

KI-Ethik und Vertrauenswürdigkeit

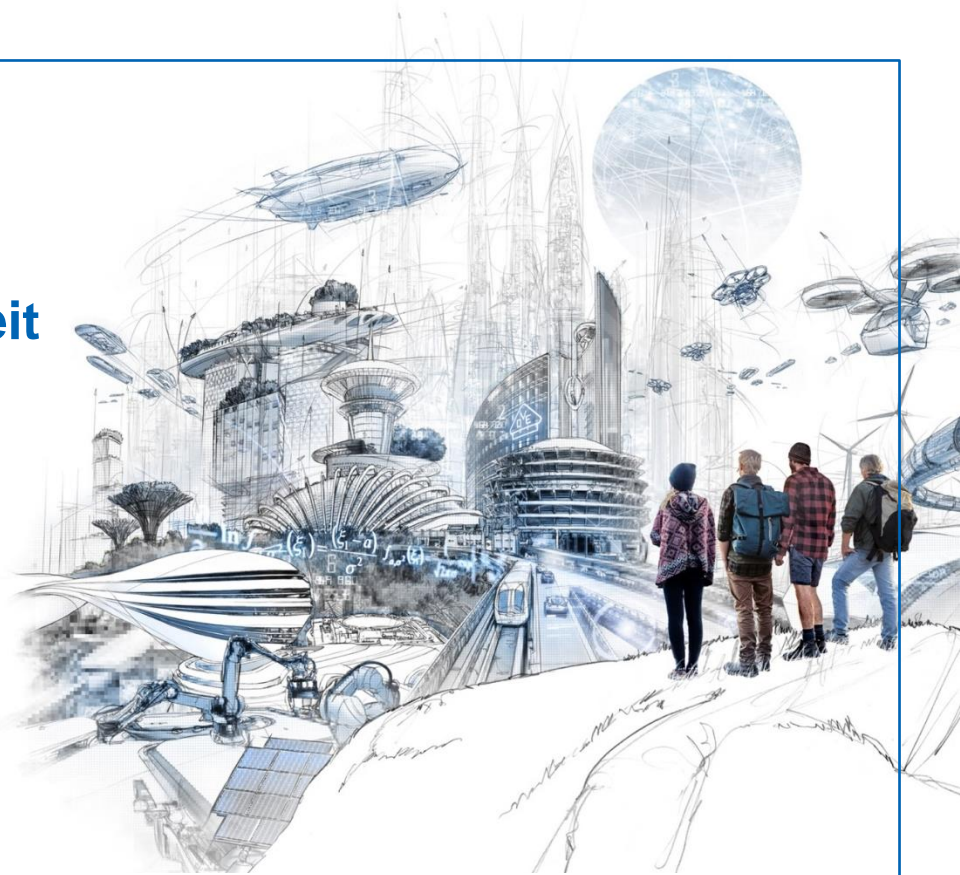
Regulierungs- und Standardisierungsrahmen

Andreas Hauschke

Projektmanager für vertrauenswürdige KI

Digitalisierung der Arbeitswelt

DGUV Congress Dresden



VDE

*“This proposal aims to implement the second objective for the development of an ecosystem of **trust** by proposing a legal framework for **trustworthy AI**. ”*

&

*“It supports the objective of the Union being a global leader in the development of secure, **trustworthy** and **ethical** artificial intelligence [...] and ensures the protection of **ethical principles** [...] ”*

EU AI Regulation – Proposal

Was bedeutet vertrauenswürdige KI?

