KI ABSICHERUNG

Safe AI for Automated Driving

11th March 2021, Online, Interim presentation

KI Absicherung

Dr. Stephan Scholz, Volkswagen AG

KI ABSICHERUNG

Safe AI for Automated Driving

Consortium Lead: Volkswagen AG

Co - Lead: Fraunhofer IAIS

Budget: €41M

Funding: €19.2M

24 Partners

Duration: 36 months

01.07.2019 – 20.06.2022

Partner

VOLKSWAGEN AKTIENGESELLSCHAFT, Audi, BMW GROUP, Opel, BOSCH Invented for life, Continental, Valeo, ZF, EFS, ASTech Automotive Safety Technologies, intel, Luxoft A DXC Technology Company, MACKEVISION, MERANTIX, qualityminds, umlaut, DFKI Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, DLR, Fraunhofer IAIS, Fraunhofer IKS, FZI, Bergische Universität Wuppertal, Technische Universität München TUM, Universität Heidelberg

External techonolgy partners

BIT TECHNOLOGY SOLUTIONS, neurocat, understand.ai, ejct

1

Project vision and goals


Making the safety of AI-based function modules for highly automated driving verifiable
KI ABSICHERUNG
Safe AI for Automated Driving
Pedestrian detection

Challenge

AI Land

Promising new technology with unimagined possibilities

Established safety processes cannot be applied

KI ABSICHERUNG
Safe AI for Automated Driving

Safety Land

Safe, trustworthy driving function

Industry consensus (Safe AI): Methodology for joint safety argumentation

Main goals

1. Methods for training and testing of AI-based functions

KI Absicherung develops and investigates means and methods for verifying AI-based functions for highly automated driving.

2. Safety argumentation

For the pedestrian detection use case, the project is developing an exemplary safety argumentation and methods for verifying a complex AI function.

3. Communication with standardization bodies on AI certification

The project‘s results will be used in the exchange with standardization bodies to support the development of a standard for safeguarding AI-based function modules.

The KI Familie and its projects

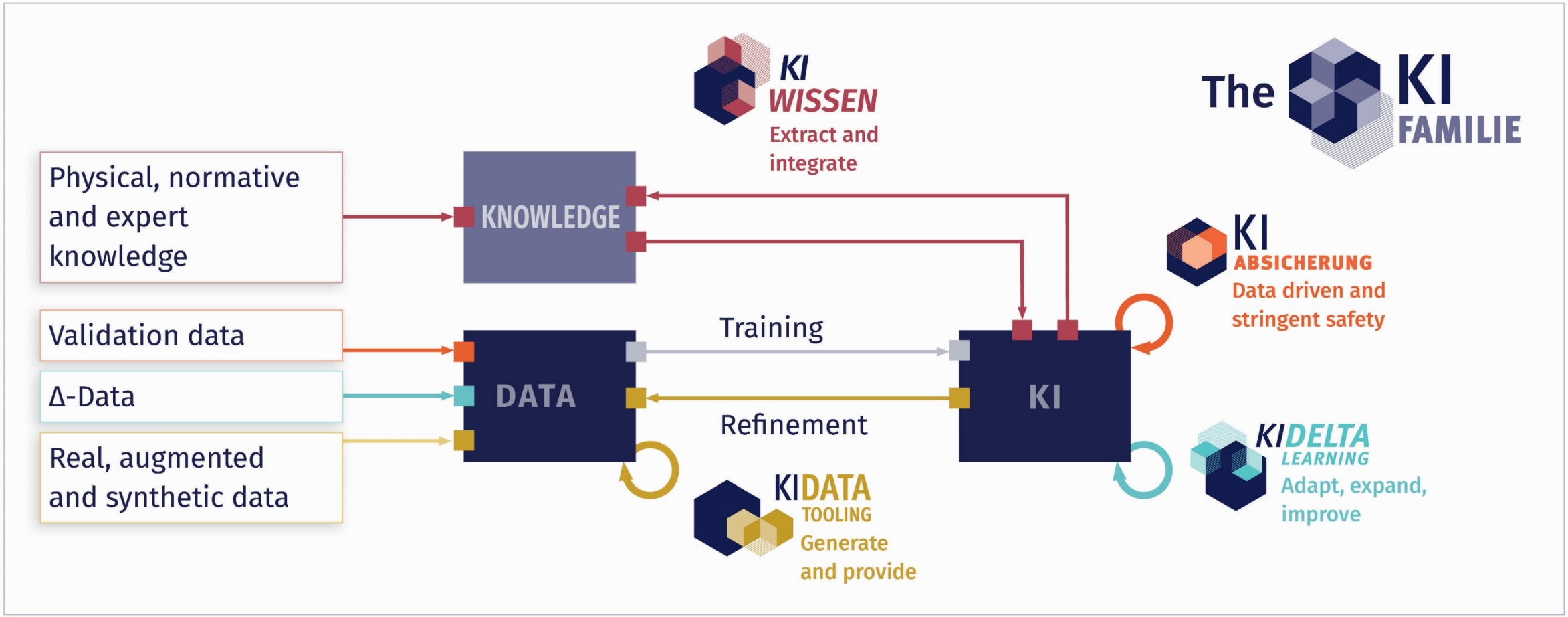
KI WISSEN Development of methods for the integration of knowledge into machine learning

