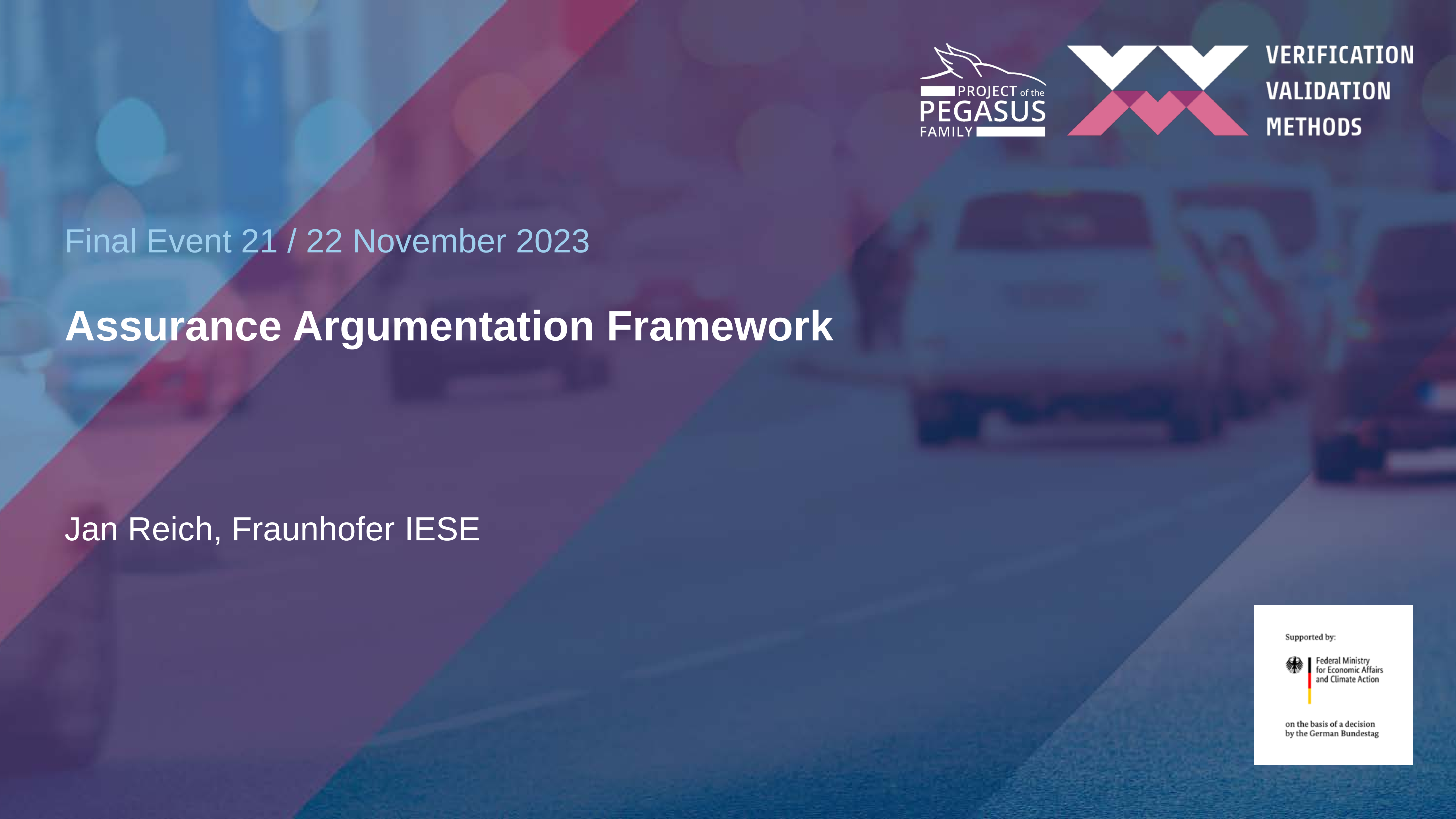


PROJECT of the PEGASUS FAMILY

VERIFICATION
VALIDATION
METHODS

Final Event 21 / 22 November 2023

Assurance Argumentation Framework

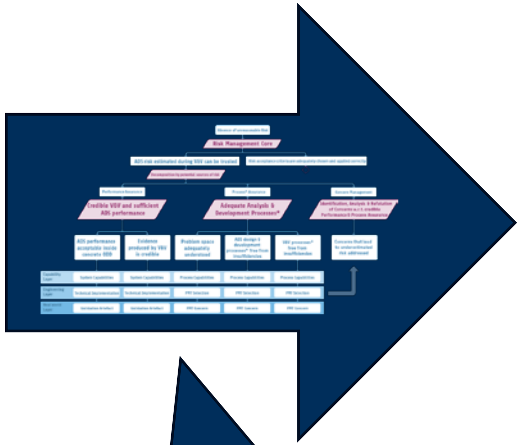
Jan Reich, Fraunhofer IESE

Supported by:

Federal Ministry
for Economic Affairs
and Climate Action

on the basis of a decision
by the German Bundestag

Safety cases tell a system‘s safety story!

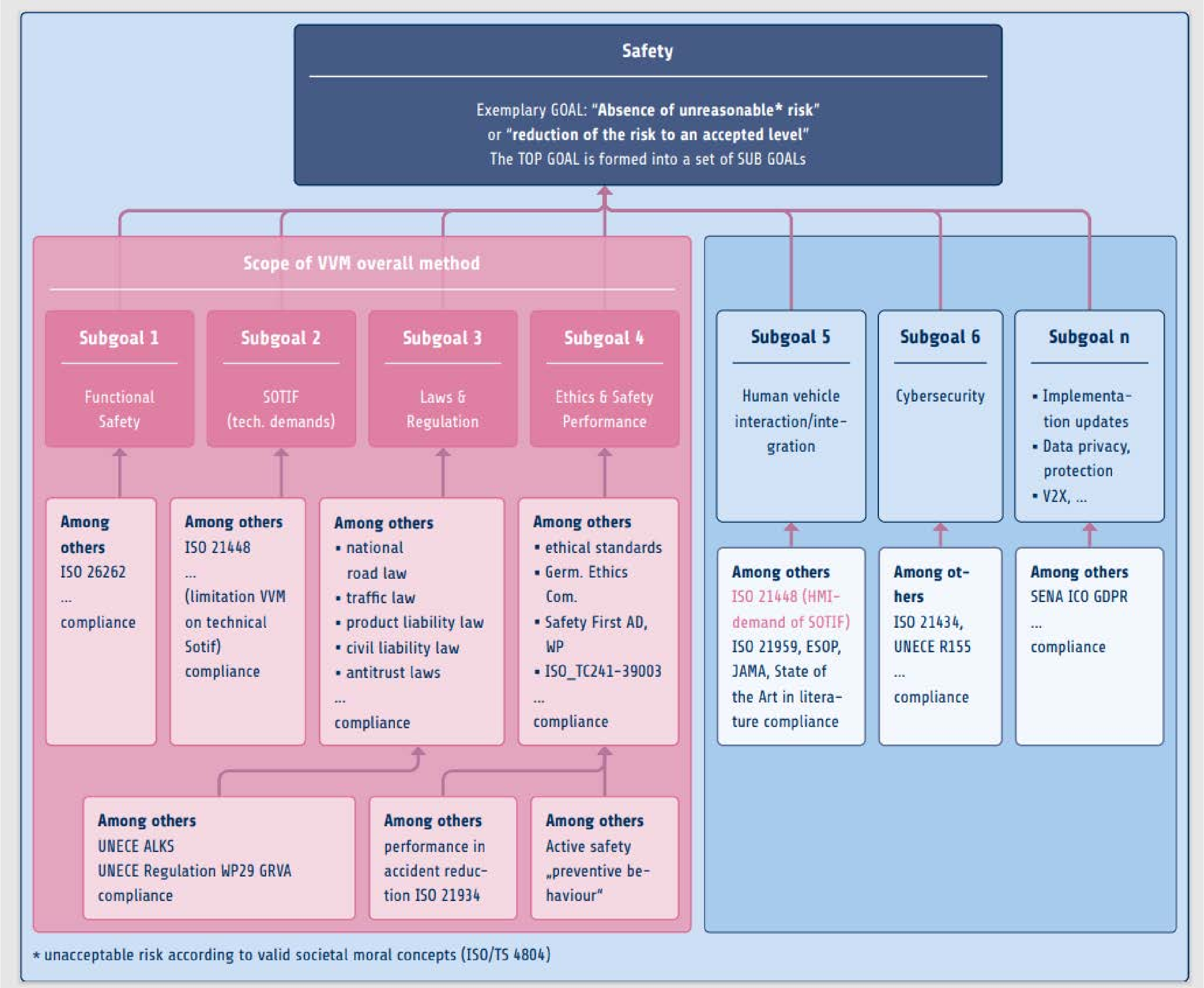


Why is this ADS safe?

Here‘s our argument and evidence!

Kraftfahrt-Bundesamt

Safety assurance aspects requiring argumentation



Safety

Exemplary GOAL: "**Absence of unreasonable* risk**"
or "**reduction of the risk to an accepted level**"
The TOP GOAL is formed into a set of SUB GOALs

Scope of VVM overall method

Subgoal 1
Functional Safety

Subgoal 2
SOTIF (tech. demands)

Subgoal 3
Laws & Regulation

Subgoal 4
Ethics & Safety Performance

Among others
ISO 26262
...
compliance

Among others
ISO 21448
...
(limitation VVM on technical Sotif)
compliance

Among others
- national road law
- traffic law
- product liability law
- civil liability law
- antitrust laws

...
compliance

Among others
- ethical standards
- Germ. Ethics Com.
- Safety First AD, WP
- ISO_TC241-39003

...
compliance

Among others
UNECE ALKS
UNECE Regulation WP29 GRVA
compliance

Among others
performance in accident reduction ISO 21934

Among others
Active safety „preventive behaviour"

Subgoal 5
Human vehicle interaction/integration

Subgoal 6
Cybersecurity

Subgoal n
- Implementation updates
- Data privacy, protection
- V2X, ...

