Universität Stuttgart
Institute of Industrial Automation and Software Engineering

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

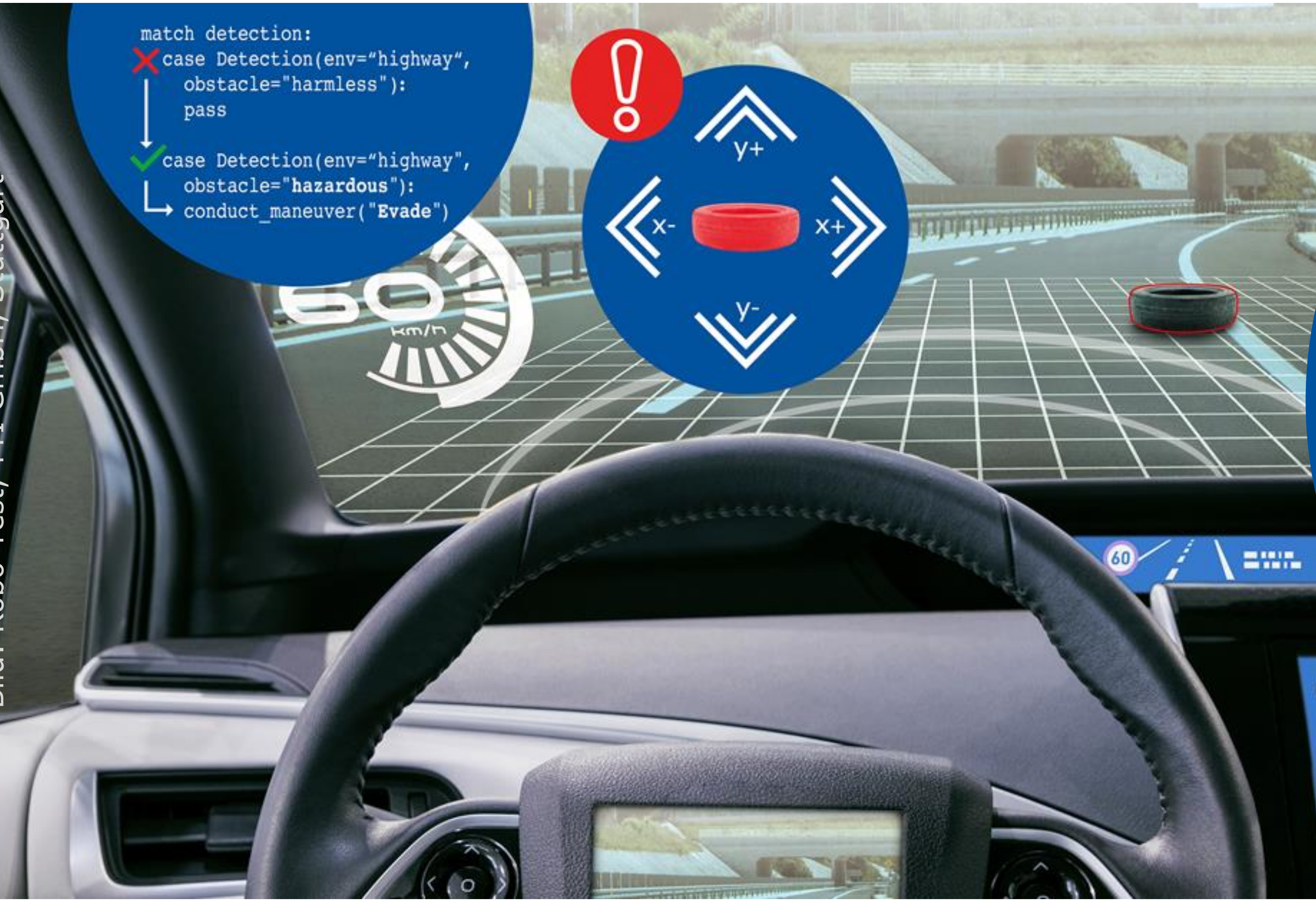


Bild: Robo-Test, TTI GmbH, Stuttgart

Hey K.I.T.T. can I trust you?

Validation and Verification of Autonomous Systems

Prof. Michael Weyrich

University of Stuttgart, Institute of Industrial Automation and Software Engineering

Smart Robot Car „K.I.T.T.“ Sci-fi television series "Knight Rider" from 1982

Autonomous vehicles (with driver) with resistant body, controlled by an AI.

K.I.T.T. und K.A.R.R. demonstrate confidence in autonomous vehicles and fear of deception.

- K.I.T.T. is designed to protect human life and demonstrates a cooperative, trusting relationship with its driver
- In contrast, K.A.R.R. is focused on self-preservation

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

Bild: Universal Television

Content

- Introduction
- Evolutionary Stages in System Development (from CI/CD and the Data Loop)
- Validation of autonomous driving functions
- Current and future work

Content

- **Introduction**
- Evolutionary Stages in System Development (from CI/CD and the Data Loop)
- Validation of autonomous driving functions
- Current and future work

Examples of fully automated and autonomous systems

Autonomous systems have a wide range of applications, from autonomous driving and logistics systems to modern production systems and robots.

Fully automated and autonomous vehicles (SAE Level 4 / 5)

auto motor sport

10.10.2019, auto motor sport

Bild: Fiat Chrysler

UPDATE FOR WAYMO-CARS

Safety drivers stay home

Handelsblatt

24.08.2022

Bild: Waymo

AUTONOMOUS DRIVING

Waymo and Daimler Truck test fully automated trucks in Texas

WirtschaftsWoche

24.07.2022

FIVE AI FROM CAMBRIDGE

Bild: PR

Will this start-up help autonomous driving achieve a breakthrough?

