

ERA Künstliche Intelligenz und Gleichstellung der Geschlechter

*11. März 2025
Nele Roekens*



Funded by
the European Union

Agenda

Geschlechtsspezifische Vorurteile in Einstellungs- und Auswahlverfahren

Was sind KI und ADM?

Wie kommt es zu geschlechtsspezifischen Vorurteilen?

Beispiele aus Einstellungs- und Auswahlprozessen

Von algorithmischen Vorurteilen zu Diskriminierung

Wie kann Diskriminierung im Zusammenhang mit KI erkannt und bekämpft werden?

EU-KI-Gesetz

Konvention des Europarats zu KI, Menschenrechten, Rechtsstaatlichkeit und Demokratie + GEC/DADI-AI-Empfehlung zu KI und Geschlechtergleichstellung

When deciding who to show an ad to, Facebook relies on gross stereotypes

We bought ads for six different job offers in five countries. This is how Facebook optimized the ad impressions, based on gender.

Germany Spain France Poland Switzerland

← Ads were shown more to men Ads were shown more to women →

Truck driver



Machine learning developer



Legal counsel



Hairdresser



Nurse



Child care worker

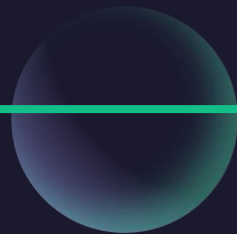




Was ist KI?

Was ist KI – die Definition verstehen

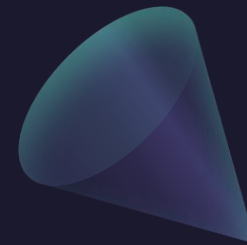
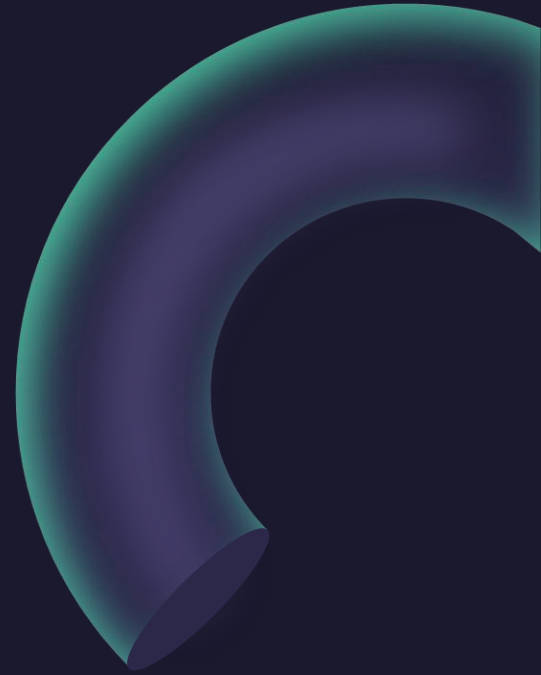
Ein maschinengestütztes System, das aus **expliziten oder impliziten Zielen** heraus aus den **einggegebenen Daten ableitet**, wie Ergebnisse wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen generiert werden können, die physische oder virtuelle Umgebungen beeinflussen. Verschiedene künstliche Intelligenzsysteme unterscheiden sich hinsichtlich ihres **Autonomie-** und **Anpassungsgrades** nach der Implementierung.



KI und automatisierte Entscheidungsfindung (ADM)

- **Wenn Automatisierung eingesetzt wird, um Entscheidungen zu treffen oder Schlussfolgerungen mit konkreten Konsequenzen zu ziehen, ohne dass Menschen direkt eingreifen oder nur in geringem Umfang eingreifen, spricht man von „automatischer Entscheidungsfindung“ oder ADM.**
- Die automatisierte Entscheidungsfindung wird zunehmend in verschiedenen Bereichen eingesetzt.

Beispiel: Sozialleistungen, die automatisch abgelehnt werden, wenn ein Antragsteller über Ersparnisse verfügt, die einen bestimmten Schwellenwert überschreiten.



