

## ADEX

Autonomous-Driving Examiner

<b>Programm / Ausschreibung</b>	IKT der Zukunft, IKT der Zukunft, IKT der Zukunft - 8. Ausschreibung (2019)	<b>Status</b>	abgeschlossen
<b>Projektstart</b>	01.11.2020	<b>Projektende</b>	29.02.2024
<b>Zeitraum</b>	2020 - 2024	<b>Projektlaufzeit</b>	40 Monate
<b>Keywords</b>	autonomous driving, verification and validation, testing, machine learning, reinforcement learning		

### Projektbeschreibung

Intelligente Mobilität ist eine der großen technologischen Herausforderungen im Kontext von KI und Internet der Dinge. Unsere Gesellschaft verspricht sich vom Autonomen Fahrer (hier verwendet für Systeme zur Steuerung hochautomatisierter Fahrzeuge), die Anzahl an Verkehrstoten und die Umweltbelastung drastisch zu reduzieren. Im letzten Jahrzehnt wurden bereits viele Fortschritte erzielt, z.B. ersichtlich an Fahrzeugen der Firma Tesla, die im Autopilot-Modus auf verschiedenen Straßen innerhalb der USA fahren. Jedoch zeigen Unfälle eben dieser Autos deutlich Schwachstellen des derzeitigen Stands der Technik auf. Anerkannte Grundsätze für das Design fehlertoleranter Systeme scheinen derzeit nicht ausreichend berücksichtigt zu werden; Zertifizieren beispielsweise der Objekterkennung gilt als schwierig, und die Psychologie von Fußgängern und Autofahrern im gemischten Verkehr zusammen mit autonomen Autos ist noch weitgehend ungeklärt. Werden wir jemals in der Lage sein, autonomen Fahrern voll zu vertrauen? Was ist nötig, damit wir unsere Kinder sorgenfrei einem hochautomatisierten Fahrzeug anvertrauen können, das sie unbeaufsichtigt durch die Stadt transportiert? Menschen absolvieren eine Führerscheinprüfung, in der sie Prüfern ihr Können demonstrieren müssen, um so Erlaubnis für selbstständiges Fahren zu erlangen. Analog dazu schlagen wir in ADEX eine umfassende, das menschliche Verhalten berücksichtigende Fahrprüfung für autonome Fahrer mit multidisziplinärem Ansatz vor, welche designagnostisch (d.h. unabhängig vom Konzept des autonomen Fahrers) dessen Schwachstellen und versteckte Fehler aufdeckt. Diese Fahrprüfung arbeitet mit einer großen Menge realistisch simulierbarer Verkehrssituationen („scenarios“), mit psychologiebasierten Verhaltensmodellen für Fußgänger und Fahrer, und mit einer Vielzahl an Wetter- und Straßenbedingungen. Analyse und Test des Verhaltens des autonomen Fahrers in übliche Verkehrssituationen spielt eine wichtige Rolle; besonderes Augenmerk wird auf Sonderfälle („edge-cases“) gelegt, welche im normalen Verkehrsalltag selten auftreten, den autonomen Fahrer aber in Grenzbereiche bringen, um dadurch zuvor unerkannte Probleme zu entdecken. Aus der österreichischen Unfalldatenbank und systematischer Unfallforschung werden zusätzlich realistische, kritische Test-Szenarien synthetisch generiert. Oft werden Unfälle durch bestimmte technische und menschliche (psychologische und physiologische) Faktoren verursacht; dieses Zusammentreffen wird im Projekt mithilfe holistischer Unfallforschung bei besonderen Häufungen („hot spots“) untersucht. Das Verhalten des getesteten autonomen Fahrers wird detailliert bewertet. Daraus abgeleitete „rewards“ ermöglichen es reinforcement-learning Methoden, vermehrt Szenarien zu generieren, die Schwachstellen (betreffend Sicherheit, Reglement, Komfort) des autonomen Fahrers zu Tage treten lassen. Gleich einer

