



KI

ABSICHERUNG

Safe AI for Automated Driving

23.2.2021, Safetrain Project Meeting

Experiences from KI-Absicherung

Michael Mock (Fraunhofer IAIS)





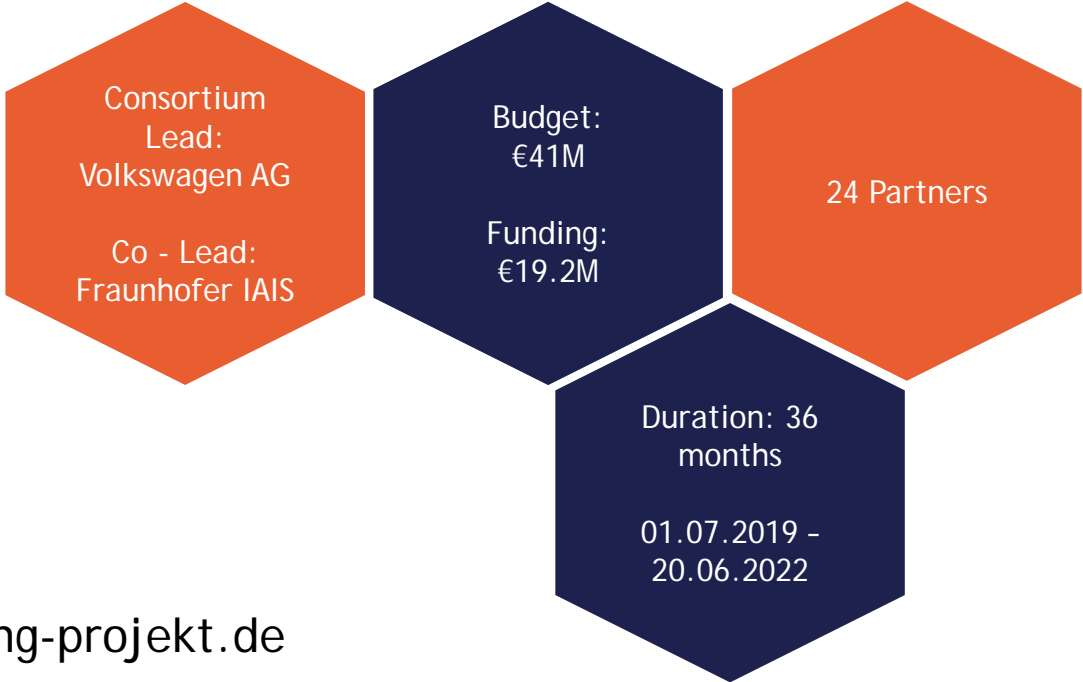
1

Project vision and goals



KI ABSICHERUNG

Safe AI for Automated Driving



»» <https://www.ki-absicherung-projekt.de>



External technology partners

https://www.ki-absicherung-projekt.de/fileadmin/KI_Absicherung/KI_Absicherung_Zwischenpraesentation/KI-A_Zwischenpraesentation_Einfuehrung_final.pdf



Making the safety of AI-based
function modules for highly
automated driving verifiable

KI ABSICHERUNG

Safe AI for Automated Driving

Pedestrian detection

[https://www.ki-absicherung-projekt.de/
veroeffentlichungen/ki-absicherung-zwischenpraesentation](https://www.ki-absicherung-projekt.de/veroeffentlichungen/ki-absicherung-zwischenpraesentation)

Challenge



AI Land



Pixabay

Promising new technology with unimagined possibilities

Established safety processes cannot be applied



Safety Land



Pixabay

Safe, trustworthy driving function



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



2

Methodology for a Safety Argumentation



An Integrated Approach to a Safety Argumentation for AI-based Perception Functions in Automated Driving

Michael Mock¹, Stephan Scholz², Frédéric Blank³, Fabian Hüger², Andreas Rohatschek³, Loren Schwarz⁴, Thomas Stauner⁴

