Federal Office
for Information Security

Informal document **GRVA-21-27**
21st GRVA, 20-24 January 2025
Provisional agenda item 3

AI/ National German Activity: BSI Technical Guideline "Process Guidelines for Derivation and Practical Evaluation of AI Security Requirements in Automotive"

Federal Office for Information Security [BSI], Germany)

WP.29/GRVA 21st session, January 20th 2025, Palais des Nations Geneva Switzerland

Introduction and Motivation

- In Germany we have identified the **challenge of operationalizing horizontal AI-regulations**, e.g. the EU AI Act, in the automotive domain (noting possible corresponding developments worldwide, e.g. in China and US)

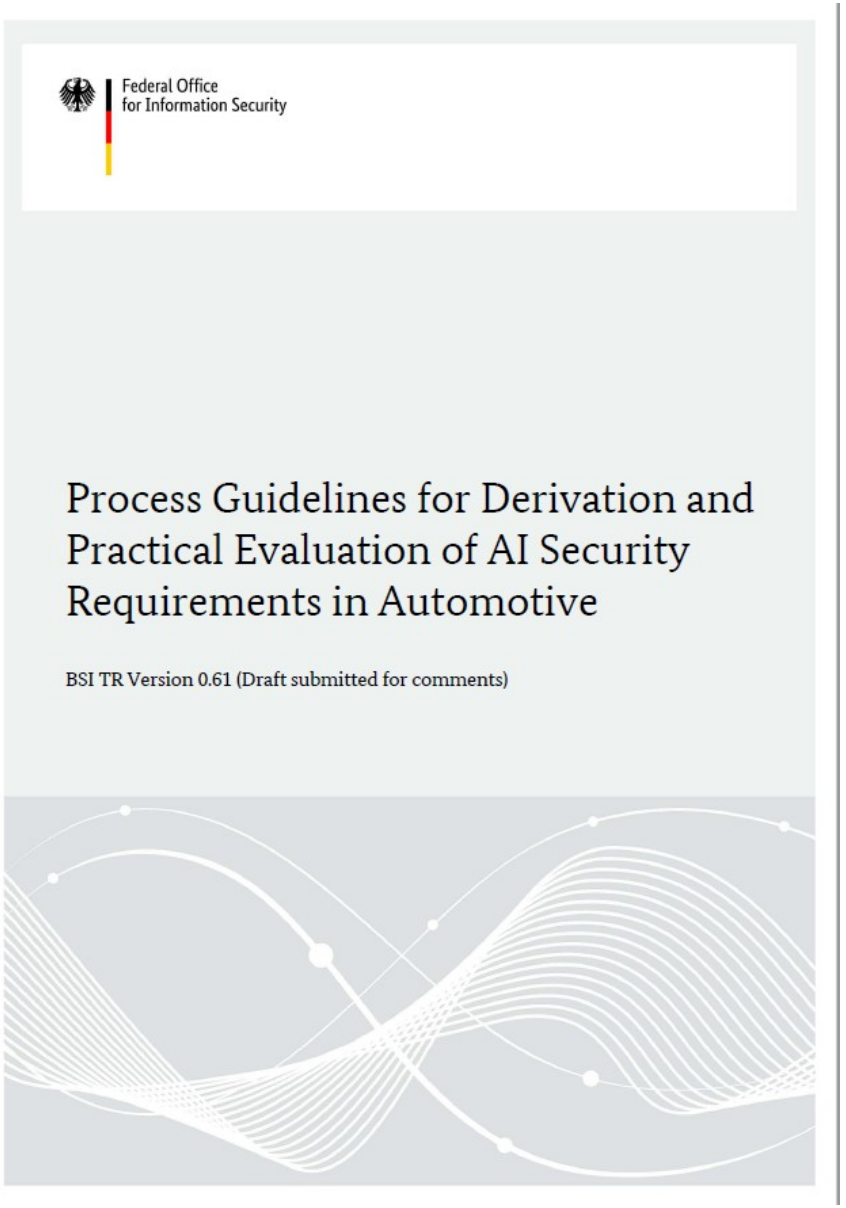
- In the project AIMobilityAudit we **worked out a process to translate existing domain-specific security and safety standards to generic and use-case-specific AI requirements and audit procedures** and presented this process togehter with exemplary audit results at the 20th GRVA session (09/2024)

- Here we present a **Draft Technical Guideline** based on the project results with a call for comments and **as a contribution to actively shape the auditing and regulation of AI systems within the automotive sector**

