, Nutzer, Wizard) und die integrierte Hardware im WoZ Fahrzeug, unterteilt in

- Messtechnik zur Datenaufzeichnung im Innenraum (Innenraum- und Fahrerkeras, Eye-Tracking, Mikrophon und Wearables)
- Messtechnik zur Aufzeichnung fahrzeugspezifischer Daten (CAN-Bus, Arduino zur Geschwindigkeitsmessung)
- Hardware zur Datenspeicherung (Car PCs, Laptop, Tablet)
- HMI zur Kommunikation mit den Studienteilnehmenden (Tablet)

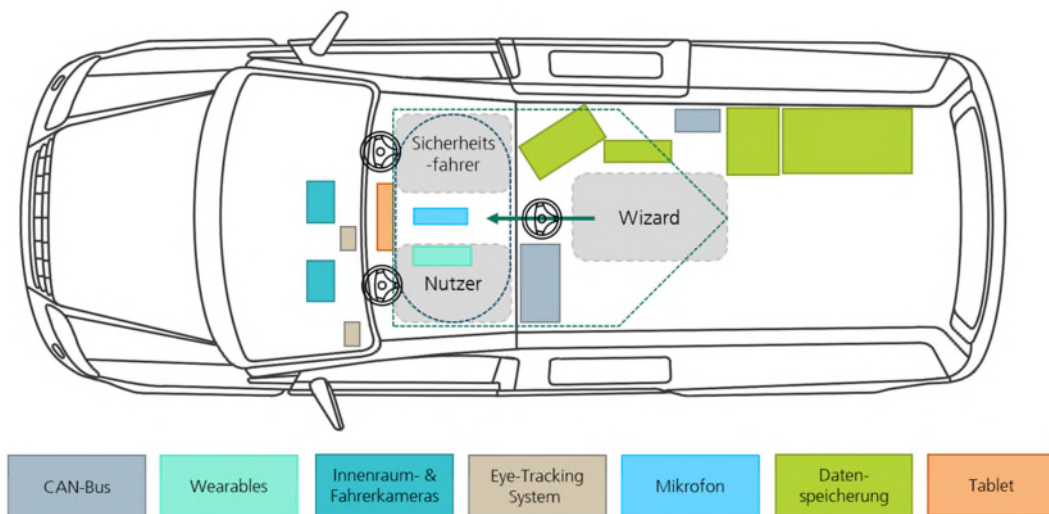


Abbildung 8: Hardwareübersicht des Wizard-of-Oz Fahrzeug

Im nachfolgenden wird die Hardware sowie die damit erhobenen Daten und der Verwendungszweck näher erläutert (Tabelle 2).

Messtechnik	Zweck der Datenerhebung	Projektpartner
Head- Eye Tracking	Erhebung von Trainingsdaten zum Nutzerverhalten während der Ausführung verschiedener FFT zur Bestimmung von levelkonformen oder nicht levelkonformen Verhalten.	Continental
Innenraumkamera	Erhebung von Trainingsdaten zur Erkennung von FFT während verschiedener Automatisierungsstufen.	Fraunhofer IOSB
Eye-Tracking SmartEye Pro dx	Erfassung von Blickrichtung und Pupillendurchmesser als ergänzende Datenquelle und Input zum Head-Eye-Tracking System, sowie zur Detektion von MS.	Fraunhofer IAO
Empatica Embrace Plus	Erhebung physiologischer Daten (Herzrate, Hautleitfähigkeit) als objektives Maß zur Detektion von MS.	Fraunhofer IAO, Universität Stuttgart IAT
CAN-Bus	Erfassung von fahrzeugspezifischen Daten (Lenkwinkel und Bremsverhalten) zur	Fraunhofer IAO

	Untersuchung des Einflusses von MS auf die Fahrtüchtigkeit.	
Mikrofon	Erhebung von mündlicher Bewertung des Wohlbefindens und der Fahrfähigkeit im Rahmen der MS Studie, sowie Beurteilung von Ansprachezeitpunkten und Interaktion mit dem Sprachassistenten im Rahmen der LKV-Datenerhebung.	Fraunhofer IAO, Universität Stuttgart IAT
Tablets	Darstellung von Missionen während der LKV-Datenerhebung sowie eines kognitiven Tests zur Leistungsfähigkeit während der MS Studie. Triggern der Wizard-of-Oz Toolbox zur Sprachinteraktion.	Fraunhofer IAO

Tabelle 2: Messtechnik und Verwendungszweck für LKV-Datenerhebung und MS Studiendurchführung

Um eine Kommunikation zwischen den verschiedenen Hardwarekomponenten, sowie eine zeitliche Synchronisierung der einzelnen Datenströme zu gewährleisten, bedarf es einer geeigneten Middleware, welche als Vermittler zwischen den verschiedenen Datenquellen dient und deren Kommunikation und Datenmanagement ermöglicht. Nach einer eingehenden Recherche wurde die Software Blickshift Recorder ausgewählt, welche verschiedene Datenströme aus verschiedenen Quellen erfassen kann und deren Kommunikation über TCP/IP Anbindungen ermöglicht (Blickshift, 2021). Des Weiteren bietet diese Middleware eine grafische Oberfläche zur Einbindung unterschiedlicher Datenquellen, sowie dem Start und Stopp der Datenaufzeichnung. Somit konnte die Datenaufzeichnung der verschiedenen Messtechnikkomponenten zeitgleich gestartet und beendet werden. Zur zeitlichen Synchronisierung der Datenströme wurde ein Arduino aufgesetzt und in Blickshift Recorder integriert, welcher in regelmäßigen Abständen eine fortlaufende Zahl sendet. Diese diente als „Zeitstempel“ und konnte mit Hilfe des Blickshift Recorders an die verschiedenen Messtechnikkomponenten gesendet und in die jeweiligen Datenfiles gespeichert werden. Des Weiteren ist diese Middleware kompatibel mit dem bereits genutzten Tool Blickshift Analytics zur Datenauswertung und Visualisierung.

2.1.1.4 Arbeitspaket AP 700 (UAP 730)

Der Fokus des Fraunhofer IAO in AP700 lag in der Demonstration des Datenerhebungsprozesses im WoZ Fahrzeug. Hierfür wurde die Pipeline, sowie die integrierte Messtechnik der LKV-Datenerhebung über verschiedene Demonstrationen auf der Abschlusspräsentation im September 2024 erlebbar gemacht (Abbildung 9). Die Gestaltung der Demonstrationen wurde zusammen mit den Projektpartnern Continental und Fraunhofer IOSB abgestimmt und erarbeitet. Ziel war es, zum einen den Ablauf der Datenerhebung über ein Video der Strecke, zum anderen die erhobenen Daten und deren Verwendungszweck visuell darzustellen.



Abbildung 9: Demonstration des WoZ-Fahrzeugs auf der AbP

2.1.1.5 Arbeitspaket AP 800 (UAP 810, 820)

Das Fraunhofer IAO nahm regelmäßig an den Applikationsmeeting LKV sowie an den Steuerkreismetings teil. Als Verantwortliche des AP200 fanden auch hier regelmäßige Meetings zur Absprache der in UAP230 definierten Inhalte zu Labeling und Datenaufbereitung statt. Im Rahmen der Ergebnisverbreitung wurden Inhalte der LKV-Datenerhebung, mit Unterstützung des Projektbüros, über LinkedIn geteilt. Des Weiteren wurde das KARLI Projekt und der Einsatz des WoZ Fahrzeugs im Rahmen einer Veranstaltung auf dem Fraunhofer IZS Campus in Stuttgart präsentiert (Abbildung 10), sowie mehrere Publikationen, in Zusammenarbeit mit der Universität Stuttgart IAT oder dem Fraunhofer IOSB, auf wissenschaftlichen Konferenzen eingereicht und präsentiert (siehe Kapitel 2.6). So wurden KARLI Projektergebnisse aus Datenerhebungen und Studien auf den folgenden wissenschaftlichen Konferenzen und Tagungen präsentiert (Abbildungen 11-13):

- Stuttgarter Symposium für Produktentwicklung, Stuttgart, Mai 2023 (SSP2023)
- International Conference on Applied Human Factors and Ergonomics, Juli 2023, San Francisco (AHFE2023)
- GfA-Frühjahrskongress – "Arbeitswissenschaft in-the-loop", März 2024, Stuttgart, (GfA2024)
- ACM Conference „Conversational User Interfaces“, Juli 2024, Luxemburg (CUI2024)
- ACM Conference „Mensch und Computer“, September 2024, Karlsruhe (MuC2024)



Abbildung 10: Demonstration des Projekts KARLI und den Einsatz des WoZ Fahrzeugs auf dem Fraunhofer IZS Campus in Stuttgart



Abbildung 11: Präsentation von KARLI Studienergebnissen auf der AHFE2023 in der Session „Vehicle Automation“



Abbildung 12: Präsentation von KARLI Studienergebnissen auf der CUI2024



Abbildung 13: Präsentation von KARLI Studienergebnissen auf der MuC 2024

2.1.1.6 AP-Koordination, Bereitstellung von Verfahren

Im Rahmen der Applikation LKV fanden alle 4 Wochen Regelmeetings statt, in denen der aktuelle Arbeitsstand der jeweiligen Projektpartner berichtet wurde, sowie die nächsten Schritte definiert wurden. Des Weiteren fanden im Rahmen des UAP230 alle 6 Wochen Regelmeetings mit den Projektpartnern statt, die Datenerhebungen und Labeling durchführten. Dies diente dazu, Anforderungen an die Datenerhebungen sowie an das Datenformat zu definieren und ein

einheitliches Labelingkonzept festzulegen. Neben den Regelmeetings fanden alle 6 Monate Konsortialmeetings online oder in Präsenz bei einem der Projektpartner statt.

Das Fraunhofer IAO brachte in KARLI zwei Laborumgebungen ein, einen Fahrsimulator mit der Fahrsimulationssoftware SILAB sowie das WoZ Fahrzeug zur Studiendurchführung im Realverkehr. In beiden Laborumgebungen fanden Nutzerstudien und Datenerhebungen für alle drei Applikationen statt, wobei das WoZ Fahrzeug für die Studiendurchführung und Datenerhebung innerhalb der Applikation MS und LKV eingesetzt wurde. Im Fahrsimulator fanden umfangreiche Nutzerstudien zum Thema proaktive Ansprachezeitpunkte der Universität Stuttgart IAT für die Applikation KI-Interaktion statt.

2.1.2 Fraunhofer IOSB

Meilenstein	Erzielte Ergebnisse	Ggf. angepasste Ziele oder Referenz Meilensteine	
Datenaufzeichnungssystem für das WoZ Fahrzeug des Fraunhofer IAO ist umgesetzt. (MS1)	Das Datenaufzeichnungssystem wurde umgesetzt.	--	Kapitel 2.1.2.3 und 2.1.2.4
Labelingkonzept für die videobasierte Innenraumanalyse ist umgesetzt und erste Daten sind gelabelt. (MS2)	Labelingkonzept wurde umgesetzt und Daten gelabelt.	--	Kapitel 2.1.2.2
Erste Verfahren sind entwickelt und demonstriert. (MS2)	Verfahren wurden entwickelt und demonstriert	--	Kapitel 2.1.2.7
Verfahren sind durch Semi-supervised/Unsupervised Learning verbessert. (MS3)	Verfahren wurden durch Semi-supervised und unsupervised Learning verbessert.	--	Kapitel 2.1.2.4
Finalen Verfahren sind in die Versuchsträger integriert und können demonstriert werden. (MS4)	Finalen Verfahren wurden in die Versuchsträger integriert und demonstriert.	Nach Projektverlängerung um 3 Monate im Monat 39 erreicht.	Kapitel 2.1.2.7

Tabelle 3: Meilensteintabelle Fraunhofer IOSB

2.1.2.1 Arbeitspaket AP 100 (UAP 120)

Im Rahmen mehrerer Abstimmungsmeetings wurden detaillierte Szenarien für die Datenerhebung auf Basis der drei relevanten Versuchsträger entwickelt und konsensual validiert, um eine einheitliche Grundlage für die anschließende Datenerhebung zu etablieren. Gleichzeitig wurden die Entwicklung und Fertigstellung von Personas und User Journeys in Zusammenarbeit mit dem Projektpartner Studiokurbos abgeschlossen. Diese Use Cases dienten als Basis für die Datenerhebung und die Demonstrationen auf der Abschlusspräsentation.



