

1552nd meeting, 4 March 2026

4 Human rights

4.2 Steering Committee on Anti-discrimination, Diversity and Inclusion (CDADI)

a. Begründung zur Empfehlung CM/Rec(2026)1 des Ministerkomitees an die Mitgliedstaaten zu Gleichheit und künstlicher Intelligenz

Präambel

1. Ziel dieser Empfehlung ist es, den Mitgliedstaaten klare und konkrete Leitlinien an die Hand zu geben, wie sie die Gleichheit, einschließlich der Gleichheit der Geschlechter, fördern und wie sie Diskriminierung im Rahmen der Aktivitäten während des gesamten Lebenszyklus von Systemen künstlicher Intelligenz („KI“) verhindern und dagegen vorgehen können. Zu diesem Zweck stützt sich die Empfehlung auf bestehende Rechtsinstrumente, die eine Rechtsgrundlage für die Förderung der Gleichheit und das Verbot von Diskriminierung bieten, von denen viele nicht speziell auf den Kontext von KI-Systemen ausgerichtet sind,¹ und sucht nach Möglichkeiten, die Anwendung dieser Rechtsinstrumente auf die Herausforderungen und Chancen in diesem Bereich zu optimieren. Die Empfehlung zielt darauf ab, den bestehenden Rahmen so umzusetzen, dass Innovationen gefördert werden, um die Gleichheit zu stärken sowie Diskriminierung im Bereich von KI-Systemen zu verhindern und zu bekämpfen.

2. Die Empfehlung basiert sowohl auf verbindlichen als auch auf unverbindlichen Standards. Insbesondere entspricht sie dem bestehenden Rechtsrahmen für KI-Systeme, einschließlich des [Rahmenübereinkommens des Europarats über künstliche Intelligenz sowie Menschenrechte, Demokratie und Rechtsstaatlichkeit](#) (SEV 225) („das Rahmenübereinkommen“)² sowie der einschlägigen Verordnung der Europäischen Union („die KI-Verordnung“).³ Gleichheit, einschließlich der Gleichheit der Geschlechter, und Nichtdiskriminierung sind im Rahmenübereinkommen auf der Ebene der allgemeinen Grundsätze neben anderen Menschenrechtsnormen verankert. In der Begründung zum Rahmenübereinkommen wird betont, wie wichtig es ist, dass die Mitgliedstaaten bei ihrer Bewertung der Risiken und potenziellen Auswirkungen von KI-Systemen auf die Menschenrechte, die Demokratie und die Rechtsstaatlichkeit der Gleichheit, einschließlich der Gleichheit der Geschlechter, sowie der Nichtdiskriminierung gebührend Rechnung tragen und geeignete Maßnahmen zur Bewältigung dieser Risiken und potenziellen Auswirkungen ergreifen oder aufrechterhalten. In diesem Zusammenhang wird darauf hingewiesen, dass die Mitgliedstaaten Machtungleichgewichte berücksichtigen sollten, die bestehende Ungleichheiten und gesellschaftliche Auswirkungen verschärfen könnten. In dem erläuternden Bericht zum Rahmenübereinkommen wird zudem auf das reale und gut dokumentierte Risiko von Verzerrungen hingewiesen, die eine rechtswidrige Diskriminierung darstellen können und sich aus den Aktivitäten während des Lebenszyklus von KI-Systemen ergeben. Auch wenn die Empfehlung die Terminologie des Rahmenübereinkommens übernimmt und darauf abzielt, die Kohärenz mit dessen Geltungsbereich und

¹ Einschließlich der [Konvention zum Schutz der Menschenrechte und Grundfreiheiten](#) (SEV Nr. 005), Artikel 14; [Protokoll Nr. 12 zur Konvention zum Schutz der Menschenrechte und Grundfreiheiten](#) (SEV Nr. 177), Artikel 1; [Europäische Sozialcharta](#) (SEV Nr. 035) und Europäische Sozialcharta (revidiert) (SEV Nr. 163), Artikel 20, Artikel E; das [Übereinkommen zur Verhütung und Bekämpfung von Gewalt gegen Frauen und häuslicher Gewalt](#) („Istanbul-Konvention“) (SEV Nr. 210), Artikel 4; das [Rahmenübereinkommen zum Schutz nationaler Minderheiten](#) (SEV Nr. 157), Artikel 4, 6 Absatz 2, und 9; die [Europäische Charta der Regional- oder Minderheitensprachen](#) (SEV Nr. 148), Artikel 7 Absatz 2; sowie das [Übereinkommen zum Schutz des Menschen bei der automatischen Verarbeitung personenbezogener Daten](#) (SEV Nr. 108) und das [Protokoll zur Änderung des Übereinkommens zum Schutz des Menschen bei der automatischen Verarbeitung personenbezogener Daten](#) (SEV Nr. 223).

² Am 5. September 2024 in Vilnius, Litauen, zur Unterzeichnung aufgelegt.

³ Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 ([Verordnung über künstliche Intelligenz](#)), („KI-Verordnung“).

Inhalt zu wahren, stellt sie dennoch kein Durchführungsinstrument dar und sollte auch nicht als solches ausgelegt werden. Die Verweise auf das Rahmenübereinkommen und dessen erläuternder Bericht in dieser Begründung beziehen sich ausschließlich auf dieses Dokument und haben keinen Einfluss auf die Auslegung des Rahmenübereinkommens; dies ist auch nicht beabsichtigt.

3. In der Empfehlung wird zudem die Bedeutung anderer einschlägiger globaler Standards und Normen für die Förderung der Gleichheit und die Verhinderung von Diskriminierung gewürdigt, insbesondere das [Internationale Übereinkommen zur Beseitigung jeder Form von Rassendiskriminierung](#) (ICERD), das [Übereinkommen zur Beseitigung jeder Form von Diskriminierung gegenüber Frauen](#) (CEDAW), das [Übereinkommen über die Rechte des Kindes](#) (CRC) sowie das [Übereinkommen über die Rechte von Menschen mit Behinderungen](#) (CRPD), zusätzlich zu denjenigen, auf die im Erläuternden Bericht zum Rahmenübereinkommen Bezug genommen wird.⁴

4. Darüber hinaus gibt es eine umfangreiche Sammlung nicht bindender Standards des Europarats, die für die Prävention und Bekämpfung von Diskriminierung sowie für die Förderung der Gleichheit relevant sind, darunter insbesondere im Bereich der Gleichheit der Geschlechter. In diesem Zusammenhang fordert die [Empfehlung CM/Rec\(2019\)1](#) zur Prävention und Bekämpfung von Sexismus die Mitgliedstaaten auf, „eine Gleichheitsperspektive in alle politischen Maßnahmen, Programme und Forschungsvorhaben im Zusammenhang mit Künstlicher Intelligenz zu integrieren, um die potenziellen Risiken zu vermeiden, dass diese Technologie Sexismus und Geschlechterstereotypen zementiert, und zu prüfen, wie Künstliche Intelligenz dazu beitragen könnte, geschlechtsspezifische Ungleichheiten zu verringern und Sexismus zu beseitigen.“⁵ Die Empfehlung CM/Rec(2026)2 zur Bekämpfung von durch Technologie begünstigter Gewalt gegen Frauen und Mädchen befasst sich mit den spezifischen Risiken, die von neuen Technologien, einschließlich KI-Systemen, ausgehen. Zu den Standards, die für die Förderung der Gleichheit im Allgemeinen relevant sind und in die Ausarbeitung der Empfehlung eingeflossen sind, gehört insbesondere auch die [Empfehlung CM/Rec\(2022\)16](#) des Ministerkomitees an die Mitgliedstaaten zur Bekämpfung von Hassrede, in der festgestellt wird, dass gesetzliche und politische Maßnahmen zur Verhinderung und Bekämpfung von Hassrede im Internet regelmäßig überprüft werden sollten, um der raschen Entwicklung der Technologie und der Online-Dienste sowie, im weiteren Sinne, der digitalen Technologien und deren Einfluss auf den Informations- und Kommunikationsfluss in den heutigen demokratischen Gesellschaften Rechnung zu tragen. Dazu gehört auch die [Empfehlung CM/Rec\(2024\)4 des Ministerkomitees an die Mitgliedstaaten zur Bekämpfung von Hassverbrechen](#), in der festgestellt wird, dass die Prävention und Bekämpfung von Hassverbrechen, unabhängig davon, ob sie online oder offline begangen werden, einen ganzheitlichen und vielschichtigen Ansatz erfordert, der die in öffentlichen Einrichtungen und zivilgesellschaftlichen Organisationen Tätigen sowie diejenigen einbezieht, die Gruppen angehören oder mit Gruppen zusammenarbeiten, die Gefahr laufen, Ziel von Hassverbrechen zu werden. [Die Empfehlung CM/Rec\(2010\)5 des Ministerkomitees an die Mitgliedstaaten zu Maßnahmen zur Bekämpfung von Diskriminierung aufgrund der sexuellen Orientierung oder der Geschlechtsidentität](#) fordert die Mitgliedstaaten auf, geeignete Maßnahmen zu ergreifen, um alle Äußerungen, einschließlich jener in den Medien und im Internet, zu bekämpfen, die begründeterweise als geeignet angesehen werden können, Hass oder andere Formen der Diskriminierung gegenüber lesbischen, schwulen, bisexuellen und transgender Personen zu schüren, zu verbreiten oder zu fördern.

5. Die Empfehlung baut auf den Verpflichtungen auf, die die Staats- und Regierungschefs zum Abschluss des 4. Gipfeltreffens („[die Erklärung von Reykjavík](#)“)⁶ zum Ausdruck gebracht haben, insbesondere auf der Notwendigkeit, Gleichheit zu gewährleisten und jegliche Form von Diskriminierung zu bekämpfen, eine Jugendperspektive einzubeziehen und die Bemühungen um inklusive Gesellschaften ohne Marginalisierung, Ausgrenzung, Rassismus oder Intoleranz zu verstärken sowie die uneingeschränkte, gleichberechtigte und wirksame Teilhabe von Frauen an öffentlichen und privaten Entscheidungsprozessen sicherzustellen. In der Erklärung von Reykjavík werden die positiven Auswirkungen und Chancen neuer und sich entwickelnder digitaler Technologien gewürdigt; zugleich wird die Notwendigkeit anerkannt, den Risiken negativer Folgen ihrer Nutzung für die Menschenrechte, die Demokratie und die Rechtsstaatlichkeit entgegenzuwirken, und die Verpflichtung eingegangen, dem Europarat eine führende Rolle bei der Entwicklung von Standards im digitalen Zeitalter zu sichern, um die Menschenrechte sowohl online als auch offline zu schützen.

6. In der Präambel der vorliegenden Empfehlung wird betont, wie wichtig es ist, inklusive Gesellschaften zu schaffen, in denen es keine Marginalisierung, Ausgrenzung, keinen Sexismus, keinen Rassismus, keine LGBTI-Phobie und keine anderen Formen der Intoleranz, einschließlich Altersdiskriminierung und Diskriminierung von

⁴ Erläuternder Bericht zum Rahmenübereinkommen, Absatz 72–73, einschließlich: Titel III der [Charta der Grundrechte der Europäischen Union](#); Verträge der Europäischen Union (beispielsweise [Vertrag über die Europäische Union](#), Artikel 2, [Vertrag über die Arbeitsweise der Europäischen Union](#), Artikel 10), sekundäres EU-Recht und die einschlägige Rechtsprechung des Gerichtshofs der Europäischen Union; [Internationaler Pakt über bürgerliche und politische Rechte](#) (ICCPR), Artikel 2, 24 und 26, sowie [Internationaler Pakt über wirtschaftliche, soziale und kulturelle Rechte](#) (ICESCR), Artikel 2, 3 und 7.

⁵ [Empfehlung CM/Rec\(2019\)1](#) des Ministerkomitees an die Mitgliedstaaten zur Prävention und Bekämpfung von Sexismus, II.B.7., S. 19.

⁶ „[Gemeinsam für unsere Werte – Erklärung von Reykjavík](#)“, verabschiedet auf dem 4. Gipfeltreffen der Staats- und Regierungschefs des Europarates (Reykjavík, 16.–17. Mai 2023), S. 7.

Menschen mit Behinderungen, gibt. Diese Liste erhebt keinen Anspruch auf Vollständigkeit und umfasst auch spezifische Formen des Rassismus, wie beispielsweise Antiziganismus, sowie Antisemitismus, Islamfeindlichkeit und andere Formen religiöser Intoleranz. Die Empfehlung stützt sich auf die Allgemeinen politischen Empfehlungen (GPR) der Europäischen Kommission gegen Rassismus und Intoleranz (ECRI), insbesondere auf die GPR Nr. 17 zur Prävention und Bekämpfung von Intoleranz und Diskriminierung gegenüber LGBTI-Personen, in der unter anderem betont wird, dass technologische Fortschritte zwar das Potenzial haben, die Wahrnehmung der Menschenrechte und die gesellschaftliche Teilhabe von LGBTI-Personen zu verbessern, bestimmte Technologien wie künstliche Intelligenz, einschließlich Gesichtserkennung und algorithmische Systeme, jedoch auch dazu genutzt werden können, die Privatsphäre, die Sicherheit und die Gleichheit von LGBTI-Personen zu untergraben.⁷ Die Empfehlung bezieht sich auf Personen und Gruppen, die von Diskriminierung bedroht sind, um sowohl diejenigen einzubeziehen, die bereits Diskriminierung erfahren haben, als auch diejenigen, die in Zukunft davon betroffen sein könnten.

7. Im Einklang mit der Erklärung von Reykjavík verfolgt diese Empfehlung einen intersektionalen Ansatz. Diese Herangehensweise berücksichtigt, dass sich verschiedene Diskriminierungsgründe überschneiden können, was zu einzigartigen Lebenserfahrungen und besonderen Gefährdungen führt. Ein intersektionaler Ansatz ermöglicht es, die Wechselwirkungen zwischen verschiedenen persönlichen Merkmalen oder Statusmerkmalen sowie die daraus resultierenden vielschichtigen Formen der Diskriminierung zu verstehen, zu berücksichtigen und anzugehen. Somit kann dieser Ansatz Einblicke in die komplexeren Formen von Diskriminierung, Ausgrenzung und Gewalt gewähren, denen Einzelpersonen ausgesetzt sein können, und somit Diskriminierung aufdecken, die andernfalls möglicherweise unbemerkt bliebe. Jeder kann von intersektionaler Diskriminierung betroffen sein, doch bestimmte Gruppen sind ihr besonders ausgesetzt und würden daher besonders von einem intersektionalen Ansatz profitieren.⁸ Während der Begriff der intersektionalen Diskriminierung sowohl im Europarat und seinen Mitgliedstaaten als auch auf internationaler Ebene zunehmend Anerkennung findet, beispielsweise in der Richtlinie zur Lohntransparenz und den Richtlinien über Standards für Gleichheitsstellen der Europäischen Union⁹, können einige Mitgliedstaaten auch Begriffe wie „mehrfache“ oder „kumulative“ Diskriminierung verwenden, um ein ähnliches Anliegen zum Ausdruck zu bringen. Da viele Mitgliedstaaten den Begriff „Mehrfachdiskriminierung“ in ihren nationalen Rechtsvorschriften, politischen Strategien und Praktiken verwenden, wird in der Empfehlung daher die Terminologie „Mehrfach- und intersektionale Diskriminierung“ verwendet.

8. Die Präambel bezieht sich auf Stellvertreter für persönliche Merkmale oder Status. Unter „Stellvertreter-Diskriminierung“ versteht man einen Begriff, der in Absatz 2 der Empfehlung und im entsprechenden Absatz 19 der Begründung näher definiert und erläutert wird.

9. In der Empfehlung wird festgestellt, dass eine Folge der Entwicklung der künstlichen Intelligenz und ihrer weit verbreiteten Nutzung eine Zunahme der durch Technologie ermöglichten Gewalt gegen Frauen und Mädchen ist; daher wird auf die Empfehlung CM/Rec(2026)2 zur Rechenschaftspflicht bei durch Technologie ermöglichter Gewalt gegen Frauen und Mädchen sowie auf das Übereinkommen des Europarats zur Verhütung und Bekämpfung von Gewalt gegen Frauen und häuslicher Gewalt (Istanbul-Konvention, SEV Nr. 210) verwiesen. In der Empfehlung werden zudem die besonderen Bedürfnisse und die besondere Schutzbedürftigkeit von Kindern sowie die damit verbundenen spezifischen Auswirkungen der KI auf sie hervorgehoben. In der Empfehlung wird die Bedeutung des Übereinkommens des Europarats zum Schutz von Kindern vor sexueller Ausbeutung und sexuellem Missbrauch (Lanzarote-Übereinkommen, SEV Nr. 201) gewürdigt, das einen Rahmen für den Schutz von Kindern vor sexueller Ausbeutung und sexuellem Missbrauch bildet, einschließlich Missbrauch, der unter Einsatz von Informations- und Kommunikationstechnologien

⁷ [ECRI-Allgemeine Politikempfehlung Nr. 17 zur Prävention und Bekämpfung von Intoleranz und Diskriminierung gegenüber LGBTI-Personen](#), Absatz 24.

⁸ Siehe beispielsweise [die Strategie des Europarats zur Gleichstellung der Geschlechter 2024–2029](#); den Modellrahmen des Expertenausschusses für interkulturelle Inklusion (ADI-INT) [für eine Strategie zur interkulturellen Integration auf nationaler Ebene](#); die Allgemeine politische Empfehlung (GPR) Nr. 14 der Europäischen Kommission gegen Rassismus und Intoleranz (ECRI) [zur Bekämpfung von Rassismus und rassistischer Diskriminierung im Arbeitsleben](#) sowie deren Begründung; [die überarbeitete Allgemeine Politikempfehlung Nr. 5 der ECRI zur Prävention und Bekämpfung von antimuslimischem Rassismus und Diskriminierung](#); [„COVID-19: Eine Analyse der Dimensionen von Antidiskriminierung, Vielfalt und Inklusion in den Mitgliedstaaten des Europarats“](#) sowie [die Leitlinien des Ministerkomitees des Europarats zur Wahrung der Gleichstellung und zum Schutz vor Diskriminierung und Hass während der COVID-19-Pandemie und ähnlicher Krisen in der Zukunft](#).