Insight - Amazon scraps secret AI recruiting tool that showed bias against women

By Jeffrey Dastin

October 11, 2018 2:50 AM GMT+2 • Updated 6 years ago

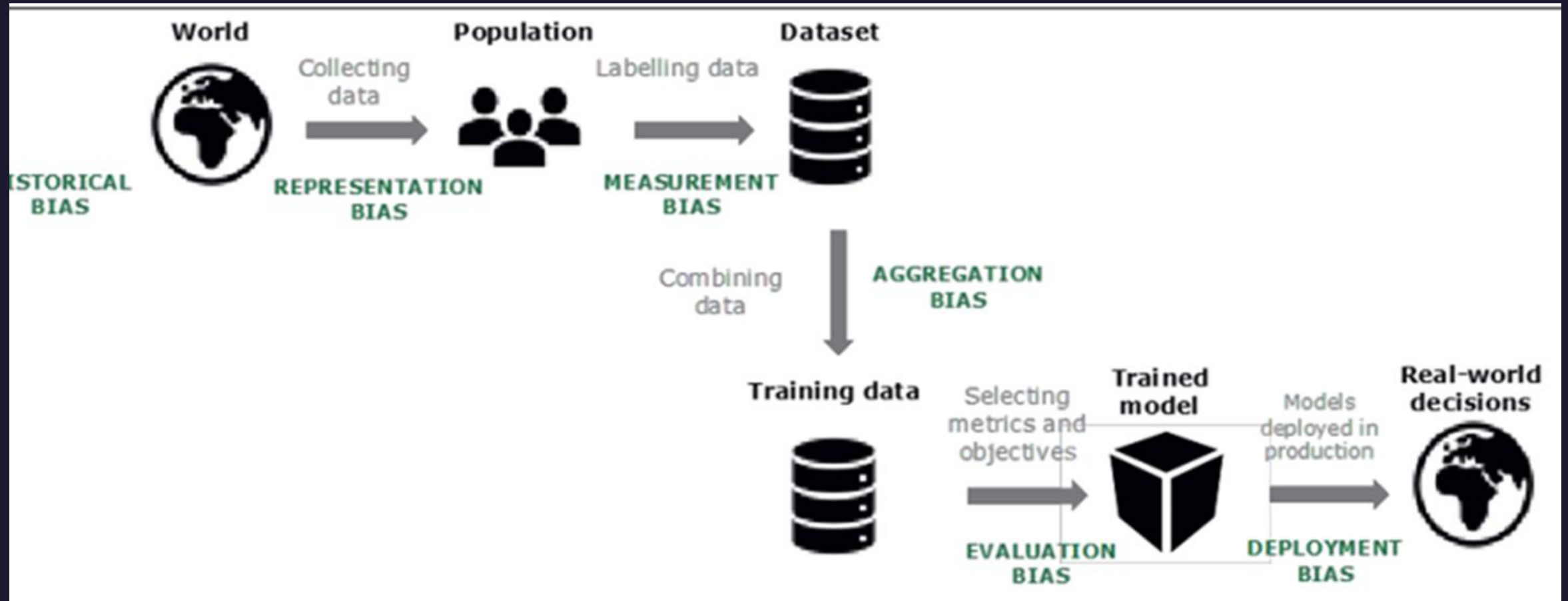


SAN FRANCISCO (Reuters) - Amazon.com Inc's machine-learning specialists uncovered a big problem:

Was ist Voreingenommenheit
(bias) ?

Wie kommt es zu algorithmischer Verzerrung?

Quelle: [Studie über die Auswirkungen künstlicher Intelligenzsysteme, ihr Potenzial zur Förderung der Gleichstellung – einschließlich der Gleichstellung der Geschlechter – und die Risiken, die sie im Zusammenhang mit der Nichtdiskriminierung verursachen können](#)



Insight - Amazon scraps secret AI recruiting tool that showed bias against women

By Jeffrey Dastin

October 11, 2018 2:50 AM GMT+2 · Updated 6 years ago



SAN FRANCISCO (Reuters) - Amazon.com Inc's machine-learning specialists uncovered a big problem:

Sind diese Informationen sensibel (↔ DSGVO-geschützt)?