Among others
ISO 21448 (HMI-demand of SOTIF)
ISO 21959, ESOP, JAMA, State of the Art in literature compliance

Among others
ISO 21434,
UNECE R155
...
compliance

Among others
SENA ICO GDPR
...
compliance

* unacceptable risk according to valid societal moral concepts (ISO/TS 4804)

Argumentation principles are realized in argumentation structures, which are communicated via documented safety cases

VERIFICATION
VALIDATION
METHODS

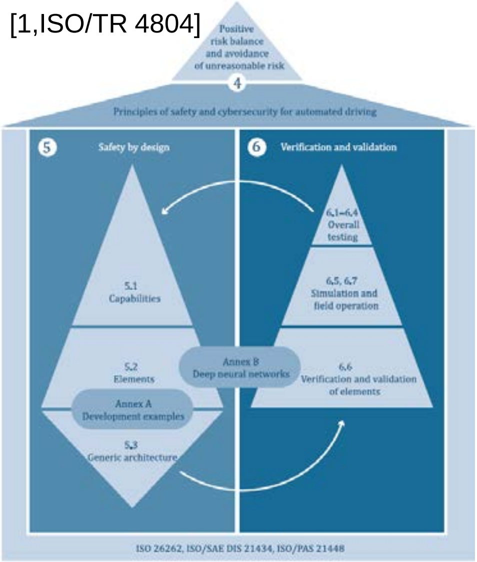
Performance (What?)

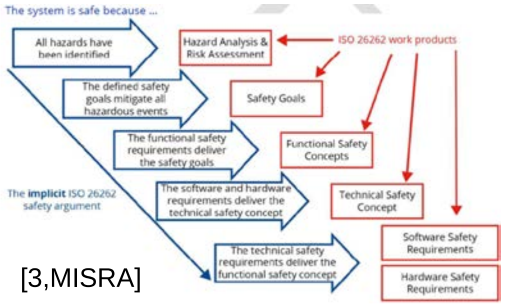
Process (How?)

Confidence (Why?)

Dialectics (Why not?)

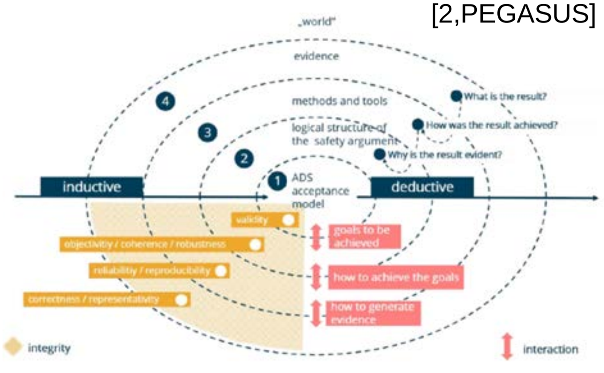
[1,ISO/TR 4804]

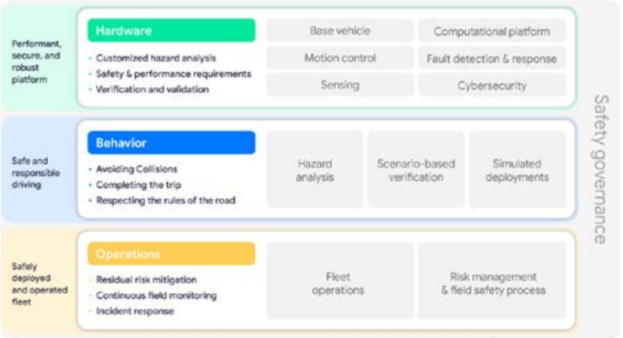

Positive risk balance and avoidance of unreasonable risk
4
Principles of safety and cybersecurity for automated driving
5
Safety by design
6
Verification and validation
5.1
Capabilities
5.2
Elements
Annex A
Development examples
5.3
Generic architecture
Annex B
Deep neural networks
6.1–6.4
Overall testing
6.5, 6.7
Simulation and field operation
6.6
Verification and validation of elements
ISO 26262, ISO/SAE DIS 21434, ISO/PAS 21448


The system is safe because ...
All hazards have been identified
The defined safety goals mitigate all hazardous events
The functional safety requirements deliver the safety goals
The software and hardware requirements deliver the technical safety concept
The technical safety requirements deliver the functional safety concept
The implicit ISO 26262 safety argument
ISO 26262 work products
Hazard Analysis & Risk Assessment
Safety Goals
Functional Safety Concepts
Technical Safety Concept
Software Safety Requirements
Hardware Safety Requirements

[3,MISRA]

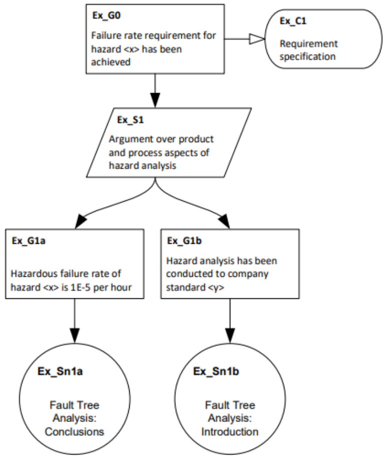
[2,PEGASUS]


„world"
evidence
methods and tools
logical structure of the safety argument
1
ADS acceptance model
2
3
4
What is the result?
How was the result achieved?
Why is the result evident?
inductive
deductive
validity
objectivity / coherence / robustness
reliability / reproducibility
correctness / representativity
goals to be achieved
how to achieve the goals
how to generate evidence
integrity
interaction


Performant, secure, and robust platform
Hardware
Customized hazard analysis
Safety & performance requirements
Verification and validation
Base vehicle
Computational platform
Motion control
Fault detection & response
Sensing
Cybersecurity
Safe and responsible driving
Behavior
Avoiding Collisions
Completing the trip
Respecting the rules of the road
Hazard analysis
Scenario-based verification
Simulated deployments
Safely deployed and operated fleet
Operations
Residual risk mitigation
Continuous field monitoring
Incident response
Fleet operations
Risk management & field safety process
Safety governance

[4,Waymo]

[5,GSN]

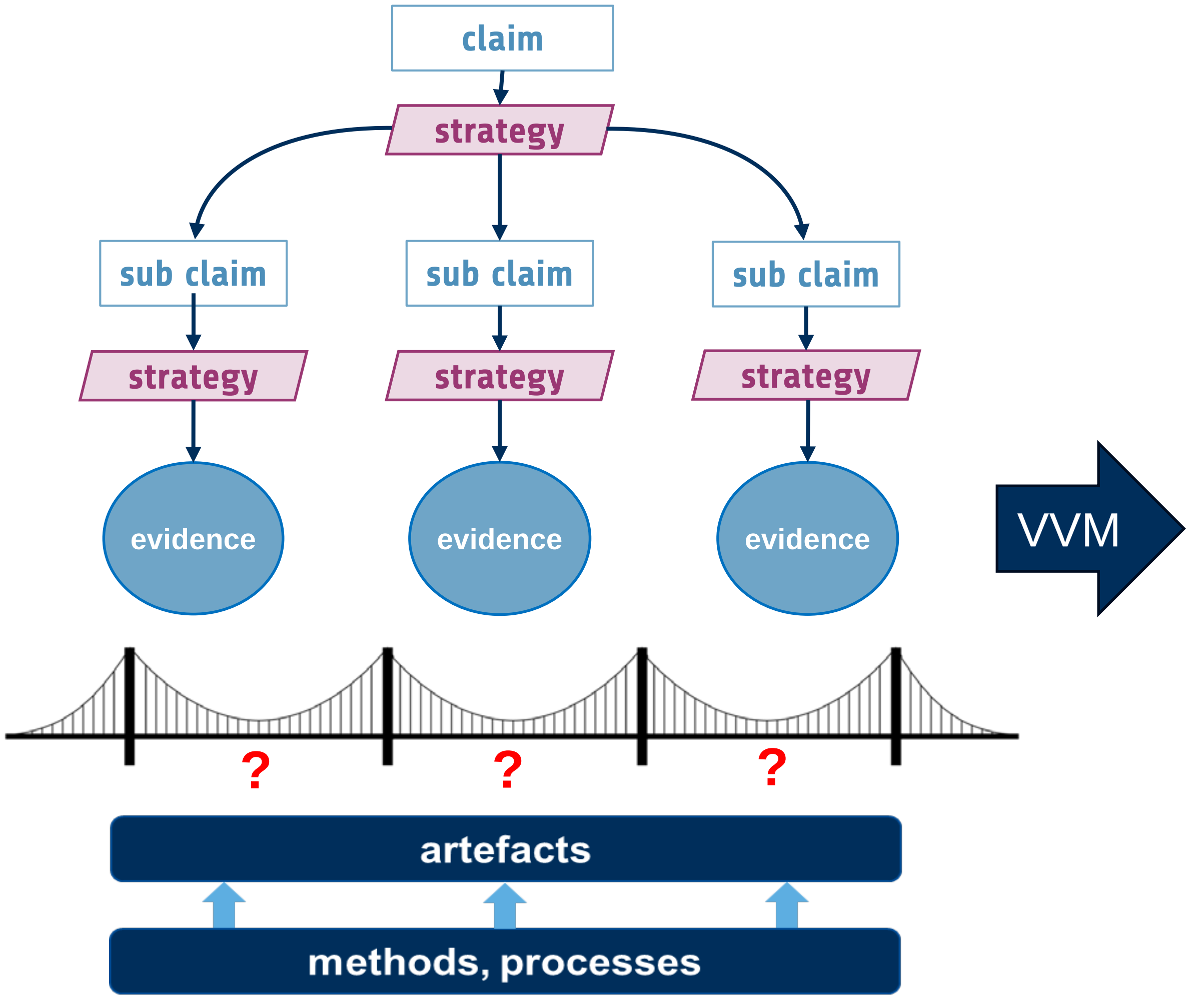

Ex_G0
Failure rate requirement for hazard <x> has been achieved
Ex_C1
Requirement specification
Ex_S1
Argument over product and process aspects of hazard analysis
Ex_G1a
Hazardous failure rate of hazard <x> is 1E-5 per hour
Ex_G1b
Hazard analysis has been conducted to company standard <y>
Ex_Sn1a
Fault Tree Analysis: Conclusions
Ex_Sn1b
Fault Tree Analysis: Introduction

