Universität Stuttgart
Institut für Automatisierungstechnik und Softwaresysteme

Prof. Dr.-Ing. Dr. h. c. M. Weyrich

Bild: Robo-Test, TTI GmbH, Stuttgart

Hey K.I.T.T. – kann ich Dir trauen?

Validierung und Verifikation autonomer Systeme

Prof. Michael Weyrich

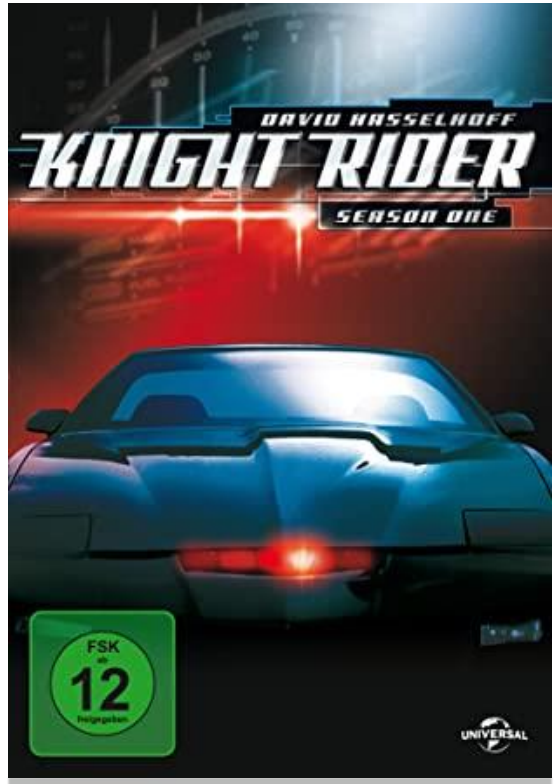
Univ. Stuttgart, Institut für Automatisierungstechnik und Softwaresysteme

Intelligentes Roboterauto „K.I.T.T.“
Science Fiction Fernsehserie „Knight Rider“ ab 1982

Autonome Fahrzeuge (mit Fahrer) mit widerstandsfähiger Karosserie gesteuert von einer AI.

K.I.T.T. und K.A.R.R. demonstrieren das Vertrauen in autonome Fahrzeuge und die Angst vor Täuschungen

- K.I.T.T. wurde entwickelt, um menschliches Leben zu schützen und demonstriert eine kooperative, vertrauensvolle Beziehung zu seinem Fahrer
- Gegenspieler K.A.R.R. ist auf Selbsterhaltung ausgerichtet



K.I.T.T. (“Knights Industries Two Thousand”, 1982 – 1986, Produzent Glen A. Larson; weitere Filme: 1991, 1997-1998

Bild: Universal Television

Inhalt

- **Einführung**
- **Evolutionsstufen in der Systementwicklung (von CI/CD und dem Data Loop)**
- **Validierung Autonomer Fahrfunktionen**
- **Aktuelle und zukünftige Arbeiten**

Inhalt

- **Einführung**
- Evolutionsstufen in der Systementwicklung (von CI/CD und dem Data Loop)
- Validierung Autonomer Fahrfunktionen
- Aktuelle und zukünftige Arbeiten

Beispiele vollautomatischer und autonomer Systeme

Autonome Systeme haben ein breites Anwendungsfeld, das von autonomen Fahr- und Logistiksystemen bis hin zu modernen Produktionssystemen und Robotern reicht.

Vollautomatisierte und autonome Fahrzeuge (SAE Level 4 / 5)

auto motor sport

10.10.2019, auto motor sport

Bild: Fiat Chrysler

UPDATE FÜR WAYMO-AUTOS

Sicherheitsfahrer bleiben zu Hause

Handelsblatt

24.08.2022

Bild: Waymo

AUTONOMES FAHREN

Waymo und Daimler Truck testen vollautomatische Lkw in Texas

WirtschaftsWoche

24.07.2022

Bild: PR

FIVE AI AUS CAMBRIDGE

Verhilft dieses Start-up dem autonomen Fahren zum Durchbruch?

Vergleichstest: Automatisierte Mähroboter

stern Bild: Stern.de

Autonomes Containerfahrzeug

BASF We create chemistry

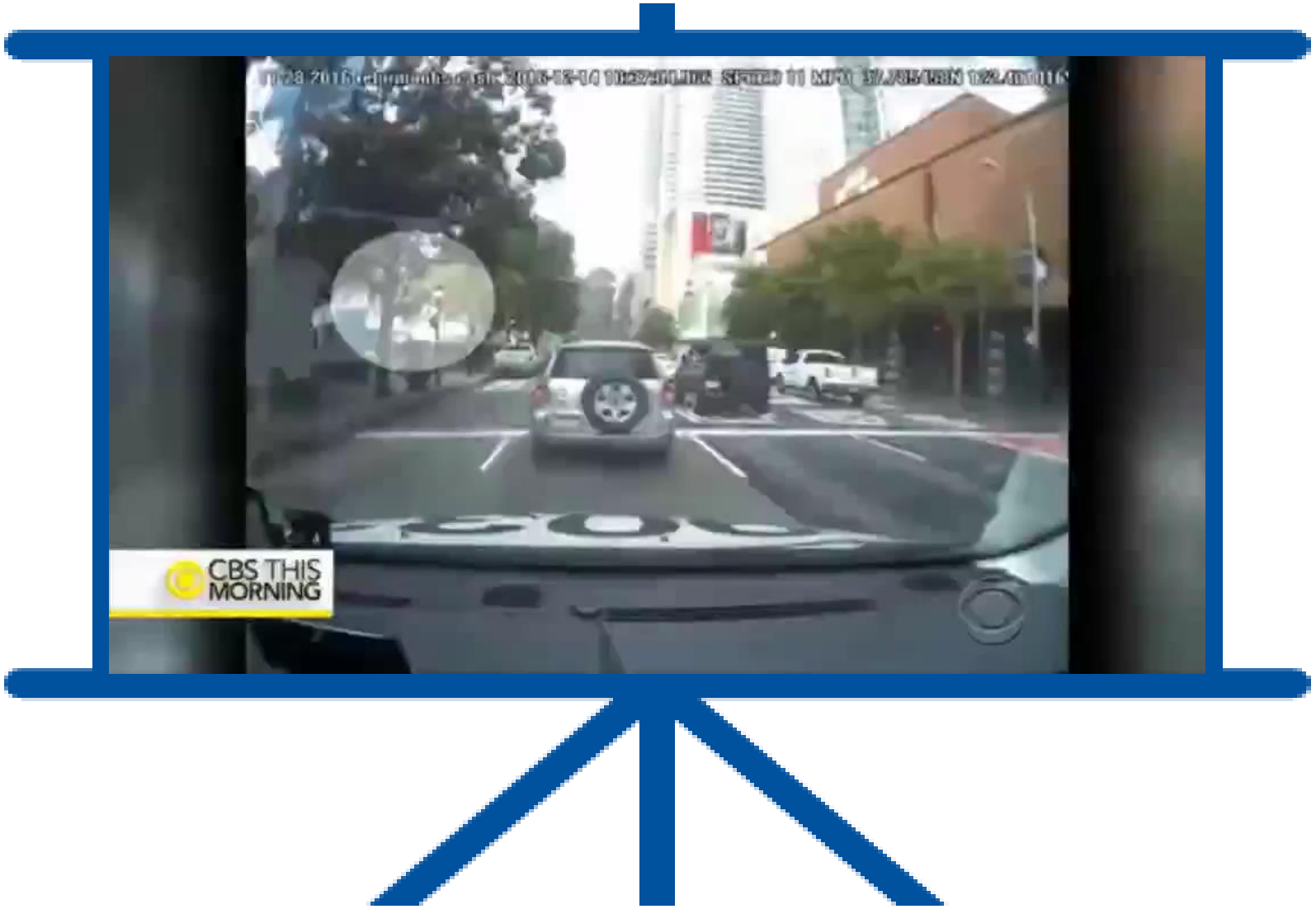
Bild: BASF, C. Schäfer

Landwirtschaft der Zukunft

JOHN DEERE

Bild: John Deere

Aber leider passieren Fehler …



Was ist Vertrauen und wie stellt man es her?