Abbildung 14: Personas die im Rahmen des KARLI Projekts entwickelt wurden

2.1.2.2 Arbeitspaket AP 200 (UAP 210, 220, 230)

Im Rahmen der Entwicklung eines effektiven Labeling-Workflows für das KARLI-Projekt wurde zunächst eine umfassende Analyse verschiedener Open-Source-Labeling-Tools durchgeführt. Im Rahmen der Evaluierung wurden Lösungen untersucht, die die Annotation von Videodaten in Bezug auf Zeitabschnitte, Bounding Boxes und Keypoints ermöglichen. Es zeigte sich, dass keines der geprüften Tools alle spezifischen Anforderungen von KARLI vollständig erfüllte. Zeitgleich wurde ein detailliertes Labeling-Konzept entwickelt, wobei erste Spezifikationen der Anforderungen in Zusammenarbeit mit Continental erarbeitet wurden. Im Anschluss wurden fünf geeignete Tools einem Vergleich unterzogen, woraus zwei potenzielle Tools zur Beschaffung ausgewählt wurden. Ein intensiver Austausch mit VICOMTECH aus Spanien diente der Vorbereitung, insbesondere durch die Analyse eines relevanten Papers sowie die Diskussion geeigneter Datensätze für Innenraumaufnahmen im Fahrzeugkontext.

Für das KARLI Projekt wurde das WebLabel-Tool von VICOMTECH als Labeling-Tool für das Konsortium ausgewählt und eine Nutzung über die Projektdauer mit VICOMTECH vereinbart. Der WebLabel Editor (GitHub: Vicomtech/weblabel) ist eine vielseitige Webanwendung, die für die Erstellung, Bearbeitung und Visualisierung von Anmerkungen (Labels oder Tags) im OpenLABEL-Format entwickelt wurde. Dieses Format wird häufig im Bereich des maschinellen Lernens und der Computer Vision verwendet, um gekennzeichnete Daten zu organisieren und Anwendungen wie Objekterkennung, semantische Segmentierung und zeitliche Ereignisverfolgung zu erleichtern.

Zu den Hauptmerkmalen des WebLabel Editors gehört die Unterstützung verschiedener Anmerkungstypen. Dazu gehören:

1. **2D-Begrenzungsrahmen in Bildern und Videos:** Ideal für Objekterkennungsaufgaben, da Benutzer rechteckige Bereiche um Objekte in Bildern oder Videobildern herum beschriften können.

2. **3D-Quader und Punktwolkensegmentierung:** Unverzichtbar für 3D-Datenanwendungen wie autonomes Fahren, bei denen räumliche Daten von LIDAR- oder Tiefensensoren mithilfe von 3D-Quadern beschriftet und in identifizierbare Bereiche segmentiert werden.
3. **2D-/3D-Punkte, -Polylinien und -Polygone:** Ermöglicht komplexere Formanmerkungen und eignet sich daher für Anwendungen, die feine Details erfordern, wie z. B. semantische Segmentierung und Erkennung von Gesichtsmerkmalen.
4. **Zeit-Ereignisse und Aktionen:** Unterstützt die zeitliche Kennzeichnung von Aktivitäten oder Ereignissen, was bei Videoannotationsaufgaben nützlich ist, um Veränderungen im Laufe der Zeit zu erfassen.
5. **Zeitkonsistente Objekte (mit eindeutigen IDs):** Erleichtert die Verfolgung desselben Objekts über verschiedene Frames oder Szenen hinweg, indem eindeutige Kennungen (UIDs) zugewiesen werden, um die Kontinuität zu wahren und eine erneute Identifizierung des Objekts zu ermöglichen.

WebLabel Editor ist für OpenLABEL-konforme JSON-Dateien optimiert und ermöglicht eine nahtlose Interoperabilität mit anderen Tools und Frameworks, die diesem Standard entsprechen. Es dient als leistungsstarkes Tool in der Datenaufbereitungs pipeline und erfüllt verschiedene Beschriftungsanforderungen in 2D- und 3D-Datenbereichen.

Auf Basis der Anforderungen der Use Cases "Motion Sickness" und "Levelkonformes Verhalten" und den Möglichkeiten des WebLabel Editor wurde eine Labeling-Hierarchie definiert, welche Keypoints, Aktivitäten und Objekte umfasst. Die annotierten Videodaten wurden für Aktivitäten manuell gelabelt, während Körperposen und Objekte automatisiert markiert wurden. Dadurch konnte eine konsistente und effiziente Datenbasis für das KARLI-Projekt gewährleistet werden.



Abbildung 15: Gastvortrag von Vicomtech zum Thema Labelingtools und Semi-automatisches Labeling

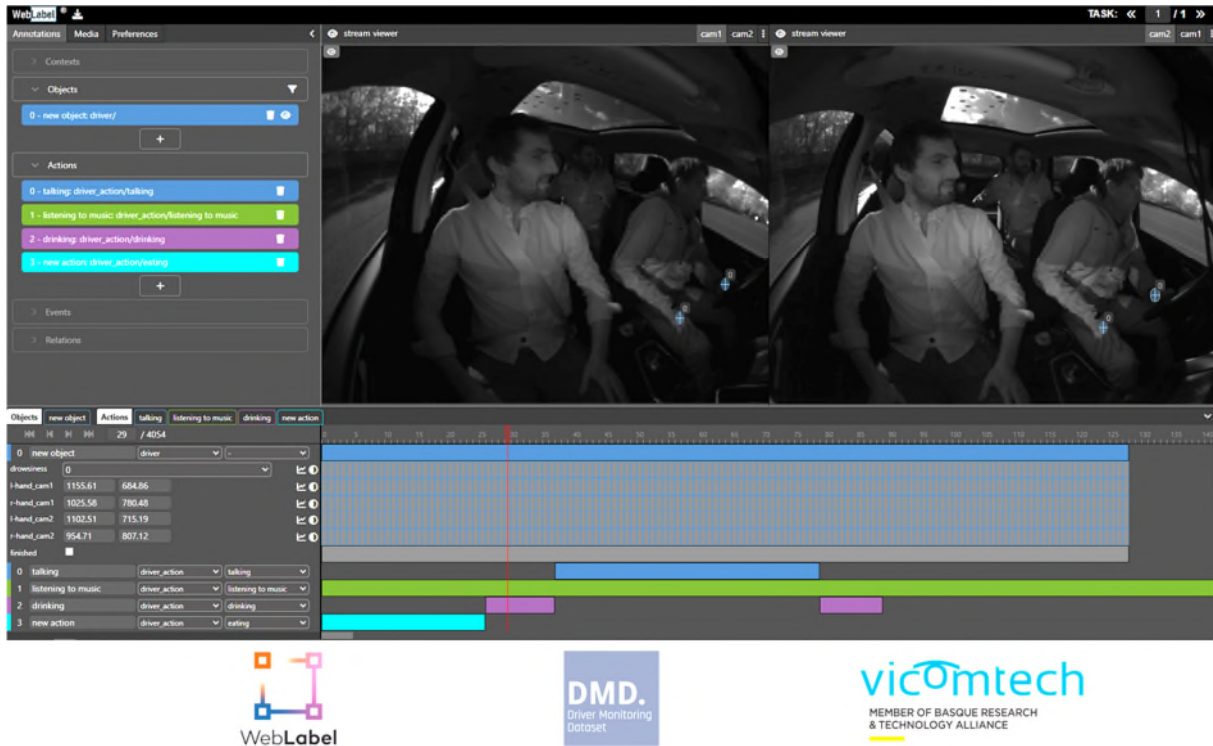


Abbildung 16: Screenshot aus dem Vicomtech Labeling-Tool. Für jedes Frame im Video lassen sich Aktivitäten, Objekte und wichtige Punkte festlegen.

2.1.2.3 Arbeitspaket AP 300 (UAP 312, 321)

Zur Konzeption eines Innenraumerfassungssystems für KARLI wurde die Eignung unterschiedlicher Kameratechnologien wie Tiefenkameras, Mono- und Stereo- sowie Infrarot- und Farbkameras analysiert. Ein Stereosystem mit Nahinfrarotbeleuchtung wurde für die Verwendung in KARLI ausgewählt, wobei augensichere Beleuchtung in Abstimmung mit Continental spezifiziert, und die erforderliche Hardware beschafft wurde. Parallel entstand ein Algorithmenkonzept zur Erfassung spezifischer Parameter auf Basis der Use Cases, das in einer Prototypenversion live auf dem M12-Meeting präsentiert wurde.

Zur Weiterentwicklung von Lernverfahren für die videobasierte Innenraumanalyse wurde zunächst eine Literaturrecherche zum Stand der Technik, insbesondere zur Erkennung von Nebentätigkeiten bei automatisierter Fahrt, durchgeführt. Basierend auf dieser Analyse wurden Verbesserungsmöglichkeiten in den überwachten Lernverfahren identifiziert, was zur Implementierung und zum Training von Modellen führte, die sukzessive durch semi- und selbstüberwachte Ansätze erweitert wurden. Die Transformation der Systeme von rein überwachten hin zu schwach und selbstüberwachten Lernverfahren mündete in der Entwicklung eines Small2Big Data-Ansatzes. Dabei erfolgte das Training der Modelle teils auf öffentlichen Datensätzen, bevor schließlich eine Optimierung auf projektspezifischen Daten durchgeführt wurde.



Abbildung 17: Aufzeichnungssystem des Fraunhofer IOSB integriert im Versuchsträger des Fraunhofer IAO.



Abbildung 18: Innenraumkamera des Fraunhofer IOSB im Versuchsträger des Projektpartners Continental.

Im Rahmen der Weiterentwicklung vom überwachten hin zu selbstüberwachten Lernverfahren wurden mehrere Modellarchitekturen und Methoden implementiert und optimiert, insbesondere für die Anforderungen der videobasierten Innenraumanalyse. Im Rahmen von KARLI wurde ein neuartiges selbstüberwachtes Lernverfahren für skelettbasierte Daten entwickelt. Die Ergebnisse wurden auf der ITSC 2024 (IEEE International Conference on Intelligent Transportation Systems) veröffentlicht und der entsprechende Code ist öffentlich verfügbar (GitHub: davidjlerch/darssl).

Dieses Verfahren SkelDINO-SAM baut auf dem sogenannten „DINO“ Ansatz (Caron et al., 2021; GitHub: facebookresearch/dino) aus der generellen Bildverarbeitung auf.

Die vorgeschlagene SkelDINO-SAM-Methode ist ein selbstüberwachender Ansatz zur Verbesserung der skelettbasierten Aktionserkennung durch die Kombination des Selbstdestillations-Frameworks von DINO mit gezielten Augmentierungstechniken für Skelettdaten:

1. **SkeletonDINO:** Diese von DINO inspirierte Methode verwendet zwei Encoder (Schüler und Lehrer), die robuste Skelettdarstellungen erlernen. Durch die Verarbeitung verschiedener erweiterter Versionen einer Eingabeskelettsequenz wird das Schülernetzwerk dazu gebracht, die Einbettungen des Lehrers über den Kreuzentropieverlust abzugleichen. Nur das Schülernetzwerk aktualisiert seine Parameter, um die Ähnlichkeit der Merkmale mit der Ausgabe des Lehrers zu maximieren.
2. **Separated Augmentation Modules (SAM):** SAM führt zwei spezialisierte Erweiterungen ein:
 - **Central Spatial Masking (CSM)** konzentriert sich auf die Maskierung von Gelenken mit hoher Konnektivität im Skelett, wodurch es für das Modell einfacher wird, Gelenkbeziehungen zu erfassen.
 - **Motion Attention Temporal Masking (MATM)** priorisiert Frames mit höherer Bewegung, da sie mehr aktionsbezogene Informationen enthalten.
3. **SkelDINO-SAM-Integration:** Hierbei wird SkeletonDINO mit SAM kombiniert, um die Eingaben des Studentennetzwerks durch Gelenk- und Rahmenmaskierung zu diversifizieren. Der Lehrer verarbeitet nur die ursprünglichen und verauschten Skelettsequenzen, während der Schüler diese sowie zusätzliche erweiterte Versionen erhält. Durch die Minimierung der Kreuzentropie zwischen den Ausgaben von Lehrer und Schüler lernt das Studentennetzwerk eine verallgemeinerte Darstellung, die in der Lage ist, mit unterschiedlichem Rauschen und Verdeckungen umzugehen.