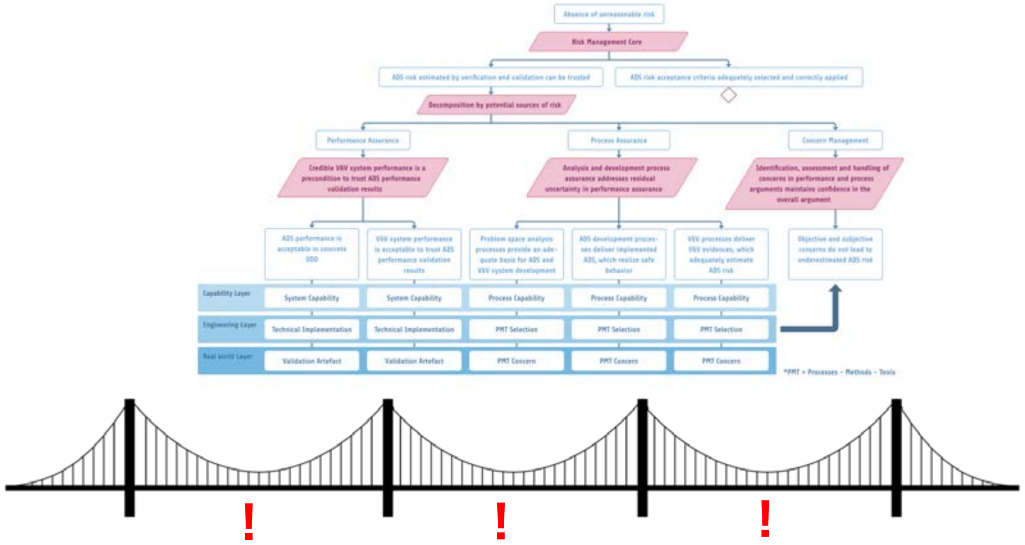
Generic Argumentation Principles

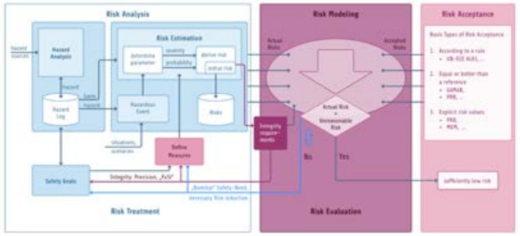
Concrete ADS Argumentation Structures

Safety Case Documentation

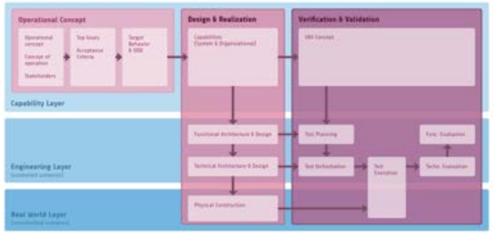
The VVM argumentation is aligned with the VVM assurance methods



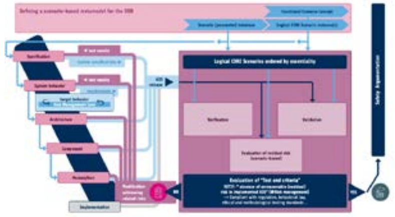

Absence of unreasonable risk
Risk Management Core
ADS risk estimated by verification and validation can be trusted
ADS risk acceptance criteria adequately selected and correctly applied
Decomposition by potential sources of risk
Performance Assurance
Process Assurance
Concern Management
Credible V&V system performance is a precondition to trust ADS performance validation results
Analysis and development process assurance addresses residual uncertainty in performance assurance
Identification, assessment and handling of concerns in performance and process arguments maintains confidence in the overall argument
ADS performance is acceptable in concrete ODD
V&V system performance is acceptable to trust ADS performance validation results
Problem space analysis processes provide an adequate basis for ADS and V&V system development
ADS development processes deliver implemented ADS, which realize safe behavior
V&V processes deliver V&V evidences, which adequately estimate ADS risk
Objective and subjective concerns do not lead to underestimated ADS risk
Capability Layer
System Capability
System Capability
Process Capability
Process Capability
Process Capability
Engineering Layer
Technical Implementation
Technical Implementation
PMT Selection
PMT Selection
PMT Selection
Real World Layer
Validation Artefact
Validation Artefact
PMT Concern
PMT Concern
PMT Concern
*PMT = Processes – Methods – Tools
!
!
!



Risk Management

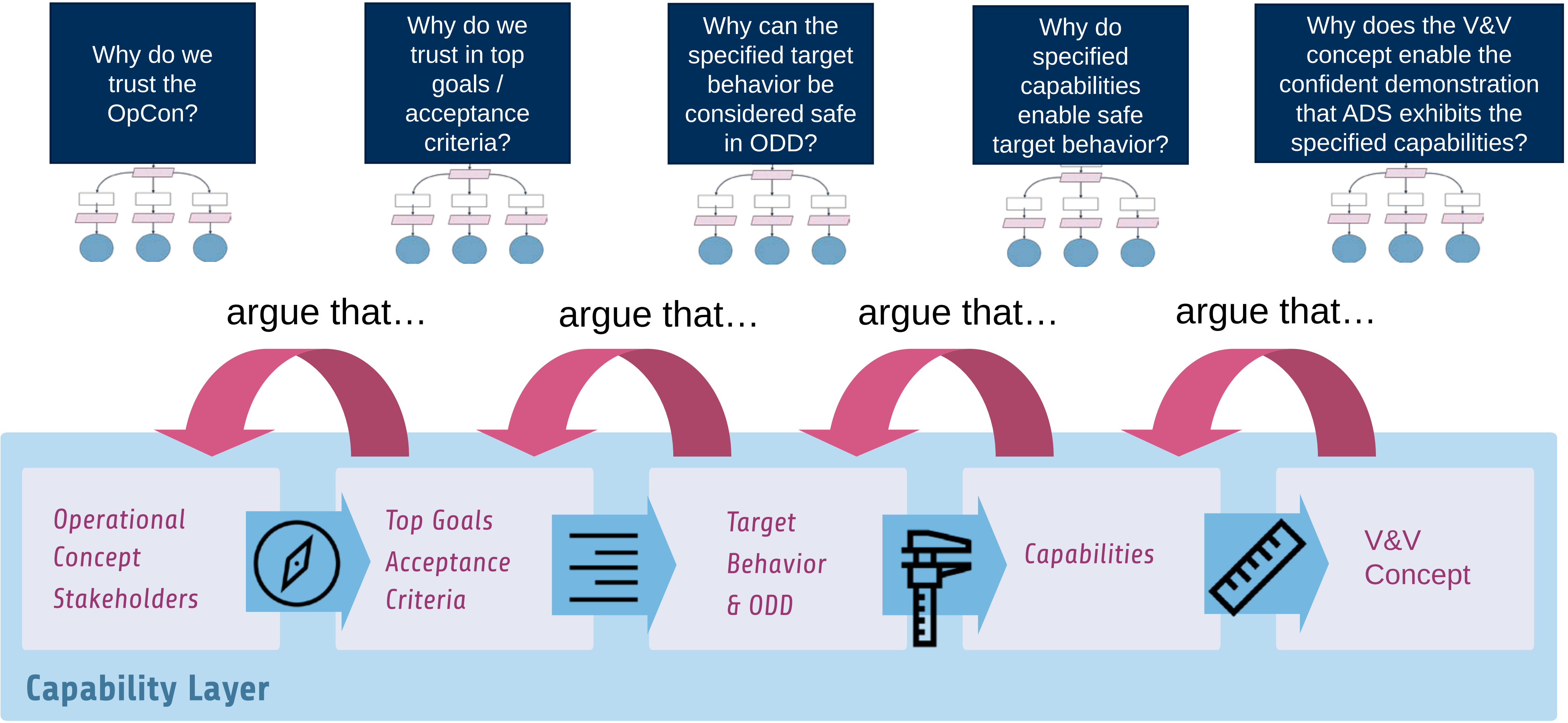


Development & Operation – Global



Scenario-based approach

Example: Divide and conquer the argumentation problem!



Why do we trust the OpCon?

Why do we trust in top goals / acceptance criteria?

Why can the specified target behavior be considered safe in ODD?

Why do specified capabilities enable safe target behavior?

