

D O R D A

Menschliche Aufsicht bei Hochrisiko-KI-Systemen

7.5.2026 – 20. IT-Rechtstag

Alexandra Ciarnau



Partnerin im IT/IP und Datenschutzteam und Co-Leiterin der Digital Industries Group

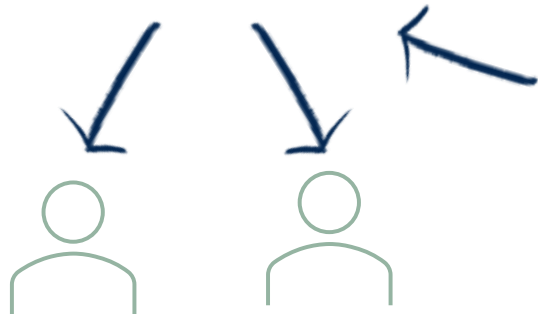
- Schwerpunkte: IT- und Datenschutzrecht, New Technology, E-Commerce
- Highly recommended für Artificial Intelligence bei Lexology 2025, Leading Associate für TMT in Legal 500 2025, Green Ambassador im Legal 500 Green Guide 2025, Trademark Star in IP Stars 2024, Empfohlen für Data:Information Technology bei Client Choice 2025
- Co-Herausgeberin der ailex (Fachzeitschrift für KI-Recht, Manz)
- Co-Autorin der Werke KI-VO Kurzkomentar (Manz), KI Praxishandbuch (Hanser), IP in der Praxis (Manz), UWG Praxishandbuchs (Linde), Handbuch Nachhaltigkeitsrecht (Manz), Handbuch Kreislaufwirtschaft (Linde), #Blockchain (LexisNexis), Praxishandbuch Nachhaltige Finanzierung (Linde)
- Vortragende an zahlreichen (Fach-)Hochschulen sowie bei diversen Seminaranbietern
- Präsidentin von Women in AI Austria

Alexandra Ciarnau

alexandra.ciarnau@dorda.at

Menschliche Aufsicht – Problemaufriss

Art 14, Art 26 Abs 2
KI-VO



Anbieter – Betreiber
Hochrisiko-KI-Systeme

Beispiele:



Bonität/Kreditw.



HR



Kritische Infrastruktur

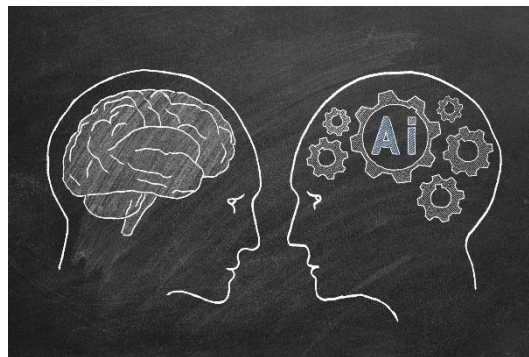


Funkanlagen



Medizinprodukte

Praxisrelevanz



Potentiell positive
Wechselwirkungen
zur **DSGVO**, DSA,
Plattformarbeits-RL

Sorgfältiger Einsatz
&
Haftungs-
minimierung

(unabhängig von
Risikoklasse)

Was steckt hinter der Verpflichtung?

Menschliche Aufsicht

▶ Kernanforderungen lange vor der KI-VO



- ▶ Menschenzentriertes Design
- ▶ Berücksichtigung menschlichen Kompetenzerhalts (zB Kontrollfunktionen, wie Human in the loop)

- ▶ Vorrang menschlichen Handelns
- ▶ Drei relevante Modelle
 - ▶ Human in the loop
 - ▶ Human on the loop
 - ▶ Human in command

Menschliche Aufsicht

▶▶ Was steckt dahinter?