Die gezielten Erweiterungen der Methode verbessern die Fähigkeit des Modells, komplexe skelettbasierte Daten zu verallgemeinern. Allerdings sind die Aktivitäten im Fahrzeuginnenraum unausgewogen verteilt. Dies führt zu Problemen im Training neuronaler Netze. Um dieses Problem zu adressieren hat das Fraunhofer IOSB in KARLI eine Methode entwickelt, um unausgewogenen Klassenverteilungen im Training entgegenzuwirken.

Die **Clustered Feature Weighting (CFW)**-Methode ist eine Strategie zum markierungsfreien Laden von Daten, die das Klassenungleichgewicht in Datensätzen, insbesondere bei der Erkennung von Ablenkungen des Fahrers, durch die Kombination von Transferlernen, unüberwachtem Clustering und gewichtetem Sampling behebt. CFW arbeitet in vier Phasen: Merkmalsextraktion, Varianzanalyse, Clustering und gewichtete Zufallsstichproben. Jede Phase verbessert schrittweise das Klassenverhältnis innerhalb von Stapeln und fördert so ein effektiveres Modelltraining.

1. **Bildkodierung und Merkmalsextraktion:**
 - Ein vorab trainierter Vision Transformer (ViT), der in der Regel auf einem großen Datensatz wie ImageNet trainiert wird, wird verwendet, um Merkmalseinbettungen aus den Eingabebildern zu extrahieren. Diese Einbettungen erfassen wesentliche klassenspezifische Informationen ohne Labels und nutzen die Fähigkeit von ViT zur Darstellung umfangreicher Merkmale.
 - Um festzustellen, ob das ViT-Modell für die Merkmalsextraktion geeignet ist, werden Varianzen innerhalb und zwischen den Klassen berechnet. Die **Intraklassenvarianz** quantifiziert die Merkmalskonsistenz innerhalb jeder Klasse (eine geringe Varianz deutet auf eine gute Merkmals-homogenität hin), während

die **Interklassenvarianz** die Trennung zwischen den Klassenmitten misst. Eine höhere Interklassenvarianz deutet darauf hin, dass verschiedene Klassen besser unterscheidbar sind.

- Die Abstände zwischen den Klassenmittelpunkten werden außerdem mit dem euklidischen Abstand gemessen, wodurch eine quantitative Bewertung der Merkmalsseparation möglich ist. Diese Metriken bilden die Grundlage für die Auswahl eines geeigneten Modells, um eine effektive Klassentrennung bei der anschließenden Clusterbildung zu gewährleisten.

2. **Feature Clustering:**

- Mit dem **HDBSCAN**-Algorithmus werden Einbettungen auf der Grundlage der Kosinus-Ähnlichkeit geclustert, wodurch Gruppen erstellt werden, die zugrunde liegende Klassenmuster ohne Labels darstellen. HDBSCAN wurde aufgrund seiner Fähigkeit ausgewählt, mit verrauschten Daten umzugehen und Cluster mit variabler Dichte zu entdecken, wodurch es für Datensätze mit komplexer Struktur effektiv ist.
- Die Kosinus-Ähnlichkeitsmatrix für jeden Stapel wird berechnet, um Beziehungen zwischen Einbettungen zu erfassen, und in eine Kosinus-Abstandsmatrix umgewandelt, die dann als Eingabe für das HDBSCAN-Clustering dient.
- Durch das Clustering auf diese Weise können Muster in den Daten aufgedeckt werden, wodurch implizit Pseudoklassen oder Gruppen erstellt werden, die der ursprünglichen Klassenverteilung nahekommen.

3. **Gewichtetes Zufalls-Sampling ohne Labels:**

- Sobald Cluster gebildet wurden, werden den Stichproben auf der Grundlage der Clustergröße Gewichte zugewiesen, wobei kleinere Cluster höhere Gewichte erhalten. Dieser Gewichtungsmechanismus gleicht das Ungleichgewicht der Klassen aus, indem er eine gleichmäßige Verteilung über die Cluster hinweg fördert und sicherstellt, dass auch Minderheitsklassen häufiger beprobt werden.
- Ausreißer, die von HDBSCAN identifiziert werden, erhalten ein minimales Gewicht, um ihre Präsenz im Trainingsprozess aufrechtzuerhalten, während gleichzeitig verhindert wird, dass sie das Modelltraining verzerren. Die Gewichtung für Ausreißer und der Parameter „min sample“ in HDBSCAN werden als Hyperparameter behandelt, die so abgestimmt werden, dass optimale Ergebnisse erzielt werden.
- Der generierte Gewichtsvektor dient als Grundlage für die gewichtete Zufallsstichprobe und gewährleistet einen ausgewogenen und vielfältigen Stichprobenprozess, der unterschiedliche Klassendarstellungen in jedem Trainingsbatch berücksichtigt.

4. **Anwendung zur Erkennung von Ablenkung beim Fahren:**

- Der CFW-Datenlader wird anhand mehrerer Aufgaben (binäre und mehrklassige Erkennung von Ablenkung beim Fahren) sowohl mit überwachten als auch mit selbstüberwachten Lernansätzen evaluiert. Beim überwachten Lernen erkennt ein vorab trainiertes ViT-Modell mit einem linearen Klassifikator Ablenkungstypen, während beim selbstüberwachten Lernen der DINOv2-Encoder als eingefrorenes Rückgrat fungiert und sich auf das Lernen von Fahrerreaktionen aus nicht gekennzeichneten Daten konzentriert.
- Durch die Ausbalancierung der Stichproben innerhalb jedes Stapels verbessert CFW die Darstellung von Minderheitsklassen erheblich und steigert die Modellleistung bei der Erkennung von Ablenkungen beim Fahren. Zur Bewertung der Wirksamkeit von CFW werden verschiedene Encoder-Architekturen, darunter

CNNs und Vision Transformers, verwendet, wobei bei beiden Modelltypen Verbesserungen beobachtet wurden.

Die CFW-Methode verbessert die Modellleistung durch Erhöhung der Klassenvielfalt innerhalb von Stapeln, was besonders bei unausgewogenen, nicht beschrifteten Daten von Vorteil ist. Die Kombination von unüberwachtem Clustering mit gewichtetem Sampling ermöglicht einen effektiven Ansatz zur Behebung von Datenungleichgewichten in kennzeichnungsfreien Umgebungen.

2.1.2.4 Arbeitspaket AP 500 (UAP 510, 520)

Im Rahmen eines Integrationsworkshops erfolgte in Zusammenarbeit mit dem Projektpartner Ford die Planung der technischen Umsetzung zur Lieferung, Integration und Software-Schnittstellen der IOSB-Kamera im Ford-Testfahrzeug. In ähnlicher Weise wurden Vorbereitungen für die Integration der Kamera in das Testfahrzeug des Fraunhofer IAO getroffen, welches zur Erhebung von Daten für Motion-Sickness-Studien genutzt wird. Die Konfiguration und der Test der Kameraerfassungssysteme sind für beide Fahrzeuge abgeschlossen. Die Integration wurde ebenfalls abgeschlossen und auf der Abschlusspräsentation demonstriert. Parallel dazu wurden erste Schritte zur Konzeption einer echtzeitfähigen Datenverarbeitung sowie die Beschaffung zusätzlicher Kamerahardware eingeleitet, um den simultanen Aufzeichnungsbedarf bei den Partnern zu decken.

Im Rahmen der Entwicklung der videobasierten Innenraumanalyse erfolgte zunächst die Konzeption überwachter Lernverfahren. Dies beinhaltete eine umfassende Literaturrecherche zum Stand der Technik in der Fahrzeuginnenraumanalyse, insbesondere mit Fokus auf die Erkennung von Aktivitäten, die Nebentätigkeiten bei automatisierter Fahrt sowie Motion Sickness auslösen können. Auf Basis der Recherche wurden Optimierungspotenziale identifiziert und in ersten Anpassungen bestehender Systeme des Fraunhofer IOSB umgesetzt. Gleichzeitig wurde eine Literaturrecherche zu semi-supervised Learning durchgeführt, um die vorhandenen Algorithmen in Richtung schwach überwachter Lernverfahren weiterzuentwickeln. Die Integration der Kamerasysteme in die Testfahrzeuge ist nun abgeschlossen, ebenso wie die Konzeption für eine Verarbeitung in Echtzeit.

Es wurde eine internationale Online-Umfrage durchgeführt, um Aktivitäten zu ermitteln, die Personen aufgrund ihrer persönlichen Erfahrungen mit Motion Sickness während Autofahrten in Verbindung bringen. Die auf der Qualtrics-Plattform entwickelte Umfrage umfasste sowohl häufige Aktivitäten wie Lesen und die Nutzung elektronischer Geräte als auch weniger häufig genannte Aktivitäten wie das Schminken und das Fotografieren. Diese zusätzlichen Aktivitäten wurden aufgrund ihrer Erkennung durch moderne Insassenüberwachungssysteme ausgewählt, um den Teilnehmern eine breitere Palette von Szenarien zur Verfügung zu stellen. Neben der Berichterstattung über Aktivitäten wurden die Teilnehmer auch gebeten, persönliche Methoden oder Umgehungslösungen zur Reduzierung der Motion Sickness bei der Ausführung dieser Aufgaben mitzuteilen.

Um Zugänglichkeit und Genauigkeit zu gewährleisten, wurde die Umfrage mit DeepL in die Hauptsprache jedes Ziellandes übersetzt, wobei die Übersetzungen von Muttersprachlern überprüft wurden. Die Teilnehmer wurden informell über soziale Netzwerke, E-Mails und Chatgruppen rekrutiert, um eine vielfältige internationale Stichprobe zu erhalten. Die Umfrage war so konzipiert, dass sie schnell durchgeführt werden konnte und weniger als zwei Minuten dauerte. Es wurden keine demografischen Daten erhoben, sodass eine DSGVO-konforme, anonyme

Teilnahme möglich war. Für eine aussagekräftige Analyse wurden nur Antworten aus Ländern mit mindestens 50 Teilnehmern in den endgültigen Datensatz aufgenommen.

Die Umfrage bestand aus zwei Hauptfragen. Zunächst wählten die Teilnehmer aus einer Liste von Aktivitäten aus, die ihrer Erfahrung nach zu Motion Sickness führen. Sie konnten mehrere Optionen auswählen und weitere, nicht aufgeführte Aktivitäten hinzufügen. Anschließend gaben die Teilnehmer offene Antworten zu Strategien, die sie zur Linderung oder Vermeidung von Symptomen der Motion Sickness anwenden.

Die Antworten von insgesamt 493 Teilnehmern aus sechs Ländern zeigten eine Reihe von Aktivitäten, die mit Motion Sickness in Verbindung gebracht werden. Lesen wurde in allen Regionen am häufigsten als Auslöser genannt, gefolgt von der Nutzung elektronischer Geräte, dem Anschauen von Videos und dem Blick nach hinten. Seltener genannte Aktivitäten wie Schminken oder Rasieren wurden als weniger relevant für Motion Sickness angesehen. Während die Reaktionsmuster von Land zu Land leicht variierten – deutsche Befragte wählten beispielsweise eher mehrere Aktivitäten aus – waren die allgemeinen Trends bemerkenswert konsistent, was auf gemeinsame globale Erfahrungen mit Auslösern von Motion Sickness hindeutet.