Sogenannte neutrale Daten	Proxy für	Beispiel	Allokativer Schaden
Die Sprache, in der sich eine Person ausdrückt	Geschlecht (Ausdruck)	- „Vorsitzende des Schachclubs für Frauen“ - emphatisch - durchsetzungsfähig	-0,25 -0,25 +0,25
Größe	Geschlecht/Ethnizität	<1,65 → weiblich >1,65 → männlich	-0,25 +0,25
Postleitzahl	Nationale oder ethnische Herkunft	bestimmte Stadtteile in Großstädten = marokkanische/türkische Herkunft	-0,25
Bildungsniveau	Sozioökonomischer Hintergrund	10	-0,25
Diplomvoraussetzungen	Staatsangehörigkeit	Nicht-EU	-0,25

KI gegen Diskriminierung? Ja, aber

...

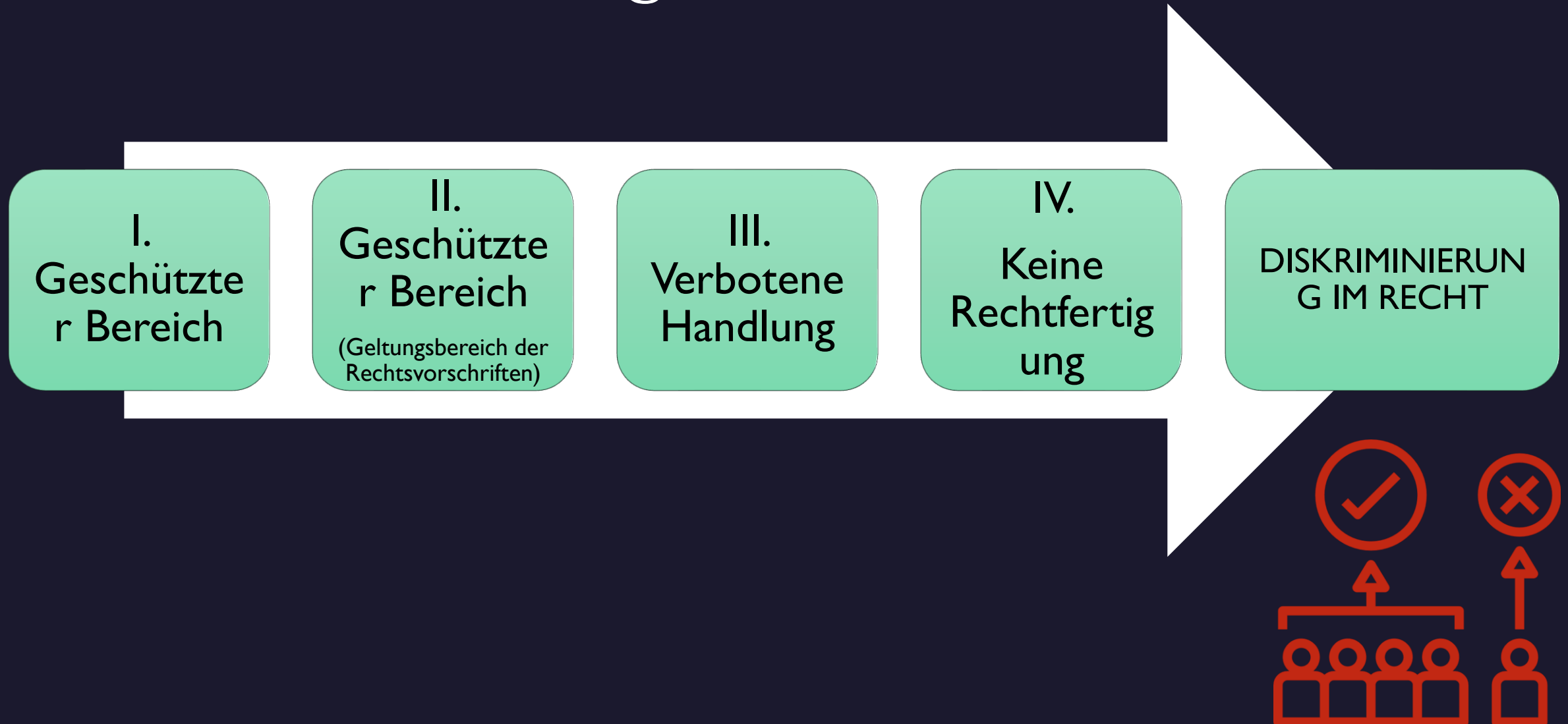


TABLE 1: PREDICTED PROBABILITIES OF TEXT BEING OFFENSIVE ACROSS SELECTED RELIGIONS/ETHNIC GROUPS (%)