- Zuverlässigkeit, Verfügbarkeit, Wartbarkeit
- Robustheit
- Benutzerfreundlichkeit
- Funktionale Sicherheit
- Informationssicherheit
- Datenschutz
- Ethik



Figure 9 – Aspects forming the meta-term “Trustworthiness”

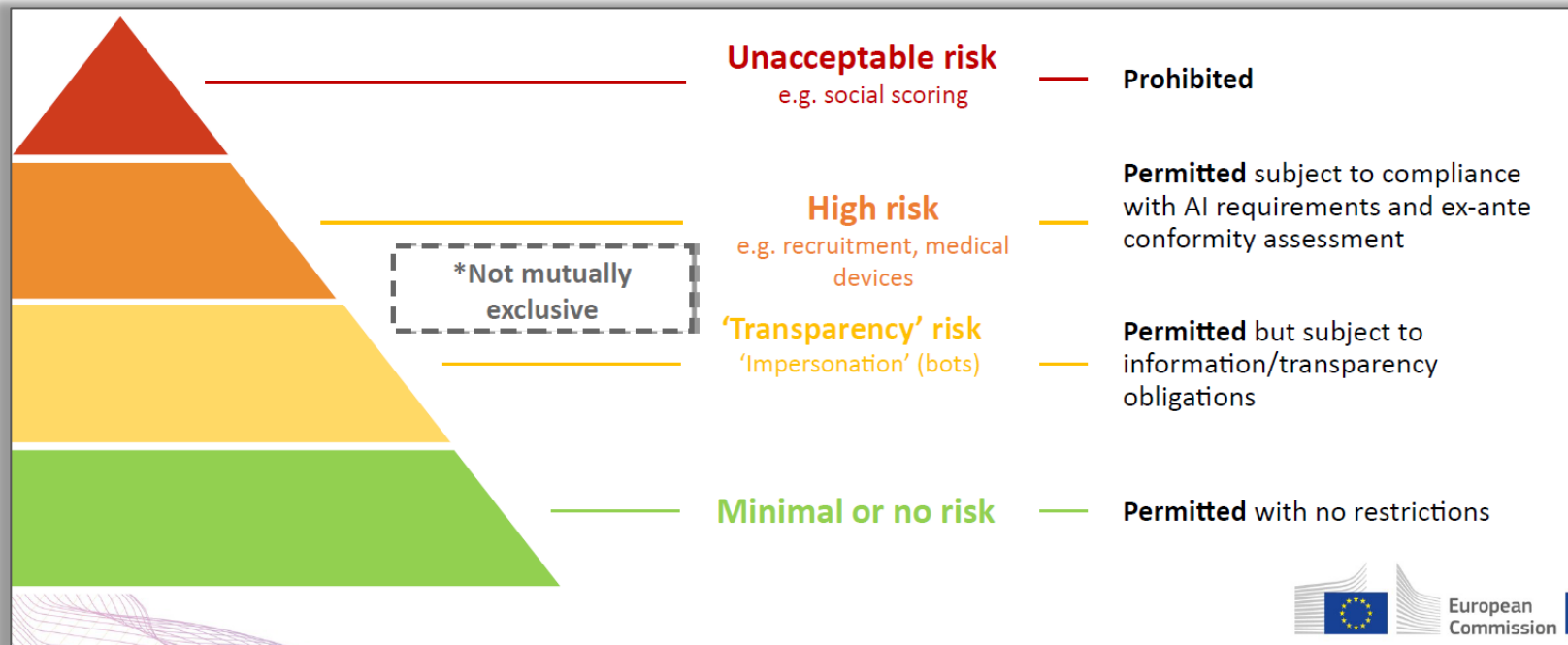
VDE-AR-E 2842-61

Entwicklung und Vertrauenswürdigkeit von
autonom/kognitiven Systemen

Wie können wir vertrauenswürdige KI erreichen?

