

Gleichstellungsrelevante Aspekte im europäischen Entwurf für ein Gesetz über Künstliche Intelligenz

Hintergrund

Die EU-Kommission hat Anfang des Jahres 2021 einen [Entwurf für ein Europäisches Gesetz \(sog. Verordnung\) über Künstliche Intelligenz](#) ("Artificial Intelligence Act", AIA/KI-VO-E) vorgelegt. Im Rahmen der Europäischen Ratspräsidentschaft unter damals slowenischem Vorsitz wurde Ende 2021 ein erster [Kompromissvorschlag zu einem Teil der Verordnung](#) erarbeitet. Die Weiterarbeit an einem Kompromissvorschlag zu den restlichen Teilen läuft gegenwärtig unter dem derzeit französischen Vorsitz im Rat der Europäischen Union weiter.

Bereits Ende 2022 soll der finale Kompromissvorschlag im Trilog zwischen Rat, Parlament und Kommission verhandelt werden. Am Ende beschließen der Rat der Europäischen Union und das Europäische Parlament die Verordnung, die dann mit einem zweijährigen Übergangszeitraum als nationales Recht auch in Deutschland in Kraft tritt.

Mit dem Gesetzentwurf soll zum einen Rechtssicherheit für Unternehmen, Organisationen und öffentliche Institutionen geschaffen werden, die sogenannte Künstliche-Intelligenz-Systeme entwickeln, vertreiben und einsetzen. Zum anderen sollen Innovationspotentiale für die wirtschaftliche Entwicklung der EU abgesichert werden. Es ist wesentliches Ziel, dass die Entwicklung und Nutzung von KI-Systemen mit den persönlichen, bürgerlichen, politischen, wirtschaftlichen und sozialen Rechten der Menschen in der Europäischen Union im Einklang stehen soll.

Gleichstellungsrelevante Aspekte im Entwurf der EU-Kommission für ein Gesetz über Künstliche Intelligenz

Im risikobasierten Gesetzentwurf sind KI-Systeme entlang von vier Gefährdungsstufen reguliert: (1) Systeme mit unannehmbarem Risiko, (2) Systeme mit hohem Risiko, (3) Systeme mit geringem Risiko, aber speziellen Pflichten zur Offenlegung, und (4) Systeme mit minimalem oder keinem Risiko. Der Kern des Gesetzentwurfs ist das Risikomanagement von Systemen mit hohem Risiko (2). Für diese Künstliche-Intelligenz-Hochrisikosysteme werden umfangreiche Kontroll-, Registrierungs- und Transparenzpflichten festgelegt.

Darüber hinaus werden einige algorithmische Systeme mit unannehmbarem Risiko verboten (1), bspw. bestimmte Systeme zur biometrischen Gesichtserkennung von Personen im öffentlichen Raum oder Systeme zur Verhaltensmanipulation, die zu physischen oder psychischen Schäden führen. Zu letzteren könnten etwa algorithmische Systeme gehören, die Nutzer*innen möglichst lange als Konsument*innen von Werbung in einem sozialen Medium binden sollen und suchterzeugend sind.

Chatbots oder Bild-, Ton- oder Videomanipulationssysteme gehören zur Kategorie (3). Hier muss offengelegt werden, wenn bspw. eine Filmsequenz künstlich erzeugt wurde oder ein Gesprächsgegenüber ein Bot ist.

Die prinzipielle Stoßrichtung des Entwurfs für ein Europäisches Gesetz über Künstliche Intelligenz – insbesondere der risikoadaptierte Regulierungsansatz – ist im Sinne des Dritten Gleichstellungsberichts der Bundesregierung. Eine geschlechtergerechte Digitalisierung setzt die gesetzliche Regulierung algorithmischer Systeme voraus. Hierbei müssen auch die Risiken für die Gleichstellung berücksichtigt und minimiert werden. Im Detail gehen die Handlungsempfehlungen des Gutachtens zum Gleichstellungsbericht jedoch an vielen Punkten über die derzeitigen im Gesetzentwurf vorgesehenen Regelungen hinaus.

Es folgt eine Übersicht der wesentlichen gleichstellungs- bzw. diskriminierungsrelevanten Aspekte des Gesetzentwurfs über Künstliche Intelligenz sowie die entsprechenden Empfehlungen im Dritten Gleichstellungsbericht zu diesen Themen.

