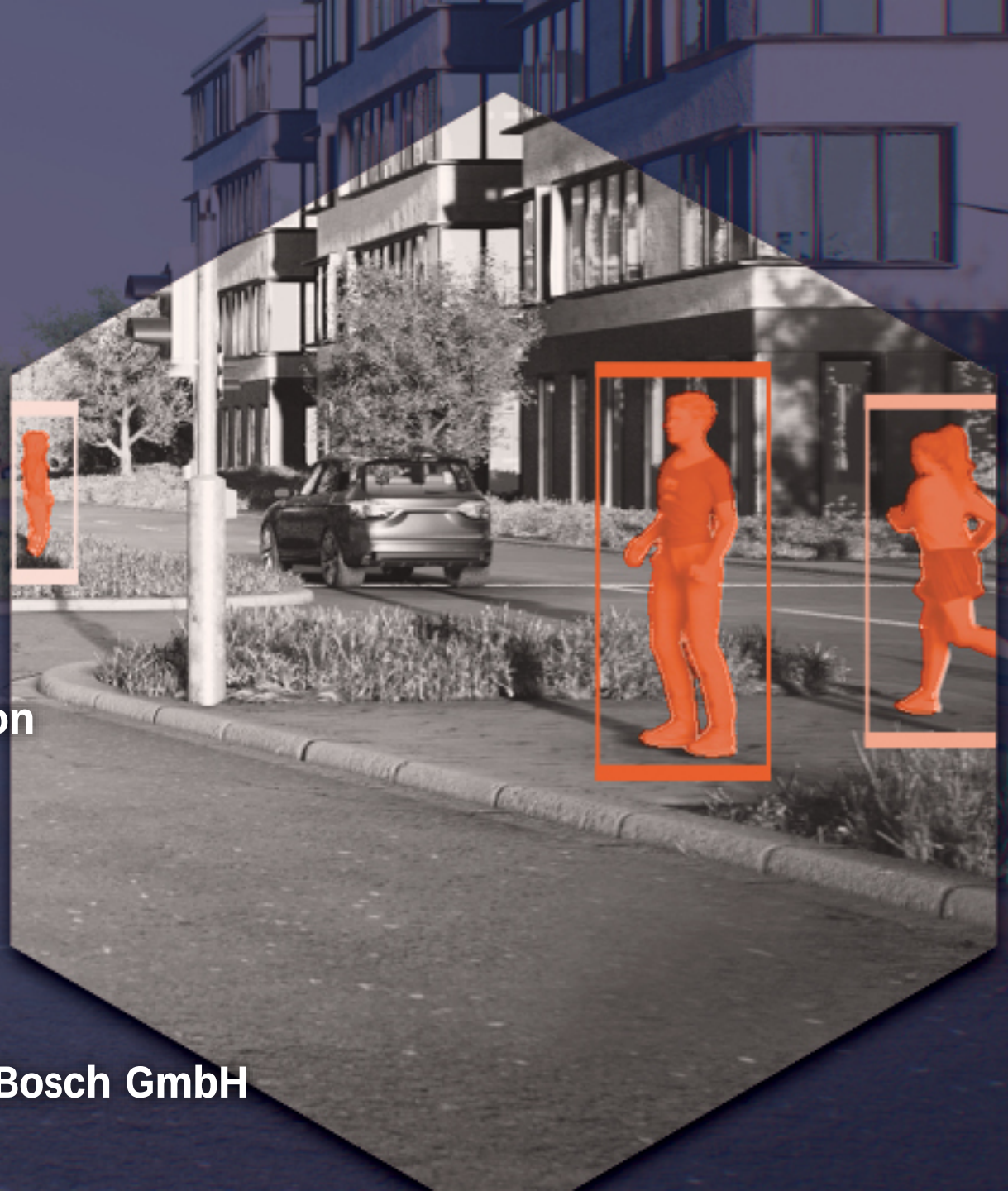
KI ABSICHERUNG

Safe AI for Automated Driving



11th March 2021, Online, Interim Presentation

Assurance Strategy for AI

Frédérik Blank, Andreas Rohatschek, Robert Bosch GmbH

Assurance Strategy for AI

Work package „Assurance Strategy for AI“ located in the (simplified) value chain

AP4.1

Structuring of Data input space

- How to describe the context in which the system is operated?
- What level of detail is needed?

TP2

Synth. Data specification & generation

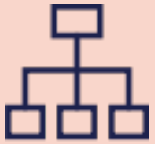
TP1

DNN-Development (incl. AI function specification)

TP3

DNN assurance methods & measures development

AP4.2/AP4.3



Safety Analysis, Structure of safety argumentation & Evidence strategies

- What are the safety goals and safety requirements to argument?
- How to build an evidence based safety argumen-tation and visualize it using a GSN?