Comparison test: Automated robotic mower

stern Bild: Stern.de

Autonomous container truck

BASF We create chemistry

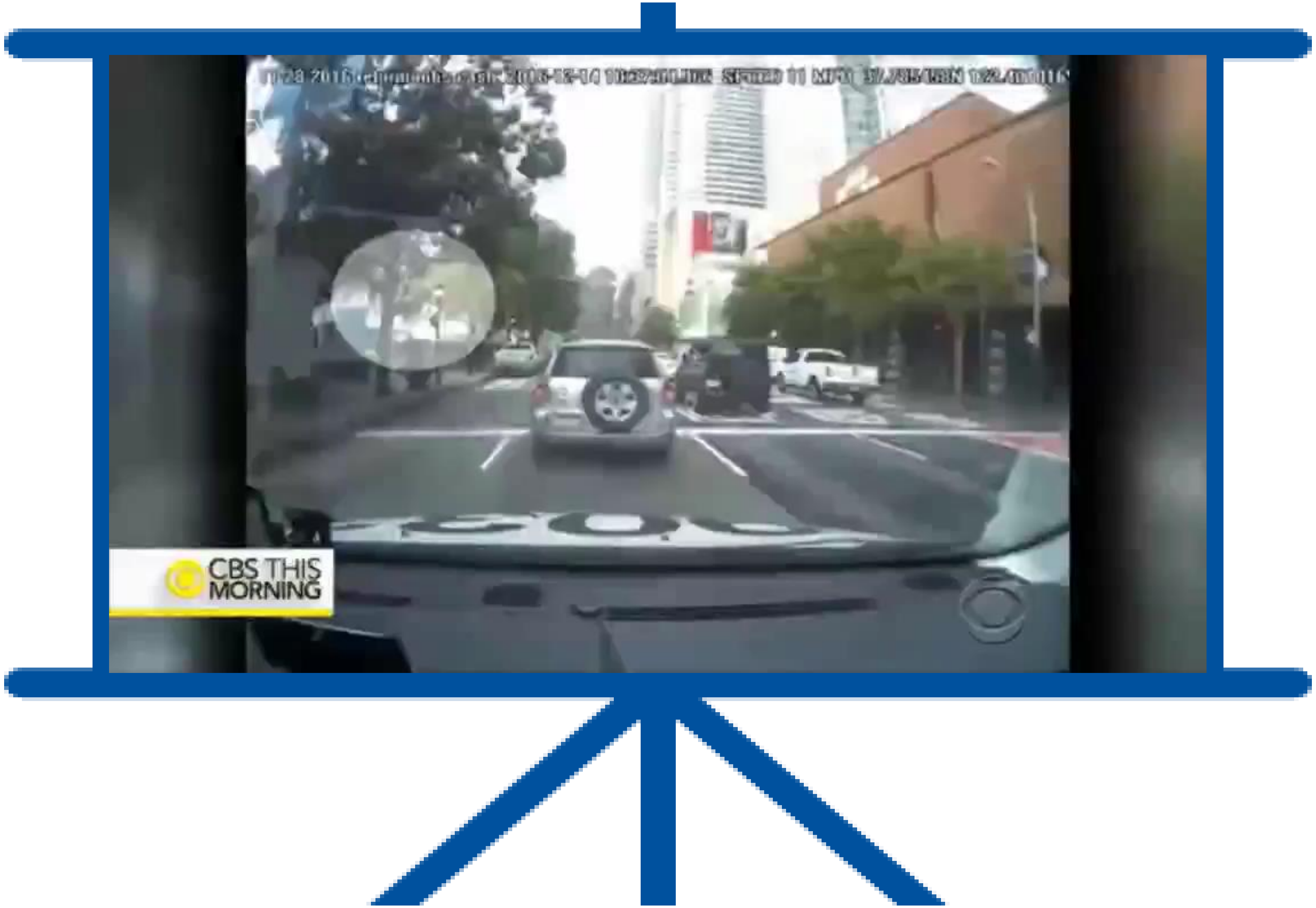
Bild: BASF, C. Schäfer

Future agriculture

JOHN DEERE

Bild: John Deere

Unfortunately mistakes happen...



What is trust and how to establish it?

“Trust is the ability to enter into a relationship / interact with a person or institution despite uncertainty and unmanageability.”
Ch. Stückelberger, Prof. for Ethics, Basel, 2009“

Ethicists define basic issues:
[Stü2009]

- Who trusts whom or what? (subjects, objects)
- How long? (duration)
- Why? (Reason for trust)
- By what means? (How is it achieved?)
- Encounter statement like: "Trust needs mistrust”

Characterization of AI trustworthiness
(VDE SPEC 90012, April 2022; [Hal2020])

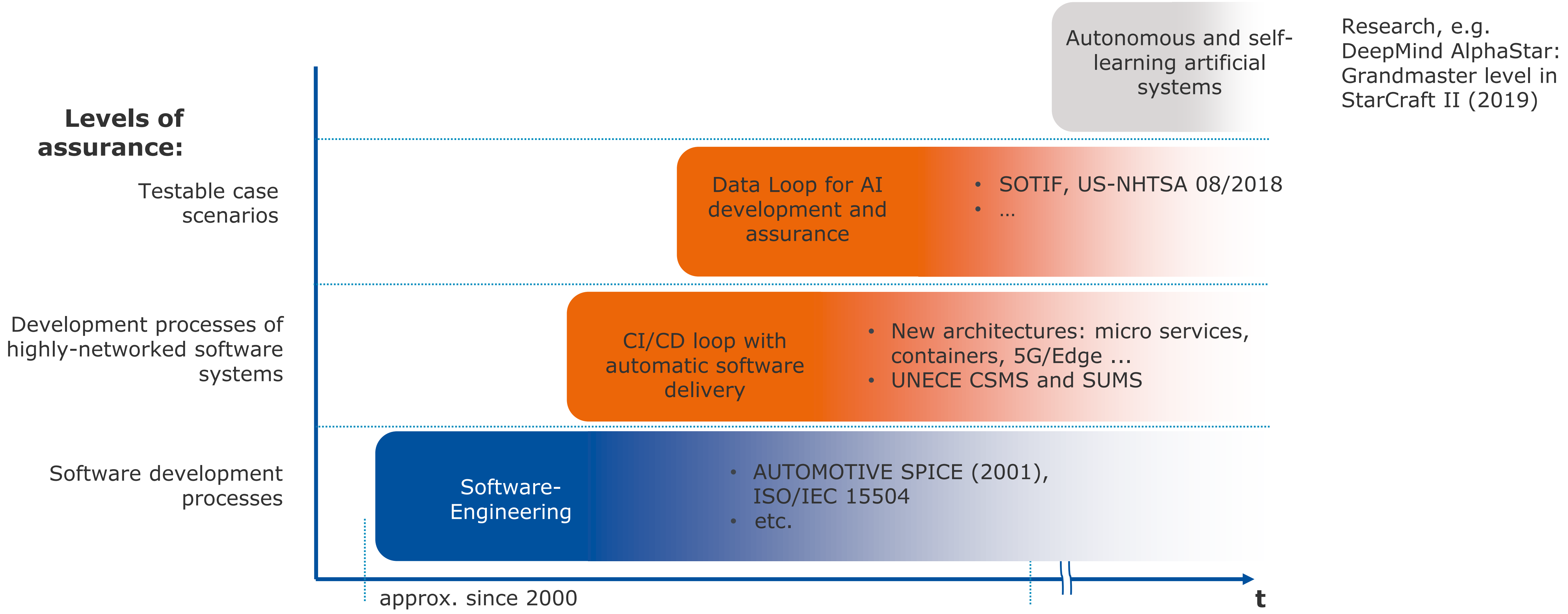
- Transparency (documentation, accessibility, comprehensibility)
- Accountability (responsibility, liability)
- Privacy (processing and protection of personal data)
- Fairness (appropriate metrics and assessments, sustainability)
- Reliability (robustness and dependability)
- Also: Safety and Security [Dor2017].