- **Zuverlässigkeit, Verfügbarkeit, Wartbarkeit**
 - **Robustheit**
 - **Benutzerfreundlichkeit**
 - **Funktionale Sicherheit**
 - **Informationssicherheit**
 - **Datenschutz**
 - **Ethik**
- Anerkannte Entwicklungsprozesse
 - Verifikations- und Validierungsmethoden
 - Interpretationsmethoden
 - Normung und Zertifizierung
 - Security by Design
 - Generalisierung statt Überanpassung
 - Repräsentative Datensätze
 - Einfluss von Domänenwissen

EU-Regulierung: risikobasierter Ansatz

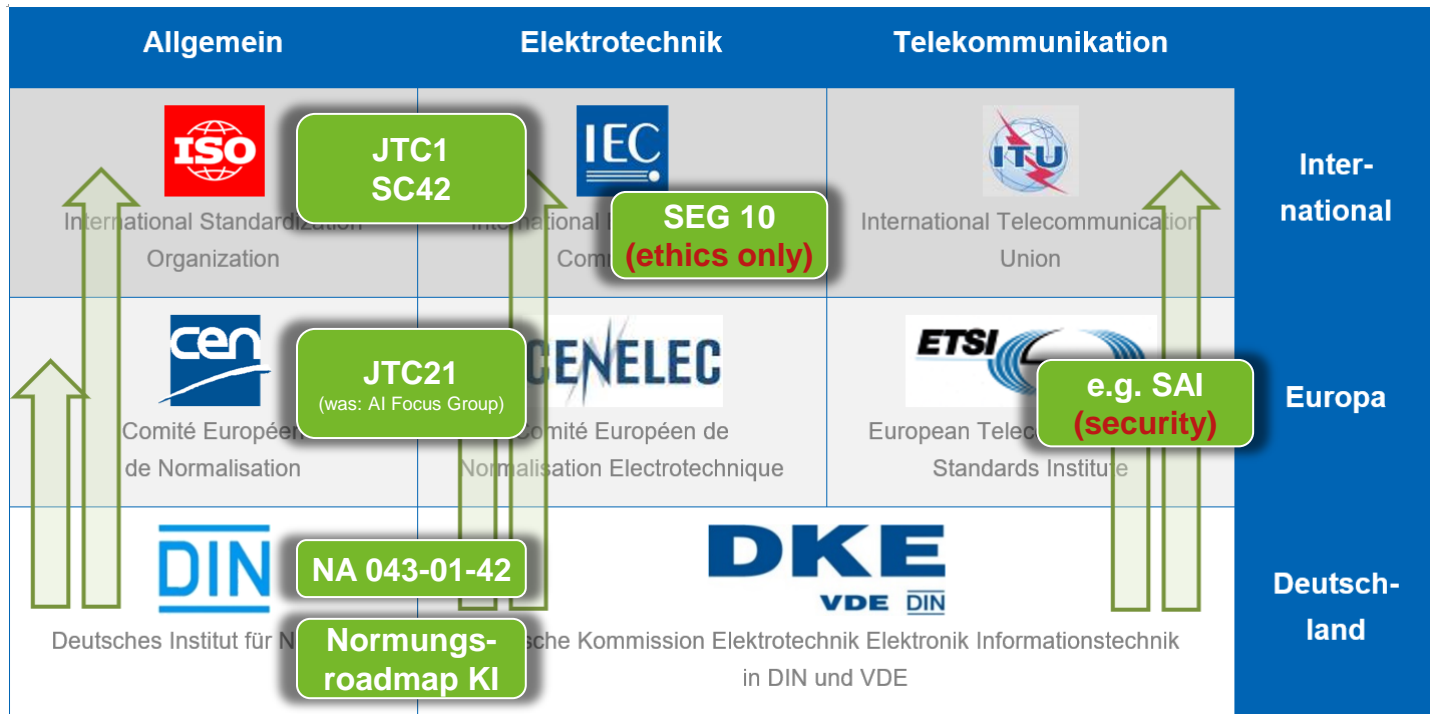


AI standardisation: core structures and interface to regulation



21.04.
Draft Regulation

New
Legislative
Framework



Auftrag

- Erstellung von **Normungsdokumenten für KI** und die Nutzung der damit verbundenen Daten
- Erwägung der **Übernahme internationaler Normen**, die für Europa relevant sind
- Berücksichtigung der **Bedürfnisse des europäischen Marktes und der Gesellschaft**,
- Untermauerung der **EU-Gesetzgebung**, Politik, Grundsätze und Werte

Zahlreiche Schnittstellen, z.B. als Liaison, Beobachterstatus etc.