KI DELTA LEARNING
Development of methods and tools for the efficient expansion and transformation of existing AI modules in autonomous vehicles to meet the challenges of new domains or more complex scenarios

KI ABSICHERUNG Methods and measures to safeguard AI-based perception functions for automated driving

KI DATA TOOLING Methods and tools for the generation and refinement of training, validation and safeguarding data for AI functions in autonomous vehicles

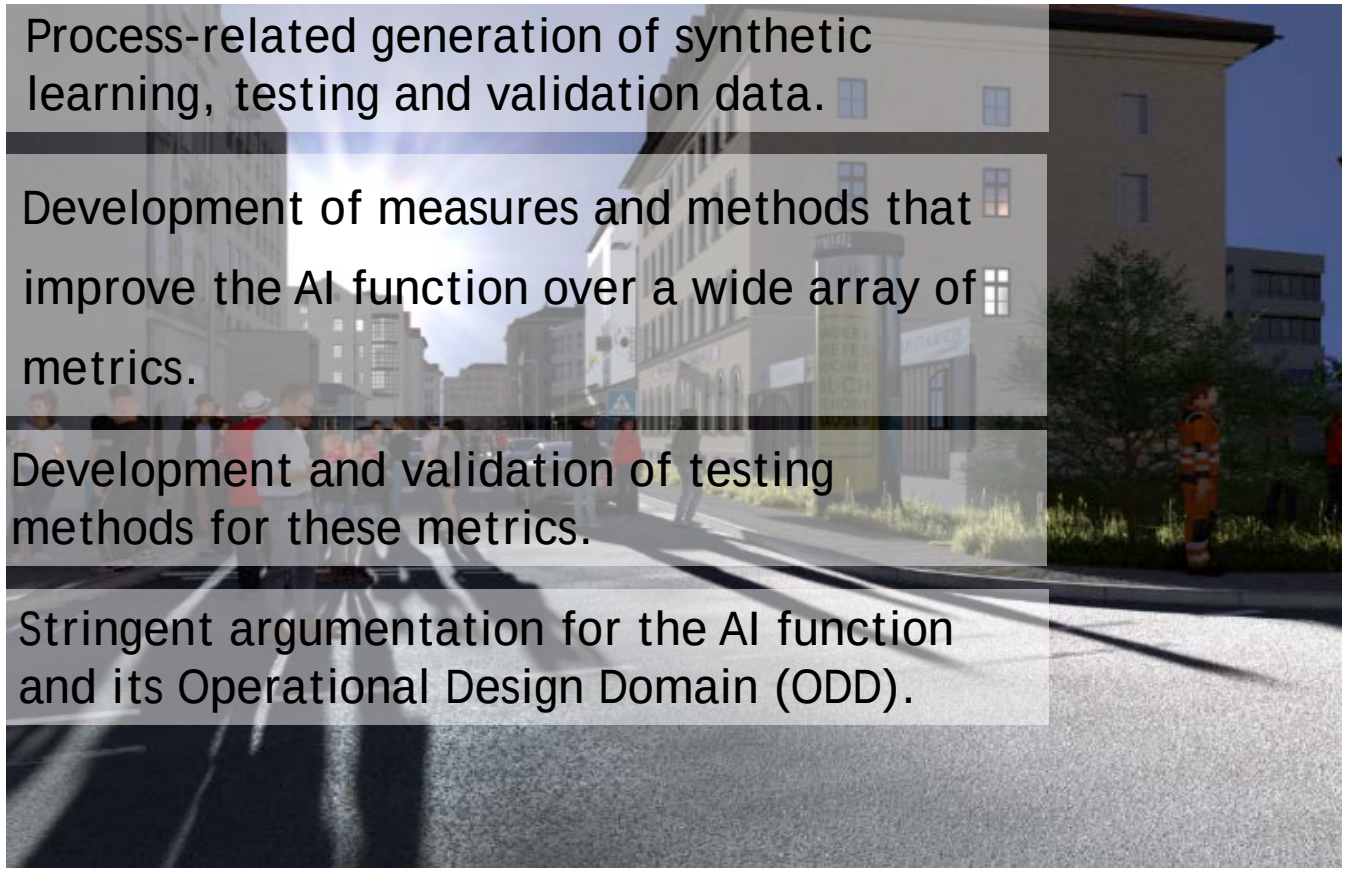
KI Absicherung in connection with the KI Familie