Die Teilnehmer teilten auch eine Reihe von Strategien zur Verringerung der Motion Sickness mit, die in sieben Hauptthemen unterteilt wurden. Zu den ernährungsbasierten Lösungen gehörten das Trinken von Wasser, das Vermeiden fettiger Lebensmittel und der Verzehr von Ingwer oder sauren Lebensmitteln wie Orangen. Auch Medikamente speziell gegen Motion Sickness wurden häufig erwähnt. Andere beliebte Gegenmaßnahmen waren Schlafen, aus dem Fenster schauen und das Auto lüften. Häufig wurden organisatorische Anpassungen empfohlen, wie z. B. die Positionierung von Geräten auf Augenhöhe, aufrechtes Sitzen und die Reduzierung der Gerätenutzung. Alternative Ansätze, wie das Hören von Musik, Hörbüchern oder das Führen von Gesprächen, dienten der Ablenkung. Darüber hinaus nannten einige Befragte physische Techniken wie Akupressur an bestimmten Punkten.

Die Ähnlichkeit der Ergebnisse in den verschiedenen Ländern spiegelt wahrscheinlich gemeinsame Muster bei der Fahrzeugnutzung und die gemeinsame Physiologie der Motion Sickness wider. Die Studie stieß jedoch auf einige Einschränkungen. Die Reihenfolge, in der die Aktivitäten aufgelistet wurden, könnte die Auswahl beeinflusst und eine potenzielle Verzerrung der Reihenfolge verursacht haben. Der Rekrutierungsansatz, der sich auf Freundes- und Familiennetzwerke stützte, könnte zu einer Verzerrung der Stichprobe geführt haben, während das anonyme Format einige Teilnehmer dazu ermutigt haben könnte, schnell und ohne sorgfältige Überlegung zu antworten. Unterschiede bei der Übersetzung könnten sich ebenfalls auf subtile Weise auf das Antwortverhalten ausgewirkt haben, da deutsche Teilnehmer eher dazu neigten, mehrere Optionen auszuwählen als Teilnehmer aus anderen Ländern.

Trotz dieser Einschränkungen bot die breite Beteiligung an der Umfrage einen umfassenden Einblick in Aktivitäten, die zur Motion Sickness beitragen, sowie in wirksame Strategien zu ihrer Linderung. Die Ergebnisse zeigen das Potenzial für die Integration dieser Erkenntnisse in Assistenzsysteme zur Verbesserung des Komforts von Fahrzeuginsassen auf, insbesondere da fortschrittliche Überwachungs- und Erkennungssysteme sich ständig weiterentwickeln.

Für die Fahrzeuginnenraumanalyse wurde das Vision-language Model MobileVLM V2 (GitHub: Meituan-AutoML/MobileVLM) eingesetzt. Das Modell MobileVLM V2 (Chu et al., 2024) ist ein multimodales System für die Verarbeitung von Bild und Sprache, das aus drei Hauptkomponenten

besteht: einem Bildkodierer, einem Sprachmodell und einem leichtgewichtigen Downsample-Projektor (LDPv2). Hier eine Zusammenfassung der wichtigsten Komponenten und Funktionen:

1. **Vision Encoder:** Das Modell verwendet einen CLIP-basierten ViT-L/14-Vision-Encoder, der Bilder in semantische Einbettungen umwandelt. Der Encoder wird vorab auf einem großen Datensatz von Bild-Text-Paaren trainiert, um eine starke Bild-Sprache-Ausrichtung zu gewährleisten. Er gibt visuelle Merkmale auf hoher Ebene aus den Eingabebildern aus, die dann in die nächsten Phasen des Systems eingespeist werden.
2. **Sprachmodell:** MobileLLaMA dient als Sprachverarbeitungs-komponente des Modells. MobileLLaMA wurde für die Effizienz auf mobilen und ressourcenbeschränkten Geräten entwickelt, verarbeitet tokenisierte Texteingaben und ist während des mehrstufigen Trainingsprozesses vollständig trainierbar. Diese Struktur ermöglicht die effektive Integration von Text- und visuellen Merkmalen, die kombiniert werden, um Antworten in einem Konversations- oder Q&A-Format zu generieren.
3. **Lightweight Downsample Projector (LDPv2):** Diese Komponente gleicht die Dimensionalität visueller Merkmale durch punkt- und tiefenweise Faltungen, Durchschnittspooling und Positionskodierung mit dem Sprachmodell ab. Der aktualisierte LDPv2 ist so optimiert, dass er die Parameter im Vergleich zu früheren Versionen um fast 99,8 % reduziert, wodurch sowohl die Ausrichtung als auch die Verarbeitungsgeschwindigkeit verbessert werden.
4. **Trainingsstrategie:** MobileVLM V2 durchläuft ein zweiphasiges Training. Im Vorab-Training wird das Modell einem umfangreichen Datensatz von Bild-Text-Paaren ausgesetzt, in dem es die Bild-Text-Ausrichtung erlernt. Im anschließenden Multitask-Training verwendet es mehrere spezialisierte Datensätze (z. B. Visual Dialog, Text-VQA), um sich in verschiedenen Bild-Sprach-Aufgaben wie Dialog, Szenenverständnis und textbasierte Fragen und Antworten zu üben. Diese umfangreiche Multitask-Phase bereitet das Modell auf komplexe nachgelagerte Anwendungen vor.

2.1.2.5 Arbeitspaket AP 600 (UAP 620, 630, 640)

Im Rahmen eines Abstimmungsmeetings, an dem drei Teilnehmer beteiligt waren, fand eine Präsenzveranstaltung am Fraunhofer IAO statt. Im Rahmen der Veranstaltung wurde das "Wizard-of-Oz"-Fahrzeug präsentiert und eine Live-Demonstration der aktuellen technischen Möglichkeiten durchgeführt. Im Vorfeld der Präsentation wurden mehrere Architekturkonzepte für das Fahrzeug entwickelt, welche den Einsatz und die Integration der Systeme in das Testfahrzeug umfassten. Die zu erfassenden und zu synchronisierenden Datenquellen wurden definiert: Sowohl die Sensoren von Continental als auch diejenigen des Fraunhofer IOSB werden für die Datenerfassung herangezogen und miteinander synchronisiert.

Die Systemarchitektur für das Messsystem im Fraunhofer-IAO-Fahrzeug wurde seitens des IOSB finalisiert. Zu diesem Zweck erfolgt eine Verbindung des IOSB-Systems über eine Schnittstelle mit dem Conti-System, wobei die Synchronisierung ebenfalls durch das IOSB realisiert wird. Die hierfür definierten Hardwareanforderungen wurden spezifiziert und befinden sich aktuell in der Beschaffungsphase.

In enger Abstimmung mit den Projektpartnern Continental und Fraunhofer IOSB wurden alle Details zur Integration der Innenraumkameras festgelegt. Des Weiteren wurde die Spezifikation und Beschaffung der benötigten Workstation initiiert, wobei sich einige Arbeitsschritte aufgrund von Verzögerungen im Fahrzeugaufbau noch in der Umsetzung befinden. Die Kalibrierung der Innenraumkameras im Fahrzeug ist abgeschlossen und die Integration des

Aufzeichnungssystem wurde erfolgreich in das Wizard-of-Oz-Fahrzeug des Fraunhofer IAO integriert.

Im Rahmen der Testträger für spezifische Anwendungsfälle konnte die Innenraumanalyse für die Studien zu Motion Sickness sowie für levelkonformes Verhalten erfolgreich integriert werden. Während der laufenden Datenerhebungskampagnen wird technische Unterstützung kontinuierlich bereitgestellt. Die Weiterentwicklung der Verfahren zur Innenraumanalyse erfolgt in iterativen Schritten und wird kontinuierlich in die Versuchsträger integriert, um eine fortlaufende Optimierung der Systemleistung zu gewährleisten.

Die Implementierung des Aufzeichnungskonzepts in der Software ist vollständig abgeschlossen, einschließlich der Entwicklung und Bereitstellung von Exportfunktionen für die aufgezeichneten Daten. Gleichzeitig wurde mit der Optimierung der maschinellen Lernverfahren aus den Arbeitspaketen AP 300 und AP 500 begonnen, um die Echtzeitfähigkeit der Algorithmen sicherzustellen. Diese Optimierungen werden fortlaufend weiterentwickelt, um die Anforderungen im Fahrzeugbetrieb zu erfüllen.

Die Implementierung von Schnittstellen erfolgte mit dem Ziel einer nahtlosen Integration in die bestehende Fahrzeuginfrastruktur. Dieser Prozess wurde in sämtlichen Fahrzeugen erfolgreich abgeschlossen. Des Weiteren wurden die Schnittstellen zur eCAL-Architektur von Continental erweitert, wodurch eine optimierte Kommunikation und Synchronisation der Datenströme zwischen den Systemen gewährleistet wird. Diese Maßnahmen sind von entscheidender Bedeutung, um die Echtzeitanforderungen des Gesamtsystems zu erfüllen und eine stabile, effiziente Datenerhebung und -verarbeitung sicherzustellen. Mithilfe dieser Schnittstelle konnte das Fraunhofer IOSB Erkennungssystem in die Erkennungssysteme von Continental integriert werden. Dadurch sind Schnittstellen für zukünftige Zusammenarbeiten bereits vorhanden.

2.1.2.6 Arbeitspaket AP 700 (UAP 720, 730)

Die in den Tabellen dargestellten Ergebnisse demonstrieren signifikante Unterschiede in den Reaktionszeiten und Fehlerquoten zwischen dem generierten User Interface (GenUI) und dem statischen Interface. Die Ergebnisse der Abbildung 18 legen nahe, dass das GenUI in vier von fünf Fällen schneller als das statische Interface reagierte. Lediglich ein Proband wies eine geringfügig langsamere Reaktionszeit bei GenUI auf. Dies lässt den Schluss zu, dass das generierte Interface eine höhere Anpassungsfähigkeit und Bedienfreundlichkeit aufweist, was für Anwendungen im Fahrzeuginnenraum von besonderem Wert ist, da dadurch schnelle und reibungslose Interaktionen ermöglicht werden.

Die Ergebnisse der Fehleranalyse in Abbildung 19 zeigen, dass beide Interfaces insgesamt eine geringe Fehlerquote aufweisen. Beim generierten Interface wurde lediglich ein Fehler registriert. Dies unterstreicht die Zuverlässigkeit und Benutzerfreundlichkeit des generierten Interfaces, was insbesondere im Kontext sicherheitskritischer Anwendungen von Bedeutung ist.