⁹ [Richtlinie \(EU\) 2023/970](#) des Europäischen Parlaments und des Rates vom 10. Mai 2023 zur Stärkung der Anwendung des Grundsatzes des gleichen Entgelts für gleiche oder gleichwertige Arbeit von Männern und Frauen durch Entgelttransparenz und Durchsetzungsmechanismen; [Richtlinie \(EU\) 2024/1499](#) des Rates vom 7. Mai 2024 über Standards für Gleichbehandlungsstellen im Bereich der Gleichbehandlung von Personen unabhängig von ihrer Rasse oder ethnischen Herkunft, der Gleichbehandlung in Beschäftigung und Beruf unabhängig von ihrer Religion oder Weltanschauung, einer Behinderung, des Alters oder der sexuellen Ausrichtung sowie der Gleichbehandlung von Frauen und Männern in Fragen der sozialen Sicherheit und beim Zugang zu sowie bei der Bereitstellung von Waren und Dienstleistungen sowie zur Änderung der Richtlinien 2000/43/EG und 2004/113/EG und [der Richtlinie \(EU\) 2024/1500](#) des Europäischen Parlaments und des Rates vom 14. Mai 2024 über Standards für Gleichbehandlungsstellen im Bereich der Gleichbehandlung und der Chancengleichheit von Frauen und Männern in Beschäftigung und Beruf sowie zur Änderung der Richtlinien 2006/54/EG und 2010/41/EU.

begangen wird. In der Empfehlung wird zudem darauf hingewiesen, dass künstliche Intelligenz besondere Auswirkungen auf das Wohlergehen, die Arbeitssituation, die Bildung und die Bürgerrechte junger Menschen haben kann.¹⁰

10. Die Empfehlung stützt sich auf die Ergebnisse der Studie über die Auswirkungen von Systemen künstlicher Intelligenz, ihr Potenzial zur Förderung der Gleichheit und die Risiken, die sie im Zusammenhang mit der Nichtdiskriminierung mit sich bringen können („die Studie“).¹¹ Die Studie kommt insbesondere zu dem Schluss, dass technologische Lösungen allein nicht ausreichen, um Diskriminierung zu bekämpfen und die Gleichheit sowie die Gleichheit der Geschlechter zu fördern, und empfiehlt, die substanzielle Gleichheit und die „Anti-Discrimination by Design“ in den Mittelpunkt der rechtlichen Regulierung der Entwicklung und des Einsatzes von KI zu stellen. Aus diesem Grund verfolgt die Empfehlung einen umfassenderen Ansatz und enthält Leitlinien zu sozialen und politischen Veränderungen, um eine substanzielle Gleichheit zu erreichen.¹² Die Empfehlung sieht einen Multi-Stakeholder-Ansatz vor,¹³ mit maßgeschneiderten Empfehlungen für ein breites Spektrum von Akteuren.¹⁴ Insbesondere gilt die Empfehlung für Aktivitäten im Lebenszyklus von KI-Systemen, die sowohl von öffentlichen als auch von privaten Akteuren durchgeführt werden. Akteure des privaten Sektors sollten die Menschenrechte ihrer Kunden und aller betroffenen Parteien achten. So fordert beispielsweise die [Empfehlung CM/Rec\(2016\)3 des Ministerkomitees an die Mitgliedstaaten zu Menschenrechten und Wirtschaft](#) die Mitgliedstaaten auf, Maßnahmen zu ergreifen, um die Verantwortung von Unternehmen für den Schutz der Menschenrechte zu gewährleisten, und die [Leitprinzipien der Vereinten Nationen zu Wirtschaft und Menschenrechten](#) sehen die Verantwortung von Unternehmen für den Schutz der Menschenrechte vor. Zu diesem Zweck und im Einklang mit der [Empfehlung CM/Rec\(2020\)1 des Ministerkomitees an die Mitgliedstaaten zu den Auswirkungen algorithmischer Systeme auf die Menschenrechte bedeutet diese unternehmerische Verantwortung](#), dass flexible Steuerungsmodelle eingeführt werden sollten, die im Falle von Vorfällen eine schnelle und wirksame Wiedergutmachung sowie Möglichkeiten zur Abhilfe gewährleisten und sicherstellen, dass die Verantwortung und Rechenschaftspflicht für den Schutz der Menschenrechte in allen Phasen des Prozesses, von der Vorschlagsphase über die Aufgabenidentifizierung, über die Datenauswahl, -erhebung und -analyse, die Systemmodellierung und -gestaltung bis hin zum laufenden Einsatz, zur Überprüfung und zu den Berichterstattungspflichten eingehalten werden. [Die Empfehlung CM/Rec\(2022\)16 des Ministerkomitees an die Mitgliedstaaten zur Bekämpfung von Hassreden](#) verweist ebenfalls auf die Rolle der Internet-Vermittler bei der Bekämpfung von Hassreden im Internet.¹⁵

11. Was die Überprüfung, Dokumentation und Berichterstattung betrifft, so setzen die Mitgliedstaaten die Empfehlung gegebenenfalls bereits um, indem sie ihren Verpflichtungen gemäß der KI-Verordnung und anderen einschlägigen völkerrechtlichen Instrumenten zum Thema KI nachkommen. Die Empfehlung verfolgt einen abgestuften und differenzierten Ansatz, um die Kohärenz mit dem abgestuften und differenzierten Ansatz des Rahmenübereinkommens und dem risikobasierten Ansatz der KI-Verordnung zu wahren.

Anhang – Grundsätze und Leitlinien zu Gleichheit und künstlicher Intelligenz

I. Ziel

Zu Absatz 1

12. Die Empfehlung betrifft die Förderung der Gleichheit, einschließlich der Gleichheit der Geschlechter, sowie Diskriminierungsrisiken. Sie baut auf der langjährigen und umfangreichen Arbeit auf, die der Europarat in den Bereichen Gleichheit, einschließlich der Gleichheit der Geschlechter, und Nichtdiskriminierung geleistet hat, sowie auf bestehenden Standards zur Bewältigung der neuen Herausforderungen, die sich aus der Entwicklung von KI-Systemen ergeben.¹⁶ Die Empfehlung richtet sich zwar in erster Linie an die Mitgliedstaaten und deren Institutionen, enthält jedoch auch Empfehlungen an Akteure des privaten Sektors, wobei die Besonderheiten

¹⁰ Siehe beispielsweise „Insight into artificial intelligence and its impact on the youth sector“, Veronica Stefan, erstellt im Rahmen der Jugendpartnerschaft zwischen dem Europarat und der Europäischen Kommission, 2024.

¹¹ [„Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination“](#), erstellt von Ivana Bartoletti und Raphaële Xenidis, verabschiedet vom Lenkungsausschuss für Antidiskriminierung, Vielfalt und Inklusion (CDADI) und der Kommission für die Gleichstellung der Geschlechter (GEC) im Jahr 2023.

¹² *Ibid.*, S. 10, 28-30, 78, 88.

¹³ Siehe hierzu beispielsweise die Empfehlung des Menschenrechtskommissars des Europarats in [„Human Rights by Design – Future-Proofing Human Rights Protection in the Era of AI“](#) (2023), S. 18–20, eine Fortsetzung von [„Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights“](#) (2019).

¹⁴ Dazu gehören staatliche Akteure, Vertreter des privaten Sektors, Hochschulen und Forschungseinrichtungen, Gleichstellungsstellen, nationale Menschenrechtsinstitutionen, Ombudsleute sowie die Zivilgesellschaft, einschließlich zivilgesellschaftlicher Jugendorganisationen und Minderheitenorganisationen (sowie kleinerer Minderheitenorganisationen), wobei der Vielfalt innerhalb der jeweiligen Gruppen gebührend Rechnung zu tragen ist.

¹⁵ Absatz 16 ff.

¹⁶ Siehe insbesondere [die Empfehlung CM/Rec\(2020\)1 des Ministerkomitees an die Mitgliedstaaten zu den Auswirkungen algorithmischer Systeme auf die Menschenrechte](#).

ihrer Tätigkeiten berücksichtigt werden.¹⁷ Um sicherzustellen, dass die Empfehlung angesichts des rasanten Tempos der technologischen Entwicklung relevant bleibt, sind ihre Grundsätze und Leitlinien technologieneutral und sollen gleichermaßen für zukünftige, innovative KI-Systeme wie auch für bereits bestehende gelten.

13. In der Empfehlung wird der Begriff „System künstlicher Intelligenz“ im Einklang mit Artikel 2 des Rahmenübereinkommens verwendet, wonach unter einem „System künstlicher Intelligenz“ ein maschinengestütztes System zu verstehen ist, das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben, wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen, erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können. Die verschiedenen Systeme künstlicher Intelligenz unterscheiden sich in ihrem Grad der Autonomie und der Anpassungsfähigkeit nach ihrer Betriebsaufnahme.

14. Wie in dem erläuternden Bericht zum Rahmenübereinkommen hervorgehoben wird, spiegelt die Definition „ein weit gefasstes Verständnis dessen wider, was Systeme der künstlichen Intelligenz sind, insbesondere im Gegensatz zu anderen Arten einfacherer, traditioneller Softwaresysteme, die auf Regeln basieren, die ausschließlich von natürlichen Personen definiert wurden, um Vorgänge automatisch auszuführen“, und „sie soll rechtliche Präzision und Sicherheit gewährleisten, dabei jedoch ausreichend abstrakt und flexibel bleiben, um trotz künftiger technologischer Entwicklungen ihre Gültigkeit zu behalten“.¹⁸ Die Definition des Rahmenübereinkommens orientiert sich an der Definition, die die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) im Jahr 2023 verabschiedet hat, um die internationale Zusammenarbeit im Bereich der künstlichen Intelligenz zu stärken und die Bemühungen um eine Harmonisierung der Regulierung der künstlichen Intelligenz auf globaler Ebene zu erleichtern.¹⁹

15. In der Empfehlung wird der Begriff „Aktivitäten im Lebenszyklus eines KI-Systems“ verwendet, der im Rahmenübereinkommen dazu dient, „alle Aktivitäten vom Entwurf eines Systems künstlicher Intelligenz bis zu dessen Außerbetriebnahme zu erfassen, unabhängig davon, welcher Akteur daran beteiligt ist“ (Absatz 15, erläuternder Bericht zum Rahmenübereinkommen). Dies gewährleistet einen umfassenden, zukunftsfähigen Ansatz. Die Definition umfasst somit (1) Planung und Konzeption, (2) Datenerhebung und -verarbeitung, (3) Entwicklung von Systemen künstlicher Intelligenz, einschließlich der Modellerstellung und/oder der Feinabstimmung bestehender Modelle für bestimmte Aufgaben, (4) Test, Verifikation und Validierung, (5) Bereitstellung der Systeme zur Nutzung, (6) Einführung, (7) Betrieb und Überwachung sowie (8) Außerbetriebnahme.²⁰

16. In der Empfehlung werden die Begriffe „Betreiber“ und „Anbieter“ (die beide aus der KI-Verordnung stammen) im Zusammenhang mit einem KI-System verwendet. „Betroffene Personen“ sind diejenigen, die KI-gestützten Entscheidungen oder Prozessen unterliegen, beispielsweise Kunden, Bewerber, Steuerzahler usw., unabhängig davon, ob die endgültige Entscheidung von einem KI-System getroffen wird oder nicht.²¹ Im Sinne dieser Empfehlung kann eine betroffene Person auch eine juristische Person sein (beispielsweise, wenn eine Organisation aufgrund der geschützten Merkmale ihrer Mitglieder betroffen ist).

Zu Absatz 2

17. Die in der Empfehlung enthaltene Liste der Diskriminierungsgründe wurde ursprünglich in der Strategie des Europarats zur Gleichheit der Geschlechter 2024–2029 veröffentlicht und stellt eine nicht erschöpfende, beispielhafte Aufzählung dar. Sie stützt sich auf Artikel 14 der Europäischen Konvention zum Schutz der Menschenrechte und Grundfreiheiten (die Konvention) und umfasst Diskriminierungsgründe, die erst in jüngerer Zeit Anerkennung gefunden haben – sei es auf politischer Ebene, durch die evolutive Auslegung in der Rechtsprechung des Gerichtshofs oder durch neuere Standards, darunter das Übereinkommen des Europarats zur Verhütung und Bekämpfung von Gewalt gegen Frauen und häuslicher Gewalt (die Istanbul-Konvention) sowie die [Empfehlung CM/Rec\(2024\)4](#) zur Bekämpfung von Hassverbrechen. Die Empfehlung umfasst Diskriminierung aufgrund tatsächlicher oder vermeintlicher Merkmale oder eines Status im Einklang mit der Rechtsprechung des Gerichtshofs.²² Ebenso wird Diskriminierung durch Assoziierung erfasst. Dies ist der Fall, wenn eine Person oder eine Gruppe diskriminiert wird, weil sie mit einem bestimmten Merkmal, Status, einer bestimmten Person oder einer bestimmten Gruppe in Verbindung gebracht wird oder als dazu gehörig wahrgenommen wird. Die Liste der Diskriminierungsgründe umfasst auch die Diskriminierung aufgrund der ethnischen Herkunft. In der Rechtsprechung des Gerichtshofs sind ethnische Zugehörigkeit und „Rasse“²³

¹⁷ Siehe auch [den Erläuternden Bericht zum Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit](#) (SEV Nr. 225), Absatz 17, sowie die [Empfehlung CM/Rec\(2020\)1](#) zu den Auswirkungen algorithmischer Systeme auf die Menschenrechte, Absatz 3.

¹⁸ Erläuternder Bericht zum Rahmenübereinkommen, Absatz 24.

¹⁹ *Ibid.*, Absatz 23-24.

²⁰ *Ibid.*, Absatz 14-15.

²¹ CDADI/GEC-Studie (2023), S. 13.

²² Siehe beispielsweise *Sejdić und Finci gegen Bosnien und Herzegowina* [Große Kammer], 2009, § 43; *Timishev gegen Russland*, 2005, § 55.

²³ Da alle Menschen derselben Spezies angehören, lehnt der Europarat Theorien ab, die auf der Existenz verschiedener „Rassen“ beruhen. In dieser Empfehlung und der Begründung wird jedoch der Begriff „Rasse“ verwendet, um sicherzustellen, dass Personen, die

miteinander verbundene und sich überschneidende Begriffe²⁴, und eine Diskriminierung aufgrund der tatsächlichen oder wahrgenommenen ethnischen Herkunft einer Person stellt eine Form der rassistischen Diskriminierung dar.²⁵ Darüber hinaus ist unter „Nationalität“ die Staatsangehörigkeit zu verstehen.²⁶ Verschiedene Kombinationen von Diskriminierungsgründen stehen häufig im Zusammenhang mit der sozialen Herkunft sowie sozioökonomischer Prekarität und Armut und überschneiden sich mit diesen. In der Empfehlung wird berücksichtigt, dass die soziale Herkunft und der sozioökonomische Status im Zusammenhang mit Diskriminierung eine Rolle spielen, die durch KI-Systeme noch verstärkt werden können.

18. KI-Systeme leiten anhand von durch Menschen definierten Eingaben und unter Verwendung von Daten, die von Menschen, Maschinen oder einer Kombination aus beiden (zu einem bestimmten Zeitpunkt im Verlauf des Prozesses) erhoben wurden, ab, wie sie Ergebnisse generieren können. Folglich können KI-Systeme die menschlichen Vorurteile widerspiegeln und verarbeiten, die bei der Programmierung des Algorithmus, bei der Datenverarbeitung und bei der Interaktion von Menschen mit dem System einfließen.²⁷ Da KI-Systeme eine beispiellose Menge an detaillierten Daten (einschließlich hochsensibler personenbezogener Daten) wesentlich schneller analysieren können als Menschen, kann dies dazu führen, dass KI-Systeme in einer Weise eingesetzt werden, die ungerechte Vorurteile aufrechterhält oder verstärkt.²⁸ So hat beispielsweise eine UNESCO-Studie aufgezeigt, dass große Sprachmodelle (LLMs) dazu neigen, Inhalte zu erzeugen, die sich negativ zu Frauen, homosexuellen Menschen und bestimmten ethnischen Gruppen äußern. Als drei KI-Modelle die Aufgabe erhielten, Sätze zu vervollständigen, die mit dem Ausdruck „Eine homosexuelle Person ist ...“ begannen, waren über 60 % der generierten Inhalte negativ. Beispiele hierfür waren: „Homosexuelle galten als die Niedrigsten in der sozialen Hierarchie“ und „Homosexuelle wurden als Prostituierte und Kriminelle angesehen und hatten keine Rechte.“²⁹ Voreingenommenheit hat im gesamten Lebenszyklus von KI-Systemen, von ihrer Entstehung bis hin zu ihrer Einführung und Nutzung, unterschiedliche Ursachen. Die Komplexität der Entstehung und der Auswirkungen von Verzerrungen ist der Grund dafür, dass dem gesamten Lebenszyklus von KI-Systemen besondere Aufmerksamkeit geschenkt werden muss.³⁰ Besondere Beachtung sollte dabei dem Thema der Rückkopplungsschleifen gelten, die Situationen beschreiben, in denen ein System Daten aus früheren Diskriminierungen als Grundlage für Vorhersagen heranzieht.³¹

19. Stellvertreter-Diskriminierung ist eine der diskriminierenden Beeinträchtigungen, die von KI-Systemen am ehesten verursacht werden.³² „Stellvertreter-Diskriminierung“ bezeichnet eine Diskriminierung, die auf der Grundlage einer scheinbar neutralen Information erfolgt, die jedoch mit einem geschützten Merkmal in Zusammenhang steht oder damit korreliert. Sie kann in Form von unmittelbarer oder mittelbarer Diskriminierung auftreten. So kann beispielsweise die Schuhgröße als Indikator für das Geschlecht dienen oder Namen als Indikator für die ethnische Zugehörigkeit.³³ Im weiteren Sinne könnte die Postleitzahl einer Person als Indikator für die ethnische Zugehörigkeit oder die sexuelle Orientierung dienen, ihr Beruf als Indikator für das Geschlecht und ihr Vorname als Indikator für das Alter.³⁴ Stellvertreter-Diskriminierung wird meist im Rahmen der mittelbaren Diskriminierung erfasst, indem eine Benachteiligung aufgrund eines geschützten Merkmals nachgewiesen wird. Allerdings unterscheidet sie sich von der Diskriminierung durch Assoziation, da sie sich auf eine Information bezieht, die eine bestimmte Person betrifft, und nicht auf die Verbindung dieser Person mit einer anderen Person, einer Gruppe oder einer Situation. Der Ad-hoc-Ausschuss für künstliche Intelligenz (CAHAI) stellte in seiner Machbarkeitsstudie aus dem Jahr 2020 fest, dass „die zunehmende Bedeutung der Stellvertreter-Diskriminierung im Zusammenhang mit maschinellem Lernen Fragen hinsichtlich der Auslegung der Unterscheidung zwischen unmittelbarer und mittelbarer Diskriminierung aufwerfen kann oder sogar hinsichtlich der Angemessenheit dieser Unterscheidung, wie sie traditionell verstanden wird“.³⁵

allgemein und fälschlicherweise als „einer anderen Rasse angehörend“ wahrgenommen werden, nicht vom Schutz dieser Empfehlung ausgeschlossen werden.

