



ILIdenT-Projektvorstellung





Projektvorstellung

- ILIdent = Intelligente Luftfahrttaugliche Ident-Technologien für die Supply Chain
- Durch die IFB Hamburg gefördertes Verbundprojekt zur Entwicklung verschiedener Identifikations- und Dokumentationslösungen für die Luftfahrtbranche
- Projektpartner:
- Use Cases, u.a. Erkennung verschiedener Bauteile (z.B. Fahrwerksbuchsen, Triebwerksrohre) & Defekte
- <https://ilident.hamburg/>



Einordnung der Thematik

Responsible/Trustworthy AI

Forschung & Anwendung

- Ansätze, um z.B. Erklärbarkeit, Robustheit von KI zu erhöhen
- Vor- & Nachteile der Ansätze

Rechtliche Anforderungen

- insb. AIA

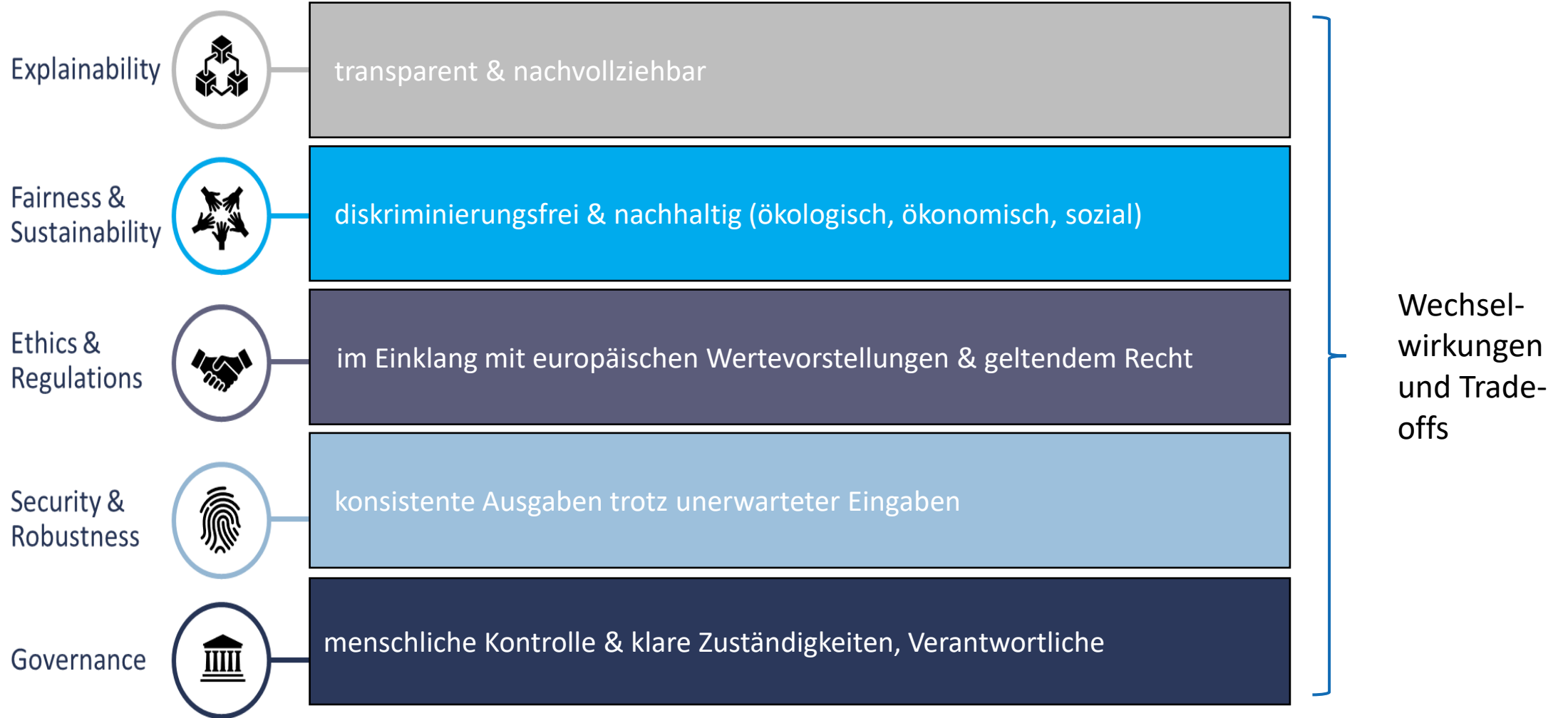
Standards

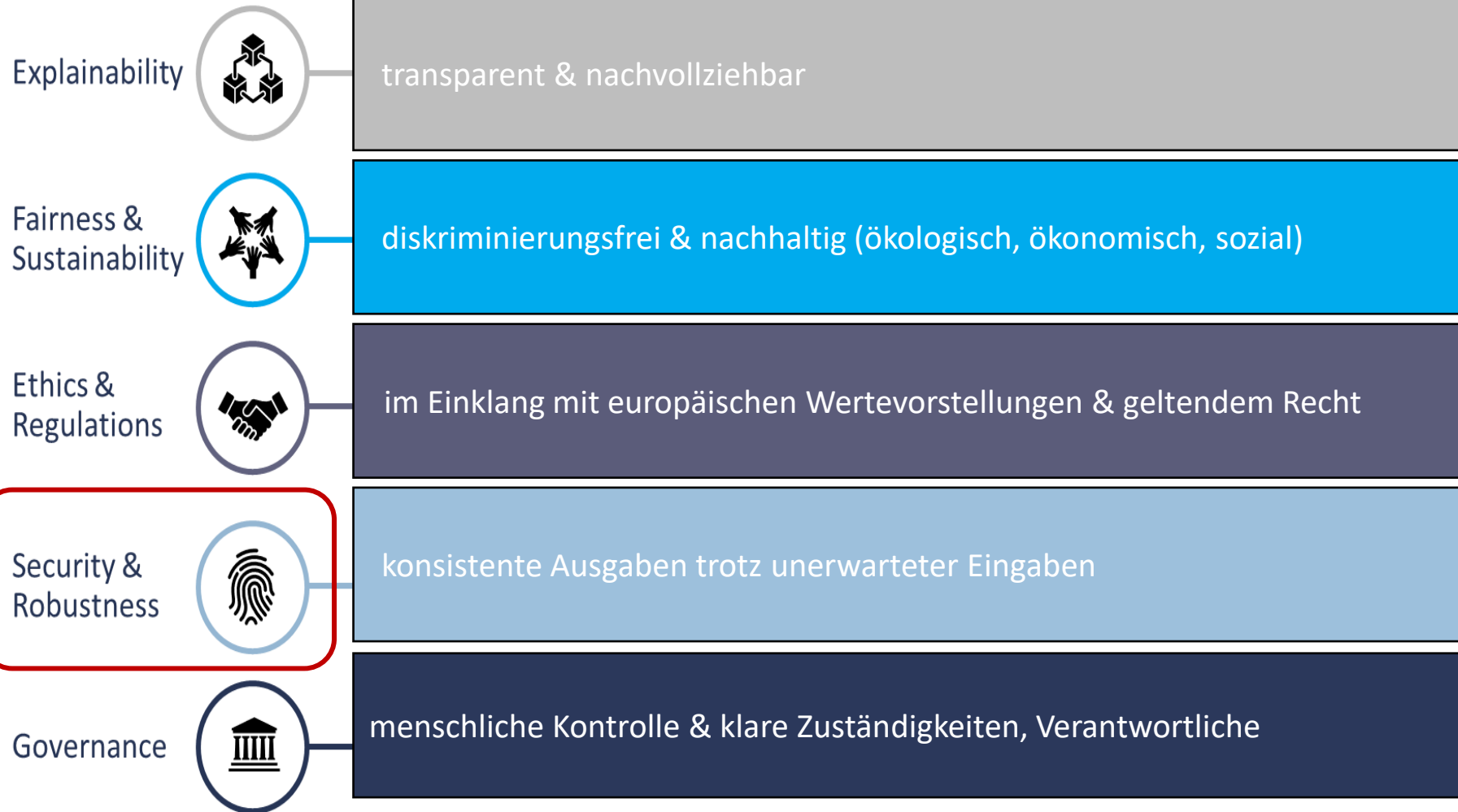
- Europäische Normen (CEN, Cenelec)
- EASA Fahrpläne & Konzeptpapiere

