

Eine wertebasierte Perspektive auf den Entwurf zur KI-Verordnung

Weiterführende Informationen und Quellen

Tim Vallée | iRights.Lab

11. Speyerer Forum zur digitalen Lebenswelt |
13. September 2022

Relevante Werte für den Umgang mit KI

- „Menschenzentrierung & Gemeinwohl“,
- „Fairness & Nichtdiskriminierung“,
- „Erklärbarkeit & Transparenz“,
- „Privatsphäre & Persönlichkeitsschutz“,
- „Sicherheit & Robustheit“,
- „Intervenierbarkeit & Verantwortung“ und
- „ökologische Nachhaltigkeit & Ressourcenschonung“

Erläuterungen finden sich in der Wertegrundlage für den KI-Einsatz in der Arbeits- und Sozialverwaltung, [hier](#) abrufbar

Diskussionspunkte bei einer Analyse des Entwurfs aus Wertesicht

- ökologische Nachhaltigkeit / Umwelt kein Schutzgut (vgl. dazu die Studie Vallée/Scheiber Die KI-Verordnung als Ausdruck digital-souveräner und nachhaltiger Politik? i.E.)
 - Methodologie, Metrik, Standards und Prozesse zur Abschätzung von Risiken und zur Prüfung und Umsetzung von Risikominimierungsmaßnahmen fehlen
 - teils erheblichen Diskriminierungsrisiken kann so kaum angemessen begegnet werden (zu Diskriminierungsrisiken vgl. etwa die Studie von Orvat im Auftrag der ADS, [hier](#))
 - Anwendungsbereich für Hochrisiko-Systeme zu eng -> es fehlen z.B. Algorithmen, die die Newsfeeds und Diskurse bei „sozialen“ Medien (mit)bestimmen (vgl. Anhang 3 der KI-Verordnung zum Anwendungsbereich von Hochrisiko-Systemen, auf den Folien am Ende)
 - Defizite beim Datenschutz, dazu die gemeinsame Stellungnahme der Datenschutzbehörden, [hier](#)
 - Informations- und Beschwerdemöglichkeiten für Betroffene fehlen
 - Effektive Aufsicht durch staatliche Stellen bleibt herausfordernd -> national zu benennende Marktüberwachungsbehörden müssen mit ausreichend Ressourcen und politischem Willen ausgestattet werden
-

Relevante Dokumente auf EU-Ebene

- Darstellung der EU-KI-Strategie einschließlich einer Übersicht über das bisherige Verfahren [hier](#)
 - Der Entwurf einschließlich Dokumenten aus den Ausschüssen [hier](#)
 - Ethik-Leitlinien für eine vertrauenswürdige KI (High-Level Expert Group für KI der Europäische Kommission) [hier](#)
-

Anhang 3: (Acht) Anwendungsbereiche für Hochrisiko-Systeme

Als Hochrisiko-KI-Systeme gemäß Artikel 6 Absatz 2 gelten die in folgenden Bereichen aufgeführten KI-Systeme:

1. Biometrische Identifizierung und Kategorisierung natürlicher Personen:

a) KI-Systeme, die bestimmungsgemäß für die biometrische Echtzeit-Fernidentifizierung und nachträgliche biometrische Fernidentifizierung natürlicher Personen verwendet werden sollen;

2. Verwaltung und Betrieb kritischer Infrastrukturen:

a) KI-Systeme, die bestimmungsgemäß als Sicherheitskomponenten in der Verwaltung und im Betrieb des Straßenverkehrs sowie in der Wasser-, Gas-, Wärme- und Stromversorgung verwendet werden sollen;

3. Allgemeine und berufliche Bildung:

a) KI-Systeme, die bestimmungsgemäß für Entscheidungen über den Zugang oder die Zuweisung natürlicher Personen zu Einrichtungen der allgemeinen und beruflichen Bildung verwendet werden sollen;

b) KI-Systeme, die bestimmungsgemäß für die Bewertung von Schülern in Einrichtungen der allgemeinen und beruflichen Bildung und für die Bewertung der Teilnehmer an üblicherweise für die Zulassung zu Bildungseinrichtungen erforderlichen Tests verwendet werden sollen;

4. Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit:

a) KI-Systeme, die bestimmungsgemäß für die Einstellung oder Auswahl natürlicher Personen verwendet werden sollen, insbesondere für die Bekanntmachung freier Stellen, das Sichten oder Filtern von Bewerbungen und das Bewerten von Bewerbern in Vorstellungsgesprächen oder Tests;

b) KI-Systeme, die bestimmungsgemäß für Entscheidungen über Beförderungen und über Kündigungen von Arbeitsvertragsverhältnissen, für die Aufgabenzuweisung sowie für die Überwachung und Bewertung der Leistung und des Verhaltens von Personen in solchen Beschäftigungsverhältnissen verwendet werden sollen;