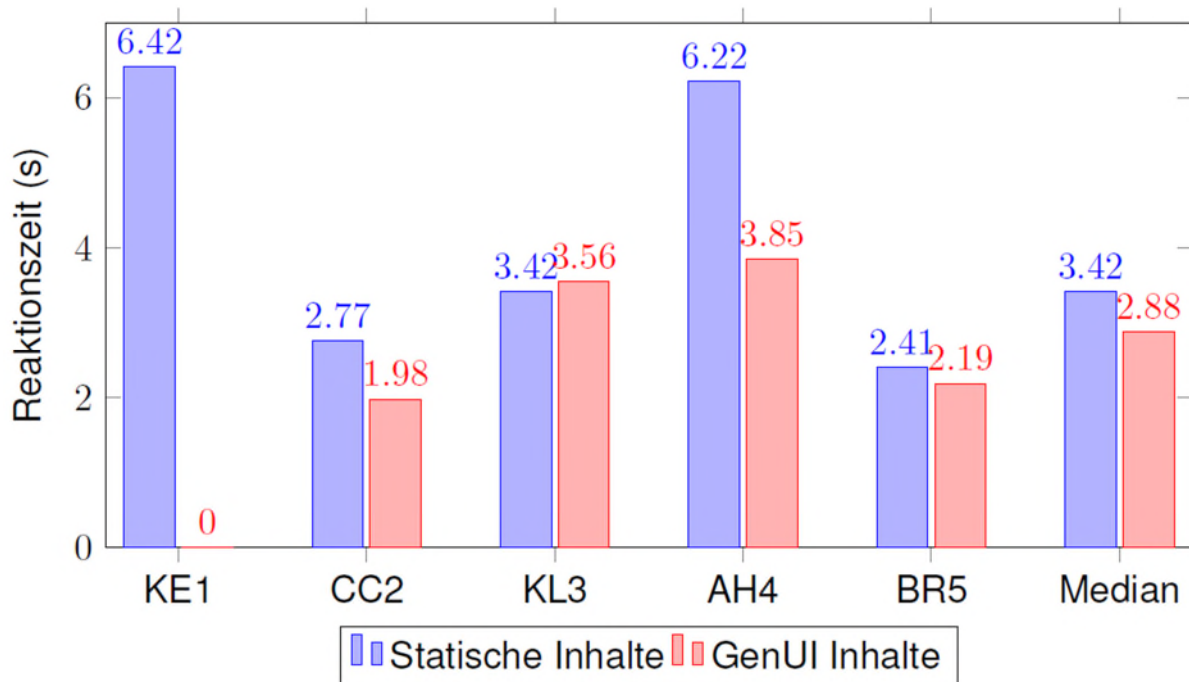


Abbildung 19: Evaluation der Reaktionszeiten für ein statisches und ein generiertes User Interface

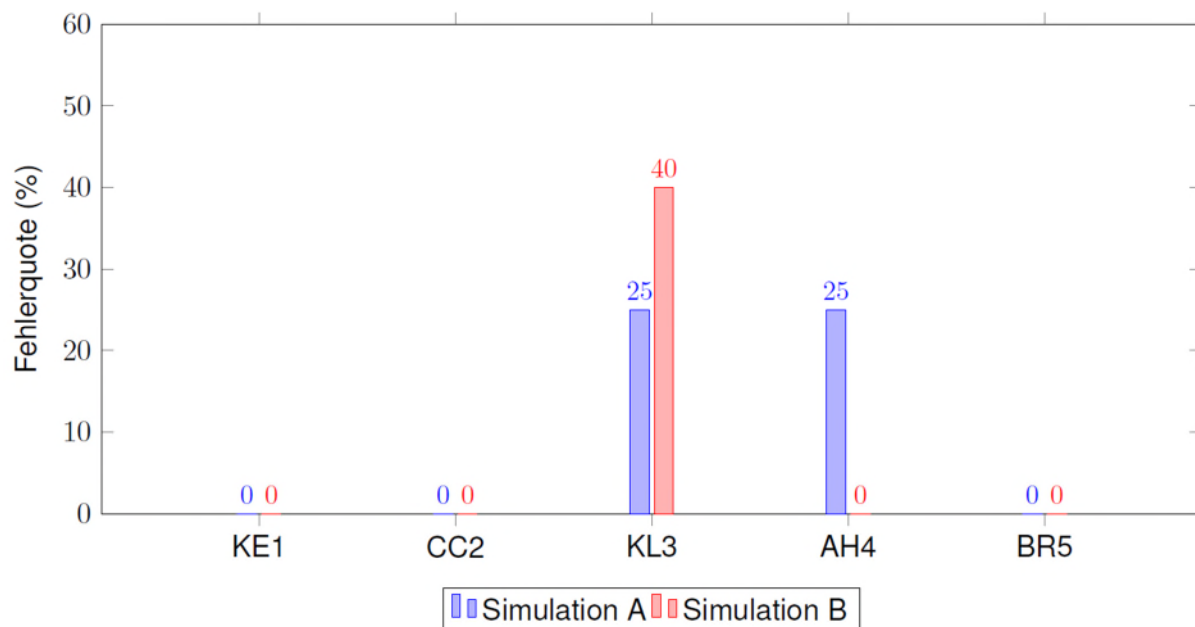


Abbildung 20: Evaluation der Fehlerquoten für ein statisches Interface (Simulation A) und ein generiertes Interface (Simulation B)

Die Analyse der Aktivitätenerkennung, wie sie in den Abbildungen 21 und 22 dargestellt ist, verdeutlicht die hohe Präzision des Systems bei der Identifikation spezifischer Aktivitäten, beispielsweise "telefonieren" und "trinken". Die Erkennung von Telefonaten weist eine hohe Übereinstimmung mit der tatsächlichen Aktivität auf. Allerdings zeigt sich in dieser Simulation,

dass die Robustheit noch weiter verbessert werden kann, um möglichen Interferenzen besser standzuhalten. Die Erkennung der Aktivität "Trinken" zeigt eine hohe Robustheit, wengleich mit einer leichten Zeitverzögerung. Dies lässt den Schluss zu, dass weitere Optimierungen erforderlich sind, um die Latenzzeiten zu minimieren und eine schnellere Reaktionsfähigkeit zu gewährleisten. Die insgesamt vielversprechenden Ergebnisse der Aktivitätenerkennung legen nahe, dass das System für den Einsatz in Fahrzeugen zunehmend geeignet wird, insbesondere durch gezielte Verbesserungen der Robustheit und Reaktionszeit.

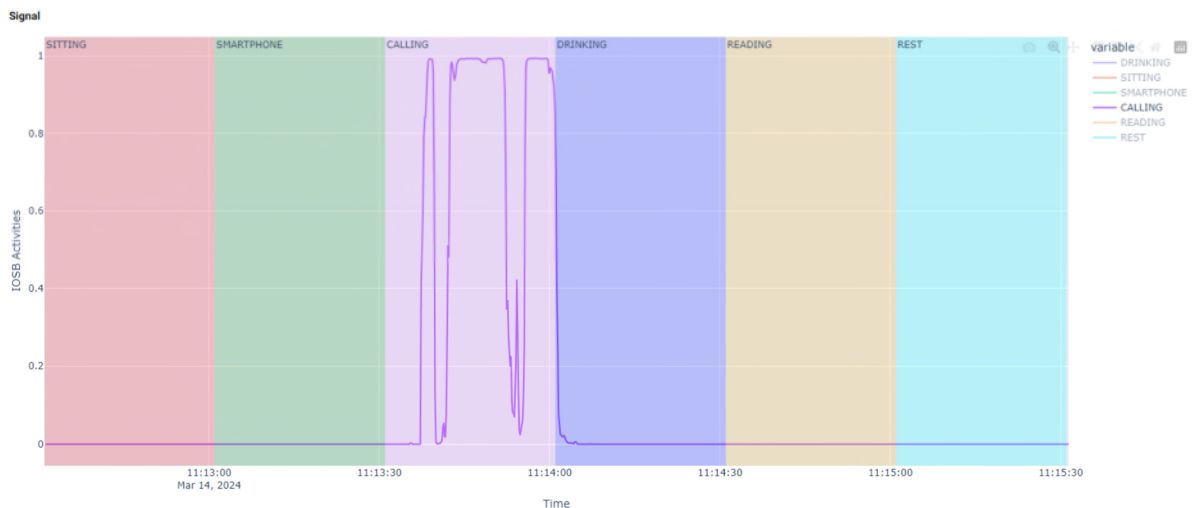


Abbildung 21: Auswertung der Aktivitätenerkennung für „telefonieren“ vom Projektpartner Continental

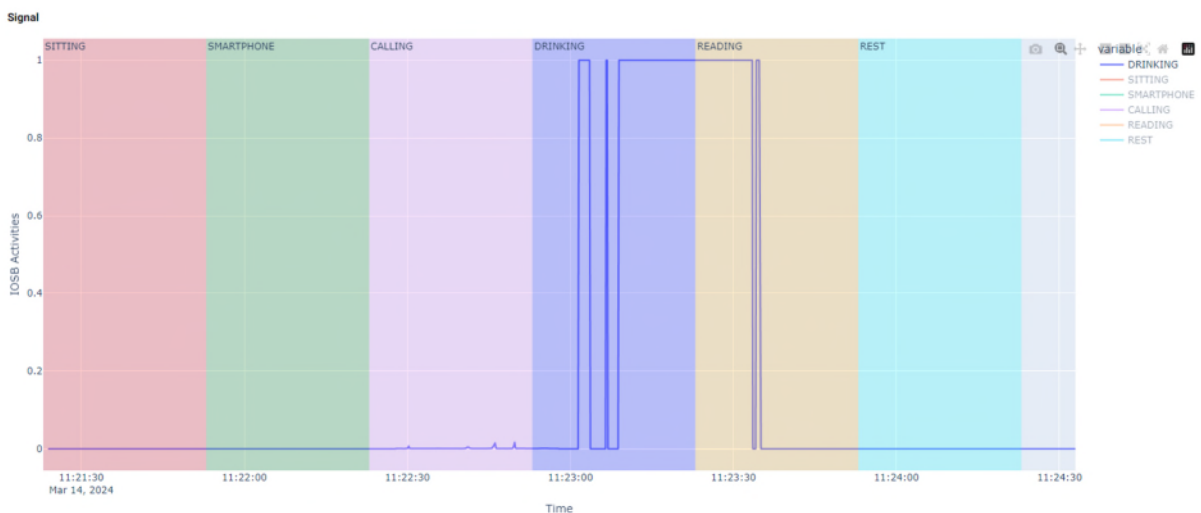


Abbildung 22: Auswertung der Aktivitätenerkennung für „trinken“ vom Projektpartner Continental

2.1.2.7 Arbeitspaket AP 800 (UAP 810, 820)

Das Fraunhofer IAO nahm regelmäßig an den Applikationsmeetings LKV und MS, sowie an den Steuerkreismetings teil. Im Rahmen der Ergebnisverbreitung wurden zum einen Inhalte der Algorithmenentwicklung aber auch der Datenerhebungen, mit Unterstützung des Projektbüros, über LinkedIn geteilt. Als wissenschaftlicher Partner hat das Fraunhofer IOSB stets Innovationen, die zum Erreichen der Projektziele hilfreich sind, aufbereitet und vorgestellt.

Für das Fraunhofer Magazin hat das Fraunhofer IOSB einen Artikel über die intelligente Interaktion im Auto der Zukunft verfasst. In diesem Artikel geht es darum, wie in KARLI KI-Assistenten entwickelt wurden. Mit der Kombination von Innenraumsensorik und Vision-Language-Modellen (VLM), die visuelle Daten mit Sprachmodellen verknüpfen, soll das Fahrzeug der Zukunft wie ein digitaler "Butler" agieren, der die Insassen kontextsensitiv unterstützt, ohne aufdringlich zu wirken. Die Technologien berücksichtigen dabei die verschiedenen Freiheitsgrade, die je nach Automatisierungslevel möglich sind: Während bei niedrigen Levels der Fokus auf dem Verkehr liegt, können sich Fahrer bei höheren Levels auch anderen Tätigkeiten widmen. Warnhinweise und Sicherheitsmeldungen informieren den Fahrer situationsgerecht und in passender Form, wie etwa bei bevorstehenden Straßenbedingungen oder der Notwendigkeit, das Fahren selbst zu übernehmen.

Im Folgenden werden die verschiedenen Präsentationen und Demonstrationen der im Rahmen des KARLI-Projekts entwickelten Technologien und deren Fortschritte vorgestellt, welche auf wichtigen Konferenzen und internen Veranstaltungen präsentiert wurden.

Die Abbildung 23 veranschaulicht die Präsentation der KARLI-Ergebnisse zur Fahreraktivitätenerkennung auf der renommierten European Conference on Computer Vision (ECCV) 2022. Die ECCV stellt eine Plattform für den Austausch zentraler technischer Fortschritte im Bereich Computer Vision dar. Des Weiteren wurden die Präsentationen zur Mensch-Maschine-Interaktion und Aktivitätenerkennung im Jahr 2023 und 2024 auf internationalen Konferenzen und Branchenevents wie der CarHMI (vgl. Abbildungen 24 und 26), der InCabin Konferenz (vgl. Abbildungen 25 und 28) sowie der ITSC 2024 (vgl. Abbildung 27) präsentiert.



