

KURZÜBERSICHT: EMPFEHLUNGEN FÜR EINE MENSCHENRECHTSKONFORME REGULIERUNG VON KÜNSTLICHER INTELLIGENZ IN DER EU

KI-ANWENDUNGEN, DIE EIN NICHT VERTRETbares RISIKO FÜR DIE MENSCHENRECHTE MIT SICH FÜHREN, SOLLEN IN DER KI-VERORDNUNG VERBOTEN WERDEN. AMNESTY INTERNATIONAL FORDERT:

- Ein **umfassenderes Verbot von biometrischer Überwachung wie beispielsweise Gesichtserkennung** im öffentlichen Raum, das nicht nur den Einsatz in Echtzeit, sondern auch nachträglich untersagt, das den Einsatz durch private Akteure verbietet und das nicht durch weitreichende Ausnahmen ausgehöhlt wird;
- Ein **Verbot vom Einsatz biometrischer Daten für andere Zwecke**, wie beispielsweise Emotionserkennung oder andere biometrische Kategorisierungssysteme;
- Eine **Ausweitung des Verbots von „Social Scoring“** durch Staaten auf den Einsatz durch private Akteure und auf Praktiken von „Social Profiling“;
- Genauere Definitionen von **manipulativen KI-Systemen** und ein Verbot jeglicher verdeckter Manipulation und Konditionierung mittels KI, die schädliche Auswirkungen auf die Menschenrechte oder andere materielle oder immaterielle Güter haben.

DER MENSCHENRECHTSSCHUTZ MUSS VON DER ENTWICKLUNG BIS ZUM EINSATZ VON KI IM MITTELPUNKT STEHEN. AMNESTY INTERNATIONAL FORDERT:

- **Größtmögliche Transparenz** und Verständlichkeit über den Einsatz von KI, sodass die Auswirkungen für Einzelne und Aufsichtsbehörden nachvollziehbar und überprüfbar sind;
- Menschenrechtliche **Sorgfaltspflichten und verpflichtende Folgenabschätzungen**, sowohl für Entwickler*innen als auch für Nutzer*innen, die KI einsetzen.

FÜR DIE EINHALTUNG DES VORGESEHENEN MENSCHENRECHTSSCHUTZES BRAUCHT ES EINE WIRKSAMSE DURCHSETZUNG UND UNABHÄNGIGE KONTROLLE. AMNESTY INTERNATIONAL FORDERT:

- Das Beseitigen und Eindämmen von weitgefassten **Ausnahmeregelungen**, die ganze Kategorien oder Anwendungsbereiche von KI vom Geltungsbereich der KI-Verordnung ausschließen;
- Hinreichend klare **Kriterien für die Risiko-Kategorisierung** von KI-Systemen und die Schaffung von Mechanismen, die eine **Aktualisierung und Ergänzung** solcher Kategorisierungen ermöglichen;
- **Konformitätsprüfungen** durch unabhängige Dritte anstelle von Selbsteinschätzungen, insbesondere bei als hochriskant eingestuften KI-Systemen;
- **Wirksame Kontrollmechanismen** und **unabhängige Instanzen** mit den notwendigen Befugnissen, Kapazitäten, Zugang zu Informationen und Daten sowie (personellen) Ressourcen für die Durchsetzung der Verordnung;
- **Abhilfe- und Klagemechanismen** für Personen, Gruppen und gemeinnützige Organisationen.



ERLÄUTERUNGEN

Mit der Verordnung zur Regulierung von Künstlicher Intelligenz (KI) in der EU („KI-Verordnung“) hat die Europäische Union die Chance, durch die Verabschiedung des **ersten verbindlichen Gesetzes zur Regulierung von KI weltweit zum Vorbild zu werden**. Die EU hat dadurch die Chance einen **hohen Standard für den Menschenrechtsschutz** zu setzen, der unsere Gesellschaft vor den negativen Auswirkungen durch KI schützt. Der Entwurf der EU-Kommission, auf dessen Grundlage auch Rat und Parlament ihre Positionen aushandeln, unterteilt KI-Systeme in Risikokategorien, nach denen sich ihre Regulierung richtet. Diese reichen von Verboten von KI-Systemen mit nicht vertretbaren Risiken bis hin zu Anforderungen an die Entwicklung oder Transparenzpflichten für KI-Systeme mit hohen oder begrenzten Risiken.