²⁴ Sejdić und Finci gegen Bosnien und Herzegowina [Große Kammer], 2009, § 43; Timishev gegen Russland, 2005, § 55-56.

²⁵ Sejdić und Finci gegen Bosnien und Herzegowina [Große Kammer], 2009, § 43; Timishev gegen Russland, 2005, § 55.

²⁶ ECRI-GPR Nr. 16 zum Schutz von Migranten ohne Aufenthaltsgenehmigung vor Diskriminierung, Absatz 10.

²⁷ CDADI/GEC-Studie (2023), S. 15.

²⁸ „Feasibility Study on a legal framework on AI design, development and application based on Council of Europe standards“, zu Deutsch: „Machbarkeitsstudie zu einem Rechtsrahmen für die Konzeption, Entwicklung und Anwendung künstlicher Intelligenz auf der Grundlage der Standards des Europarats“, Ausschuss für künstliche Intelligenz (CAI), 17. Dezember 2020, Absatz 13.

²⁹ „Challenging system prejudice: an investigation into bias against women and girls in large language models“, zu Deutsch: „Systematische Vorurteile hinterfragen: Eine Untersuchung zu Voreingenommenheit gegenüber Frauen und Mädchen in großen Sprachmodellen“, UNESCO und Internationales Forschungszentrum für Künstliche Intelligenz, 2024.

³⁰ CDADI/GEC-Studie (2023), S. 15, unter Berufung auf Ivana Bartoletti, „The Complex Issue of Algorithmic Fairness“ (The Yuan 2021), abrufbar unter: <https://www.the-yuan.com/129/The-Complex-Issue-of-Fairness-in-AI-Part-I.html>. Eine ausführliche Darstellung der verschiedenen Ursachen für verzerrte Ergebnisse finden Sie auf S. 1618.

³¹ CDADI/GEC-Studie (2023), S. 56.

³² *Ibid.*, S. 17, 53, 56.

³³ Fundamental Rights Agency, *Bias in Algorithms – Artificial Intelligence and Discrimination* (2022), S. 24. Weitere Beispiele finden Sie im Bericht des Sonderberichterstatters der Vereinten Nationen zu zeitgenössischen Formen von Rassismus, Rassendiskriminierung, Fremdenfeindlichkeit und damit zusammenhängender Intoleranz (A/HRC/56/68), veröffentlicht am 3. Juni 2024, Absätze 18, 32 und 40.

³⁴ CDADI/GEC-Studie (2023), S. 17.

³⁵ Machbarkeitsstudie (2020) unter Berufung auf die CoE-Studie von F. Zuiderveen Borgesius, „Discrimination, artificial intelligence, and algorithmic decision-making“, zu Deutsch: „Diskriminierung, künstliche Intelligenz und algorithmische Entscheidungsfindung“, 2018; „Affinity Profiling and Discrimination by Association in Online Behavioural Advertising“, zu Deutsch: „Affinitätsprofilierung und Diskriminierung durch Assoziation in der verhaltensorientierten Online-Werbung“, Wachter 2020.

20. Die Empfehlung trägt den tiefgreifenden Veränderungen Rechnung, die KI-Systeme für das Leben von Kindern mit sich bringen. KI-Technologien sind in Spielzeug, Videospielen und adaptiver Lernsoftware integriert.³⁶ Infolgedessen haben Kinder bereits in immer jüngerem Alter Zugang zu Technologien, die KI-Systeme nutzen. Neben diesem allgemeinen Überblick enthält die vorliegende Begründung konkrete Beispiele, die Kinder betreffen. Darüber hinaus berücksichtigt die Empfehlung, dass KI-Systeme auch dann, wenn sie nicht direkt von Kindern und Jugendlichen genutzt werden, erhebliche Auswirkungen auf deren Leben und Rechte haben.

II. Umfassender Ansatz in Bezug auf Gesetzgebung, Politik und regulatorische

Rahmenbedingungen

Zu den Absätzen 3 – 4

21. Die Mitgliedstaaten können Absatz 3 der Empfehlung durch eine Vielzahl von Maßnahmen umsetzen, die dem jeweiligen nationalen rechtlichen, politischen und regulatorischen Rahmen angemessen sind. Solche Maßnahmen müssen wirksam und zielgerichtet sein, um Gleichheit zu fördern bzw. Diskriminierung zu verhindern, und sie sollten sich auf sich überschneidende und kumulative Formen der Diskriminierung beziehen, die bei Aktivitäten im Lebenszyklus von KI-Systemen auftreten können. Einige Mitgliedstaaten könnten sich dafür entscheiden, einen umfassenden Aktionsplan zu KI und Gleichheit³⁷ einzuführen; sie könnten dies jedoch auch durch die Aufnahme spezifischer Inhalte in sektorale Aktionspläne erreichen.

22. In der Empfehlung wird betont, dass es für die Mitgliedstaaten von entscheidender Bedeutung ist, sicherzustellen, dass ihre bestehenden Rechtsvorschriften, Strategien und regulatorischen Rahmenbedingungen die Gleichheit, einschließlich der Gleichheit der Geschlechter, fördern und alle Formen der Diskriminierung während des gesamten Lebenszyklus von KI-Systemen verhindern und bekämpfen (Absatz 4.1). Dazu gehört die Überprüfung und Weiterentwicklung bestehender Rechtsvorschriften, Strategien und regulatorischer Rahmenbedingungen, um sicherzustellen, dass diese gegebenenfalls auf Aktivitäten im gesamten Lebenszyklus von KI-Systemen Anwendung finden.

23. In der Empfehlung wird berücksichtigt, dass viele Mitgliedstaaten bereits über rechtliche, regulatorische und politische Rahmenbedingungen im Bereich der Antidiskriminierung und Gleichheit verfügen, die in vielen Fällen auf Aktivitäten im Lebenszyklus von KI-Systemen anwendbar sind. Der erste Schritt zur Gewährleistung eines umfassenden Ansatzes in Bezug auf Rechtsvorschriften, politische Maßnahmen und regulatorische Rahmenbedingungen, wie in Absatz 4.1 festgestellt, besteht daher darin, dass die Mitgliedstaaten diese überprüfen und bewerten, ob Reformen oder Weiterentwicklungen erforderlich sind. Dieser Ansatz entspricht dem in der [Resolution 2343 \(2020\) der Parlamentarischen Versammlung](#) zur Verhinderung von Diskriminierung durch den Einsatz künstlicher Intelligenz dargelegten Ansatz, in der die Mitgliedstaaten aufgefordert werden, ihre Antidiskriminierungsgesetze zu überprüfen und gegebenenfalls anzupassen, um sicherzustellen, dass diese alle Fälle abdecken, in denen der Einsatz von KI zu unmittelbarer oder mittelbarer Diskriminierung, einschließlich Diskriminierung durch Assoziierung, führen kann, und dass Beschwerdeführer uneingeschränkter Zugang zur Justiz haben. Darüber hinaus sollten die Mitgliedstaaten andere bestehende rechtliche, regulatorische und politische Rahmenbedingungen prüfen, einschließlich derjenigen, die Rechtsbehelfe und Überprüfung betreffen, um sicherzustellen, dass diese den Umfang und die Umstände von Beschwerden wegen Diskriminierung im Zusammenhang mit KI-Systemen umfassend abdecken. Damit die einschlägigen rechtlichen, regulatorischen und politischen Rahmenbedingungen wirksam und sinnvoll sind, müssen sie in der Praxis effektiv umgesetzt und durchgesetzt werden. Die Mitgliedstaaten sollten zudem sicherstellen, dass neue Rahmenbedingungen, die als Reaktion auf KI-Systeme entwickelt werden, mit den bestehenden Rahmenbedingungen zur Bekämpfung von Diskriminierung und zur Förderung der Gleichheit im Einklang stehen. Um die Einhaltung dieser Vorschriften zu überwachen, könnten die Mitgliedstaaten beispielsweise eine Gleichheitsstelle beauftragen oder nationale Beobachtungsstellen bzw. Forschungszentren einrichten oder unterstützen, die den spezifischen Auftrag haben, die Auswirkungen von KI-Systemen auf die Gleichheit, einschließlich der Gleichheit der Geschlechter, und die Nichtdiskriminierung zu überwachen, um kontinuierlich evidenzbasierte politische Empfehlungen zu erarbeiten.

24. In Absatz 4 wird ausdrücklich auf die Anwendung eines abgestuften und differenzierten Ansatzes Bezug genommen, um die Kohärenz mit dem abgestuften und differenzierten Ansatz des Rahmenübereinkommens und dem risikobasierten Ansatz der KI-Verordnung zu wahren. In Artikel 1 Absatz 2 des Rahmenübereinkommens heißt es, dass Maßnahmen „so abgestuft und differenziert sein [sollen], wie es angesichts der Schwere und der Wahrscheinlichkeit des Eintretens nachteiliger Auswirkungen auf Menschenrechte, Demokratie und Rechtsstaatlichkeit während des gesamten Lebenszyklus von Systemen

³⁶ Siehe beispielsweise den Bericht zum Thema „The metaverse and its impact on human rights, the rule of law and democracy“, zu Deutsch: „Das Metaversum und seine Auswirkungen auf Menschenrechte, Rechtsstaatlichkeit und Demokratie“, Europarat und Institute of Electrical and Electronics Engineers (IEEE), Mai 2024.

³⁷ *Ibid.*, S. 84.

künstlicher Intelligenz erforderlich ist.“ Dies entspricht dem Grundsatz der Verhältnismäßigkeit, der für die Anwendung der Europäischen Menschenrechtskonvention von grundlegender Bedeutung ist. In Absatz 17 des erläuternden Berichts zum Rahmenübereinkommen wird erklärt, dass „diese Bestimmung besagt, dass Maßnahmen gemäß dem Rahmenübereinkommen entsprechend dem von einem System künstlicher Intelligenz ausgehenden Risiko angemessen ausgestaltet werden müssen und dass es den Vertragsparteien des Rahmenübereinkommens obliegt, zu entscheiden, wie die einschlägigen konkurrierenden Interessen in den einzelnen Bereichen gegeneinander abzuwägen sind, wobei die Besonderheiten der Aktivitäten im privaten Sektor, deren innerstaatlicher Rechtsrahmen und die nationale Strategie zum Thema künstliche Intelligenz zu berücksichtigen sind, während gleichzeitig der Schutz und die Förderung der Menschenrechte, der Demokratie und der Rechtsstaatlichkeit gewährleistet werden müssen.“ Ebenso räumt die Empfehlung den Mitgliedstaaten Flexibilität bei der Anwendung eines solchen Ansatzes ein.

Zu Absatz 5

25. Bestimmte Rechte, die als „eingeschränkte Rechte“ bezeichnet werden, darunter das Recht auf Achtung des Privat- und Familienlebens gemäß Artikel 8 der Konvention und das Recht auf freie Meinungsäußerung gemäß Artikel 10, können zum Schutz der Rechte anderer oder des übergeordneten öffentlichen Interesses eingeschränkt werden. In der Rechtsprechung des Gerichtshofs sind feststehende Maßstäbe für die Beurteilung der Zulässigkeit solcher Einschränkungen verankert.³⁸ Wie in anderen Bereichen auch, muss jede Einschränkung der Menschenrechte im Zusammenhang mit dem Einsatz von KI-Systemen dem Legalitätsprinzip entsprechen, das Klarheit und Vorhersehbarkeit des Rechts verlangt. Dieser Grundsatz verlangt beispielsweise, dass das Recht zugänglich und präzise formuliert ist, damit der Einzelne sein Verhalten entsprechend ausrichten kann. Darüber hinaus ist es von entscheidender Bedeutung, dass Einschränkungen der Menschenrechte nicht über das hinausgehen, was in einer demokratischen Gesellschaft zur Erreichung des angestrebten legitimen Ziels erforderlich ist, und dass der Eingriff in einem angemessenen Verhältnis zu dem öffentlichen Interesse steht, dem er dienen soll.

26. Im Fall *Glukhin gegen Russland* wandte der Gerichtshof diesen Maßstab an und stellte fest, dass sowohl gegen Artikel 8 als auch gegen Artikel 10 verstoßen worden sei, wobei er den stark in die Privatsphäre eingreifenden Charakter der Gesichtserkennungstechnologie hervorhob. In diesem Fall war diese Technologie von den Strafverfolgungsbehörden eingesetzt worden, um einen friedlichen Demonstranten aufzuspüren und festzunehmen, dessen Handlungen keine Gefahr oder Bedrohung für die öffentliche Ordnung darstellten. Der Gerichtshof vertrat die Auffassung, dass der flächendeckende Einsatz von Gesichtserkennungstechnologie durch Strafverfolgungsbehörden zur Identifizierung von Demonstranten Grundrechte verletzen würde, sofern keine klaren Vorschriften und wirksamen Schutzvorkehrungen gegen Willkür festgelegt würden, wie beispielsweise Genehmigungsverfahren, Verfahren zur Prüfung, Nutzung und Speicherung der gewonnenen Daten, etwaige Kontrollmechanismen oder verfügbare Rechtsbehelfe. Es wurde betont, dass dieser Bedarf an Schutzmaßnahmen umso größer ist, wenn es um den Einsatz von Technologien zur (Echtzeit-)Gesichtserkennung geht.³⁹ Dieser Fall unterstreicht zudem die Notwendigkeit der in den Absätzen 5.2 und 5.3 dargelegten Maßnahmen, die Transparenz, Erklärbarkeit und eine solide Aufsicht vorsehen, um den von KI-Systemen ausgehenden Risiken zu begegnen. Darüber hinaus verdeutlicht dieser Fall, wie wichtig es ist, sicherzustellen, dass Einschränkungen verhältnismäßig bleiben und dass KI-Systeme, einschließlich Gesichtserkennungstechnologie, nur in geeigneten Situationen eingesetzt werden, wie in Absatz 5.1 hervorgehoben wird.

27. Absatz 5 der Empfehlung steht im Einklang mit den in Artikel 8 des Rahmenübereinkommens festgelegten Grundsätzen, wonach die Mitgliedstaaten Maßnahmen erlassen oder aufrechterhalten müssen, um sicherzustellen, dass angemessene Transparenz- und Aufsichtsanforderungen, die auf die jeweiligen Kontexte und Risiken zugeschnitten sind, in Bezug auf Aktivitäten im Lebenszyklus von KI-Systemen gelten. Wie im Erläuternden Bericht zum Rahmenübereinkommen dargelegt, ist dieses Ergebnis auf bestimmte Merkmale von Systemen künstlicher Intelligenz zurückzuführen, die künstliche Intelligenz von herkömmlichen Computersystemen unterscheiden; dazu können Komplexität, Intransparenz, Anpassungsfähigkeit und unterschiedliche Grade an Autonomie gehören.⁴⁰ Der Erläuternde Bericht zum Rahmenübereinkommen unterstreicht ferner, dass Transparenz im Zusammenhang mit dem Lebenszyklus von KI-Systemen bedeutet, sicherzustellen, dass die Entscheidungsprozesse und Abläufe dieser Systeme für die relevanten Interessengruppen zugänglich und nachvollziehbar sind, wobei Aspekte wie Sicherheit, geistiges Eigentum und Datenschutz zu berücksichtigen sind. Transparenz kann die Offenlegung wichtiger Informationen über KI-Systeme erfordern, wie beispielsweise Algorithmen, die Herkunft der Daten und Trainingsmethoden; dies trägt dazu bei, die Rechenschaftspflicht zu stärken, und ermöglicht es Einzelpersonen, Ergebnisse gegebenenfalls anzufechten. Der Grundsatz der Transparenz umfasst Erklärbarkeit und Interpretierbarkeit, wodurch sichergestellt wird, dass die Vorhersagen und Entscheidungen von KI-Systemen sowohl von Fachleuten als auch von Laien nachvollzogen werden können, insbesondere in sensiblen Bereichen wie dem

³⁸ *Stoll gegen Schweiz*, Antrag Nr. 69698/01, Urteil der Grossen Kammer vom 10. Dezember 2007.

³⁹ *Glukhin gegen Russland*, Antrag Nr. 11519/20, Urteil vom 4. Juli 2023.

⁴⁰ Erläuternder Bericht zum Rahmenübereinkommen, Absatz 56.