„Vertrauen ist die Fähigkeit, in einer Beziehung / Interaktion zu einer Person oder Institution zu treten trotz Ungewissheit und Unüberschaubarkeit."
Ch. Stückelberger, Prof. für Ethik, Basel, 2009"

Ethiker definieren Grundfragen:
[Stü2009]

- Wer vertraut wem oder was? (Subjekte, Objekte)
- Wie lange? (Zeit)
- Warum? (Ursache des Vertrauens)
- Mit welchen Mitteln? (Wie wird es erzielt?)
- Treffen Aussage wie: „Vertrauen braucht Misstrauen"

Charakterisierung der Vertrauenswürdigkeit von AI
(VDE SPEC 90012, April 2022; [Hal2020])

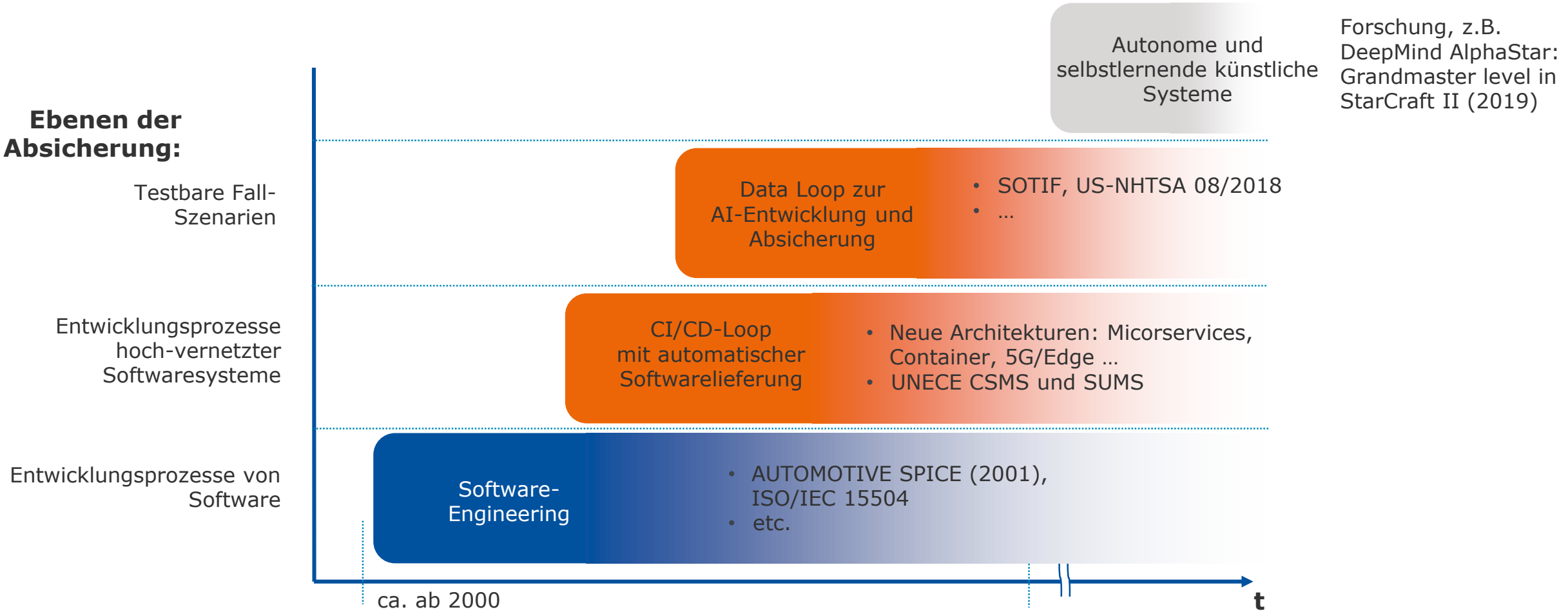
- Transparency (Dokumentation, Zugänglichkeit, Verständlichkeit)
- Accountability (Rechenschaft, Verantwortung, Haftung)
- Privacy (Verarbeitung und Schutz persönlicher Daten)
- Fairness (angemessene Metriken und Würdigungen, Nachhaltigkeit)
- Reliability (Robustheit und Zuverlässigkeit)
- Außerdem: Safety and Security (Sicherheit) [Dor2017]

Inhalt

- Einführung
- **Evolutionsstufen in der Systementwicklung (von CI/CD und dem Data Loop)**
- Validierung Autonomer Fahrfunktionen
- Aktuelle und zukünftige Arbeiten

Evolutionsstufen in der Systementwicklung

Neue Formen des Software-Engineering: Über Software-definierte Systeme mit „Continuous Integration and Deployment“ (CI/CD) hin zum „Data Loop“ für die AI-Systementwicklungen.

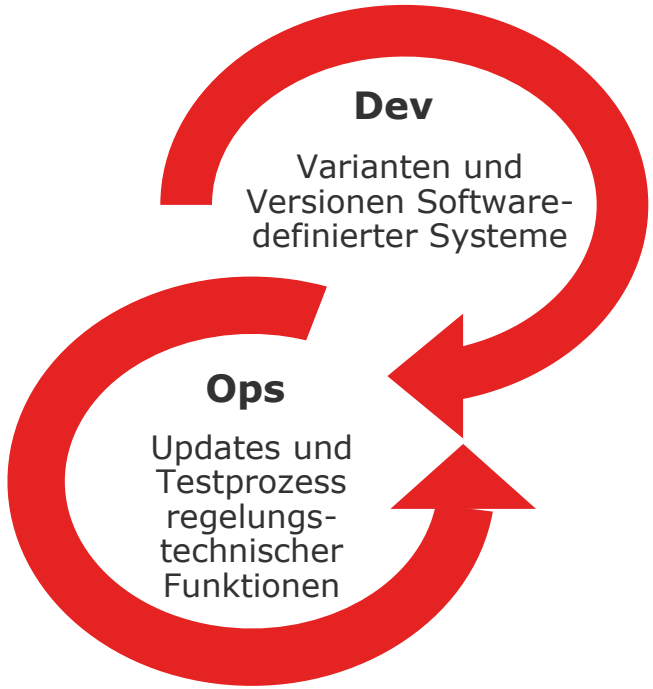


CI/CD-Loop mit automatischer Softwarelieferung

Die Verbindung während des Betriebs hochkonfigurierbarer softwaredefinierte Mobilitätssysteme im Betrieb schafft neue Möglichkeiten für die (Weiter-)Entwicklung.

„Online" Verbindung zu den Fahrzeugen im Betrieb:

Entwicklung neuer Funktionalitäten für das **Re-Deployment von Software** und einer sinnvollen **Einbeziehung des „Backends"**



Einsatz von standardisierten Testverfahren:

- UNECE SUMS No. 156 - Software update and software update management system
- UNECE CSMS No. 155 - Provisions for cyber security and cyber security management system
- Verfahren zum Umgang mit komplexen Software-Produktlinien
- ...

Vernetze Kommunikation und allgegenwärtige Sensorik

Die Kommunikation der Fahrzeuge untereinander mit der Infrastruktur und einem „Back-end“, stellt neue Funktionen bereit und wirft neue Fragen in Punkto Sicherheit auf.

- Sensoren im Fahrzeug erfassen vielfältige **personenbezogene Informationen**
- Sensoren außerhalb des Fahrzeugs **erfassen das Umfeld**
- Fahrzeuge und die Infrastruktur **tauschen Daten aus**
- Es besteht Informationsaustausch zwischen Fahrzeugen und einem **„Back-end“**

IAS-Teammitglieder in der Arena 2036 (Bild: Univ. Stuttgart, IAS)

5G Testfeld für hoch-vernetzte Fahrzeuge

Fahrzeuge und deren Komponenten können während des Betriebes im Feld miteinander, mit der Infrastruktur und mit den Entwicklungsabteilungen kommunizieren.

- **Sichere Vernetzung:** Edge und Cloud Kommunikation auf Internet-Protokollen zwischen Fahrzeugen und Verkehrsinfrastruktur
- **Management und zuverlässiges (Re-) Deployment** für variantenreiche Software mit Software-Produktlinien
- **Analyse und verlässliche Synchronisation von Daten** und Informationen mit dem Digitalen Zwilling
- **Remote Funktionserbringung (mit Collective Perception)** für das Fahrzeug aus dem Back-end

