

Europäischer Ansatz für Künstliche Intelligenz

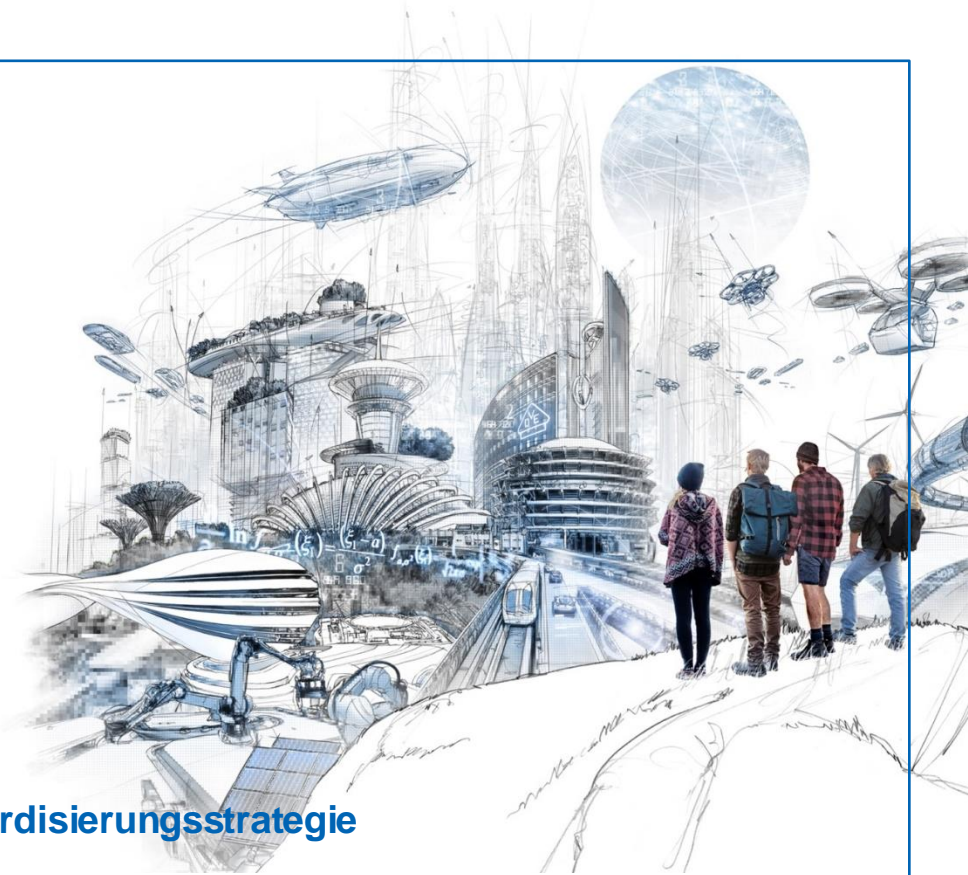
Dr. Sebastian Hallensleben

Head of Digitalisation & AI at VDE e.V.