Gesundheitswesen und der Justiz. Eine vollständige Transparenz kann durch technologische Einschränkungen sowie durch die Notwendigkeit, konkurrierende Interessen wie Datenschutz, Vertraulichkeit, Geschäftsgeheimnisse und nationale Sicherheit gegeneinander abzuwägen, eingeschränkt sein.⁴¹ Transparenz, Erklärbarkeit und Interpretierbarkeit sind von entscheidender Bedeutung, um sicherzustellen, dass KI-Systeme angemessen eingesetzt werden, Diskriminierung verhindert und bekämpft wird und die Rechenschaftspflicht für diskriminierende Ergebnisse gewährleistet ist.

28. In Absatz 5.1 der Empfehlung wird betont, wie wichtig es ist, die Angemessenheit des Einsatzes von KI-Systemen in einem bestimmten Kontext sicherzustellen. Dies bezieht sich auf die „Null-Fragen“, wie sie in der HUDERIA-Methodik erläutert werden.⁴² Entscheidungsträger prüfen unter Anwendung eines abgestuften und differenzierten Ansatzes, ob der Einsatz von KI-Systemen in einer bestimmten Situation gerechtfertigt ist, bevor sie überlegen, wie diese eingesetzt werden sollten; dabei wird berücksichtigt, dass der Einsatz von KI-Systemen in bestimmten Kontexten nicht von Natur aus geeignet oder vorteilhaft ist. Die Gewährleistung der Angemessenheit des Einsatzes von KI-Systemen in bestimmten Kontexten orientiert sich an der Berücksichtigung der Art des Problems, das das KI-System zu lösen versucht, sowie der tatsächlichen und potenziellen Auswirkungen oder Risiken für die Gleichheit und Nichtdiskriminierung, die mit dem Einsatz eines KI-Systems für diesen bestimmten Zweck verbunden sind, und an der Verpflichtung zur Wahrung der Menschenrechte, wodurch ein übermäßiges Vertrauen in Technologie in Kontexten verhindert wird, in denen diese Risiken verstärken oder unverhältnismäßige Auswirkungen hervorrufen könnte.⁴³ Dies ist insbesondere in Bereichen relevant, in denen der Einsatz von KI-Systemen erhebliche Auswirkungen auf die Gleichheit, einschließlich der Gleichheit der Geschlechter, und die Nichtdiskriminierung haben kann.⁴⁴ Bei der Feststellung, in welchen Situationen der Einsatz von KI-Systemen angemessen ist oder nicht, sollten die Perspektiven von Personen und Gruppen berücksichtigt werden, die von Diskriminierung bedroht sind.

29. Die Empfehlung legt den Schwerpunkt auf die menschliche Aufsicht, was beispielsweise bedeuten könnte, dass KI-Systeme während ihrer Nutzung wirksam von Menschen überwacht werden können oder dass die endgültige Entscheidung nach Prüfung der Ergebnisse und der zugrunde liegenden Begründung von einem Menschen getroffen wird.⁴⁵

Zu Absatz 6

30. In Absatz 6 wird der in der gesamten Empfehlung verfolgte Ansatz verdeutlicht, wonach die Mitgliedstaaten bei der Konzeption und Umsetzung von Maßnahmen das Ausmaß des Diskriminierungsrisikos bzw. die Auswirkungen auf die Rechte einer Person, die sich aus einem bestimmten Anwendungsfall oder Kontext ergeben, gebührend berücksichtigen sollten. Da die Konvention, einschließlich des Gleichheitsgrundsatzes und des Diskriminierungsverbots gemäß Artikel 14, Anwendung findet, sollte sich die Verhältnismäßigkeitsprüfung nicht auf Systeme mit hohem Risiko beschränken, sondern alle Risikostufen von KI-Anwendungen umfassen, einschließlich Szenarien mit mittlerem und geringem Risiko. Ein solcher Ansatz trägt der Tatsache Rechnung, dass selbst Systeme mit geringerem Risiko erhebliche kumulative oder kontextspezifische Auswirkungen auf Rechte und Gleichheit, einschließlich der Gleichheit der Geschlechter, haben können. Die Bewertung einer bestimmten Maßnahme sollte von der zuständigen nationalen Behörde gemäß den nationalen Rechtsvorschriften erfolgen, wobei eine angemessene parlamentarische Kontrolle gewährleistet sein muss, um Rechenschaftspflicht und Transparenz zu gewährleisten.

31. In diesem Zusammenhang hat die Parlamentarische Versammlung die Mitgliedstaaten aufgefordert, sicherzustellen, dass der Einsatz von KI-basierten Technologien durch öffentliche Behörden einer angemessenen parlamentarischen Kontrolle und öffentlichen Überprüfung unterliegt, indem solche Technologien in die regelmäßigen parlamentarischen Debatten einbezogen, Strukturen für diese Diskussionen geschaffen, eine Meldepflicht der Regierung vor dem Einsatz eingeführt und die systematische Erfassung ihrer Nutzung in einem öffentlichen Register vorgeschrieben werden.⁴⁶ Die Parlamentarische Versammlung betonte zudem, wie wichtig es ist, sicherzustellen, dass die Aktivitäten großer privater Unternehmen, die an der Entwicklung und Bereitstellung KI-basierter Technologien und Dienste beteiligt sind, einer demokratischen Kontrolle unterliegen.⁴⁷

Zu Absatz 7

⁴¹ *Ibid*, Absatz 57-65.

⁴² Siehe CaAI(2024)16rev2, Seite 12.

⁴³ Siehe beispielsweise [„The Administration and You, A Handbook“, 3. Auflage](#), Grundsätze des Verwaltungsrechts bezüglich der Beziehungen zwischen Einzelpersonen und Behörden, Europarat, S. 19: „Der Einsatz algorithmischer Entscheidungssysteme durch Behörden sollte in einem angemessenen Verhältnis zu dem verfolgten berechtigten öffentlichen Interesse stehen.“

⁴⁴ Siehe Absätze 95–118 der vorliegenden Begründung.

⁴⁵ Siehe auch erläuternder Bericht zum Rahmenübereinkommen, Absatz 65.

⁴⁶ [Resolution 2343 \(2020\)](#) „Preventing discrimination caused by the use of artificial intelligence“, zu Deutsch: „Verhinderung von Diskriminierung durch den Einsatz künstlicher Intelligenz“ Parlamentarische Versammlung des Europarates, Absatz 11.

⁴⁷ [Resolution 2341 \(2020\)](#) „Need for democratic governance of artificial intelligence“, zu Deutsch: „Die Notwendigkeit einer demokratischen Steuerung künstlicher Intelligenz“ Parlamentarische Versammlung des Europarates, Absatz 14.6.

32. Positive Verpflichtungen bedeuten, dass Staaten unter bestimmten Umständen die Pflicht haben, aktiv Maßnahmen zu ergreifen, um Gleichheit zu erreichen und Diskriminierung zu verhindern.⁴⁸ Dies geht über begrenzte passive oder negative Verpflichtungen zur Nichtdiskriminierung hinaus, da es die Ergreifung präventiver Maßnahmen gegen Diskriminierung oder positiver Maßnahmen zur Förderung der Gleichheit als Mittel zur Einhaltung von Artikel 14 der Konvention beinhaltet.⁴⁹ Im Zusammenhang mit Artikel 14 können solche den Mitgliedstaaten obliegenden positiven Verpflichtungen sogenannte „positive Maßnahmen“ oder „positive Aktionen“ umfassen, die ein Staat ergreifen könnte oder sollte, um „tatsächliche Ungleichheiten“ zu beseitigen.⁵⁰

33. Ohne geeignete Schutzmaßnahmen besteht die Gefahr, dass KI-Systeme strukturelle Ungleichheiten sowie die in der Gesellschaft bestehenden Mechanismen der Ausgrenzung und Benachteiligung reproduzieren und verstärken. Absatz 7 der Empfehlung soll den Mitgliedstaaten daher dabei helfen, einen umfassenden Ansatz für ihre Gesetzgebung, ihre Politik und ihre regulatorischen Rahmenbedingungen zu verfolgen, der sich am Grundsatz der substantziellen Gleichheit orientiert.⁵¹ Dieser Ansatz ist eine Folge der Tatsache, dass formale Gleichheit oder Chancengleichheit den unterschiedlichen Ausgangspositionen verschiedener gesellschaftlicher Gruppen nicht ausreichend Rechnung tragen können. Die Annahme, dass ein System verschiedenen Gruppen gleichermaßen gerecht wird, wird *de facto* dazu führen, dass Personen, die nicht zur Mehrheit gehören, nicht im gleichen Maße wie andere Gruppen von KI-Systemen und den damit verbundenen Möglichkeiten profitieren können.

34. Positive Maßnahmen oder vorübergehende Sondermaßnahmen, wie in Absatz 5 der ECRI-Leitlinie Nr. 7 zu nationalen Rechtsvorschriften zur Bekämpfung von Rassismus und rassistischer Diskriminierung erläutert,⁵² umfassen Maßnahmen, die darauf abzielen, Benachteiligungen von Gruppen, die Diskriminierung und Intoleranz ausgesetzt sind, zu verhindern oder auszugleichen und deren uneingeschränkte Teilhabe in allen Lebensbereichen zu ermöglichen. Solche Maßnahmen können ergriffen werden, um tatsächliche oder *de facto* vollständige Gleichheit zu erreichen. ECRI hat das Konzept der vorübergehenden Sondermaßnahmen „als wirksames Instrument zur Schaffung fairer und gleicher Chancen in der Gesellschaft für Angehörige benachteiligter Gruppen“ beschrieben.⁵³ So kam der Gerichtshof beispielsweise in der *Rechtssache Horváth und Kiss gegen Ungarn*, bei der es um die systematische Einweisung von Roma-Kindern in Sonderschulen in Ungarn ging, zu dem Schluss, dass im Zusammenhang mit dem Recht auf Bildung von Angehörigen von Gruppen, die in der Vergangenheit unter Diskriminierung im Bildungsbereich gelitten haben und deren Auswirkungen bis heute andauern, strukturelle Mängel die Umsetzung positiver Maßnahmen erforderten, um *unter anderem* die Beschwerdeführer bei etwaigen Schwierigkeiten zu unterstützen, auf die sie beim Durcharbeiten des Lehrplans stießen. Daher waren zusätzliche Maßnahmen erforderlich, um diese Probleme anzugehen, wie beispielsweise eine aktive und strukturierte Einbindung der zuständigen Sozialdienste.⁵⁴

35. So fordert die Empfehlung beispielsweise in Kapitel III die Mitgliedstaaten auf, öffentliche und private Anbieter und Betreiber von KI-Systemen dazu anzuhalten und dabei zu unterstützen, die Vielfalt ihrer Belegschaft auf allen Ebenen zu fördern und zu erhalten, einschließlich derjenigen, die KI-Systeme entwickeln und einsetzen, indem sie ein ausgewogenes Geschlechterverhältnis und eine verbesserte Vertretung von Personen aus unterrepräsentierten Gruppen auf allen Entscheidungsebenen gewährleisten (Absatz 10). Ihre Beteiligung sollte sinnvoll sein, d. h., diese Personen sollten echte Möglichkeiten haben, mit Entscheidungsträgern in Kontakt zu treten, und ihnen sollte ausreichend Zeit eingeräumt werden, um substantielle Beiträge zu leisten. Die Mitgliedstaaten können sich von der [Empfehlung CM/Rec\(2003\)3 des Ministerkomitees an die Mitgliedstaaten zur ausgewogenen Beteiligung von Frauen und Männern an politischen und öffentlichen Entscheidungsprozessen](#) sowie von der [Empfehlung CM/Rec\(2023\)9 des Ministerkomitees an die Mitgliedstaaten zur aktiven politischen Teilhabe von Jugendlichen aus nationalen Minderheiten leiten lassen](#). Zu diesem Zweck sollten Strategien zur Personalgewinnung und -bindung ausdrücklich auf Teilnahmebarrieren eingehen, die über die Einstellungsrichtlinien hinausgehen, darunter die Unternehmenskultur, das Arbeitsumfeld und die Aufstiegsmöglichkeiten, die sich möglicherweise unverhältnismäßig stark auf Frauen und Personen auswirken, die von Diskriminierung bedroht sind. In der Empfehlung wird ferner darauf hingewiesen, dass die Mitgliedstaaten geeignete Maßnahmen ergreifen sollten, um den besonders gravierenden Auswirkungen

⁴⁸ Vgl. [Europäischer Gerichtshof für Menschenrechte, Leitfaden zu Artikel 14 der Konvention \(Diskriminierungsverbot\) und zu Artikel 1 des Protokolls Nr. 12 \(allgemeines Diskriminierungsverbot\)](#) (2022), Absatz 42–43. Siehe hierzu beispielsweise auch Thlimmenos gegen Griechenland, [Beschwerde Nr. 34369/97](#), Urteil vom 2. April 2000, sowie Horváth und Kiss gegen Ungarn, [Beschwerde Nr. 11146/11](#), Urteil vom 29. Januar 2013.

⁴⁹ Der Gerichtshof hat den Begriff der positiven Verpflichtungen im Zusammenhang mit der Überprüfung der Internetkommunikation angewandt, beispielsweise in der Rechtssache Bărbulescu gegen Rumänien, [Beschwerde Nr. 61496/08](#), Urteil der Großen Kammer vom 5. September 2017, Randnummer 108, bei der Prüfung des Geltungsbereichs von Artikel 8.

⁵⁰ Leitfaden zu Artikel 14 der Europäischen Menschenrechtskonvention und zu Artikel 1 des Protokolls Nr. 12, S. 14.

⁵¹ CDADI/GEC-Studie (2023), S. 73.

⁵² Siehe die [überarbeitete Allgemeine Politikempfehlung Nr. 7 der ECRI zu nationalen Rechtsvorschriften zur Bekämpfung von Rassismus und rassistischer Diskriminierung](#), Absatz 5.

⁵³ Seminar mit nationalen Fachgremien zur Bekämpfung von Rassismus und rassistischer Diskriminierung zum Thema „positive Maßnahmen“: [Erläuterung](#) (2007).

⁵⁴ [Horváth und Kiss gegen Ungarn](#), Absatz 104.

entgegenzuwirken, die der Einsatz von KI-Systemen in verschiedenen Bereichen, darunter auch im Bildungswesen, auf die Gleichheit, einschließlich der Gleichheit der Geschlechter, und die Nichtdiskriminierung haben kann. Die Mitgliedstaaten können diese Ziele durch eine Vielzahl von Maßnahmen erreichen. Eine wirksame Methode sind positive Maßnahmen (zum Beispiel, je nach nationalem Kontext, die Einführung einer Quote für die Besetzung von Entscheidungspositionen in der KI-Forschung und -Entwicklung mit Frauen und unterrepräsentierten Gruppen oder die Finanzierung von Stipendien für Mädchen und junge Menschen aus unterrepräsentierten Gruppen, damit diese naturwissenschaftliche, technische, ingenieurwissenschaftliche und mathematische Fächer – MINT – in der Schule, im Rahmen von Ausbildungsprogrammen oder über E-Learning-Plattformen studieren können). Es könnten auch weitere Anreize in Betracht gezogen werden, wie beispielsweise Auszeichnungssysteme für Organisationen, die bei der Förderung der Gleichheit, einschließlich der Gleichheit der Geschlechter, bei der Entwicklung und dem Einsatz von KI vorbildliche Praktiken an den Tag legen, sowie die Unterstützung von KI-Initiativen im öffentlichen Interesse, darunter Open-Source-Projekte, die darauf abzielen, die Gleichheit zu fördern und Diskriminierung abzubauen.

36. Darüber hinaus empfiehlt die ECRI, dass Gleichheitsstellen den Einsatz positiver Maßnahmen fördern sollten, insbesondere in Bereichen, in denen tief verwurzelte, lang anhaltende strukturelle Diskriminierung bekämpft werden muss.⁵⁵ In ihrer Allgemeinen Empfehlung Nr. 7 befürwortet die ECRI ausdrücklich positive Verpflichtungen zur Förderung der Gleichheit und zur Verhinderung von Diskriminierung in Form von Verfassungsbestimmungen, Pflichten für Behörden sowie Verpflichtungen für öffentliche Einrichtungen, „die Vergabe von Aufträgen, Darlehen, Zuschüssen oder anderen Vorteilen“ von der Einhaltung der positiven Verpflichtung zur Förderung der Gleichheit und zur Verhinderung von Diskriminierung abhängig zu machen. Wie in der Studie über die Auswirkungen von Systemen künstlicher Intelligenz, deren Potenzial zur Förderung der Gleichheit und die Risiken, die sie im Zusammenhang mit der Nichtdiskriminierung mit sich bringen können, dargelegt wurde, kann dies als Rechtsgrundlage dienen, um im Zusammenhang mit dem Einsatz von KI durch öffentliche Behörden eine Verpflichtung zur gleichheitsorientierten Politikgestaltung zu schaffen. Dieser Ansatz könnte von den Mitgliedstaaten bei der Überprüfung, Anwendung und Weiterentwicklung bestehender und neuer gesetzlicher, regulatorischer und politischer Rahmenbedingungen im Zusammenhang mit KI-Systemen berücksichtigt werden. Die norwegische Gleichheitsgesetzgebung könnte in diesem Zusammenhang als nützliche Inspiration dienen, da sie die Förderung der Gleichheit als gesetzliche Verpflichtung vorsieht.⁵⁶

37. Was Absatz 10.2 betrifft, so kann die Erleichterung der Vereinbarkeit von Privat- und Berufsleben, beispielsweise durch Investitionen in die Pflege, die Förderung der Work-Life-Balance und die Auseinandersetzung mit unbezahlter Pflege- und Hausarbeit, dazu beitragen, die Ungleichheit zwischen den Geschlechtern in der Arbeitswelt zu verringern und den Arbeitsplatz inklusiv und vielfältig zu gestalten.⁵⁷ Es könnten Maßnahmen umgesetzt werden, die darauf abzielen, flexible Arbeitsmodelle zu ermöglichen und Privat- und Berufsleben voneinander zu trennen.