Aufsicht = Verantwortung im Schadensfall

Human in the loop

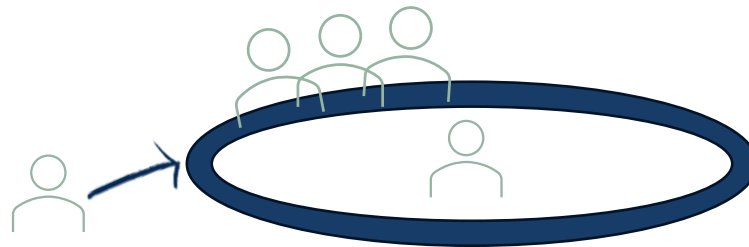
- ▶▶ Menschliche Intervention in jedem Stadium des KI-Systems

Human on the loop

- ▶▶ Menschliche Intervention während der Design- und Betriebsphase

Human in command

- ▶▶ Entscheidungen zum KI-Einsatz:
 - ▶▶ ob, wann und wie?
 - ▶▶ Interventionsmöglichkeiten
 - ▶▶ Umfang der Intervention



Menschliche Aufsicht

» Was steckt dahinter?

Aufsicht aufschlüsseln  kontrafaktisch
korrektiv

	Art des kontrafaktischen Einflusses	Art des Prozesses	Zeitpunkt des Eingreifens
Aufsicht 1. Grades	Direkt kontrafaktisch – Ergebnis hängt von menschlicher Entscheidung ab	Teilweise automatisiert	Jederzeit, auch korrektiv
Aufsicht 2. Grades	Indirekt kontrafaktisch – Ergebnis hängt von Schwellwerten ab	Vollständig automatisiert	Vor der Einführung des KI-Systems oder korrektiv nach der Entscheidung

Hafez in Windholz et al, Praxishandbuch KI-VO 2024, S 460; Laux, Institutionalised Distrust and Human Oversight of Artificial Intelligence 2023

Anbieterpflichten im Fokus

Menschliche Aufsicht – Art 14 KI-VO

► Anbieterpflicht bei Konzeption und Entwicklung

Abs 1

- Vage Grundverpflichtung
- Ermöglichung wirksamer Beaufsichtigung während der Verwendung
- Verpflichtung zur Umsetzung „geeigneter Instrumente einer **Mensch-Maschine-Schnittstelle**“

Abs 2

- Ziele** der menschlichen Aufsicht
- Verhinderung/Minimierung der Risiken für Gesundheit, Sicherheit oder Grundrechte bei **bestimmungsgemäßer Verwendung** oder **vorhersehbare Fehlanwendungen**

Abs 3

- Anforderungen an Aufsichtsmaßnahmen
- Relevante Umstände: Angemessenheit gemäß **Risiken**, Grad der **Autonomie** und **Kontext** der Nutzung
- Verpflichtende Maßnahmen (keine absolute Wahlfreiheit): **Eingebaute oder vom Betreiber umzusetzende Vorkehrungen**

Abs 4

- Anforderungen an die Verwendung durch Betreiber**
- Fähigkeiten und Grenzen, Interpretation, kontrafaktische und korrektive Möglichkeiten, „Stopp-Taste“ oder ähnliche Funktion zur Unterbrechung

Abs 5

- Besondere Anforderungen an biometrische Fernidentifizierung**
- Ausnahmen für Strafverfolgung, Migration, Grenzkontrolle und Asyl bei Unverhältnismäßigkeit

Menschliche Aufsicht – Art 14 KI-VO

► Konkrete Designanforderungen und -möglichkeiten

Mensch-Maschine-Schnittstelle (Abs 1)



Steuerung und Überwachung

Technisch machbare Vorkehrungen ab Werk (Abs 3; **oversight by design**)

Programmierungen,
Funktionalitäten, wie zB HMI

Vorkehrungen zur Umsetzung durch Betreiber (Abs 3)

Katalog an Mindestanforderungen (Abs 4)

Verständnis der Fähigkeiten und Grenzen, wie zB

Betriebsanleitung

Weitere Produkthinweise

Testmöglichkeit



Wirtschaftliche
Zumutbarkeit?

Haftung für
Fehlverhalten des
Betreibers?