¹ Fraunhofer IAIS, 53757 St. Augustin, Germany

² Volkswagen AG, 38440 Wolfsburg, Germany

³ Robert Bosch GmbH, 70469 Stuttgart, Germany

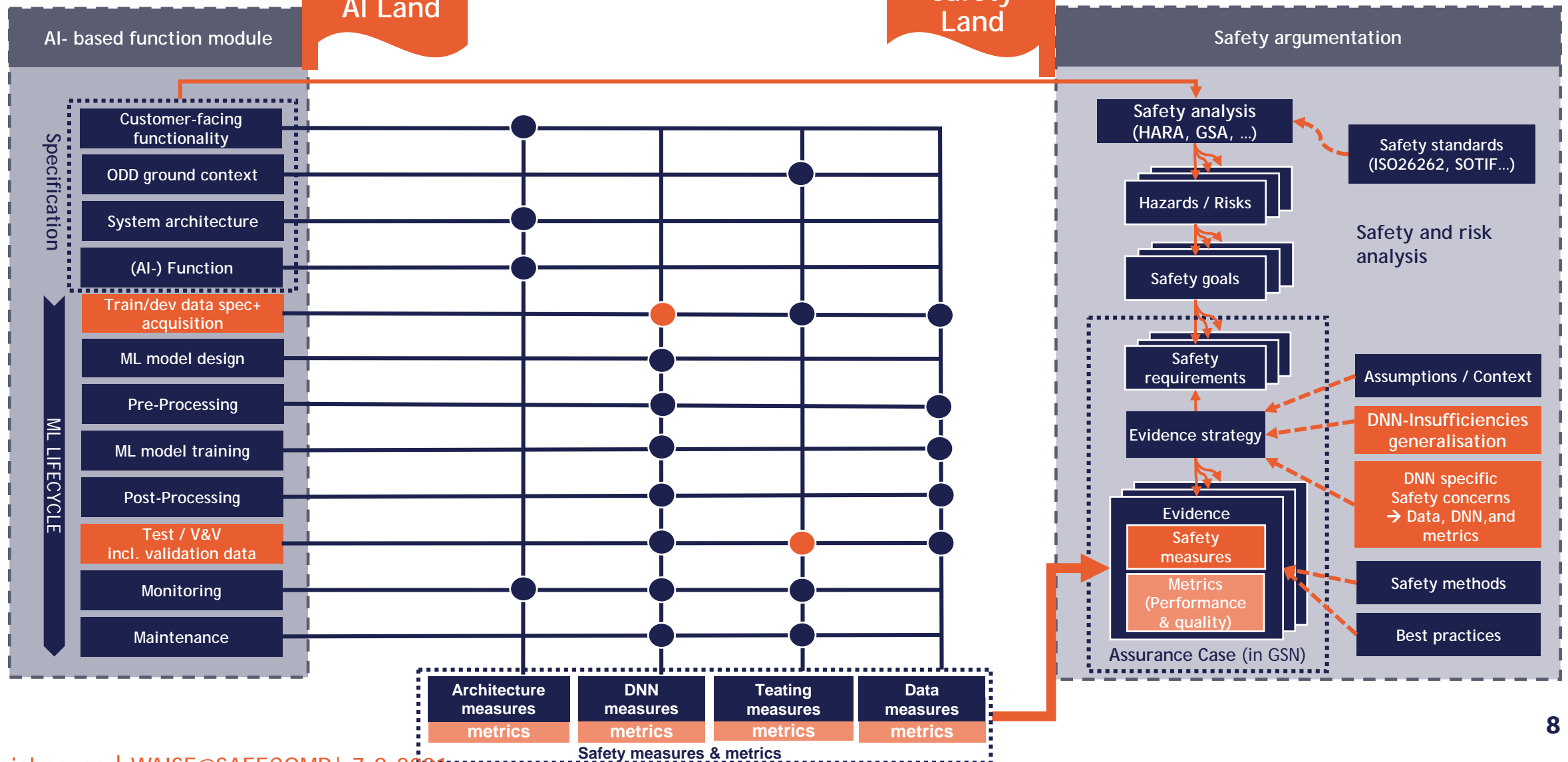
⁴ BMW AG, 80809 München, Germany

An Integrated Approach to a Safety Argumentation for AI-Based Perception Functions in Automated Driving, [Mock, Michael](#); [Scholz, Stefan](#); Blank, Frederik; [Hüger, Fabian](#); Rohatschek, Andreas; Schwarz, Loren; Stauner, Thomas, Computer Safety, Reliability, and Security. SAFECOMP 2021 Workshops

<https://ki-familie.vdali.de/ki-newsletter-nr-1/ki-absicherung-proof-of-project-concept>

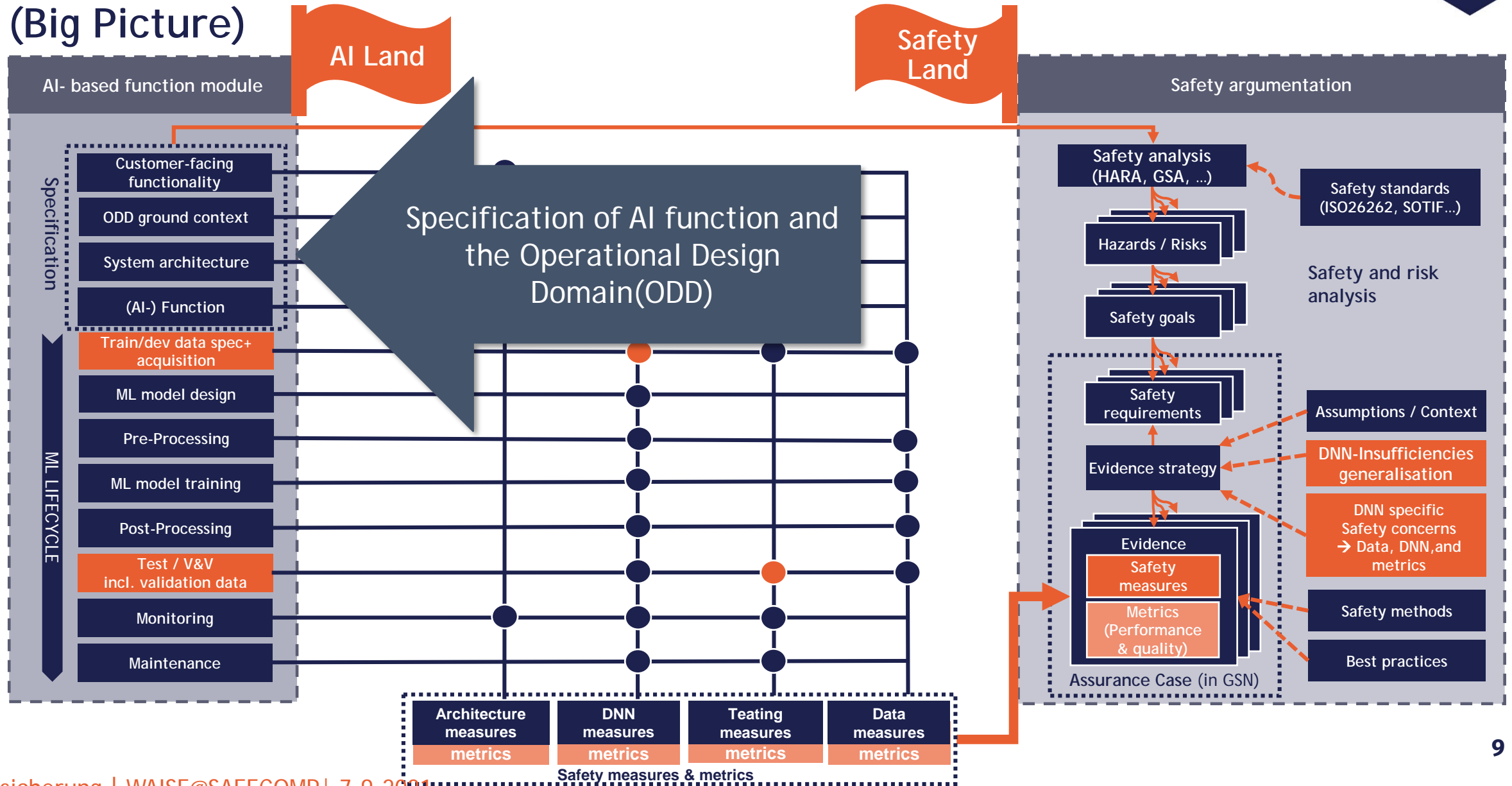
https://www.ki-absicherung-projekt.de/fileadmin/KI_Absicherung/KI_Absicherung_Zwischenpraesentation/KI-A_Zwischenpraesentation_PoPC_final.pdf

Project Approach to a Safety Argumentation for AI-based Functions (Big Picture)