Chair CEN-CENELEC JTC 21

Co-Chair Classification & Risk Assessment OECD ONE.AI

**Austrian Standards Expert Talk zur EU-Standardisierungsstrategie
2024-06-12**



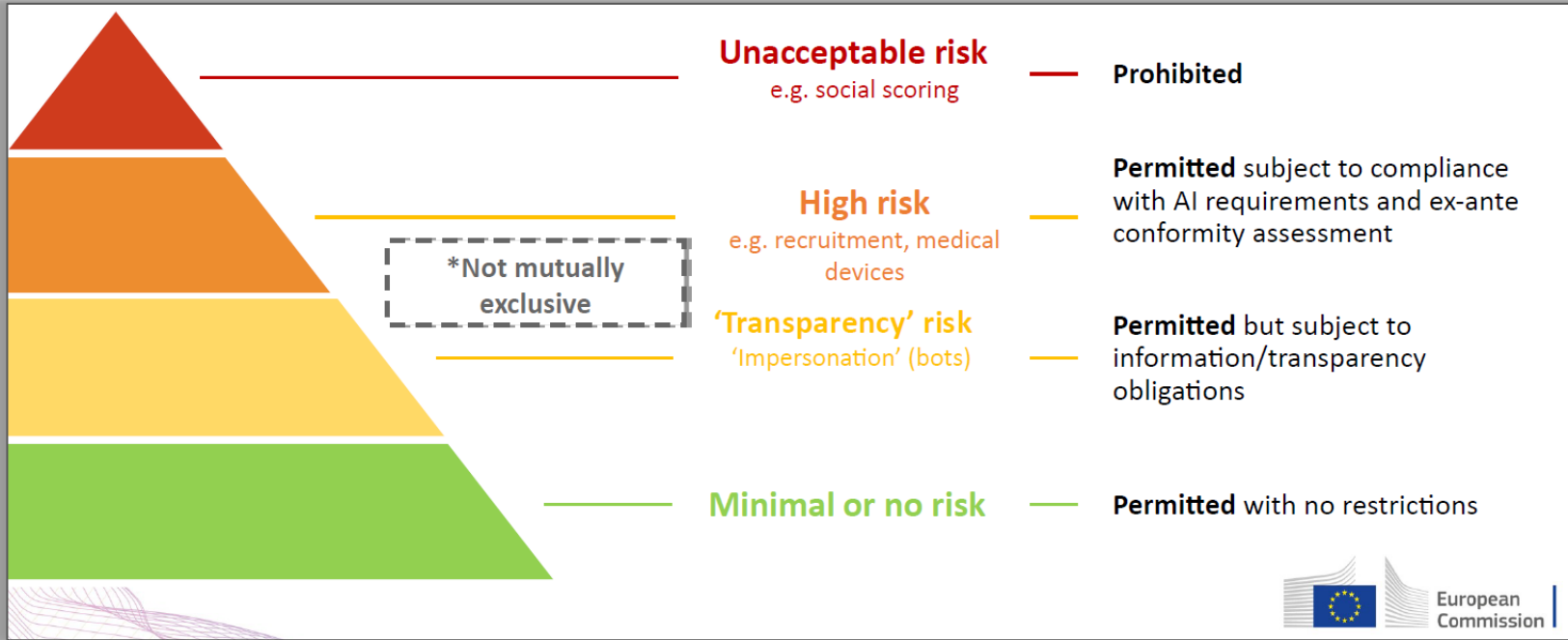
VDE

Entwicklung der harmonisierten Standards zum EU AI Act

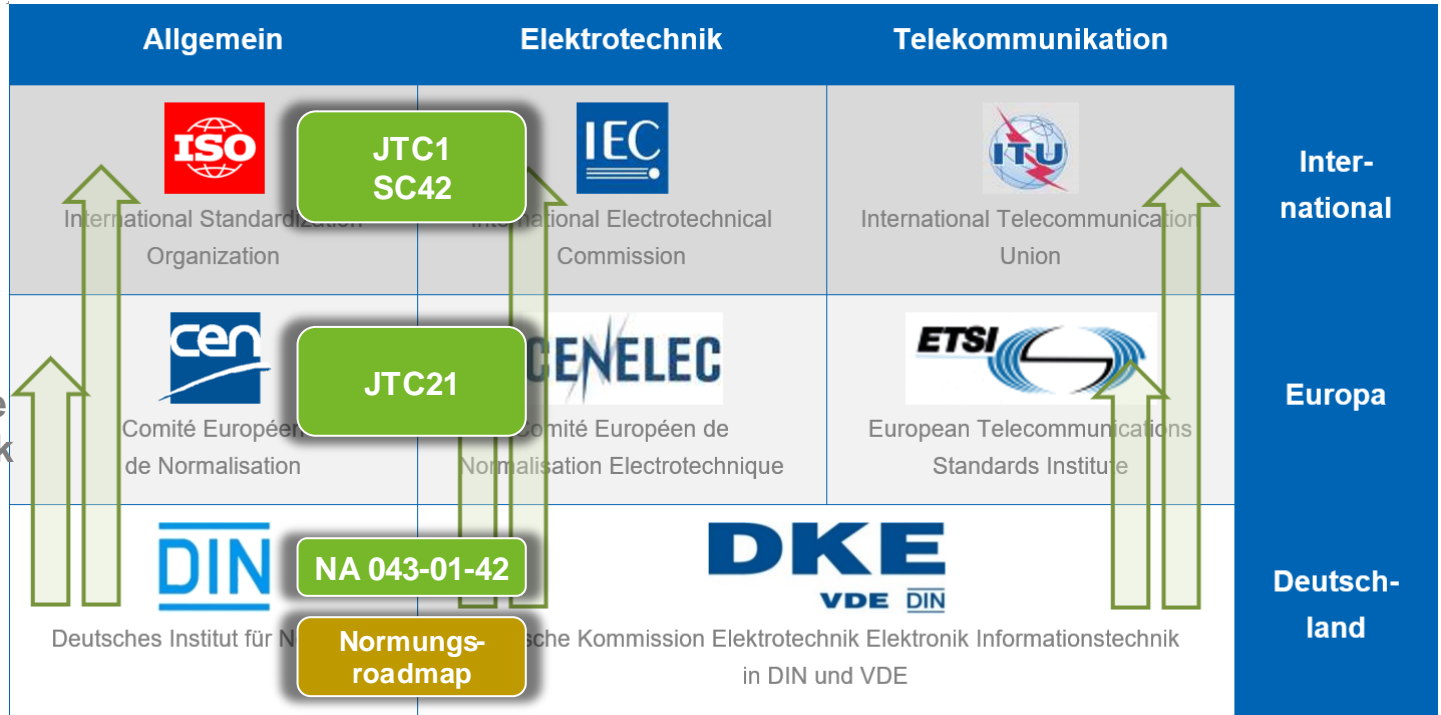
AI

EU AI Act:

High-level requirements only ⇒ Details to be defined elsewhere



JTC21 im Kontext der Standardisierungslandschaft



 New Legislative Framework

 AI Act



New Legislative Framework - Principles and structure

(as presented by the European Commission)



- ▶ **Essential requirements** designed to ensure a high-level of protection of public interests. They define the results to be attained, or the hazards to be dealt with, but do not specify the technical solutions for doing so
- ▶ **Harmonized standards** detailing technical solutions to meet the essential requirements
 - ▶ Voluntary – manufacturers can use other methods
 - ▶ Presumption of conformity with the essential requirements they cover
- ▶ **Division of responsibilities** along the value & distribution chain of the product
 - ▶ Manufacturers, importers, distributors, authorized representatives
- ▶ **Conformity assessment procedures**
 - ▶ Internal checks
 - ▶ Third-party assessment

VDE

Standardisation Request der EU

Ergänzungen angekündigt (Stand März 2024) :