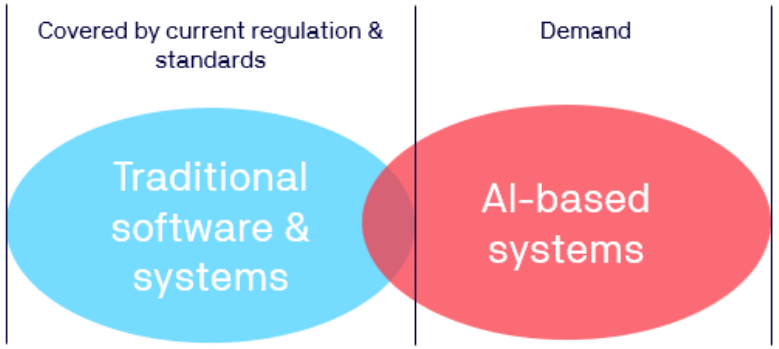
Structure of the Technical Guideline

1. Introduction
2. **Scope**
3. Terms and Definitions
4. Challenges in Compliance to Current Safety and Security Frameworks
5. **Generic Requirements for AI Systems**
6. **Generalized Audit Approach**

Appendix A.1 Exemplary Evaluation of AI Requirements based on a Use Case

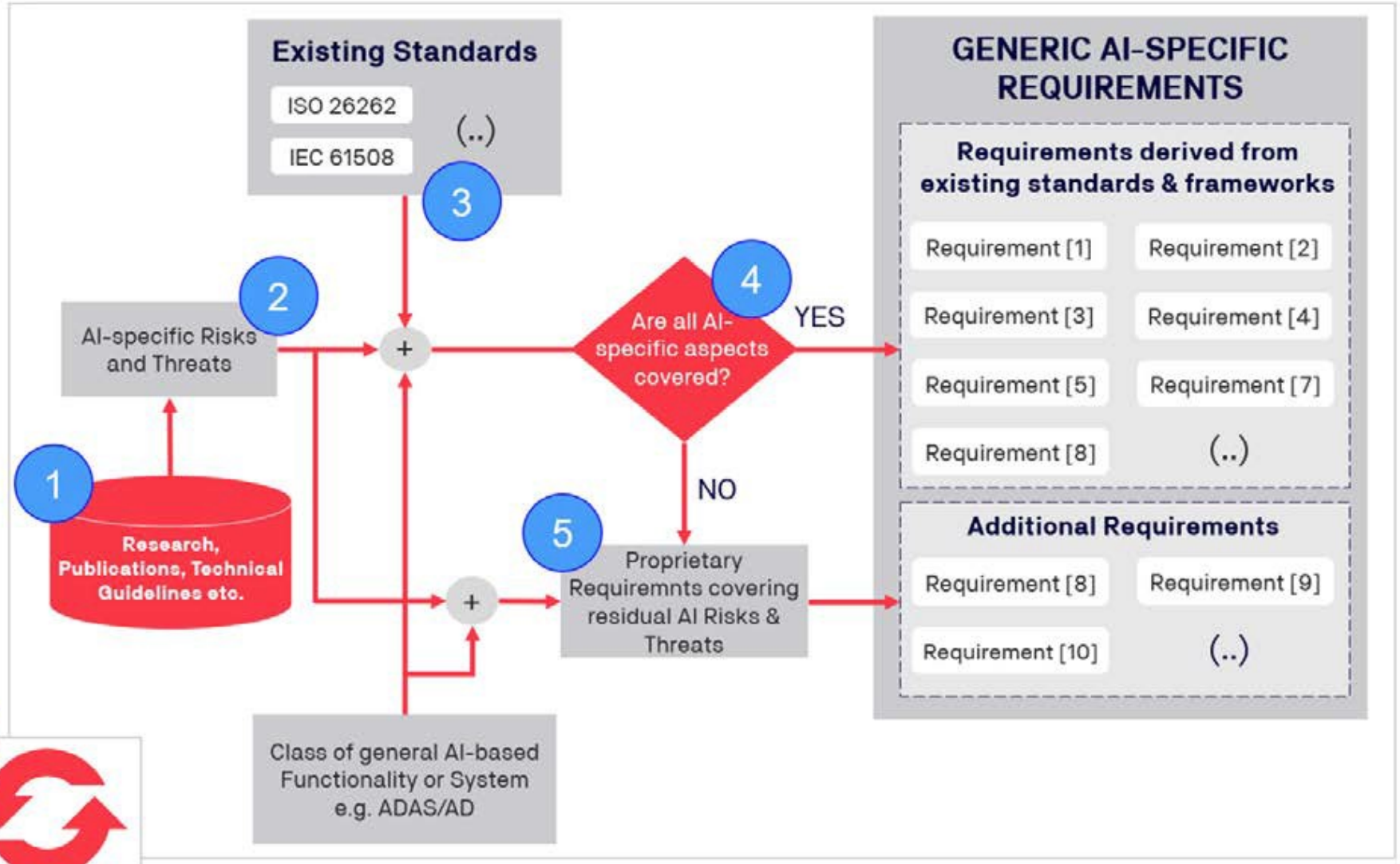


Section: Scope



- **Mapping and extension of requirements from applicable automotive safety standards** (ISO 26262, HARA methodology, ISO 21448) **for AI specific properties**
- **Robustness against AI related cybersecurity attacks** in accordance with established AI security frameworks and state-of-the-art research.
- **List of generic, use case-independent requirements**, adaptable to specific use cases and risk levels
- **Generalized iterative audit approach** to standardize and acquire practical knowledge for auditing AI systems, particularly addressing the gap of standardization and established thresholds for safety and security-critical AI systems
- **Transition from generic to specific audit requirements**, focusing on exploring methods and sources for defining and selecting thresholds and metrics for AI systems that aim to replicate non-quantifiable human behavior

Section: Generic Requirements for AI-Systems (1): Pre-requisites & Recommendation for Requirement Elicitation



Federal Office for Information Secur

Section: Generic Requirements for AI-Systems (2): List of 15 Generic Requirements for AI-Systems in Automotive

ID	Description	Life-Cycle-Categories	Category
1	The AI design and development process shall adhere to existing standards and regulations, and it shall be tracked and documented.	Design & Development	Design & Development
2	The system shall implement safety mechanisms to prevent failures of the AI component.	Design & Development, Operation & Monitoring	Design & Development
3	The least complex AI model architecture shall be chosen to limit risks and enhance explainability.	Design & Development, Verification & Validation	Design & Development
4	The datasets shall be managed according to standardized methods, and all key processes shall be well-documented.	Design & Development, Verification & Validation	Data Management
5	The datasets shall undergo quality assessments and be adequately prepared for training and testing.	Design & Development, Verification & Validation	Data Management
6	The AI system shall be developed, tested, and operated within its operational design domain.	Design & Development, Verification & Validation, Deployment, Operation & Monitoring	Performance
7	The AI model shall consistently meet performance requirements.	Verification & Validation, Deployment, Operation & Monitoring	Performance
8	The AI model and system shall be tested against test scenarios created by domain experts.	Verification & Validation	Performance