AP4.4/AP4.5

Test methods & test strategy

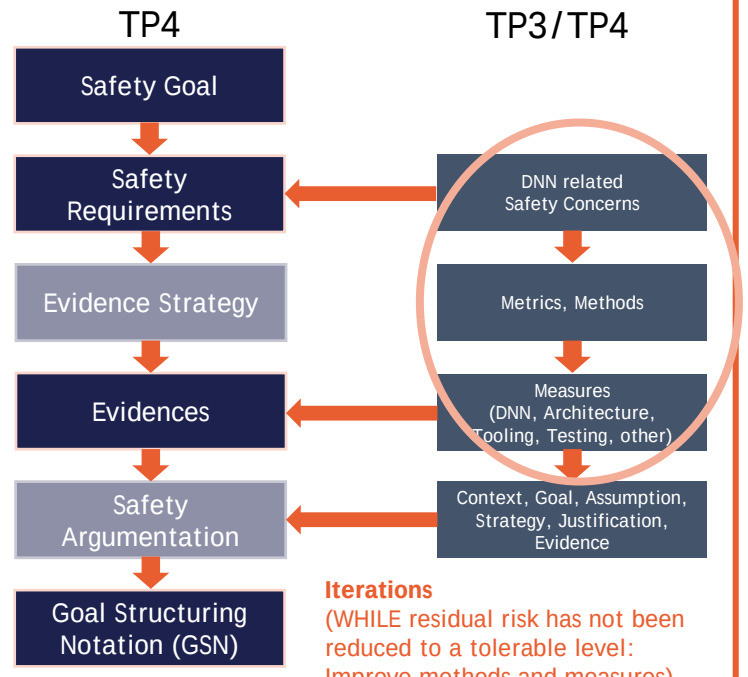
- To what extent does the training & test data cover the ODD?
- How robust & effective is the safety argumentation?

Assurance Strategy for AI
Assurance Case Strategy for AI-based Perception

The path to an evidence based Safety Argumentation

- Identify potential causes of insufficiencies in the function (in KIA: “DNN related safety concerns”)
- Introduce metrics or some form of judgment to argue that insufficiency was mitigated
- Develop methods to mitigate the insufficiencies
- Argue that the residual risk associated with the causes has been reduced to a tolerable level
- Create the evidence based safety argumentation in a Goal Structuring Notation

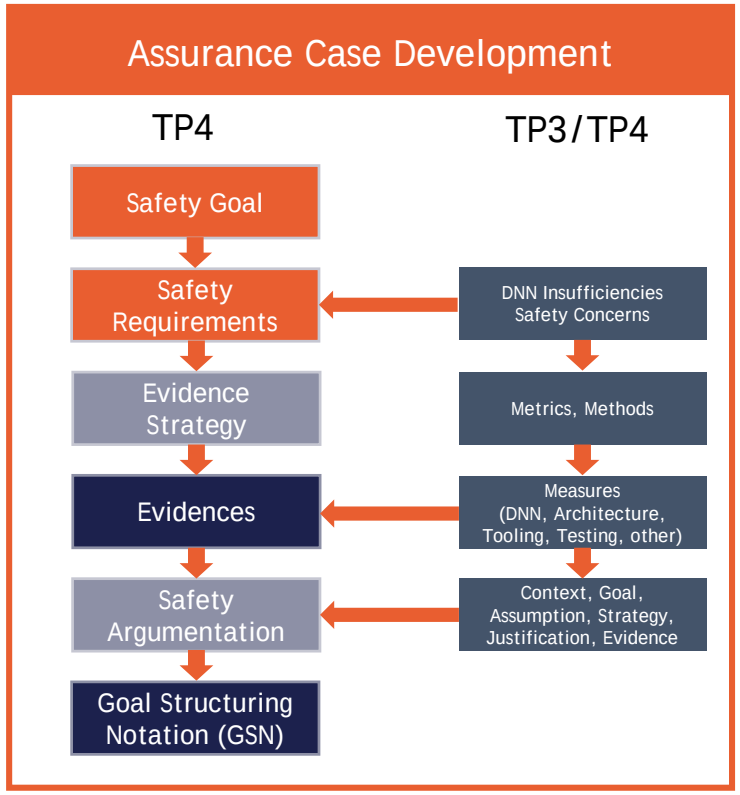
Assurance Case Development



Iterations
(WHILE residual risk has not been reduced to a tolerable level: Improve methods and measures)