11. nachhaltige KI

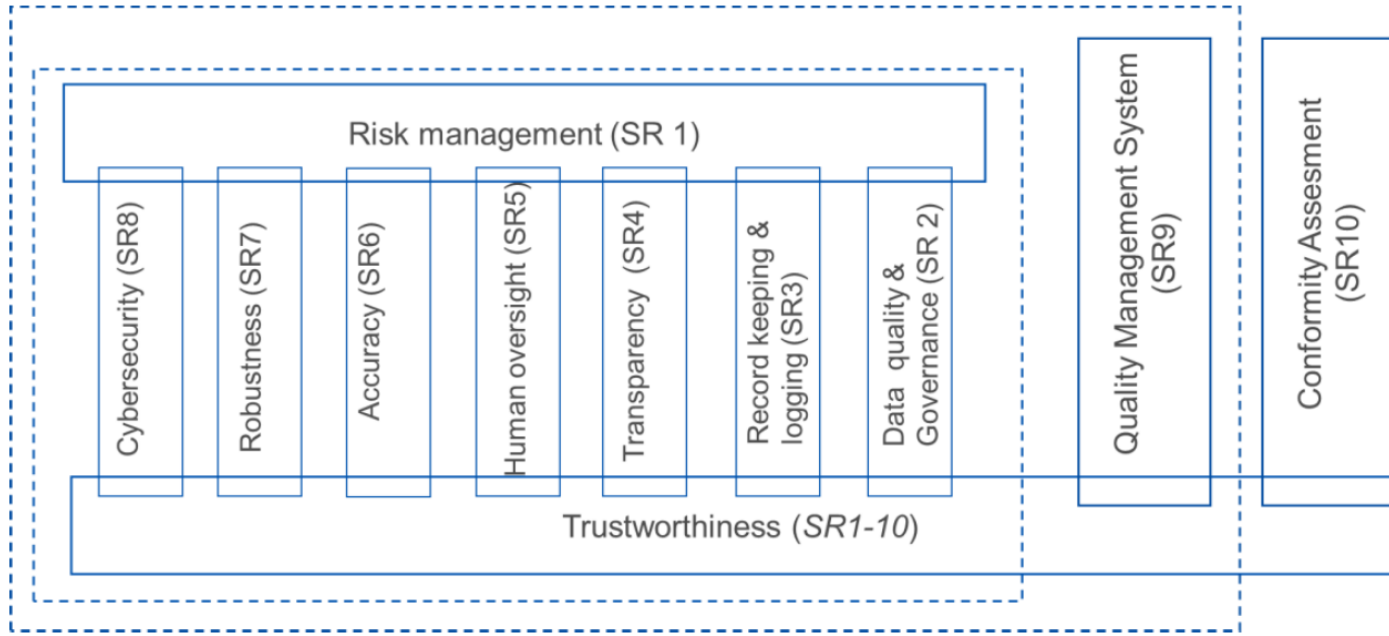
12. generative KI

1.	European standard(s) and/or European standardisation deliverable(s) on risk management system for AI systems
2.	European standard(s) and/or European standardisation deliverable(s) on governance and quality of datasets used to build AI systems
3.	European standard(s) and/or European standardisation deliverable(s) on record keeping through logging capabilities by AI systems
4.	European standard(s) and/or European standardisation deliverable(s) on transparency and information provisions to the users of AI systems
5.	European standard(s) and/or European standardisation deliverable(s) on human oversight of AI systems

6.	European standard(s) and/or European standardisation deliverable(s) on accuracy specifications for AI systems
7.	European standard(s) and/or European standardisation deliverable(s) on robustness specifications for AI systems
8.	European standard(s) and/or European standardisation deliverable(s) on cybersecurity specifications for AI systems
9.	European standard(s) and/or European standardisation deliverable(s) on quality management system for providers of AI systems, including post-market monitoring process
10.	European standard(s) and/or European standardisation deliverable(s) on conformity assessment for AI systems

Architecture of standards zur Strukturierung der inhaltlichen Arbeit

> 140 Experten
> 25 Länder



Working Groups



- **WG1: Strategic Advisory Group**
- **WG2: Operational Aspects**
- **WG3: Engineering Aspects**
- **WG4: Foundational and Societal Aspects**
- **WG5: Cybersecurity**

Zeitablauf

- Inhaltliche Arbeit für sämtliche erforderlichen Deliverables läuft, Strukturen stehen schon seit ca. Anfang 2023
- Fast die gesamte Aktivität des JTC21 zählt auf den Standardisation Request / AI Act ein
= **oberste Priorität**
- Entwürfe für die zu harmonisierenden Standards bis **Ende 2024**
- Veröffentlichung spätestens **Q4/2025**,
⇒ ca. **halbes Jahr Zeit zur Umsetzung** bis zum Wirksamwerden des AI Acts (Juni 2026)

Mitwirkung zusätzlicher Experten weiterhin erwünscht; laufende Information und Werbung

- Through your national AI mirror committee

- Through Annex 3 organisations



- Indirectly through liaisons

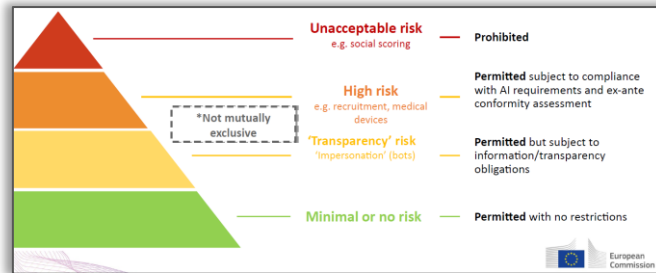
including other technical committees, associations, networks etc.

- Through ETSI

Mode 4 cooperation in place, including (but not limited to) cybersecurity

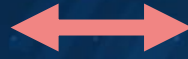
VDE

Complementing the EU AI Act: What else is needed beyond the scope of the EU AI Act?



+ ??

**Sicherheit,
Risikominimierung
und Compliance**



**Innovationsfreundlichkeit,
Wettbewerbsfähigkeit
und Einfachheit**

AI

Wie bringen wir beides unter einen Hut?

Wie kombiniert der Hersteller eines KI-gesteuerten Mähroboters Compliance und Wettbewerbsfähigkeit?



Wie wäre es mit einer einfachen, klaren und überzeugenden Antwort auf Fragen wie –

- Welche Eigenschaften hat die verbaute KI?
Wie sorgfältig wurde sie trainiert?
- Wie zuverlässig ist die KI?
Ist sie für jeden Menschen sicher?
- Was passiert mit den Daten, z.B. den Fotos aus meinem Garten?



Transparency	A	B	C	D	E	F	G
Accountability	A	B	C	D	E	F	G
Privacy	A	B	C	D	E	F	G
Fairness	A	B	C	D	E	F	G
Reliability	A	B	C	D	E	F	G

(not actual ratings for the lawnmower in the picture)

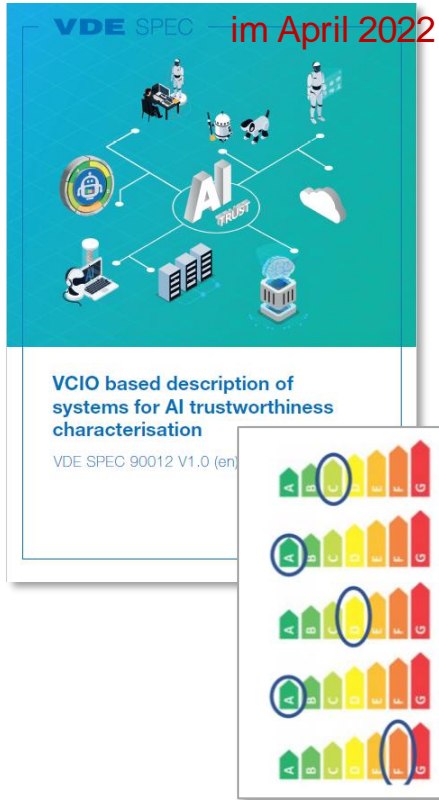
AI Trust Standard & Label – ein knappes Datenblatt für KI-Produkte



Konsortialstandard 2021/22


Version 1 veröffentlicht

im April 2022



VCIO based description of systems for AI trustworthiness characterisation

VDE SPEC 90012 V1.0 (en)



Digital Trust Forum



Transparency



T1. Disclosure of origin of data sets

T2. Accessibility

T1.1
Is the origin of the data documented?