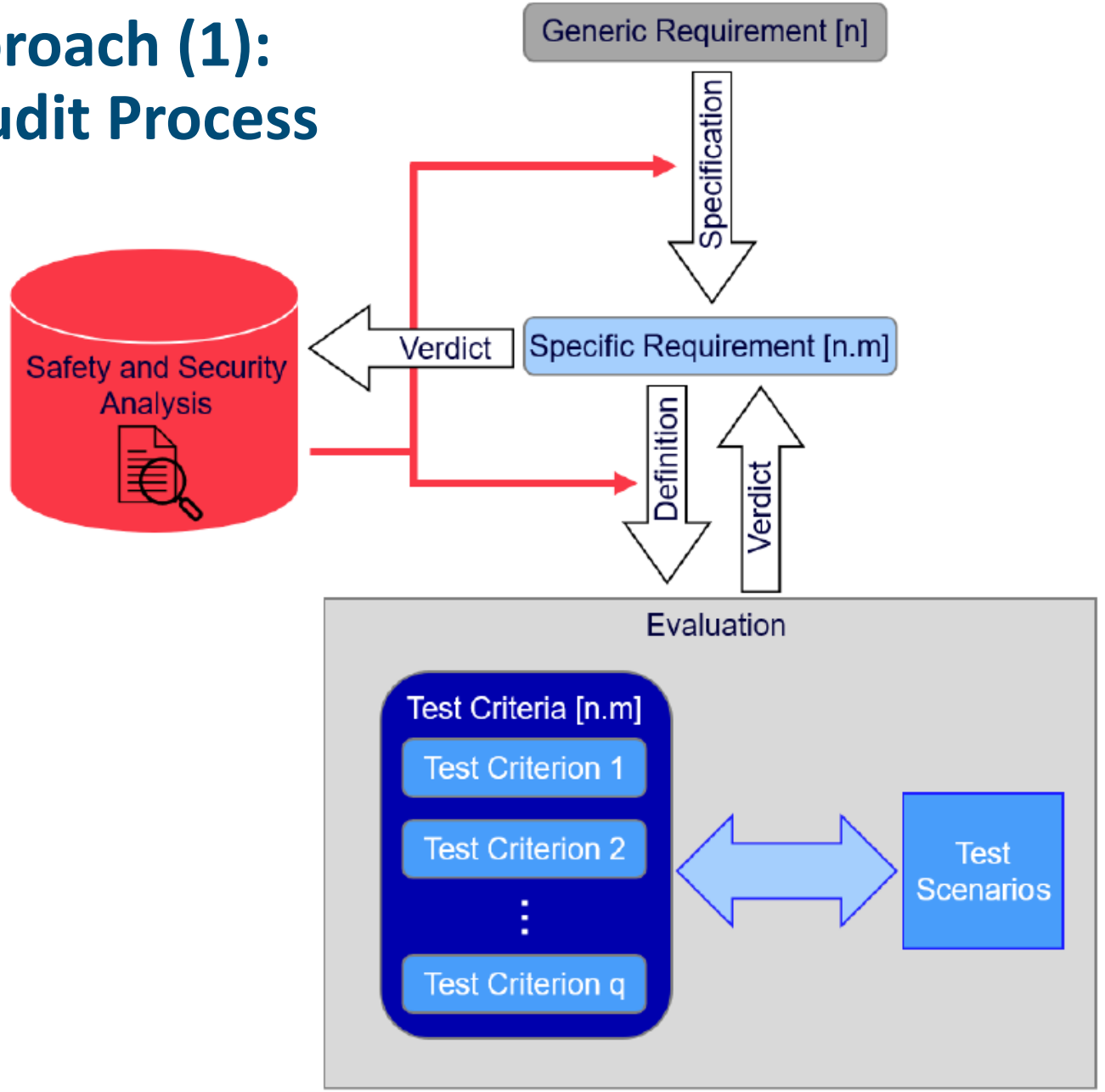
...

Federal Office for Information Security

