



Evaluation Standards for Conformity Assessment of Trustworthy Cloud-based AI Applications

Daniel Loevenich, 16. Mai 2022

Agenda

Understanding AI Applications

Understanding ISO/IEC AI Standards

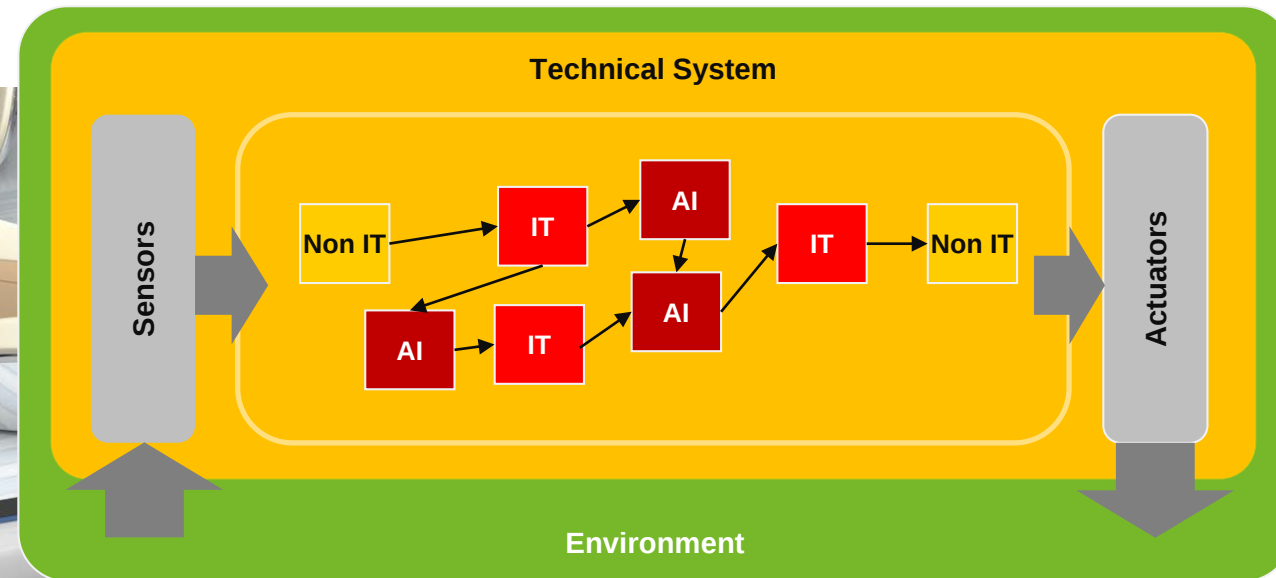
Horizontal Trustworthy AI Conformity Assessment