Why does the V&V concept enable the confident demonstration that ADS exhibits the specified capabilities?

argue that…

argue that…

argue that…

argue that…

Operational Concept Stakeholders

Top Goals Acceptance Criteria

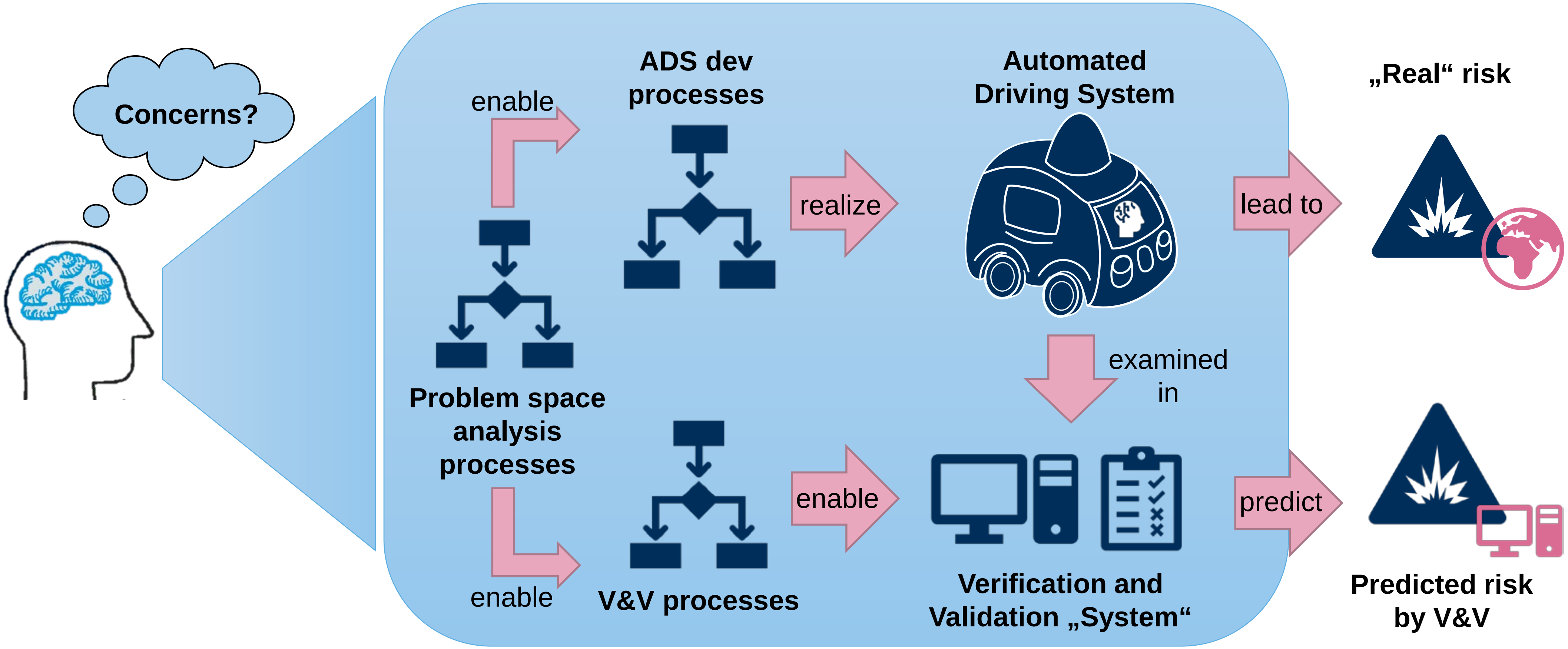
Target Behavior & ODD

Capabilities

V&V Concept

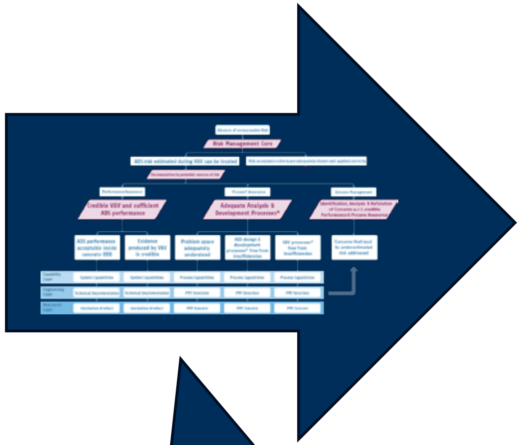
Capability Layer

Argumentation includes making explicit why we properly handle sources of inadequate risk estimation



Safety cases tell a system‘s safety story!

VERIFICATION
VALIDATION
METHODS

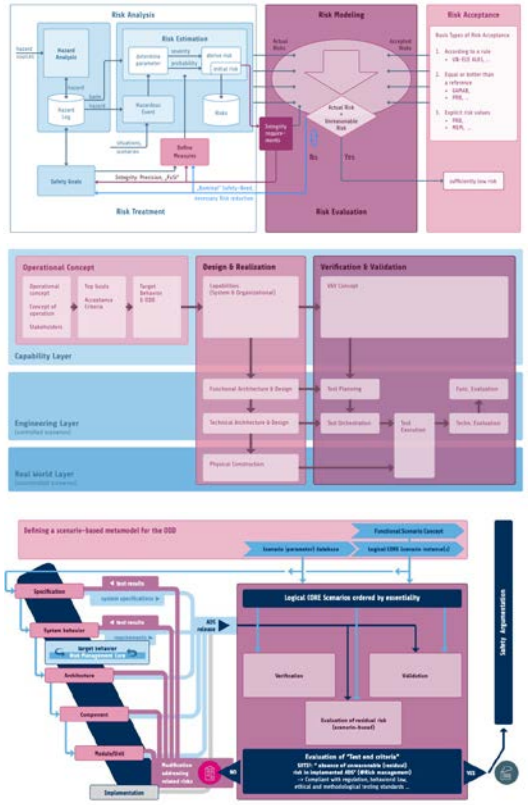


Why is this ADS safe?

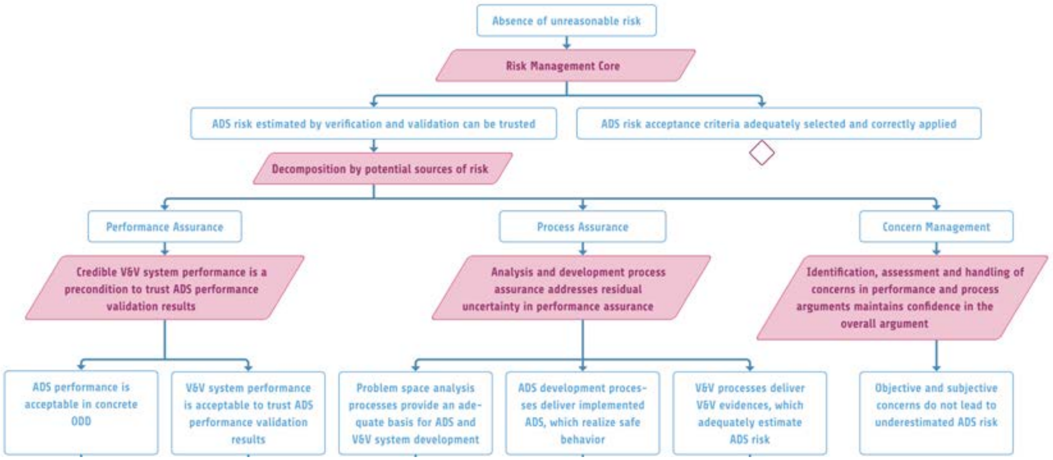
Here‘s our argument and evidence!

Kraftfahrt-
Bundesamt

The VVM minimum set of assumptions about necessary assurance artifacts and processes enables arguing normative requirements.



VVM Assurance Framework



Absence of unreasonable risk

Risk Management Core

ADS risk estimated by verification and validation can be trusted

ADS risk acceptance criteria adequately selected and correctly applied

Decomposition by potential sources of risk

Performance Assurance

Process Assurance

Concern Management

Credible V&V system performance is a precondition to trust ADS performance validation results

Analysis and development process assurance addresses residual uncertainty in performance assurance

Identification, assessment and handling of concerns in performance and process arguments maintains confidence in the overall argument

ADS performance is acceptable in concrete ODD

V&V system performance is acceptable to trust ADS performance validation results

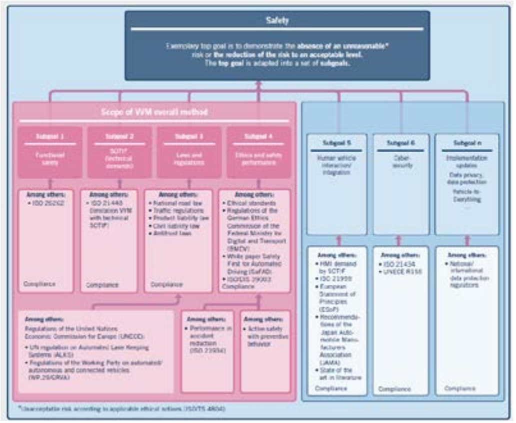
Problem space analysis processes provide an adequate basis for ADS and V&V system development