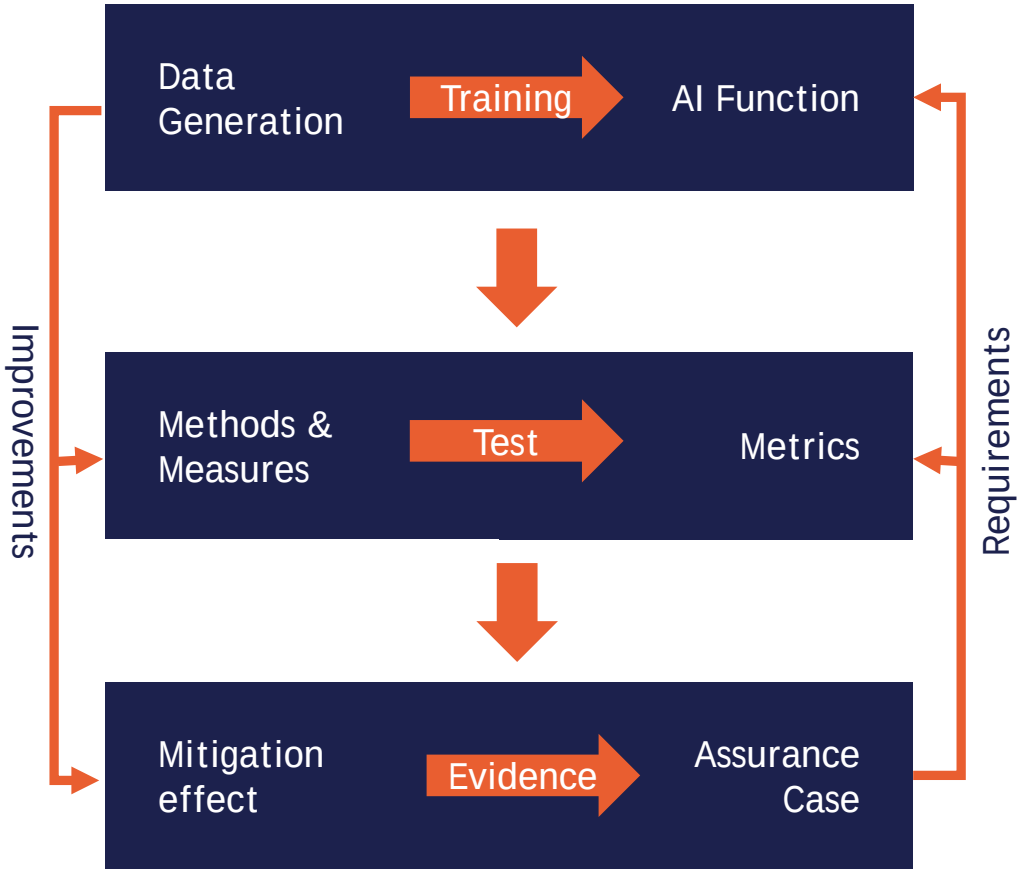
2

Methodological and conceptual approach

From a data-driven AI function to an Assurance Case



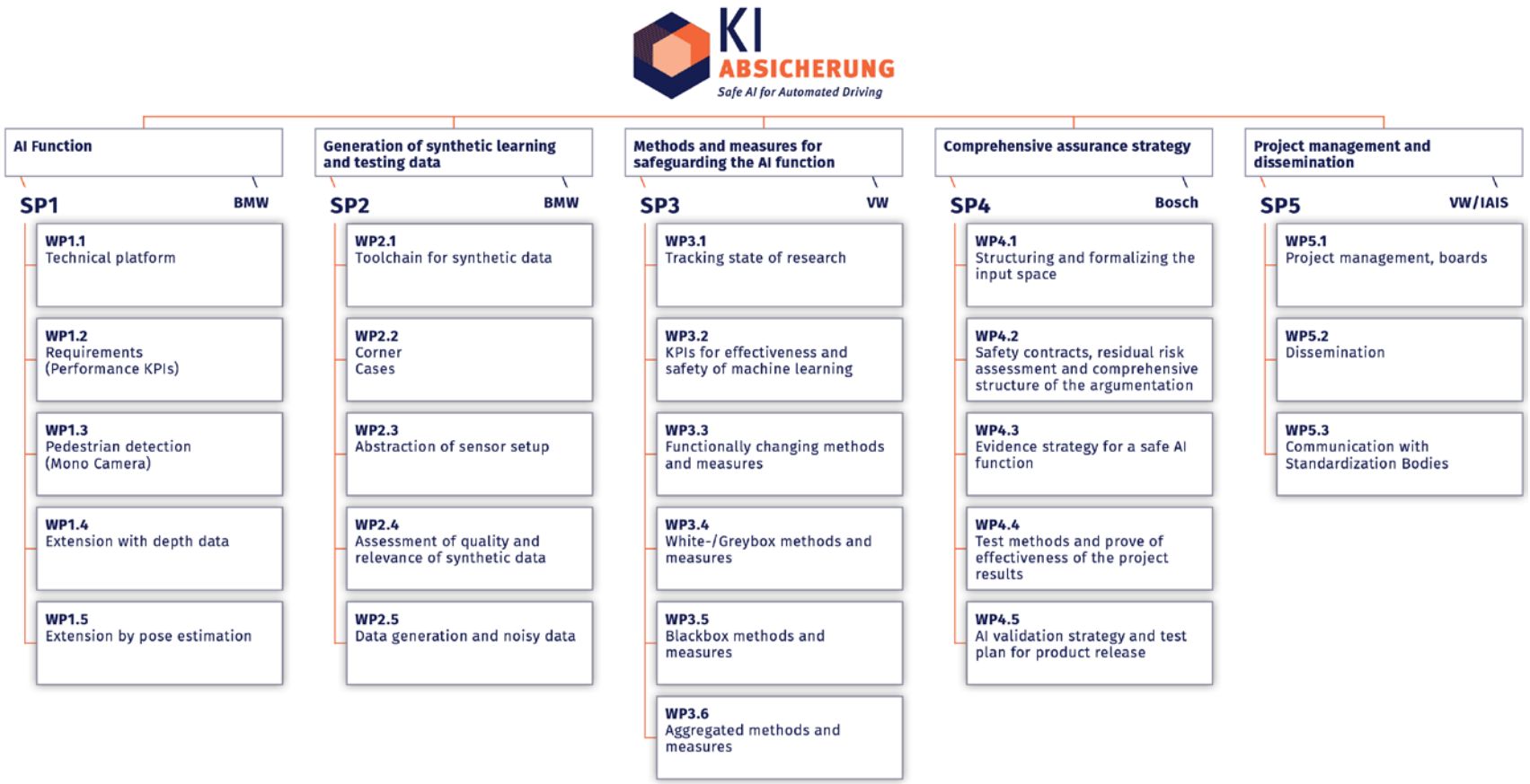
Source: BIT Technology Solutions



Conceptual approach

1. Provide the AI function for pedestrian detection.
2. Generate synthetic learning, testing and validation data.
3. Develop and evaluate measures and methods for the verification of the AI function.
4. Establish an overall safety strategy for the AI function.
5. Define and implement an Assurance Case.