Was & wie?
normativ

Was?
verbindlich

Wie?
freiwillig





Robustheit & Sicherheit

RAI

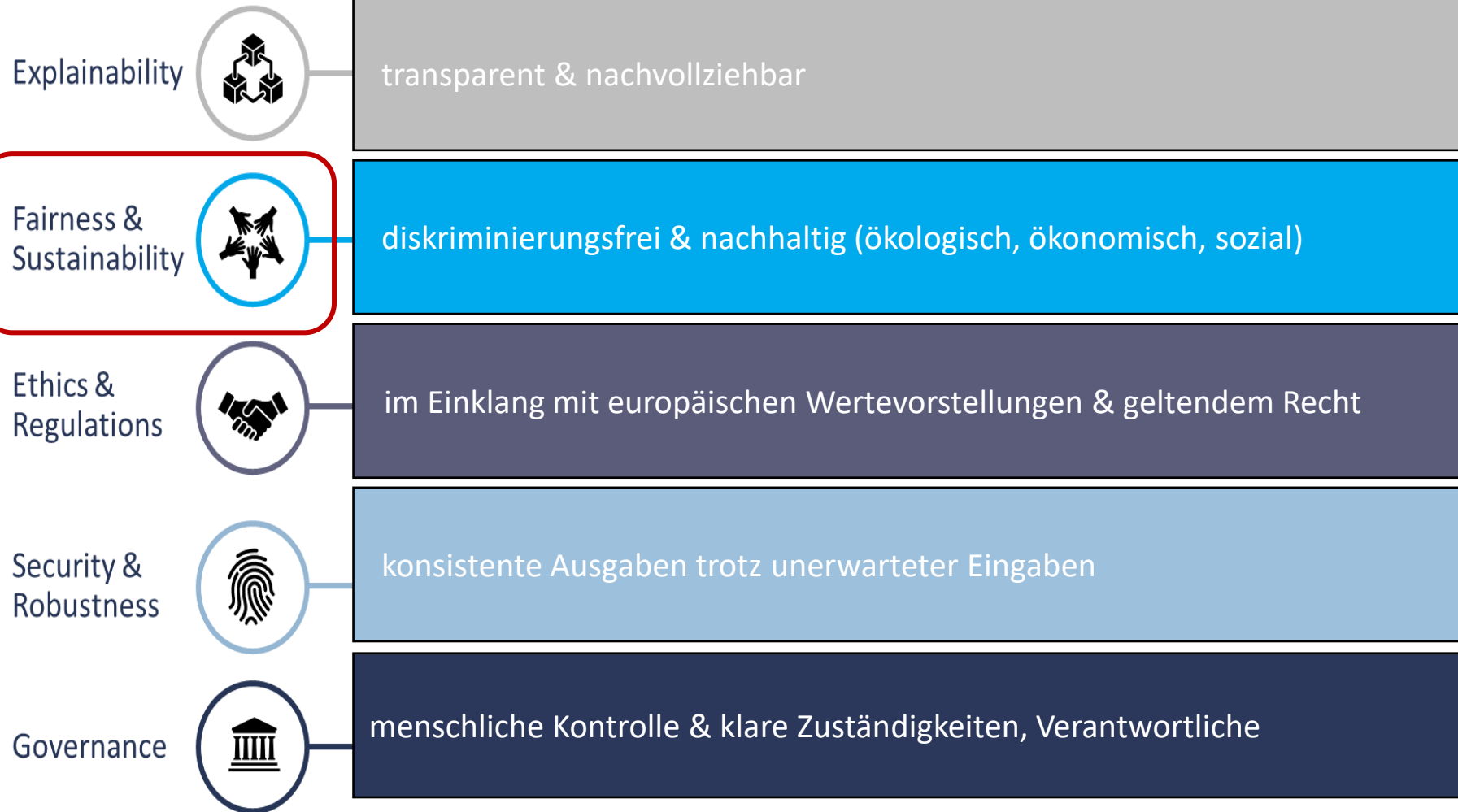
- Bandbreite/Variationen in Eingabedaten
- Generalisierbarkeit
- Drift-Phänomene im Betrieb
- KI-spezifische Sicherheitsaspekte