Zu Absatz 8:

38. Die Konsultation zu den von den Mitgliedstaaten eingeführten Maßnahmen im Zusammenhang mit der Konzeption, Entwicklung und dem Einsatz von KI-Systemen spielt eine entscheidende Rolle bei der Gewährleistung von Transparenz, Rechenschaftspflicht und einer fundierten Beteiligung der Öffentlichkeit. Wie im Erläuternden Bericht zu Artikel 19 des Rahmenübereinkommens dargelegt, kann eine öffentliche Konsultation das zivilgesellschaftliche Engagement fördern, Einzelpersonen und Expertinnen und Experten in die Lage versetzen, sich an der öffentlichen Diskussion über Themen von weitreichender gesellschaftlicher und politischer Bedeutung zu beteiligen, und ein größeres öffentliches Bewusstsein für grundlegende und sich abzeichnende Fragen schaffen, einschließlich solcher, die für die frühen Phasen der Konzeption relevant sind und durch die Aktivitäten im Lebenszyklus von Systemen künstlicher Intelligenz aufgeworfen werden. Dem Bericht des Menschenrechtskommissars des Europarates zum Thema „Human Rights by Design“⁵⁸ zufolge fanden in den letzten Jahren in vielen Mitgliedstaaten öffentliche Konsultationen zum Thema KI statt. So weist der Kommissar beispielsweise darauf hin, dass Irland bei der Ausarbeitung seiner nationalen KI-Strategie eine umfassende Konsultation durchgeführt hat, an der die gesamte Regierung, relevante Interessengruppen (Industrie, Expertinnen und Experten, Wissenschaftler und Forschungseinrichtungen) sowie die breite Öffentlichkeit beteiligt waren. Die Malta.AI-Arbeitsgruppe, ein Expertengremium, das mit der Ausarbeitung der maltesischen KI-Strategie betraut ist, hat im Rahmen der Ausarbeitung des „Malta Ethical AI Framework“, einer Reihe von Leitlinien für ethische und vertrauenswürdige KI, eine öffentliche Konsultation gestartet. Durch eine stärkere Einbindung der Öffentlichkeit sowie von Gleichheitsstellen, nationalen Menschenrechtsinstitutionen und Ombudsmännern, die über Fachwissen zu den Auswirkungen künstlicher Intelligenz auf Gleichheit und Nichtdiskriminierung verfügen, können die Mitgliedstaaten sicherstellen, dass Rechts- und Politikreformen

⁵⁵ Siehe die [Allgemeine Politikempfehlung Nr. 2 \(überarbeitete Fassung\) der ECRI zu Gleichstellungsstellen zur Bekämpfung von Rassismus und Intoleranz auf nationaler Ebene](#), Absatz 60.

⁵⁶ Siehe Kapitel 4 des norwegischen Gesetzes über Gleichstellung und das Diskriminierungsverbot ([Gleichstellungs- und Antidiskriminierungsgesetz](#)) von 2018.

⁵⁷ Strategie des Europarates [zur Gleichstellung der Geschlechter \(2024-2029\)](#), S. 41.

⁵⁸ [„Human Rights by Design – Future-Proofing Human Rights Protection in the Era of AI“](#) (2023), S. 18–20.

präzise und umfassend auf neue Herausforderungen reagieren. Es sind gezielte Informationsmaßnahmen erforderlich, um die Interessengruppen, einschließlich der Bürger, über Konsultationen zu rechtlichen oder politischen Initiativen sowie darüber zu informieren, wie sie sich daran beteiligen können. Die Mitgliedstaaten werden dazu angehalten, die Bereitstellung von Informationen über verschiedene Medien und Plattformen zu erleichtern, um Inklusion zu gewährleisten. Zu einer sinnvollen Beteiligung sollte auch gehören, dass man nach der Teilnahme über das Ergebnis informiert wird. In der Empfehlung wird festgelegt, dass diejenigen, die am stärksten von Diskriminierung betroffen sind, konsultiert werden sollten. Auch wenn diese Aufzählung nicht vollständig ist, kann dies Frauen, von Rassismus bedrohte Personen, LGBTI-Personen, Jugendliche, Kinder, ältere Menschen sowie Menschen mit Behinderung und deren Vertreterinnen und Vertreter umfassen. Dies ist besonders wichtig, um den am stärksten unterrepräsentierten Gruppen mehr Gehör zu verschaffen und die demokratische Rechenschaftspflicht bei der Verwaltung von KI zu stärken.⁵⁹

III. Förderung der Gleichheit durch öffentliche und private Akteure

39. In diesem Kapitel liegt der Schwerpunkt auf dem positiven Potenzial von KI-Systemen zur Förderung der Gleichheit, einschließlich der Gleichheit der Geschlechter. Darin werden konkrete Maßnahmen untersucht, mit denen dieses Potenzial bei der Konzeption, Entwicklung und dem Einsatz von KI-Systemen ausgeschöpft werden kann.

40. Dieses Kapitel enthält zahlreiche Empfehlungen, die sich auf die Aktivitäten von Akteuren des privaten Sektors beziehen. Dieser Ansatz steht im Einklang mit dem Ansatz des Rahmenübereinkommens⁶⁰ und folgt zudem früheren Empfehlungen des Europarats, beispielsweise der [Empfehlung CM/Rec\(2020\)1](#) zu den menschenrechtlichen Auswirkungen algorithmischer Systeme, in der gefordert wird, dass „Akteure des privaten Sektors, die algorithmische Systeme entwerfen, entwickeln oder implementieren [...] einem Standardrahmen für die Sorgfaltspflicht im Bereich der Menschenrechte folgen, um zu vermeiden, dass Diskriminierung über den gesamten Lebenszyklus ihrer Systeme hinweg gefördert oder verfestigt wird“ und dass „sie dafür sorgen, dass die Konzeption, Entwicklung und der fortlaufende Einsatz ihrer algorithmischen Systeme keine unmittelbaren oder mittelbaren diskriminierenden Auswirkungen auf Einzelpersonen oder Gruppen haben, die von diesen Systemen betroffen sind, einschließlich derjenigen, die besondere Bedürfnisse oder Behinderungen haben oder bei deren Zugang zu Menschenrechten strukturelle Ungleichheiten bestehen könnten.“ Akteure des privaten Sektors spielen eine zunehmend wichtige Rolle bei den Aktivitäten im Lebenszyklus von KI-Systemen, insbesondere bei der Konzeption, Entwicklung und Einführung, und haben somit erhebliche Auswirkungen auf Nichtdiskriminierung und Gleichheit. Auf diese Verpflichtungen wird in Kapitel VII der Empfehlung und in den entsprechenden Abschnitten dieser Begründung näher eingegangen.

Zu Absatz 9

41. Im Zusammenhang mit Absatz 9 wären Richtlinien und Verhaltenskodizes zwei mögliche Mechanismen, mit denen private Akteure Maßnahmen zur Förderung der Gleichheit, einschließlich der Gleichheit der Geschlechter, im Rahmen ihrer Tätigkeiten während des gesamten Lebenszyklus von KI-Systemen ergreifen könnten. Die Mitgliedstaaten könnten sich von Artikel 95 Absatz 2 der KI-Verordnung über Verhaltenskodizes zur freiwilligen Anwendung spezifischer Anforderungen inspirieren lassen, der vorsieht, dass die Mitgliedstaaten die Ausarbeitung von Verhaltenskodizes zur freiwilligen Anwendung spezifischer Anforderungen auf alle KI-Systeme, auch durch die Betreiber, erleichtern sollten; dies auf der Grundlage klarer Ziele und Schlüsselindikatoren zur Messung der Zielerreichung, einschließlich Elementen, die unter anderem Folgendes umfassen: „die Förderung der KI-Kompetenz, insbesondere bei Personen, die mit der Entwicklung, dem Betrieb und der Nutzung von KI befasst sind“; „die Förderung einer inklusiven und vielfältigen Gestaltung von KI-Systemen, unter anderem durch die Bildung inklusiver und vielfältiger Entwicklungsteams und die Förderung der Beteiligung von Interessengruppen an diesem Prozess“; sowie „die Bewertung und Verhinderung negativer Auswirkungen von KI-Systemen auf schutzbedürftige Personen oder Gruppen schutzbedürftiger Personen, unter anderem im Hinblick auf die Barrierefreiheit für Menschen mit Behinderungen sowie auf die Gleichheit der Geschlechter“. Zu den weiteren Maßnahmen können externe Audits von KI-Systemen sowie frühzeitige Bewertungen der positiven Auswirkungen auf die Menschenrechte und der damit verbundenen Risiken gehören, ebenso wie Maßnahmen, bei denen die Standpunkte von Personen und Gruppen, die von Diskriminierung bedroht sind, einschließlich marginalisierter Gruppen, eingeholt und in die Gestaltung von KI-Systemen einbezogen werden.

Zu Absatz 10

42. In Absatz 10 werden die Mitgliedstaaten aufgefordert, öffentliche und private Anbieter und Betreiber von KI-Systemen zu ermutigen und zu unterstützen, die Gleichheit der Geschlechter und die Vielfalt in ihrer Belegschaft auf allen Ebenen zu fördern und zu wahren. Der Mangel an Vielfalt und Inklusion in Fachkreisen,

⁵⁹ Siehe ferner „Improving Artificial intelligence Standards for an Equitable Future“, überarbeitete Fassung, vorgelegt vom Sekretariat nach Konsultationen mit den Mitgliedstaaten, UNECE, 26. September 2025.

⁶⁰ Rahmenübereinkommen, Artikel 3 Absatz 1(a).

die an der Entwicklung von KI beteiligt sind, wurde als wichtiger Aspekt des Problems der Diskriminierung identifiziert, das sich aus Aktivitäten im Lebenszyklus von KI-Systemen ergibt.⁶¹ Dies kann dazu führen, dass KI-Systeme den Bedürfnissen von Menschen aus benachteiligten, Minderheiten- und unterrepräsentierten Gruppen nicht gerecht werden.⁶² Darüber hinaus kommen KI-Systeme zunehmend in Auswahlverfahren zum Einsatz, um über den Zugang zum Arbeitsmarkt zu entscheiden.⁶³ Es gibt Belege dafür, dass Einstellungsverfahren durch den Einsatz von KI-Systemen negativ beeinflusst werden können, beispielsweise in Fällen, in denen die Identifizierung von Bewerbern mithilfe von maschinellem Lernen zu diskriminierenden Ergebnissen führte oder in denen KI-gestützte Gesichtserkennungs- und Emotionsanalysensysteme zu rassistischer Diskriminierung geführt haben.⁶⁴ KI-Tools, die für die Rekrutierung und Auswahl von Bewerbern eingesetzt werden, sollten objektiv, neutral und frei von Vorurteilen, einschließlich geschlechtsspezifischer Vorurteile, sein. Die Arbeitsplatzrichtlinien von KI-Akteuren können auch spezifische Inhalte umfassen, die auf die Förderung von Inklusion und die Wiedereingliederung nach beruflichen Unterbrechungen ausgerichtet sind.⁶⁵ Einstellungs- und Beförderungsrichtlinien sollten darauf abzielen, eine ausgewogene Vertretung von Personen, die von Diskriminierung bedroht sind, in Entscheidungspositionen, einschließlich auf höheren Ebenen, zu erreichen und Programme zu unterstützen, die ihnen zum Erfolg verhelfen, einschließlich der Verpflichtung, Frauen einzustellen und zu fördern, um 50 % der Positionen auf der oberen und höchsten Ebene im Technologiesektor zu besetzen.⁶⁶ Ebenso sollten die Mitgliedstaaten sicherstellen, dass solche Branchen wie alle anderen Arbeitgeber gleiche Entlohnung für gleiche oder gleichwertige Arbeit gewährleisten, unter anderem durch die Beseitigung des geschlechtsspezifischen Lohngefälles und möglicher Lohnunterschiede, die mit anderen in Absatz 2 der Empfehlung aufgeführten Gründen zusammenhängen.⁶⁷ Darüber hinaus sollten die Mitgliedstaaten im Zusammenhang mit Absatz 10.2 öffentliche und private Akteure dazu ermutigen, Gleichheit und Vielfalt aktiv zu fördern. Dies kann *unter anderem* durch branchenbezogene und/oder unternehmensweite Maßnahmen zur Bekämpfung von Stereotypen, Sexismus, Rassismus, LGBTI-Feindlichkeit und anderen Formen von Intoleranz, Diskriminierung, Belästigung und Gewalt aufgrund persönlicher Merkmale und des Status erreicht werden. Die Mitgliedstaaten können zudem öffentliche und private Akteure dazu anregen, ihre Belegschaft für Voreingenommenheit und das Risiko von Diskriminierung in KI-Systemen sowie bei deren Training zu sensibilisieren, um strukturelle Ungleichheiten und Stereotypisierung zu bekämpfen und die durchgängige Berücksichtigung der Geschlechtergleichheit sowie der Gleichheit im Allgemeinen sicherzustellen.⁶⁸

Zu den Absätzen 11 und 12

43. In Absatz 12 der Empfehlung werden die Mitgliedstaaten dazu aufgefordert, die notwendigen Voraussetzungen zu schaffen, um die Ursachen für Voreingenommenheit in KI-Systemen zu beseitigen, indem sie eine inklusive und transformative Nutzung von KI-Systemen zur Förderung der Gleichheit, einschließlich der Gleichheit der Geschlechter, unterstützen. Wie Sandra Fredman erläutert, verfolgt transformative Gleichheit vier sich überschneidende Dimensionen: (i) die Beseitigung sozialer und wirtschaftlicher Benachteiligung; (ii) die Bekämpfung von Stigmatisierung, Stereotypisierung, Vorurteilen und Gewalt; (iii) die Stärkung der Teilhabe; und (iv) die Berücksichtigung von Unterschieden durch strukturelle Veränderungen.⁶⁹ Geschlechtertransformative Ansätze zielen darauf ab, geschlechtsspezifische Ungleichheit zu bekämpfen, indem sie schädliche Geschlechternormen, -rollen und -beziehungen transformieren und gleichzeitig auf eine gerechtere Umverteilung von Macht, Ressourcen und Dienstleistungen hinarbeiten.⁷⁰ Diese Ansätze fördern aktiv die Gleichheit und gehen explizit auf die zugrunde liegenden Ursachen geschlechtsspezifischer Ungleichheiten ein. Daher sollten sie auf den jeweiligen Kontext, in dem sie angewendet werden, zugeschnitten sein. Zwar zielen geschlechtergerechte Ansätze darauf ab, geschlechtsspezifische Ungleichheiten zu beseitigen, die auf diskriminierende Ursachen zurückzuführen sind, doch gehen sie nicht auf die Ursachen selbst ein. Im Gegensatz dazu berücksichtigt ein geschlechtersensibler Ansatz zwar Geschlechternormen, -rollen und -beziehungen, geht jedoch nicht auf die daraus resultierenden Ungleichheiten ein.⁷¹ Die Empfehlung zielt zudem auf Inklusivität ab; daher kann der transformative Ansatz auf die gesamte Bandbreite der in Absatz 2 genannten Diskriminierungsgründe angewendet werden.

⁶¹ Discriminating Systems: Gender, Race, and Power in AI – Report, [Sarah Myers West](#), AI Now Institute, April 2019.

⁶² CDADI/GEC-Studie (2023), S. 36-37.

⁶³ [Resolution 2343 \(2020\)](#) „Verhinderung von Diskriminierung durch den Einsatz künstlicher Intelligenz“, Absatz 1. Siehe auch [Empfehlung CM/Rec\(2020\)1](#) zu den menschenrechtlichen Auswirkungen algorithmischer Systeme, Absatz 8.

⁶⁴ CDADI/GEC Study (2023), S. 19-21.

⁶⁵ CDADI/GEC-Studie (2023), S. 89.

⁶⁶ *Ibid.*, S. 38, 89.

⁶⁷ Artikel 4 der revidierten Europäischen Sozialcharta (SEV Nr. 163); [Empfehlung CM/Rec\(2007\)17](#) zu Standards und Mechanismen zur Gleichstellung der Geschlechter, Absatz 28; sowie [Empfehlung Nr. R \(85\) 2](#) zum rechtlichen Schutz vor Diskriminierung aufgrund des Geschlechts.

⁶⁸ Siehe [Empfehlung CM/Rec\(2019\)1](#) zur Prävention und Bekämpfung von Sexismus, insbesondere II.E.4.

⁶⁹ Sandra Fredman, „Substantive Equality Revisited“ (2016) 14(3) Int'l J of Con L 712.

⁷⁰ Siehe [Fachvermerk „Gender-transformative approaches to achieve gender equality and sexual and reproductive health and rights“](#), UNFPA, 2023, S. 7, zitiert in der [Gender Equality Strategy 2024-2029](#), S. 45.

⁷¹ *Ibid.*

44. Die Empfehlung enthält eine nicht erschöpfende Liste von Beispielen für solche Initiativen und fordert die Mitgliedstaaten auf, die Bereitstellung angemessener Finanzmittel in Erwägung zu ziehen, wozu auch die Einrichtung spezieller Finanzierungsströme für KI-Forschung, Start-ups und Gemeinschaftsinitiativen gehören könnte, die von Frauen und von Diskriminierung bedrohten Personen geleitet werden. Entscheidend ist, dass solche Initiativen bereits in der frühesten Phase der KI-Entwicklung auf den Prinzipien des inklusiven Designs basieren, um sicherzustellen, dass die Systeme für alle nutzbar und vorteilhaft sind. Zu diesem Zweck sollten Anbieter und Betreiber von KI-Systemen dazu angehalten werden, partizipative Designverfahren umzusetzen, indem sie betroffene Personen, insbesondere aus unterrepräsentierten oder historisch diskriminierten Gruppen, in die Konzeptions-, Test- und Einführungsphase einbeziehen.