Der aktuelle Entwurf der KI-Verordnung versucht einige **Risiken für die Menschenrechte und für unsere Gesellschaft durch Einschränkungen und Kontrollmechanismen abzufedern, jedoch gehen diese nicht weit genug, um Menschen in der EU und weltweit angemessen zu schützen**. Im Folgenden werden die Änderungen und Ergänzungen zusammengefasst, die aus Sicht von Amnesty International nötig sind, um Betroffenen wie Nutzer*innen einen Umgang mit künstlicher Intelligenz zusichern zu können, der ihre Menschenrechte wahrt.

Um einen umfangreichen Schutz der Menschenrechte zu gewährleisten, müssen KI-Systeme verboten werden, die ein nicht vertretbares Risiko für die Menschenrechte mit sich führen, prozedurale Schutzmechanismen wie Transparenz, verbindliche Folgenabschätzungen, und unabhängige Kontrollen im Gesetz verankert werden und Ausnahmeregelungen eng gefasst sein, um den Schutz nicht zu unterlaufen.

EIN VERBOT VON KI-ANWENDUNGEN MIT NICHT VERTRETBAREN RISIKEN FÜR DIE MENSCHENRECHTE

Amnesty International begrüßt, dass einige Anwendungen von KI unter Titel II des Kommissionsentwurfs zur KI-Verordnung gänzlich verboten werden, z. B. Verhaltensmanipulation, biometrische Überwachung in Echtzeit und „Social Scoring“ für Staaten. Wesentliche Technologien werden jedoch nicht angemessen erfasst, die Verbote greifen nicht weit genug oder die Definitionen sind zu unscharf formuliert.

Das vorgesehene Verbot von **Gesichtserkennung und anderen Systemen zur biometrischen Überwachung im öffentlichen Raum** muss jeglichen Einsatz verbieten, nicht nur in Echtzeit, sondern auch die nachträgliche Auswertung sowie den Einsatz durch private Akteure. Zudem sieht die Verordnung weitreichende Ausnahmen vor, sodass die Notwendigkeit und Verhältnismäßigkeit des Einsatzes in Zweifel gezogen werden kann. Auch der **Einsatz von weiteren biometrischen Daten** für andere Zwecke, wie beispielsweise **Emotionserkennung** oder andere **biometrische Kategorisierungssysteme** muss gänzlich verboten werden. Das **Verbot von Social Scoring** muss dringend auf den Einsatz durch private Akteure ausgeweitet werden. Das Verbot sollte zudem auf schädliche Praktiken von **Social Profiling** erweitert werden.

Im Kommissionsvorschlag werden **manipulative KI-Systeme** sehr ungenau definiert und sind nur dann verboten, sofern sie psychologische oder physische Schäden herbeiführen können oder eine Vulnerabilität einzelner Personen ausnutzen. Dabei sind solche Vulnerabilitäten im Kommissionsvorschlag abschließend definiert und gelistet. Die Konditionierung und Manipulation, die durch KI-Systeme möglich sind, greifen jedoch die Menschenwürde jeder Person an und gefährden unter Umständen die Gedanken- und Meinungsfreiheit. Das Verbot muss deshalb auf jede Art von verdeckt konditionierenden und manipulierenden KI-Systemen ausgeweitet werden, die schädliche Auswirkungen wie eine Verletzung der Menschenrechte verursachen.

MENSCHENRECHTSSCHUTZ IN DEN MITTELPUNKT STELLEN – VOM ENTWICKLUNGSPROZESS BIS ZUM EINSATZ VON KI

Um über die gesamte Lebensspanne von KI den Schutz der Menschenrechte zu gewährleisten, braucht es klare Vorgaben, die eine Durchsetzung der Regeln erst ermöglichen. Als übergeordneter Grundsatz ist hier insbesondere die **Transparenz** zu nennen. Die Verordnung muss eine größtmögliche Transparenz im Zusammenhang mit dem Einsatz von KI-Systemen herstellen, die auch die Verständlichkeit des Einsatzes von KI

fördert. Nur so können die Auswirkungen auf die Rechte und Freiheiten von Einzelnen, der breiten Öffentlichkeit, den Aufsichtsbehörden und unabhängigen Stellen wirksam überprüft, Verantwortlichkeiten festgelegt und Akteure zur Rechenschaft gezogen werden.