- Normungsgemeinschaft in Europa (Fokus auf CEN-CENELEC, ETSI)
- Internationale Normungsgemeinschaft (ISO/IEC JTC1/SC42, **IEC SEG10**, IEEE, ...)
- Europäische Kommission
- Anhang-3-Organisationen in Europa
- Initiativen wie



BIG DATA VALUE
ASSOCIATION



GAIA-X

Einige aktuelle Ideen als Ergänzung zum Fahrplan der AI Focus Group

- Erweiterte Ziel-Spezifikation
- AI-Konformitätsbewertung
- Grüne und nachhaltige KI
- Erkennung von Sprache
- Data Governance & Qualität für AI

Kick-off 2021-06-01
~ 100 Teilnehmer*

Fazit Juli 2021:

Es hat sich gezeigt, dass ein **substanzieller internationaler Konsens über KI-Ethik** möglich ist, wenn man die Herausforderung richtig einteilt.
Ein breiterer Ansatz "**Ethik in Normen**" wird empfohlen.

Europäische und internationale Normung



- CEN-CENELEC Focus Group for Artificial Intelligence

Roadmap Report
October 2020

- IEC SEG 10 Ethics in autonomous and artificial intelligence applications



Finaler Report Juli 2021

AI Ethics Impact Group
www.ai-ethics-impact.org

Bertelsmann Stiftung

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



INTERNATIONAL CENTER FOR
ETHICS IN THE SCIENCES AND
HUMANITIES (IZEW)