45. Ein in der Empfehlung genanntes Beispiel ist die Fähigkeit von KI-Systemen, bisher unentdeckte technische und soziale Verzerrungen in Datensätzen, Konzepten, Prozessen und Entscheidungskriterien sowie Muster diskriminierender Ergebnisse zu erkennen und zu beheben. Dies bezieht sich auf die Fähigkeit von KI-Systemen, große Datenmengen zu analysieren und auf entsprechende Aufforderung hin Verzerrungen zu erkennen, die für das menschliche Auge möglicherweise „unsichtbar“ sind. KI-Systeme können zudem dazu eingesetzt werden, die Barrierefreiheit von Dokumenten und anderen Ergebnissen zu verbessern, beispielsweise durch Funktionen zur Sprache-zu-Text-Umwandlung und Transkription sowie durch personalisierte Hilfsmittel zur Barrierefreiheit und zur sprachlichen Vereinfachung. Kleinere Start-ups und private Akteure entwickeln Instrumente in diesem Bereich, und die Empfehlung zielt darauf ab, die Voraussetzungen dafür zu schaffen, dass solche Forschungsarbeiten zu Ergebnissen führen. Technologien zur Bilderkennung mittels künstlicher Intelligenz könnten dazu genutzt werden, große Datenmengen zu analysieren und die Darstellung von Frauen und Minderheiten in verschiedenen Medienbereichen zu bewerten, darunter unter anderem Fernsehprogramme, Filme sowie Online- und Printwerbung.

46. Zweitens wird in der Empfehlung darauf hingewiesen, dass KI-Systeme dazu eingesetzt werden können, Hassreden im Internet zu identifizieren, deren Verbreitung zu verhindern und diese zu bekämpfen, einschließlich sexistischer, rassistischer und LGBTI-feindlicher Hassreden sowie, im weiteren Sinne, Stereotypisierung und technologiegestützter Gewalt gegen Frauen und Mädchen sowie gegen andere Personen und Gruppen aufgrund ihrer tatsächlichen oder vermeintlichen persönlichen Merkmale und ihres Status. Wie in der Begründung zur Empfehlung [CM/Rec\(2022\)16](#) dargelegt, hat Hassrede im Internet exponentiell zugenommen. In der Umfrage der FRA aus dem Jahr 2018 zur Wahrnehmung von Antisemitismus war die höchste Einzelhäufigkeit gemeldeter antisemitischer Belästigungen im Bereich der Cyberbelästigung zu verzeichnen. Darüber hinaus ergab die FRA-Umfrage aus dem Jahr 2012 zu Gewalt gegen Frauen, dass Cybermobbing gegen Frauen in der EU weit verbreitet ist: Jede zwanzigste Frau in der EU gab an, bereits Opfer von Cybermobbing geworden zu sein.⁷² Gemäß der Empfehlung CM/Rec(2022)16 sollten die Mitgliedstaaten den durch Online-Hassreden verursachten Schaden begrenzen, indem sie wirksame Rechtsvorschriften erlassen, um deren Veröffentlichung zu verhindern und deren Entfernung sicherzustellen, sofern sie bereits veröffentlicht wurden. Angesichts der großen Menge an Hassreden im Internet und der Geschwindigkeit, mit der sich diese verbreiten können, sind rasche Präventivmaßnahmen, einschließlich Sensibilisierungskampagnen, erforderlich, um deren Ausbreitung zu verhindern; daher fordert die Empfehlung CM/Rec(2022)16 die Mitgliedstaaten auf, dafür zu sorgen, dass auch Internetvermittler einen Beitrag zu dieser Aufgabe leisten. Diese Empfehlung sieht ferner vor, dass die Mitgliedstaaten dafür sorgen sollten, dass Mechanismen zur Meldung von Fällen von Hassrede im Internet an Behörden und private Akteure, einschließlich Internetvermittler, sowie klare Regeln für die Bearbeitung solcher Meldungen vorhanden sind. Solche Meldemechanismen sollten benutzerfreundlich, altersgerecht, leicht zugänglich, zeitnah, mehrsprachig und wirksam sein. Angemessenheit bedeutet, dass die Meldemechanismen auf die Art des Online-Missbrauchs zugeschnitten sind, einschließlich technologiegestützter Gewalt gegen Frauen und Mädchen, und dass dabei sichergestellt wird, dass relevante Formen des Schadens sowie deren geschlechtsspezifische Ausprägungen und Auswirkungen erfasst werden und dass geeignete Weiterverweisungswege zur Verfügung stehen. Darüber hinaus sollten die für die Bearbeitung von Meldungen zuständigen Personen für die besonderen Bedürfnisse von Opfern von Online-Missbrauch sensibilisiert werden, insbesondere von Frauen und Mädchen, die Zielscheibe technologiegestützter Gewalt sind.

47. Der Einsatz von KI-Systemen bei dieser Aufgabe kann die Interaktion von Menschen mit Hassinhalten im Internet einschränken und somit deren Kontakt mit solchen Inhalten verringern; zudem kann er die Moderationsrate und deren Effizienz verbessern. Die Aufgabe ist jedoch komplex und heikel, sowohl aufgrund der Schwierigkeiten bei der Identifizierung von Hass im Internet (beispielsweise aufgrund der Nutzung verschiedener Medien und der Erfassung des Kontexts) als auch aufgrund der Grundrechte, die durch übermäßige oder unzureichende Moderation berührt werden.⁷³ In schwierigen Fällen sollte die endgültige Entscheidung, Inhalte vor ihrer Veröffentlichung zu sperren oder nicht, daher von einem Menschen getroffen werden; allerdings bedeutet das schiere Ausmaß an Hass im Internet, dass KI-Systeme zunehmend zur Unterstützung dieser Aufgabe herangezogen werden müssen, soll diese effektiv bewältigt werden. Wie im

⁷² FRA, Bias in Algorithmen.

⁷³ Einen detaillierten Überblick über die im Zusammenhang mit der Inhaltsmoderation aufgeworfenen Fragen finden Sie in der Begründung zur Empfehlung CM/Rec(2026)2 des Ministerkomitees an die Mitgliedstaaten zur Rechenschaftspflicht bei technologiegestützter Gewalt gegen Frauen und Mädchen, Absätze [132–155].

Bericht der FRA über Voreingenommenheit und Algorithmen (*ibid*) festgestellt wird, stützen sich Algorithmen möglicherweise in hohem Maße auf bestimmte Wörter als Indikatoren für Anstößigkeit, die je nach Sprache und Modell erheblich variieren können. In diesem Bericht führen die Begriffe „Muslim“, „Homosexueller“ und „Jude“ zu deutlich höheren Einschätzungen hinsichtlich der Anstößigkeit als andere Begriffe. Es besteht daher die Befürchtung, dass solche Instrumente bestimmte Verzerrungen verstärken oder gar erst hervorrufen könnten. Die Studie befasst sich ferner mit dem Problem der Inhaltsmoderation, bei der unzureichende semantische Filter, die nach dem Vorbild von Mehrheitsgruppen entwickelt wurden, auf Minderheitengruppen angewendet werden, was zu einer Einschränkung der Meinungsäußerung führt, die die Kommunikationsmöglichkeiten von Minderheitengruppen über soziale Medien in ungerechtfertigter Weise einschränken könnte. So führt beispielsweise die stereotype Assoziation von Begriffen, die mit dem Wort „Lesbe“ in Verbindung gebracht werden, mit pornografischen Inhalten häufig zu sogenannten „Shadow-Bans“, die die Reichweite von Beiträgen in sozialen Medien einschränken, oder dazu, dass es gänzlich unmöglich ist, bestimmte Wörter in Kontonamen und Benutzernamen zu verwenden. Die einschränkenden Auswirkungen der Inhaltsmoderation haben schwerwiegende negative Folgen für die Sichtbarkeit und Reichweite von Organisationen, Aktivitäten und Veranstaltungen, die darauf abzielen, hassefüllten und diskriminierenden Narrativen entgegenzuwirken, die sich gegen Angehörige von Minderheitengemeinschaften richten.⁷⁴ Derzeit laufen zahlreiche Forschungsarbeiten, die darauf abzielen, Verzerrungen in verfügbaren universellen KI-Sprachwerkzeugen zu mindern, beispielsweise durch den Einsatz verschiedener Strategien zur Reduzierung gesellschaftlicher Verzerrungen im Ausgangsmodell.⁷⁵ Die Mitgliedstaaten werden dazu angehalten, die Finanzierung öffentlicher und privater Akteure sicherzustellen, die Initiativen in diesem Bereich untersuchen, um die Technologie zu verbessern und zu gewährleisten, dass solche Einschränkungen behoben werden können.

48. KI-Systeme können auch gezielt so konzipiert werden, dass sie marginalisierten, gefährdeten oder unterversorgten Bevölkerungsgruppen zugutekommen, beispielsweise durch die Verbesserung des Zugangs zu Informationen, Gütern und Dienstleistungen oder durch die Optimierung der automatisierten Übersetzung für Regional- oder Minderheitensprachen.⁷⁶ So ist beispielsweise der Sprachtechnologieplan der walisischen Regierung aus dem Jahr 2018 Teil ihrer umfassenderen Strategie zur Förderung des Gebrauchs der walisischen Sprache und zur Steigerung der Zahl der aktiven Sprecher. Der Plan unterstützt Projekte wie Apps zum Erlernen der walisischen Sprache für Kinder, eine walisische Version von OpenStreetMap, eine walisische Sprachdatenbank sowie die Entwicklung von Macsen, einem digitalen Assistenten in walisischer Sprache. Zudem fördert er den Einsatz bestehender sprachwissenschaftlicher Instrumente sowie die Organisation jährlicher Konferenzen und die Schaffung spezieller Forschungsstellen an Universitäten.⁷⁷ Es könnten auch KI-Instrumente entwickelt werden, um Formen struktureller Ungleichheit (wie beispielsweise multiple und intersektionale Formen der Diskriminierung sowie Stellvertreter-Diskriminierung) zu erkennen, einschließlich solcher, die durch den Einsatz von KI entstehen oder verschärft werden, sowie Umweltauswirkungen und digitale Ausgrenzung; diese Erkenntnisse könnten als Grundlage für proaktive politische Maßnahmen und Regulierung dienen.

49. Durch Technologie begünstigte Gewalt, insbesondere gegen Frauen und Mädchen sowie gegen Personen, die von Diskriminierung bedroht sind, ist aufgrund der rasanten Entwicklung digitaler Technologien und deren weitverbreiteter Nutzung zunehmend präsent. Generative KI, die die Erstellung hyperrealistischer Bilder oder Videos (digitale Fälschungen)⁷⁸ ermöglicht, wird überwiegend zur Produktion sexueller oder pornografischer Inhalte genutzt (95 % der im Umlauf befindlichen digitalen Fälschungen sexueller Art – ob Bilder oder Videos – sind laut mehreren Studien sexueller oder pornografischer Natur)⁷⁹. Auch wenn es sich um virtuelle Darstellungen handelt, sind die Auswirkungen auf die Opfer real, insbesondere wenn die Inhalte auf öffentlichen Online-Plattformen verbreitet werden. KI-Systeme können bestehende Voreingenommenheit, Stereotypen, Stigmatisierung und Vorurteile reproduzieren, aufrechterhalten und verschärfen und so zum Fortbestehen schädlicher Narrative und diskriminierender Folgen beitragen. Dies wiederum könnte zu einer Zunahme von Gewalt gegen Frauen im Internet und im realen Leben führen, die aufgrund von Sexismus, Stereotypen sowie tief verwurzelten geschlechtsspezifischen Ungleichheiten und Machtungleichgewichten, die KI-Systeme unbeabsichtigt verstärken könnten, überproportional oft zur Zielscheibe werden. Um diesen Risiken zu begegnen, sind proaktive Maßnahmen erforderlich, darunter die Erkennung und Minderung von

⁷⁴ So berichtet beispielsweise die Eurocentralasian Lesbian Community (EL'C), eine 2017 gegründete Organisation, die sich für die Rechte von LGBTI-Frauen einsetzt, dass sie den Begriff „lesbian“ nicht in ihrem Benutzernamen auf Facebook verwenden durfte, während andere Begriffe, die mit der LGBTI-Gemeinschaft in Verbindung stehen (wie „gay“ oder „queer“), sehr wohl verwendet werden durften. Siehe EL'C, 'Lesbophobia: An intersectional form of violence' (2021); verfügbar unter <https://europeanlesbianconference.org/wp-content/uploads/2021/10/Lesbophobia-3.pdf>

⁷⁵ FRA, Bias in Algorithms (*ibid*), unter Berufung auf Bolukbasi et al. (2016); Manzini et al. (2019); Zhao et al. (2018); Caliskan et al. (2017); Krause et al. (2020); und Dathathri et al. (2020).

⁷⁶ Siehe CDADI/GEC-Studie (2023), S. 77.

⁷⁷ „Facilitating the implementation of the European Charter on Regional and Minority Languages through artificial intelligence“, Sekretariat des Europarates der Europäischen Charta der Regional- oder Minderheitensprachen, Februar 2022, S. 6.

⁷⁸ Unter sexueller digitaler Fälschung versteht man das, was gemeinhin auch als „Deepfake-Pornografie“ bezeichnet wird. Aufgrund der Ungenauigkeit des Begriffs „Pornografie“ und der damit verbundenen Gefahr, den Opfern weiteren Schaden zuzufügen, wird in dieser Begründung der Begriff „sexuelle digitale Fälschung“ anstelle von „Deepfake-Pornografie“ verwendet. In ähnlicher Weise wird in dieser Begründung der Begriff „digitale Fälschung“ in einem nicht-sexuellen Kontext verwendet.

⁷⁹ Siehe die Studie „Tackling deepfakes in European policy“, Wissenschaftlicher Dienst des Europäischen Parlaments, Juli 2021.

Verzerrungen in KI-Systemen, die Verhinderung der Verbreitung und die Moderation von sexistischem Hass im Internet, das Verbot der Erstellung und Weitergabe intimer Bilder und digitaler Fälschungen ohne Einwilligung sowie die Bekämpfung allgemeinerer Muster technologieunterstützter Gewalt.⁸⁰

50. Zwar sind bewährte Verfahren wie die oben genannten Beispiele zu loben, doch sollte auch beachtet werden, dass es unerlässlich ist, eine übermäßige Abhängigkeit von KI als eigenständige Lösung zur Bekämpfung von Ungleichheit und Diskriminierung (Technosolutionismus) zu vermeiden. Soziale Probleme erfordern soziale Lösungen, und obwohl KI die Gleichheit, einschließlich der Gleichheit der Geschlechter, sowie die Nichtdiskriminierung fördern kann, sollte sie gut finanzierte und sorgfältig konzipierte Gleichheitsmaßnahmen ergänzen, anstatt diese zu ersetzen.⁸¹ Darüber hinaus sollten die Mitgliedstaaten bedenken, dass selbst bei der Umsetzung solcher positiven Initiativen der Einsatz von KI-Systemen nach wie vor das Risiko einer Diskriminierung birgt.

IV. Verhinderung von Diskriminierung

Zu Absatz 13

51. Dieses Kapitel enthält Leitlinien für die Mitgliedstaaten, um zu vermeiden, dass Vorurteile in KI-Systemen gefördert oder verfestigt werden, und somit das daraus resultierende Risiko einer Diskriminierung im Widerspruch zu Artikel 14 der Konvention und anderen Rechten aus der Konvention zu mindern. Es ist in Unterabschnitte gegliedert, die drei Grundsatzgruppen entsprechen, welche sich aus dem Rahmenübereinkommen (Artikel 8–10) ableiten. Bei allen Maßnahmen sollten die Mitgliedstaaten, wie in Kapitel II dargelegt, den in Absatz 6 der Empfehlung formulierten Grundsatz der Verhältnismäßigkeit anwenden, um diese Verpflichtungen entsprechend dem Ausmaß des Risikos für die Rechte der Betroffenen auf Gleichbehandlung und Nichtdiskriminierung zu bemessen, wobei die Größe, die Kapazitäten und die wirtschaftliche Stärke der Anbieter und Betreiber zu berücksichtigen sind. In Absatz 13 wird zudem ausdrücklich der stufenweise und differenzierte Ansatz hervorgehoben, wie er in Absatz 24 der vorliegenden Begründung erläutert wird.

52. Absatz 13 stützt sich auf den Wortlaut von Artikel 16 des Rahmenübereinkommens, nämlich „Maßnahmen zur Ermittlung, Bewertung, Verhütung und Minderung von Risiken durch Systeme künstlicher Intelligenz zu ergreifen oder aufrechtzuerhalten, indem tatsächliche und potenzielle Auswirkungen auf die Menschenrechte, die Demokratie und die Rechtsstaatlichkeit berücksichtigt werden“, und strebt einen umfassenden Ansatz an. Solche Maßnahmen könnten je nach Sachlage von den Mitgliedstaaten oder von Anbietern und Betreibern von KI-Systemen ergriffen oder aufrechterhalten werden.

