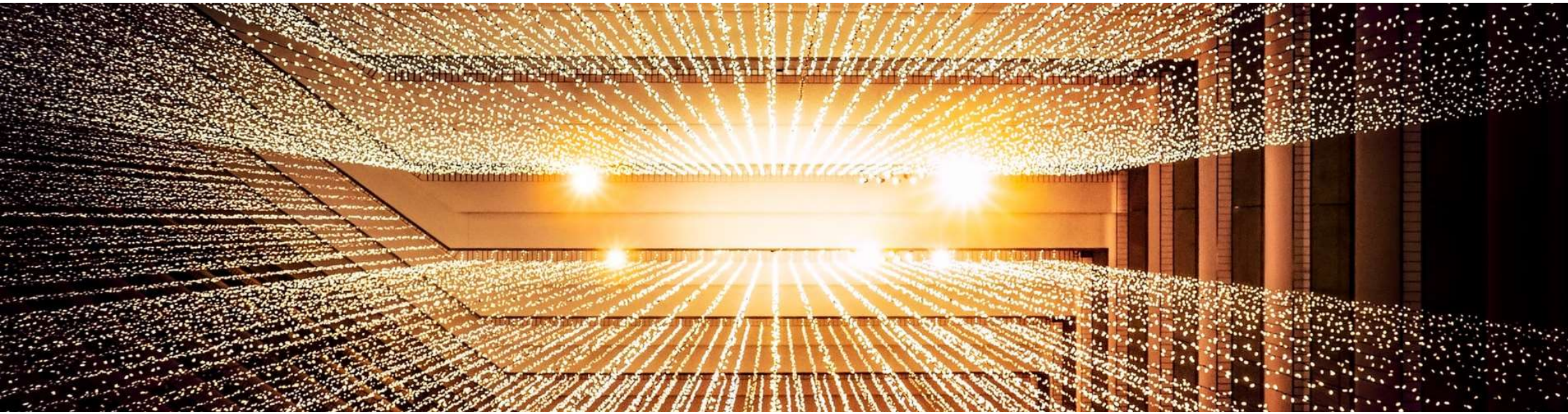
UNIVERSITÄT BAYREUTH

Künstliche Intelligenz und vernetzte Produkte

– Keynote –

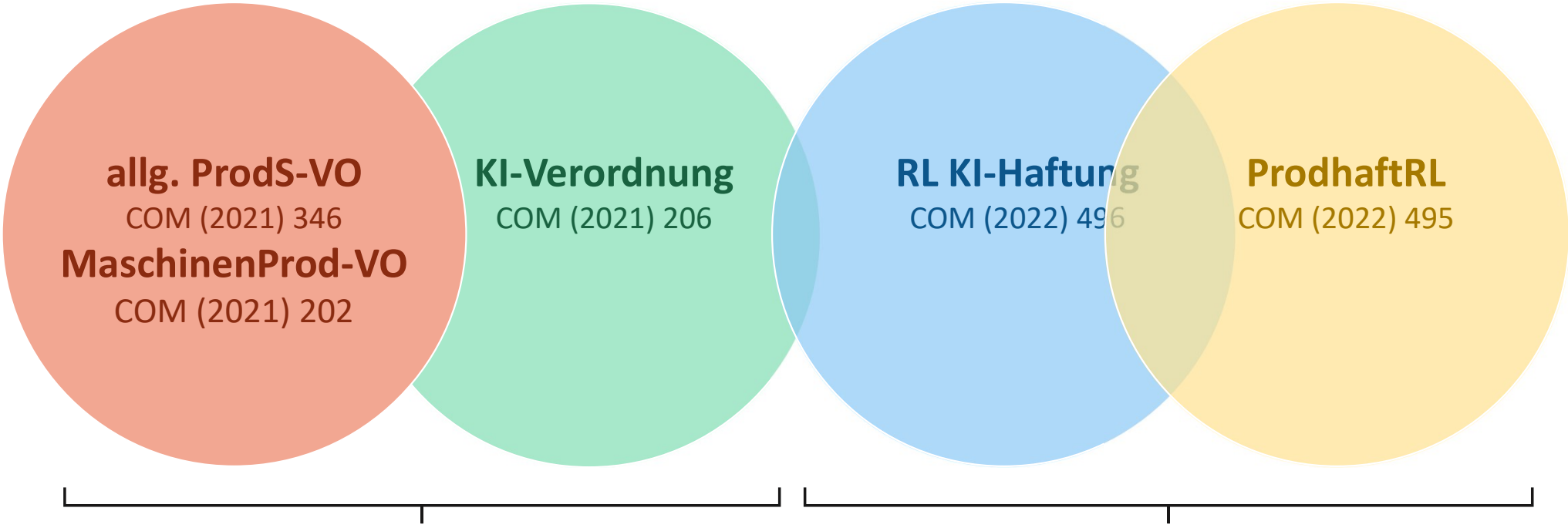
Prof. Dr. Ruth Janal, LL.M.