Ausgewählte Kernpunkte des Verordnungsvorschlags	Regulierung algorithmischer Systeme im Dritten Gleichstellungsbericht
<p>Was wird unter „Systemen der Künstlichen Intelligenz“ (KI-Systeme) verstanden?</p>	
<p>Der Begriff KI-System wird in Artikel 3 des KI-VO-E definiert und wurde bereits im oben erwähnten nochmal erweitert:</p> <p>Ein KI-System werde mit Hilfe von „Techniken und Konzepten“ wie maschinellern Lernen, logik- und wissensbasierten sowie statistischen Ansätzen entwickelt (Artikel 3, Annex I). Auf deren Basis schließe ein KI-System mittels Lernen, Urteilen oder Modellieren (“learning, reasoning or modelling“) Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen anhand vom Menschen definierter Ziele.</p> <p>Die o. g. „Techniken und Konzepte“ umfassen bspw. mit „Such- und Optimierungsmethoden“ oder „logikbasierten Ansätzen“ auch Systeme, die in der Informatik nicht zwangsläufig unter Künstliche Intelligenz fallen.</p>	<p>Die Sachverständigenkommission des Dritten Gleichstellungsberichts vermeidet den Begriff „Künstliche Intelligenz“, da er den mechanistischen Charakter der Technologie verschleierte (Dritter Gleichstellungsbericht, S. 88). Besonders in der Anpassung der Definition im Kompromissvorschlag zur Verordnung wird deutlich, wie stark vermenschlicht die Definition von KI-Systemen ist, wenn sie als lernend, urteilend und modellierend definiert werden (Artikel 3).</p> <p>Im Gleichstellungsbericht wird, wie auch im Gutachten Datenethikkommission der Bundesregierung, stattdessen der neutralere Begriff „algorithmisches System“ genutzt. Hiermit wird der funktionale Kern eines Computersystems in den Vordergrund gestellt, nämlich Rechenverfahren bzw. Algorithmen. Somit entfällt der fachliche Streit darüber, welche algorithmischen Systeme überhaupt als Künstliche Intelligenz gezählt werden können. Zum Beispiel würde ein regelbasierter Filter, mit dem Bewerbungen nach bestimmten Kategorien wie Alter oder Geschlecht priorisiert werden könnten, je nach Auslegung des Begriffs „logikbasierter Ansatz“ vielleicht gar nicht als KI-System gelten, hätte aber großes Schädigungspotential, wenn er zur Benachteiligung von Personen bestimmten Alters oder Geschlechts führte.</p>

Ausgewählte Kernpunkte des Verordnungsvorschlags	Regulierung algorithmischer Systeme im Dritten Gleichstellungsbericht
<p>An wen richtet sich die Verordnung und wen schützt sie?</p>	
<p>Der Vorschlag richtet sich vorrangig an Entwickler*innen von KI-Systemen, bezeichnet als „Anbieter“ bzw. „Provider“. Das könnte etwa eine Firma sein, die ein KI-Personalauswahlsystem entwickelt. Außerdem werden „Nutzer“ bzw. „User“ adressiert. Damit sind nicht von KI betroffene Personen gemeint, wie bspw. Schüler*innen, deren Abiturnoten ein KI-System berechnet, sondern die Schule, die ein solches KI-System einsetzt.</p> <p>Endnutzer*innen bzw. die betroffenen Personen, über die mit einem KI-System etwas entschieden wird, werden zwar mittelbar als potentiell geschädigte natürliche Personen durch die Regulierung geschützt, aber nicht als eigens definierte*r Akteur*in berücksichtigt.</p>	<p>Rechte der vom Einsatz algorithmischer Systeme betroffenen Personen fehlen im KI-Verordnungsvorschlag fast vollständig. Die betroffenen Personen verfügen jedoch in der Regel kaum über die Fähigkeit zur Überprüfung oder Kontrolle derartiger Systeme. Der Dritte Gleichstellungsbericht empfiehlt hier u. a., datenschutzrechtliche Auskunftsrechte (Art. 15 DSGVO) zu konkretisieren sowie Informations- und Offenlegungspflichten zu schaffen, mit denen betroffene Personen umfassend über den Einsatz und die Funktionsweise algorithmischer Systeme informiert werden.</p> <p>Zudem müssen etwa Arbeitnehmer*innen spezielle Rechte eingeräumt werden, eine Prüfung algorithmischer Systeme zu verlangen.</p> <p>Nicht zuletzt müssen institutionelle Vorkehrungen gestärkt werden, um zu vermeiden, dass betroffene Personen ihre Rechte individuell einklagen müssen. Dazu sind Antidiskriminierungsstellen mit Kompetenzen und Befugnissen auszustatten, konkret mit einem Verbandsklagerecht, und ein pauschalisierter Schadensersatz ist vorzusehen. Zwar beziehen sich diese Handlungsempfehlungen im Gleichstellungsbericht vor allem auf die Hochrisikosysteme in der Personalauswahl, können aber auf Hochrisiko-Systeme verallgemeinert werden (Dritter Gleichstellungsbericht, S. 169ff.).</p>