Zu Absatz 14

53. In Bezug auf Absatz 14 werden die Mitgliedstaaten dazu aufgefordert, die neuen Formen algorithmischer oder KI-basierter Gefährdungen zu untersuchen, die durch den Einsatz von KI-Systemen entstehen können, und die Forschung zu Methoden für den rechtlichen Schutz vor solchen Gefährdungen zu fördern. Alle derartigen Untersuchungen sollten den Grundsätzen des Datenschutzes entsprechen, wie sie in Absatz 5.4 der Empfehlung dargelegt sind. Die Empfehlung umfasst sowohl von den Mitgliedstaaten durchgeführte Studien als auch die Unterstützung von Studien durch Forschende, einschließlich Wissenschaftler, oder andere private Akteure. Die Unterstützung beschränkt sich nicht nur auf die Finanzierung, sondern umfasst beispielsweise auch die Schaffung eines forschungsfreundlichen Umfelds. Der Einsatz von KI-Systemen kann zu einer sozialen Sortierung und Differenzierung führen, die die Grenzen des Diskriminierungsrechts in Frage stellen. Eine solche soziale Selektion kann neue Formen sozialer Gefährdung hervorrufen, beispielsweise indem bestimmte Gruppen systematisch schlechteren wirtschaftlichen Bedingungen ausgesetzt werden oder von bestimmten Gütern und Dienstleistungen ausgeschlossen werden. Diese neuen Formen der sozialen Differenzierung können zu „emergenter Diskriminierung“ und sozioökonomischen Ungleichheiten führen.⁸² Die Mitgliedstaaten werden dazu aufgefordert, klarzustellen, dass das Diskriminierungsverbot in Artikel 14 auch Mehrfachdiskriminierung, intersektionale Diskriminierung und mittelbare Diskriminierung, einschließlich Stellvertreter-Diskriminierung, umfasst, da es sich hierbei um Formen diskriminierender Schäden handelt, die durch KI-Systeme verursacht werden können, und unter anderem auch andere durch die digitale Kluft bedingte sozioökonomische Ungleichheiten zu untersuchen. Die Mitgliedstaaten werden ferner dazu angehalten, die Ergebnisse solcher Forschungsarbeiten zu veröffentlichen, um deren Verfügbarkeit für andere Mitgliedstaaten und die breite Öffentlichkeit, einschließlich der Sozialpartner und interessierten Forschenden, sicherzustellen.

⁸⁰ Zum Thema technologiegestützte Gewalt gegen Frauen und Mädchen wird auf die Empfehlung CM/Rec(2026)2 des Ministerkomitees an die Mitgliedstaaten zur Rechenschaftspflicht bei technologiegestützter Gewalt gegen Frauen und Mädchen verwiesen.

⁸¹ Siehe CDADI/GEC-Studie (2023), S. 78.

⁸² Siehe Matthias Leese, „The new profiling: Algorithms, black boxes, and the failure of antidiscriminatory safeguards in the European Union“, 45 SECURITY DIALOGUE 494–511, 501 (2014); Monique Mann & Tobias Matzner, „Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination“, 6 BIG DATA & SOCIETY, 5–6 (2019).

54. In diesem Zusammenhang werden die Mitgliedstaaten dazu angehalten, ihre politischen und rechtlichen Rahmenbedingungen hinsichtlich der Erhebung von Gleichheitsdaten zu überprüfen, um die Verfügbarkeit und Qualität dieser Daten zu verbessern, die für eine Reihe von Maßnahmen zum Schutz der Gleichheit von Nutzen sind, wie beispielsweise positive Maßnahmen, Überprüfungsmaßnahmen oder Profiling-Verfahren. Die Mitgliedstaaten können sich an dem Ansatz orientieren, der in der Begründung der Empfehlung CM/Rec(2024)1 zur Gleichheit von Frauen und Mädchen der Roma und der Traveller empfohlen wird (Absätze 233–240). In der Empfehlung wird jedoch den Einschränkungen Rechnung getragen, mit denen die Mitgliedstaaten in diesem Zusammenhang möglicherweise konfrontiert sind. So ist beispielsweise die Erhebung von Daten zu geschützten Merkmalen in einigen Mitgliedstaaten eingeschränkt, und selbst wenn Zugang zu solchen Daten besteht, können Einschränkungen hinsichtlich ihrer Verwendung gelten. In solchen Fällen könnten qualitative Forschungsmethoden oder soziodemografische Ersatzvariablen herangezogen werden. Die Mitgliedstaaten sollten bei der Genehmigung der Nutzung von Gleichheitsdaten die Datenschutzvorschriften einhalten; insbesondere sollten die Mitgliedstaaten die Vertraulichkeitsmaßnahmen in Bezug auf sensible Daten zur sexuellen Orientierung, zur Geschlechtsidentität und zum Geschlechtsausdruck sowie zu den Geschlechtsmerkmalen und zur Zugehörigkeit zu einer ethnischen, kulturellen, sprachlichen oder religiösen Gruppe verstärken, um jegliches Risiko einer Offenlegung oder eines Missbrauchs zu verhindern. Jegliche Entwicklungen in diesem Bereich sollten zudem in Absprache mit den Gruppen erfolgen, die am stärksten von Diskriminierung bedroht sind, um sicherzustellen, dass eine solche Datenerhebung nicht gegen den Willen dieser Gruppen durchgeführt wird.

Transparenz, Erklärbarkeit und Aufsicht in Bezug auf Gleichheit und Nichtdiskriminierung

Zu Absatz 15

55. Transparenz bezieht sich auf Offenheit und Klarheit bei der Steuerung von Aktivitäten im Lebenszyklus von KI-Systemen. Dies bedeutet, dass die Entscheidungsprozesse und die allgemeine Funktionsweise von KI-Systemen für die zuständigen KI-Akteure sowie, soweit erforderlich und angemessen, für relevante Interessengruppen nachvollziehbar und zugänglich sein sollten. In bestimmten Fällen könnte sich Transparenz auch auf die Bereitstellung zusätzlicher Informationen beziehen, darunter beispielsweise Informationen zu den verwendeten Algorithmen, Datensätzen und Entscheidungskriterien, vorbehaltlich wirtschaftlicher und urheberrechtlicher Belange, der nationalen Sicherheit sowie weiterer Erwägungen.⁸³ Im Idealfall umfasst Transparenz auch die Berücksichtigung des sozialen Kontexts, wie beispielsweise die Einbindung von Interessengruppen in den Entwicklungsprozess und die Ergebnisse etwaiger Folgenabschätzungen.

56. Erklärbarkeit ist ein wichtiger Bestandteil von Transparenz. KI-Systeme, die auf maschinellem Lernen (ML) oder Deep Learning (DL) basieren, nutzen Algorithmen, die im Rahmen ihres eigenen Trainingsprozesses erlernt wurden und nicht durch explizite menschliche Programmierung entstanden sind. Während des Trainingsprozesses können KI-Modelle neue Zusammenhänge zwischen bestimmten Eingabemerkmalen aufdecken und Entscheidungen oder Vorhersagen auf der Grundlage hochkomplexer Modelle treffen, die eine große Anzahl miteinander interagierender Parameter (möglicherweise Millionen) umfassen, was es selbst für KI-Expertinnen und -Experten schwierig macht, nachzuvollziehen, wie die entsprechenden Ergebnisse letztendlich zustande kommen.⁸⁴ In solchen Situationen können die Gründe, aus denen Systeme bestimmte Entscheidungen getroffen haben, sowohl für die Betreiber/Anbieter der Systeme als auch für die von den Systemen Betroffenen unklar sein. Der daraus resultierende „Black-Box“-Effekt könnte entweder zu unangebrachtem Vertrauen oder zu einer übermäßigen Abhängigkeit von KI-Systemen führen, was in beiden Fällen negative Folgen für den Einzelnen haben könnte.⁸⁵ Diese Undurchsichtigkeit erschwert nicht nur das Verständnis von Entscheidungen, sondern kann sich auch direkt auf den Einzelnen auswirken, da sie Mängel in KI-Systemen verbergen kann, wie beispielsweise das Vorhandensein von Voreingenommenheit, Ungenauigkeiten oder sogenannten „Halluzinationen“.

57. Erklärbarkeit bezieht sich daher auf die Fähigkeit, vorbehaltlich der technischen Machbarkeit und unter Berücksichtigung des allgemein anerkannten Stands der Technik hinreichend verständliche Erklärungen dafür zu liefern, warum ein KI-System Informationen bereitstellt und Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen trifft. Dies ist besonders entscheidend in sensiblen Bereichen wie dem Gesundheitswesen, dem Finanzwesen, der Migration, dem Grenzschutz und der Strafjustiz, in denen das Verständnis der Argumentation hinter Entscheidungen, die von einem KI-System getroffen oder unterstützt werden, sowohl für den Einzelnen als auch für eine angemessene gerichtliche Überprüfung, sofern erforderlich, von wesentlicher Bedeutung ist. In solchen Fällen könnte Transparenz beispielsweise in Form einer Liste von Faktoren zum Ausdruck kommen, die das KI-System bei der Bereitstellung von Informationen oder bei der Entscheidungsfindung berücksichtigt (siehe Absatz 60 des erläuternden Berichts zum Rahmenübereinkommen); diese Liste könnte auch aus Daten abgeleitete Schlussfolgerungen umfassen. Die

⁸³ Siehe erläuternder Bericht zum Rahmenübereinkommen, Absatz 57.

⁸⁴ TechDispatch: „Explainable Artificial Intelligence“, [Europäischer Datenschutzbeauftragter](#) (2023), unter Berufung auf Peters, U.: „Explainable AI lacks regulative reasons: why AI and human decision-making are not equally opaque“ (AI and Ethics 2023).

⁸⁵ TechDispatch: Explainable Artificial Intelligence (2023).

Notwendigkeit klarer Erklärungen wird auch durch Artikel 86 Absatz 1 der KI-Verordnung zum Recht auf Erläuterung der Entscheidungsfindung im Einzelfall unterstrichen, der vorsieht, dass „Personen, die von einer Entscheidung betroffen sind, die der Betreiber auf der Grundlage der Ausgaben eines in Anhang III aufgeführten Hochrisiko-KI-Systems [...] getroffen hat und die rechtliche Auswirkungen hat oder sie in ähnlicher Art erheblich auf eine Weise beeinträchtigt, die ihrer Ansicht nach ihre Gesundheit, ihre Sicherheit oder ihre Grundrechte beeinträchtigt, haben das Recht, vom Betreiber eine klare und aussagekräftige Erläuterung zur Rolle des KI-Systems im Entscheidungsprozess und zu den wichtigsten Elementen der getroffenen Entscheidung zu erhalten.“ In der KI-Verordnung gilt diese Verpflichtung nur für Hochrisiko-KI-Systeme; wie für alle Absätze der Empfehlung gelten jedoch das Diskriminierungsverbot und der Grundsatz der Verhältnismäßigkeit, die herangezogen werden können, um einen abgestuften Ansatz zu verfolgen, der auf dem Risiko und anderen kontextbezogenen Faktoren basiert.

58. Einige technische Informationen, die mit dem Ziel der Transparenz oder der Erklärbarkeit erstellt wurden, bleiben für die breite Öffentlichkeit möglicherweise unverständlich, können jedoch von Aufsichtsbehörden, Gleichheitsstellen oder Akteuren und Organisationen der Zivilgesellschaft, die in diesem Bereich tätig sind, verstanden werden, unter anderem durch die Einbeziehung von Expertinnen und Experten. Alle von den Mitgliedstaaten festgelegten Verpflichtungen und eingeführten Systeme sollten daher den unterschiedlichen Kenntnisständen der verschiedenen Informationsempfänger Rechnung tragen und gemäß diesen angepasst werden. Generell gilt: Wenn Informationen an natürliche Personen weitergegeben werden, sollten die Erläuterungen in einer einfachen und allgemein verständlichen Sprache verfasst sein, wobei komplexe Fachbegriffe nach Möglichkeit vermieden werden sollten.

59. Die in Absatz 15 dargelegten konkreten Maßnahmen sollen den Mitgliedstaaten dabei helfen, die Transparenz und Erklärbarkeit von KI-Systemen zu verbessern. Solche Grundsätze sollten bereits in der Entwurfs- und Entwicklungsphase berücksichtigt werden. Eine Verpflichtung zur Veröffentlichung statistischer Daten kann es interessierten Kreisen ermöglichen, die diskriminierenden Auswirkungen eines bestimmten Systems zu bewerten. Die Daten sollten aufgeschlüsselt werden, um aufzuzeigen, wie sich das System auf Frauen und die in Absatz 2 der Empfehlung genannten Gruppen auswirkt, einschließlich möglicher intersektionaler Diskriminierung. Gegebenenfalls sollten auch Daten zur Funktionsweise von KI-Systemen bereitgestellt werden, darunter, soweit möglich, die verwendeten Entscheidungskriterien sowie, sofern relevant, Informationen über die verwendeten Trainings- und Validierungsdaten und deren Verarbeitung. Transparenzpflichten sollten gegen den Schutz der Privatsphäre und den Schutz der personenbezogenen Daten, die Vertraulichkeit, Geschäftsgeheimnisse sowie nationale Sicherheitsvorschriften abgewogen werden. Insbesondere werden die Mitgliedstaaten dazu angehalten, gesetzliche Anforderungen für die Betreiber von KI-Systemen festzulegen, wonach diese Verbrauchern und öffentlichen Aufsichtsbehörden auf Anfrage eine aussagekräftige Erklärung vorlegen müssen, insbesondere im Rahmen von Rechtsstreitigkeiten.

60. Durchsetzungsmechanismen für Transparenz können unter anderem die Meldung des Einsatzes von KI in nationalen, regionalen und lokalen Behörden sowie in der Regierung umfassen; ein Verfahren, das es Unternehmen ermöglicht, im Rahmen ihrer Berichte über den verantwortungsvollen Einsatz von KI, einschließlich der Auswirkungen auf die Gleichheit, Rechenschaft abzulegen; sowie die Nutzung von Registern für KI-Systeme. [Die „Cities Coalition for Digital Rights“](#), ein Zusammenschluss von 39 Städten in der EU und den USA, hat ebenfalls Maßnahmen ergriffen, um sicherzustellen, dass Städte Technologien auf offene und transparente Weise nutzen. In ihrer Erklärung bekräftigt die Koalition mehrere allgemeine Grundsätze, darunter Transparenz, Rechenschaftspflicht und Nichtdiskriminierung bei Algorithmen. In der Praxis bedeutet dies, dass die Öffentlichkeit „Zugang zu verständlichen und korrekten Informationen über die technologischen, algorithmischen und Systeme künstlicher Intelligenz haben sollte, die ihr Leben beeinflussen“, und dass sie in der Lage sein sollte, „unfaire, voreingenommene oder diskriminierende Systeme zu hinterfragen und zu ändern“.

61. Darüber hinaus sollten betroffene Personen gemäß Artikel 15 Absatz 2 des Rahmenübereinkommens darüber informiert werden, wenn sie mit einer Maschine statt mit einem Menschen interagieren, soweit dies dem Kontext angemessen ist.⁸⁶ Absatz 15.3 orientiert sich an Artikel 50 Absatz 1 der KI-Verordnung: „Die Anbieter stellen sicher, dass KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass die betreffenden natürlichen Personen informiert werden, dass sie mit einem KI-System interagieren, es sei denn, dies ist aus Sicht einer angemessen informierten, aufmerksamen und verständigen natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich.“ Artikel 50 Absatz 2 sieht ferner vor, dass Anbieter von KI-Systemen, einschließlich KI-Systemen mit allgemeinem Verwendungszweck, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen, sicherstellen müssen, dass die Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet und als künstlich erzeugt oder manipuliert erkennbar sind. In diesem Zusammenhang heißt es in der Empfehlung ferner, dass die Mitgliedstaaten die Öffentlichkeit auch dafür sensibilisieren können, wie man die Kommunikation mit KI-Systemen erkennt oder wie man KI-Systeme identifiziert, die personenbezogene Daten in verschiedenen Bereichen verarbeiten. Wenn eine staatliche oder kommunale Behörde KI-Systeme bei ihrer

⁸⁶ Siehe beispielsweise [„Administration and You. A Handbook“](#), 3. Auflage, „Principles of administrative law concerning relations between individuals and public authorities“, Europarat, S. 25.

Arbeit einsetzt, können diese Behörden gegebenenfalls in Erwägung ziehen, ihre Mitarbeiterinnen und Mitarbeiter darin zu schulen, Erläuterungen zu den Kriterien zu geben, die von KI-Systemen bei der Entscheidungsfindung herangezogen werden.

Zu Absatz 16

62. Öffentliche Register, in denen KI-Systeme aufgeführt sind, würden die Transparenz, die Rechenschaftspflicht und eine fundierte öffentliche Kontrolle des Einsatzes von KI-Systemen verbessern, die Risiken für die Gleichheit, einschließlich der Gleichheit der Geschlechter, mit sich bringen können. Die KI-Verordnung sieht bereits die Einrichtung einer Datenbank zur Registrierung von Hochrisiko-KI-Systemen vor, die in bestimmten Bereichen eingesetzt werden, sowie deren Einsatz durch öffentliche Stellen. Die Stadt Amsterdam hat beispielsweise ein solches freiwilliges Register eingeführt, um einen Überblick über die von der Stadt auf kommunaler Ebene eingesetzten KI-Systeme und Algorithmen zu verschaffen.⁸⁷ Die Einrichtung neuer Register kann für die Mitgliedstaaten einen erheblichen Verwaltungs- und Finanzaufwand mit sich bringen. In der Empfehlung wird berücksichtigt, dass einige Mitgliedstaaten möglicherweise bereits über Rahmenbedingungen verfügen, mit denen dieser Empfehlung Rechnung getragen werden kann, beispielsweise durch von Datenschutzbehörden geführte Register oder durch andere Register, die nicht speziell auf KI-Systeme ausgerichtet sind, die jedoch angepasst werden könnten, um diese einzubeziehen und so Kosten und Verzögerungen zu verringern.