ADS development processes deliver implemented ADS, which realize safe behavior

V&V processes deliver V&V evidences, which adequately estimate ADS risk

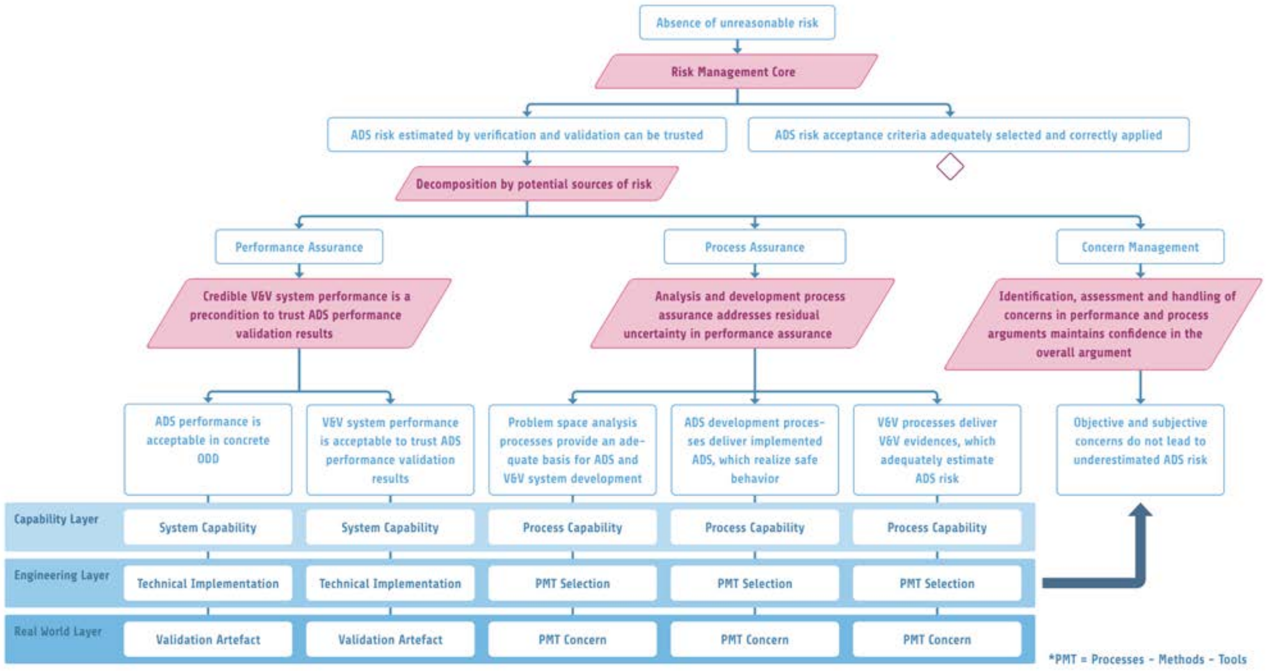
Objective and subjective concerns do not lead to underestimated ADS risk

Explicit safety argumentation structure


Safety
Exemplary top goal is to demonstrate the absence of an unreasonable* risk or the reduction of the risk to an acceptable level.
The top goal is adapted into a set of subgoals.
Scope of VVM overall method
Subgoal 1
Functional safety
Subgoal 2
SOTIF (technical demands)
Subgoal 3
Laws and regulations
Subgoal 4
Ethics and safety performance
Subgoal 5
Human vehicle interaction/ integration
Subgoal 6
Cyber-security
Subgoal n
Implementation updates
Data privacy, data protection
Vehicle to Everything
Among others:
ISO 26262
ISO 21448
National road law
Traffic regulations
Product liability law
Civil liability law
Antitrust laws
Ethical standards
ISO 21434
UNECE R156
Compliance
*Unacceptable risk according to applicable ethical notions (ISO/TS 4804)

Explainable compliance with normative requirements

Argumentation structure is the basis for project-specific concretization



Absence of unreasonable risk

Risk Management Core

ADS risk estimated by verification and validation can be trusted

ADS risk acceptance criteria adequately selected and correctly applied

Decomposition by potential sources of risk

Performance Assurance

Process Assurance

Concern Management

Credible V&V system performance is a precondition to trust ADS performance validation results

Analysis and development process assurance addresses residual uncertainty in performance assurance

Identification, assessment and handling of concerns in performance and process arguments maintains confidence in the overall argument

ADS performance is acceptable in concrete ODD

V&V system performance is acceptable to trust ADS performance validation results

Problem space analysis processes provide an adequate basis for ADS and V&V system development

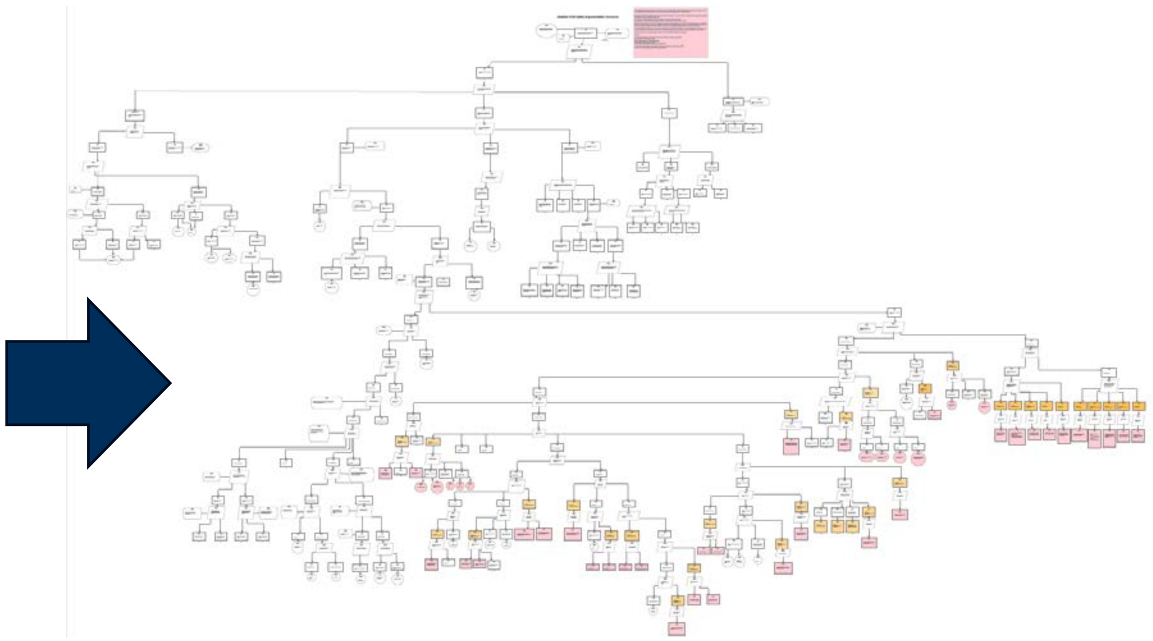
ADS development processes deliver implemented ADS, which realize safe behavior

V&V processes deliver V&V evidences, which adequately estimate ADS risk

Objective and subjective concerns do not lead to underestimated ADS risk

Capability Layer	System Capability	System Capability	Process Capability	Process Capability	Process Capability
Engineering Layer	Technical Implementation	Technical Implementation	PMT Selection	PMT Selection	PMT Selection
Real World Layer	Validation Artefact	Validation Artefact	PMT Concern	PMT Concern	PMT Concern