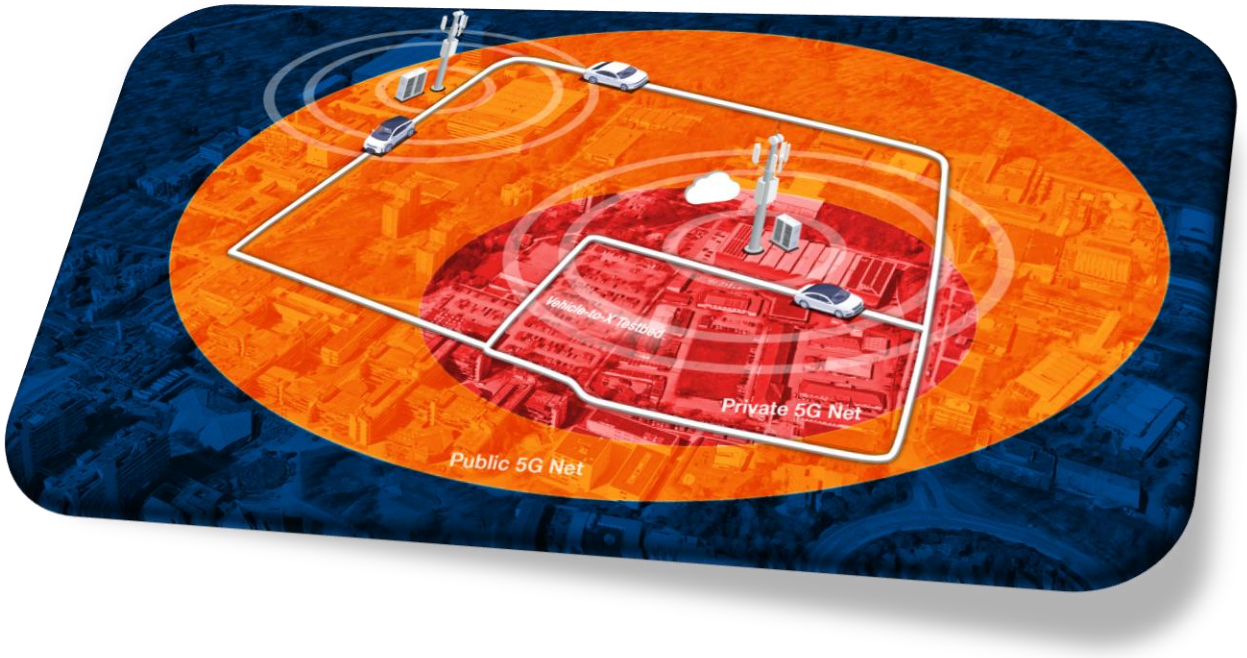


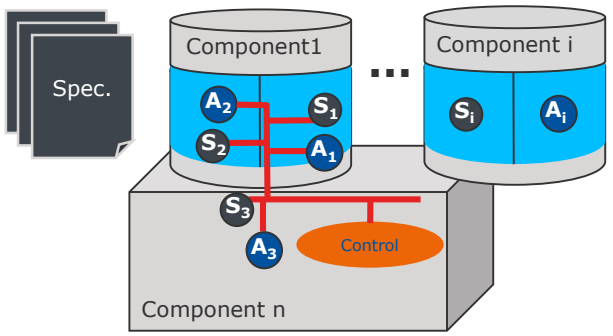
Bild: Univ. Stuttgart, IAS

5G Testfeld (Netz: private und public) mit Edge und Cloud am Campus der Univ. Stuttgart am Paffenwaldring.

Digitaler Zwilling zur Fehlererkennung im Zusammenspiel von Komponenten nach Softwareänderungen

Beispiel zur Verifikation software-definierter Funktionsänderungen (Diss. IAS [Zel2019]).

1 **Steuerungssoftware im Fahrzeug unterliegt mehr und mehr Änderungen**



Softwaresystem des Fahrzeugs

2 **Digitaler Zwilling der Softwarekomponenten dient der Analyse im Simulator**

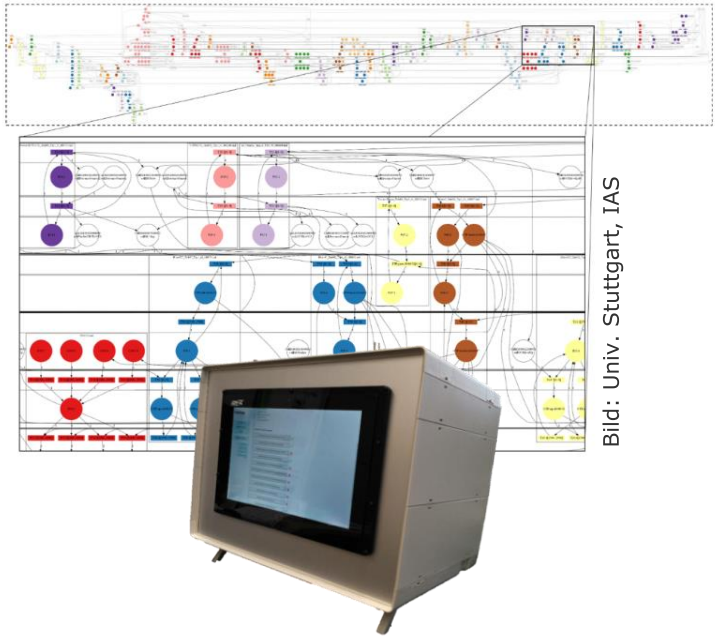
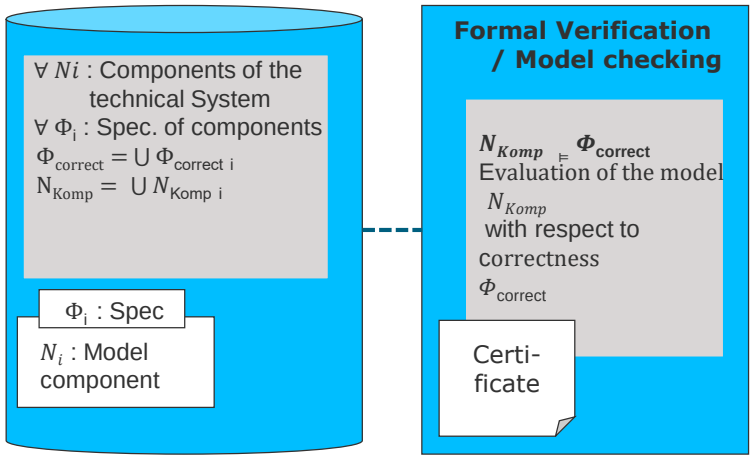


Bild: Univ. Stuttgart, IAS

Testgerät mit Modellen zur Verhaltens-simulation der Steuerungssoftware

3 **Fehler durch Auswirkungen von software-basierten Funktions-änderungen werden gefunden**



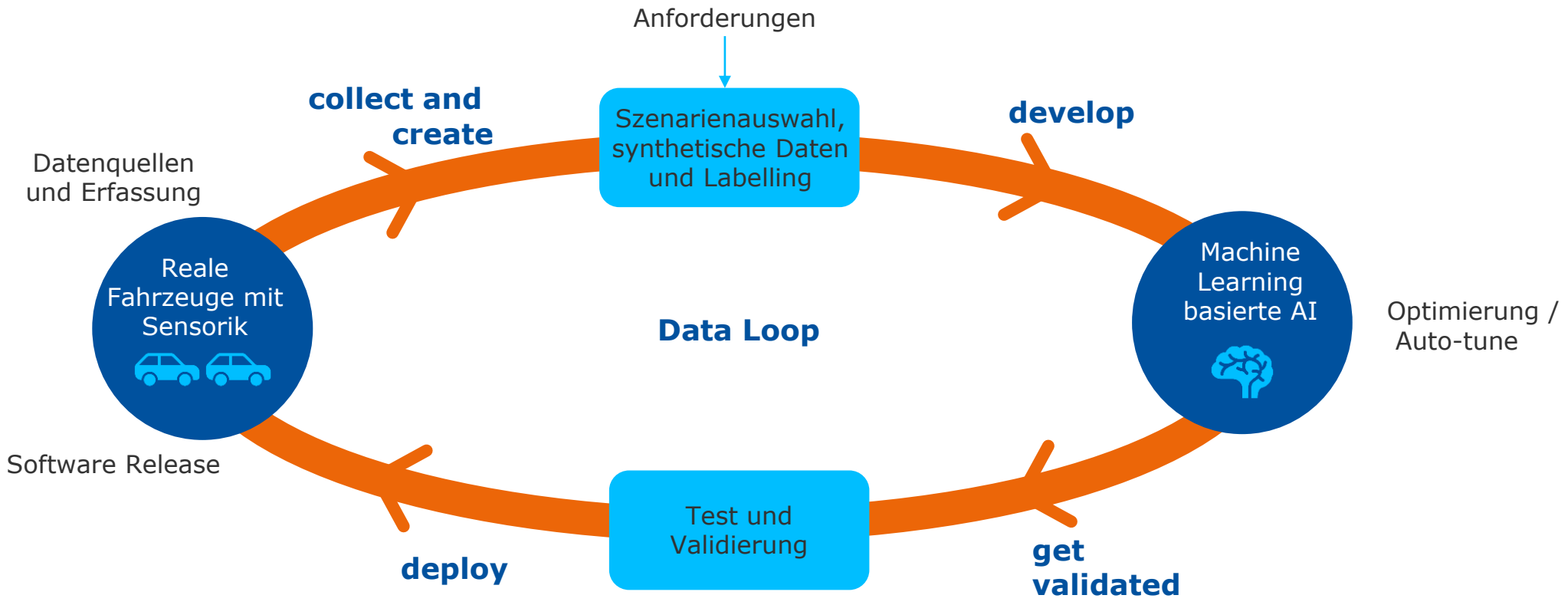
Modell-basierte Analyse und Verifikation

Inhalt

- Einführung
- Evolutionsstufen in der Systementwicklung (von CI/CD und dem Data Loop)
- **Validierung Autonomer Fahrfunktionen**
- Aktuelle und zukünftige Arbeiten

„Data Loop“ zur Entwicklung Autonomer Systeme

Bei der Entwicklung von Machine Learning Steuerungen, ist das gezielte Training mit realen und synthetischen Szenarien sowie ein Test zur Validierung entscheidend für die Qualität.