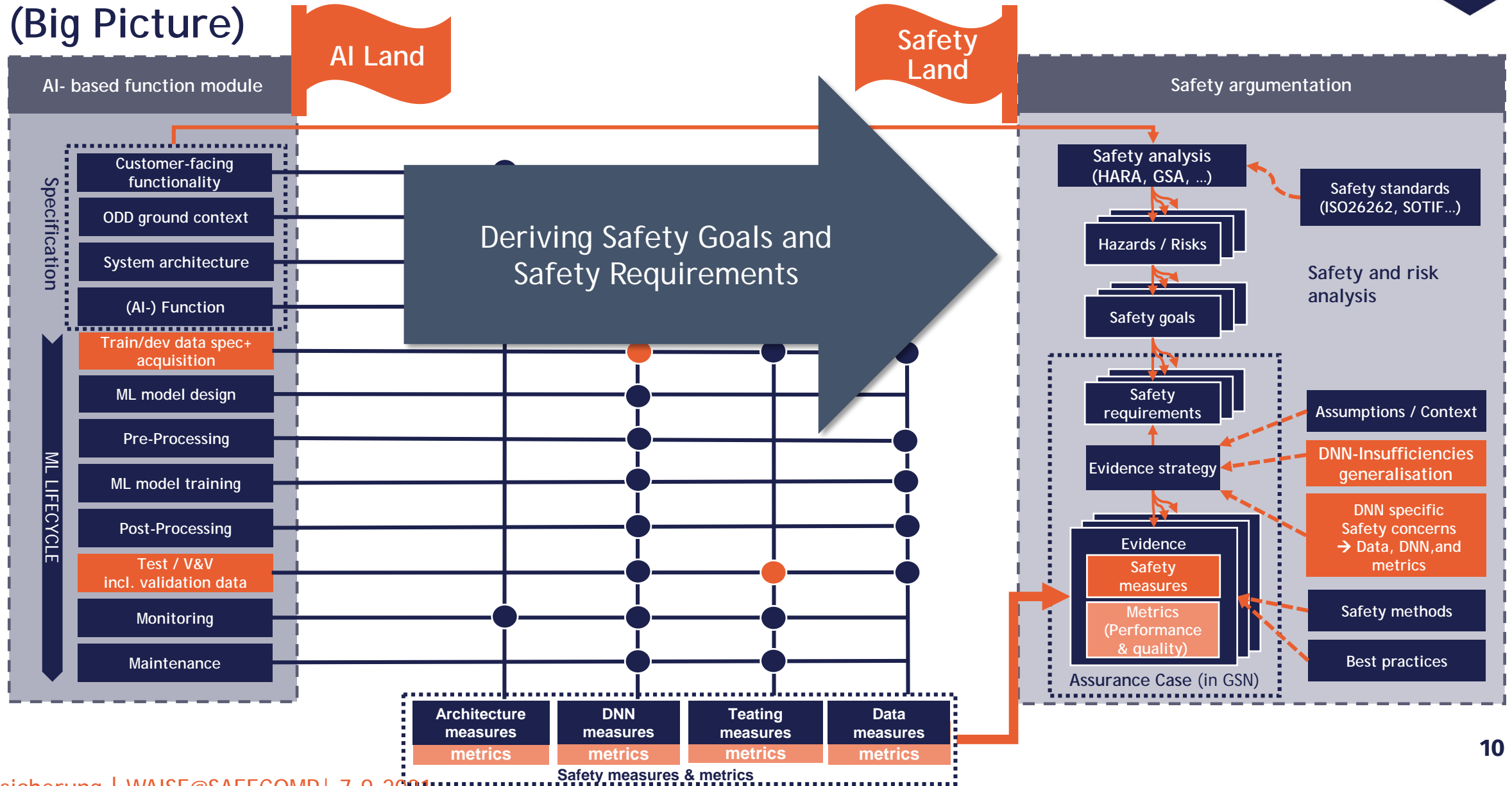




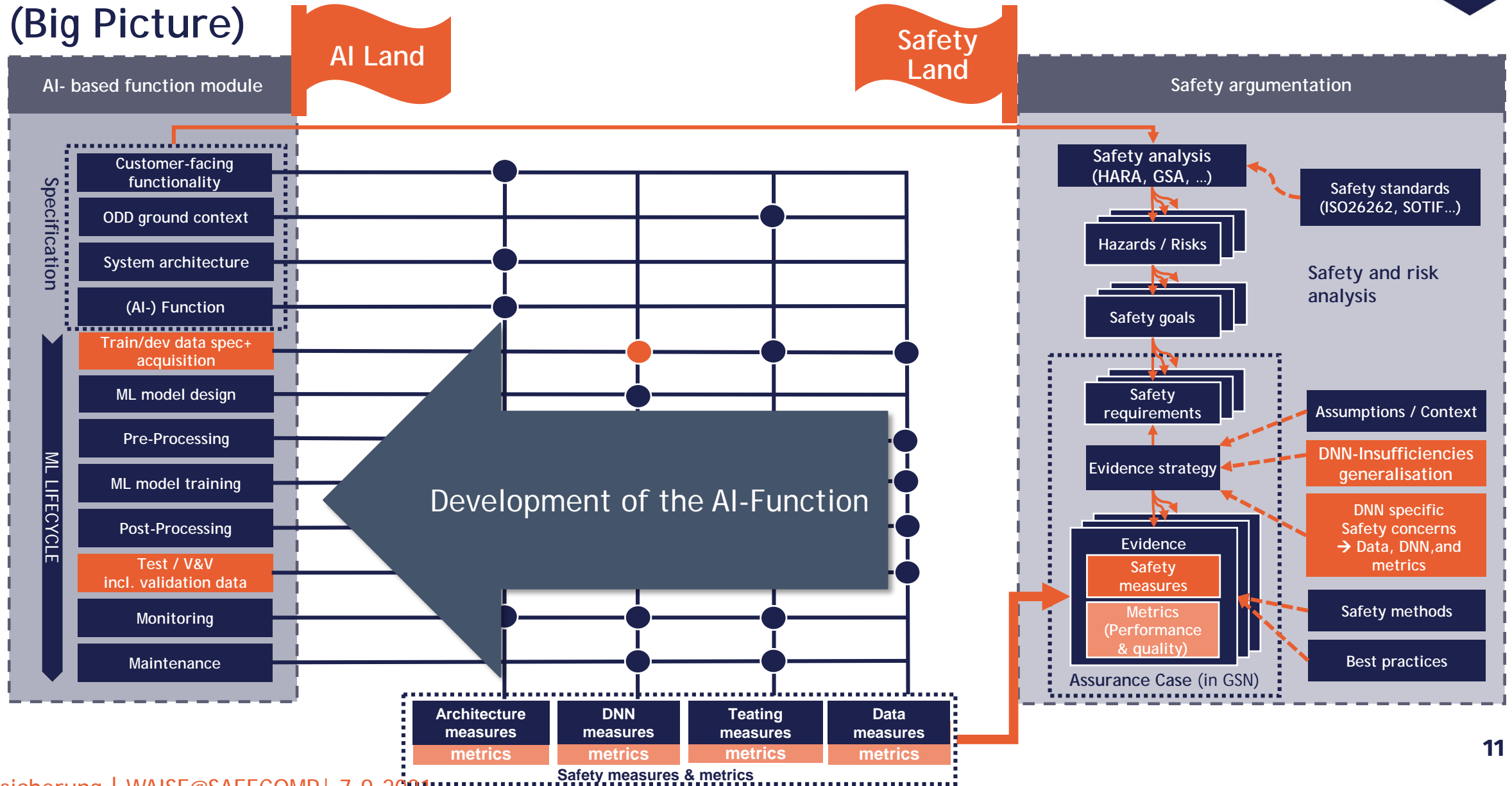
Project Approach to a Safety Argumentation for AI-based Functions (Big Picture)