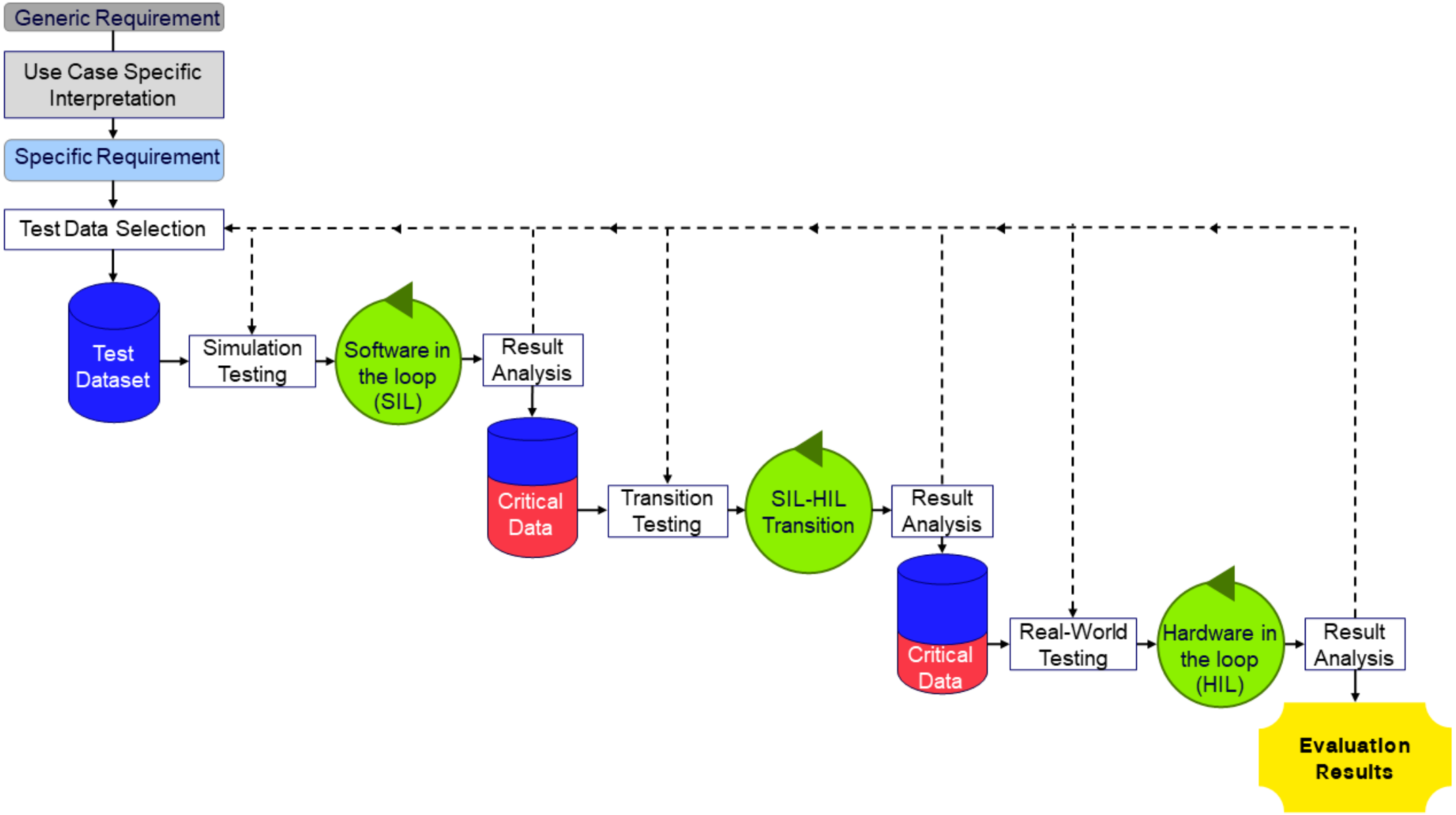
Section: Generalized Audit Approach (1): Iterative Evaluation Scheme/Audit Process