Content

- Introduction
- **Evolutionary Stages in System Development**
- **(from CI/CD and the Data Loop)**
- Validation of autonomous driving functions
- Current and future work

Evolutionary stages in system development

New forms of software engineering: via software-defined systems with "Continuous Integration and Deployment" (CI/CD) to the "Data Loop" for AI system developments.

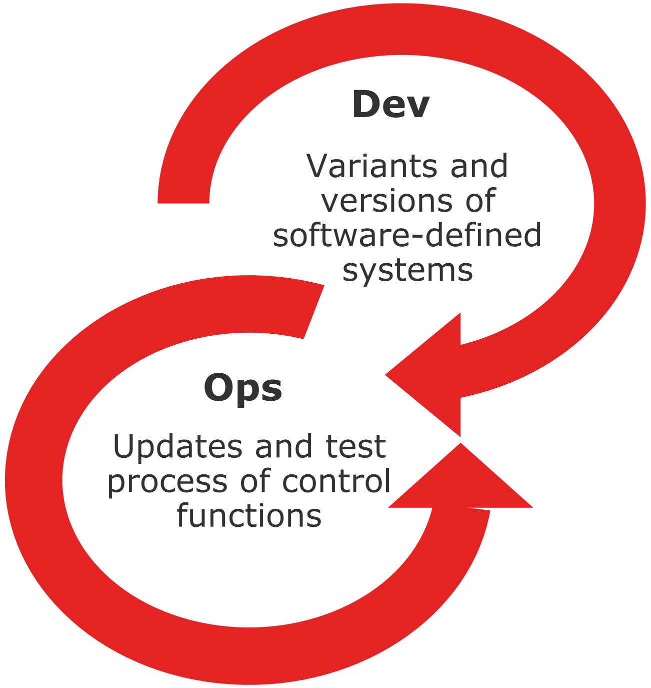


CI/CD-Loop with automatic Software deployment

New opportunities for (further) development are created due to the connection during operation of highly configurable software-defined mobility systems.

"Online" connection to vehicles in operation:

Development of new functionalities for the **re-deployment of software** and a meaningful **inclusion of the backend.**



Use of standardized testing procedures:

UNECE SUMS No. 156 - Software update and software update management system

Procedures for dealing with complex software product lines

UNECE CSMS No. 155 - Provisions for cyber security and cyber security management system

…

Networked communication and ubiquitous sensors

New functions raise questions in terms of security about the communication of the vehicles between each other, with the infrastructure and a back-end.

- Sensors in the vehicle record a wide range of **personal information**
- **Environment detection** by sensors outside the vehicle
- Vehicles and infrastructure **exchange data**
- There is an information exchange between vehicles and a **back-end**

IAS team members in the Arena 2036 (Bild: Uni of Stuttgart, IAS)

5G test field for highly-networked vehicles

During operation in the field vehicles and their components can communicate with each other, the infrastructure and the development departments.

- **Secure networking:** Edge and cloud communication on Internet protocols between vehicles, and traffic infrastructure
- **Management and reliable (re-)deployment** for variant-rich software with software product lines
- **Analysis and reliable synchronization of data** and information with the digital twin
- **Remote function provision (with Collective Perception)** for the vehicle from the back-end