Menschliche Aufsicht – Art 14 KI-VO

» Konkrete Designanforderungen und -möglichkeiten

- Ermöglichung der ordnungsgemäßen Überwachung, wie zB
 - Reports
 - Zwischenergebnisse
 - Ampelsysteme zur Verlässlichkeit von Outputs oder Sicherheit der Systeme
- Erkennung und Behebung von Anomalien, Fehlfunktionen und unerwarteter Leistungen, wie zB
 - Human in the loop oder in command-Modelle
- Bewusstsein von Automatisierungsbias, wie zB
 - Eingebaute Fragen und Hinweise
 - Peer-Reviews
 - Cooling-Off-Phasen im Betrieb
 - Eingebaute Einschränkungen über die sich ein KI-System nicht hinwegsetzen kann
 - Notwendige Begründungen und Eingaben

Menschliche Aufsicht – Art 14 KI-VO

» Konkrete Designanforderungen und -möglichkeiten

- Möglichkeiten zur richtigen Interpretation, wie zB
 - Transparenz über Datennutzung und Gewichtung
 - Entscheidungsbaum
- Möglichkeiten zur Ablehnung der Verwendung, Außerachtlassung oder Korrektur von Ausgaben, wie zB
 - Möglichkeiten zur Deaktivierung
 - Verhinderung der Verwendung bestimmter Informationen
 - Entscheidungen revidieren
- Unterbrechung, wie zB
 - „Stopp-Taste“