*PMT = Processes - Methods - Tools

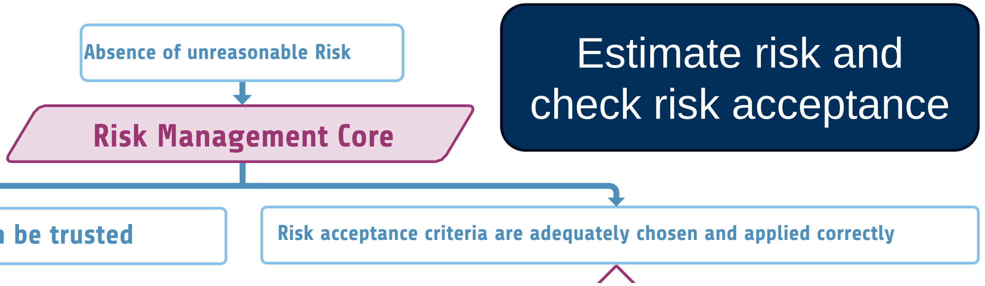


Argumentation Structure

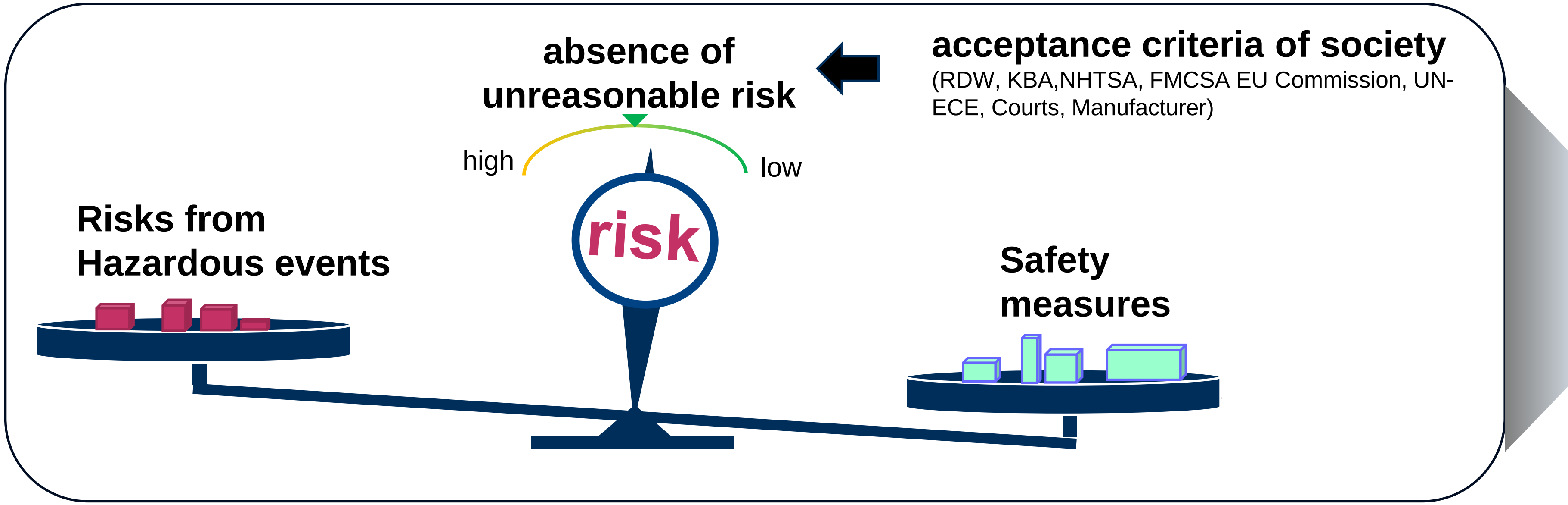
GSN Exemplification

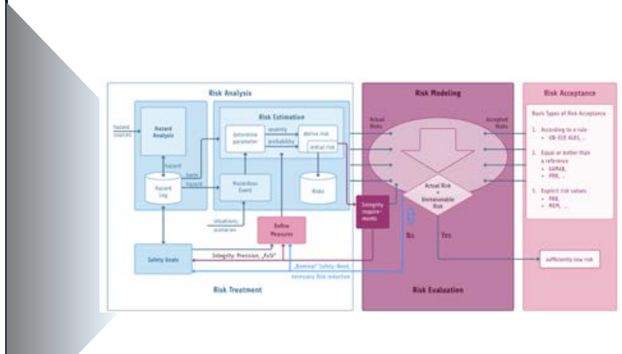
Top Level Strategy

Safety operationalization through risk management core


Absence of unreasonable Risk
Risk Management Core
ADS risk estimated during V&V can be trusted
Risk acceptance criteria are adequately chosen and applied correctly

Estimate risk and check risk acceptance


absence of unreasonable risk
acceptance criteria of society
(RDW, KBA,NHTSA, FMCSA EU Commission, UN-ECE, Courts, Manufacturer)
high
low
risk
Risks from Hazardous events
Safety measures

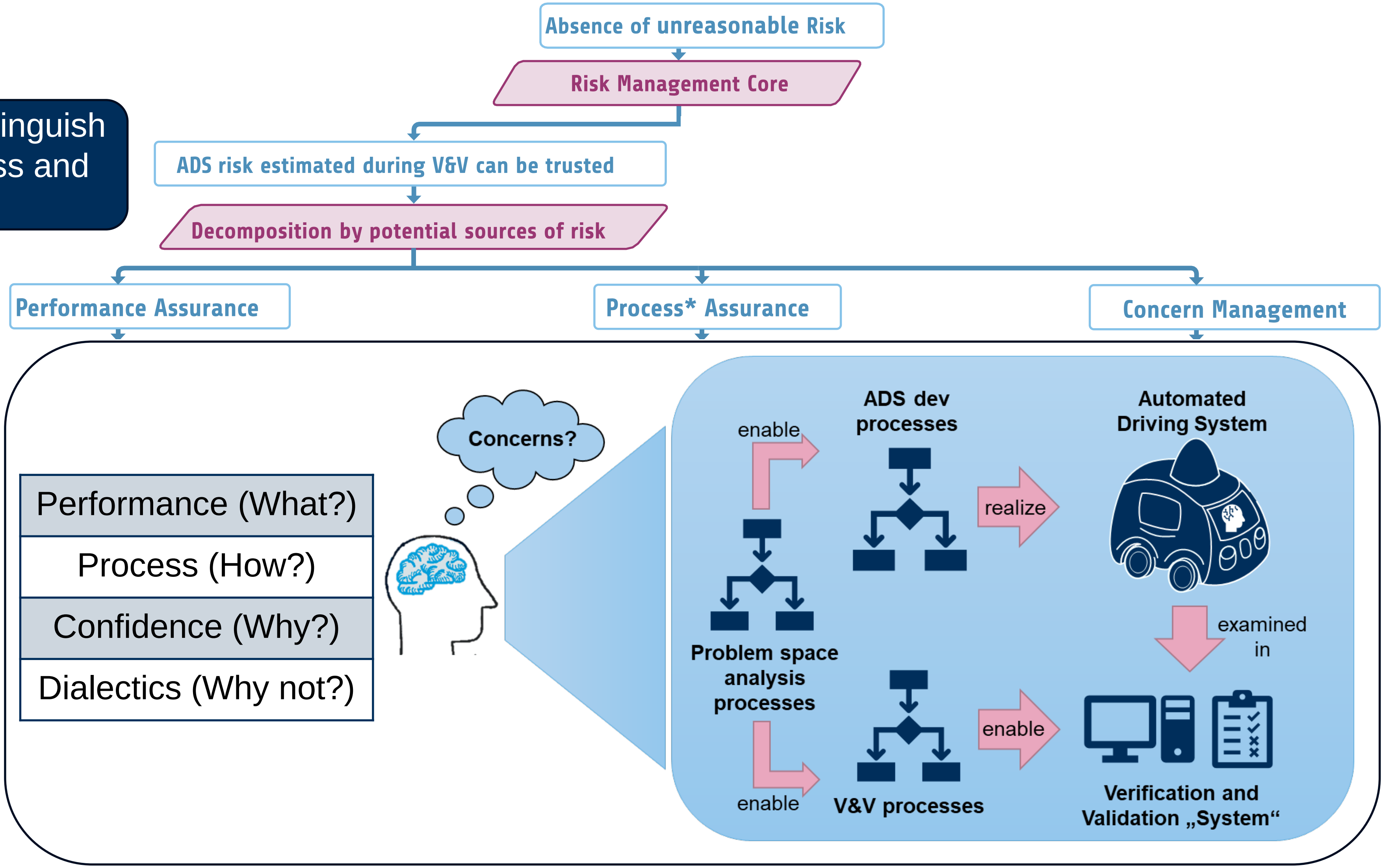


Risk Management Core

ADS Risk Estimation Decomposition Strategy

Generic argumentation principles as level 2 decomposition

What‘s relevant? Distinguish performance, process and confidence



Absence of unreasonable Risk

Risk Management Core

ADS risk estimated during V&V can be trusted

Decomposition by potential sources of risk

Performance Assurance

Process* Assurance

Concern Management

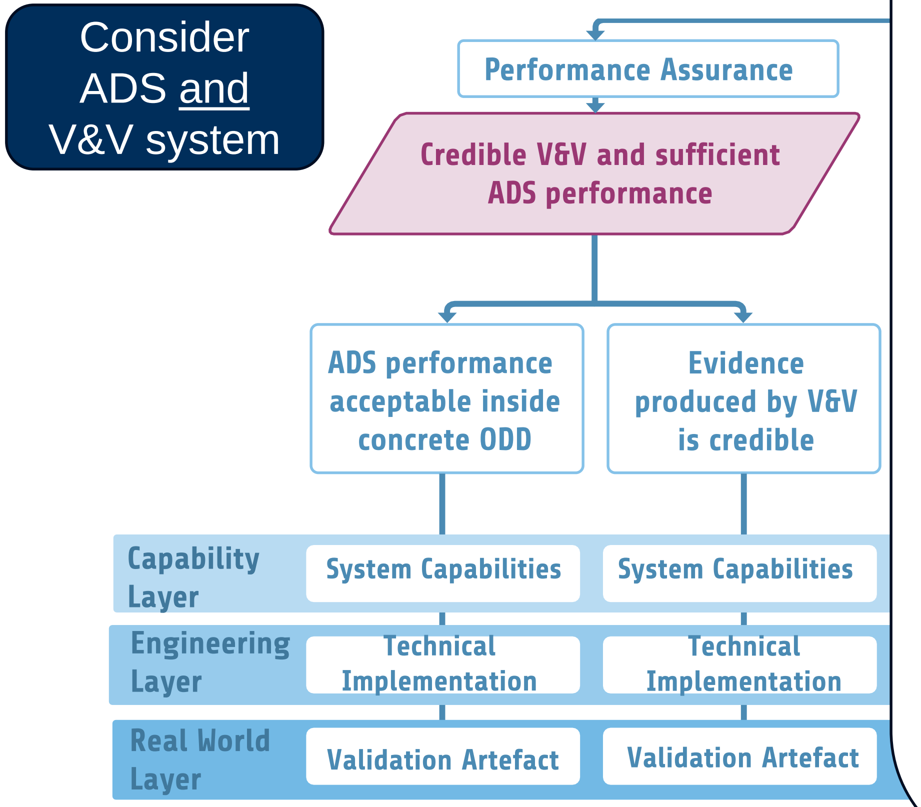
Performance (What?)

Process (How?)

Confidence (Why?)

Dialectics (Why not?)