- **Sources to support the definition process** may include input and alignment from real-world feedback, simulations (Hardware-in-the-Loop/Software-in-the-Loop), and benchmarks of human performance

- **Multiple challenges within the process of defining precise values and thresholds**, e.g. high variety of input data and human performance variability



Section: Generalized Audit Approach (2): Testing Activities

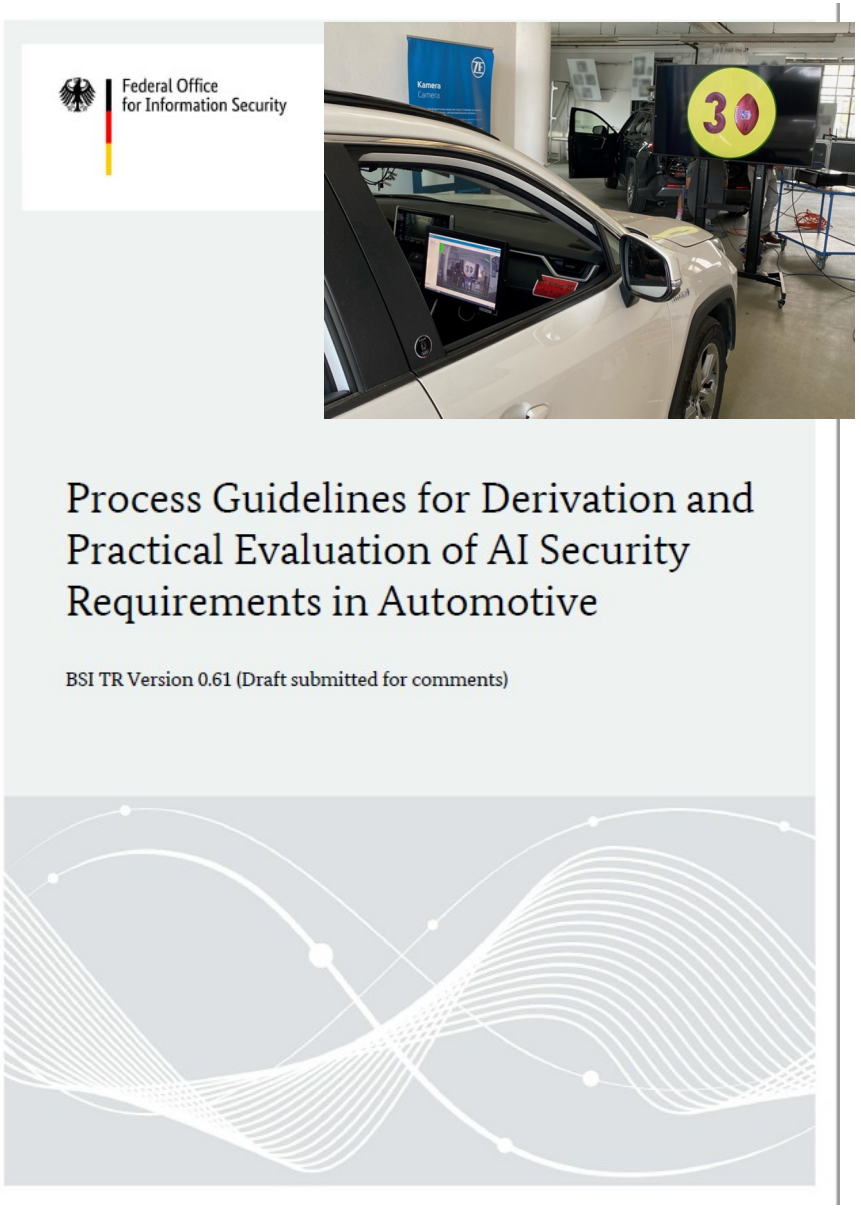


Appendix A.1 Exemplary Evaluation of AI Requirements based on a Use Case – Demonstrate the Practical Feasibility

- **Real-world automotive use case** (road user detection) **to evaluate an exemplary AI requirement** (the system shall be robust against relevant AI-related threats)
- **HARA:** 1. non-detection, 2. false detection, 3. false classification
- **Specification:** the system shall be robust against adversarial patches
- **Definition of test criteria:** the system shall be robust when exposed to a series of (1) black-box and (2) white-box adversarial body patches.
- **Dynamic & static tests** in simulation, transition testing and real world