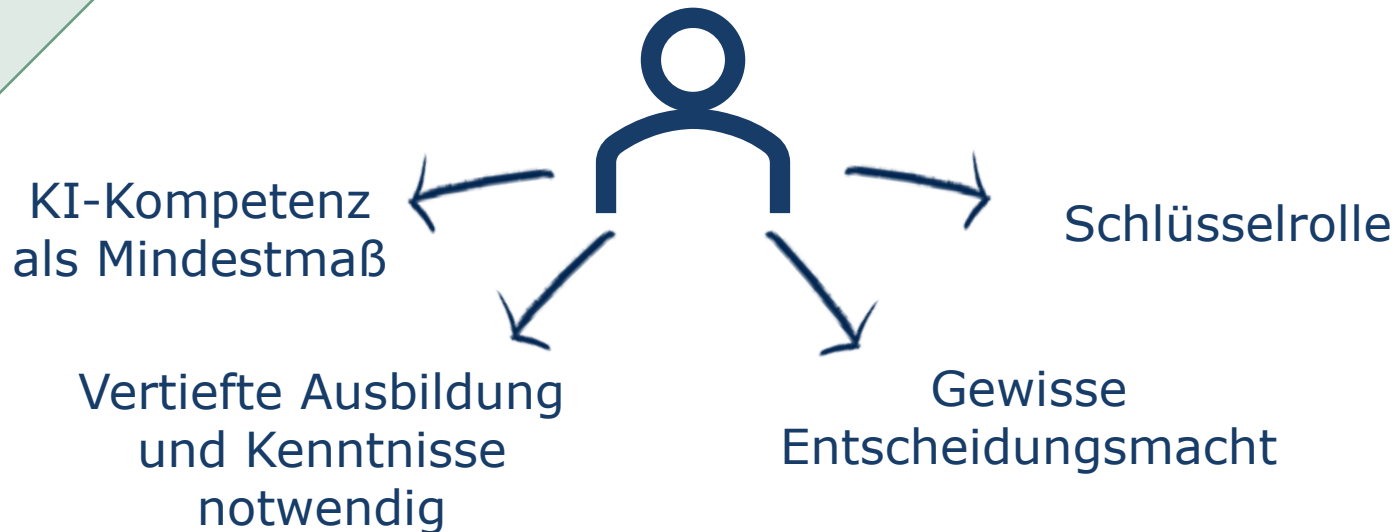
Betreiberpflichten im Fokus

Menschliche Aufsicht – Art 26 Abs 2 KI-VO

» Betreiberpflicht bei Inbetriebnahme/Verwendung

Abs 2

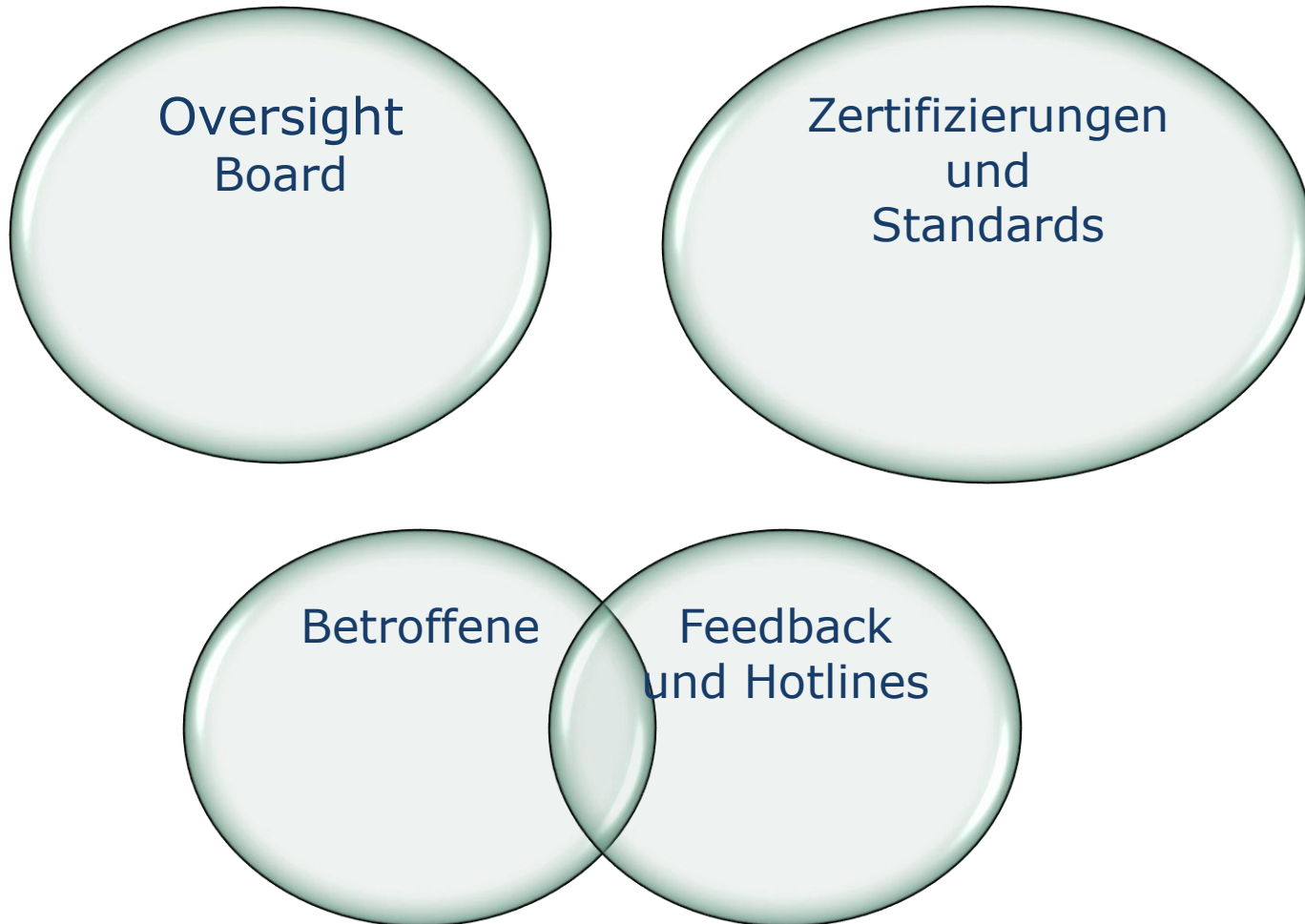
- Übertragung menschlicher Aufsicht an kompetente, ausgebildete und befugte Personen
- Erforderliche Unterstützung zur Ausübung von Aufsichtspflichten
- Verpflichtung zur Umsetzung „geeigneter Instrumente einer **Mensch-Maschine-Schnittstelle**“



Erhaltung von Aufsichtskompetenzen, wie zB durch Simulationsprozesse, regelmäßige Erneuerungen von Qualifikationen