Performance argumentation based on scenario-based approach



Consider ADS and V&V system

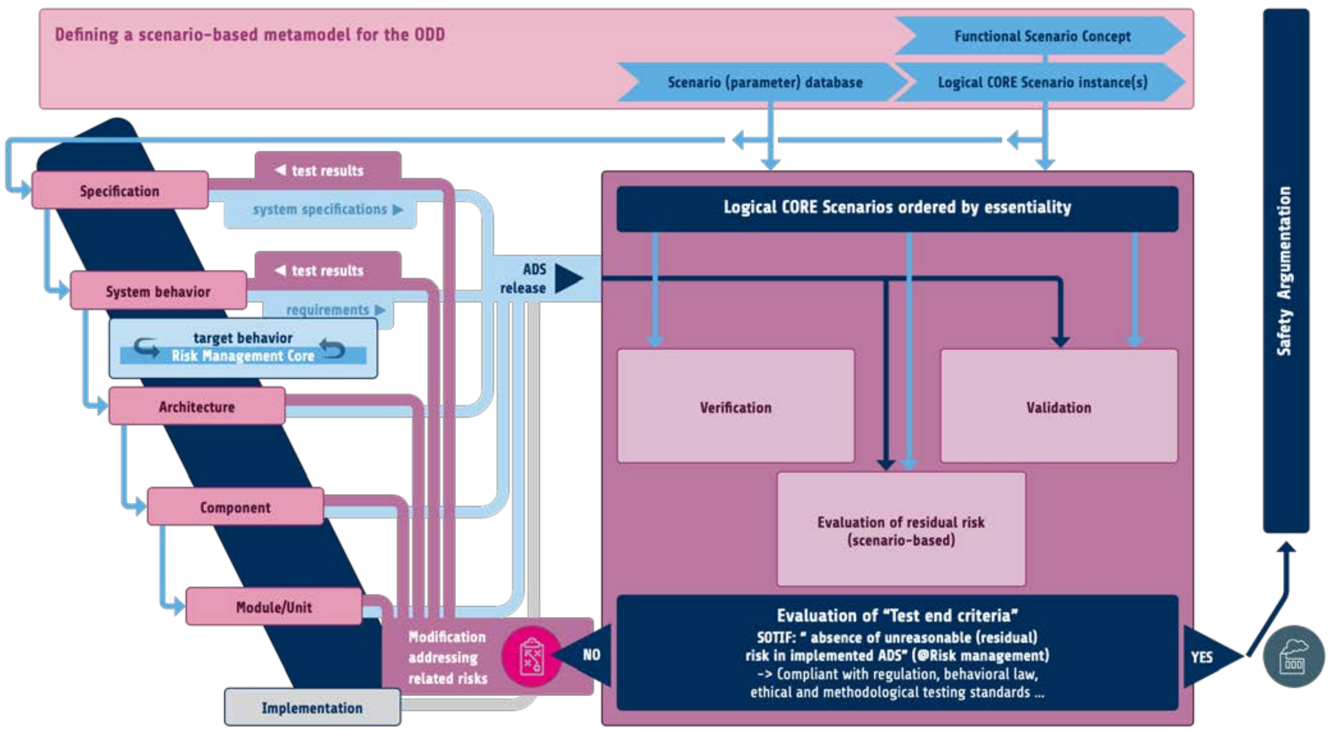
Performance Assurance

Credible V&V and sufficient ADS performance

ADS performance acceptable inside concrete ODD

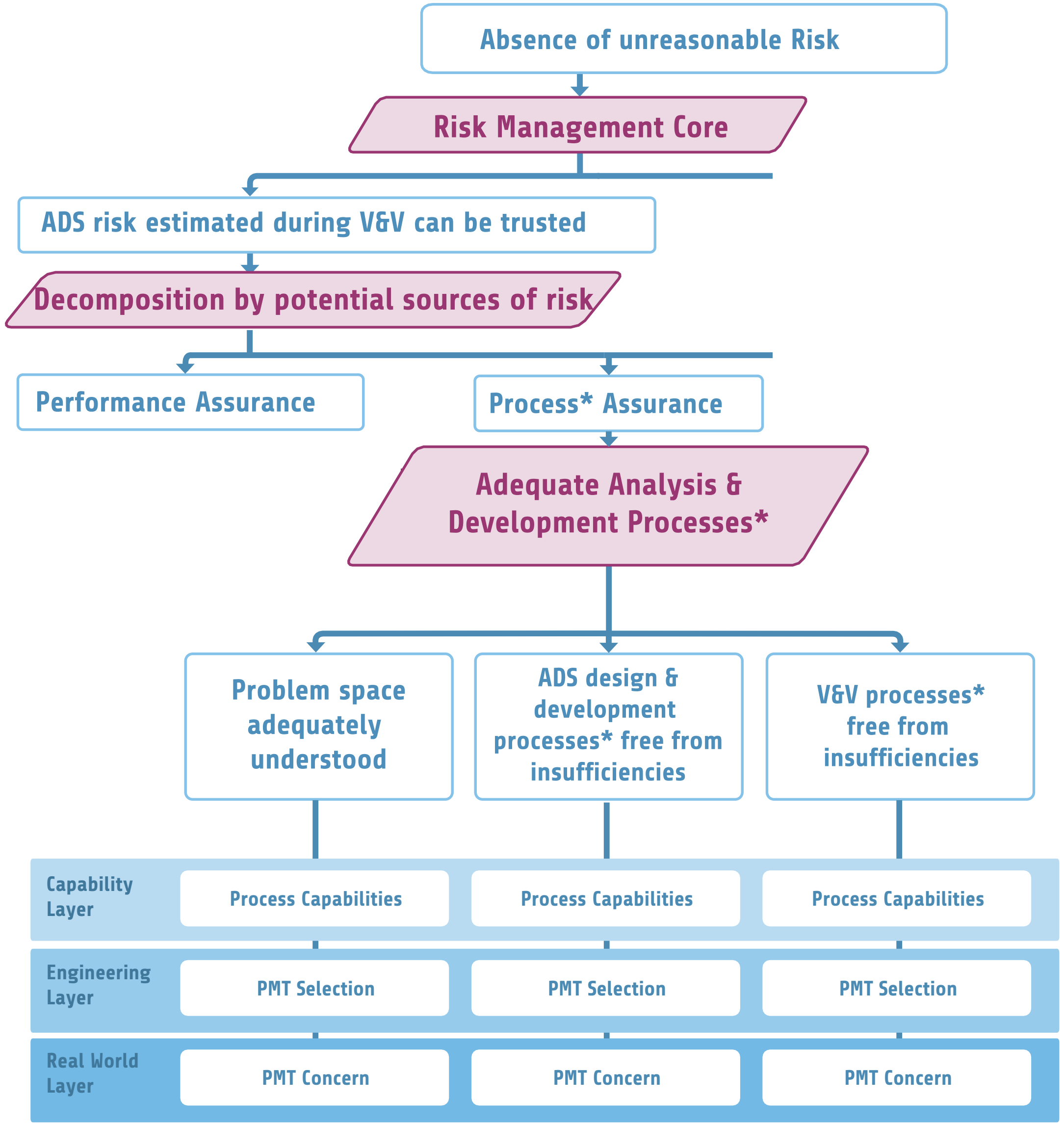
Evidence produced by V&V is credible

Capability Layer	System Capabilities	System Capabilities
Engineering Layer	Technical Implementation	Technical Implementation
Real World Layer	Validation Artefact	Validation Artefact


Defining a scenario-based metamodel for the ODD
Functional Scenario Concept
Scenario (parameter) database
Logical CORE Scenario instance(s)
Specification
test results
system specifications
System behavior
test results
requirements
target behavior
Risk Management Core
Architecture
Component
Module/Unit
Implementation
ADS release
Modification addressing related risks
NO
Logical CORE Scenarios ordered by essentiality
Verification
Validation
Evaluation of residual risk (scenario-based)
Evaluation of "Test end criteria"
SOTIF: " absence of unreasonable (residual) risk in implemented ADS" (@Risk management)
-> Compliant with regulation, behavioral law, ethical and methodological testing standards ...
YES
Safety Argumentation

Technical V&V Toolchain

Process argumentation boosts confidence in performance evidence!