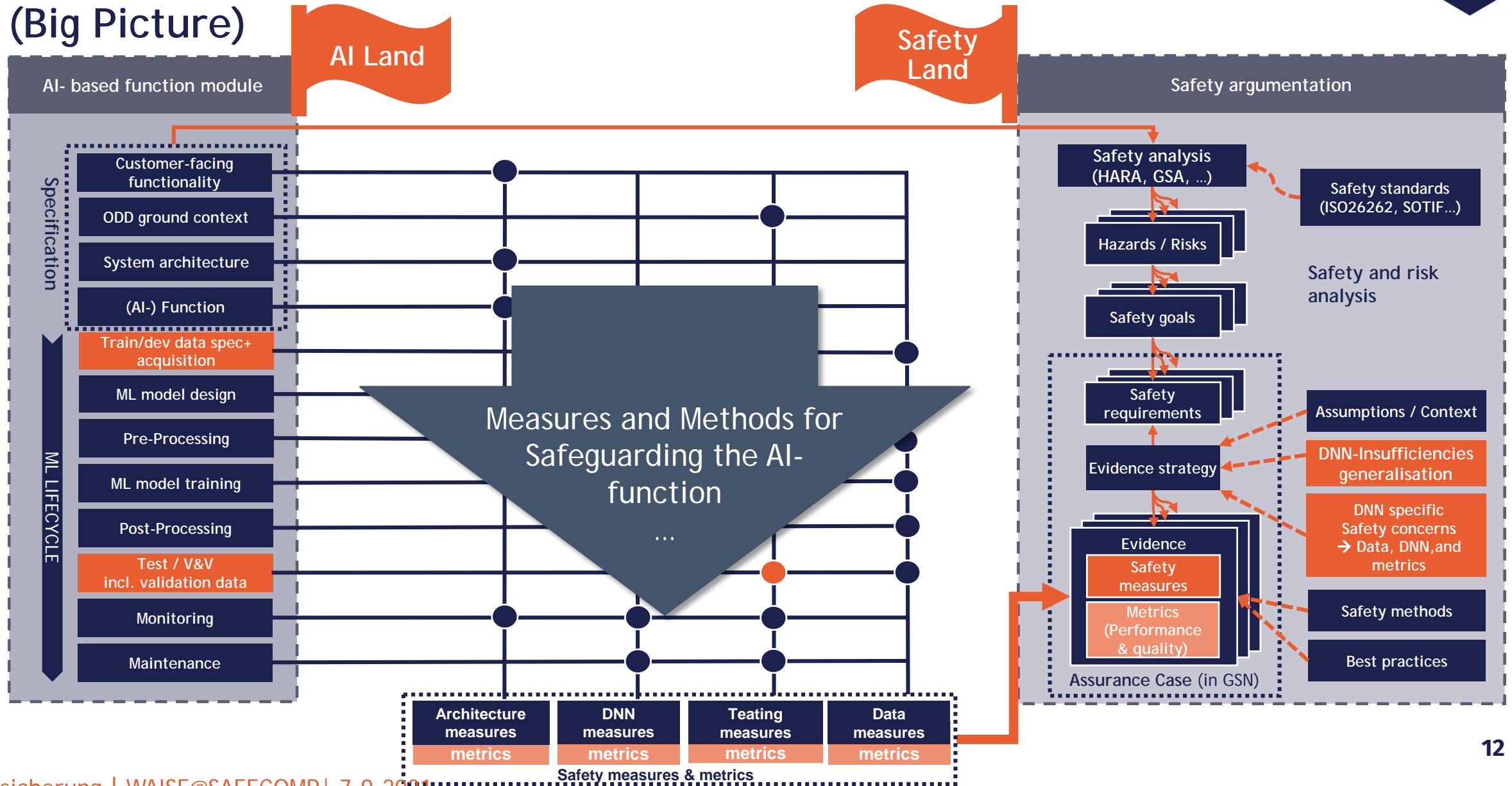
Project Approach to a Safety Argumentation for AI-based Functions (Big Picture)