Project structure with sub-projects and work packages



Our Approach: Establishment of a Holistic Safety Strategy

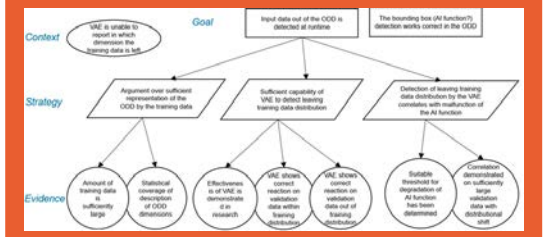
Identification: Identify potential causes of insufficiencies in the function (in KIA: “DNN related safety concerns”)

Quality metric: Introduce metrics or some form of judgment to argue that insufficiency was mitigated

Countermeasure: Develop methods to mitigate the insufficiencies

Argumentation: Argue that the residual risk associated with the causes has been reduced to a tolerable level

Formalisation: Create the evidence based safety argumentation in a Goal Structuring Notation



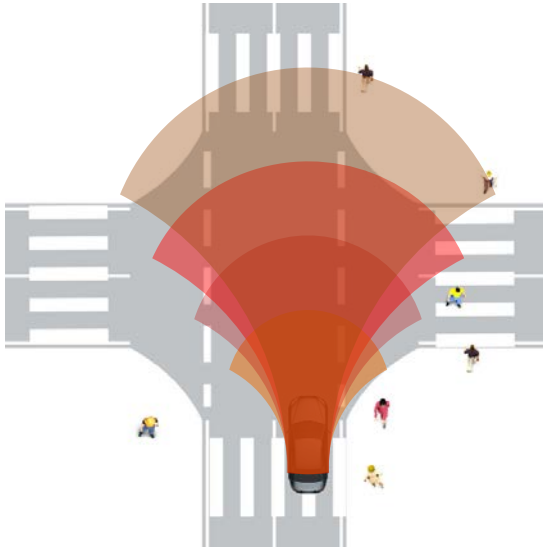
Structured safety verification
(Assurance Case)