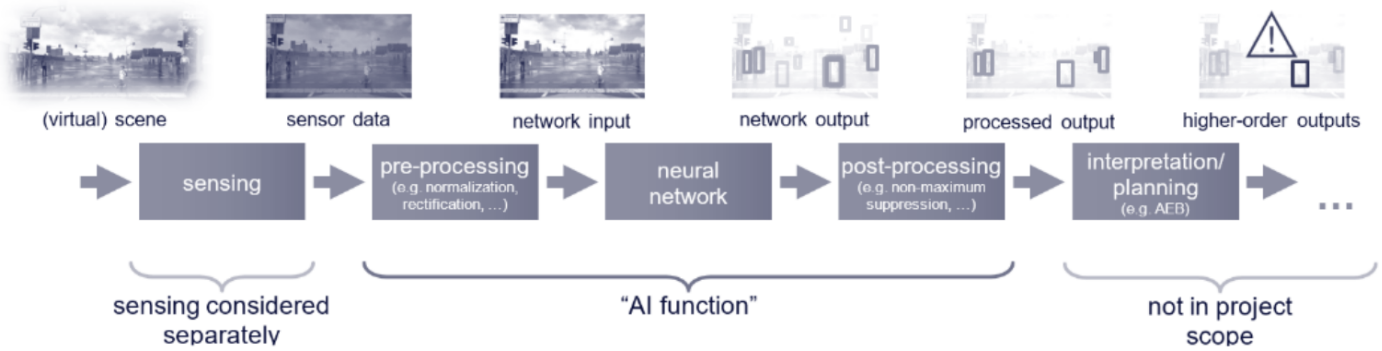
What are the causes of insufficiencies and what sources of evidence can be used to make this argument?

Assurance Strategy for AI
Safety Goal and Safety Requirements



Safety goal for the system:
"Avoid collisions with pedestrians"

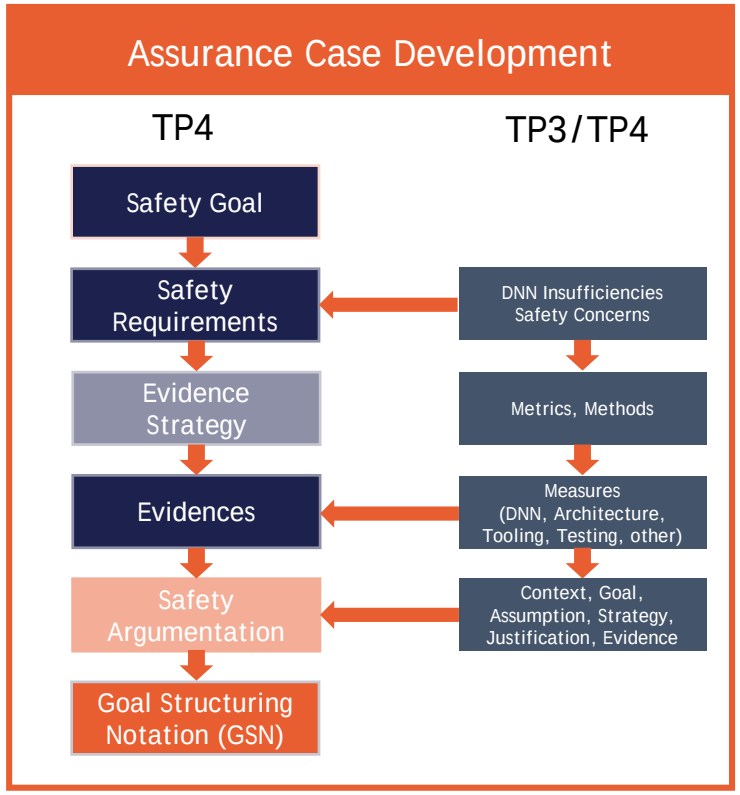
Safety goal for perception within the ODD*:
"No relevant pedestrian shall be overlooked"



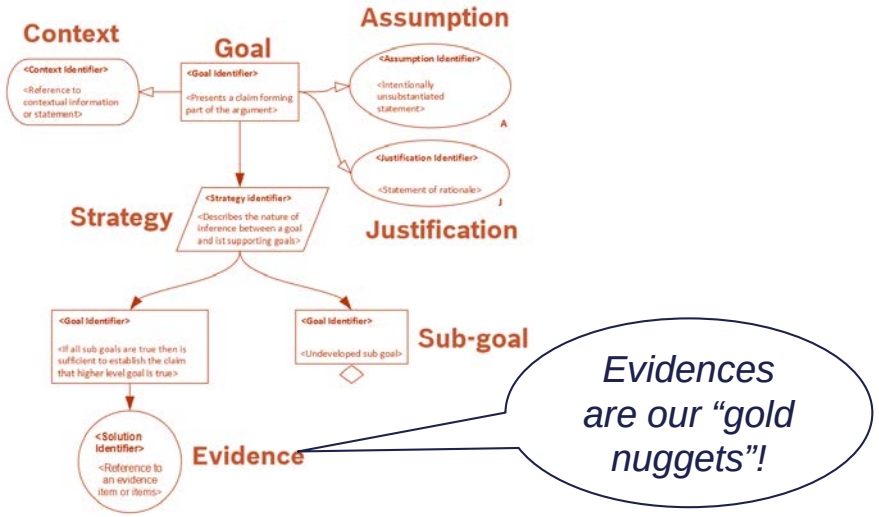
What safety requirements can be derived from the safety goal considering hazards and risks?