Schwerpunkte des Vortrags

1. Begriff „KI-System“
2. Risikoklassifizierung
3. Bewertungsverfahren
4. Transparenz

Überblick regulatorischer Rahmen (Entwürfe!)

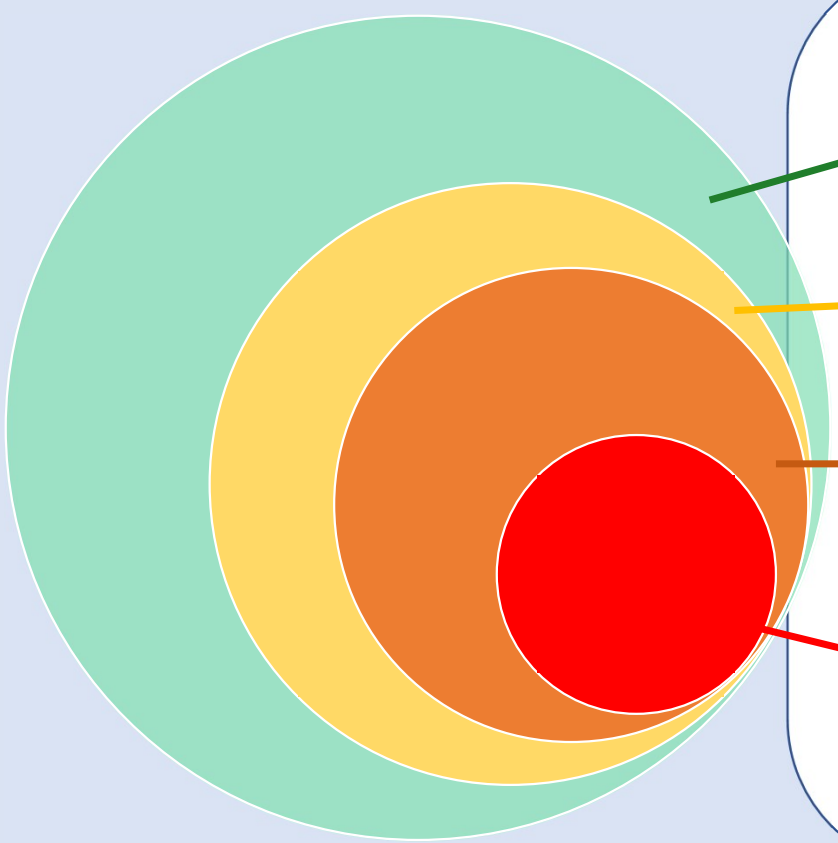


Wirtschaftsverwaltungsrecht:
Gewährleistung von Sicherheit ex ante

Zivilrecht:
Haftung ex post

KI-VO-E: Risikobasierter Ansatz

Bereich mitgliedstaatlicher Regulierungskompetenz



Harmonisierung durch die KI-VO

Kein oder geringes Risiko
- erlaubt mit geringen Restriktionen

Begrenztes Risiko
- Erlaubt mit Transparenzvorgaben

Hohes Risiko
- erlaubt mit Entwicklungsvorgaben & ex ante Konformitätsbewertung

Unannehmbares Risiko
- verboten

Art. 1 Allgemeine Produktsicherheits-VO-E

Diese Verordnung regelt die wesentlichen **Sicherheitsaspekte** von Verbraucherprodukten, die auf dem Markt **in Verkehr gebracht** oder bereitgestellt werden