Specification

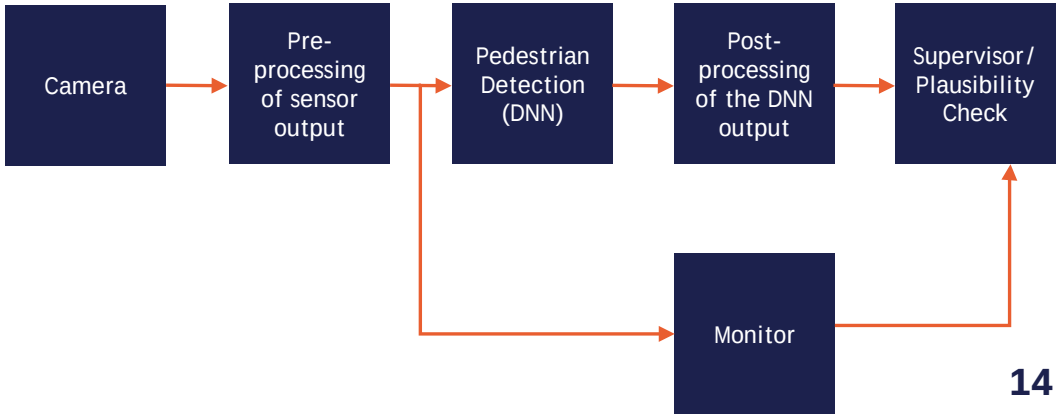
Function

Surrounding conditions

System Architecture



Source: Bosch



AI Function-Pedestrian detection

Function

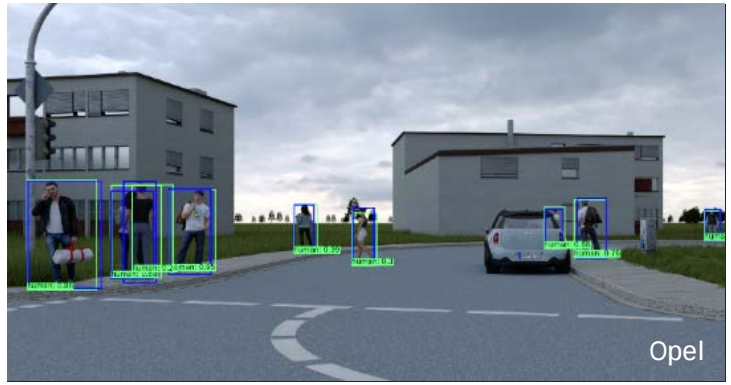
Sourrounding conditions

System architecture

AI Function

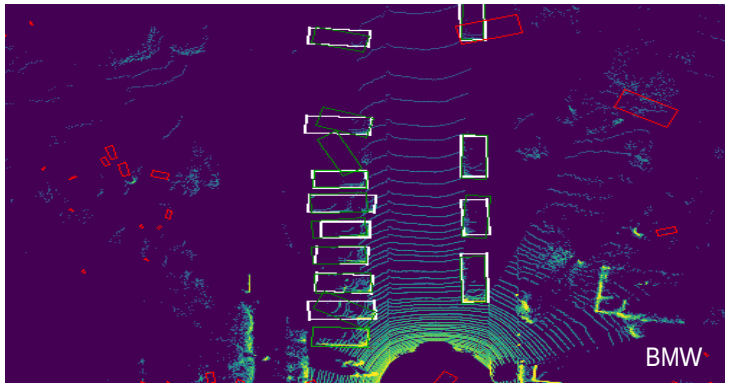
Semantic Segmentation

2D Bounding Box Detection

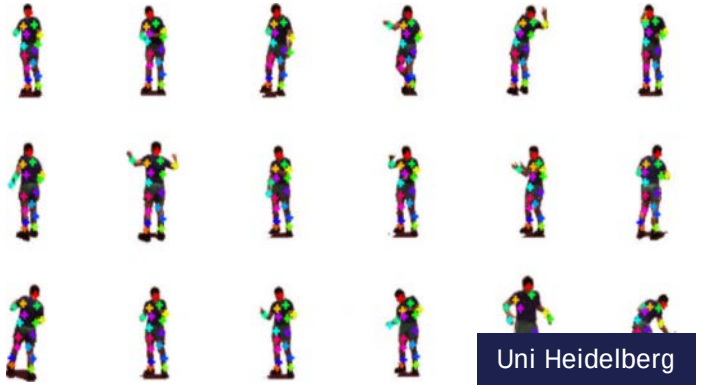


Instance Segmentation

3D Bounding Box Detection



3D Pose estimation

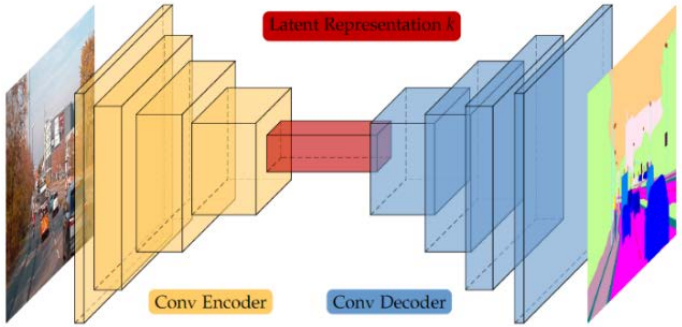


ML-Lifecycle

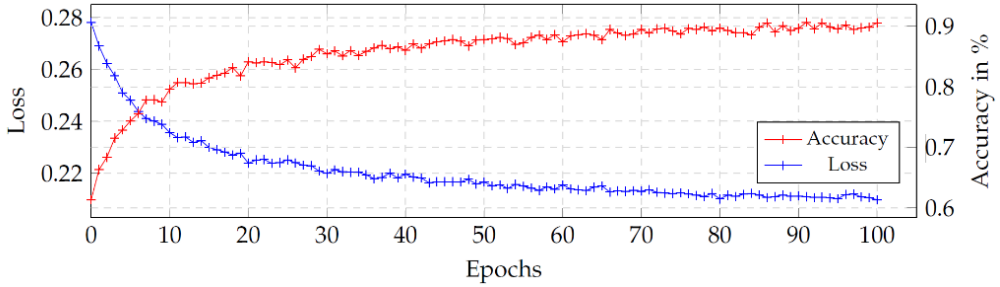
ML LIFECYCLE

- ML-Data
- Model design
- Pre-Processing
- Training
- Post-Processing
- Test/ V&V incl. Data
- Monitoring
- Maintenance