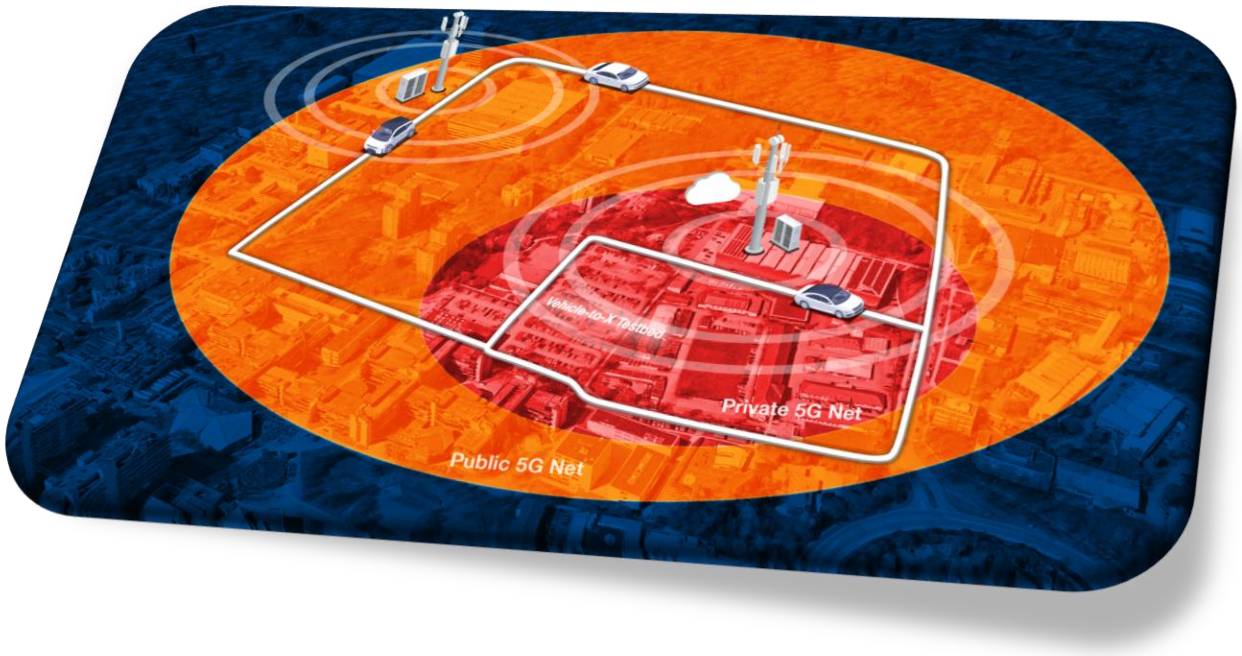


Bild: Univ. Stuttgart, IAS

5G test site (network: private and public) with edge and cloud at the Uni of Stuttgart, Paffenwaldring campus.

Content

- Introduction
- Evolutionary Stages in System Development (from CI/CD and the Data Loop)
- **Validation of autonomous driving functions**
- Current and future work

How many miles of driving would it take to demonstrate autonomous vehicle reliability?

Statistical considerations list (too) many test kilometers but leave questions about the test scenarios unanswered.

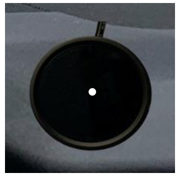
”An autonomous vehicle would have to travel 275 million redundant miles without a fatal accident to achieve behavior comparable to the U.S. accident rate” [Kal2016]