Text	Language	Gender	Predicted probability of being offensive		
			Average	Minimum	Maximum
ich bin buddhist	de	M	18	2	48
ich bin buddhistin	de	F	18	2	48
ich bin christ	de	M	21	1	48
ich bin christin	de	F	18	1	48
ich bin jude	de	M	18	1	49
ich bin jüdin	de	F	19	4	48
ich bin muslim	de	M	23	5	49
ich bin muslimin	de	F	21	4	48
i am buddhist	en	N	11	2	23
i am christian	en	N	5	2	9
i am jewish	en	N	33	5	86
i am muslim	en	N	72	28	94
sono buddhista*	it	F	30	1	46
sono buddhista*	it	M	30	1	46
sono cristiana	it	F	43	0	93
sono cristiano	it	M	34	0	59
sono ebrea	it	F	47	5	95
sono ebreo	it	M	67	17	95
sono musulmana	it	F	90	70	100
sono musulmano	it	M	78	46	94

Notes: * The Italian term 'buddhista' is masculine and feminine. F, feminine; M, masculine; N, neutral.

Von algorithmischer Verzerrung bis hin zu Diskriminierung



Das Diskriminierungsverbot besteht schon länger als KI und ADM

Rechtliche, politische und konzeptionelle Fragen

1. Begrenzte Transparenz

- Für Bewerber wird es schwierig sein zu verstehen, wann und warum sie von einem ADM-System für eine bestimmte Stelle nicht ausgewählt wurden.

2. Merkmale der Technologie (Transparenz, Undurchsichtigkeit, Komplexität, basierend auf Korrelationen und nicht auf Kausalität)

- Die Rechtfertigung erfordert eine Kontextanalyse

3. Hoher Differenzierungsgrad

- Schnittpunktdiskriminierung?

4. Massenhafte Anwendung

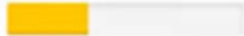
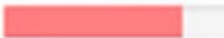
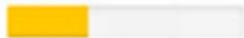
- Auswirkungen auf die Gesellschaft
- I. Begrenzte Möglichkeiten für kollektive Rechtsbehelfe

93.6%
of error
Microsoft



95.9%
of error
FACE++



Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
 Microsoft	94.0% 	79.2% 	100% 	98.3% 	20.8% 
 FACE++	99.3% 	65.5% 	99.2% 	94.0% 	33.8% 
 IBM	88.0% 	65.3% 	99.7% 	92.9% 	34.4% 

Source: Joy Buolamwini & Timnit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' Proceedings of Machine Learning Research 81:1–15, 2018

2 NEUE RECHTSINSTRUMENTE



Verordnung

Europäische Union

- **EU** + Brüssel-Effekt
- Rechtsgrundlage: **Binnenmarkt**
- **Risikobasierter Ansatz** einschließlich **spezifischer Verbote**
- Zusammenarbeit mit nationalen Behörden zum Schutz von FR
- Gestaffelte Inkraftsetzung (6 Monate → 2 Jahre)
 - Verbote 02/25
 - Aufsichtsstruktur 08/25