Source: BIT Technology Solutions



Volkswagen AG



Volkswagen AG

ML-Lifecycle-Validation data

ML LIFECYCLE

- ML-Data
- Model design
- Pre-Processing
- Training
- Post-Processing
- **Test/ V&V incl. Daten**
- Monitoring
- Maintenance

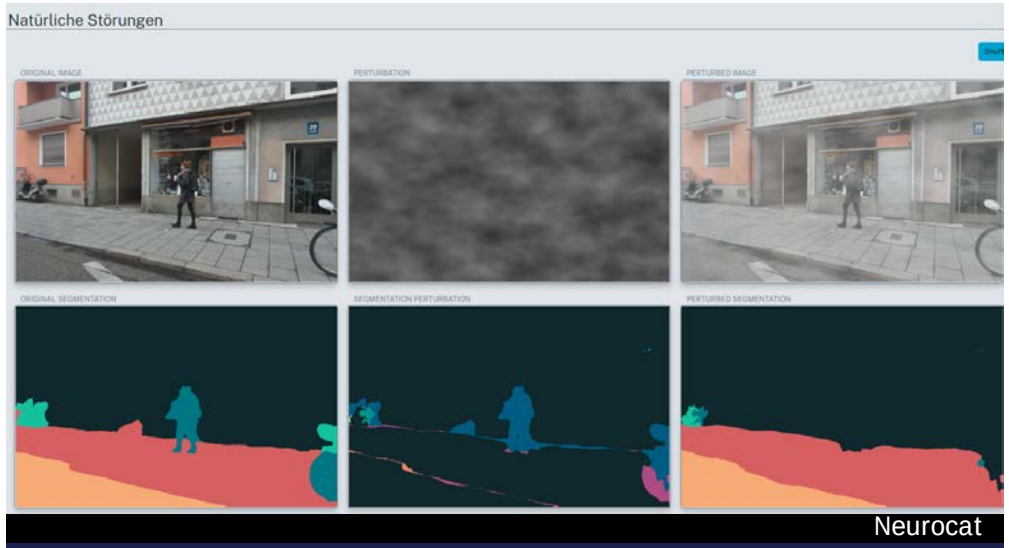
Identify, Measure and & Counteract „DNN-specific „Safety Concerns”

DNN-specific safety concern:

- False positive / negative: Pedestrian detection is incorrect resp. not robust enough

Method:

- Assessment of uncertainty: Stochastic evaluation of a multitude of model variations (Monte Carlo Dropout)