Wie sollen Versuche zur Validierung autonomer Fahrzeuge zukünftig durchgeführt werden?

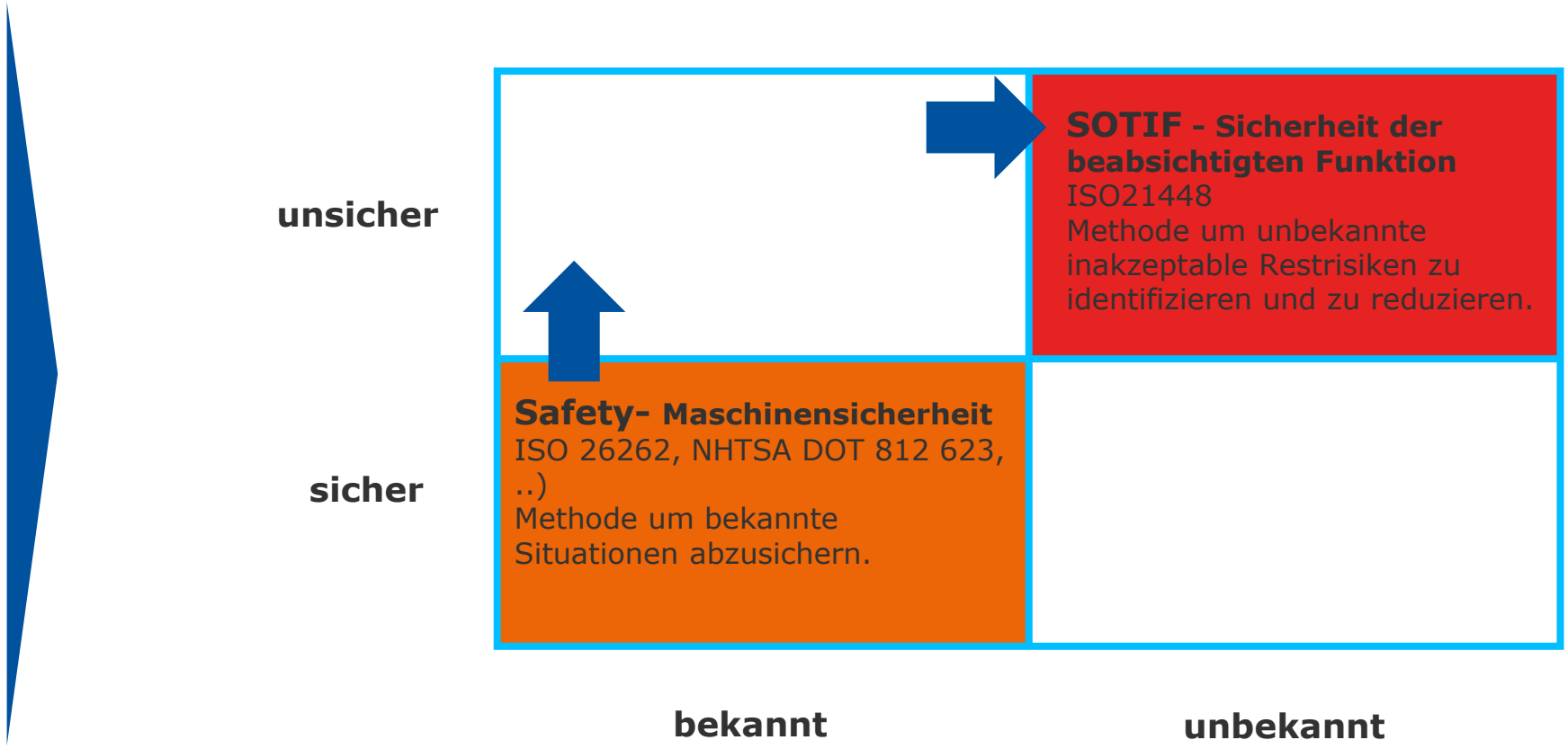
Normung, Standardisierung und Forschung sind gefordert.

Mobilität

Off-Road

Logistik

Bilder: Robo-Test



Wie viele Kilometer muss man fahren, um die sichere Funktion nachzuweisen?

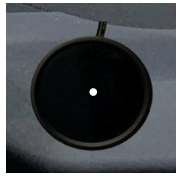
Statistische Überlegungen führen (zu)viele Testkilometer auf und lassen Fragen nach der Teststrecke unbeantwortet.

„Ein autonomes Fahrzeug müsste **<u>275 Millionen</u> redundante Meilen** ohne tödliche Unfälle zurücklegen, um ein mit der amerikanischen Unfallrate vergleichbares Verhalten zu erreichen“ [Kal2016]