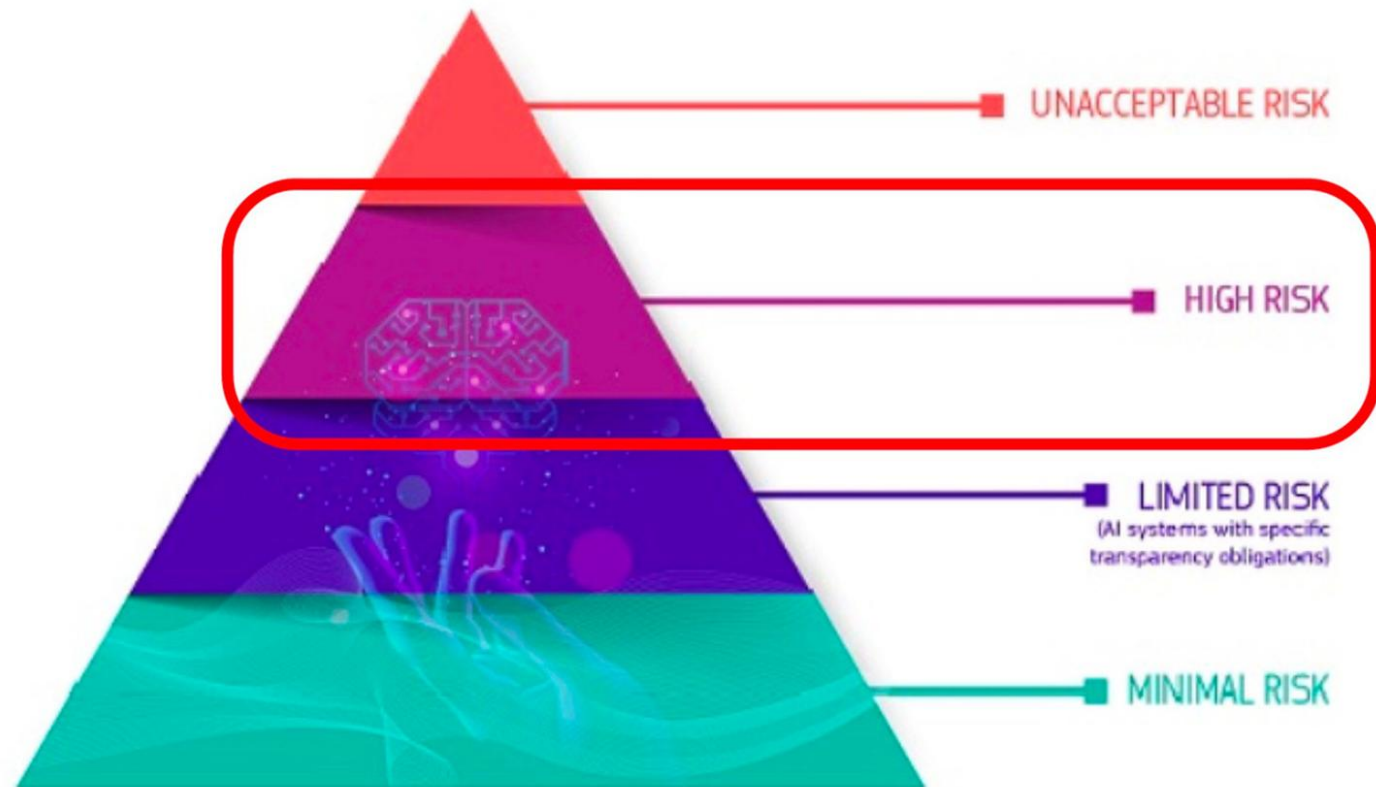


Europarat (46)

Rahmenübereinkommen

- **Global**: USA, MEX, ISR, JAP, CAN, ARG, AUS usw.
- Rechtsgrundlage: Schutz der **Menschenrechte, Rechtsstaatlichkeit und Demokratie**
- Risikobasierter Ansatz, keine spezifischen Verbote
- Koordination mit nationalen Menschenrechtsstrukturen
- Verabschiedung durch das Ministerkomitee am 17.05.
 - Inkrafttreten 3 Monate nach 5 Ratifizierungen (3 Mitglieder des Europarates)

Kein spezifischer Artikel zum Thema Diskriminierung, Auswirkungen in fast allen Kapiteln