*) ODD = Operational Design Domain

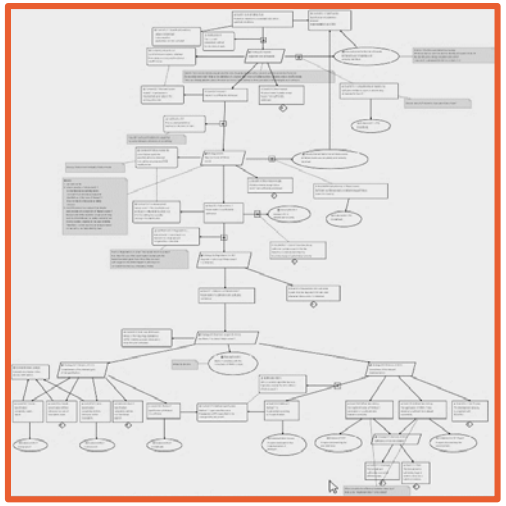
Assurance Strategy for AI
Safety Argumentation represented in a Goal Structuring Notation (GSN)



- **Graphical notation** that represents the elements of an assurance case and the relationships between them
- Shows how **goals** (claims) can be broken down into **sub-goals** until they can be supported by direct reference to available **evidence**
- Principle aim is to improve the comprehension of the **assurance case** thus enabling rigorous review and analysis



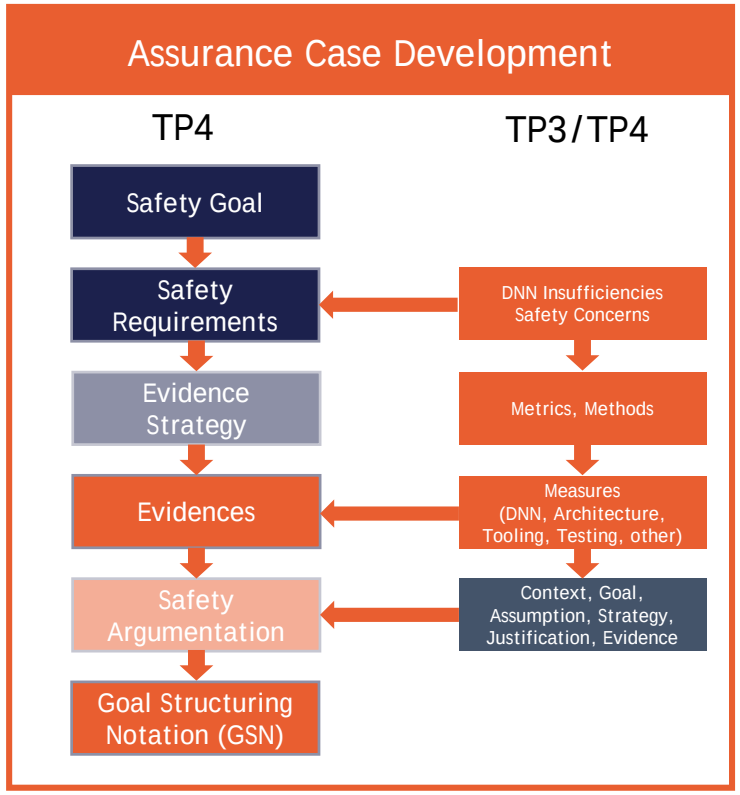
Source: Goal structuring notation, community standard version2