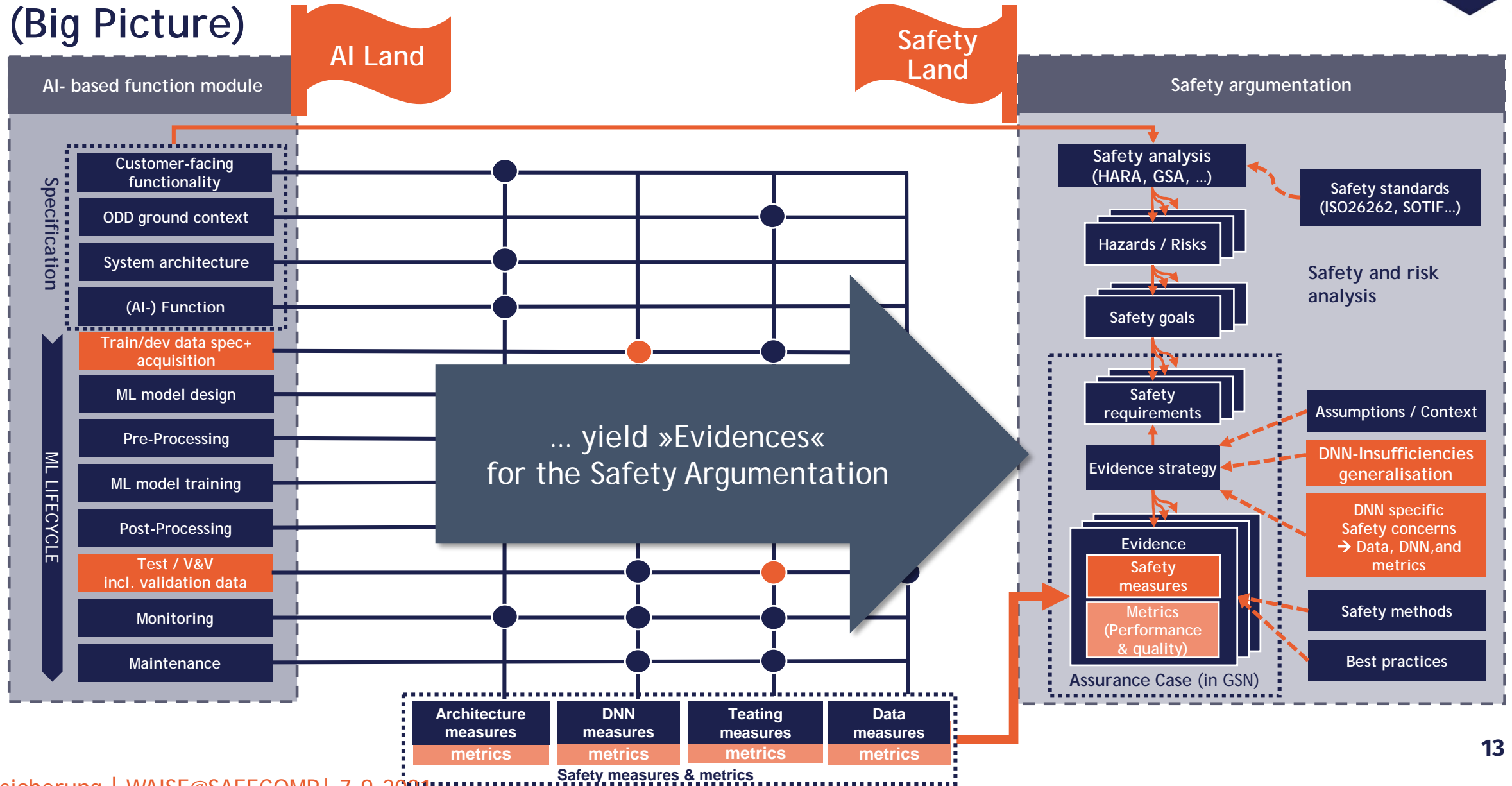
Project Approach to a Safety Argumentation for AI-based Functions (Big Picture)



Project Approach to a Safety Argumentation for AI-based Functions (Big Picture)



Project Approach to a Safety Argumentation for AI-based Functions (Big Picture)