Yes, comprehensive logging of all training and operating data, version control of data sets etc.

Yes, logging and version control through an intermediary (e.g. data supplier)

No logging. Data used is not controlled or documented in any way

T1.2
Is it for each individual use plausible, which data is being used?

Yes, the use of data and the individual application are intelligible

Yes, it is intelligible on an abstract, not case specific level, which data is being used

No, but a summary on the data usage is available

No

T1.3
Are the characteristics of the training data set documented and disclosed? Are the data sheets to the data sets comprehensive?

Yes and the data sheets are comprehensive

Yes, but the data sheet contains few or missing information

No

T2.1
Are the modes of interpretability oriented toward the needs of the target groups and developed with them?

Yes

Yes, but without participation of the target groups

Yes, but only toward one target group

No, only one mode of interpretability is developed without regard to target groups' needs

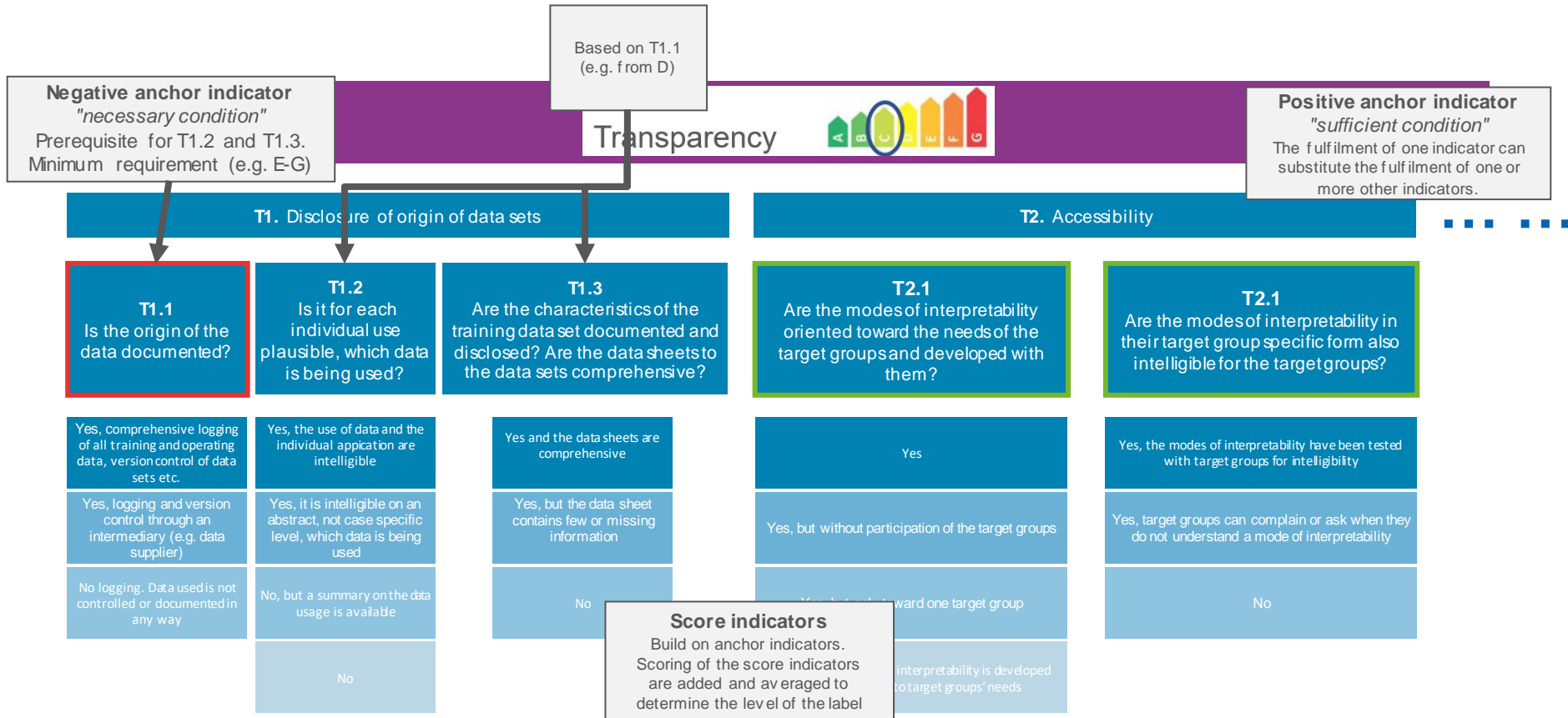
T2.1
Are the modes of interpretability in their target group specific form also intelligible for the target groups?

Yes, the modes of interpretability have been tested with target groups for intelligibility

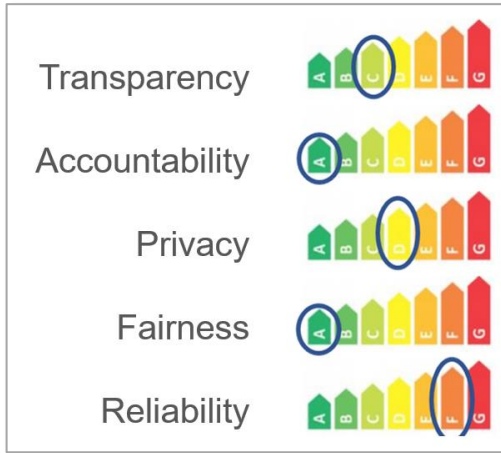
Yes, target groups can complain or ask when they do not understand a mode of interpretability

No





Der AI Trust Standard unterstützt AI Act Compliance und Wettbewerbsfähigkeit

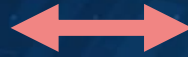


- ✓ erlaubt **positive Differenzierung** im Markt
- ✓ fördert **fairen Wettbewerb**
- ✓ gibt Herstellern und Einkäufern eine **gemeinsame Sprache zur Spezifizierung**
- ✓ ist **kompatibel mit dem AI Act** und entwickelt sich wechselseitig konsistent mit den harmonisierten Standards