Bild: Robo-Test

Detected car in safe distance

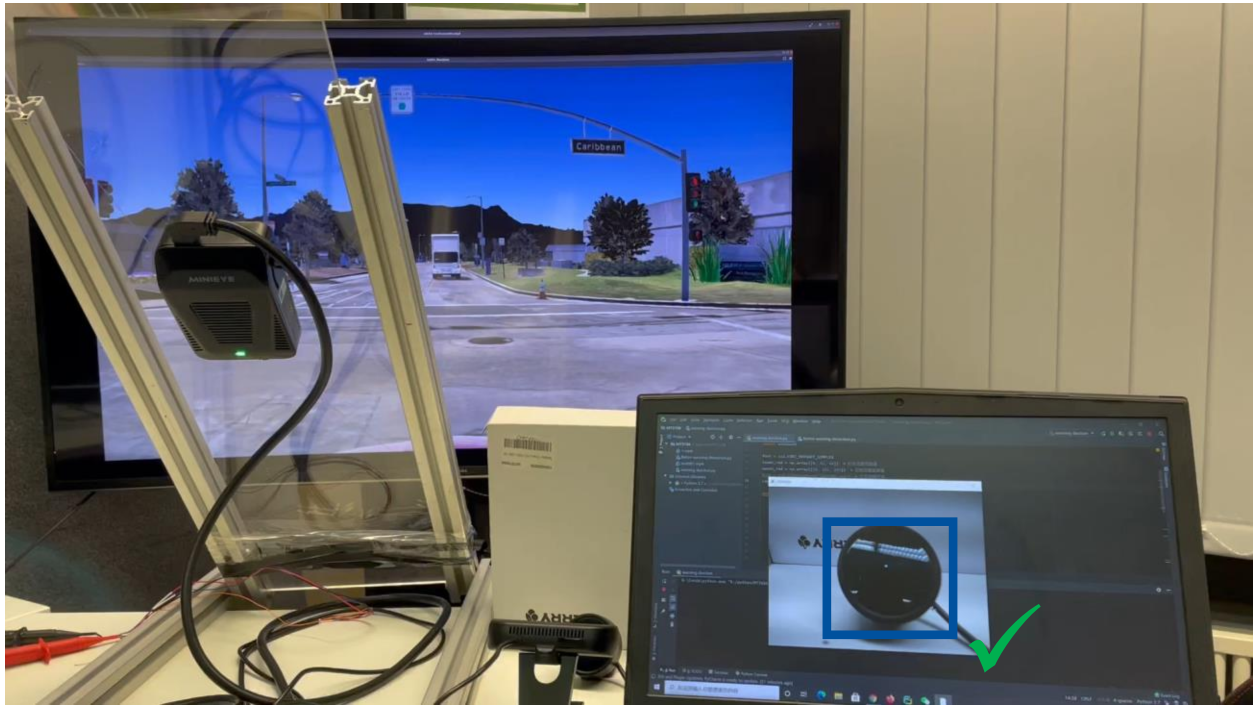
Detected in dangerous distance

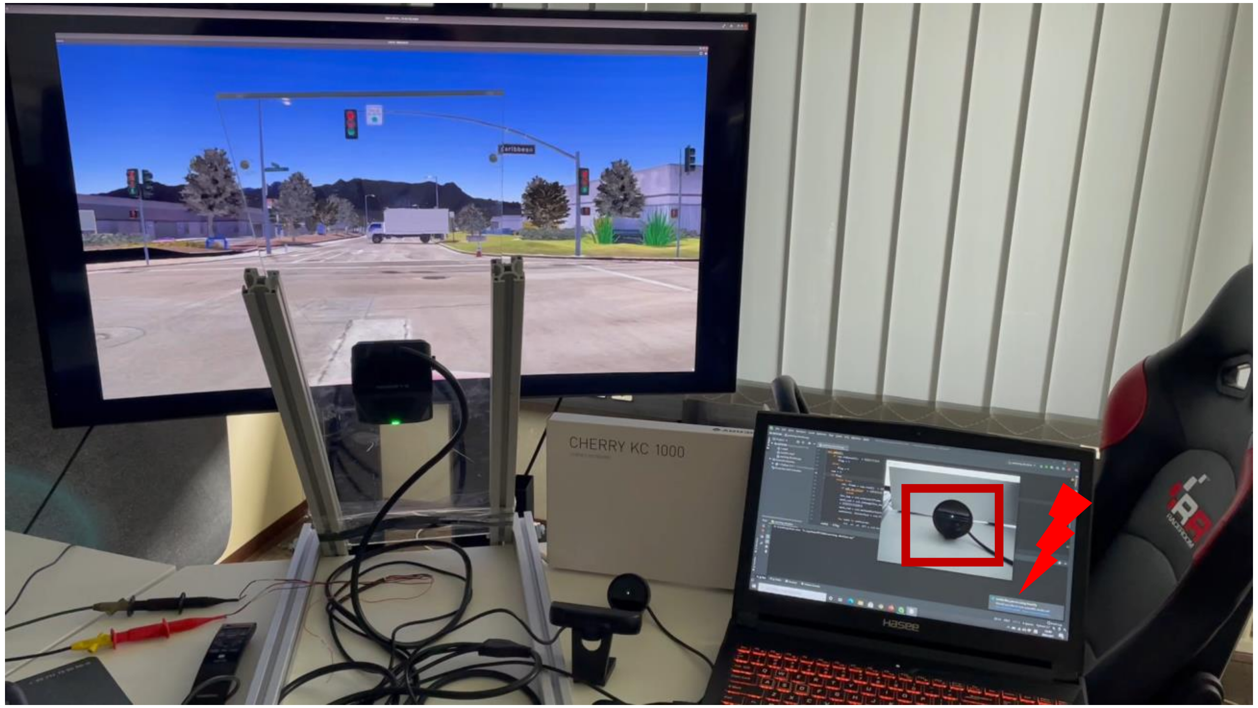
Detected nothing in front



Do driver assistance systems on the market malfunction?

A scenario test with synthetic data – a start-up company validates driver assistance systems.





Images: **Robo-Test**

Truck detected

Truck **not** detected

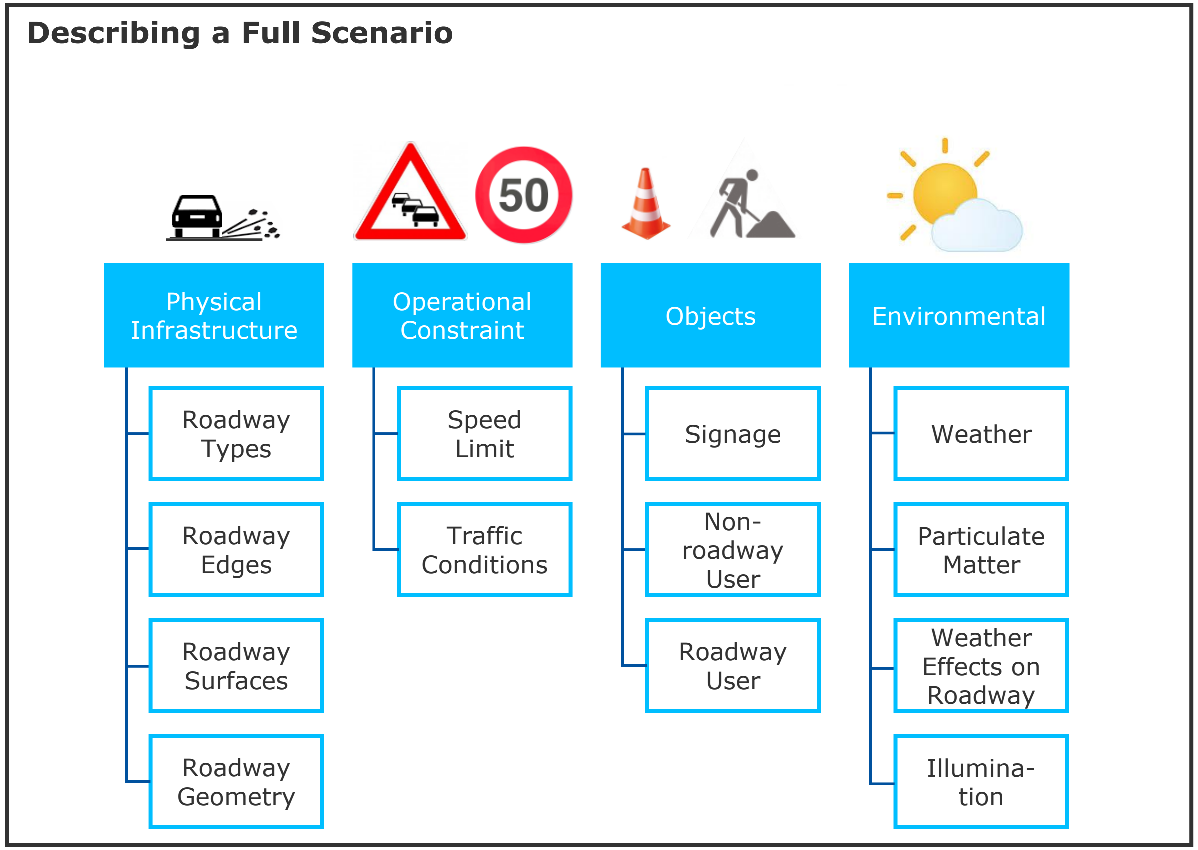
A variety of validation scenarios are required

For which scenarios are the autonomous functions to be designed and safeguarded, so that they function safely without restrictions?

Operational Design Domain (ODD)

ODDs define scenarios over time intervals and geographic regions

- **Environmental conditions** are to be defined with which an autonomous system must be able to deal with
- ODD scenarios should describe suitable and difficult **combinations**
- **ODD test management** must be able to model many scenarios



ODD Topologie nach US NHTSA DOT HS 812 623

Automatic generation of scenarios with ontologies

Dimensions and parameters of scenarios are varied in the simulation, using existing test cases and optimizing them for coverage and criticality.

Scenario:
Sudden obstacle
A person runs in front of the vehicle.

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

Scenario:
Overtaking maneuver
Overtaking in snow.