AIA

- Artikel 9: Risikomanagementsystem
- Artikel 10: Daten und Daten-Governance
- Artikel 15: Genauigkeit, Robustheit und Cybersicherheit

Bewertungsmaßstäbe

- BM-1 bis BM-20
- u.a. zu Training & Testmethoden, (statistische) Validierungsmaße, systematische Erfassung von Risiken & Ableitung von Maßnahmen, Restrisiko



Fairness & Nachhaltigkeit

RAI

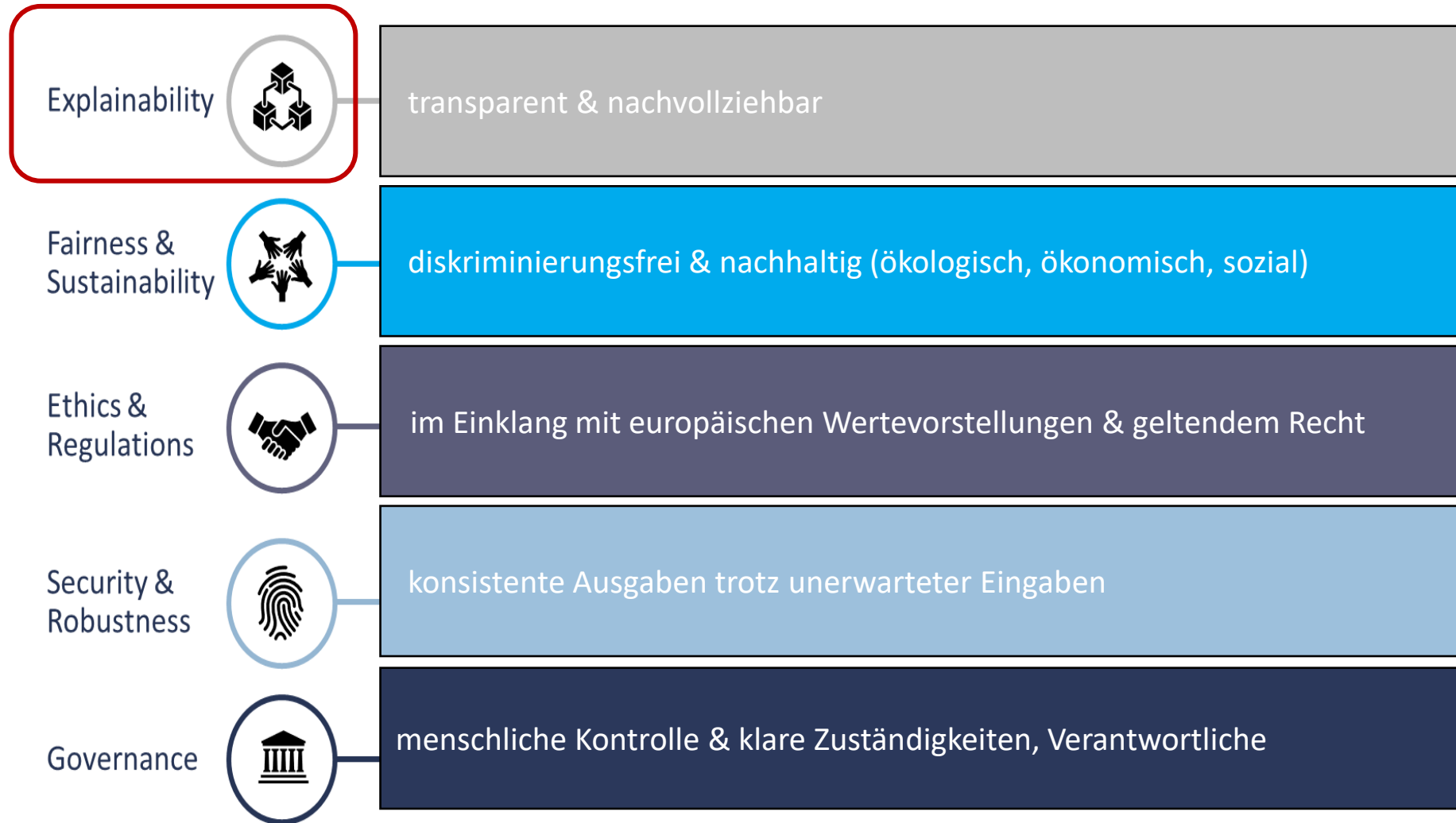
- Definition & Ausdruck von Fairness in Form messbarer Größen
- Bias & Quellen (daten- & interpretationsgetrieben)
- Fairness in Anwendungsfällen ohne personenbezogene Daten