Absence of unreasonable Risk

Risk Management Core

ADS risk estimated during V&V can be trusted

Decomposition by potential sources of risk

Performance Assurance

Process* Assurance

Adequate Analysis & Development Processes*

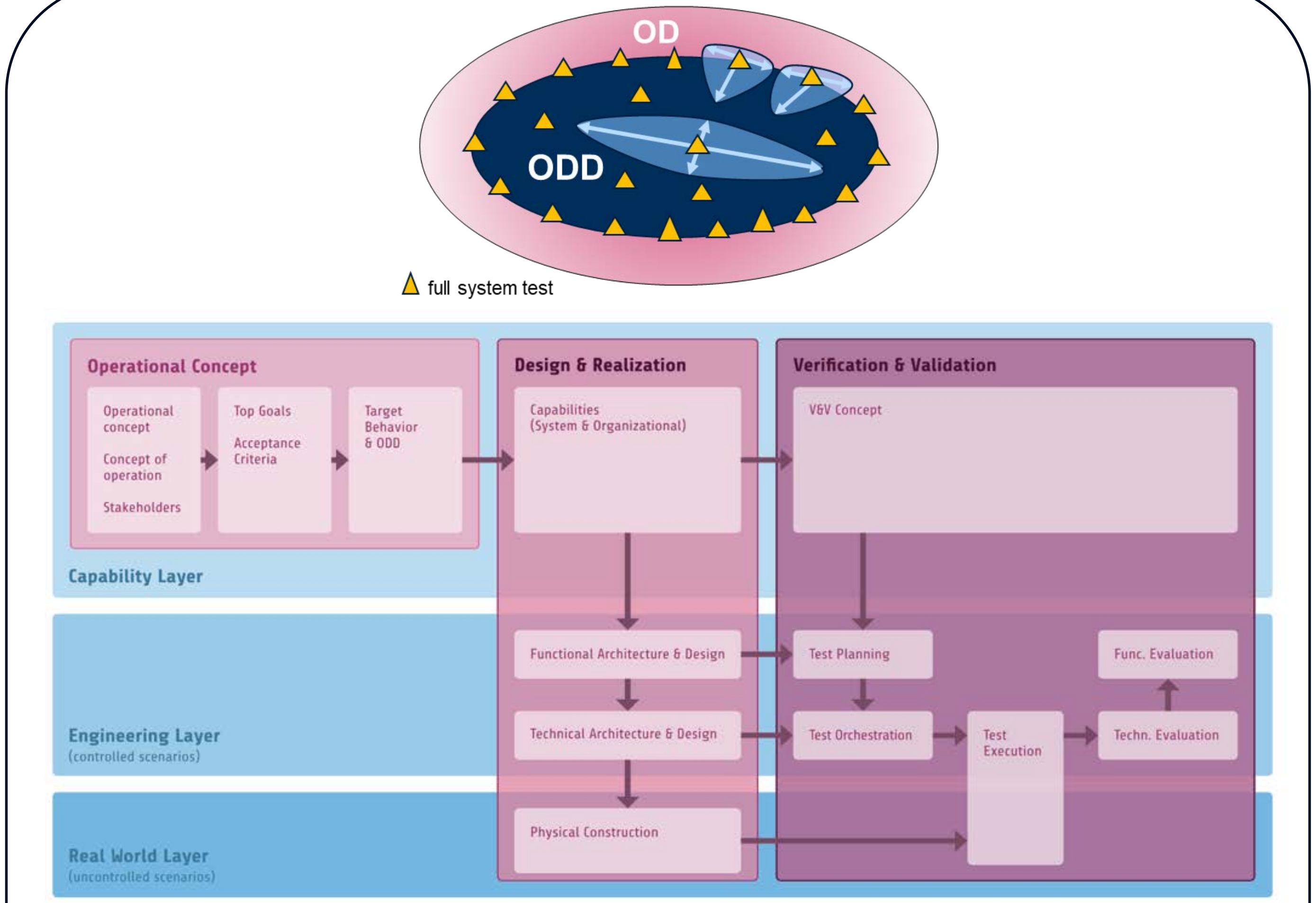
Problem space adequately understood

ADS design & development processes* free from insufficiencies

V&V processes* free from insufficiencies

Capability Layer	Process Capabilities	Process Capabilities	Process Capabilities
Engineering Layer	PMT Selection	PMT Selection	PMT Selection
Real World Layer	PMT Concern	PMT Concern	PMT Concern

PMT = Processes - Methods - Tools



Operational Concept

Operational concept
Concept of operation
Stakeholders

Top Goals
Acceptance Criteria

Target Behavior & ODD

Design & Realization

Capabilities (System & Organizational)

Verification & Validation

V&V Concept

Capability Layer

Functional Architecture & Design

Technical Architecture & Design

Physical Construction

Test Planning

Test Orchestration

Test Execution

Func. Evaluation

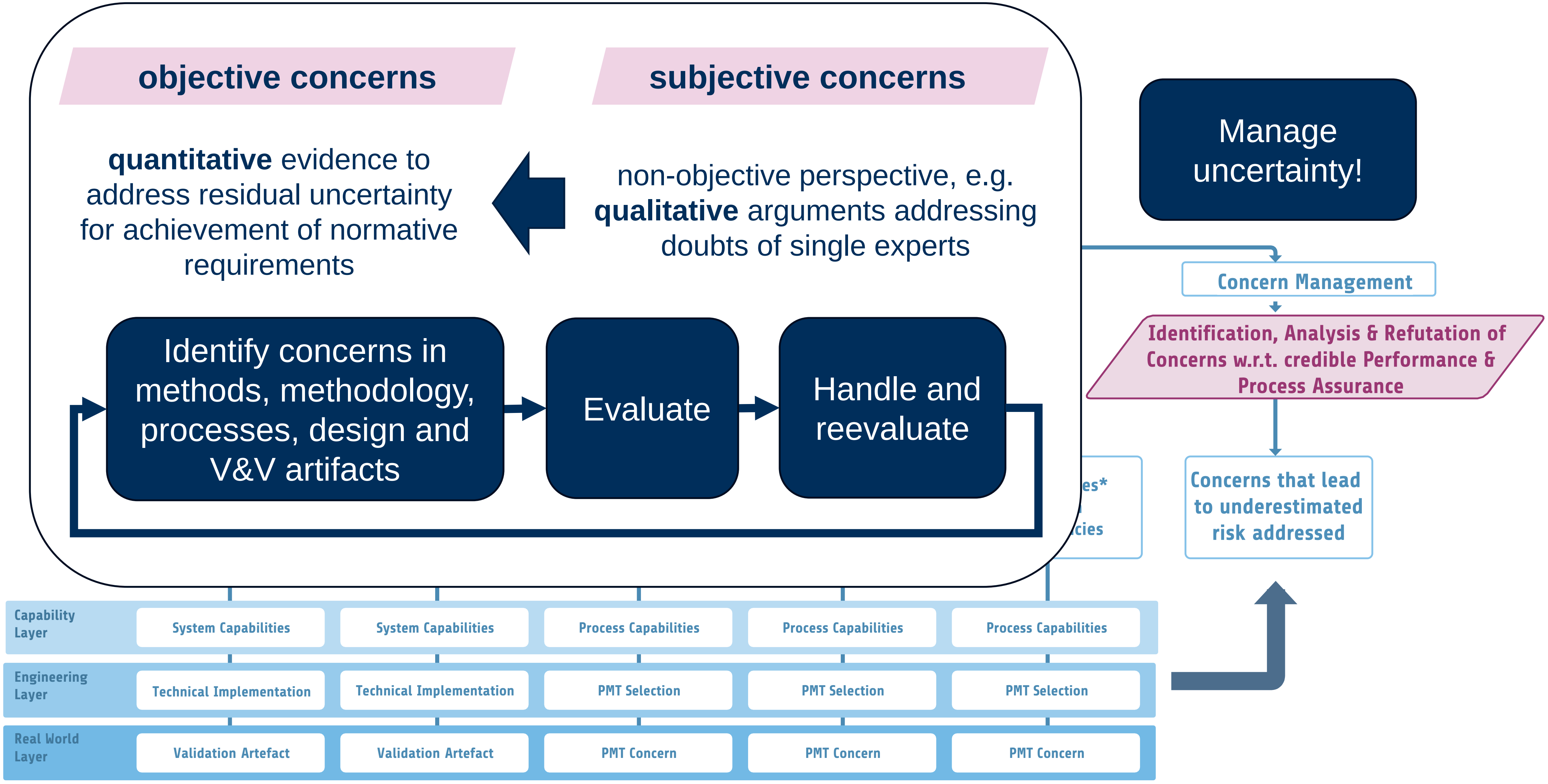
Techn. Evaluation

Engineering Layer (controlled scenarios)

Real World Layer (uncontrolled scenarios)

* processes are core in-house know how, VVM focusses on methods and problem space description. This leaves concrete process arguments in the hands of individual organizations.

Argument validity = Why we are right and why aren‘t we wrong!



objective concerns

quantitative evidence to address residual uncertainty for achievement of normative requirements

subjective concerns

non-objective perspective, e.g. **qualitative** arguments addressing doubts of single experts

Identify concerns in methods, methodology, processes, design and V&V artifacts

Evaluate

Handle and reevaluate

Manage uncertainty!

Concern Management

Identification, Analysis & Refutation of Concerns w.r.t. credible Performance & Process Assurance

Concerns that lead to underestimated risk addressed

es*

cies

Capability Layer	System Capabilities	System Capabilities	Process Capabilities	Process Capabilities	Process Capabilities
Engineering Layer	Technical Implementation	Technical Implementation	PMT Selection	PMT Selection	PMT Selection
Real World Layer	Validation Artefact	Validation Artefact	PMT Concern	PMT Concern	PMT Concern

PMT = Processes - Methods - Tools

Summary and further material

- **The VVM safety argumentation structure**
  - enables telling an explicit ADS safety story for explainable compliance with normative requirements
  - helps justifying deployment decisions based on the VVM assurance framework
  - provides a starting point for argumentation teams in concrete ADS projects

- **Deep dives into VVM argumentation today in Stream 1**
  - The role of Argumentation – Overview Argumentation Strategies
  - Building Blocks of the Argumentation: ODD Coverage in the Scenario-based Approach
  - Building Blocks of the Argumentation: Behavior Specification
- Tomorrow: Embedding in international context and the look into the future

- Feel free to come over to discuss safety argumentation in breaks at posters **3.8 – 3.11** behind your seats ☺

VERIFICATION
VALIDATION
METHODS

Thank you!

Jan Reich, Fraunhofer IESE

jan.reich@iese.fraunhofer.de

Joint work with: Marcus Nolte (TU Braunschweig), Tino Brade (Robert Bosch GmbH), Marco Fistler (IAV GmbH contracted by BMW AG), Nayel Fabian Salem (TU Braunschweig) Christian Lalitsch-Schneider (ZF Friedrichshafen AG)

PROJECT of the PEGASUS FAMILY

VDA LEITINITIATIVE

A project developed by the VDA Leitinitiative
autonomous and connected driving

Supported by:

Federal Ministry
for Economic Affairs
and Climate Action

on the basis of a decision
by the German Bundestag

Literature referenced in this presentation

- [1] ISO/TR 4804:2020 “Road vehicles - Safety and cybersecurity for automated driving systems- Design, verification and validation”
- [2] PEGASUS Project (2019), „PEGASUS Safety Argumentation“, https://www.pegasusprojekt.de/files/tmpl/pdf/PEGASUS%20Safety%20Argumentation.pdf
- [3] Favaro et al. (2023) „Building a Credible Case for Safety: Waymo's Approach for the Determination of Absence of Unreasonable Risk”. www.waymo.com/safety
- [4] MISRA (2019) „Guidelines for Automotive Safety Arguments“
- [5] Goal Structuring Notation Community Standard (Version 3) (2019) https://goalstructuringnotation.info/