Führerscheinprüfungen soll diese Fahrprüfung für autonome Fahrer unserer Gesellschaft ermöglichen, Vertrauen in die Zuverlässigkeit der Systeme zu fassen und sie auf unseren Straßen zu akzeptieren. Das interdisziplinäre Konsortium umfasst Experten zu KI, Sicherheit, fehlertolerantes Design, CPS, menschliche Psychologie in Verkehrssituationen und Standardisierung. Ein unabhängiges Beratergremium (IAB) aus Experten für rechtliche und ethische Aspekte der KI, der virtuellen autonomen Fahrsimulation und Vertretern von Führerscheinschulen sowie von Regulierungs- und Zertifizierungsbehörden hilft in der Projektsteuerung.

## **Abstract**

Smart mobility is one of the grand technological challenges of cyber-physical systems in the era of artificial intelligence and the Internet-of-Things. More specifically, autonomous-driving controllers are a major priority of our modern society with the great promise of dramatically reducing human fatalities and to decrease pollution. While many advances have already been achieved over the past decade, such as Tesla's cars driving in Autopilot mode on various roads within the US, the very accidents of some of these cars clearly expose the vulnerabilities of the current state-of-the-art. In particular, highly automated vehicles appear to ignore some of accepted principles for the design of fault-tolerant systems; certification of object-recognition techniques is causing major concerns, and the psychology of pedestrians and car drivers in mixed-traffic together with autonomous cars is largely unexplored.

Will we ever be able to fully trust autonomous-driving controllers? What is needed so we can light-heartedly entrust our children to such highly automated vehicle, having them shuttled unattended across town?

Human drivers have to prove their abilities to a driving examiner, before being allowed to drive autonomously. Following this analogy, we propose in ADEX a multidisciplinary approach to the design and synthesis of a comprehensive, human-centric, autonomous-driving examination, which is agnostic to the particular design (i.e., independent of its underlying concept) of an autonomous-driving controller, to expose its vulnerabilities or unknown flaws. The driving examination will consist of a large set of realistically simulated traffic situations, called scenarios, with psychology-based human behavioral models for pedestrians and drivers, and various weather and road conditions. While the analysis and testing of the behavior of the autonomous-driving controller in normal traffic conditions will play a very important role, a special emphasis will be put on edge cases rarely occurring in normal traffic, to allow uncovering hidden issues. For the autonomous-driver such edge cases might represent system limiting situations. Analysis of real-world traffic accidents will be used to synthetically generate new, realistic and critical traffic scenarios for testing. Causes of accidents are often resulting from the interaction of technical and human (psychological and physiological) factors and both will be studied using holistic accident analysis on hot spots. Actions of the autonomous-driving controller will be thoroughly quantified in form of rewards, that will be exploited by reinforcement-learning techniques, to more and more generate traffic situations in which the autonomous-driving controller would fail (towards safety, regulation and comfort criteria). As for the human-driving examination, if an autonomous-driving controller passes it, we as a society will achieve a greater confidence in the system's reliability and will be more inclined to accept it on our roads. The proposed work will be driven by an interdisciplinary consortium, with experts in artificial intelligence, safety, fault-tolerant design, cyber-physical systems, human psychology in traffic situations and standardization efforts in autonomous driving. The project directions will be steered, and the outcomes followed by an Independent Advisory Board (IAB), consisting of experts in legal and ethical aspects of artificial intelligence, virtual autonomous driving simulation and representatives from driving license schools as well as regulatory and certification bodies. ADEX aims at developing a trustworthy framework for automatically examining autonomous-driving controllers.

## **Projektkoordinator**

- AIT Austrian Institute of Technology GmbH

## **Projektpartner**

- Technische Universität Wien
- SV Univ.Prof.DI.Dr.Ernst Pflieger
- AVL List GmbH