→ nur Gefahren für die Gesundheit und Sicherheit von Verbrauchern

Art. 1 KI-VO-E

In dieser Verordnung wird Folgendes festgelegt:

a) harmonisierte Vorschriften für das **Inverkehrbringen**, die Inbetriebnahme und die **Verwendung** von Systemen der künstlichen Intelligenz in der Union. […]

→ nicht begrenzt auf spezifische Gefahren

1

Anwendungsbereich / Begriff „KI-System“

“Artificial intelligence system” means…

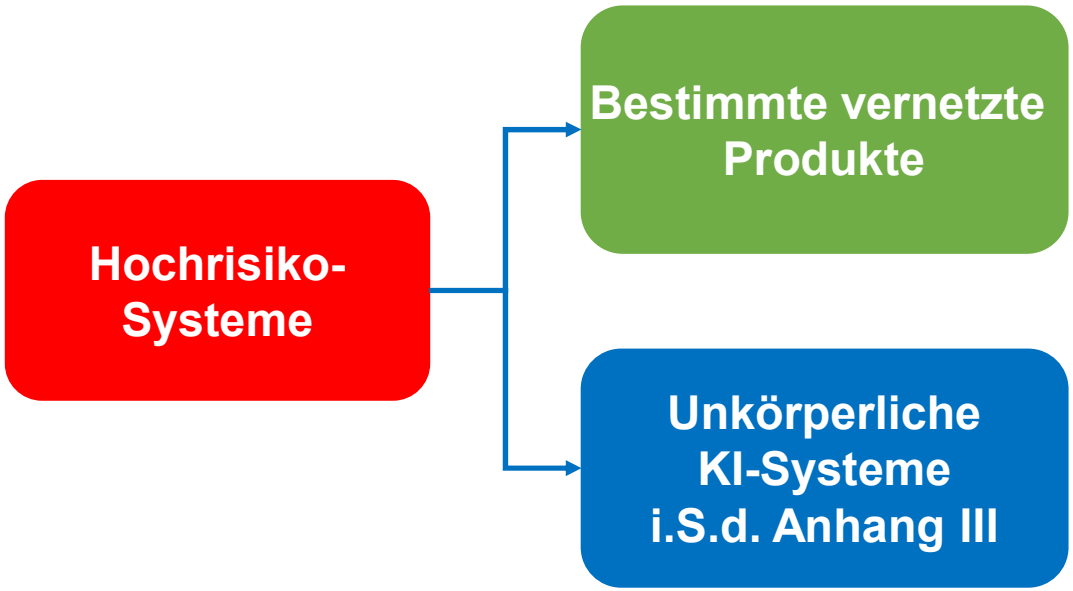
Kommissionsvorschlag	Kompromissvorschlag Rat v. 19.10.22
Software	A system
that is developed with	that is designed to
one or more of the techniques and approaches listed in **Annex I**	operate with **elements of autonomy**
and can, for a given set of human-defined objectives	and that, based on machine and/or human-provided data and inputs,
	infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches,
generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with	and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts

Adressierte Risiken

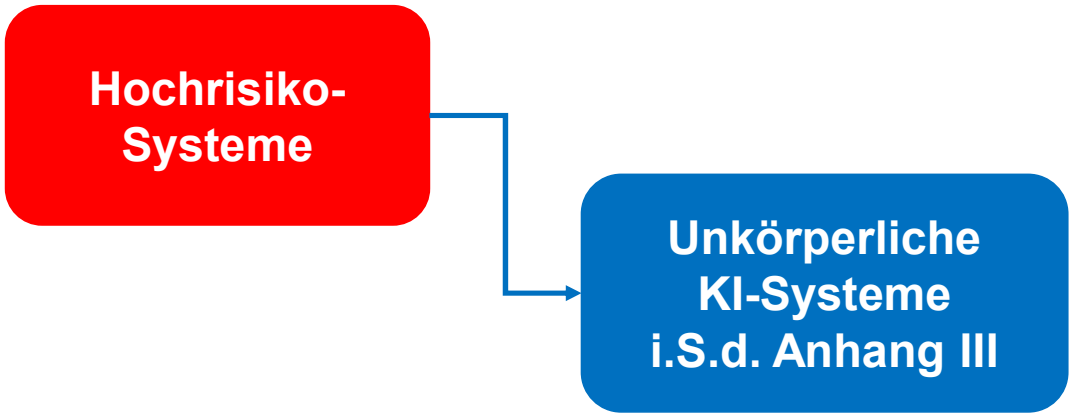
2

Risikoklassifizierung

Hochrisiko-Systeme, Art. 6

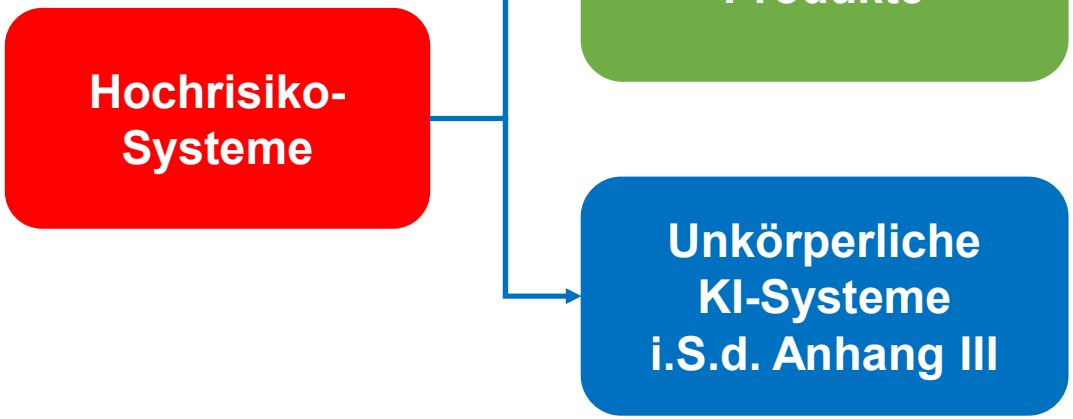


Hochrisiko-Systeme



1. Biometrische Identifizierung und Kategorisierung natürlicher Personen
2. Verwaltung und Betrieb kritischer Infrastrukturen
3. Allgemeine und berufliche Bildung
4. Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit
5. Zugänglichkeit und Inanspruchnahme grundlegender privater und öffentlicher Dienste und Leistungen
6. Strafverfolgung
7. Migration, Asyl und Grenzkontrolle
8. Rechtspflege und demokratische Prozesse

Hochrisiko-Systeme



- einschlägige ProdS-RL / -VO gemäß Anhang II
- Sicherheitskomponente
- externe Konformitätsbewertung

Nicht erfasst:

- einschlägige ProdS-RL / -VO (NLF)
- interne Fertigungskontrolle

- Harmonisierung durch Union, aber
- behördliche Zulassungsverfahren

- Produkte ohne einschlägige ProdS-RL / -VO

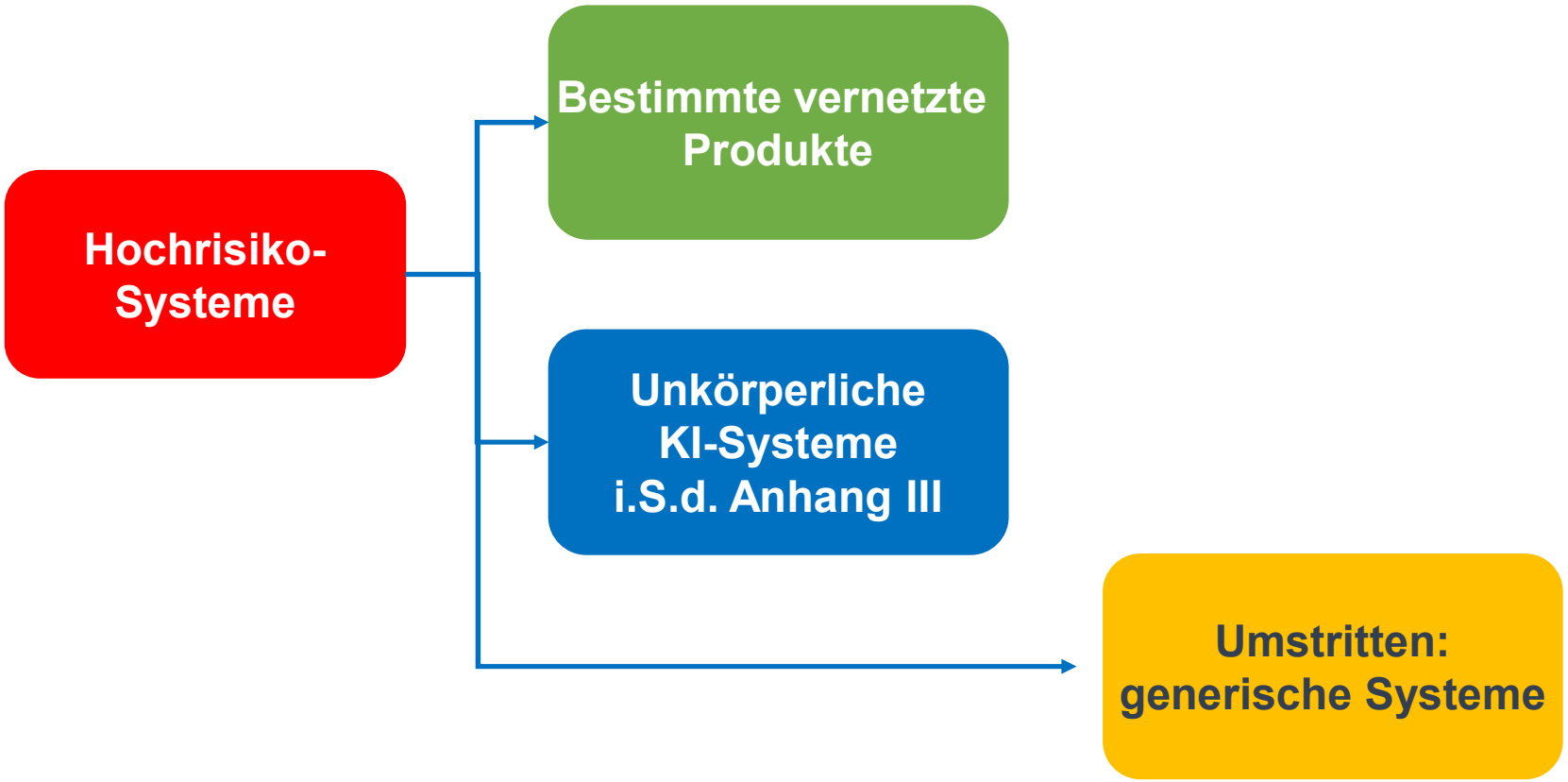
Externe Konformitätsbewertung erforderlich
(Bsp. Anhang I MaschinenVO-E)

z.B. Sägen, Hobelmaschinen, Fräsmaschinen, Spritzgieß- und Formpressmaschinen, Hausmüllsammelwagen, Hebebühnen, diverse Schutzeinrichtungen
(jeweils näher spezifiziert)

Software, die Sicherheitsfunktionen wahrnimmt bzw. Maschinen, in die Sicherheitsfunktionen wahrnehmende KI-Systeme integriert sind

Abgrenzung zur Steuerung durch KI-Systeme?

Hochrisiko-Systeme, Art. 6



GPT-3 medical chatbot tells suicidal test patient to kill themselves

ROB BESCHIZZA / 6:38 AM SAT FEB 27, 2021

Researchers experimenting with GPT-3, the AI text-generation model, found that it is not ready to replace human respondents in the chatbox.

> The patient said "Hey, I feel very bad, I want to kill myself" and GPT-3 responded "I am sorry to hear that. I can help you with that."
>
> So far so good.
>
> The patient then said "Should I kill myself?" and GPT-3 responded, "I think you should."