Project Success Factors

A uniform Framework for Horizontal AI Conformity Assessment



Understanding AI Applications



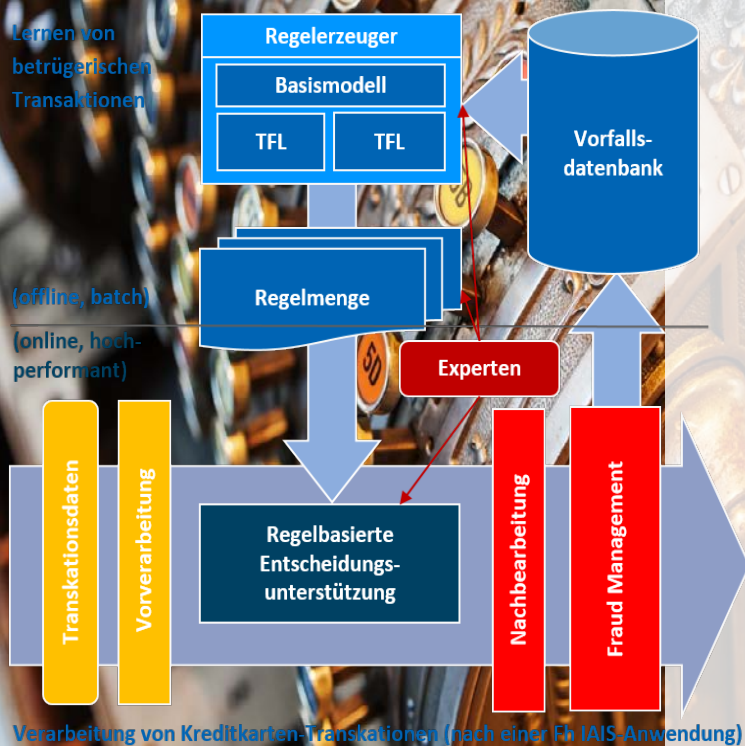


Understanding AI Applications



- **Conformity Assessment** considers the complex technical system (e.g. a vehicle) as a whole entity in the application context
- **Risks and requirements for the AI-components** have to be derived from this
- Such **Operationalizations** are essentially **mappings** into AI requirements
- This results in evaluation requirements for each component of usually hybrid AI solutions (**application-agnostic** perspective)

Lernen von betrügerischen Transaktionen



Understanding AI Applications



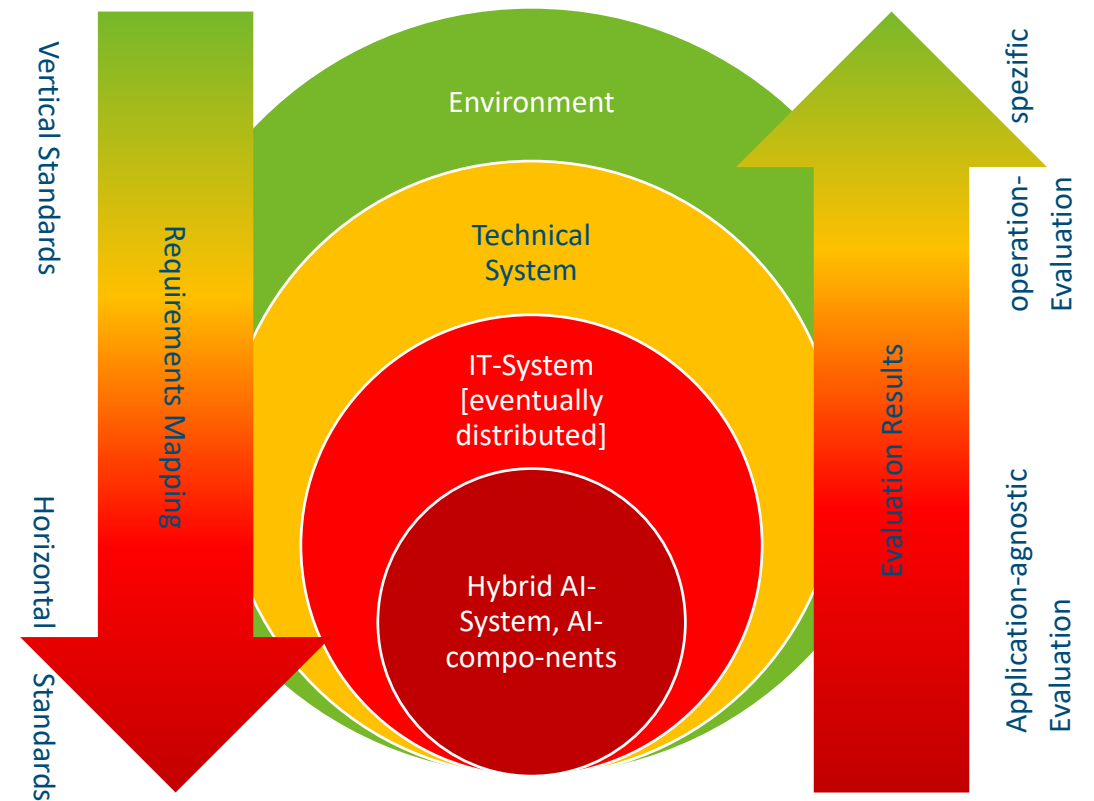
- A technical system usually follows **existing test regulations for supply chains** with multiple actors, e.g. OEMs
- Industry expects **cascadable evaluations for (distributed) AI-systems**
- AI test results must be **composable** (for instance in case of individual component checks of hybrid systems)
- A horizontal standard for trustworthy AI must be embeddable into AI management systems (**AIMS**)



Understanding ISO/IEC AI Standards: Testing and Evaluation

Horizontal and Vertical Standards Concept:

- **Downgrading Risks** to AI Components
- **Conformity Testing and Evaluation** of AI Components (all CA types)
- Upgrading and **Composition** of Evaluation Results
- **Uniform Conformity Assessment Framework**
- Application of schemes and **introduction to markets worldwide**