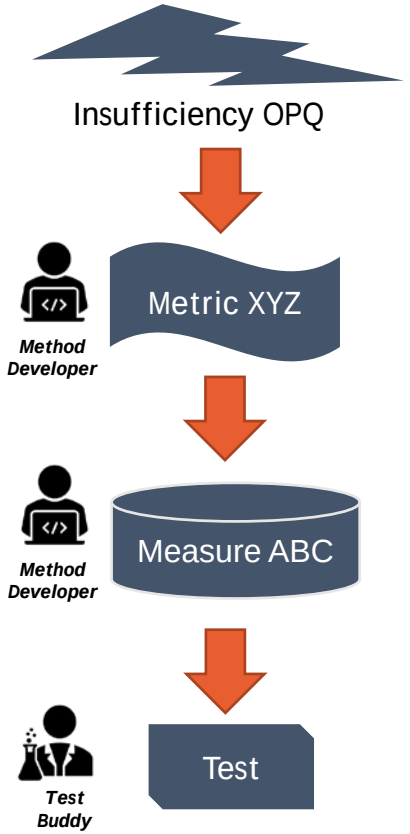


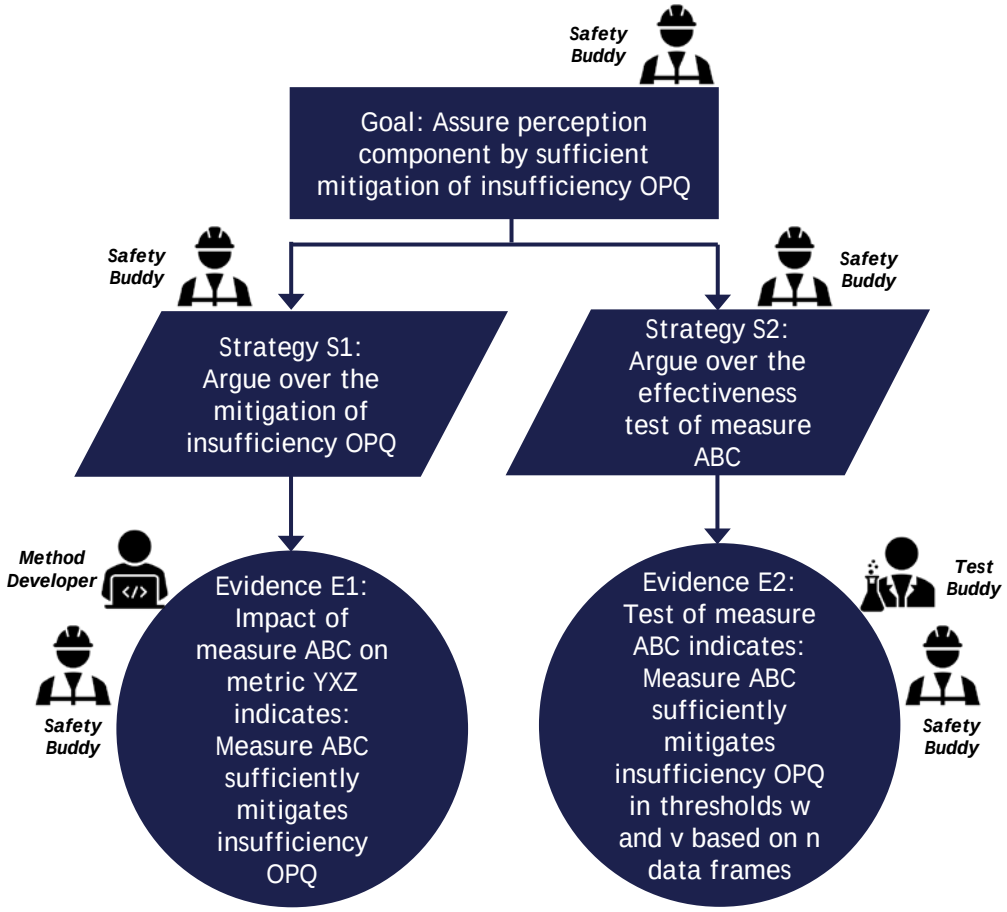
current example from AP4.3

Our evidence strategies: Specification, data, implementation, general approach

Assurance Strategy for AI
How to create Evidences from Methods and Tests

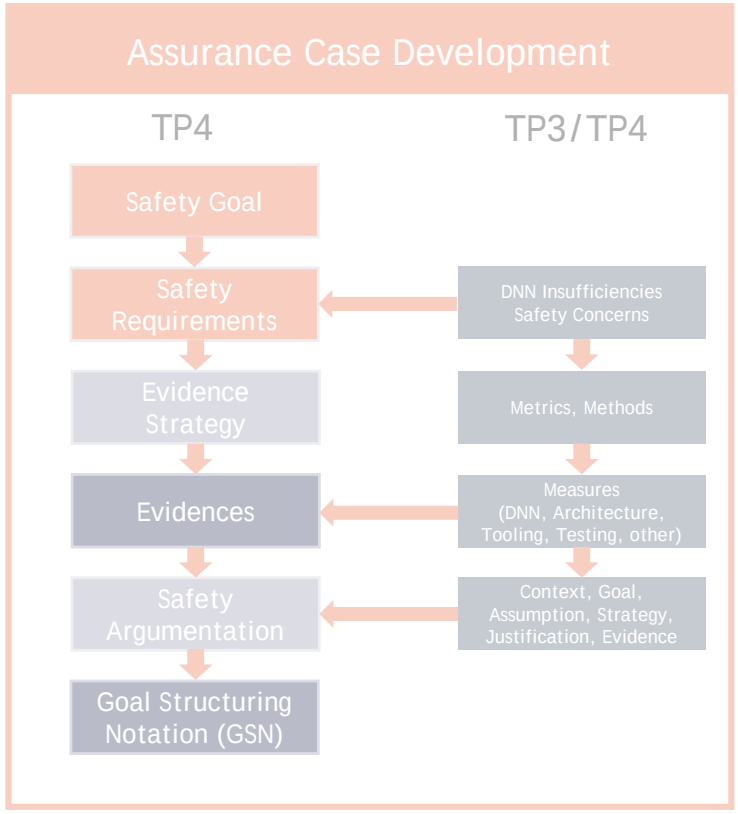






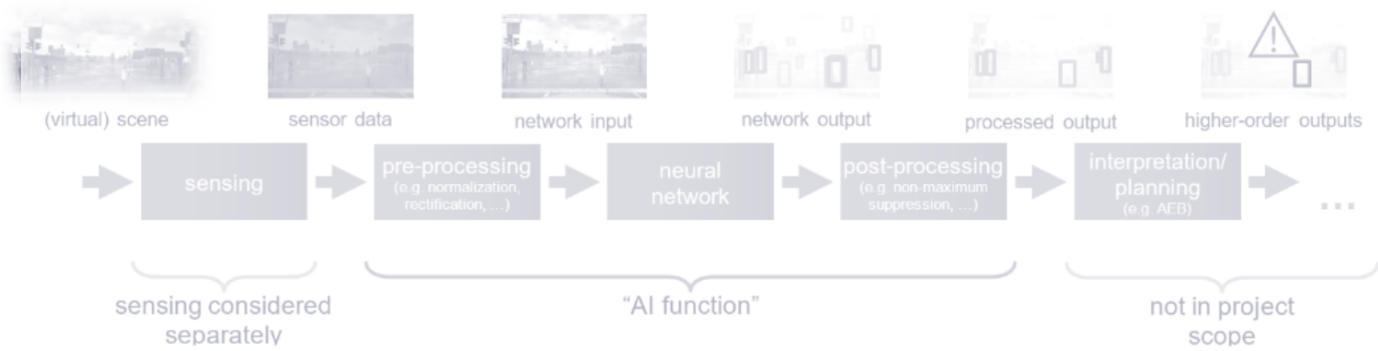