AIA

- Artikel 10: Daten und Daten-Governance
- Artikel 69: Verhaltenskodizes (freiwillig)

Bewertungsmaßstäbe

- BM-21 bis BM-30
- u.a. systematische Erfassung von Bias(quellen), Identifizierung von Untergruppen & Untersuchung des Modellverhaltens für Untergruppen, Sampling, fehlende/fehlerhafte Daten



Erklärbarkeit

RAI

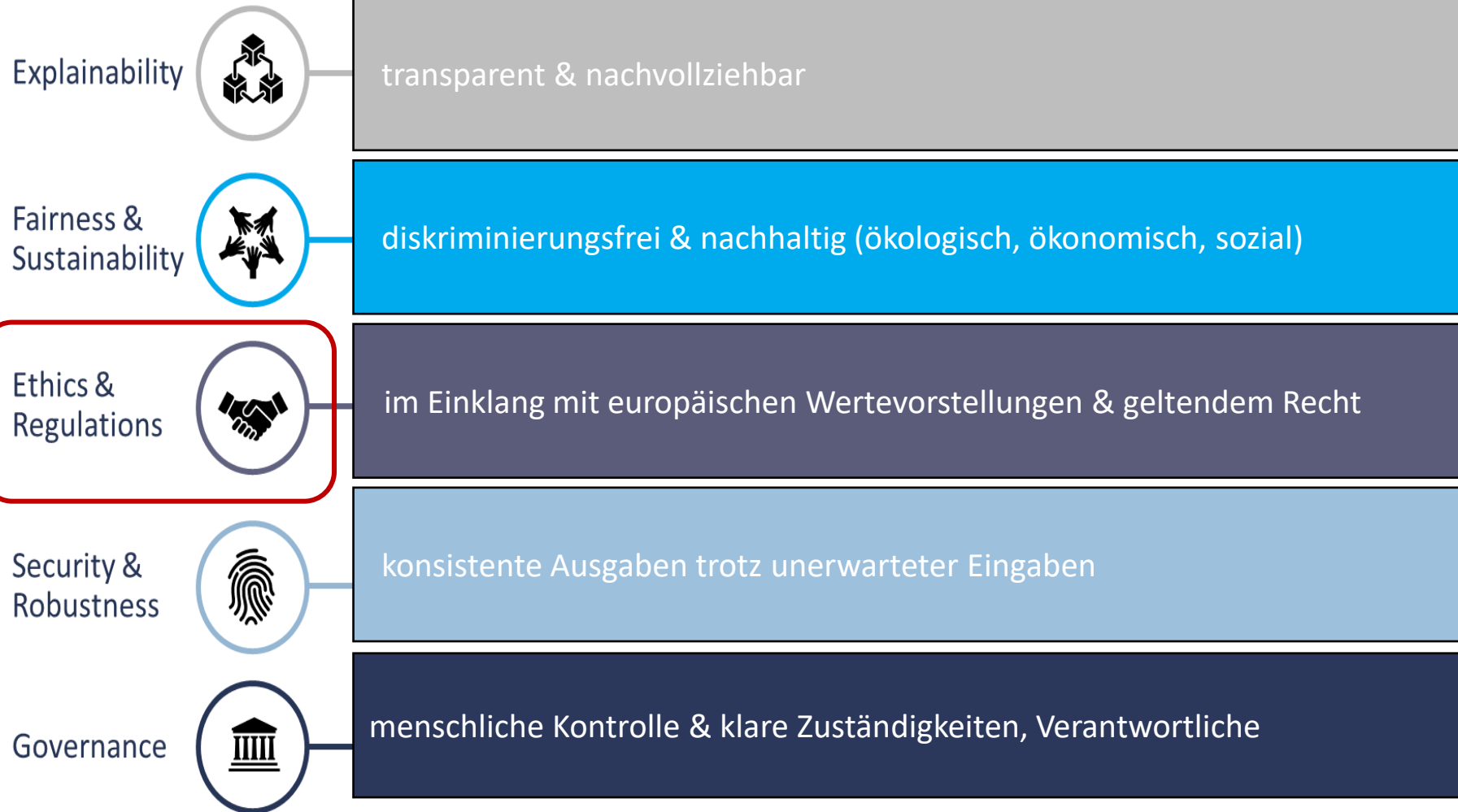
- modell-inhärente vs. post-hoc Erklärbarkeit
- Vor- & Nachteile verschiedener XAI-Ansätze