Menschliche Aufsicht

▶ Einbindung Externer denkbar und marktüblich



Wechselwirkungen zu Art 22 DSGVO

Menschliche Aufsicht – Art 14 KI-VO

» Konkrete Designanforderungen und -möglichkeiten

- Möglichkeiten zur richtigen Interpretation, wie zB
 - Transparenz über Datennutzung und Gewichtung
 - Entscheidungsbaum
- Möglichkeiten zur Ablehnung der Verwendung, Außerachtlassung oder Korrektur von Ausgaben, wie zB
 - Möglichkeiten zur Deaktivierung
 - Verhinderung der Verwendung bestimmter Informationen
 - Entscheidungen revidieren
- Unterbrechung, wie zB
 - „Stopp-Taste“

Automatisierte Einzelentscheidung

▶ Kriterium – „ausschließlich“ automatisierte Verarbeitung

Umfang der menschlichen Intervention?

- Parlamentsentwurf: "*based solely or predominantly on automated processing*" → nicht in Finalfassung übernommen
- Telos: Verbot von Entscheidungen "*ohne jegliches menschliches Eingreifen*" (ErwGr 71)

Unzureichend	Ausreichend
Pro-forma Prozesse (WP215 rev 01, 22)	Menschliche Intervention nach Handlungsanweisungen, Richtlinien, Schulungen (BVwG)
Ex-post Kontrolle auf Betroffenenantrag (Art 22 Abs 3 DSGVO <i>e contrario</i>)	Menschliche Intervention ohne detaillierte Kenntnis der Funktionsweise und Logik des Algorithmus
Programmierung des Algorithmus durch natürliche Person	Notwendigkeit von Begründungen und sonstigen Eingaben
Cut off-Entscheidungen (Scoring ohne eigene Überlegungen, Eingriffswille und -möglichkeit)	Menschliche Intervention aufgrund von Fach- und Entscheidungskompetenzen
	Bloße Vorauswahl durch System

Automatisierte Einzelentscheidung

▶ Ausschluss einer ADM bei menschlicher Aufsicht nach KI-VO?

Ausschluss von ADM je nach Umständen denkbar!