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

Example: A cognitive test tool for managing training and test cases

Creation of real and synthetic ODD scenarios has great significance

Robo-Test Platform for ODD training and validation testing

- Use of **different simulation systems** (e.g. VTD, Unreal Engine) for vision and lidar
- Special simulation **GNSS**
- Inclusion of **real data**
- Creation of **libraries** with standard scenarios

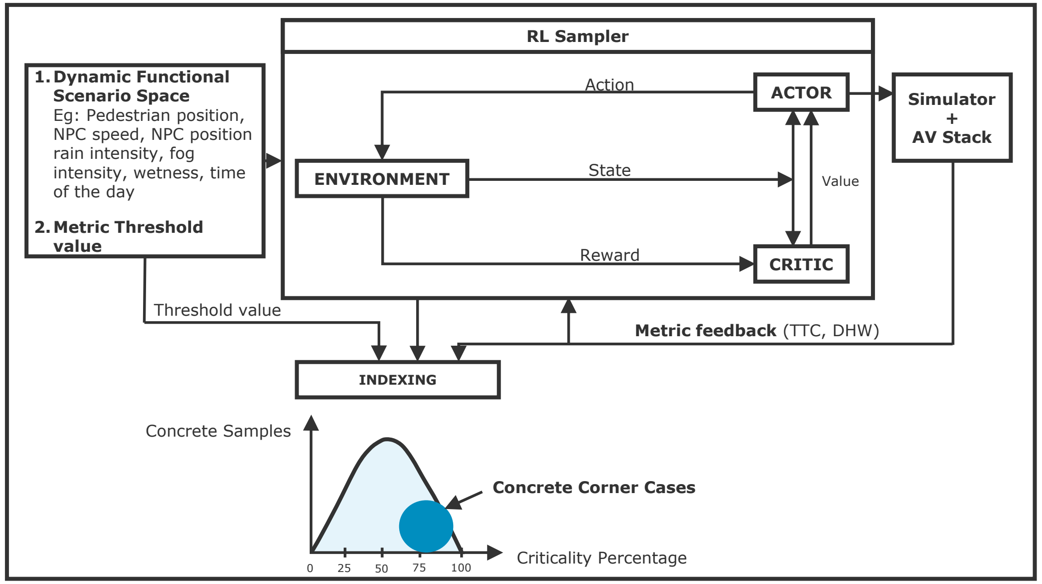
- Definition of scenario parameters
- **Ontologies** for systematic generation when generating scenarios
- Efficiency through tools to optimize **coverage**

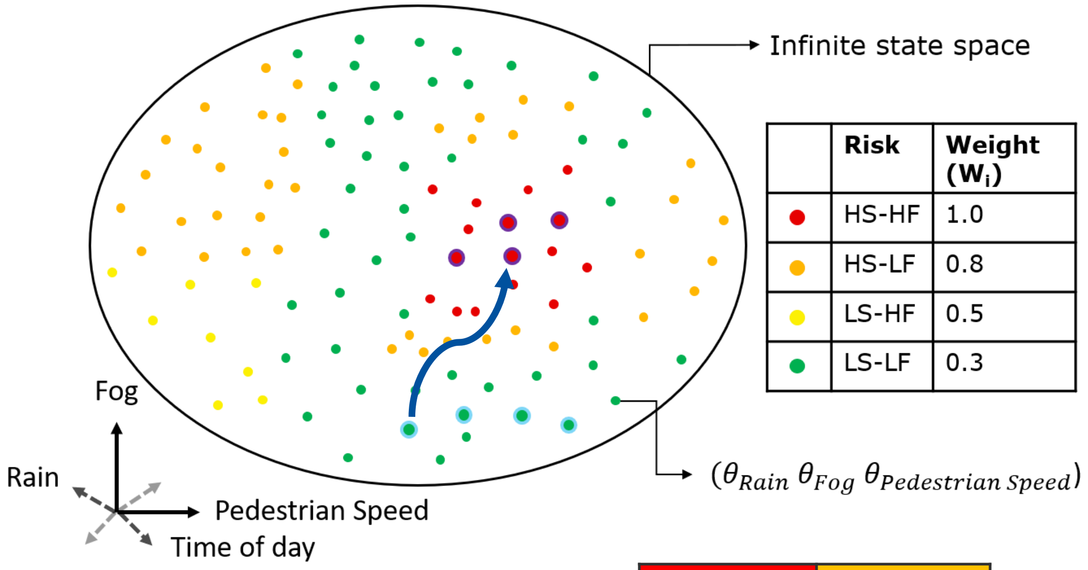
Method: How can the critical dimensions be identified?

In the simulation, the critical dimensions and parameters of a scenario can be assessed and identified.

The dimensions of the scenarios are run through with **re-enforcement learning** to cover critical areas.

The **risks of the individual dimensions and their parameters** in the scenarios become assessable.





	Risk	Weight (W_i)
● (red)	HS-HF	1.0
● (orange)	HS-LF	0.8
● (yellow)	LS-HF	0.5
● (green)	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

Method: Omit non-essential main components

Further focus of the scenarios through a Principal Component Analysis (PCA) based on the high-risk test cases from the re-enforcement learnings.

Approach: Main components (PCs) that are non necessary are omitted. In the example: "Fog" is considered particularly critical, whereas "rain" has little influence and can be omitted.

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

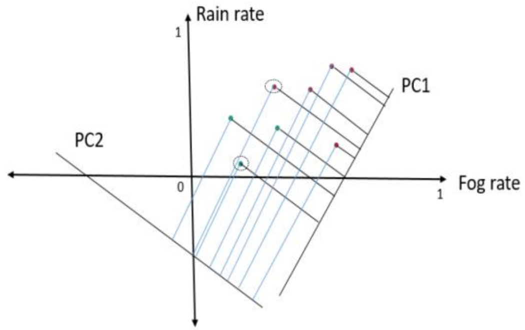
m

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~~	
$\sum$	**2,0**	**1,0**	~~**0,8**~~	