3

Assurance Case Definition



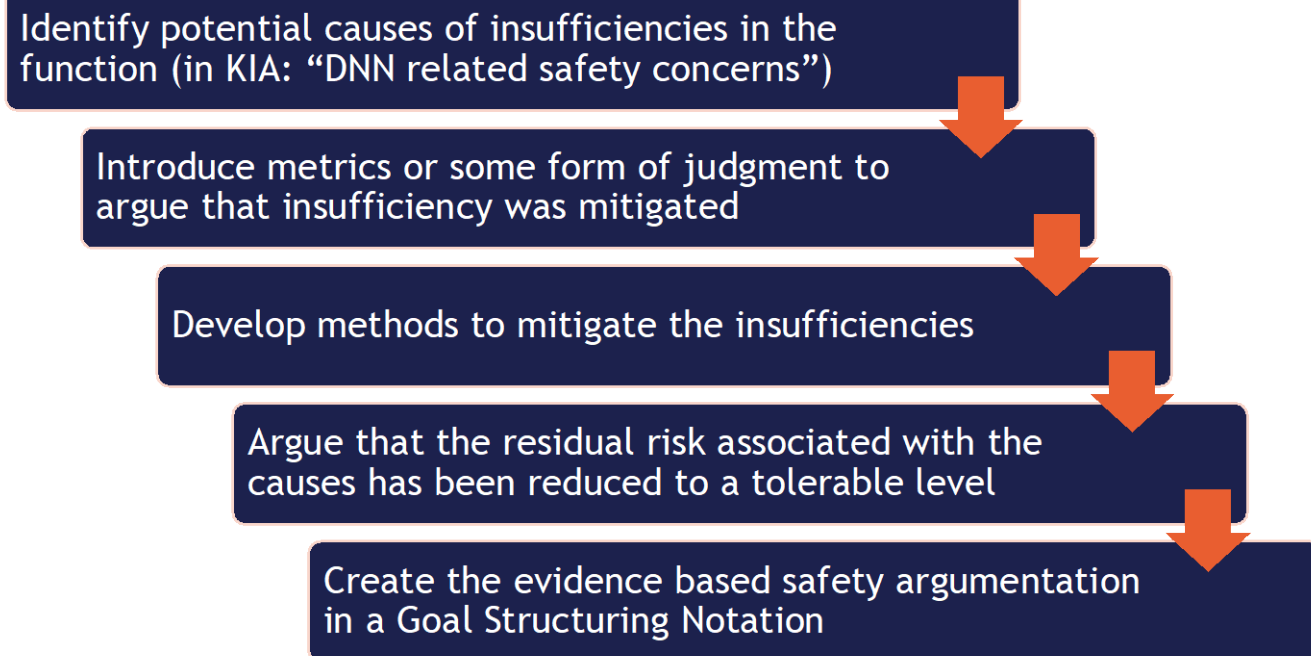
<p>FI-1 INSUFFICIENT GENERALIZATION CAPABILITY Wrong outputs by an AI-based function that was trained on a limited database. Erroneous input to output mapping or wrong approximation.</p>	<p>SC-2.2 INADEQUATE SEPARATION OF TEST AND TRAINING DATA Test data might be correlated to training data which might induce overfitting on test data.</p>	 <p>Work-in-progress</p> <p>Based on:</p> <p>O. Willers, S. Sudholt, S. Raafatnia, S. Abrecht: Safety Concerns and Mitigation Approaches Regarding the Use of Deep Learning In Safety-Critical Perception Tasks</p> <p>T. Sämann, P. Schlicht, F. Hüger: Strategy to Increase the Safety of a DNN-based Perception for HAD Systems</p> <p>G. Schwalbe, B. Knie, T. Sämann, T. Dobberphul, L. Gauerhof, S., V. Rocco: Structuring the Safety Argumentation for Deep Neural Network Based Perception In Automotive Applications</p> <div data-bbox="1984 678 2147 1112"> <p>Functional Insufficiencies</p> <p>DNN-characteristics-related concerns</p> <p>Data-related concerns</p> <p>Metric-related concerns</p> </div>
<p>SC-1.1 UNRELIABLE CONFIDENCE INFORMATION DNNs tend to be overconfident in their predictions under certain conditions or in general outputting unreliable confidence information.</p>	<p>SC-2.3 DEPENDENCE ON LABELLING QUALITY Labelling quality can directly affect the resulting model performance. Moreover, due to missing labelling quality, evaluation results might be misleading.</p>	
<p>SC-1.2 BRITTLINESS OF DNNs Non-robustness against common perturbations such as noise or certain weather conditions as well as targeted perturbations known as adversarial examples</p>	<p>SC-2.3.1 MISSING LABEL DETAILS OR META-LABELS Missing meta-labels or label details possibly leads to improper data selection or insufficient training objectives.</p>	
<p>SC-1.2.1 LACK OF TEMPORAL STABILITY Detection results rapidly changing in time whereas little change occurs in the ground truth</p>	<p>SC-2.4 SPECIFICATION OF THE ODD An incomplete or incorrect ODD specification leads to incomplete data records for training and testing.</p>	
<p>SC-1.3 INCOMPREHENSIBLE BEHAVIOUR Inability to explain exactly how DNNs come to a decision.</p>	<p>SC-2.5 DISTRIBUTIONAL SHIFT OVER TIME A DNN is trained and tested at a certain point in time. Changes will occur naturally and therefore can potentially harm the performance of DNNs.</p>	
<p>SC-1.4 INSUFFICIENT PLAUSIBILITY AI based functions usually lack basic plausibility checks, which are intended to identify detections of the perception function that violate physical laws.</p>	<p>SC-2.6 UNKNOWN BEHAVIOUR IN RARE CRITICAL SITUATIONS The long tail problem describes the fact that there exists an enormous amount of possibly safety-critical street scenes that have a low occurrence probability.</p>	
<p>SC-2.1 DATA DISTRIBUTION IS NOT A GOOD APPROXIMATION OF REAL WORLD The distribution of data used in the development should be a valid approximation of the ODD in the real world.</p>	<p>SC-3.1 SAFETY-AWARE METRICS Some state-of-the-art metrics only evaluate the average performance of DNNs. Safety-aware metrics are required to sophisticatedly evaluate the performance of DNNs.</p>	

Assurance Strategy for AI

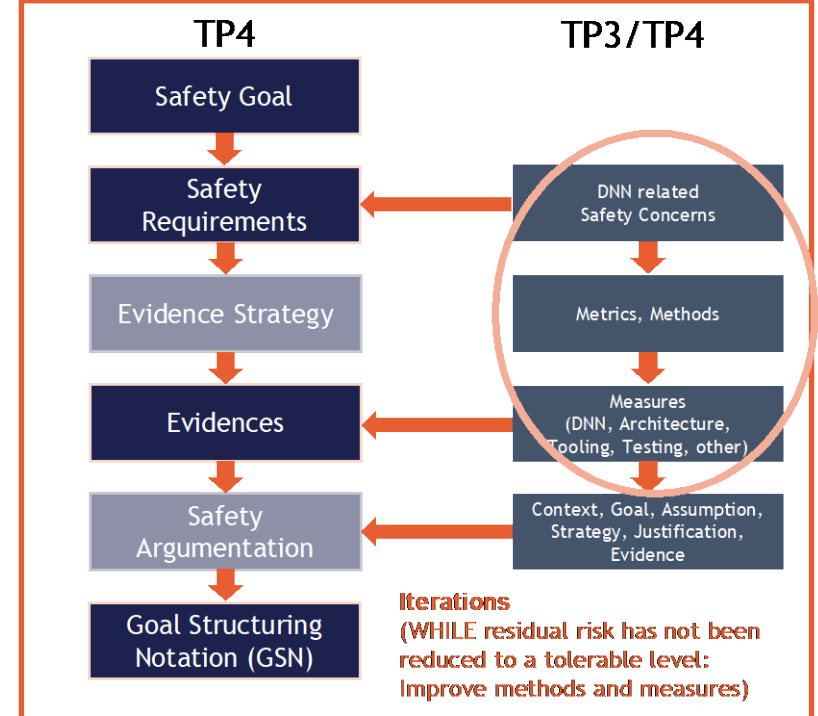
Assurance Case Strategy for AI-based Perception