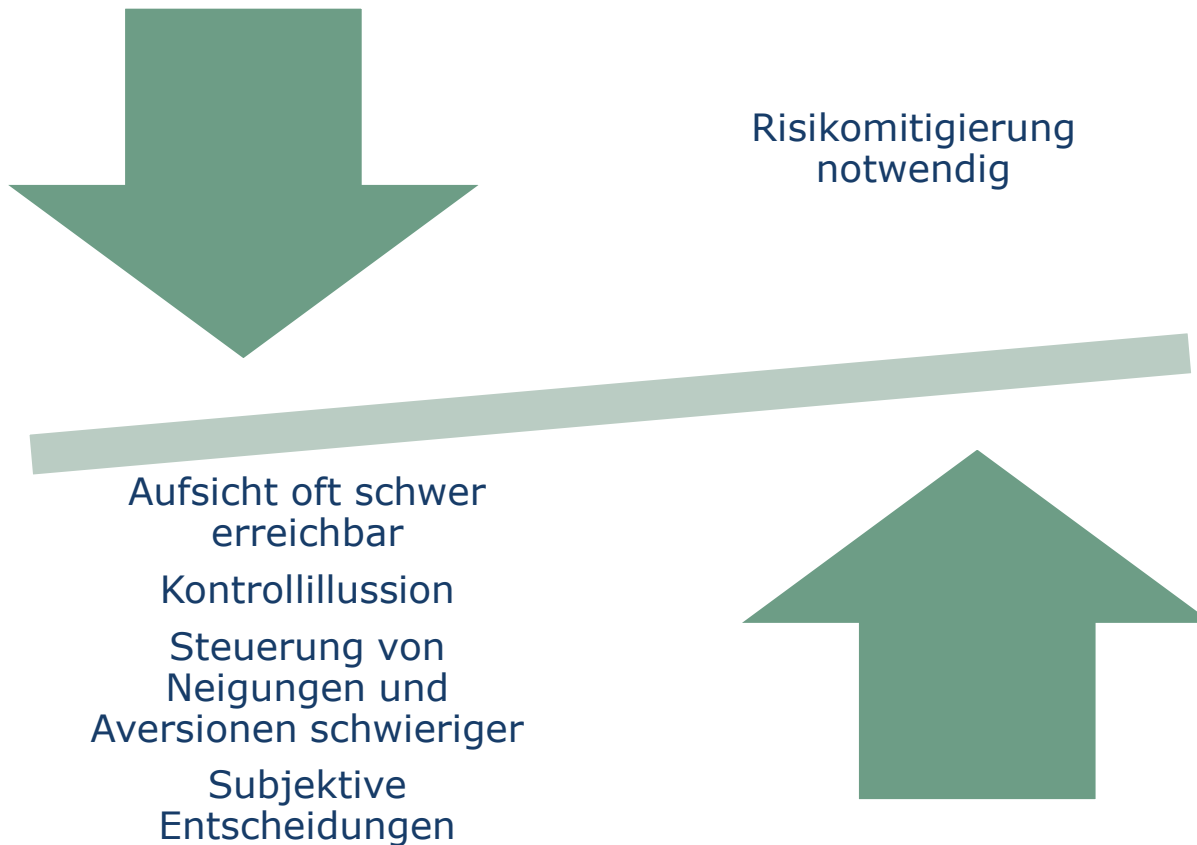
Unzureichend	Ausreichend
Mensch-Maschine-Interaktion	Menschliche Intervention nach Betriebsanleitung
Bloße Hinweise zum besseren Verständnis der Grenzen des KI-Systems	Zwischenergebnisse und menschliche Entscheidung
Programmierung des Algorithmus durch natürliche Person	Menschliche Intervention durch Human in the loop
Bloße Ampelsysteme zur Verlässlichkeit von Outputs, wenn keine menschliche Würdigung erfolgt	Eingebaute Einschränkungen über die sich ein KI-System nicht hinwegsetzen kann

Widerlegung der Hochrisiko-KI-Klassifizierung nach Art 6 Abs 3 kann uE zum Ausschluss einer ADM führen!

Gesamtschau der Umstände ausschlaggebend!

Kritik an der menschlichen Aufsicht

Menschliche Aufsicht als Spannungsfeld



VIELEN DANK FÜR IHRE AUFMERKSAMKEIT

Mag Alexandra Ciarnau

T: +43 1 533 47 95 – 23

alexandra.ciarnau@dorda.at

Client Choice Award Lexology 2022-2025: Blockchain

TIER 1 Legal500 2007-2026: TMT

TIER 1 Legal500 2020-2026: Data Privacy & Data Protection

TIER 1 Legal500 2021-2026: Intellectual Property

BAND 1 Chambers Europe 2008-2026: TMT:IT

BAND 1 Chambers Europe 2026: TMT:Data Protection

DORDA Rechtsanwälte GmbH · Universitätsring 10 · 1010 Wien · www.dorda.at



The Legal 500 (2025)
Axel Anderl (TMT)
Hall of Fame



The Legal 500 (2025)
TMT
Tier 1



The Legal 500 (2025)
Data Privacy & Data
Protection
Tier 1

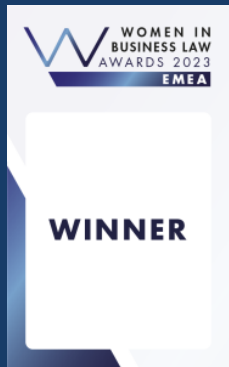


Trend Anwaltsranking (2024)
Axel Anderl
Data Protection, IP and Media
Top 1 overall ranking



Managing IP (2025)
**Austrian Copyright Firm of
the Year**

D O R D A



Austria Firm of the Year 2024
**Gender Diversity National Firm
2024**



**Client Choice winner
IT & Internet**

Client Choice Awards 2025



Who's Who Legal (2025)
Axel Anderl (Data Privacy & Protection)
Thought Leader Global Elite



Chambers Europe (2025)
TMT:IT
Band 1

