

FFG MISSION EXPERT TALK

KI ZWISCHEN FORTSCHRITT UND VERANTWORTUNG

Risiken des KI-Einsatzes für Staat und Gesellschaft

 Dr.in Ronke Babajide

Juni 2026 | Wien, Österreich

ABOUT ME

Dr.in Ronke Babajide

-  Quereinsteigerin aus den Naturwissenschaften in die IT
-  Über 30 Jahre erfolgreich in der Technologiebranche tätig
-  Managerin Presales Systems Engineering bei Fortinet
-  Host des Podcasts "Women in Technology Spotlight" & Mentorin

WARUM DIESES THEMA HEUTE?

Ein Jahr später: Der Wandel

-  Mein letzter Talk war vor einem Jahr. In diesem Jahr hat sich die KI-Welt rasant weiterentwickelt.
-  Wir sehen heute reale, messbare Schäden in Unternehmen und Staat, wo wir früher oft nur theoretische Risiken diskutiert haben.

Befund 2026: "Responsible AI hält nicht Schritt mit den wachsenden Fähigkeiten der KI. Sicherheits-Benchmarks hinken hinterher, Vorfälle steigen scharf an."

 [Quelle: International AI Safety Report 2026](#)

“

We can't assume that handing responsibilities over to a machine will eliminate human biases. Artificial intelligence, after all, is constructed and taught by humans.

– JOY BUOLAMMINI

⚙️ DIE MECHANIK DER VERZERRUNG



Technik wird von Menschen gemacht

Kein Mensch ist frei von Vorurteilen. Jeder Entwickler ist ein Produkt seiner Konditionierung, die unbewusst in die Algorithmen einfließt.



Unsere Vorurteile stecken bereits in den Daten

Die riesigen Datenmengen aus dem Internet, die zum Training genutzt werden, spiegeln historische Ungleichheiten und gesellschaftliche Klischees wider.



WEIRD Data Samples

Trainingsdaten stammen meist aus westlichen, gebildeten, industrialisierten, reichen und demokratischen Gesellschaften (Western, Educated, Industrialized, Rich, Democratic).

Der Verstärkungs-Effekt

Der trainierte Algorithmus lernt diese Muster nicht nur, sondern wendet sie systematisch an und verstärkt sie in neuen Entscheidungen.

NIEDERLANDE

Kinderbetreuungsskandal (2013–2019)

- ! Ein KI-Algorithmus stempelte 26.000 Familien fälschlicherweise als Betrüger ab.
- ! Menschen mit doppelter Staatsbürgerschaft wurden überproportional häufig markiert.
- ! Viele Familien wurden in den finanziellen Ruin getrieben; Hunderte Kinder wurden ihren Eltern weggenommen.

🔗 **Politische Konsequenz:** Premierminister Mark Rutte und sein gesamtes Kabinett traten 2020 zurück.

AUSTRALIEN

Der "Robodebt"-Skandal (2015–2019)

- ! Ein automatisiertes System glich Einkommensdaten ab und forderte zu Unrecht Geld von Sozialhilfeempfängern zurück.
- ! Über 400.000 fehlerhafte Schuldbescheide wurden ohne menschliche Überprüfung verschickt.
- ! Eine königliche Untersuchungskommission bezeichnete das System als "illegal" und "grausam".

💰 **Rechtliche Konsequenz:** Im September 2025 wurde eine historische Rekordentschädigung an die Opfer ausgezahlt.

DEEPPFAKE-KRISE

Grok (Januar 2026)

- ⚠ Ein Update des Aurora-Modells ermöglichte die massenhafte Generierung nicht-einvernehmlicher Bilder realer Personen.
- 📈 Schätzungen zufolge wurden in nur 11 Tagen rund 3 Millionen sexualisierte Bilder generiert, darunter 23.000 von Kindern.
- 🚫 Massiver internationaler Backlash: Malaysia und Indonesien verboten das Tool; EU und UK starteten Untersuchungen.