The path to an evidence based Safety Argumentation



Assurance Case Development

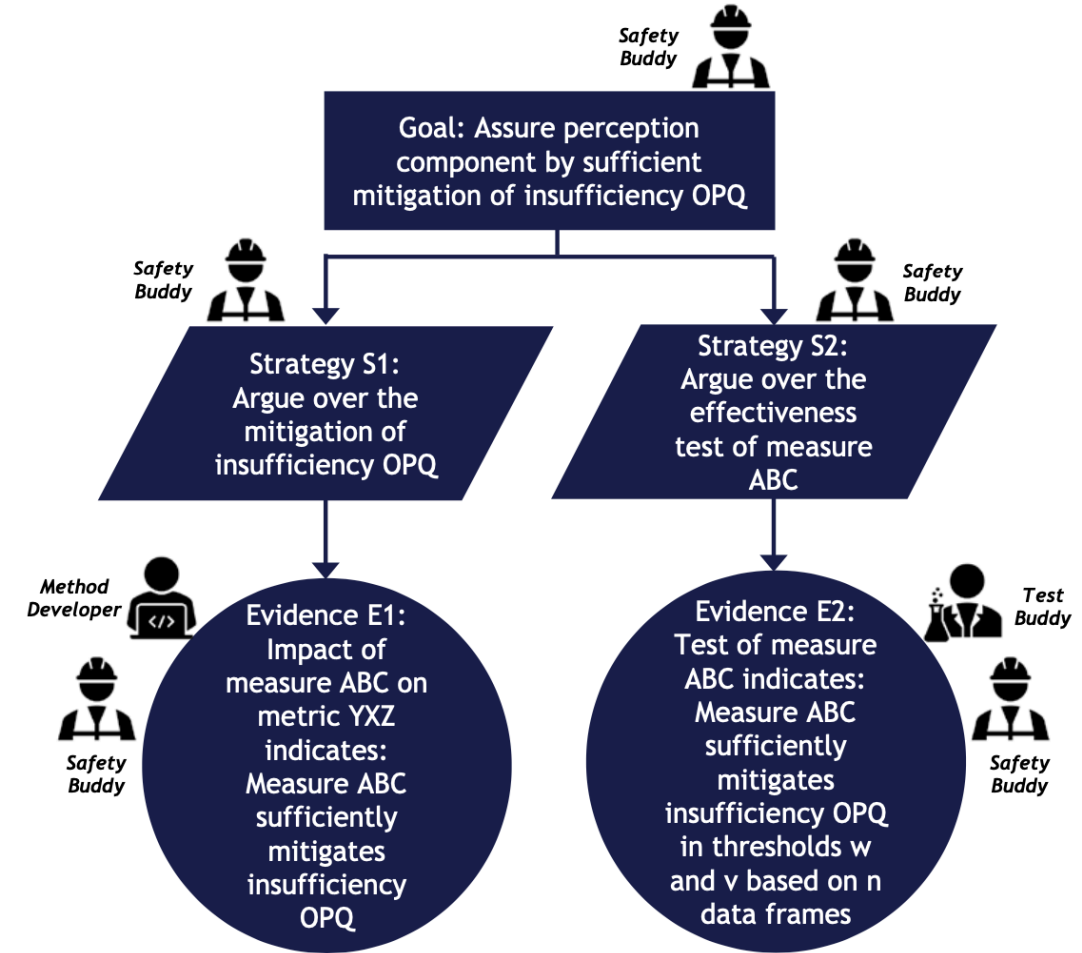
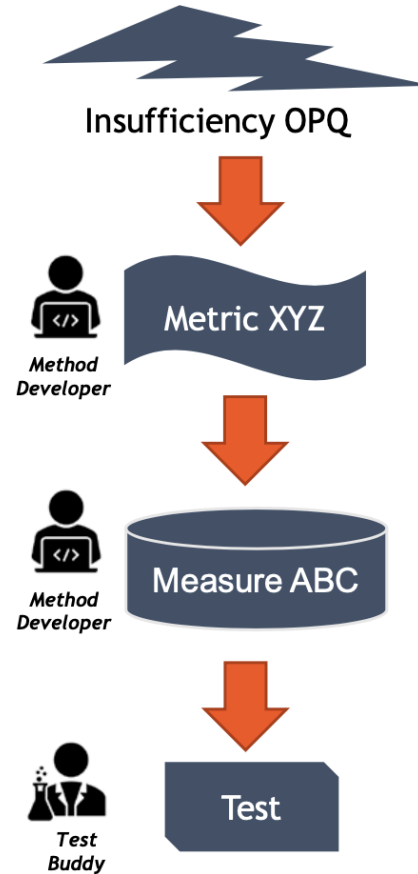
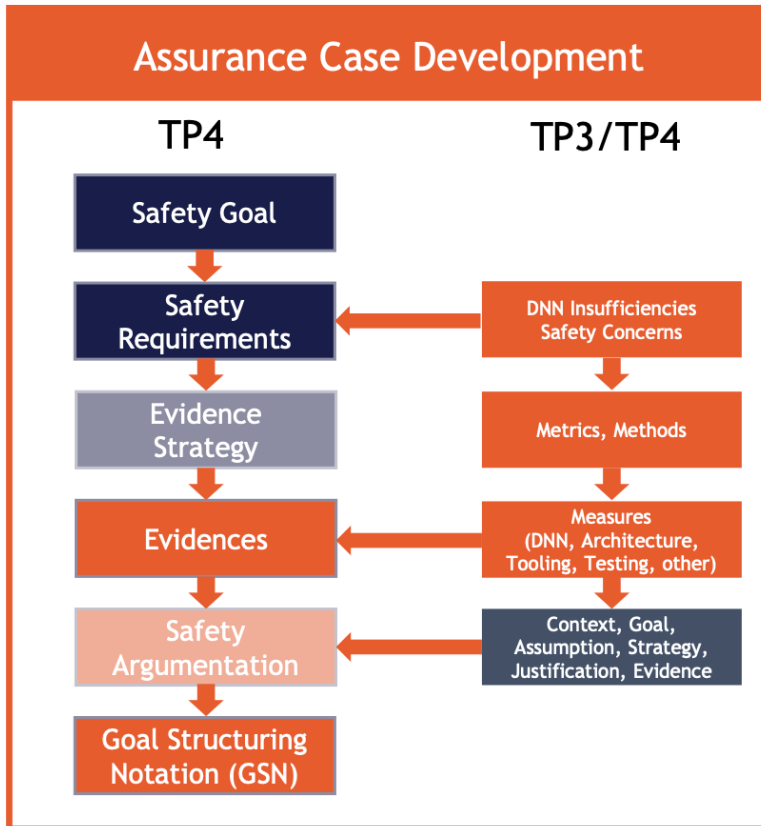


What are the causes of insufficiencies and what sources of evidence can be used to make this argument?