Neurocat

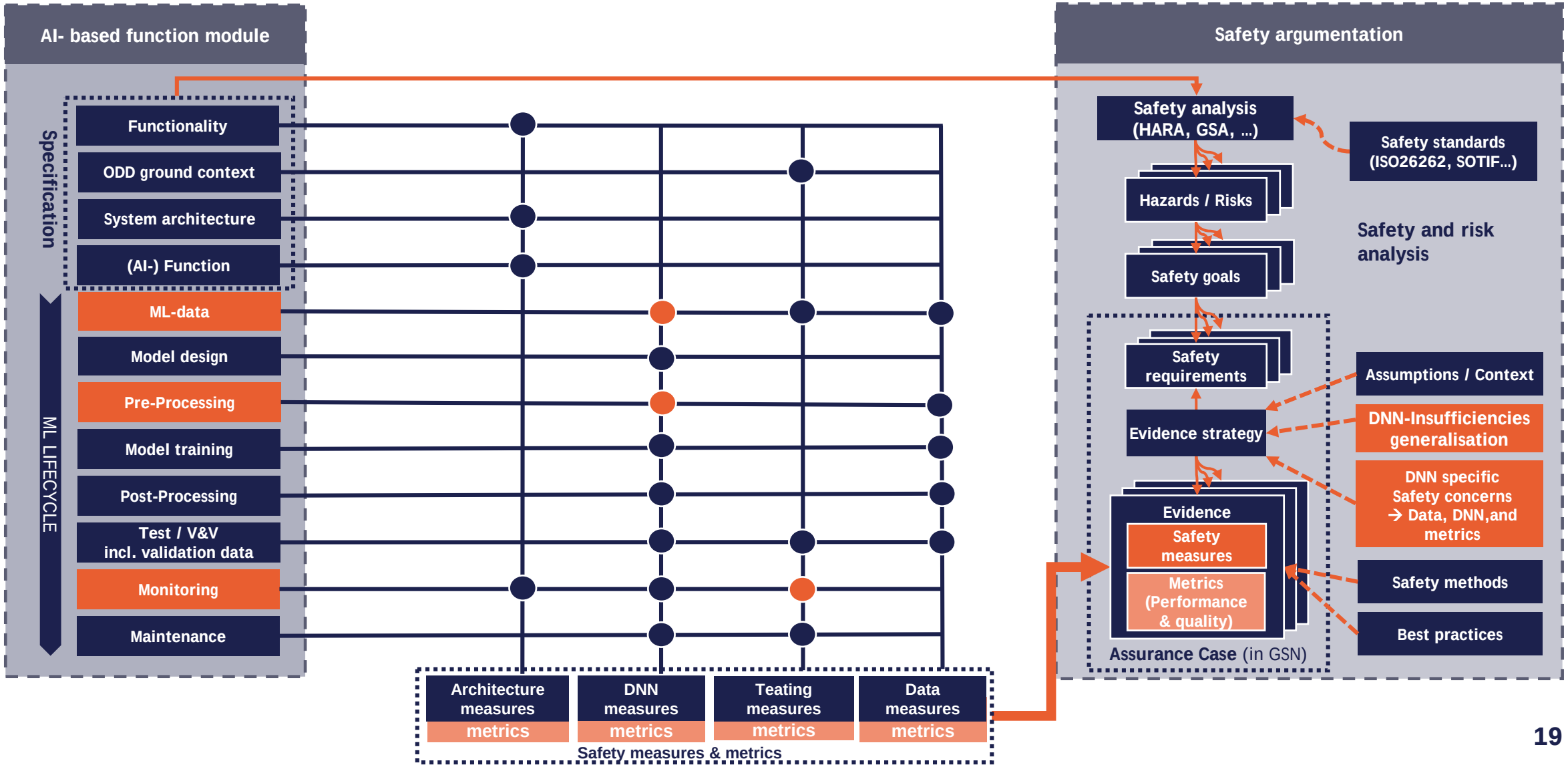
DNN-specific safety concern:

- Adversarial error: leads to overlooking (false negative)

Method:

- Systematic analysis of adversarial errors
- Methods to evaluate adversarial resilience
- Counter mechanisms during training or operation

AI Specific Evidence-Based Safety Argumentation



Parallel Activities – Available for other committees
Bringing our results „on the road“

- Establishing contacts with TÜV
  - VDTÜV Meeting with the BSI on AI-Safety
  - TÜV@Conference 2020 – Meet the Expert

- Collaboration with „Certified AI“
  - KI.NRW Initiative driven by Fraunhofer IAIS and the BSI (starting in 2021)
  - Projekt Letter of Intent for collaboration

- Contributions to „DIN-Roadmap AI“
  - Leading the „Quality and Certification“ Workgroup (Fraunhofer IAIS)
  - Contributing to the „Mobility“ Workgroup (BMW)

KI ABSICHERUNG

Safe AI for Automated Driving

Project coordination: Dr. Stephan Scholz, Volkswagen AG

Co-lead and scientific coordination: PD. Dr. Michael Mock, Fraunhofer IAIS

Email: ki-absicherung-konsortialfuehrung@eict.de

KI Absicherung ist ein Projekt der KI Familie und wurde aus der VDA Leitinitiative autonomes und vernetztes Fahren heraus entwickelt.

KI FAMILIE

Gefördert durch:

Bundesministerium für Wirtschaft und Energie

aufgrund eines Beschlusses des Deutschen Bundestages

www.ki-absicherung.vdali.de @KI_Familie KI Familie