Künstliche Intelligenz: Mit Standards zu Compliance und Wettbewerbsfähigkeit

Sicherheit,
Risikominimierung
und Compliance



Innovationsfreundlichkeit,
Wettbewerbsfähigkeit
und Einfachheit



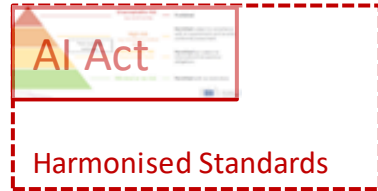
New
Legislative
Framework

JTC21



Completing the European AI ecosystem

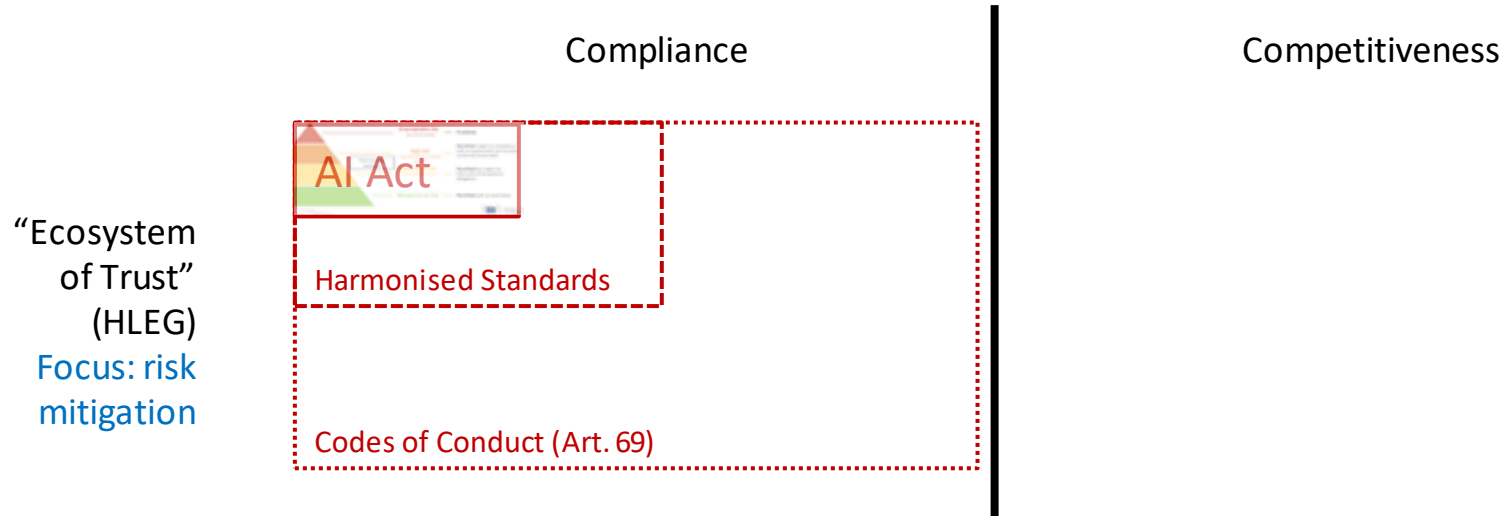
“Ecosystem
of Trust”
(HLEG)
Focus: risk
mitigation