Ausgewählte Kernpunkte des Verordnungsvorschlags	Regulierung algorithmischer Systeme im Dritten Gleichstellungsbericht
In welche Risikostufen werden Systeme Künstlicher Intelligenz eingeteilt und was bedeuten diese?	
<p>In der Verordnung werden KI-Systeme in verschiedene Risikostufen eingeordnet (siehe Abschnitt „Hintergrund“). Demnach werden Systeme mit unannehmbarem Risiko verboten (Artikel 5). Das umfasst beispielsweise ein mit diversen Ausnahmen behaftetes Verbot biometrischer Echtzeit-Identifikation von Personen im öffentlichen Raum, wie dies etwa bei Videoüberwachung mit sofortiger Gesichtserkennung der Fall wäre.</p> <p>Zu einer weiteren Stufe gehören Hochrisiko-KI-Systeme (Artikel 6), für die umfassende Kontroll-, Test-, Dokumentations-, Konformitäts- sowie Registrierungsanforderungen gelten. Eine der acht hier zugehörigen Gruppen sind zum Beispiel Systeme im Bereich „Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit“ (z. B. Stellenanzeigen, Sichtung von Bewerbungen).</p> <p>Schließlich werden KI-Systeme mit geringem oder minimalem Risiko in einer dritten Stufe erfasst. Dazu zählen bspw. Programme zur Bildmanipulation („Deepfakes“), die fiktive Filmaufnahmen von Personen täuschend echt erscheinen lassen. Solche künstlichen Veränderungen sollen gekennzeichnet werden (Artikel 52).</p>	<p>Im Gleichstellungsbericht wird eine Risikoprüfung von algorithmischen Systemen empfohlen, die sich an den von der Datenethikkommission empfohlenen Kritikalitäts-Stufen orientiert (Dritter Gleichstellungsbericht, S. 168). Es werden verbindliche Kontroll-, Informations- und Transparenzpflichten für algorithmische Systeme in sensiblen Bereichen wie Personalwesen oder Soziale Medien eingefordert. Allgemein wird eine generelle Risikoprüfung und Technikfolgenabschätzung algorithmischer Systeme hinsichtlich Datenschutz- und Diskriminierungsgefahren empfohlen.</p> <p>Im Gegensatz zum Entwurf der Verordnung werden alle algorithmengestützten Personalauswahlsysteme und biometrische Systeme insgesamt als riskanter eingeschätzt. Daher wird auch für diese Bereiche ein absolutes Verbot nicht ausgeschlossen.</p>

Ausgewählte Kernpunkte des Verordnungsvorschlags	Regulierung algorithmischer Systeme im Dritten Gleichstellungsbericht
<p>Wie wird Diskriminierung bei der Regulierung algorithmischer Systeme berücksichtigt?</p>	
<p>Der Entwurf thematisiert Diskriminierungsrisiken von KI-Systemen nur in den Erwägungsgründen (ErwG): Zum Beispiel wird erläutert, dass KI-Systeme im Personalbereich „historische Diskriminierungsmuster fortschreiben, beispielsweise gegenüber Frauen, bestimmten Altersgruppen und Menschen mit Behinderungen oder Personen mit einer bestimmten rassischen oder ethnischen Herkunft oder sexuellen Ausrichtung“ (ErwG 36).</p> <p>Die im Verordnungsvorschlag verankerte Anforderung an Provider von Hochrisiko-Systemen, ihre Trainings-, Validierungs- und Testdatensätze im Hinblick auf mögliche Verzerrungen (Bias) zu untersuchen (Artikel 10), kann einige solcher Diskriminierungsmuster minimieren und entspricht vergleichbaren Empfehlungen im Dritten Gleichstellungsbericht.</p>	<p>Oft fußen Diskriminierungen bestimmter Personengruppen durch algorithmische Systeme auf einer Datenbasis, die nicht alle späteren End-Nutzer*innen repräsentiert oder benachteiligende Geschlechterverhältnisse, etwa der Arbeitswelt, widerspiegelt.</p> <p>Sowohl der Dritte Gleichstellungsbericht als auch das Gutachten der Datenethikkommission empfehlen beim Einsatz algorithmischer Systeme mit hohem Diskriminierungsrisiko konkrete Verfahrensregelungen mit Bezug zu Diskriminierungen. Für die entsprechenden Verpflichtungen in der Datenschutzgrundverordnung (DSGVO) oder die nationalen Regelungen zum Schutz vor Diskriminierung, wie das Allgemeine Gleichbehandlungsgesetz (AGG), werden im Gleichstellungsbericht Erweiterungen vorgeschlagen, die insbesondere den Arbeitnehmer*innenschutz vor Diskriminierung durch algorithmische Systeme erhöhen (Dritter Gleichstellungsbericht, S. 168 ff.).</p> <p>Außerdem werden gleichstellungsorientierte Gesetzes-, Technik- sowie Datenschutzfolgenabschätzungen empfohlen (Dritter Gleichstellungsbericht, S. 229 ff.). Diese Verfahren finden sich in der KI-Verordnung nicht in den Anforderungen an die Anbieter*innen wieder.</p>