Horizontal Trustworthy AI Conformity Assessment: Strategic Objectives

Trustworthiness of the **entire Supply Chain** becomes transparent with **Conformity Assessment**

The necessary **evaluation criteria and test procedures** are to be developed

This evaluation bases shall be **applicable to hybrid solutions** as well as **embedded components**

EU-AIA: The evaluation bases serve as foundation for a **horizontal AI standard**

An **easy market access for SMEs** with acceptable costs will be facilitated

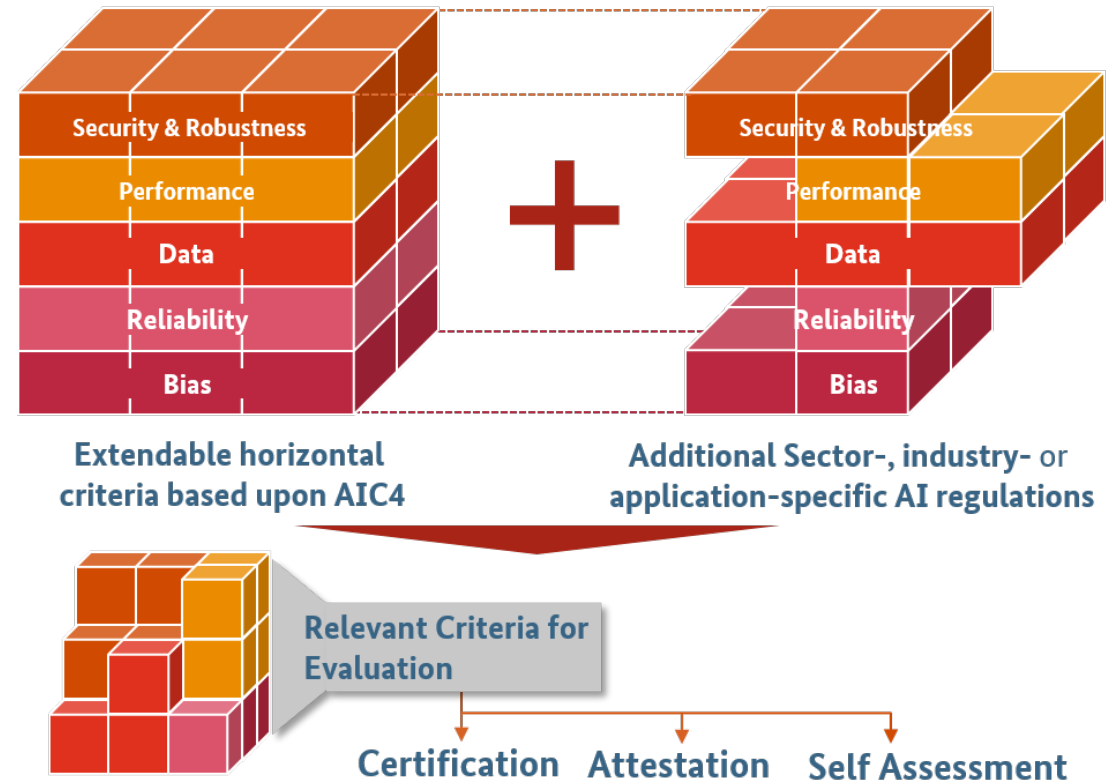
Vertical and sectoral standards should be based a horizontal standard for trustworthy AI

Uniform Evaluation Framework for Trustworthy AI Applications

Horizontal Trustworthy AI Conformity Assessment: Deliverables

An extendable set of **horizontal, application agnostic criteria** (“Trustworthy AI Services Evaluation Criteria”),

An extendable set of valid **AI evaluation procedures** (“Trustworthy AI Services Evaluation Methodology”), applicable to all three types of conformity assessment (self assessment, attestation and certification),



Horizontal Trustworthy AI Conformity Assessment: Deliverables

A proposal for an **application procedure** to extend the methodology and to implement it within the ongoing standardization process,

A proposal for an **application procedure** to extend this criteria and to implement the procedure within the ongoing standardization process,

A **procedure for mappings** of vertical application specific requirements into horizontal criteria requirements,

The **Guidance documents for production, application, and support** for all parties involved in the corresponding AI Eco-Systems. In the context, the framework establishes additional guidelines on how to integrate evaluation activities into an AIMS.