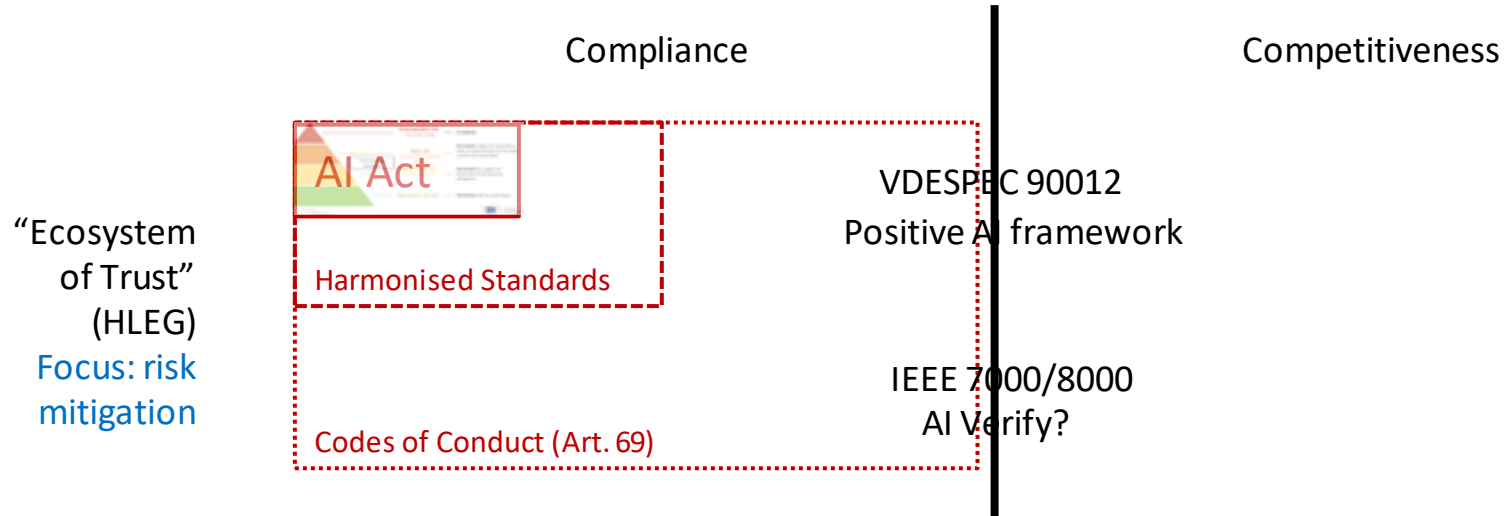
Compliance

Competitiveness

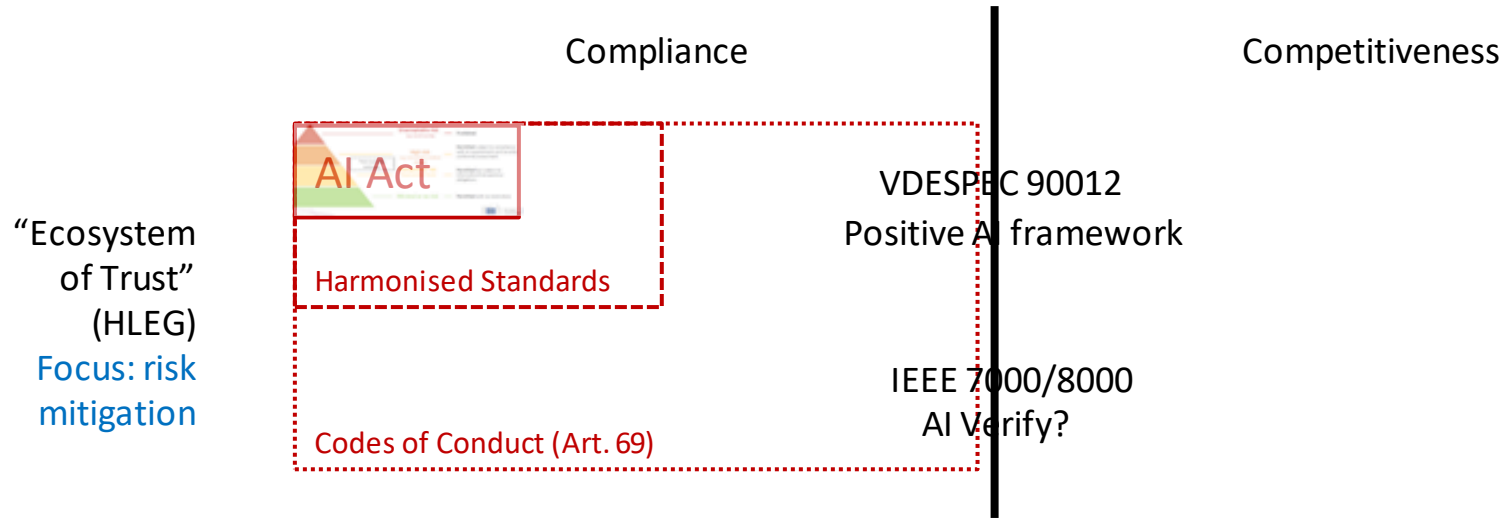
Completing the European AI ecosystem



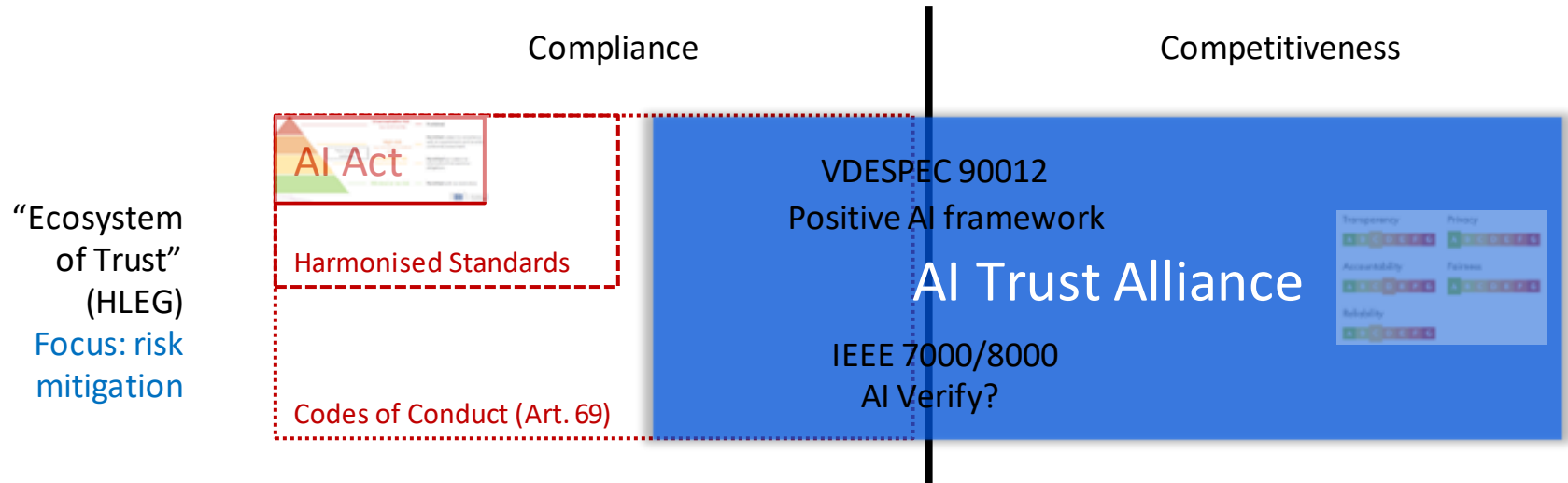
Completing the European AI ecosystem



Completing the European AI ecosystem



Completing the European AI ecosystem: AI Trust Alliance



VDE

POSITIVE AI



VDE

4 tracks of collaboration



Measuring the characteristics of products/organizations/people
(Specifications)



Communicating characteristics
(Labels)



Proving that standards are followed and labels are justified
(Certification, Auditing paths)



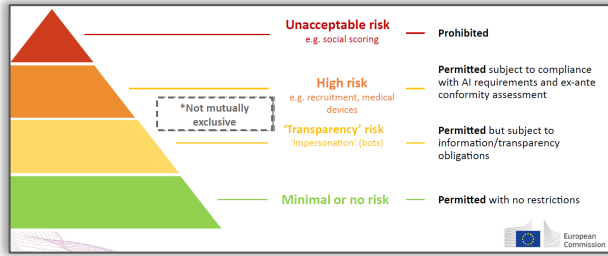
Implementing the label and **achieving** good ratings
(Tools, Automation, Training)

VDE

AI Trust Alliance: current stakeholders in the discussions



The future



Once we have the „optimal“ AI regulation and standards in place:
Are we done?

No.