Ausgewählte Kernpunkte des Verordnungsvorschlags	Regulierung algorithmischer Systeme im Dritten Gleichstellungsbericht
<p>Welche Prüf- und Transparenzvorschriften werden für Hochrisiko-KI-Systeme festgeschrieben?</p>	
<p>KI-Systeme mit hohem Risiko müssen spezielle Verfahrensorderungen erfüllen. Dazu gehört z. B. die Einführung eines Risikomanagementsystems mit konkreten Maßnahmen zur Risikobewältigung. Systemtests müssen während des gesamten Entwicklungsprozesses und in jedem Fall vor dem Einsatz des Systems erfolgen und kontinuierlich fortgesetzt werden (Artikel 9). Artikel 10 enthält zudem Anforderungen an in das KI-System bspw. zu Optimierungs- und Trainingszwecken eingeegebene Datensätze. In den Artikeln 12, 13 und 14 sind Dokumentations-, Aufzeichnungspflichten und Transparenzpflichten formuliert. Die Anbieter*innen von Hochrisikosystemen müssen ihre Systeme zudem einem Konformitätsbewertungsverfahren nach Artikel 43 unterziehen. Bis auf Anwendungen im Bereich „biometrische Systeme“ können sie diese Konformitätsbewertung selbst vornehmen.</p> <p>Die Artikel 16 bis 29 enthalten weitere Pflichten der Provider und User, aber auch z. B. der Händler*innen oder Importeure. Daran anschließend werden in den Artikeln 30 bis 39 Kriterien für Behörden oder Akkreditierungsstellen festgelegt, die für die Konformitätsbewertungen für Hochrisiko-KI-Systeme zuständig sind. Die Anforderungen an Konformitätsbewertungen und Registrierung werden schließlich in den Artikeln 40 bis 51 festgelegt.</p> <p>Ein bedeutendes Instrument zum Herstellen öffentlicher Transparenz ist die Schaffung einer EU-Datenbank für eigenständige Hochrisiko-KI-Systeme (Artikel 60).</p> <p>Nach Artikel 69 wird für Systeme mit minimalem Risiko angeregt, Verhaltenskodizes aufzustellen.</p>	<p>Aus Gleichstellungssicht ist die Integration von prüfbaren Kriterien geschlechtergerechter und partizipativer Technikgestaltung wichtig, wie sie etwa das im Dritten Gleichstellungsbericht empfohlene Gender-Extended-Research-and-Development-Modell (GERD) enthält (Dritter Gleichstellungsbericht, S. 103ff.). Solche Kriterien werden im KI-Verordnungsentwurf nicht angeführt.</p> <p>Ähnlich der Datenethikkommission empfiehlt die Sachverständigenkommission im Dritten Gleichstellungsbericht, diskriminierende Effekte durch Verfahren wie Risikofolgenabschätzung und Output-Analysen zu identifizieren, die u. a. die Verletzung von Diskriminierungsverboten einbeziehen (Dritter Gleichstellungsbericht, S. 219).</p> <p>Die Sachverständigenkommission fordert außerdem eine umfassende externe Kontrolle algorithmischer Systeme sicherzustellen, unabhängig davon ob algorithmische System durch private oder staatliche Akteur*innen eingesetzt werden. Hierbei muss insbesondere ein Augenmerk auf Diskriminierungsrisiken gelegt werden. Zudem muss gewährleistet werden, dass die kontrollierenden Stellen auch über die notwendigen Kompetenzen zu Diskriminierungsrisiken verfügen (Dritter Gleichstellungsbericht, S. 169; 219). Eine Kontrolle durch Dritte ist aber für Hochrisiko-Systeme im KI-Verordnungsvorschlag nur in sehr wenigen Fällen bei biometrischen Systemen vorgesehen, sonst reicht eine anforderungskonforme interne Kontrolle.</p>