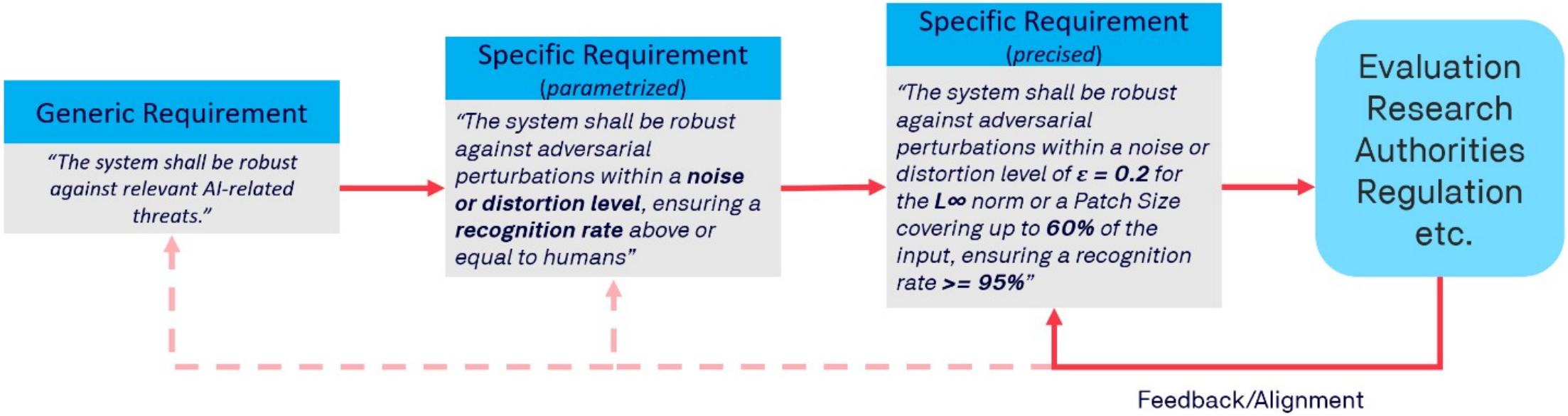
Next Steps

- **Coordinate draft technical guideline** (TG) with interested parties from government, industry and research
- **Apply TG to other use cases** and update according to evaluation results
- **Use TG as a contribution to shape the process of AI regulation for the automotive sector from within the sector**
- **Align TG with other approaches** and discuss open questions, e.g. in joint working group



Exemplary Open Questions

- How can reproducability and comparability of the generic approach be achieved as e.g. in the type approval process? How to align thresholds to achieve comparability?
- How can the iterative approach with continouos refinement be integrated into the process of type approaval and market surveillance?



Thank you for your attention!

Contact

Dr. Arndt von Twickel

Head of Division „Cybersecurity for Intelligent Transport Systems and Industry 4.0“

arndt.twickel@bsi.bund.de

Federal Office for Information Security (BSI)
Godesberger Allee 87
53175 Bonn
www.bsi.bund.de

Bild: ©Adobe Stock, Bee