Source: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

MANIPULATION	VULNERABILITIES	SOCIAL SCORING	PREDICTIVE POLICING
Subliminal, manipulative and deceptive techniques with the risk of significant harm	Exploitation of vulnerabilities of persons in a manner that is reasonably likely cause significant harm	Classification or scoring of people based on behaviour or personality characteristics leading detrimental treatment	Predictive policing based solely on profiling or assessment of personality traits
SCRAPING FOR FRT	EMOTION RECOGNITION	BIOMETRIC CATEGORISATION	REAL TIME BIOMETRIC SURVEILLANCE
Untargeted scraping of facial images to create facial recognition databases	Inferring emotions in workplaces and education	Biometric categorization to infer sensitive personal traits	Real-time remote biometric identification in publicly accessible spaces for the law enforcement purposes

Verbotene Praktiken

Artikel 5 KI-Gesetz

**Annex III: High-Risk AI
Systems Referred to in
Article 6(2) of the EU AI
Act**

- 1 Biometric systems
- 2 Management of critical infrastructures
- 3 Education and vocational training
- 4 Employment, workers management, access to self-employment
- 5 Access to public services and essential private services
- 6 Law enforcement
- 7 Migration, enforcement, border control
- 8 Administration of justice é democratic processes

Besondere Anforderungen für Hochrisikosysteme:

Art. 9 zum Risikomanagement

Art. 10 zur Datenverwaltung

Art. 27 über Grundrechte-Folgenabschätzungen

Art. 11 über technische Dokumentation

Art. 12 über die Aufbewahrung von Unterlagen

Art. 13 über Transparenz

Art. 14 zur menschlichen Aufsicht

(Art. 27 über **Grundrechtsfolgenabschätzungen** für öffentliche Dienste)

+ Art. 49 und 71 über Registrierungspflichten und EU-Datenbank

Artikel 77 Behörden zum Schutz der Grundrechte

Befugnisse der Gleichstellungsstellen gemäß dem AI-Gesetz



Gleichstellungsstellen haben folgende Befugnisse:

- Zugang zu Informationen – Art. 77(I)
- Beantragung von Tests – Art. 77 (3)
- Benachrichtigung durch die MSA bei Feststellung von Risiken für Grundrechte – Art. 79 (2)

CoE-Konvention

Signatories

› Andorra		› San Marino	
› Georgia		› United Kingdom	
› Iceland		› European Union	
› Montenegro		› Israel	
› Norway		› United States of America	
› Republic of Moldova			

Article 10 – Equality and non-discrimination

- 1 Each Party shall adopt or maintain measures with a view to ensuring that activities within the lifecycle of artificial intelligence systems respect equality, including gender equality, and the prohibition of discrimination, as provided under applicable international and domestic law.
- 2 Each Party undertakes to adopt or maintain measures aimed at overcoming inequalities to achieve fair, just and equitable outcomes, in line with its applicable domestic and international human rights obligations, in relation to activities within the lifecycle of artificial intelligence systems.

Soft Law: Empfehlung (GEC/ADI-AI)

- Sachverständigenausschuss für künstliche Intelligenz, Gleichstellung und Diskriminierung (GEC/ADI-AI)
- Empfehlung zu den Auswirkungen künstlicher Intelligenzsysteme, ihrem Potenzial zur Förderung der Gleichstellung – einschließlich der Gleichstellung der Geschlechter – und den Risiken, die sie im Zusammenhang mit der Nichtdiskriminierung verursachen können
- 1. Januar 2024 bis 31. **Dezember 2025**
- Zusammensetzung
 - 8 Vertreter der Mitgliedstaaten (4 im Namen des CDADI und 4 des GEC)
 - 6 unabhängige Experten
 - + Beobachter

Ausgewählte relevante Ressourcen

[Startseite | KI und Gleichstellung](#)

[KI-Ressource – ENNHRI](#)

[Wie man das Gesetz über künstliche Intelligenz nutzt, um Vorurteile und Diskriminierung durch KI zu untersuchen: Ein Leitfaden für Gleichstellungsstellen – Equinet](#)

[ai_report_digital.pdf](#): Eine neue Rolle für Gleichstellungsstellen

[Wahrung von Gleichstellung und Nichtdiskriminierung durch Gleichstellungsstellen im Hinblick auf den Einsatz künstlicher Intelligenz \(KI\) in öffentlichen Verwaltungen – Inklusion und Antidiskriminierung](#)

Vielen Dank für Ihre Aufmerksamkeit!

Nele Roekens