Consider bad actors, especially for generative AI

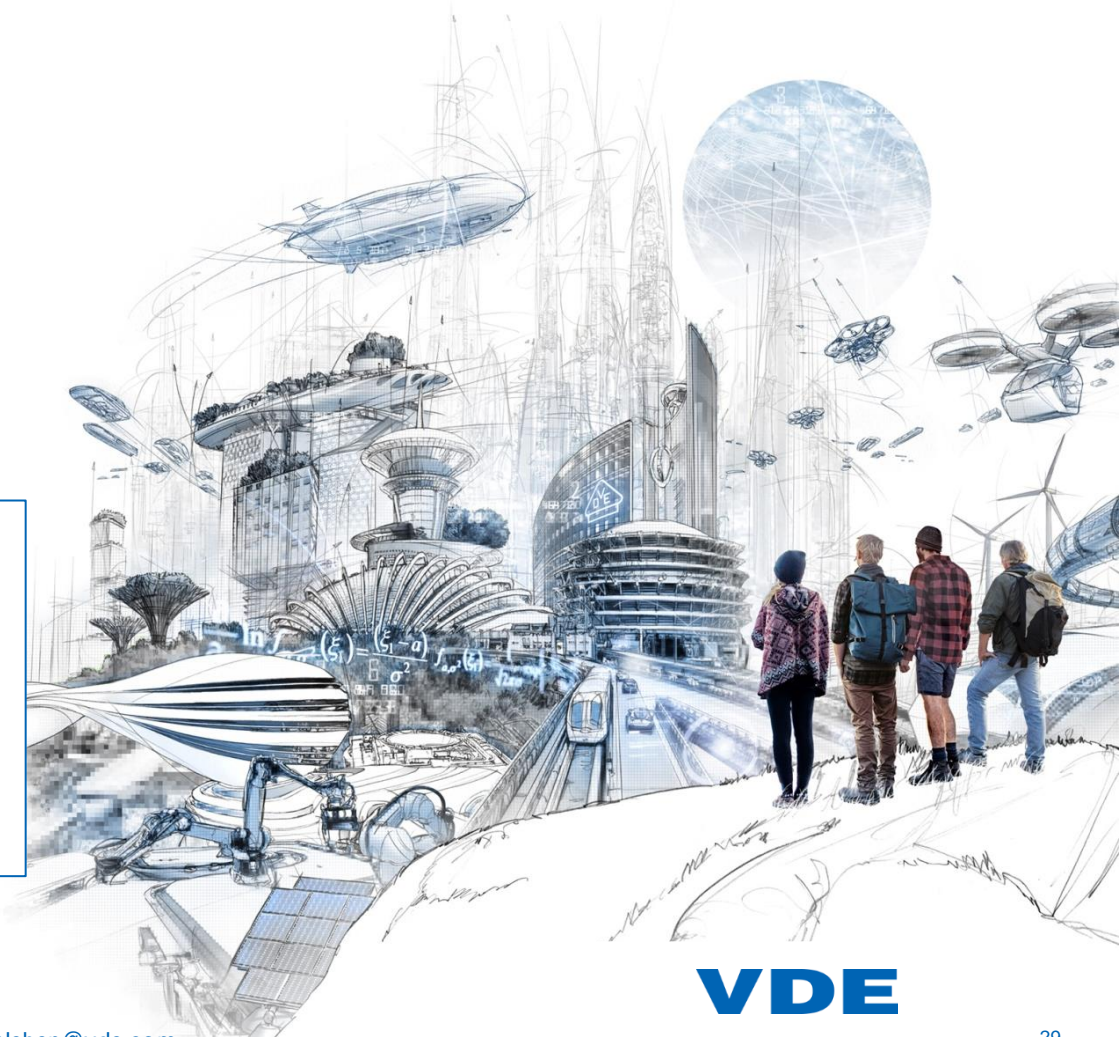
Build infrastructures and protocols for a **resilient** digital space,
especially around **trust** and **identity**



Vielen Dank!

Dr Sebastian Hallensleben
Head of Digitalisation and AI

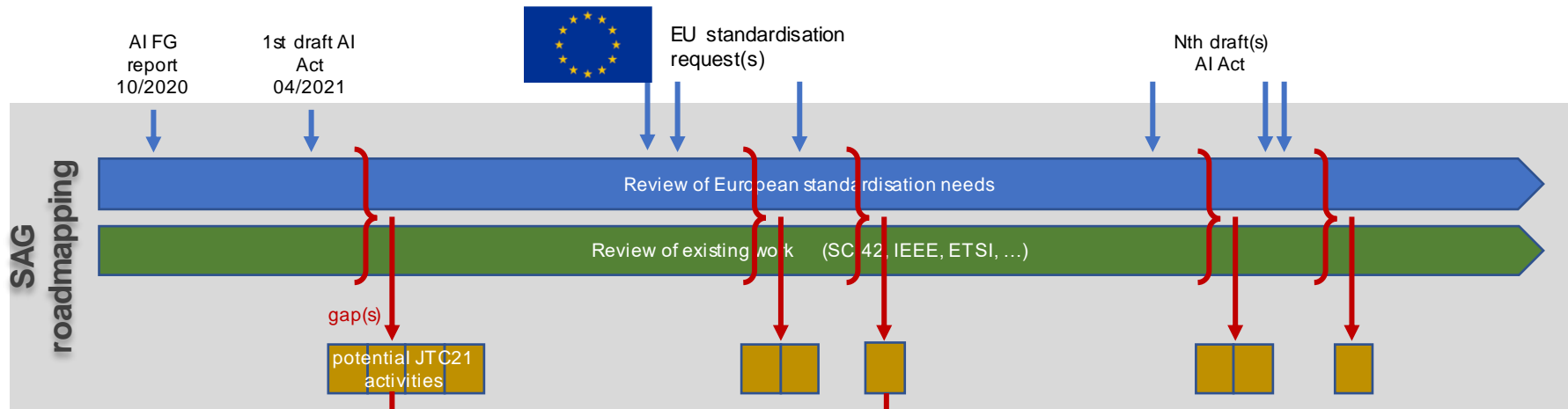
Tel: +49 170 791 6306
E-Mail: sebastian.hallensleben@vde.com