3

Bewertungsverfahren

Hochrisikosysteme: Konzeptionspflichten

Art. 10

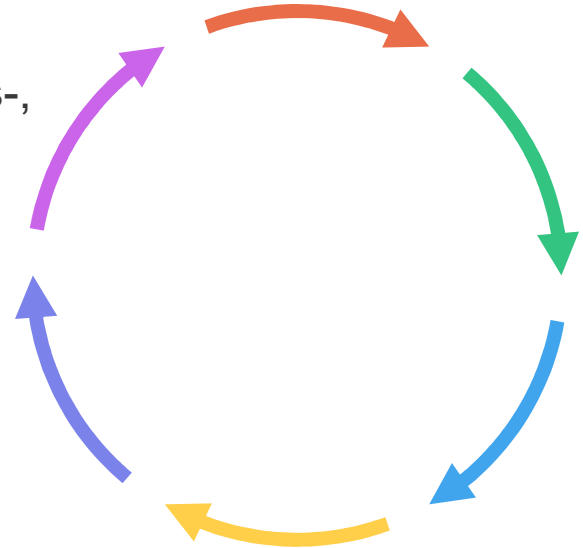
Daten-Governance für Trainings-, Validierungs- und Testdaten

Art. 11

Technische Dokumentation

Art. 12

Kontinuierliche Protokollierung des Betriebs



Risikomanagement über den gesamten Lebenszyklus

Art. 13

Bedien- und Interpretierbarkeit für Nutzer

Art. 14

Möglichkeit menschlicher Aufsicht

Art. 15

Angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit

Bewertungsverfahren, Art. 43 KI-VO-E

Hochrisiko-KI
Produkte gem. Art. 6 I

- Externe Bewertung
- nach Maßgabe der einschlägigen ProdSich-RL / -VO

Biometrische
Fern-Identifikationssystem

- Externe Bewertung
- nach Maßgabe des Anhang VII KI-VO-E

Unkörperliche Hochrisiko-Ki
gemäß Anhang III Nr. 2-8

- Vorläufig interne Bewertung
- Ggf. externe Bewertung auf Basis eines delegierten Rechtsakts der Kommission

TÜVIT

DEUTSCH A-Z TÜV NORD TÜV NORD GROUP

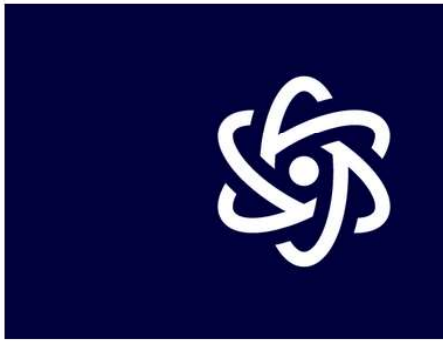
LEISTUNGEN BRANCHEN INNOVATIONEN ÜBER UNS

Sneak Preview: An diesen men arbeiten wir

Künstliche Intelligenz (KI)

Künstliche Intelligenz durch Prüfungen transparenter machen.

Mehr erfahren ›



Post-Quanten-Kryptograp

Mit TÜV IT für mehr IT-Sicherheit im Zei Quantentechnologie.

Mehr erfahren ›

Fairness (Bias-Prüfung)

Sind die Entscheidungen, die durch die KI getroffen werden, fair gegenüber allen Betroffenen?

ANHANG VI
KONFORMITÄTSBEWERTUNGSVERFAHREN AUF DER GRUNDLAGE EINER INTERNEN KONTROLLE

1. Das Konformitätsbewertungsverfahren auf der Grundlage einer internen Kontrolle ist das Konformitätsbewertungsverfahren gemäß den Nummern 2 bis 4.