5. Zugänglichkeit und Inanspruchnahme grundlegender privater und öffentlicher Dienste und Leistungen:

a) KI-Systeme, die bestimmungsgemäß von Behörden oder im Namen von Behörden verwendet werden sollen, um zu beurteilen, ob natürliche Personen Anspruch auf öffentliche Unterstützungsleistungen und -dienste haben und ob solche Leistungen und Dienste zu gewähren, einzuschränken, zu widerrufen oder zurückzufordern sind;

b) KI-Systeme, die bestimmungsgemäß für die Kreditwürdigkeitsprüfung und Kreditpunktbewertung natürlicher Personen verwendet werden sollen, mit Ausnahme von KI-Systemen, die von Kleinanbietern für den Eigengebrauch in Betrieb genommen werden;

c) KI-Systeme, die bestimmungsgemäß für die Entsendung oder Priorisierung des Einsatzes von Not- und Rettungsdiensten, einschließlich Feuerwehr und medizinischer Nothilfe, verwendet werden sollen;

6. Strafverfolgung:

- a) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden für individuelle Risikobewertungen natürlicher Personen verwendet werden sollen, um das Risiko abzuschätzen, dass eine natürliche Person Straftaten begeht oder erneut begeht oder dass eine Person zum Opfer möglicher Straftaten wird;
- b) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden als Lügendetektoren und ähnliche Instrumente oder zur Ermittlung des emotionalen Zustands einer natürlichen Person verwendet werden sollen;
- c) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden zur Aufdeckung von Deepfakes gemäß Artikel 52 Absatz 3 verwendet werden sollen;
- d) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden zur Bewertung der Verlässlichkeit von Beweismitteln im Zuge der Ermittlung oder Verfolgung von Straftaten verwendet werden sollen;
- e) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden zur Vorhersage des Auftretens oder erneuten Auftretens einer tatsächlichen oder potenziellen Straftat auf der Grundlage des Profils natürlicher Personen gemäß Artikel 3 Absatz 4 der Richtlinie (EU) 2016/680 oder zur Bewertung von Persönlichkeitsmerkmalen und Eigenschaften oder vergangenen kriminellen Verhaltens natürlicher Personen oder von Gruppen verwendet werden sollen;
- f) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden zur Erstellung von Profilen natürlicher Personen gemäß Artikel 3 Absatz 4 der Richtlinie (EU) 2016/680 im Zuge der Aufdeckung, Ermittlung oder Verfolgung von Straftaten verwendet werden sollen;
- g) KI-Systeme, die bestimmungsgemäß zur Kriminalanalyse natürlicher Personen eingesetzt werden sollen und es den Strafverfolgungsbehörden ermöglichen, große komplexe verknüpfte und unverknüpfte Datensätze aus verschiedenen Datenquellen oder in verschiedenen Datenformaten zu durchsuchen, um unbekannte Muster zu erkennen oder verdeckte Beziehungen in den Daten aufzudecken;

7. Migration, Asyl und Grenzkontrolle:

- a) KI-Systeme, die bestimmungsgemäß von zuständigen Behörden als Lügendetektoren und ähnliche Instrumente oder zur Ermittlung des emotionalen Zustands einer natürlichen Person verwendet werden sollen;
- b) KI-Systeme, die bestimmungsgemäß von zuständigen Behörden zur Bewertung eines Risikos verwendet werden sollen, einschließlich eines Sicherheitsrisikos, eines Risikos der irregulären Einwanderung oder eines Gesundheitsrisikos, das von einer natürlichen Person ausgeht, die in das Hoheitsgebiet eines Mitgliedstaats einzureisen beabsichtigt oder eingereist ist;
- c) KI-Systeme, die bestimmungsgemäß von zuständigen Behörden zur Überprüfung der Echtheit von Reisedokumenten und Nachweisunterlagen natürlicher Personen und zur Erkennung unechter Dokumente durch Prüfung ihrer Sicherheitsmerkmale verwendet werden sollen;
- d) KI-Systeme, die bestimmungsgemäß zuständige Behörden bei der Prüfung von Asyl- und Visumanträgen sowie Aufenthaltstiteln und damit verbundenen Beschwerden im Hinblick auf die Feststellung der Berechtigung der den Antrag stellenden natürlichen Personen unterstützen sollen;

8. Rechtspflege und demokratische Prozesse:

- a) KI-Systeme, die bestimmungsgemäß Justizbehörden bei der Ermittlung und Auslegung von Sachverhalten und Rechtsvorschriften und bei der Anwendung des Rechts auf konkrete Sachverhalte unterstützen sollen.

Kontakt

Tim Vallée, Ass. iur., M.A., LL.M.
Forschung & Projekte
iRights.Lab
t.vallee@irights-lab.de
www.irights-lab.de