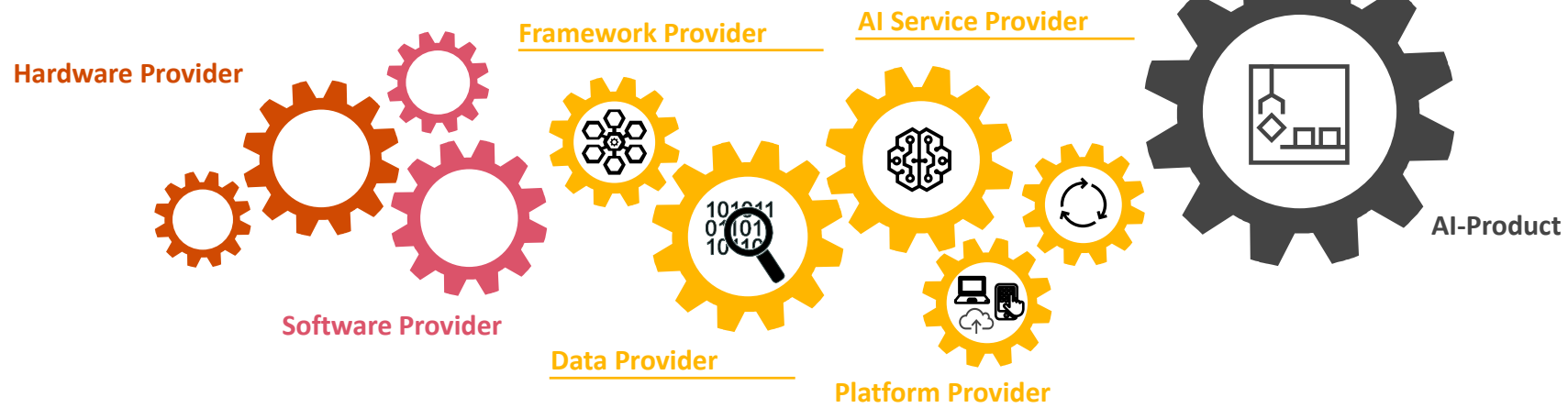
Horizontal Trustworthy AI Conformity Assessment: Project Roadmap

Definition of evaluation bases	Applicability and Market Penetration	Publication Phase	International Standardization
Q3 - 2022	Q1 - 2023	Q4 22/Q3 23	2025
Development of criteria, methodology, scheme with use cases	Validation and Acceptance on basis of relevant Use Cases	Publication within three Workshops - Europe, USA, Asia	Transfer and Harmonization in hEN/ISO Standard



Project Success Factors: Drivers for Technology Transfer

Cloud Providers drive AI technologies within the complete AI supply chain. They offer full service enterprise customer support with AI experts for development, IDEs, Frameworks and quality measurement tools and operate AI solutions continuously.



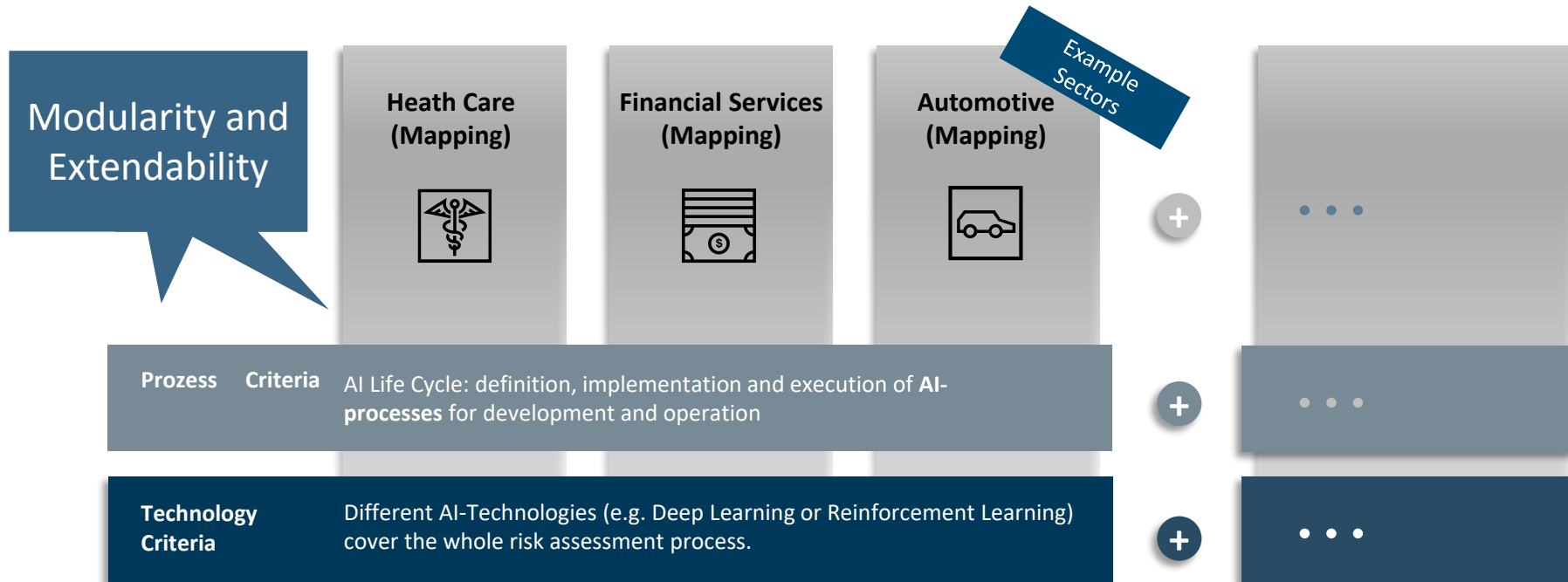
A project for an horizontal standard for trustworthy AI shall use these resources.