🔗 Quelle: Reuters (2026)

PREISDISKRIMINIERUNG

Instacart (Dezember 2025)

- 💰 Einsatz eines KI-Preistests, der verschiedenen Kunden unterschiedliche Preise für identische Produkte im selben Laden anzeigte.
- 👁️ Vollständiger Mangel an Transparenz: Kunden wurden nicht über die dynamische, algorithmische Preisanpassung informiert.
- 🔄 Nach heftigem Protest von Verbraucherschützern und rechtlichem Druck stellte das Unternehmen das Programm ein.

🔗 Quelle: CNBC (2025)

MENTALE GESUNDHEIT

Character.AI (Januar 2026)

- 🧑 Ein 14-jähriger Schüler beging Suizid nach monatelanger, emotionaler Abhängigkeit von einem KI-Begleiter.
- 👤 Die Familie klagte; Google und Character.AI einigten sich im Januar 2026 auf einen weitreichenden Vergleich.

Die US-Behörde FTC startete eine umfassende Untersuchung zu den psychologischen Risiken von KI-Companion-Apps.

🔗 Quelle: CNN / FTC (2026)

ALBANIEN: DIE KI-MINISTERIN

”Diella” für öffentliches Beschaffungswesen

 Einführung im September 2025 durch Premierminister Edi Rama zur vollautomatisierten Steuerung aller öffentlichen Ausschreibungen.

Das Versprechen: Ein zu 100 % korruptionsfreies Vergabeverfahren durch Ausschaltung des menschlichen Faktors.

 **Das Problem:** Wenn die KI eine Fehlentscheidung trifft, kann sie weder ihre Gründe erklären, noch kann sie rechtlich zur Rechenschaft gezogen werden.

 [Quelle: CEPA \(September 2025\)](#)

USA: DOGE & MILITÄR-KI

Automatisierung & Rüstungskonflikte

 Die Regierungsbehörde **DOGE** nutzt KI-Modelle, um massiv Bundesregulierungen zu streichen und sensible Staatsdaten einzuspeisen.

 **Pentagon-Blacklist (Februar 2026):** Anthropic wurde gesperrt, weil sich das Unternehmen weigerte, ethische Schranken für autonome Waffensysteme zu entfernen.

 **Die Folge:** Konkurrenten wie OpenAI und xAI füllten die Lücke und stimmten ”allen rechtmäßigen militärischen Nutzungen” zu.

 [Quellen: Tech Policy Press / Crescendo AI \(2025/2026\)](#)

DEMOKRATISCHE KERNFRAGE

KI-Systeme im öffentlichen Sektor agieren als ”Black Boxes”. Da sie ihre Entscheidungsfindung nicht transparent begründen können, verletzen sie das fundamentale Prinzip der demokratischen Rechenschaftspflicht und staatlichen Haftung.

DER BLINDE FLECK: ÖKOLOGIE DER KI

07

⚡ STROMBEDARF

+17 %

Anstieg des Stromverbrauchs von Rechenzentren in 2025. Bis 2026 prognostizierter Verbrauch von **650–1.050 TWh** – das entspricht dem Strombedarf des fünftgrößten Landes der Erde.

💧 WASSERVERBRAUCH

760+ Mrd. L

Wasserverbrauch für die Kühlung von KI-Servern im Jahr 2025. Das übersteigt das gesamte weltweit konsumierte Flaschenwasser und belastet lokale Ökosysteme massiv.

☁️ CO₂-FUSSABDRUCK

33–80 Mio. t

Geschätzter CO₂-Ausstoß von KI-Systemen weltweit im Jahr 2025. Die Hardware-Produktion und der Dauerbetrieb konterkarieren globale Klimaschutzbemühungen.

⚠️ ZIELKONFLIKT FÜR DIE EU-MISSIONEN

Wir nutzen KI, um den Klimawandel zu erforschen, Böden zu analysieren und Städte zu optimieren. Doch die physische Infrastruktur der KI beschleunigt die Erderwärmung und verbraucht kostbares Wasser. Die nationalen Aktionspläne müssen diesen Ressourcenverbrauch zwingend in ihre CO₂- und Umweltbilanzen einrechnen.

📅 UMSETZUNGS-FAHRPLAN

Aug 2024

Inkrafttreten: Der EU AI Act wird offiziell Gesetz.

Feb 2025

Verbote greifen: Verbot von Social Scoring, biometrischer Kategorisierung und manipulativer KI.