Abbildung 23: Ergebnisverbreitung bei der ECCV 2022



Abbildung 24: Ergebnisverbreitung bei der CarHMI 2023



Abbildung 25: Ergebnisverbreitung bei der InCabin Konferenz 2024



Abbildung 26: Ergebnisverbreitung bei der CarHMI 2024

Des Weiteren präsentiert Abbildung 29 eine interne Demonstration des GenU-In am Fraunhofer IOSB in Karlsruhe. Die Abbildungen 30 und 31 veranschaulichen die Funktionalität des KI-Assistenten, der Motion Sickness induzierende Aktivitäten identifiziert und Empfehlungen zur Reduzierung von Reiseübelkeit bereitstellt. Die Abschlusspräsentation des Projekts (Abbildung 31) präsentiert die jüngsten Fortschritte des Motion-Sickness-Assistenten und demonstriert dessen Integration in das gesamte Assistenzsystem.



Abbildung 27: Ergebnisverbreitung bei der ITSC 2024



Abbildung 28: Ergebnisverbreitung bei der InCabin.Sensing 2024



Abbildung 29: Demonstration eines KI-Assistenten für den Fahrzeuginnenraum beim KARLI-Konsortialmeeting in Karlsruhe am Fraunhofer IOSB.

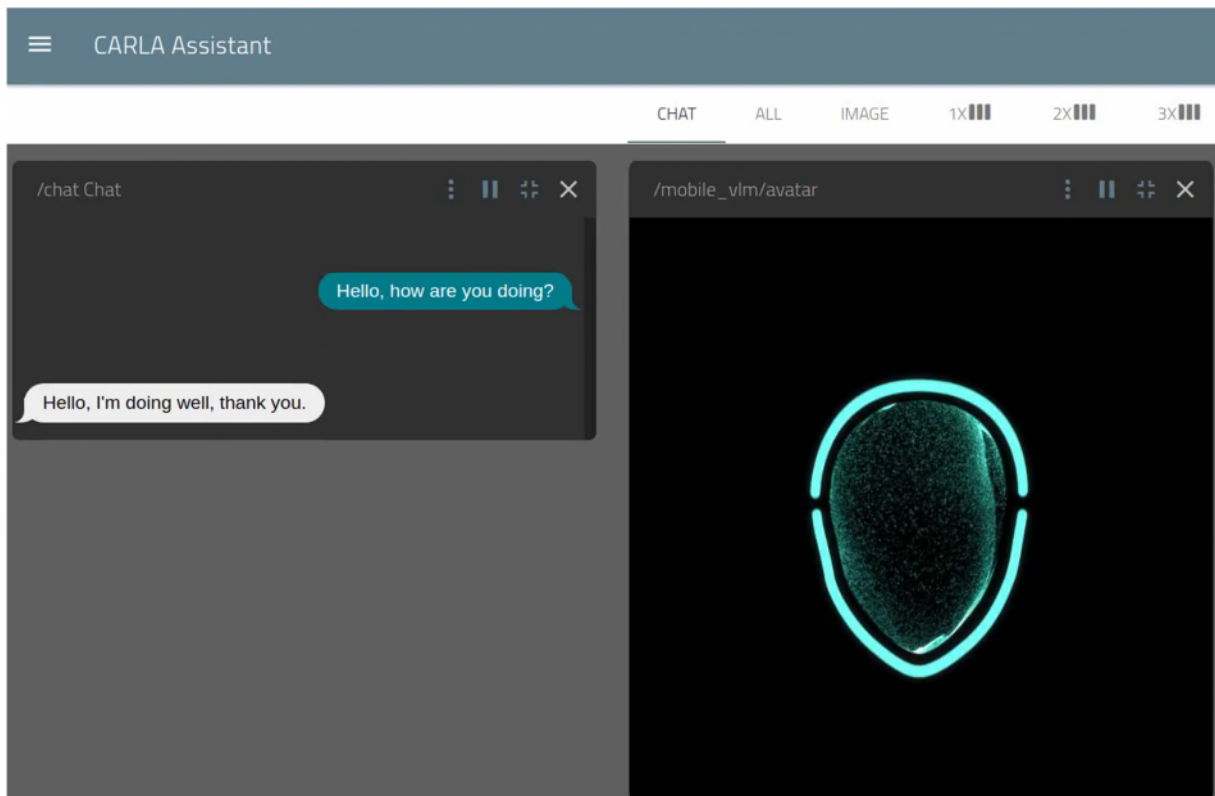


Abbildung 30: Screenshot des KI-Assistenten mit dem KARLI Avatar des Projektpartners Studiokurbos.

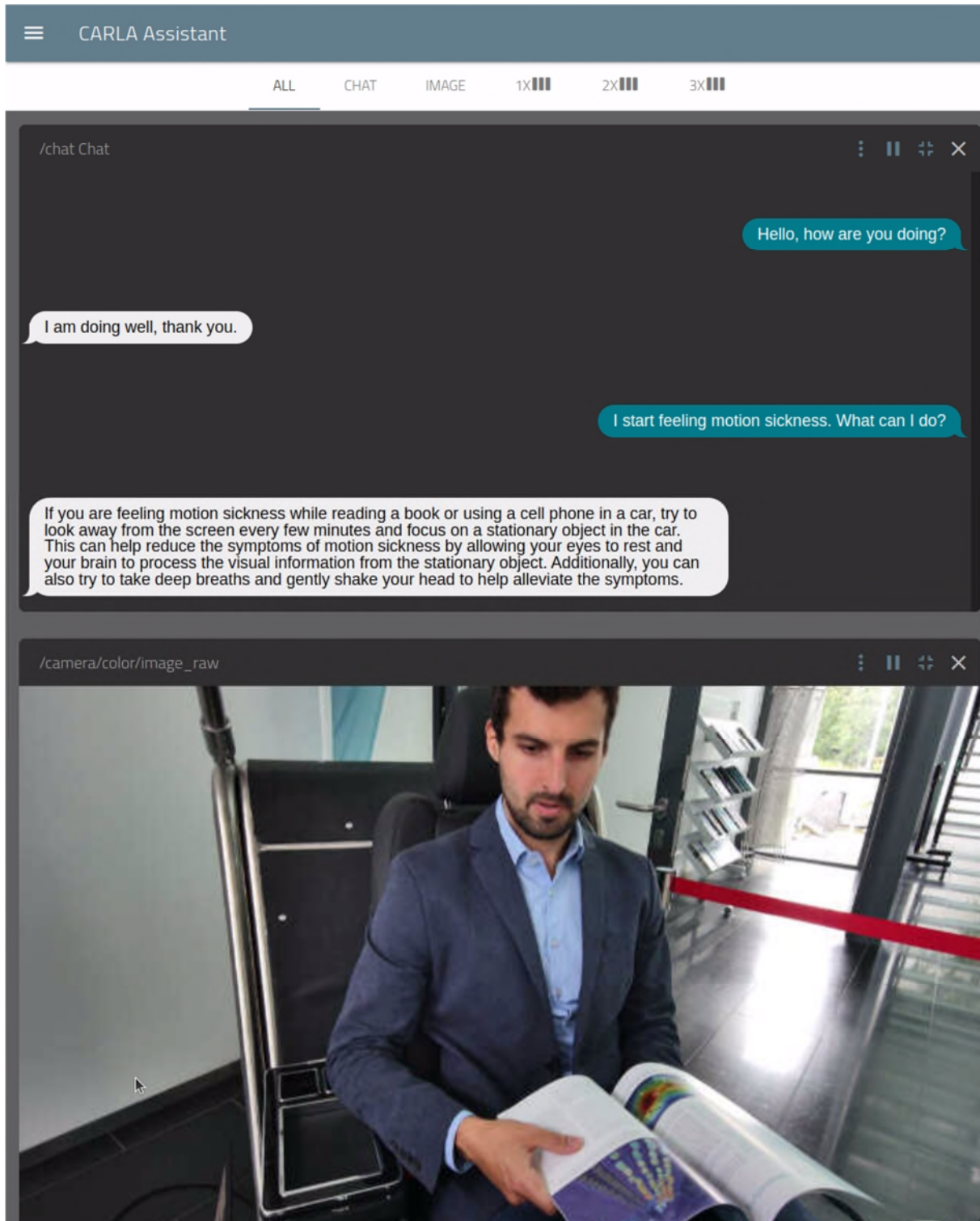


Abbildung 31: Screenshot des KI-Assistenten für den Fahrzeuginnenraum. Der Assistent erkennt Motion Sickness induzierende Aktivitäten und gibt Hilfestellungen basierend auf der aktuellen Aktivität.



Abbildung 32: Demonstrator des Motion Sickness Assistenten auf der KARLI Abschlusspräsentation



Abbildung 33: Screenshot der Demonstration des KI-Assistenten auf der KARLI Abschlusspräsentation. Der Assistent erkennt die Motion Sickness induzierende Aktivität „lesen“ und warnt den Fahrer vor dem Motion Sickness induzierenden „Kreisverkehr“.

2.1.2.8 AP-Koordination, Bereitstellung von Verfahren

Im Rahmen der Applikationen LKV und MS wurden jeweils in einem vierwöchigen Rhythmus Regelmeetings durchgeführt, in deren Verlauf die jeweiligen Projektpartner ihre aktuellen Arbeitsstände präsentierten und auf Basis dessen die nächsten Projektschritte gemeinsam festlegten. Des Weiteren wurden regelmäßig Koordinationsmeetings abgehalten, in deren Rahmen Algorithmen und Schnittstellen abgestimmt und definiert sowie der Entwicklungsfortschritt erörtert wurden. Des Weiteren wurden für das Modul UAP230 ebenfalls alle sechs Wochen Regelmeetings mit den Projektpartnern abgehalten, die mit der Datenerhebung und dem Labeling betraut waren. Ziel dieser Treffen war die Festlegung spezifischer Anforderungen an die Datenerhebungen und das Datenformat sowie die Entwicklung eines einheitlichen Labelingkonzepts.

Des Weiteren wurden alle sechs Monate Konsortialtreffen abgehalten, entweder in Form einer Online-Konferenz oder in Präsenz bei einem der Projektpartner. Die wissenschaftliche Koordination des Projekts wurde durch das Fraunhofer IOSB gewährleistet. Darüber hinaus stellte das Fraunhofer IOSB einen Fahrsimulator mit SILAB-Anbindung sowie eine speziell ausgestattete Sitzkiste für Demonstrationen zur Verfügung. In dieser Sitzkiste erfolgte die abschließende Evaluation der GenUI-Komponente.

2.2 Zahlenmäßiger Nachweis

2.2.1 Fraunhofer IAO

Der zahlenmäßige Nachweis liegt gesondert vor.

2.2.2 Fraunhofer IOSB

Der zahlenmäßige Nachweis liegt gesondert vor.

2.3 Notwendigkeit und Angemessenheit der geleisteten Arbeit

2.3.1 Fraunhofer IAO

KI-basierte Assistenzsysteme werden in Fahrzeugen immer präsenter und sind vor allem bei der Entwicklung automatisierter Fahrzeuge von enormer Bedeutung. Für den Entwicklungsprozess solcher Systeme ist es zum einen wichtig, möglichst reale und variantenreiche Datensätze zum Training der KI basierten Applikationen zu erheben. Zum anderen ist es wichtig, schon frühzeitig Evaluationsstudien mit den Nutzenden unter möglichst realen Bedingungen durchzuführen. Auf Grund der geringen Verfügbarkeit automatisierter Fahrzeuge zu Forschungszwecken ist es wichtig, eine möglichst immersive Testumgebung zu schaffen, um somit den Studienteilnehmenden den Eindruck zu verleihen mit einem realen System zu interagieren.

Wie die Ergebnisse der LKV- Datenerhebung zeigen, bringt vor allem eine iterative Datenerhebung im Realverkehr einen Mehrwert für die Trainingsdaten zur Entwicklung KI-basierter Systeme. Hierbei werden nicht nur verschiedene Wetter- und Verkehrsbedingungen abgedeckt, sondern auch Gewohnheiten der Nutzer erfasst, welche sich durch eine wiederholte Interaktion mit dem System ergeben. Diese Ergebnisse können in weiterführenden Datenerhebungen im Realverkehr genutzt werden, um die Qualität der Trainingsdaten zu erhöhen.