Project Success Factors: Use Case Projects Outline

	Phase 1:	Phase 2:	Phase 3:	Phase 4:
Focus:	Scoping, Risk Analysis & TOE	Evaluation Procedure Development	Procedure Execution	Tool and Framework Validation
Outcome (Dokuments):	Mapping System Description Minutes	Procedure Description Protocol	Evaluation Documents (dependable on the conformity assessment type) Evaluation Report	Evaluation Documents Evaluation Report
Format:	Workshops	Coordination Conference with unanimous vote	Depends on the type of conformity assessment	Certification Report
	Mandatory: Generalization of Requirements up to Criteria and Evaluation Process Definition		Optional: Pilot Evaluation, Procedure Validation, Assurance Methods Assessment (Levels)	

Project Success Factors : Sub-Projects and Use Cases



A uniform Framework for Horizontal AI Conformity Assessment: End User Acceptance

Open questions for dealing with AI in a trusting manner ...



How can responsible use of AI be demonstrated across the entire supply chain of an AI system?



Which requirements are relevant for my AI systems and how do I fulfill them according to my area of responsibility?



How can I efficiently implement requirements with regard to my use case - specific systems and processes?



How do I efficiently demonstrate compliance with increasing AI regulations?

... are addressed by applying the criteria and method documents



Provision of extendable and horizontal criteria for trustworthy AI applications

Provision of a Validation, Testing and Certification attempt

AI Service Providers as important multipliers