Ausgewählte Kernpunkte des Verordnungsvorschlags	Regulierung algorithmischer Systeme im Dritten Gleichstellungsbericht
Woran orientiert sich die Risikoeinstufung?	
<p>Die Risikoeinstufung der Systeme orientiert sich an ihrer Zweckbestimmung. Die Prüfmaßnahmen zielen insgesamt darauf ab, Risiken innerhalb der „bestimmungsgemäßen Verwendung“ (“intended purpose“) eines Hochrisiko-KI-Systems offenzulegen und zu minimieren. „Vernünftigerweise vorhersehbare Fehlanwendungen“ (“reasonably foreseeable misuse“) sollen im Rahmen des Risikomanagements abgeschätzt werden. Die User müssen die Systeme gemäß ihrer Zweckbestimmung nutzen und bei bestimmten Fehlfunktionen die Nutzung aussetzen (Artikel 29). Schwerwiegende Vorfälle müssen den Anbietern bzw. den Marktüberwachungsbehörden der Mitgliedsstaaten gemeldet werden (Artikel 29 bzw. 62).</p> <p>Der og. Kompromissvorschlag des Rates sieht zudem vor, universell einsetzbare KI-Systeme nicht den Regularien zu unterwerfen (Artikel 52a). Solche Systeme sind etwa noch nicht auf konkrete Daten trainierte Maschinenlernalgorithmen. Erst wenn sie für einen bestimmten Zweck genutzt werden, unterlägen sie den Regeln.</p>	<p>Die Sachverständigenkommission unterstreicht in ihrem Gutachten die Notwendigkeit, algorithmische Systeme aus einer soziotechnischen Perspektive zu betrachten (Dritter Gleichstellungsbericht, S. 87). Das bedeutet, sie eingebettet in ihren gesellschaftlichen Kontext zu bewerten und ihre Wechselwirkung mit diesem Kontext zu beachten.</p> <p>Das heißt auch, die potentiellen Einsatzmöglichkeiten zu antizipieren, die auch nicht bestimmungsgemäß sein können, und Wechselwirkungen der Systeme mit ihrer Umwelt und der sie benutzenden Menschen in die Risikobewertung sowohl vor der Inbetriebnahme als auch während des laufenden Betriebs einzubeziehen. Hier greifen bereits oben erwähnte gleichstellungsorientierte Technikfolgenabschätzungen, die explizit Diskriminierungsrisiken abschätzen.</p>
Welche Anwendungsbereiche werden von der Verordnung ausgenommen?	
<p>Systeme, „die ausschließlich für militärische Zwecke entwickelt oder verwendet werden“, sind von der Verordnung ausgenommen (Artikel 2). Der o. g. Kompromissvorschlag hat zudem alle KI-Systeme im Bereich der Nationalen Sicherheit sowie ausschließlich im Bereich Forschung und Entwicklung eingesetzte Systeme aus dem Anwendungsbereich der Verordnung entfernt.</p>	

Ausblick

In den anstehenden parlamentarischen Debatten und im Trilog zwischen Parlament, Kommission und Rat um die KI-Verordnung der EU werden Fragen der **Antidiskriminierung und Geschlechtergerechtigkeit** nur dann eine angemessene Rolle spielen, wenn **Akteur*innen mit Sachverstand in Gleichstellungsfragen** ebenfalls einbezogen werden. Hierzu gehören zum Beispiel einschlägige Vertreter*innen der EU-Institutionen, die sich mit Gleichstellung und Antidiskriminierung beschäftigen, wie beispielsweise der Ausschuss für die Rechte der Frauen und die Gleichstellung der Geschlechter des Europäischen Parlaments (FEMM), das Institut für Gleichstellungsfragen (EIGE) oder die Rechte-Agentur der EU (FRA). Gleichermaßen müssen die Belange potentiell durch KI diskriminierter Betroffener, die durch zivilgesellschaftliche Gruppen, Arbeitnehmer*innenvertretungen, Behindertenverbände, Interessengruppen von Senior*innen und Erwerbslosen usw. repräsentiert sind, strukturell und finanziell ermöglicht und eingefordert werden.

Es ist im Sinne des Gleichstellungsberichts, wenn am Ende eine Verordnung entstanden sein wird, die algorithmische Systeme risiko- und kontextbasiert reguliert. Eine solche Verordnung bezieht die sozio-technische Perspektive ein und unterwirft umfassend und nachhaltig sämtliche algorithmischen Systeme mit Schädigungspotential für die Gesellschaft gesetzlichen Regeln.