2.3.2 Fraunhofer IOSB

Die fortschreitende Fahrzeugautomatisierung bedingt die Präsenz von Fahrerassistenzsystemen, welche in Abhängigkeit der jeweiligen Situation Unterstützung bieten und eine sichere Mensch-Maschine-Interaktion gewährleisten. Die Kombination aus datenbasierten KI-Modellen und innovativer Sensorik dient der Entwicklung grundlegender Spezifikationen für zukünftige Fahrzeugarchitekturen sowie der Erschließung des Potenzials von Big Data aus Serienfahrzeugen. Die gezielte Förderung dieses Konsortiums stärkt die Innovationskraft von Unternehmen in Deutschland und erhöht die Sichtbarkeit von KMU im Bereich KI-basierter Fahrzeuginteraktion. Dies trägt zur Stärkung der Wettbewerbsposition des Standorts Deutschland in der globalen Automobilindustrie bei.

Das Fraunhofer IOSB leistet wesentliche Beiträge zu KARLI durch die Entwicklung KI-gestützter Fahrerüberwachungssysteme sowie der Integration von Video-Sprachmodellen (Video Language Models, VLM). Die genannten Technologien ermöglichen die semantische Erkennung und Analyse von Aktivitäten im Fahrzeuginnenraum, welche für eine levelkonforme Interaktion von essenzieller Bedeutung sind. Die präzise Insassenüberwachung und situationsangepasste Reaktionsmöglichkeiten, welche durch die Forschenden entwickelt wurden, bilden die Grundlage für eine adaptive MMI und unterstützen den sicheren Übergang zwischen verschiedenen Automationsleveln. Dabei werden Big-Data-Strukturen und skalierbare KI-Methoden eingesetzt, sodass der Einsatz der entwickelten Systeme auch in Serienfahrzeugen der nächsten Generation möglich ist. Die Umsetzung der Vision der adaptiven Interaktion ist dadurch gewährleistet.

2.4 Voraussichtlicher Nutzen, Verwertbarkeit der Ergebnisse

2.4.1 Fraunhofer IAO

Durch die beschriebenen Arbeitsinhalte und Ergebnisse konnte sich das Fraunhofer IAO intensiv mit den Themen der Datensammlung und -aufbereitung, mit dem Fokus auf Datensammlung im Realverkehr, auseinandersetzen. Das im Rahmen der Projektlaufzeit angepasste WoZ Fahrzeug ermöglicht eine Datenerhebung in einer hoch immersiven Versuchsumgebung, was auch für nachfolgende Projekte genutzt werden kann, um Fragestellungen zur Mensch-Technik Interaktion in automatisierten Fahrzeugen zu untersuchen. Des Weiteren können Datenerhebungen für die Entwicklung von KI-basierten Systemen im Realverkehr durchgeführt werden. Der Umbau des Fahrzeugs, sowie die Definition einer Datenerhebungspipeline und die Integration geeigneter Soft- und Hardware ermöglicht es außerdem, das Fahrzeug an die jeweilige Fragestellung anzupassen und externe Sensorik zu integrieren.

Für die Ergebnisverbreitung wirkte das Fraunhofer IAO im Rahmen von KARLI an insgesamt 8 Veröffentlichungen mit, die auf wissenschaftlichen Konferenzen präsentiert wurden. Die im Rahmen von Studien erhobenen Daten fließen außerdem in einer Doktorarbeit ein, die in Kooperation mit der Universität Stuttgart erfolgt.

2.4.2 Fraunhofer IOSB

Die im Rahmen des Projekts erzielten Ergebnisse flossen bereits während der Projektlaufzeit in verschiedene Industrieprojekte ein. Es ist zu erwarten, dass sich dieser Trend fortsetzt, da durch die Verabschiedung von Vorschriften für automatisierte Fahrfunktionen die Erfassung des Innenraums im Fahrzeug verpflichtend wird. Die Fortschritte in der robusteren Erfassung der 3D-Körperposen sowie in der Aktivitätserkennung – sowohl zur Erkennung levelkonformen Verhaltens

als auch von Aktivitäten, die Motion Sickness induzieren können – tragen dazu bei, die Expertise des Fraunhofer IOSB am Markt zu etablieren. Auch die Fortschritte bei der Entwicklung generativer Benutzeroberflächen mithilfe von Vision-Language-Modellen tragen zu dieser Entwicklung bei. Da die Bewältigung dieser Herausforderungen noch nicht abgeschlossen ist, plant das Fraunhofer IOSB, diese Arbeiten in weiteren Forschungsprojekten zu verwerten und fortzusetzen.

Die wissenschaftlichen Ergebnisse wurden bereits während der Projektlaufzeit in Form von Veröffentlichungen und Vorträgen genutzt. Zudem fließen sie direkt in Dissertationen ein, die im Rahmen des Projekts fortgeführt werden.

Die Verfahren zum unüberwachten Training der Aktivitätserkennung und zur Nutzung von Vision-Language-Models können jedoch nicht nur im Fahrzeuginnenraum eingesetzt werden. Am Fraunhofer IOSB finden ähnliche Systeme beispielsweise Anwendung in der manuellen Montage oder in der Assistenz in der Pflege. Auch hier sind weitere Projekte, insbesondere zur Unterstützung von Pflege- oder Klinikpersonal, geplant.

Die gewonnenen Erkenntnisse bilden zudem die Grundlage für weitere Anträge auf öffentlich geförderte Projekte im Bereich Mobilität sowie in anderen Anwendungsgebieten, in denen ein Verständnis menschlicher Gesten bzw. moderne Mensch-Maschine-Interaktionen erforderlich sind.

2.5 Fortschritt bei anderen Stellen

Das spanische Forschungszentrum VICOMTECH hat einen umfangreichen multimodalen Datensatz zur Fahrerüberwachung (DMD) sowie ein Labeling-Tool veröffentlicht, das speziell zur Analyse von Aufmerksamkeit und Wachsamkeit im Fahrzeug entwickelt wurde (Ortega et al., 2020). Die bei der Erstellung des Datensatzes gewonnenen Erkenntnisse fließen direkt in das KARLI-Projekt ein und unterstützen die Entwicklung KI-gestützter Fahrerassistenzsysteme.

Parallel dazu nutzt das Fraunhofer IOSB neue Entwicklungen in der Sprachmodell-Technologie, insbesondere Video-Sprachmodelle (Video Language Models, VLM), die mit großen Sprachmodellen (Large Language Models, LLM) verbunden sind. Diese Technologie ermöglicht eine semantische Situationsanalyse im Fahrzeuginnenraum und wurde beim Arbeitsmeeting VI in Babenhausen den Projektpartnern vorgestellt. Das Fraunhofer IOSB setzt VLM ein, um die Erkennung von Aktivitäten im Innenraum zu optimieren und präzise Fahrer- und Passagierverhalten zu analysieren (Zhao et al., 2023; Lin et al., 2023; Zhu et al., 2023).

In Kooperation mit der Universität Stuttgart IAT forscht das Fraunhofer IAO zudem an der Nutzung von LLMs für kontextbasierte Sprachinteraktion im Fahrzeug sowie der Entwicklung von intelligenten Sprachassistenten unter Nutzung von generativer KI. LLMs bieten hier neue Potenziale u.a. für die synthetische Datenerzeugung, deren Qualität z.B. durch Human-in-the-Loop Verfahren sichergestellt werden (Yeh et al., 2024), sowie für die Entwicklung persönlicher LLM-basierter Assistenten, die in Zukunft neben der Datenverarbeitung auch logische Denkfähigkeiten mitbringen sollen, um als persönliche Assistenten zu fungieren (Li et al., 2024).

2.6 Erfolgte und geplante Veröffentlichungen

Mathis, L.-A., Bubeck, C. B., & Peissner, M. (2024). When to Approach the User: Investigating Suitable Context Factors for Proactive Voice Assistance in Automated Cars. In *Proceedings of Mensch und Computer 2024* (pp. 299-309).

Mathis, L. A., Günes, C., Entz, K., Lerch, D., Diederichs, F., & Widloither, H. (2024, July). Generating Proactive Suggestions based on the Context: User Evaluation of Large Language Model Outputs for In-Vehicle Voice Assistants. In *Proceedings of the 6th ACM Conference on Conversational User Interfaces* (pp. 1-7).

Mathis, L.-A., Piechnik, D., Bubeck, C. B., Layer, S., & Widloither, H. (2024). "Ist jetzt ein guter Zeitpunkt?": Proaktive Ansprachen durch einen Sprachassistenten bei fahrfremden Tätigkeiten im Realverkehr. In GfA (Hrsg.), *Arbeitswissenschaft in-the-loop. Mensch-Technologie-Integration und ihre Auswirkung auf Mensch, Arbeit und Arbeitsgestaltung*. GfA-Press.

Diederichs, F., Herrmanns, A., Lerch, D., Zhong, Z., Piechnik, D., Mathis, L. A., ... & Rausch, J. (2024, June). Activities that Correlate with Motion Sickness in Driving Cars—An International Online Survey. In *International Conference on Human-Computer Interaction* (pp. 3-12). Cham: Springer Nature Switzerland.

Mathis, L. A., Werner, K., & Schmidt, G. J. (2023). Proactive voice assistance in automated vehicles - Understanding when to engage the user. *Stuttgarter Symposium für Produktentwicklung SSP 2023*. Stuttgart.

Piechnik, D., Mathis, L. A., Lerch, D., Diederichs, F., & Widloither, H. (2023). Technical setup of a Wizard-of-Oz vehicle for on-road AI data collection. *ITS European Congress 2023*. Lisbon.

Mathis, L. A., Werner, K., & Widloither, H. (2023). Exploring use cases and user perception of a proactive voice assistant in automated vehicles. *Human Factors in Transportation. AHFE International. AHFE International*.

Diederichs, F., Wannemacher, C., Faller, F., Mikolajewski, M., Martin, M., Voit, M., ... & Piechnik, D. (2022, June). Artificial intelligence for adaptive, responsive, and level-compliant interaction in the vehicle of the future (KARLI). In *International Conference on Human-Computer Interaction* (pp. 164-171). Cham: Springer International Publishing.

Lerch, D., El Bachiri, Y., Martin, M., Diederichs, F., Stiefelhagen, R. (2024). 3D Skeleton-based Driver Activity Recognition using Self-supervised Learning. *IEEE Intelligent Transportation Systems Conference (ITSC)*.

Bhardwaj, S., Lerch, D., Martin, M., Diederichs, F., Stiefelhagen, R. (2024). Self-supervised Driver Distraction Detection for Imbalanced Datasets. Eingereicht bei: *IEEE/CVF Winter Conference on Applications of Computer Vision 2025*.