Assurance Strategy for AI

How to create Evidences from Methods and Tests



Interaction of Method Developer, Safety Buddy and Test Buddy leads to evidences for the safety argumentation

https://www.ki-absicherung-projekt.de/fileadmin/KI_Absicherung/KI_Absicherung_Zwischenpraesentation/KI-A_Zwischenpra__sentation_TP4_final.pdf



4

ODD & Data Definition

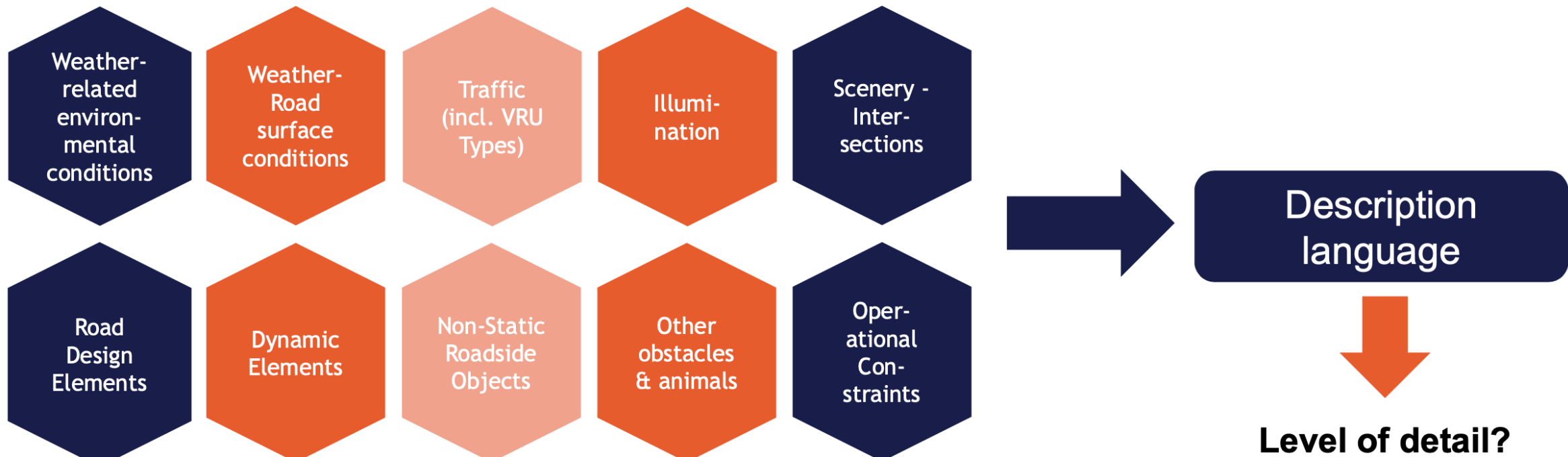


Assurance Strategy for AI

Existing Taxonomy as of PAS 1883:2020



- An ODD describes / specifies operating conditions under which a given driving automation system or feature is specifically designed to function [...]
- Taxonomy and Definitions for Terms Related to Driving Automation Systems (examples)



https://www.ki-absicherung-projekt.de/fileadmin/KI_Absicherung/KI_Absicherung_Zwischenpraesentation/KI-A_Zwischenpra__sentation_TP4_final.pdf



Assurance Strategy for AI

“No relevant pedestrian shall be overlooked within defined ODD*”

A description language & input space modeling is needed to...

Complexity of language



be able to describe / **specify operating conditions** (and edges of ODD*) as of PAS 1883:2020 and others



systematically capture important knowledge and describe the (expected) **key input space dimensions** and their **possible variations** having an influence on the functional performance of a DNN-based function (→ Zwicky Boxes & Ontology)



perform training and assurance **data coverage estimations** for data driven AI-based systems



describe **Corner cases / rare critical situations** to be considered in training / test data sets



for synthetic perception data production & meta-data: describe data dimensions that should be varied & **incrementally generate new data** by analyzing coverage and generating missing combinations

DNN-specific Safety Concerns (examples)

Data distribution is not a good approximation to real world

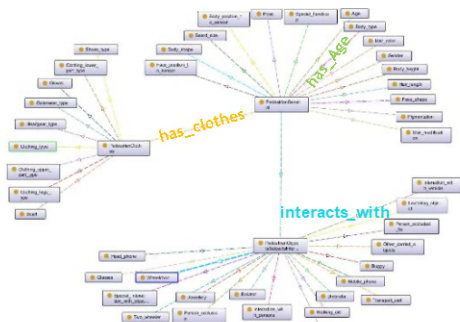
Unknown behavior in rare critical situations



Assurance Strategy for AI

Data representations of the data input space aligned to ontology

Ontology Graph (Relations)



Excerpt of ontology

Asset & Object descriptions for data analytics



Source: Mackevision

- PedestrianGeneral
- PedestrianGeneral:Age "adult"
- PedestrianGeneral:Gender "female"
- PedestrianGeneral:Body shape "normal"
- PedestrianGeneral:Body type
- PedestrianGeneral:Body height "160cm-200cm"
- PedestrianGeneral:Pigmentation "low"
- PedestrianGeneral:Skin modification "no"
- PedestrianGeneral:Hair length "short"
- PedestrianGeneral:Hair color
- PedestrianGeneral:Beard size "no"
- PedestrianGeneral:Special handicap "nothing"

Systematic Combination of variations

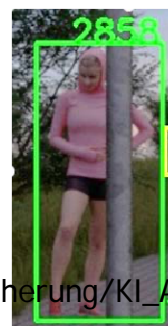
Dimension	Person1	Person2	Person3	...
Age	Child	Teenager	Adult	
Gender	Male	Female	Male	
Body height	80-120 cm	120-160 cm	160-200 cm	
Pose	Running	Lying	Walking	
Pedestrian Location	Middle of street	Left side walk	Right side walk	
...	

Representations of variations

DAYTIME	morning	day	evening	night	
HAZE/FOG	no		yes		
STREET CONDITION	dry	wet	icy	snow	broken
SKY	cloudy	no		clear	
RAIN	no		yes		
REFLECTION ON ROAD	no		yes		
SHADOW ON ROAD	no		yes		
VRU TYPE	adult		child		
VRU POSE	pedestrian	jogger	cyclist		
VRU CONTRAST TO BG	low		high		

Zwicky Box - Discretized variations of important dimensions

Object Annotations for DNN-Training



- Occlusion_level: medium
- Occluded_body_part: arm
- Occlusion_object: lamp

Source: BIT Technology Solutions

Systematically identify and describe the (known / expected) **key input space dimensions** and their **possible variations & combinations** having an influence on the functional performance of a DNN-based function



KI
ABSICHERUNG
Safe AI for Automated Driving

PD Dr. Michael Mock, Fraunhofer IAIS
Consortium Co-Lead and Scientific Coordinator
michael.mock@iais.fraunhofer.de

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

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



KI
FAMILIE

Supported by:



on the basis of a decision
by the German Bundestag