AIA

- Artikel 10: Daten und Daten-Governance
- Artikel 13: Transparenz und Bereitstellung von Informationen für die Nutzer
- Artikel 14: Menschliche Aufsicht

Bewertungsmaßstäbe

- BM-28 bis BM-42
- u.a. Datenvorverarbeitung, Feature Engineering & Selection, Auswahl eines geeigneten Modelltyps (Annahmen & Limitationen), XAI-Maßnahmen & Einschränkungen, Wiederholbarkeit, Form/Häufigkeit der Erklärung



Recht & Ethik

RAI

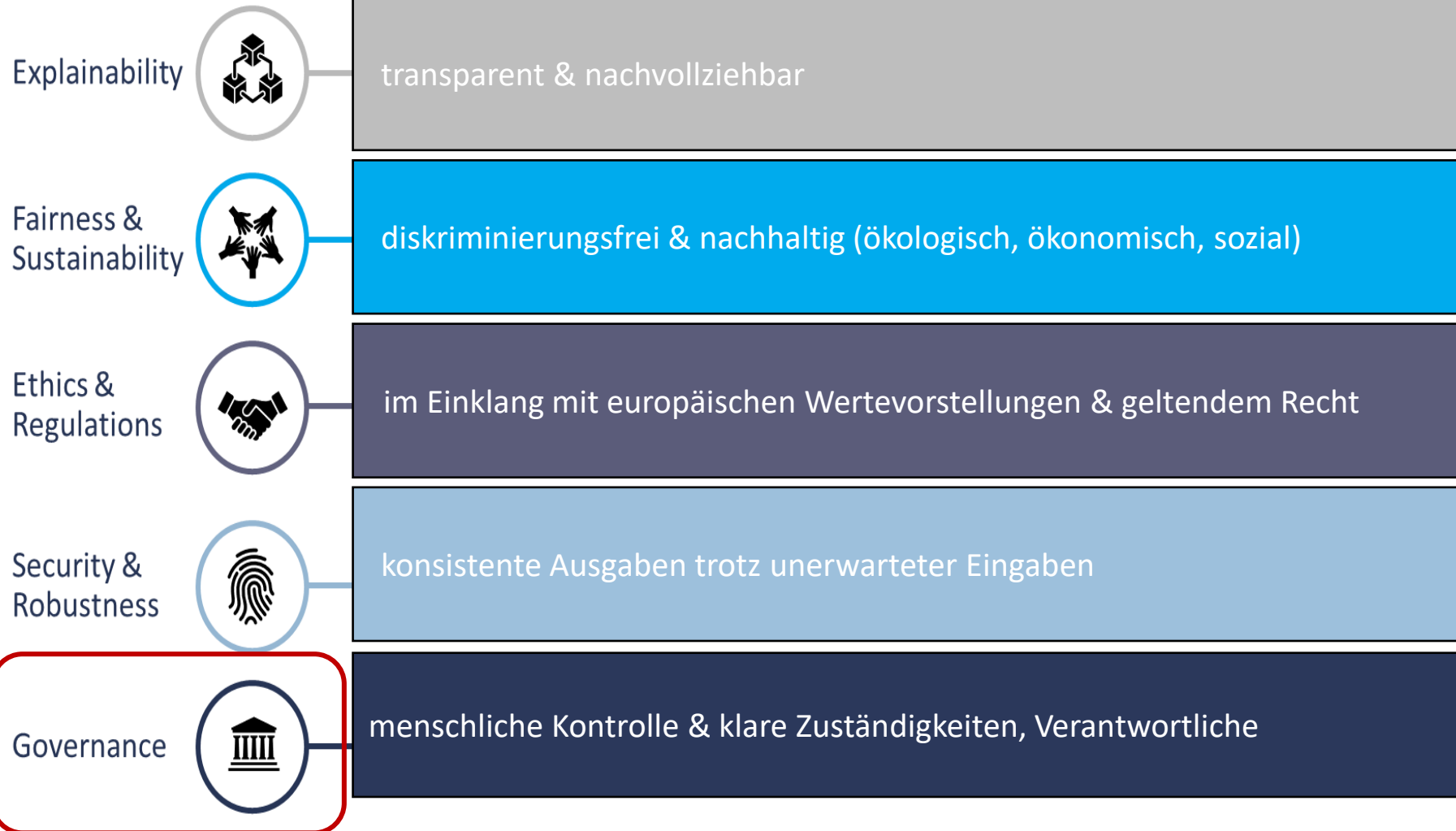
- Umgang mit Daten
- Mensch-Maschine-Interaktion
- Abwägungen zw. verschiedenen Gütern
- Schutz der Grundrechte

AIA

- Artikel 10: Daten und Daten-Governance
- Artikel 14: Menschliche Aufsicht

Bewertungsmaßstäbe

- BM-43 bis BM-47
- u.a. Datenschutz, Recht an Daten, systematische Erfassung der Risiken für Gesundheit, Sicherheit & Grundrechte



Governance

RAI

- Prozesse (Automatisierungsgrad/ human in the loop/ on the loop/ in command; Umgang mit Fehlerfällen)
- Verantwortlichkeiten
- Dokumentation