Die menschenrechtlichen Verpflichtungen der KI-Verordnung adressieren in erster Linie die Entwickler*innen von KI-Systemen. Während sich viele Gefahren durch die Gestaltung von KI-Systemen ergeben, stellt in vielen Fällen erst ihr Einsatz ein Risiko dar. Darum muss die Verordnung **sowohl für die Entwicklung, als auch für den Einsatz** von KI-Systemen, Verpflichtungen definieren und Rechenschaft fordern. Dazu gehören **menschenrechtliche Sorgfaltspflichten und Folgenabschätzungen** sowie Konformitätsprüfungen bezüglich der Regelungen der KI-Verordnung, sowohl während der Entwicklung als auch beim Einsatz von KI-Systemen.

WIRKSAME DURCHSETZUNG UND KONTROLLE

Damit die KI-Verordnung zum größtmöglichen Schutz der Menschenrechte führt, müssen Mechanismen für eine wirksame Durchsetzung geschaffen werden und Schlupflöcher geschlossen werden, durch die der hohe Standard der Regulierung unterlaufen wird.

Weitgefasste **Ausnahmeregelungen** drohen, den angestrebten Schutz zu unterlaufen. Dazu gehören Ausnahmen für KI-Systeme für den Einsatz im Grenzschutz oder die nationale Sicherheit, sowie KI-Systeme, die bereits vor Inkrafttreten der Verordnung genutzt werden und somit trotz ihrer Kategorisierung als risikobehaftet vom Geltungsbereich der Verordnung ausgenommen sind. Auch sogenannte Testverfahren und „**Sandboxes**“, die für Innovationsräume und Rechtssicherheit bei der frühen Entwicklung von KI-Systemen geschaffen werden sollen, dürfen nicht als Freibrief für die Nichtbeachtung bestehender Regelungen – einschließlich der Datenschutzgrundverordnung – verstanden werden und müssen transparent gestaltet werden.

Ein weiteres Problem stellt die **Kategorisierung** und das damit verbundene Verbot von KI-Systemen mit einem nicht hinnehmbaren Risiko dar. Die vorgesehene Aufzählung ist abschließend und sieht keinen Mechanismus vor, um diese Liste zu ergänzen. Um die **Zukunftsfähigkeit** der KI-Verordnung sicherzustellen, muss ein Verfahren geschaffen werden, um alle Listen und Kategorisierungen von KI-Systemen zu **aktualisieren und hinreichend klare Kriterien für die Kategorisierung** müssen festgelegt werden.

Anders als im Kommissionsvorschlag vorgesehen, dürfen Konformitätsprüfungen von als hochriskant klassifizierten KI-Systemen nicht überwiegend auf der Grundlage interner Kontrollen erfolgen. Vielmehr sollten **Konformitätsprüfungen durch Dritte** stattfinden. Die EU und ihre Mitgliedsstaaten müssen **wirksame Kontrollmechanismen** gewährleisten und eine behördlich angeordnete Einhaltung der Regeln der KI-Verordnung ermöglichen. Da eine hinreichende Expertise und Verständnis von KI-Systemen Voraussetzung für eine wirksame Kontrolle sind, empfiehlt Amnesty die Einrichtung von **unabhängigen Instanzen**, die über die notwendigen Befugnisse, Kapazitäten, Zugang zu Informationen und Daten sowie (personellen) Ressourcen verfügen.

Die KI-Verordnung erteilt weder Einzelpersonen noch Personengruppen **Abwehrrechte** und sieht keinerlei **Abhilfe- oder Klagemechanismen** vor. Amnesty fordert, dass die Verordnung das Recht für Einzelne und Gruppen auf wirksame Rechtsbehelfe verankert sowie einen Klagemechanismus für gemeinnützige Organisationen vorsieht, die u.a. eine Untersuchung von KI-Systemen durch die nationalen Aufsichtsbehörden veranlassen können.