Aug 2025

General Purpose AI: Erste strenge Transparenzregeln für große Modelle (wie ChatGPT, Claude).

Aug 2026

Vollständige Anwendung: Strenge Auflagen für **High-Risk-Systeme** treten in Kraft.

⚠️ HIGH-RISK IM ÖFFENTLICHEN SEKTOR



Strafverfolgung & Justiz: KI zur Risikobewertung von Straftätern oder zur automatisierten Fallanalyse.

Sozialleistungen: Systeme zur automatisierten Vergabe, Kürzung oder Überprüfung von Sozialhilfe (z.B. Robodebt).



Personalwesen: KI-gestützte Bewerberauswahl und Leistungsbewertung im öffentlichen Dienst.



Biometrische Identifizierung: Echtzeit- und nachträgliche Gesichtserkennung im öffentlichen Raum.



[EU AI Act Portal](#)



[High-Risk Rules \(Art. 6\)](#)

MEDIZIN & GESUNDHEIT

Wissenschaftliche Durchbrüche

 **Antibiotika-Entdeckung (August 2025):** Forscher des MIT haben mithilfe generativer KI völlig neue Verbindungen entwickelt, die resistente Superkeime wie MRSA abtöten können.

 **Krebsforschung (Anfang 2026):** KI transformiert das Wirkstoffdesign und beschleunigt klinische Studien, indem sie Bilddaten in direkt umsetzbare Therapievorschlüsse übersetzt.

[Quelle: MIT News \(2025\)](#)

KLIMA & UMWELT

Erforschung & Optimierung

 **Präzise Wettermodelle (Dezember 2025):** Die NOAA hat drei neue KI-gestützte globale Wettermodelle im operativen Dienst eingeführt, die Extremwetterereignisse drastisch schneller vorhersagen.

 **Ressourcen-Effizienz:** Intelligente Netze (Smart Grids) nutzen KI, um die Einspeisung erneuerbarer Energien in Echtzeit zu steuern und Energieverluste um bis zu 19 % zu senken.

[Quelle: NOAA / ETC Journal \(2026\)](#)

ÖFFENTLICHER SEKTOR

Bürokratieabbau & Schutz

 **Entlastung der Verwaltung (September 2025):** OECD-Berichte zeigen, dass KI repetitive administrative Aufgaben automatisiert, wodurch Beamte mehr Zeit für komplexe Einzelfälle haben.

 **Geo-KI im Katastrophenschutz:** Einsatz von lokal angepassten Drohnen- und Satellitendaten zur automatisierten, sekundenschnellen Schadensbewertung nach Naturkatastrophen.

[Quelle: OECD Report \(2025\)](#)

☰ FÜNF HANDLUNGSFELDER

-  **Menschliche Kontrolle garantieren:** Keine automatisierte Entscheidung im öffentlichen Sektor ohne menschliche Überprüfung und Verantwortung.
-  **Trainingsdaten bereinigen:** Datensätze aktiv auf historische Verzerrungen prüfen, bereinigen und diversifizieren.
-  **Teams diversifizieren:** Nur diverse Entwicklungsteams können blinde Flecken in der Algorithmen-Entwicklung erkennen.
-  **Transparenz einfordern:** Einsatz von erklärbarer KI (Explainable AI) statt intransparenter "Black-Box"-Entscheidungen.
-  **Ökologischen Fußabdruck mitdenken:** Ressourcenverbrauch von KI-Infrastrukturen in die CO₂-Bilanz der EU-Missionen einrechnen.

💬 DREI DISKUSSIONSTHESEN

THESE 1: DAS LETZTE WORT

"Eine KI darf niemals das letzte Wort haben. Im öffentlichen Sektor muss jede automatisierte Entscheidung von einem Menschen verantwortet werden."

THESE 2: DER KLIMAKONFLIKT

"Der KI-Boom gefährdet unsere Klimaziele. Die nationalen Aktionspläne für Boden, Wasser und Klima müssen den Ressourcenverbrauch von KI-Infrastrukturen einrechnen."

THESE 3: DIVERSITÄT ALS SICHERHEIT

"Diversität in der Entwicklung ist kein ethischer Luxus, sondern eine Sicherheitsanforderung. Ohne sie bauen wir Diskriminierung direkt in den Staat ein."