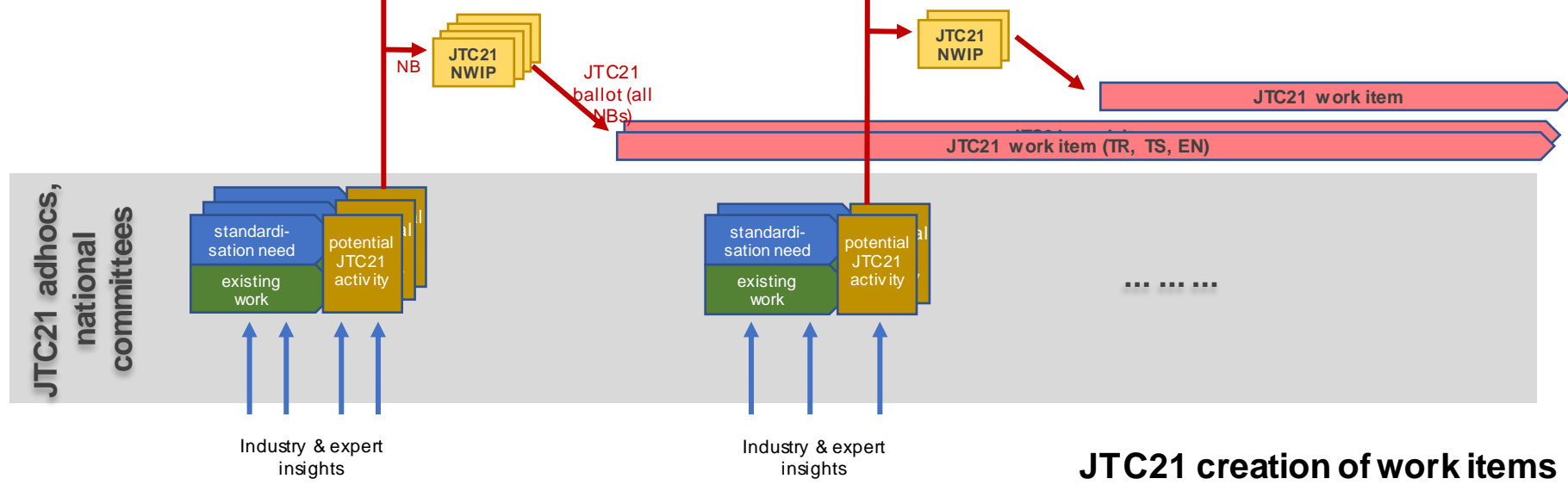
VDE

Backup

TOP DOWN



BOTTOM UP



JTC21 creation of work items

Inspired by the EU NLF: Constructing a binding AI framework at the international level

- AI Risk Management
- G7 Hiroshima / Generative AI

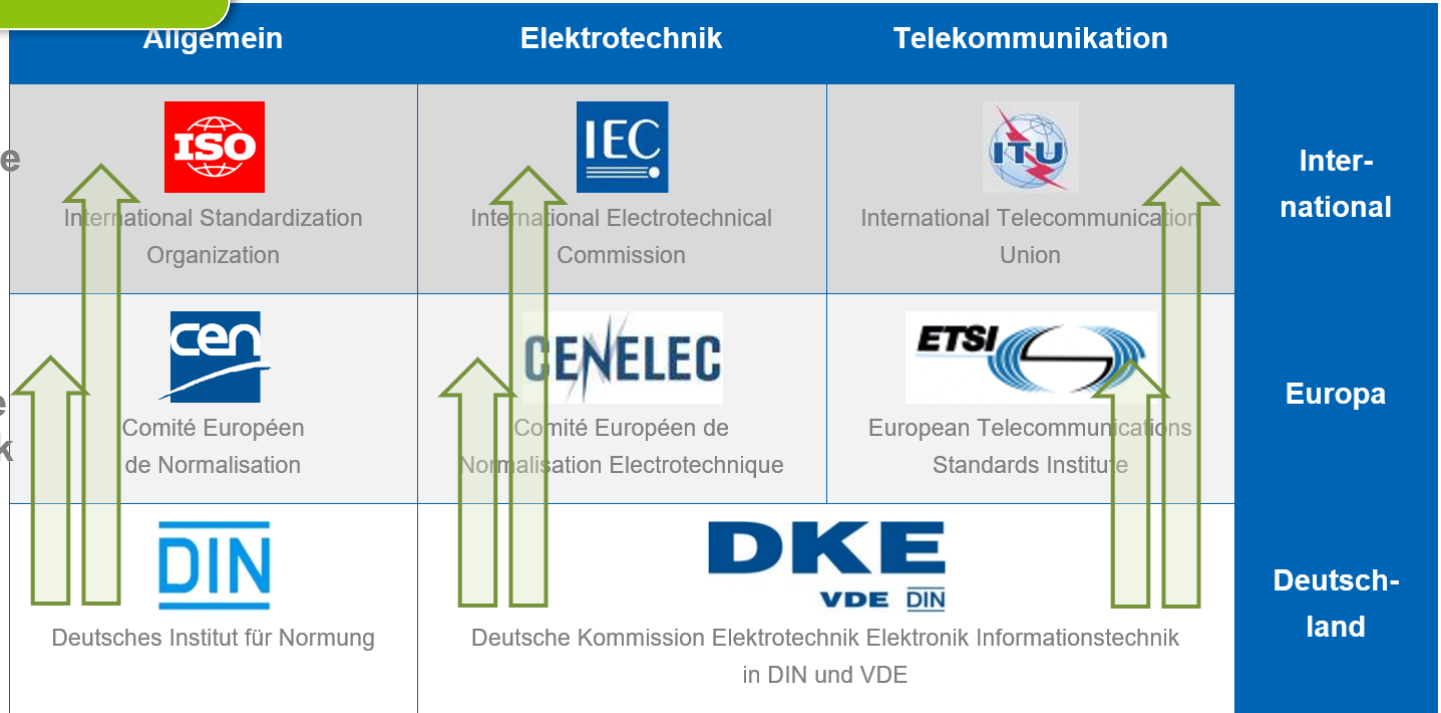


DDG for Responsible Business Conduct



New Legislative Framework

AI Act



Risiken und deren Mitigierung

Risiken / Herausforderungen	Mitigierung
Endliche Bandbreite in der europäischen Normung	<ul style="list-style-type: none">▪ Aufbau auf internationale Standards wo immer möglich▪ Direct adoption, wo ausreichend▪ Modified adoption als innovativer Prozess, von CEN-CENELEC bestätigt
Verzögerungen und inhaltliche Unwägbarkeiten in der internationalen Normung	<ul style="list-style-type: none">▪ Sparsame Verwendung von „parallel development“▪ Adoption nur, wenn intl. Standard bereits (fast) fertiggestellt, d.h. inhaltlich und zeitlich stabil sind
Hoher Anspruch des AI Acts; Interpretationsspielräume; HAS Assessment	<ul style="list-style-type: none">▪ Aktive Rolle der EU-Kommission (DG CNECT) in JTC21 bei Plenary Meetings sowie in den Arbeitsgruppen ⇒ frühzeitige Ausräumung von Unsicherheiten und Kontroversen▪ Verzicht auf den HAS-Prozess, stattdessen direkte Beurteilung durch DG CNECT
z.T. schwierige Konsensbildung durch divergierende Interessen	<ul style="list-style-type: none">▪ Regelmäßige Bereitstellung von Analysen durch den Joint Research Council▪ Intensive politische Unterstützung und Flankierung durch die EU-Kommission▪ Verdeutlichung der allseits unerfreulichen Auswirkungen von „common specifications“