Interaction of Method Developer, Safety Buddy and Test Buddy leads to evidences for the safety argumentation

Assurance Strategy for AI
Safety Goal and Safety Requirements



Safety goal for the system:
"Avoid collisions with pedestrians"

Safety goal for perception within the ODD*:
"No relevant pedestrian shall be overlooked"

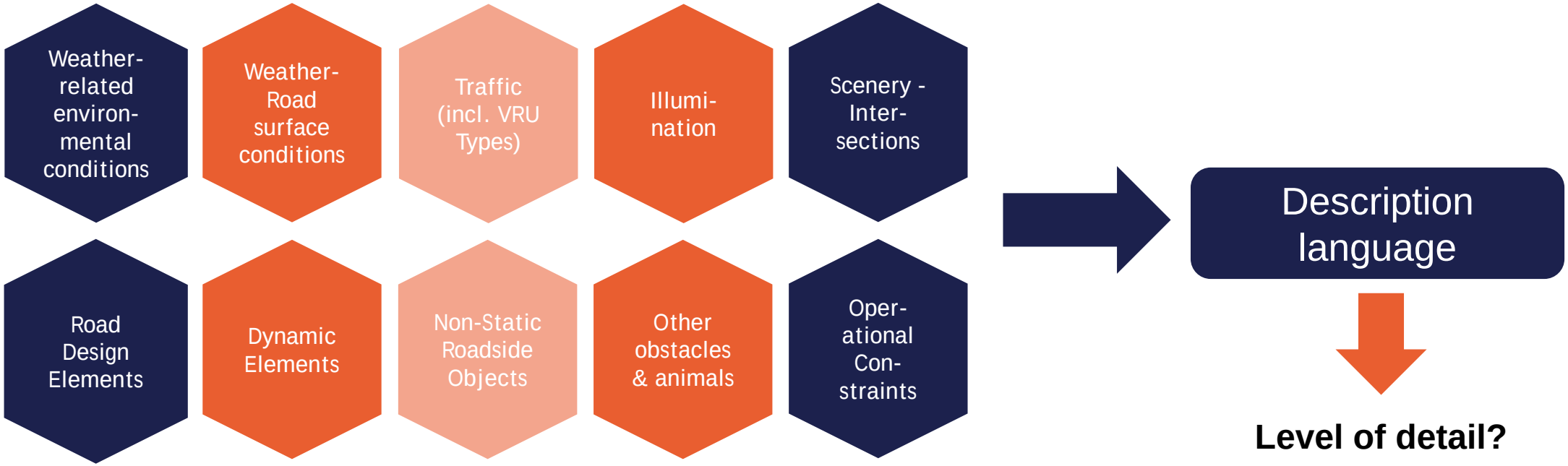


Deep dive: How to describe the ODD and evaluate test data coverage?