Bild: **Robo-Test**

Detected car in safe distance

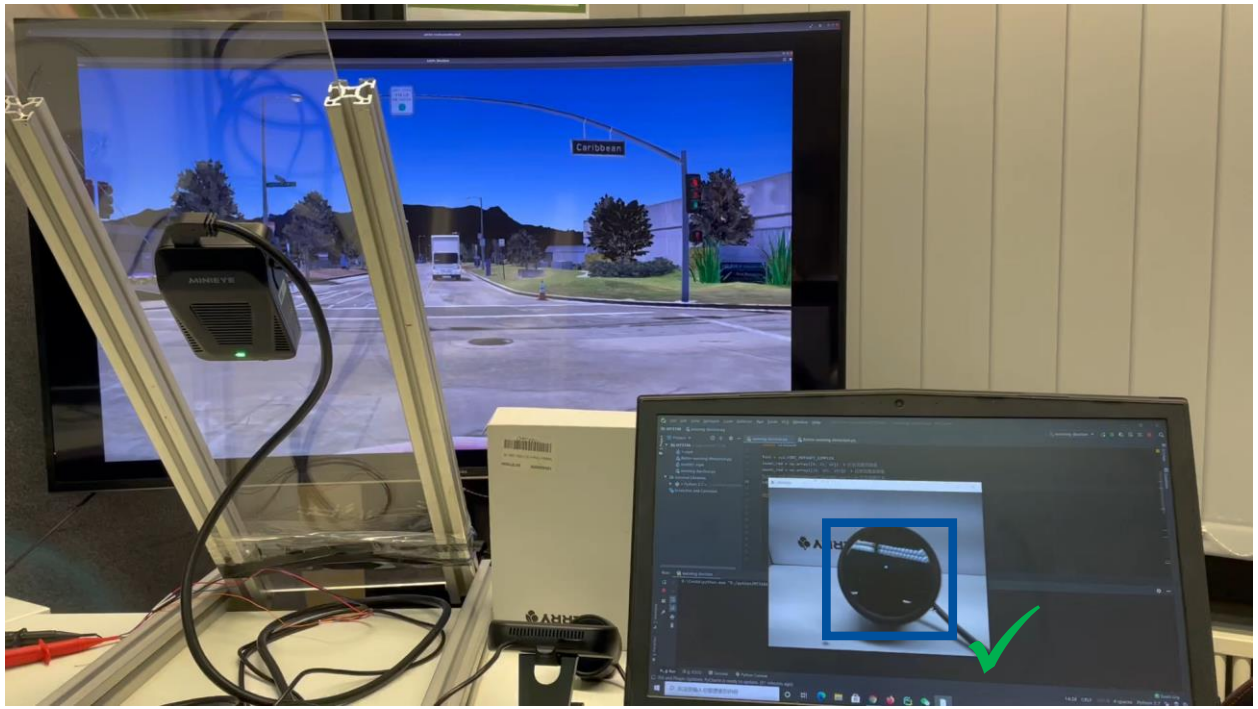
Detected in dangerous distance

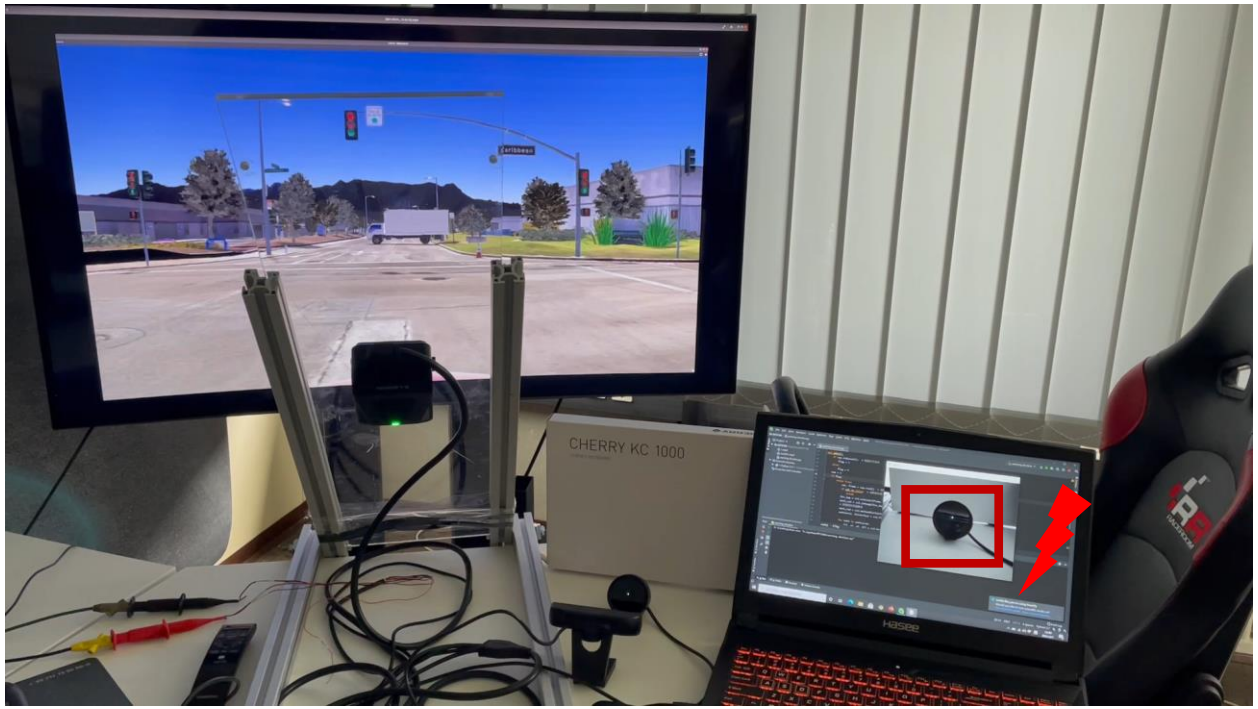


Detected nothing in front

Haben am Markt befindliche Fahrerassistenzsysteme Fehlfunktionen?

Ein Szenarientest mit synthetischen Daten - Start-up validiert Fahrerassistenzsysteme.





Bilder: Robo-Test

LKW wird erkannt

LKW wird **nicht** erkannt

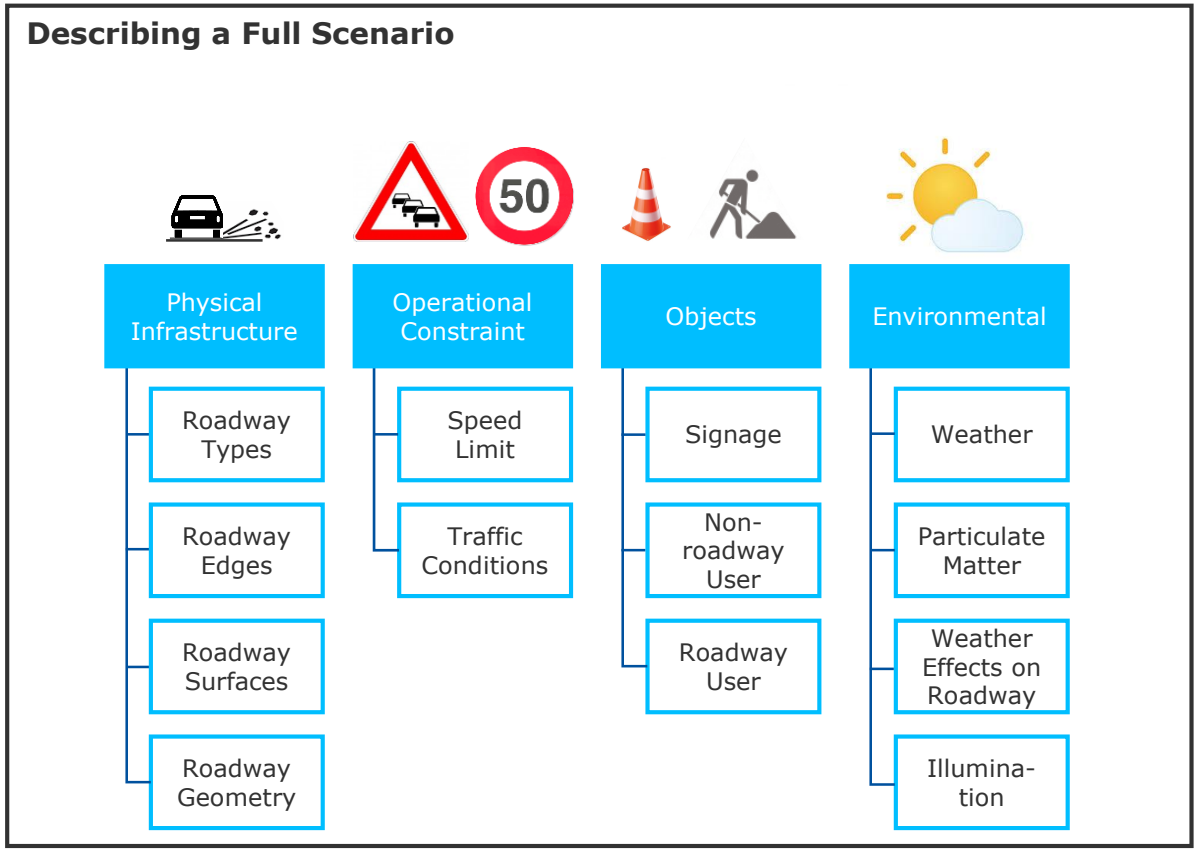
Benötigt werden vielfältige Validierungsszenarien

Für welche Szenarien sind die autonomen Funktionen auszulegen und abzusichern, damit diese möglichst ohne Einschränkungen sicher funktionieren.

Operational Design Domain (ODD)

ODDs definieren Szenarien über Zeitintervalle und geografische Regionen

- Es sollen **Umweltbedingungen** definiert werden, mit denen ein autonomes System umgehen können muss
- ODD-Szenarien sollen geeignete und schwierige **Kombinationen beschreiben**
- Das **ODD-Test-Management** muss in der Lage sein, viele Szenarien zu modellieren



ODD Topologie nach US NHTSA DOT HS 812 623

Automatische Generierung von Szenarien mit Ontologien

Dimensionen und Parameter von Szenarien werden in der Simulation variiert, wobei vorhandene Testfälle genutzt und auf Abdeckung und Kritikalität optimiert werden.

Szenario: **Plötzliches Hindernis**
Ein Mensch läuft vor das Fahrzeug.

Dimension	Begin	End
Vehicle Type	Car	Car
Vehicle Distance Diff.	20	20
Vehicle Lane	Sidewalk	Sidewalk
Ped. Gender	Male	Male
Ped. Age	12	12
Ped. Clothes color	White	White
Ped. Distance Diff.	22	22
Ped. Lane	Sidewalk	Same

Szenario: **Überhol-manöver**
Überholen bei Schnee.

Dimension	Begin	End
Urban	True	True
Number of Lanes	2	2
Vehicle Type	Car	Car
Vehicle Distance Diff.	-30	20
Vehicle Lane	Different	Same
Snow	70	70

Beispiel: Ein kognitives Testtool zur Verwaltung von Trainings- und Testfällen

Erzeugen von realen und synthetischen ODD-Szenarien hat große Bedeutung

Robo-Test Plattform
für ODD Training- und Validierungstest

- Einsatz **unterschiedlicher Simulationssysteme** (z.B. VTD, Unreal Engine) für Vision und Lidar
- Spezielle Simulation **GNSS**
- Einbeziehen von **Realdaten**
- Erzeugen von **Bibliotheken** mit Standard-Szenarien

Definition von **Szenarien Parametern**

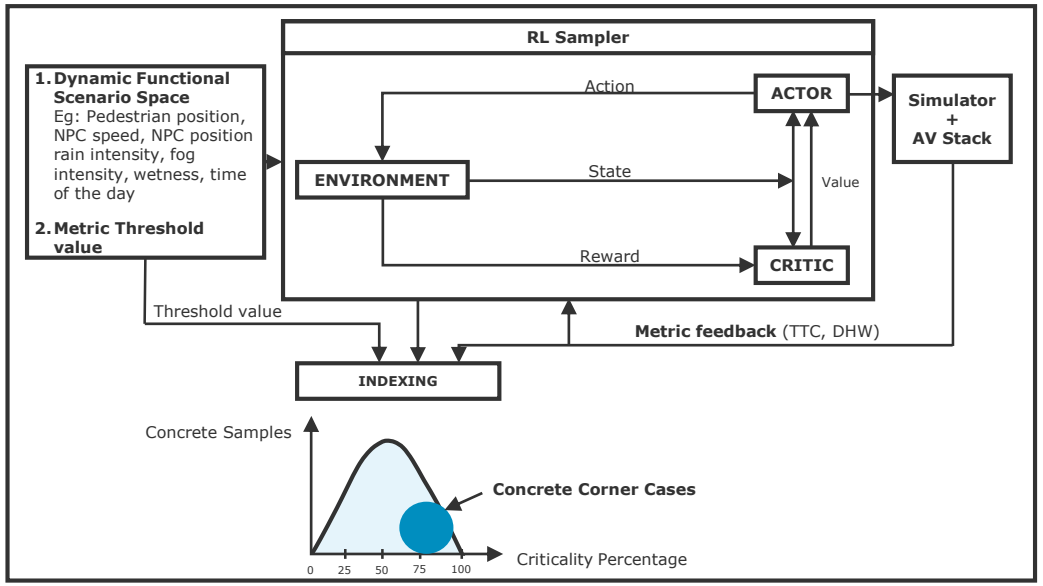
Ontologien zur systematischen Erzeugung bei der Generierung von Szenarien