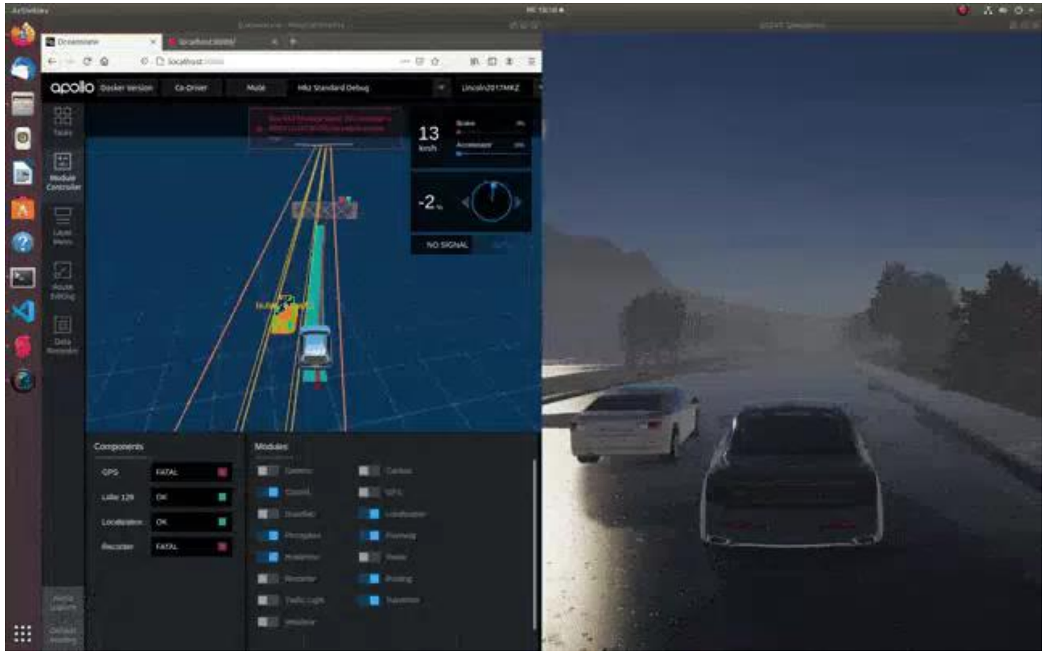
Inv. PCA

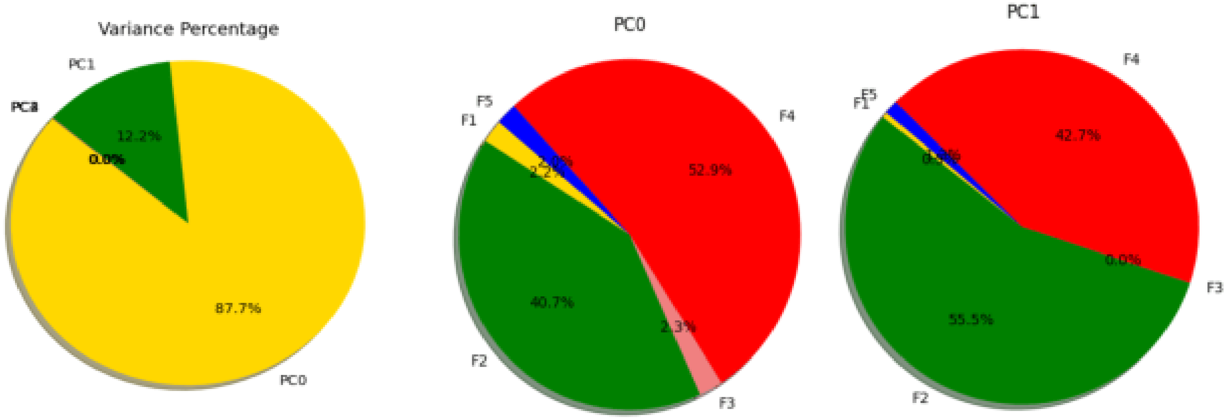
	Fog	Pedestrian speed	Rain
PC1	60%	30%	10%
PC2	14%	70%	16%



Method: Example of an optimised scenario

Result analysis: Functional scenario of an overtaking situation, shows the risky dimensions and parameters in a risk modelling.