*) ODD = Operational Design Domain

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)



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 variated & incrementally generate new data
by analyzing coverage and generating missing combinations

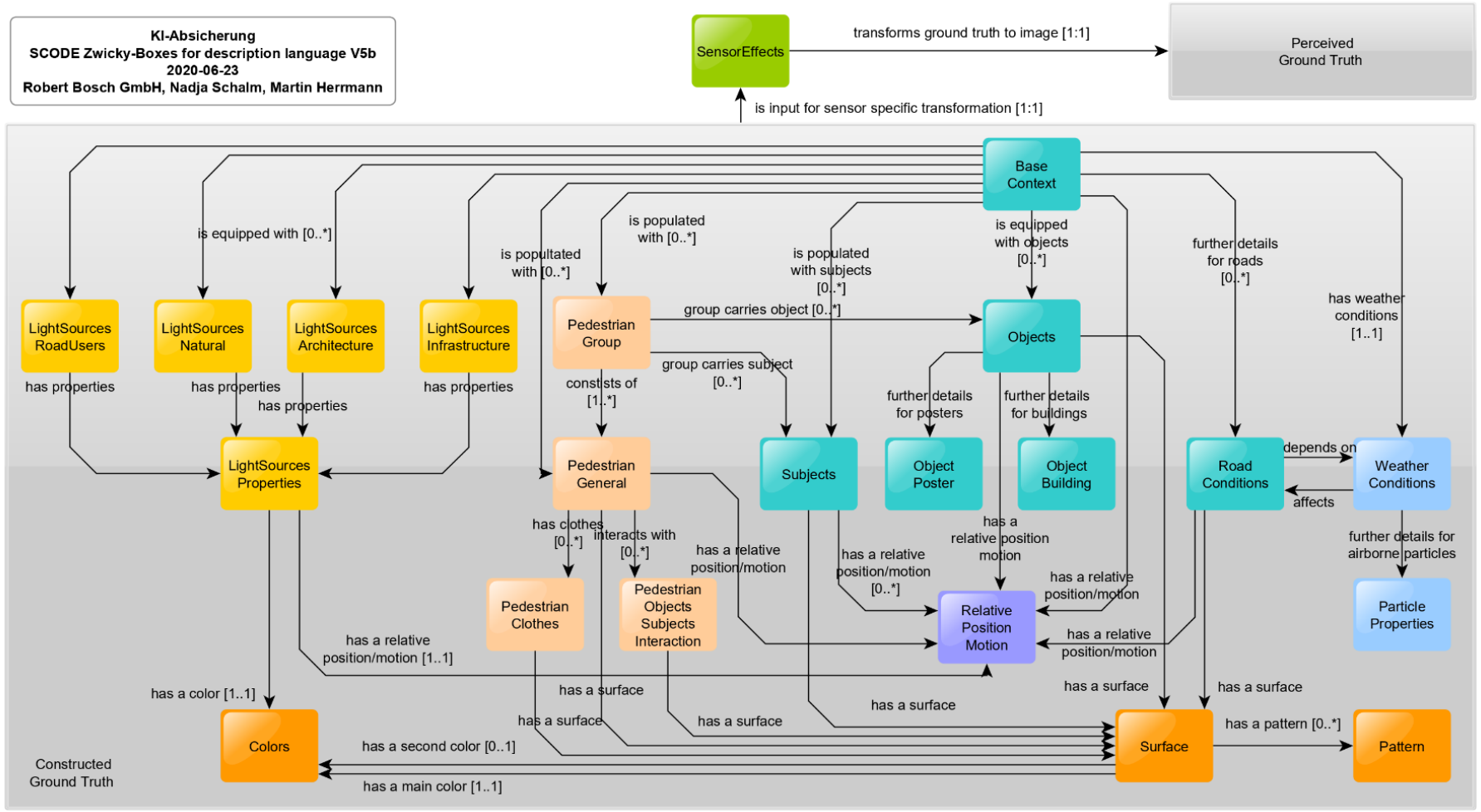
DNN-specific Safety Concerns (examples)

Data distribution is not a good approximation to real world

Unknown behavior in rare critical situations

*) ODD = Operational Design Domain

Assurance Strategy for AI
High Level view of Domain model derived from SCODE Zwicky-Boxes