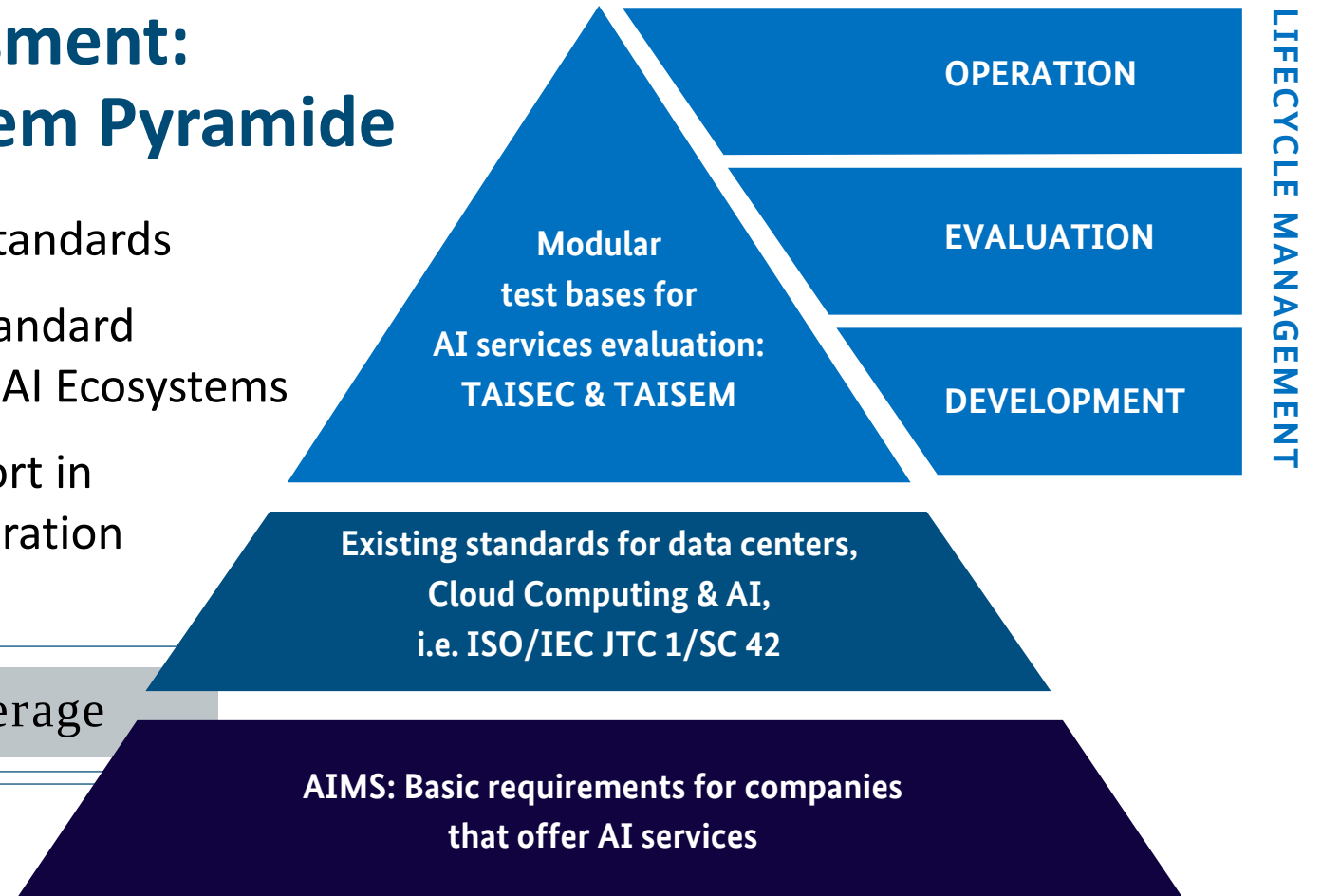
Bundesamt
für Sicherheit in der
Informationstechnik

Deutschland
Digital•Sicher•BSI

A uniform Framework for Horizontal AI Conformity Assessment: Standardization System Pyramide

Transfer of AI Trustworthiness Standards
Conformity Assessment of AI-Standard
Solutions lead to Acceptance of AI Ecosystems
Worldwide End Customer Support in
regulated sectors: Market Penetration

Complete AI Ecosystem Coverage



Project Partners

Development, Validation, Market Penetration

Hyperscaler:

- AWS
- Alibaba
- Google
- Huawei
- Microsoft
- Oracle
- SAP
- Dt. Telekom

Research Projects:

- DFKI
- Fh IAIS (Program „Certified AI“)
- With Use Cases: GAIA-X



Standardization

Evaluation, Certification and Accreditation:

- DIN/VDE
- CEN/CENELEC AI JTC21
- **ISO/IEC JTC 1/SC42**
- AI Quality und Testing Hub
- DFKI Osnabrück
- Fh IAIS
- PricewaterhouseCoopers
- Verband der TÜV



Thank you for your attention!

Deutschland
Digital•Sicher•BSI

Contact:

Daniel Loevenich

daniel.loevenich@bsi.bund.de

Tel. +49 (0) 228 9582 5395

Fax +49 (0) 228 10 9582 5395

Mobil +49 (0)171 30 29 824

Bundesamt für Sicherheit in der Informationstechnik (BSI)
Referat TK 23: Grundsatz, Strategie und Nachweise in der
Künstlichen Intelligenz

Godesberger Allee 185-189
53175 Bonn
www.bsi.bund.de



Das BSI als die Cyber-Sicherheitsbehörde des Bundes
gestaltet Informationssicherheit in der Digitalisierung
durch Prävention, Detektion und Reaktion
für Staat, Wirtschaft und Gesellschaft.



Bundesamt
für Sicherheit in der
Informationstechnik