VII Literaturverzeichnis

- Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., & Joulin, A. (2021). Emerging properties in self-supervised vision transformers. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 9650-9660).
- Blickshift (2021). <https://www.blickshift.com/products-services/blickshift-recorder/?lang=de>, aufgerufen am 28.10.2024
- Carsten, O., Kircher, K., Jamson, S. (2013). Vehicle-based studies of driving in the real world: the hard truth? *Accident; analysis and prevention* 58, 162–174.
- Cheng, Y.-B., Chen, X., Chen, J., Wei, P., Zhang, D., & Lin, L. (2021). Hierarchical transformer: Unsupervised representation learning for skeleton-based human action recognition. 2021 IEEE International Conference on Multimedia and Expo (ICME).
- Chu, Xiangxiang, et al. "MobileVlm v2: Faster and stronger baseline for vision language model." *arXiv preprint arXiv:2402.03766* (2024).
- Diederichs, F., Knauss, A., Wilbrink, M., Lilis, Y., Chrysochoou, E., Anund, A., ... & Bischoff, S. (2020). Adaptive transitions for automation in cars, trucks, buses and motorcycles. *IET Intelligent Transport Systems*, 14(8), 889-899.
- Du, Y., Fu, Y., & Wang, L. (2015). Skeleton based action recognition with convolutional neural network. In *IAPR Asian Conference on Pattern Recognition*.
- Eriksson, A., & Stanton, N. A. (2017). Driving performance after self-regulated control transitions in highly automated vehicles. *Human factors*, 59(8), 1233-1248.
- Fraser, N. M., Gilbert, G. N. (1991). Simulating speech systems. *Computer Speech & Language* 5(1), 81–99.
- Habibovic, A., Andersson, J., Nilsson, M., Lundgren, V. Malmsten, Nilsson, J. (2016). Evaluating interactions with non-existing automated vehicles: three Wizard of Oz approaches. 2016 IEEE Intelligent Vehicles Symposium (IV). p. 32–37.
- Kim, B., Chang, H. J., Kim, J., & Choi, J. Y. (2022). Global-local motion transformer for unsupervised skeleton-based action learning. In *Computer Vision – ECCV 2022*.
- Kiss, M., Schmidt, G., Babbel, E. (2008). Das Wizard of Oz Fahrzeug: rapid prototyping und usability Testing von zukünftigen Fahrerassistenzsystemen. 3. Tagung Aktive Sicherheit durch Fahrerassistenz.
- Li, N., Jain, J. J., Busso, C. (2013). Modeling of Driver Behavior in Real World Scenarios Using Multiple Noninvasive Sensors. *IEEE Transactions on Multimedia* 15 (5), 1213–1225.
- Li, Y., Wen, H., Wang, W., Li, X., Yuan, Y., Liu, G., Liu, J., Xu, W., Wang, X., Sun, Y., Kong, R., Wang, Y., Geng, H., Luan, J., Jin, X., Ye, Z., Xiong, G., Zhang, F., Li, X., . . . Liu, Y. (2024,

January 10). *Personal LLM Agents: Insights and Survey about the Capability, Efficiency and Security*. <http://arxiv.org/pdf/2401.05459v2>

Lin, B., Zhu, B., Ye, Y., Ning, M., Jin, P., & Yuan, L. (2023). Video-LLaVA: Learning United Visual Representation by Alignment Before Projection. *arXiv Preprint arXiv:2311.10122*.

Lu, Z., & de Winter, J. C. (2015). A review and framework of control authority transitions in automated driving. *Procedia Manufacturing*, 3, 2510-2517

Martin, M., Voit, M., & Stiefelhagen, R. (2020). Dynamic interaction graphs for driver activity recognition. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*.

Ortega, J. D., Kose, N., Cañas, P., Chao, M. A., Unnervik, A., Nieto, M., ... & Salgado, L. (2020, August). Dmd: A large-scale multi-modal driver monitoring dataset for attention and alertness analysis. In *European Conference on Computer Vision* (pp. 387-405). Springer, Cham.

Paravan (2022). <https://www.schaeffler-bywire.com/de/space-drive-2/>, aufgerufen am 28.10.2024.

Qiu, H., Hou, B., Ren, B., & Zhang, X. (2022). Spatiotemporal tuples transformer for skeleton-based action recognition. *arXiv preprint arXiv:2201.02849*.

Rao, H., Xu, S., Hu, X., Cheng, J., & Hu, B. (2021). Augmented skeleton based contrastive action learning with momentum lstm for unsupervised action recognition. *Information Sciences*, 569.

Shi, L., Zhang, Y., Cheng, J., & Lu, H. (2019). Two-stream adaptive graph convolutional networks for skeleton-based action recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*.

Shi, E., Gasser, T. M., Seeck, A., Auerswald, R. (2020). The Principles of Operation Framework: A Comprehensive Classification Concept for Automated Driving Functions. *SAE International Journal of Connected and Automated Vehicles*, Volume 3 (1)

Song, S., Lan, C., Xing, J., Zeng, W., & Liu, J. (2017). An end-to-end spatio-temporal attention model for human action recognition from skeleton data. *Proceedings of the AAAI Conference on Artificial Intelligence*.

Su, K., Liu, X., & Shlizerman, E. (2020). Predict & cluster: Unsupervised skeleton-based action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Wu, W., Hua, Y., Zheng, C., Wu, S., Chen, C., & Lu, A. (2022). SkeletonMAE: Spatial-temporal masked autoencoders for self-supervised skeleton action recognition.

Xu, S., Rao, H., Hu, X., Cheng, J., & Hu, B. (2023). Prototypical contrast and reverse prediction: Unsupervised skeleton-based action recognition. *IEEE Transactions on Multimedia*.

Yeh, C., Ren, D., Assogba, Y., Moritz, D., & Hohman, F. (2024). Exploring Empty Spaces: Human-in-the-Loop Data Augmentation. *arXiv preprint arXiv:2410.01088*.

Zhao, Y., Misra, I., Krähenbühl, P., & Girdhar, R. (2023). Learning Video Representations from Large Language Models. CVPR.

Zheng, N., Wen, J., Liu, R., Long, L., Dai, J., & Gong, Z. (2018). Unsupervised representation learning with long-term dynamics for skeleton based action recognition. Proceedings of the AAAI Conference on Artificial Intelligence.

Zhou, Y., Duan, H., Rao, A., Su, B., & Wang, J. (2023). Self-supervised action representation learning from partial spatio-temporal skeleton sequences. In Proceedings of the AAAI Conference on Artificial Intelligence.

Zhu, B., Lin, B., Ning, M., Yan, Y., Cui, J., Wang, H., ... Others. (2023). LanguageBind: Extending Video-Language Pretraining to N-modality by Language-based Semantic Alignment. *arXiv Preprint arXiv:2310.01852*.

Berichtsblatt

1. ISBN oder ISSN	2. Berichtsart (Schlussbericht oder Veröffentlichung) Schlussbericht
3. Titel Abschlussbericht „Künstliche Intelligenz für Adaptive, Responsive und Levelkonforme Interaktion im Fahrzeug der Zukunft (KARLI)“	
4. Autor(en) [Name(n), Vorname(n)] Widloirther, Harald Mathis, Lesley-Ann Diederichs, Frederik Lerch, David Martin, Manuel	5. Abschlussdatum des Vorhabens 30.09.2024
	6. Veröffentlichungsdatum
	7. Form der Publikation Bericht
8. Durchführende Institution(en) (Name, Adresse) Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO) Nobelstr. 12 70569 Stuttgart Fraunhofer Institut für Optronik, Systemtechnik und Bildauswertung (IOSB) Fraunhoferstr. 1 76131 Karlsruhe	9. Ber. Nr. Durchführende Institution
	10. Förderkennzeichen 19A21031F
	11. Seitenzahl 51
12. Fördernde Institution (Name, Adresse) Bundesministerium für Wirtschaft und Klimaschutz (BMWK) 53107 Bonn	13. Literaturangaben 31
	14. Tabellen 3
	15. Abbildungen 33
16. Zusätzliche Angaben	
17. Vorgelegt bei (Titel, Ort, Datum)	
18. Kurzfassung Das Projekt KARLI befasste sich unter anderem mit der Entwicklung von KI-basierten Funktionen zur Fahrerzustandserkennung für automatisiertes Fahren. Automatisierte Systeme ermöglichen es dem Nutzer die Fahrzeugsteuerung für vordefinierte Streckenabschnitte oder sogar während der gesamten Fahrt an das System zu übertragen. Dadurch kann sich der Nutzer fahrfremden Aktivitäten widmen und die Aufmerksamkeit muss, je nach Automatisierungsgrad, mehr oder weniger auf die Straße gerichtet sein. Aufgrund intransparenter Spezifikationen levelkonformen und nicht levelkonformen Verhaltens ist es für den Nutzer jedoch oft schwierig sich richtig zu verhalten. Im Rahmen der Applikation levelkonformes Verhalten lag der Fokus daher darauf, ein System zu entwickeln, welches den Nutzer (z.B. durch Aktivitätenerkennung) bestmöglich in verschiedenen Automatisierungsstufen unterstützt. Ein essenzieller Bestandteil bei der Entwicklung solcher KI-basierten Systeme sind möglichst variantenreiche und realitätsnahe Trainingsdaten. Der Fokus des Fraunhofer IAO lag daher auf der Durchführung einer iterativen Datensammlung im Realverkehr. Hierfür wurde ein Wizard-of-Oz Fahrzeug aufgebaut, welches die Simulation von verschiedenen Automatisierungsstufen und Übergaben während der Fahrt ermöglicht. Durch die Ausstattung des Fahrzeugs mit verschiedener Messtechnik, unter anderem auch von anderen Projektpartnern, konnten variantenreiche Datensätze erhoben werden. Diese dienten als Trainingsdatensätze für die Entwicklung KI-basierter Assistenzsysteme im Fahrzeug zur Gewährleistung von levelkonformen Verhalten. Der Fokus des Fraunhofer IOSB lag auf der Weiterentwicklung des kamerabasierten Advanced Occupant Monitoring Systems. Es wurde ein System mit seriennaher Hardware aufgebaut und in drei unterschiedliche Realfahrzeuge integriert. Dadurch konnten generische Forschungsdaten im realen Fahrbetrieb bei Projektpartnern erhoben werden. Ein Realfahrdatensatz mit qualitativ hochwertigen Labels wurde erstellt. Auf dieser Grundlage entwickelte das Fraunhofer IOSB das Advanced Occupant Monitoring System weiter. Das Fraunhofer entwickelte ein neues Modul für die KI-Fusion mit anderen Sensoren der Partner.	
19. Schlagwörter Künstliche Intelligenz, Driver Monitoring, Wizard-of-Oz, Mensch-Fahrzeug Interaktion, Realfahrtstudien, levelkonformes Verhalten	
20. Verlag	21. Preis

Document Control Sheet

1. ISBN or ISSN	2. type of document (e.g. report, publication) report
3. title Report „Künstliche Intelligenz für Adaptive, Responsive und Levelkonforme Interaktion im Fahrzeug der Zukunft (KARLI)“	
4. author(s) (family name, first name(s)) Widloirther, Harald Mathis, Lesley-Ann Diederichs, Frederik Lerch, David Martin, Manuel	5. end of project 30.09.2024
	6. publication date
	7. form of publication
8. performing organization(s) (name, address) Fraunhofer Institut for Industrial Engineering (IAO) Nobelstr. 12 70569 Stuttgart Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB) Fraunhoferstr. 1 76131 Karlsruhe	9. originator's report no.
	10. reference no. 19A21031F
	11. no. of pages 51
12. sponsoring agency (name, address) Bundesministerium für Wirtschaft und Klimaschutz (BMWK) 53107 Bonn	13. no. of references 31
	14. no. of tables 3
	15. no. of figures 33
16. supplementary notes	
17. presented at (title, place, date)	
18. abstract The KARLI project focused, among other things, on the development of AI-based functions for driver state detection in automated driving. Automated systems allow users to transfer vehicle control to the system for predefined segments of the journey or even for the entire trip. This enables users to engage in non-driving activities, and their attention needs to be directed more or less towards the road, depending on the level of automation. However, due to opaque specifications of level-compliant and non-level-compliant behavior, it is often challenging for users to behave appropriately. Therefore, within the application of level-compliant behavior, the focus was on developing a system that best supports the user (e.g., through activity recognition) at various levels of automation. An essential component in the development of such AI-based systems is a diverse and realistic training dataset. The focus of Fraunhofer IAO was therefore on conducting iterative data collection in real traffic. For this purpose, a Wizard-of-Oz vehicle was set up, which enables the simulation of different automation levels and take-overs during the drive. By equipping the vehicle with various measurement technologies, including those from other project partners, diverse datasets could be collected. These served as training datasets for the development of AI-based assistance systems in the vehicle to ensure level-compliant behavior. Fraunhofer IOSB focused on the further development of the camera-based Advanced Occupant Monitoring System. A system with close-to-production hardware was set up and integrated into three different real vehicles. This enabled generic research data to be collected from project partners in real driving conditions. A real driving data set with high-quality labels was created. On this basis, Fraunhofer IOSB further developed the Advanced Occupant Monitoring System. Fraunhofer developed a new module for AI fusion with other sensors from the partners.	
19. keywords Artificial intelligence, Driver monitoring, Wizard-of-Oz, human-vehicle interaction, on road data collection, level compliant behavior	
20. publisher	21. price