Total

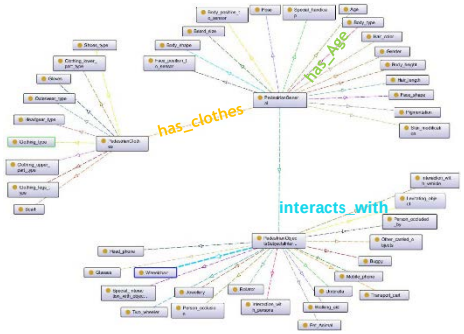
- ~250 dimensions
- ~1000 alternatives
- Several Sub-domains

Approach

- Review of public data sources / existing standards
- Brainstorming with experts
- Expert interviews
- Iterative refinement
- Needs to be challenged / extended by identified corner cases

Assurance Strategy for AI
Data representations of the data input space aligned to ontology

Ontology Graph (Relations)



Excerpt of ontology

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	

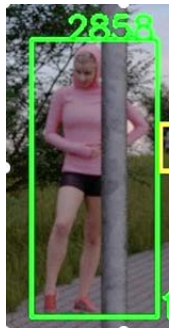
Zwicky Box – Discretized variations of important dimensions

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"

Object Annotations for DNN-Training



Occlusion_level: medium
Occluded_body_part: arm
Occlusion_object: lamp

Source: BIT Technology Solutions

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	
...	...	...	...	

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

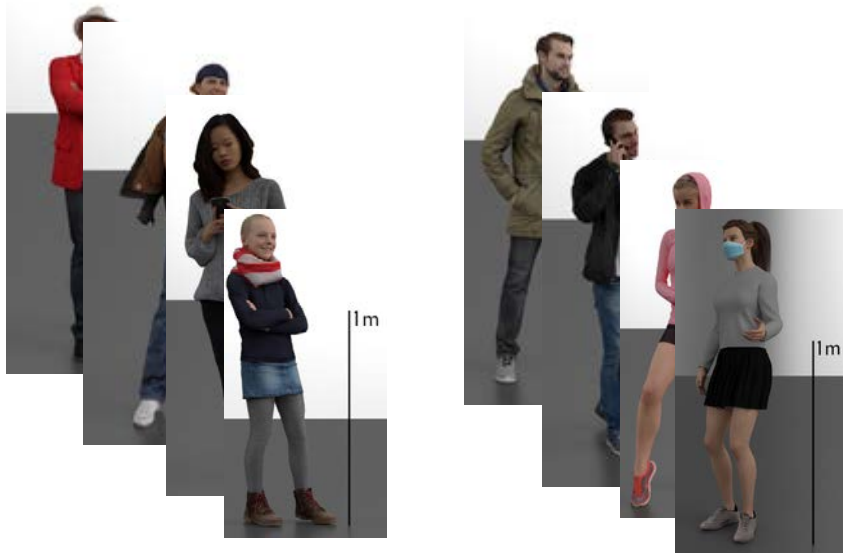
Assurance Strategy for AI
Application of combinatorial testing on part of domain model (Example)

Age	child		teenager		adult
Gender	male		female		
Body shape	thin	normal	muscular	obese	
Body height	80cm-120cm		120cm-160cm	160cm-200cm	
Pigmentation	high		medium	low	
Pose	standing	walking	running	sitting	lying

Zwicky Box – Discretized variations of important pedestrian dimensions (reduced)

Source: Mackevision

- The available 19 assets cover approx. 37 of these pair combinations (combinatorial testing), thus 63% are not covered.
- To cover all pairs further 28 asset proposals (as input to synth. data generation) were generated as a systematic combinations of combinatorial pairs



Source: BIT Technology Solutions

Next step: Add additional assets to test data set based on findings from data coverage analysis. Check DNN detection performance on these assets, especially regarding False Negatives & Uncertainties

Assurance Strategy for AI

Work package „Assurance Strategy for AI“ located in the (simplified) value chain

AP4.1

Structuring of Data input space

- How to describe the context in which the system is operated?
- What level of detail is needed?

TP2

Synth. Data specification & generation

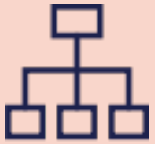
TP1

DNN-Development (incl. AI function specification)

TP3

DNN assurance methods & measures development

AP4.2/AP4.3



Safety Analysis, Structure of safety argumentation & Evidence strategies

- What are the safety goals and safety requirements to argument?
- How to build an evidence based safety argumen-tation and visualize it using a GSN?

AP4.4/AP4.5

Test methods & test strategy

- To what extent does the training & test data cover the ODD?
- How robust & effective is the safety argumentation?

KI ABSICHERUNG

Safe AI for Automated Driving

Frédérik Blank, Andreas Rohatschek, Robert Bosch GmbH

Frederik.Blank@de.bosch.com

Andreas-Juergen.Rohatschek@de.bosch.com

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