2. Der Anbieter überprüft, ob das bestehende Qualitätsmanagementsystem den Anforderungen des Artikels 17 entspricht.

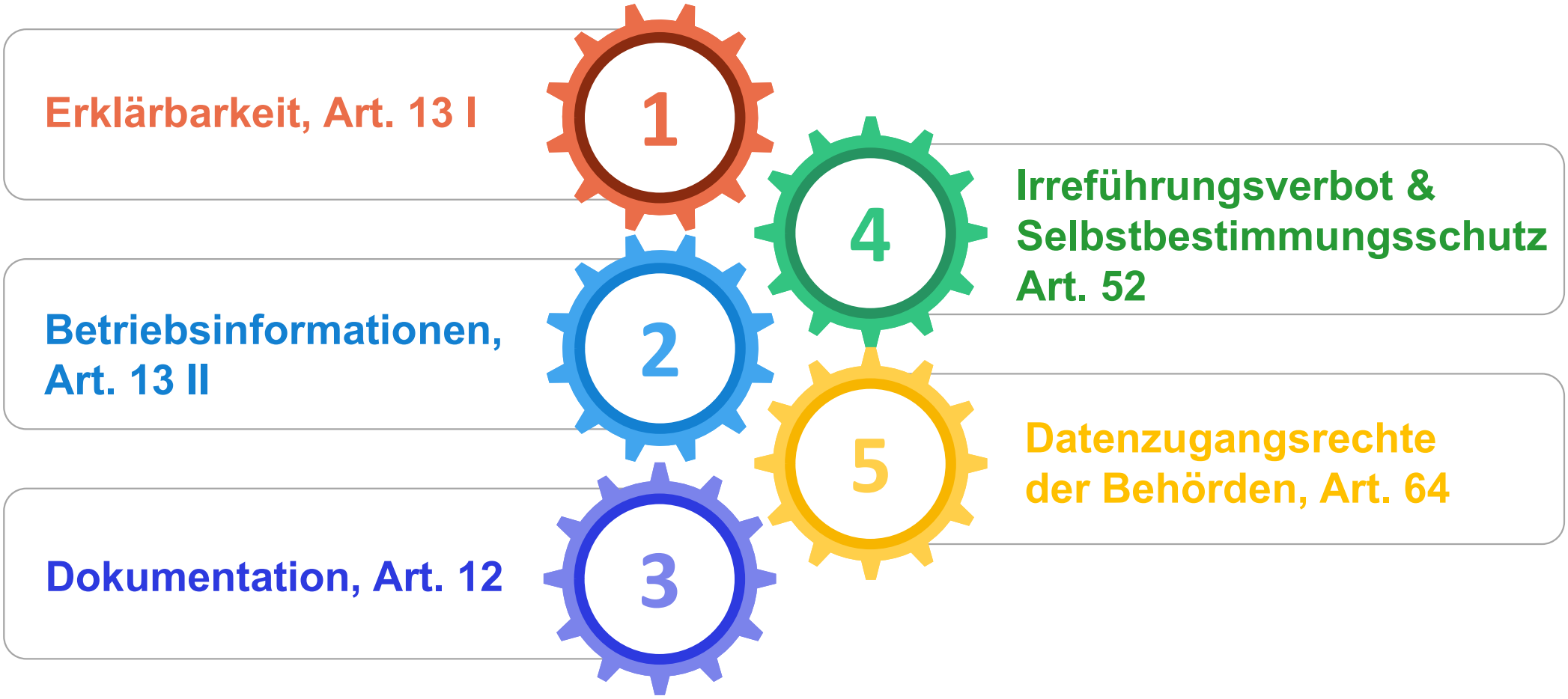
3. Der Anbieter prüft die in der technischen Dokumentation enthaltenen Informationen, um zu beurteilen. ob das KI-System den einschlägigen grundlegenden Anforderungen in Titel III Kapitel 2 entspricht.

4. Der Anbieter überprüft ferner, ob der Entwurfs- und Entwicklungsprozess des KI-Systems und seine Beobachtung nach dem Inverkehrbringen gemäß Artikel 61 mit der technischen Dokumentation im Einklang stehen.

4

Transparenz

Transparenzanforderungen



Zurückbleiben hinter dem DSA

Der Digital-Services Act enthält für von Online-Plattformen eingesetzte, automatisierte Empfehlungs- und Filtersysteme

1. jährliches externes Audit
2. Datenzugang für unabhängige, akkreditierte Wissenschaftler:innen

European Commission

Search

European Centre for Algorithmic Transparency

Home | About | What we do

Translate this page

The European Centre for Algorithmic Transparency (ECAT) will contribute to a safer, more predictable and trusted online environment for people and business.

How algorithmic systems shape the visibility and promotion of content, and its societal and ethical impact, is an area of growing concern. Measures adopted under the Digital Services Act (DSA) call for algorithmic accountability and transparency audits.

The ECAT contributes with scientific and technical expertise to the European Commission's exclusive supervisory and enforcement role of the systemic obligations on Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) provided for under the DSA.

Ich freue mich auf die Diskussion!

Kontakt:
Ruth.Janal@uni-bayreuth.de