Effizienz durch Tools zur Optimierung der „**Coverage**"

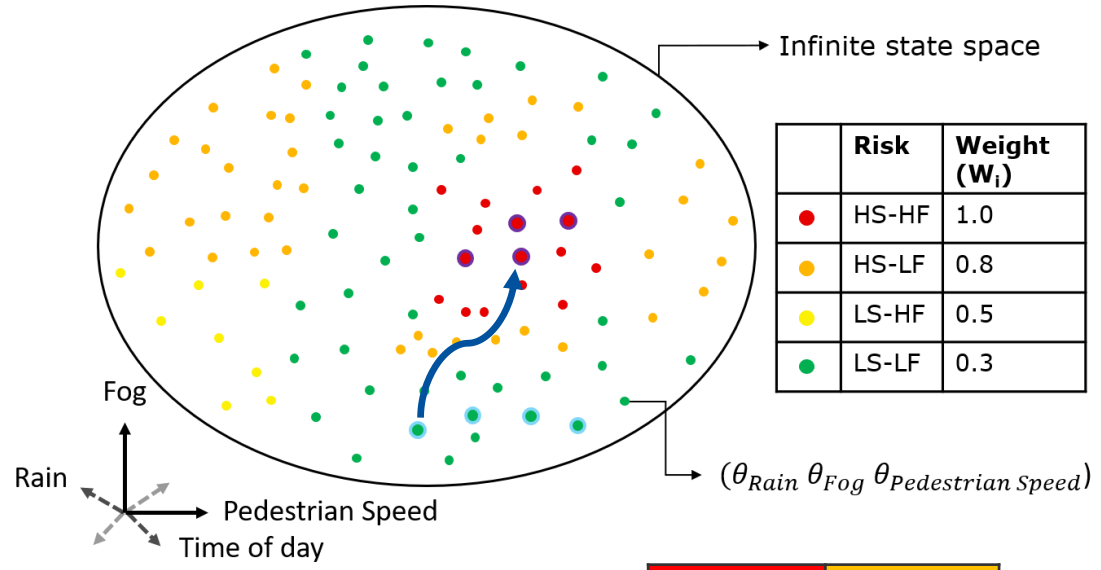
Methode:
Wie lassen sich die kritischen Einstellungen identifizieren?

In der Simulation können die kritischen Dimensionen und Parameter eines Szenarios beurteilt und repräsentative Fälle identifizieren werden.

Die Dimensionen der Szenarien werden mit **Re-enforcement Learning** durchlaufen, um die kritischen Bereichen abzudecken.



Die **Risiken der einzelnen Dimensionen und deren Parameter** in den Szenarien werden beurteilbar.



	Risk	Weight (W_i)
●	HS-HF	1.0
●	HS-LF	0.8
●	LS-HF	0.5
●	LS-LF	0.3

High frequency of occurrence **High** severity	**Low** frequency of occurrence **High** severity
High frequency of occurrence **Low** severity	**Low** frequency of occurrence **Low** severity

Methode:
Unwesentliche Hauptkomponenten werden weggelassen

Weitere Konzentration der Szenarien durch eine Principal Component Analysis (PCA), die auf die risikoreichen Testfälle aus dem Re-enforcement learnings aufsetzt.

Ansatz: Nicht notwendige Hauptkomponenten (PC) werden weggelassen.
Im Beispiel: „Nebel“ wird als besonders kritisch angesehen, hingegen hat „Regen“ wenig Einfluss und kann entfallen.

Specific Test Cases			Criticality
Fog	Pedestrian speed	Rain	
$\theta_{Fog,1}$	$\theta_{Speed,1}$	$\theta_{Rain,1}$	0,2
$\theta_{Fog,2}$	$\theta_{Speed,2}$	$\theta_{Rain,2}$	0,5
$\theta_{Fog,3}$	$\theta_{Speed,3}$	$\theta_{Rain,3}$	0,7
$\theta_{Fog,4}$	$\theta_{Speed,4}$	$\theta_{Rain,4}$	0,8

i

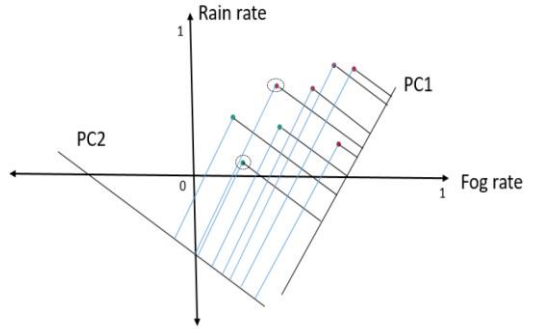
m

PCA

	PC1	PC2	PC3	
	0,3	0,2	0,2	
	0,6	0,4	0,3	...
	0,9	0,3	0,1	
	0,2	0,1	0,2	
Σ	**2,0**	**1,0**	**0,8**	

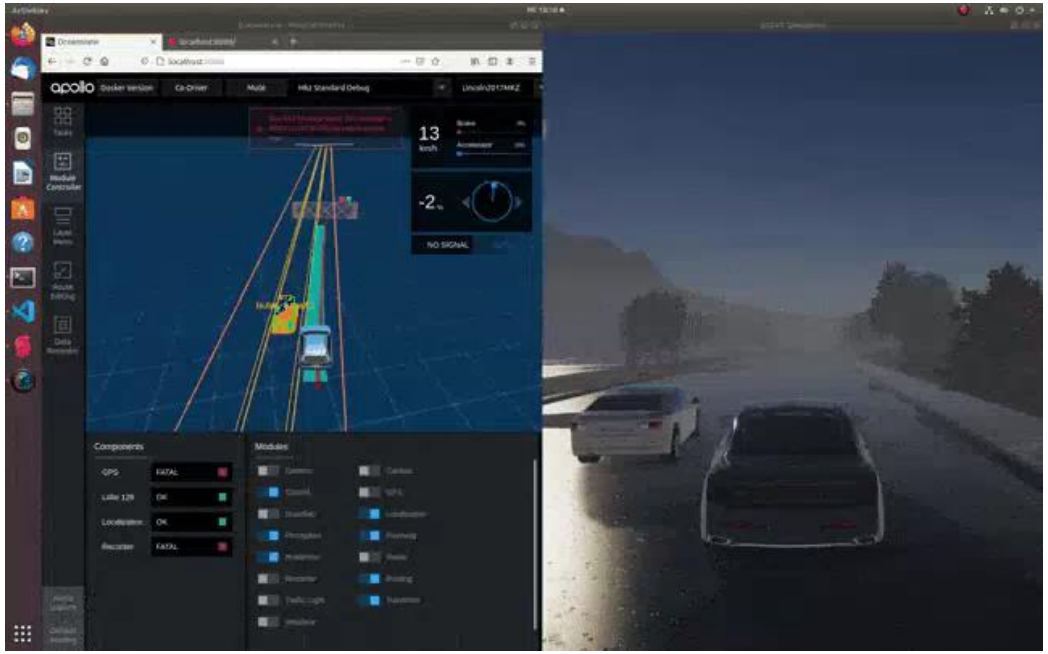
Inv. PCA

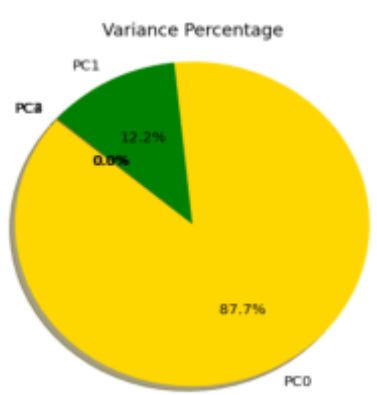
	Fog	Pedestrian speed	Rain
PC1	60%	30%	10%
PC2	14%	70%	16%



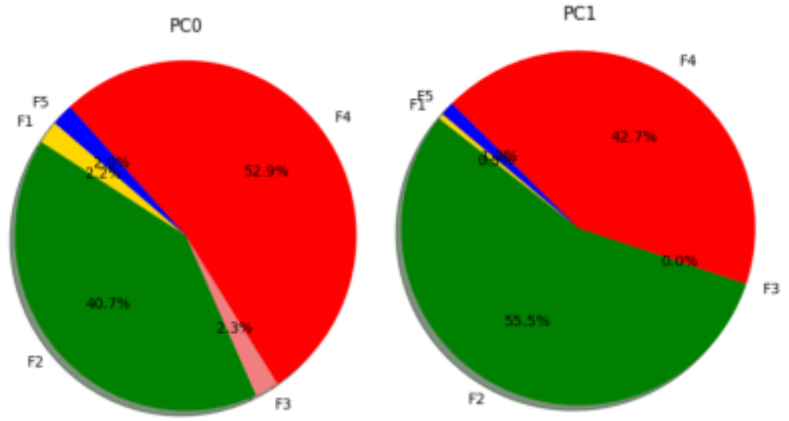
Methode:
Beispiel eines optimierten Szenarios

Ergebnisanalyse: Funktionales Szenario eines Überholvorganges zeigt die riskanten Dimensionen und Parameter in einer Risikomodellierung.