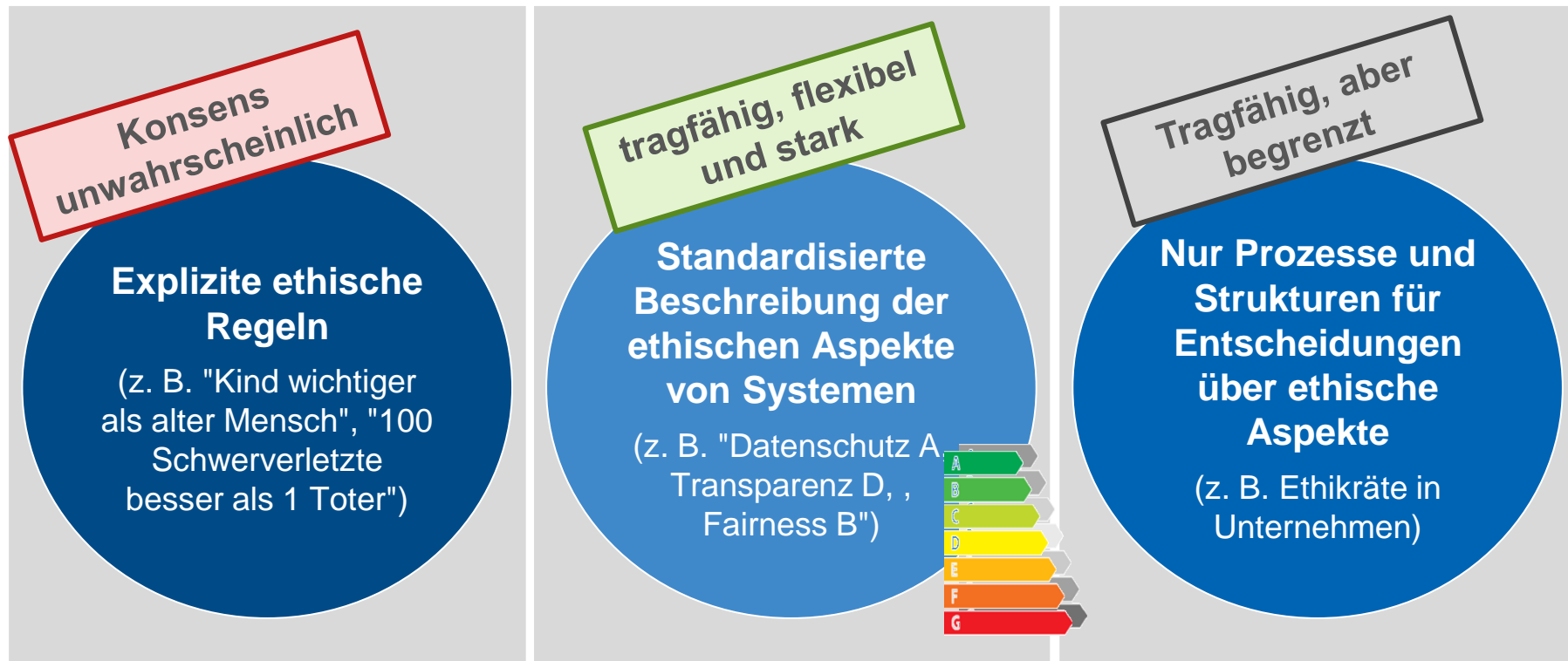


Zielsetzung:

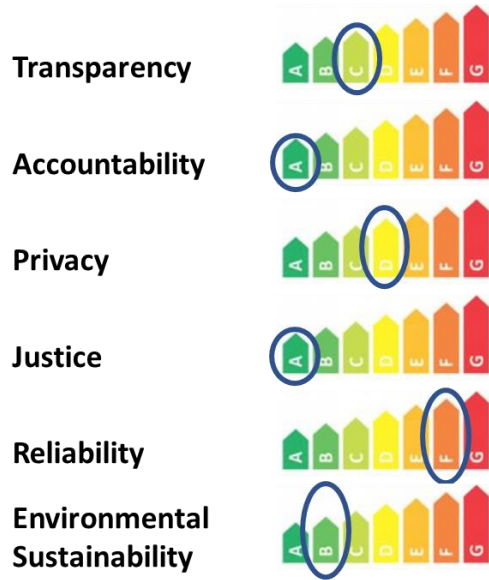
Operationalisierung der KI-Ethik ...

- ... mit einem Ansatz, der für die Industrie, die Regulierungsbehörden und die Verbraucher/Bürger praktikabel ist
- ... und Ethik messbar und durchsetzbar macht

Wie kann man KI-Ethik standardisieren?



Zusammenführung von Forschung, Zivilgesellschaft, Normung und Regulierung: Erste Überlegungen



1. Welche Kategorien sollten einbezogen werden?
2. Wie kann Transparenz, Rechenschaftspflicht usw. messbar gemacht werden?
3. Welches Niveau ist für eine bestimmte Anwendung akzeptabel?

Zusammenführung von Forschung, Zivilgesellschaft, Normung und Regulierung: Erste Überlegungen

Transparency



Accountability



Privacy



Justice



Reliability



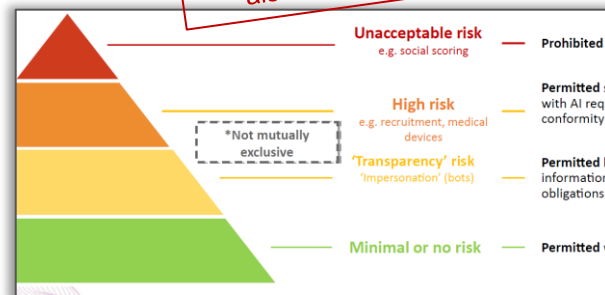
Environmental Sustainability



1. Welche Kategorien sollten einbezogen werden?
2. Wie kann Transparenz, Rechenschaftspflicht usw. messbar gemacht werden?
3. Welches Niveau ist für eine bestimmte Anwendung akzeptabel?



"codes of conduct" als Brücke?



Thank you!

Andreas Hauschke

Projectmanager for trustworthy AI

Tel: +49 69 6308-423

E-Mail: Andreas.Hauschke@vde.com



VDE