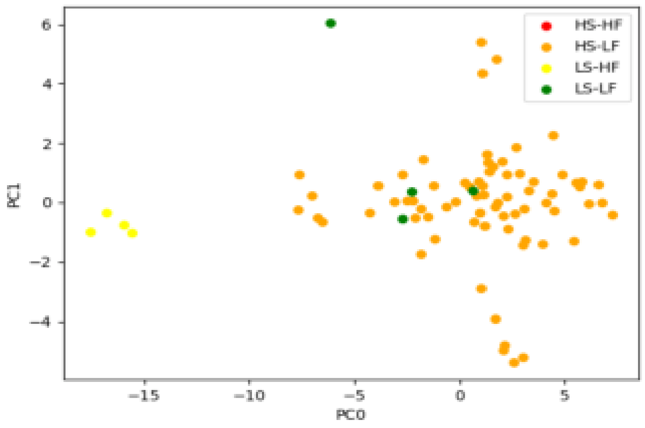




Dimensional reduction to two main components

Time of day (red) and fog (green) are the most critical influences

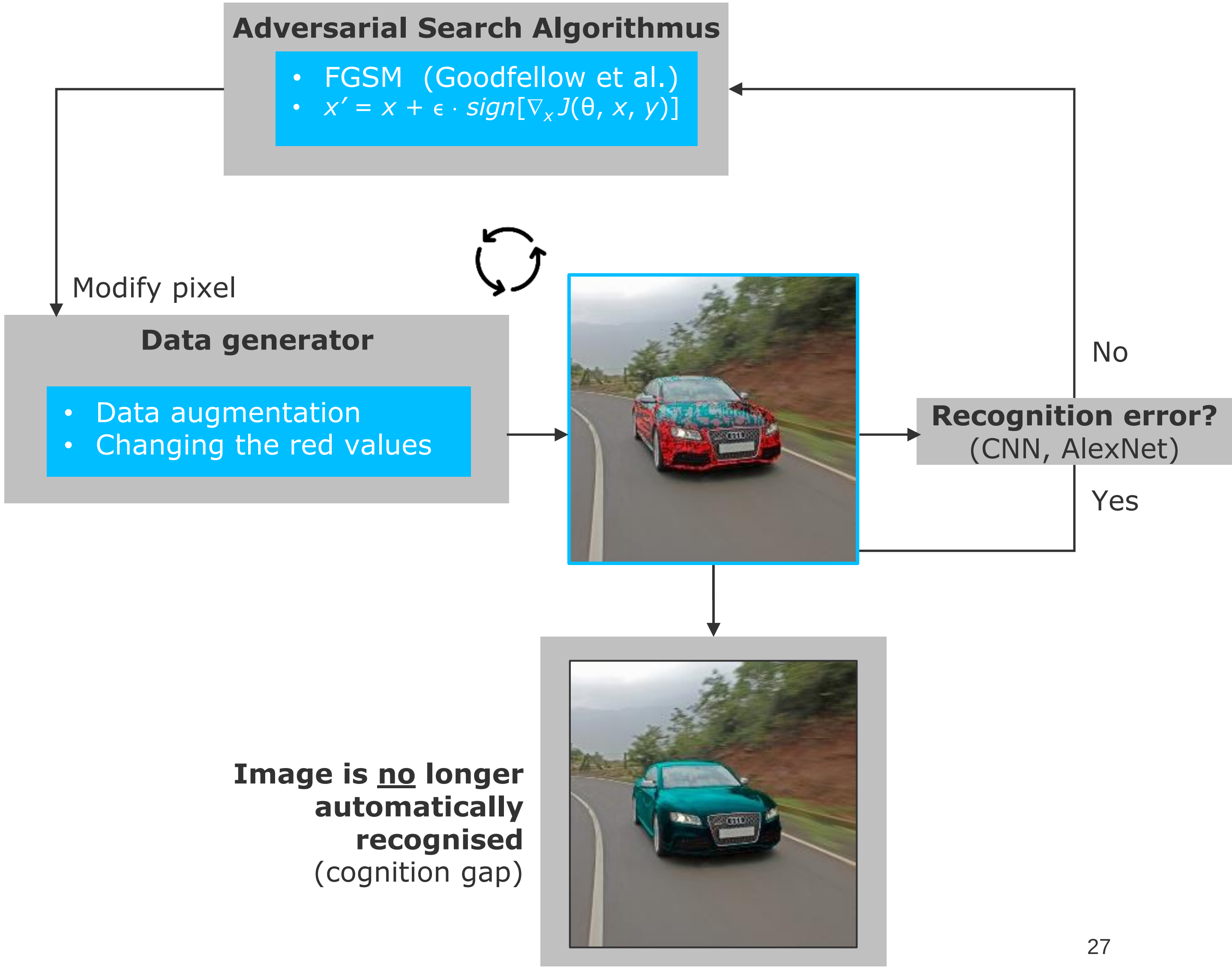
Risk modelling



Method: How can unknown faults be found?

Example of systematically finding gaps in cognition (in CNN).

- ML-based image recognition methods have recognition gaps that humans cannot comprehend
- Such gaps can be found through "Adversarial Search" [Vie2021].
- If the gaps are known, then they can be corrected through post-training



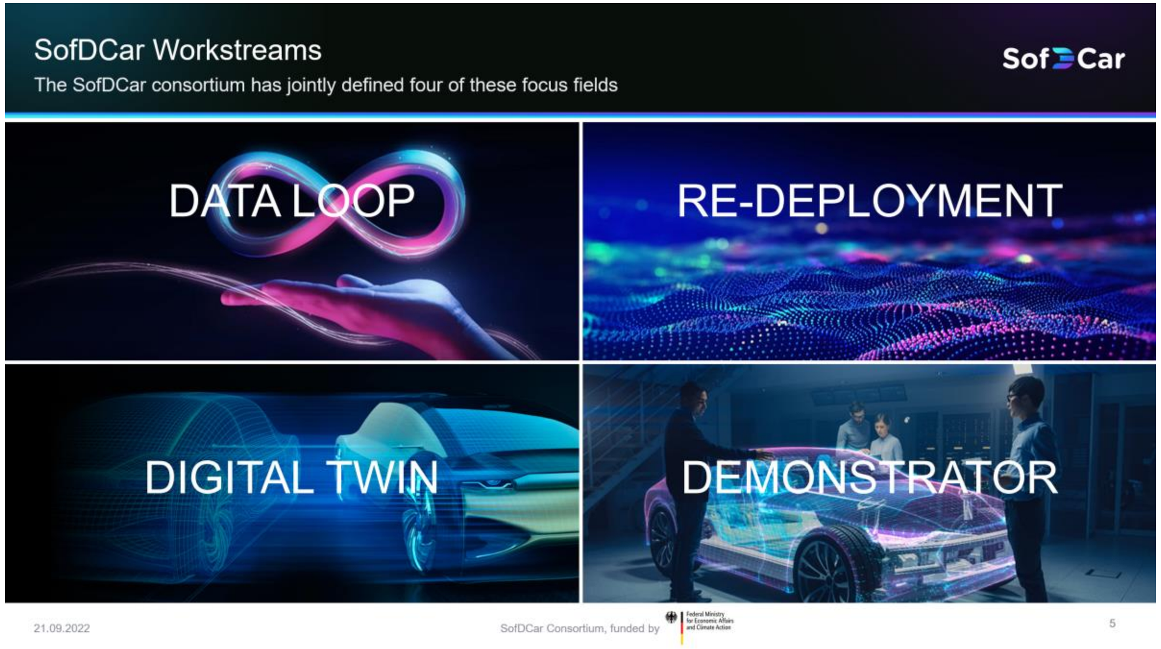
Content

- Introduction
- Evolutionary Stages in System Development (from CI/CD and the Data Loop)
- Validation of autonomous driving functions
- **Current and future work**

Software-defined vehicles (SofDCar lead project)

SofDCar consortium addresses the challenges of future E/E and software architecture in vehicles.

- Vehicles are considered as part of a **network of all vehicles and infrastructure**
- **Digital twins** based on efficient data structures form a virtual image of the physical vehicles
- A **data loop** enables a connection between the vehicle in operation and development, e.g. for the re-deployment of software.



Systematic validation of autonomous systems

Processes and tools to define and deliver training and testing.

Needed are:

- Appropriate **standards** create clarity of the procedure and scopes
- **Clever data analysis**, algorithms and co-simulation help with testing
- **Cognitive testing** with consideration of boundary conditions and exceptional cases form a basis
- KPIs for comprehensible training and **optimised coverage** for testing

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

The following team members assisted in the preparation:

Andreas Löcklin

Alexander Schuster

Hannes Vietz

Manuel Müller

Iman Sonji

Bibliography / References

[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