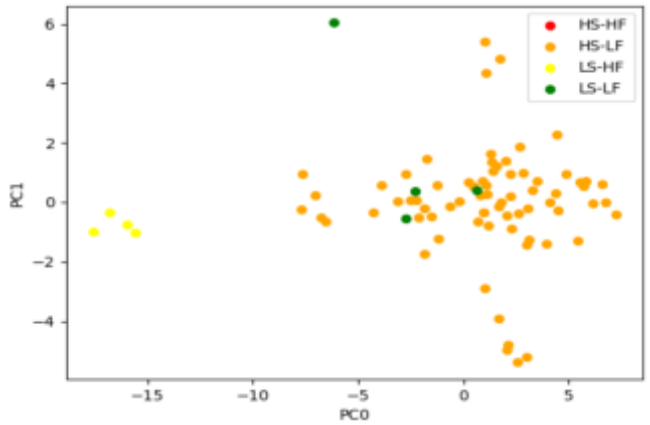


Dimensionsreduktion auf zwei Hauptkomponenten



Tageszeit (rot) und Nebel (grün) sind die kritischsten Einflüsse

Risiko-Modellierung



Beispiel: Anwendungsszenarien im Bereich Off-Road

Aufgrund fehlender Eckdaten werden Corner Cases oft auch synthetisch nachgebildet.

Deponie

z.B. Oberflächen:

Form
- Grube
- Halde
- Böschung

Beschaffenheit
- Fahrspuren
- Schotter
- Pflanzen

beweglich
- Kipper
- Bagger
- Walze

unbeweglich
- Container
- Spundwand

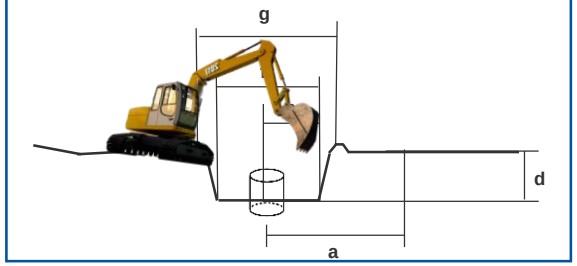
z.B. Objekte

Bild: Th. Egloffstein, 2021, ICP mbH

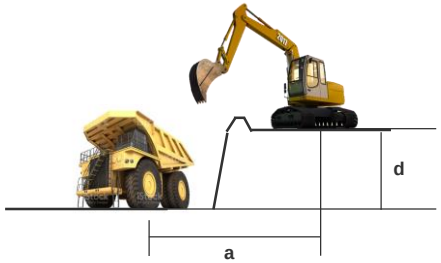
Abgeleitete Szenarien

Navigation in unwegsamem Gelände (z.B. an steilen Hängen)

Graben einer Baugrube zum Entfernen von Material



Beladung eines Kippers mit Material



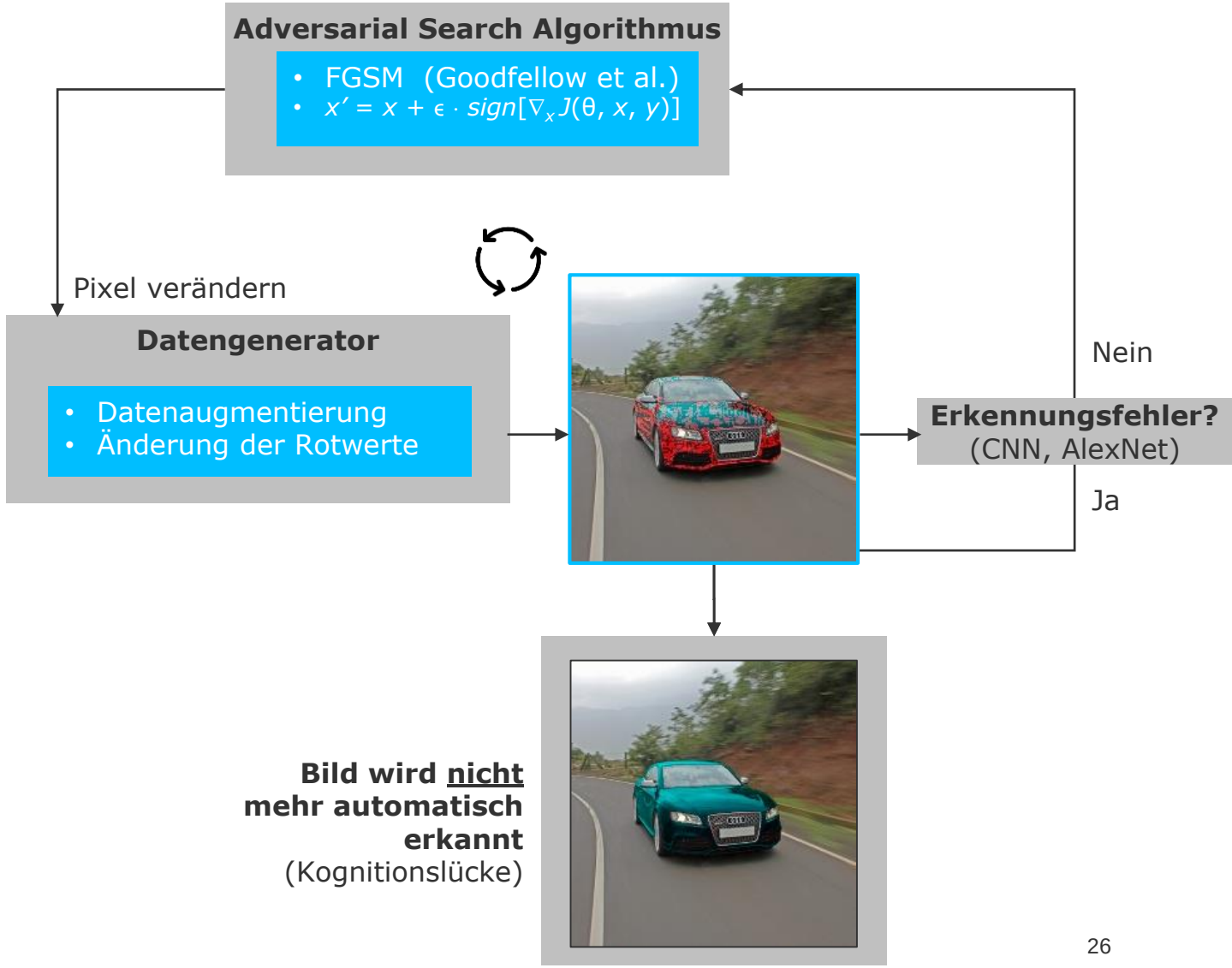
Betrieb bei extremen Bedingungen (z.B. Regen, schlammiger Boden...)

Bilder: Univ. Stuttgart, IAS

Methode:
Wie lassen sich unbekannte Fehler finden?

Beispiel zum systematischen Auffinden von Kognitionslücken (in CNN).

- ML-basierte Verfahren zur Bilderkennung weisen Erkennungslücken auf, die Menschen nicht nachvollziehen können
- Durch „Adversarial Search" können solche Lücken gefunden werden [Vie2021]
- Sind die Lücken bekannt, dann können diese durch ein Nachtraining behoben werden