AIA

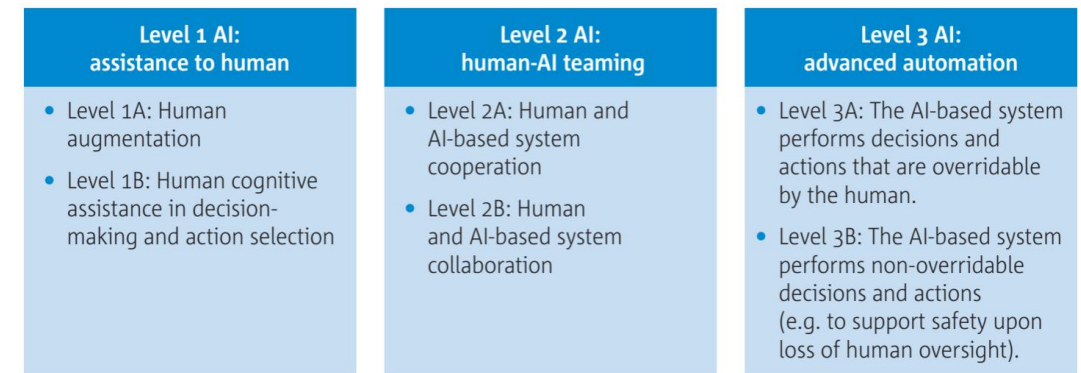
- Art. 9: Risikomanagementsystem
- Art. 10: Daten und Daten-Governance
- Art. 11: Technische Dokumentation
- Art. 12: Aufzeichnungspflichten
- Art. 13: Transparenz und Bereitstellung von Informationen für die Nutzer
- Art. 14: Menschliche Aufsicht
- Art. 17: Qualitätsmanagementsystem

Bewertungsmaßstäbe

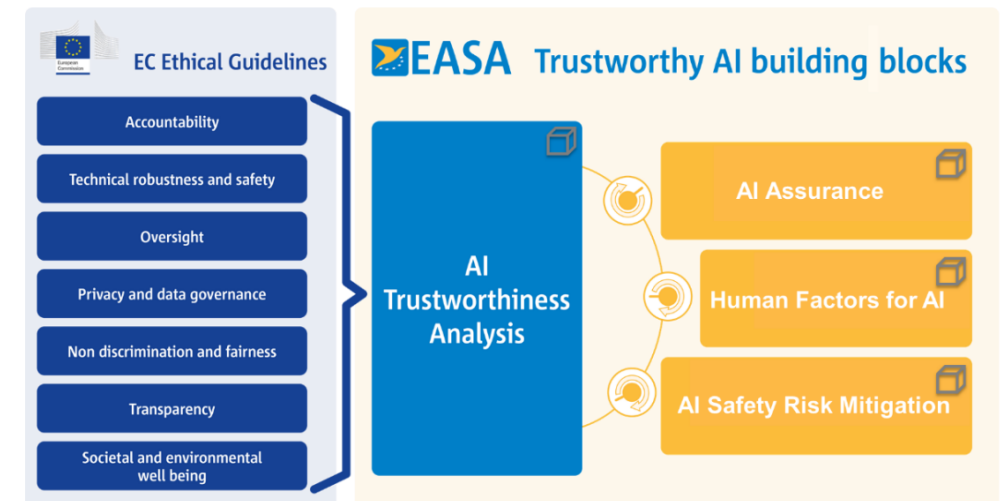
- BM-48 bis BM-61
- u.a. geeignetes Maß an Automatisierung/menschlicher Aufsicht, Information & Schulung, Gewöhnungs-/Ermüdungseffekte, Gestaltung des User-Interfaces, Feedback-Mechanismus, Monitoring-Struktur, Ressourcenplan, Reaktion auf Fehlerfälle

Standards & Compliance

- Die Agentur der Europäischen Union für Flugsicherheit (EASA) hat erste Dokumente veröffentlicht, die bereits angedachte Anleitungen & annehmbare Nachweisverfahren beinhalten
- Unterscheidung von drei KI-Leveln auf Basis des Automatisierungsgrads / Levels an menschlicher Aufsicht
- AI trustworthiness framework identifiziert vier Bausteine, wobei der erste Baustein als Schnittstelle zu den sieben HLEG-Kriterien für ethische KI dient:



EASA AI Roadmap 2.0, Abbildung 4



EASA Concept Paper proposed issue 02, Abbildung 2



Thank You!

Artificial Intelligence Center Hamburg
(ARIC) e.V.
Van-der-Smissen-Straße 9
22767 Hamburg

Office: +49 40 797 244 42
E-Mail: info@aric-hamburg.de

www.aric-hamburg.de