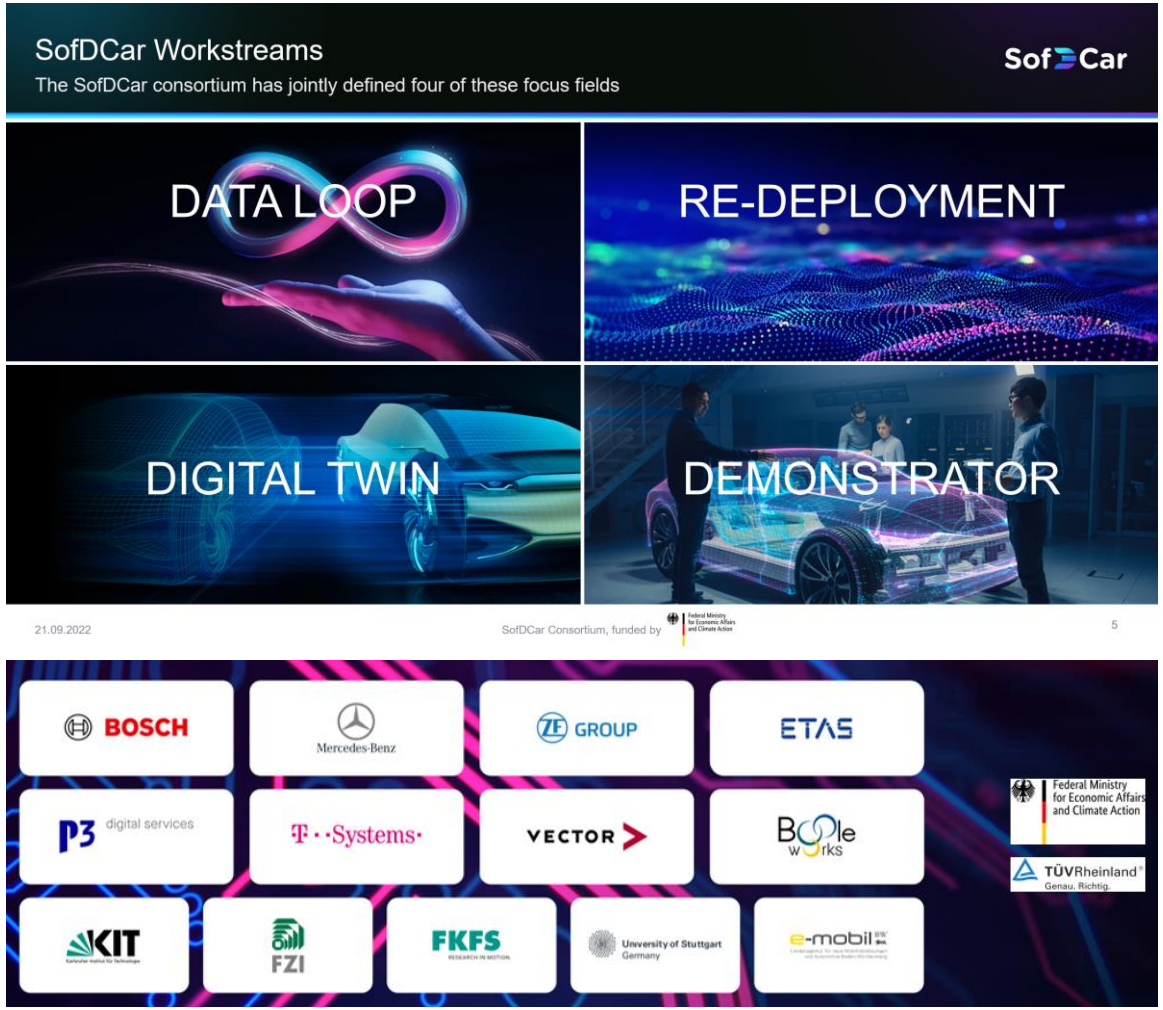
Inhalt

- Einführung
- Evolutionsstufen in der Systementwicklung (von CI/CD und dem Data Loop)
- Validierung Autonomer Fahrfunktionen
- **Aktuelle und zukünftige Arbeiten**

Software-definierte Fahrzeuge (Leitprojekt SofDCar)

SofDCar Alliance beschäftigt sich mit den Herausforderungen der zukünftigen E/E- und Software-Architektur in Fahrzeugen.

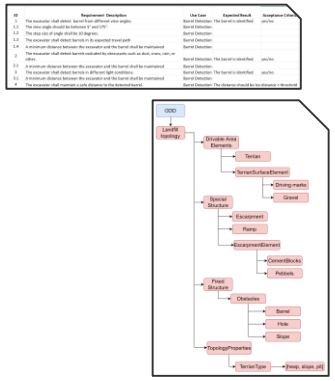
- Die Fahrzeuge werden als Teil eines **Netzwerks aus allen Fahrzeugen** und der Infrastruktur betrachtet
- **Digitale Zwillinge** auf Basis effizienter Datenstrukturen bilden ein virtuelles Abbild der physischen Fahrzeuge
- Ein **Data Loop** ermöglichet eine Verbindung zwischen dem Fahrzeug im Betrieb und der Entwicklung, z.B. zum Re-Deployment von Software



Projekt SynDAB: Synthetische Daten für die Entwicklung

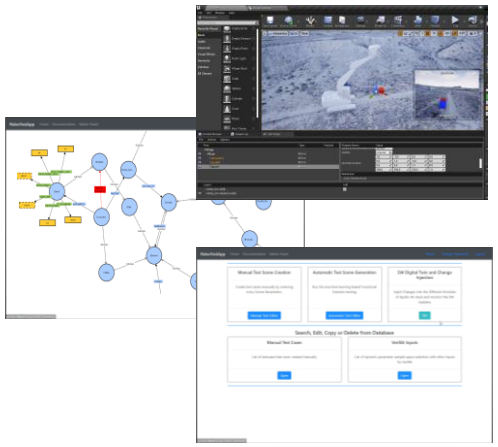
Die Erzeugung von synthetischen Daten für kritische Zustände ist bei Bau- und Arbeitsmaschinen unumgänglich, um voll-automatische Funktionen erstellen zu können.

Requirement



Assistance System for Management

Optimized Scenario Definition



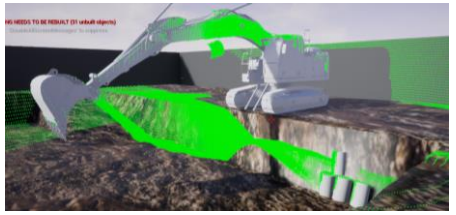
Test Case Generator

Automatic Scenario Creation

UNREAL ENGINE · USD · JSON · rosbridge

Training and Test

Hybrid Data



ROS

Bilder: Univ. Stuttgart, IAS, Fhg IOSB, Robo-Test)

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung
Associated project of
IntCDC CLUSTER OF EXCELLENCE

Systematische Validierung autonomer Systeme

Prozesse und Tools zur Definition und Durchführung von Trainings und Tests.

Benötigt werden:

- Geeignete **Standards** schaffen Klarheit zum Vorgehen und den Umfängen
- **Pfiffige Datenanalyse**, Algorithmen und Co-Simulation helfen bei der Überprüfung
- **kognitives Testen** mit Berücksichtigung von Randbedingungen und Ausnahmefällen bilden eine Basis
- KPIs für ein nachvollziehbares Training und **optimierte Abdeckung** für Test

IAS-Teammitglieder (Bild: Univ. Stuttgart)

University of Stuttgart
Institute of Industrial Automation
and Software Engineering

Prof. Dr.-Ing. Dr. h. c. Michael Weyrich

e-mail michael.weyrich@ias.uni-stuttgart.de
web www.ias.uni-stuttgart.de
phone +49 711 685 67301

Bei der Erstellung unterstützt hat das folgende Team:

Andreas Löcklin

Alexander Schuster

Hannes Vietz

Manuel Müller

Iman Sonji

Literaturverzeichnis / Referenzen

[Dor2017]	Derek Doran, Sarah Schulz, Tarek R. Besold: What Does Explainable AI Really Mean? A New Conceptualization of Perspectives. CoRR abs/1710.00794 (2017); https://doi.org/10.48550/arXiv.1710.00794
[Stü2009]	Stückelberger, Ch.: Welche Ethik schafft Glaubwürdigkeit im gemeinnützigen (und kommerziellen) Wirken?, Vortrag Swiss Philanthropiy Forum, 05.03.2009 https://www.slideserve.com/mandell/philanthropie-und-vertrauen-welche-ethik-schafft-glaubw-rdigkeit-im-gemeinn-tzigen-und-kommerziellen-wirken (abgerufen 13.09.2022)
[Hal2020]	Sebastian Hallensleben, Carla Hustedt (Hrsg.): „From Principles to Practice – An interdisciplinary framework to operationalise AI ethics", Juni 2020, VDE, Bertelsmann-Stiftung https://www.ai-ethics-impact.org/resource/blob/1990526/c6db9894ee73aefa489d6249f5ee2b9f/aieig---report---download-hb---en-data.pdf (abgerufen 13.09.2022)
[VDE900012]	VDE SPEC 900012 V1.0 (en): VCIO based description of systems for AI trustworthiness characterization, April 2022 https://www.vde.com/resource/blob/2176686/a24b13db01773747e6b7bba4ce20ea60/vde-spec-vcio-based-description-of-systems-for-ai-trustworthiness-characterisation-data.pdf (abgerufen 13.09.2022)
[Zel2019]	A. Zeller, "Absicherung von verteilten Automatisierungssystemen nach Änderungen der Steuerungssoftware – Modellkomposition zur Nutzung der funktionalen Verifikation", Institut für Automatisierungstechnik und Softwaresysteme der Universität Stuttgart, 2019. (abgerufen 24.09.2022)
[Kal2016]	N. Kalra and S. M. Paddock, "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?," Transportation Research Part A: Policy and Practice, vol. 94, pp. 182–193, 2016.
[Vie2021]	H. Vietz, T. Rauch, A. Löcklin, N. Jazdi, und M. Weyrich, „A Methodology to Identify Cognition Gaps in Visual Recognition Applications Based on Convolutional Neural Network", in 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE), Lyon, France, 23-27 August 2021