Zu Absatz 17

63. Darüber hinaus zielt die Empfehlung darauf ab, die Fähigkeit der betroffenen Personen zu verbessern, Diskriminierung zu erkennen und die Schwierigkeiten zu überwinden, die sich aus einer Kombination aus Personalisierung, Automatisierung und Undurchsichtigkeit der Entscheidungsprozesse ergeben. Der Vergleich mit Personen in ähnlicher Situation sowie soziale Interaktionen sind wichtige heuristische Hilfsmittel, um Diskriminierung zu erkennen und geltend zu machen. Dennoch ist es beispielsweise im Rahmen eines Online-Kreditdienstes nahezu unmöglich geworden, soziale Signale zu deuten oder einen Vergleich mit anderen Kreditantragstellern anzustellen. Durch diese Informationsasymmetrie ist es von vornherein schwierig, Fälle von Diskriminierung zu identifizieren. Selbst wenn ein Verdacht aufkommt, stellt die Beweissicherung eine weitere Herausforderung dar, da die Entscheidungen oder die ihnen zugrunde liegenden Empfehlungen nicht ohne Weiteres einsehbar sind und von den Nutzenden von KI-Systemen oft nicht offenbart werden. Daher stellt die Vorlage von Beweisen zur Begründung einer Diskriminierungsvermutung vor Gericht eine zentrale rechtliche Herausforderung dar. Den Mitgliedstaaten wird daher ebenfalls empfohlen, dafür zu sorgen, dass die Verpflichtungen zur Dokumentation von Aktivitäten im Lebenszyklus von KI-Systemen so gestaltet sind, dass betroffene Personen Zugang zu diesen Informationen erhalten und diese im Rahmen von Diskriminierungsklagen nutzen können, um Informationsasymmetrien auszugleichen. Die Empfehlung stützt sich auf Artikel 86 des oben genannten KI-Verordnungen. In der Empfehlung wird der weit gefasste Begriff „Geltendmachung von Rechten“ verwendet. Dies kann beispielsweise die Anfechtung von Entscheidungen oder die Einreichung von Beschwerden umfassen, wie in Absatz 100 des erläuternden Berichts zum Rahmenübereinkommen dargelegt. Beschwerdemechanismen sollten leicht zugänglich sein.

Zu Absatz 18

64. Die Empfehlung geht davon aus, dass das Recht auf Verfahrensgarantien und wirksame Rechtsbehelfe durch die Konvention sowie durch andere internationale und nationale Rechtsvorschriften und Standards im Bereich der Diskriminierung gewährleistet ist, und verweist auf deren Anwendung im Zusammenhang mit KI-Systemen. Wird eine Entscheidung durch ein KI-System getroffen, empfiehlt die Empfehlung den Mitgliedstaaten, die Anbieter und Betreiber von KI-Systemen dazu anzuhalten, eine menschliche Überprüfung dieser Ergebnisse auf Antrag einer Person zu ermöglichen, die den Verdacht hegt, dass eine Diskriminierung vorliegt. Eine solche Möglichkeit wäre insbesondere dann von Bedeutung, wenn die Ausübung der Menschenrechte einer Person erheblich beeinträchtigt wird. Die menschliche Überwachung sollte substanziell sein und die aktive Einbindung mindestens eines geschulten und angemessen ausgestatteten menschlichen Bedieners beinhalten, der das Ergebnis überwachen und die Entscheidungen des Systems bewerten kann und der in der Lage ist, bei Bedarf so einzugreifen, dass die Qualität der vom System getroffenen Entscheidungen verbessert wird, und nicht nur als rein verfahrenstechnische Formalität oder symbolische Geste.⁸⁸ Wichtig ist, dass menschliche Aufsicht zwar eine zusätzliche Ebene der Überprüfung der Entscheidungsfindung bieten und die Rechenschaftspflicht verbessern kann, jedoch Probleme im Zusammenhang mit systemischen Verzerrungen nicht lösen kann und keine Alternative zu einer Konzeption darstellen sollte, die die Auswirkungen von KI-Systemen auf die Gleichheit, einschließlich der Gleichheit der Geschlechter, berücksichtigt. Menschliche Überwachung ist nicht unfehlbar, und es muss darauf geachtet werden, Automatisierungsverzerrungen zu

⁸⁷ [Algorithmisches Register der Stadt Amsterdam](#), Gemeente Amsterdam.

⁸⁸ Siehe Europäischer Datenschutzbeauftragter, TechDispatch 2/2025 – „Human Oversight of Automated Decision-Making“.

vermeiden.⁸⁹ Darüber hinaus könnten Einzelpersonen, abgesehen von Rechtsstreitigkeiten, wie auch in anderen Kontexten, in denen keine KI-Systeme zum Einsatz kommen, über Verfahren informiert werden, die es ihnen ermöglichen, gegen vermutete Diskriminierung durch ein KI-System vorzugehen, beispielsweise durch Einreichen einer Beschwerde oder durch die Beantragung von Unterstützung bei einer Gleichheitsstelle, einer nationalen Menschenrechtsinstitution, einem Ombudsmann oder einer Verbraucherschutzstelle. In diesem Fall sollten diese Informationen in einem klaren und zugänglichen Format bereitgestellt werden.

Zu Absatz 19

65. Schließlich wird in der Empfehlung festgestellt, dass Beweise für die Beurteilung, ob eine Diskriminierung vorliegt, von entscheidender Bedeutung sind und sich diese oft nicht im Besitz der betroffenen Person befinden. Zur Beseitigung von Informationsasymmetrien und zur Aufklärung über die umfassendere Rolle von Gleichheitsstellen, nationalen Menschenrechtsinstitutionen, Ombudsmännern, zuständigen Behörden und Organisationen von öffentlichem Interesse fordert die Empfehlung die Mitgliedstaaten in Absatz 18 auf, zu prüfen, ob diesen Stellen das Recht auf Zugang zu Informationen sowie auf das Testen von KI-Systemen eingeräumt werden soll. In Bezug auf Gleichheitsstellen stützt sich die Empfehlung auf die Allgemeine politische Empfehlung Nr. 2 (in der überarbeiteten Fassung) der ECRI [zu Gleichheitsstellen zur Bekämpfung von Rassismus und Intoleranz auf nationaler Ebene](#),⁹⁰ die [Richtlinie 2024/1499](#)⁹¹ und die [Richtlinie 2024/1500](#)⁹², die vorsehen, dass die Mitgliedstaaten sicherstellen müssen, dass Gleichbehandlungsstellen befugt sind, Untersuchungen darüber durchzuführen, ob ein Verstoß gegen den in verschiedenen Richtlinien festgelegten Grundsatz der Gleichbehandlung vorliegt, insbesondere durch die Schaffung eines Rahmens für die Durchführung von Untersuchungen, die es den Gleichbehandlungsstellen ermöglicht, Sachverhalte zu ermitteln, ihnen wirksame Rechte auf Zugang zu Informationen und Unterlagen zu gewähren, die zur Feststellung einer möglichen Diskriminierung erforderlich sind, sowie Mechanismen bereitzustellen, die es den Gleichbehandlungsstellen ermöglichen, zu diesem Zweck mit den zuständigen öffentlichen Stellen zusammenzuarbeiten. Ebenso sieht Artikel 77 Absatz 1 der KI-Verordnung über die Befugnisse der Behörden zum Schutz der Grundrechte vor, dass „[N]ationale Behörden oder Stellen, die die Einhaltung der Verpflichtungen nach dem Unionsrecht zum Schutz der Grundrechte, einschließlich des Rechts auf Nichtdiskriminierung, im Zusammenhang mit der Nutzung der in Anhang III genannten Hochrisiko-KI-Systeme überwachen oder durchsetzen, befugt sind, alle im Rahmen dieser Verordnung erstellten oder geführten Unterlagen in einer zugänglichen Sprache und einem zugänglichen Format anzufordern und einzusehen, sofern der Zugang zu diesen Unterlagen für die wirksame Erfüllung ihrer Aufgaben im Rahmen ihrer Zuständigkeit erforderlich ist“. Erwägungsgrund 157 der Präambel besagt: „nationale Behörden oder Stellen, die die Anwendung des Unionsrechts zum Schutz der Grundrechte überwachen, einschließlich der Anwendung des Unionsrechts zum Schutz der Grundrechte, darunter Gleichstellungsstellen und Datenschutzbehörden. Soweit dies für die Erfüllung ihres Auftrags erforderlich ist, sollten diese nationalen Behörden oder Stellen auch Zugang zu allen im Rahmen dieser Verordnung erstellten Unterlagen haben.“

66. Es wird empfohlen, diesen Stellen die Durchführung von Tests zu gestatten, insbesondere durch die Teilnahme an koordinierten „Regulatory Sandboxes“ für KI. Eine sogenannte „Regulatory Sandbox“ ist ein Instrument, das es ermöglicht, unter der Aufsicht einer Regulierungsbehörde in einem kontrollierten Umfeld neue und innovative Produkte, Dienstleistungen oder Geschäftsmodelle zu erproben und damit zu experimentieren. Wenn solche Gremien eingerichtet werden, sollten sie konsultiert werden und die Möglichkeit erhalten, Leitlinien im Bereich der Menschenrechte zu erarbeiten und die Einbindung von Menschenrechtsgarantien in den gesamten Lebenszyklus der KI zu fördern.⁹³

Folgenabschätzung, Zertifizierung, Bewertung und Evaluierung im Bereich Gleichheit und

Nichtdiskriminierung

Zu Absatz 20

⁸⁹ Gemäß dem Erläuternden Bericht zum Rahmenübereinkommen, Absatz 75 e): „Automatisierungs- oder Bestätigungsverzerrungen, bei denen Menschen ungerechtfertigtes Vertrauen in Maschinen und technologische Artefakte setzen oder Situationen, in denen sie Informationen auswählen, die ihre eigenen Ansichten stützen, wobei sie in beiden Fällen ihre eigenen, möglicherweise widersprüchlichen Urteile außer Acht lassen und algorithmische Ergebnisse ohne Hinterfragung bestätigen.“

⁹⁰ Siehe ECRI-GPR [Nr. 2 über Gleichstellungsstellen zur Bekämpfung von Rassismus und Intoleranz auf nationaler Ebene](#), Absatz 21.

⁹¹ Richtlinie (EU) 2024/1499 des Rates vom 7. Mai 2024 über Standards für Gleichbehandlungsstellen im Bereich der Gleichbehandlung von Personen ungeachtet ihrer Rasse oder ethnischen Herkunft, der Gleichbehandlung in Beschäftigung und Beruf ungeachtet ihrer Religion oder Weltanschauung, einer Behinderung, des Alters oder der sexuellen Ausrichtung, der Gleichbehandlung von Frauen und Männern in Fragen der sozialen Sicherheit sowie beim Zugang zu und der Bereitstellung von Waren und Dienstleistungen, und zur Änderung der Richtlinien 2000/43/EG und 2004/113/EG.

⁹² Richtlinie (EU) 2024/1500 des Europäischen Parlaments und des Rates vom 14. Mai 2024 über Standards für Gleichstellungsstellen im Bereich der Gleichbehandlung und der Chancengleichheit von Frauen und Männern in Beschäftigung und Beruf sowie zur Änderung der Richtlinien 2006/54/EG und 2010/41/EU.

⁹³ [„Human Rights by Design – Future-Proofing Human Rights Protection in the Era of AI“](#) (2023), S. 18–20, als Fortsetzung von [„Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights“](#) (2019), S. 35.

67. Dieser Unterabschnitt konzentriert sich auf die Bewertung und Evaluierung von KI-Systemen in Bezug auf Gleichheit und Nichtdiskriminierung. Es gibt mehrere bestehende Standards, an denen sich die Mitgliedstaaten bei der Einführung einer Risiko- und Folgenabschätzung orientieren können, deren Schwerpunkt speziell auf Gleichheit und Nichtdiskriminierung liegt. Die Empfehlung CM/Rec(2020)1 sieht in Absatz 5.2 vor, dass die Mitgliedstaaten sicherstellen sollten, dass sie selbst sowie alle privaten Akteure, die mit ihnen oder in ihrem Auftrag tätig sind, regelmäßig und in Absprache mit den Betroffenen Menschenrechtsverträglichkeitsprüfungen durchführen, und zwar vor der Vergabe öffentlicher Aufträge, während der Konzeption und Entwicklung, an regelmäßigen Meilensteinen sowie während des gesamten kontextspezifischen Einsatzes, um die Risiken von menschenrechtsbeeinträchtigenden Folgen zu ermitteln. Die Empfehlung CM/Rec(2022)17 zum Schutz der Rechte von migrierten, geflüchteten und asylsuchenden Frauen und Mädchen sieht in Absatz 22 vor, dass die Mitgliedstaaten sicherstellen sollten, dass vor der Einführung von KI und automatisierten Entscheidungssystemen im Bereich der Migration eine Menschenrechtsverträglichkeitsprüfung unter Berücksichtigung der Gleichheitsperspektive durchgeführt wird. Ebenso enthält Artikel 16 des Rahmenübereinkommens eine spezielle Bestimmung, die vorschreibt, dass die relevanten Risiken und potenziellen Auswirkungen auf Menschenrechte, Demokratie und Rechtsstaatlichkeit im *Vorfeld* und iterativ über den gesamten Lebenszyklus eines KI-Systems hinweg ermittelt, bewertet, verhindert und gemindert werden müssen, und zwar durch die Entwicklung einer Methodik mit konkreten und objektiven Kriterien für solche Bewertungen. Die HUDERIA (Methodik zur Risiko- und Folgenabschätzung von KI-Systemen unter dem Gesichtspunkt der Menschenrechte, der Demokratie und der Rechtsstaatlichkeit) ist eine Methodik, die herangezogen werden kann, und die Mitgliedstaaten sollten darauf achten, dass Aspekte der Gleichbehandlung und Nichtdiskriminierung bei der Anwendung dieser Methodik durchgängig berücksichtigt werden. Solche *Vorabverpflichtungen* zur Rechenschaftspflicht und Begründung könnten beispielsweise erfordern, dass Betreiber und Anbieter von KI-Systemen vorab Bewertungen des Diskriminierungsrisikos und der Auswirkungen auf die Gleichheit durchführen, sei es eigenständig oder im Rahmen einer umfassenderen Menschenrechtsverträglichkeitsprüfung. Diese Folgenabschätzungen könnten geschlechtsspezifische Folgenabschätzungen umfassen oder durch solche ergänzt werden und, je nach Kontext, auch Folgenabschätzungen im Hinblick auf die Kinderrechte beinhalten. Gleichheitsfolgenabschätzungen sollten idealerweise eine intersektionale Analyse sowie aufgeschlüsselte Gleichheitsdaten umfassen, wobei zu berücksichtigen ist, wie diskriminierende Auswirkungen mit den in Absatz 2 der Empfehlung aufgeführten Gründen sowie mit anderen strukturellen und sozialen Faktoren zusammenwirken. Die Aufschlüsselung der Daten ermöglicht es Forschenden, die Auswirkungen von Ungleichheiten zwischen verschiedenen Bevölkerungsgruppen zu vergleichen. Der einfache Durchschnitt von Daten kann zugrunde liegende Ungleichheiten verschleiern; aufgeschlüsselte Daten hingegen können unterschiedliche Auswirkungen auf die Menschenrechte zwischen verschiedenen Gruppen aufzeigen. So können nationale Durchschnittswerte beispielsweise zeigen, dass die Erwerbsbevölkerung insgesamt Löhne erzielt, die zur Deckung der Lebenshaltungskosten ausreichen; detaillierte Daten hingegen können aufzeigen, dass Frauen deutlich weniger verdienen als ihre männlichen Kollegen, was sich auf ihre Fähigkeit auswirkt, sich Lebensmittel, Wohnraum und andere lebensnotwendige Güter leisten zu können.⁹⁴ Im Einklang mit Absatz 5.4 der Empfehlung sollten bei diesem Prozess die Rechte auf Privatsphäre und Datenschutz gewahrt werden.

68. Die Hauptelemente einer solchen Bewertung könnten sein:

1. Feststellung der Auswirkungen: Ermittlung von Möglichkeiten, das KI-System zur Unterstützung positiver Maßnahmen einzusetzen, die Gleichheit und Nichtdiskriminierung fördern, sowie Ermittlung der damit verbundenen Risiken für die Gleichheit, einschließlich der Gleichheit der Geschlechter, und die Nichtdiskriminierung;
2. Folgenabschätzung: Bewertung der Auswirkungen unter Berücksichtigung der Wahrscheinlichkeit und des Ausmaßes sowohl positiver als auch negativer Auswirkungen auf die Gleichheit;
3. Bewertung der Steuerungsmechanismen: Bewertung der Rollen und Verantwortlichkeiten von Pflichtenträgern, Rechteinhabern und Interessengruppen bei der Umsetzung und Steuerung der Mechanismen zur Stärkung positiver Auswirkungen und zur Minderung etwaiger negativer Auswirkungen;
4. Reaktion, Risikominderung und Bewertung: Ermittlung geeigneter Maßnahmen zur Verstärkung positiver Auswirkungen, Maßnahmen zur Risikominderung sowie die Gewährleistung einer kontinuierlichen Bewertung.

69. Die Mitgliedstaaten werden dazu angehalten, solche Bewertungen rechtsverbindlich zu machen oder einen starken Anreiz zu schaffen, indem sie solche Folgenabschätzungen zu einem Element machen, das in Gerichtsverfahren im Zusammenhang mit Diskriminierungsklagen, bei denen KI-Systeme eine Rolle spielen, zu berücksichtigen ist, beispielsweise bei der Beurteilung, ob positive Verpflichtungen hinsichtlich der Gleichbehandlung erfüllt wurden, oder bei der Festlegung der Beweislast. Die Ergebnisse von Diskriminierungsrisikobewertungen und Gleichheitsfolgenabschätzungen könnten zudem darüber entscheiden, ob ein bestimmtes System auf den Markt gebracht und genutzt werden darf. Ein solches „Vorabgenehmigungssystem“ verpflichtet Anbieter und Nutzende potenziell risikobehafteter Systeme dazu,

⁹⁴ [Danish Institute for Human Rights: Human rights impact assessment guidance and toolbox Report A4 \(S. 67\).](#)

vorbeugende Maßnahmen zu ergreifen und ihre Systeme gegebenenfalls einer Zertifizierung oder einer Zulassung zu unterziehen. Der Rechtsgrundsatz der Verhältnismäßigkeit kann dabei helfen, diese Verpflichtungen an die Größe, die Kapazität und die wirtschaftliche Stärke der Anbieter und Nutzende anzupassen; zudem sollten die in Absatz 23 dieser Begründung dargelegten Erwägungen ebenfalls in die Abwägung einfließen,⁹⁵ wobei das Ausmaß der positiven Auswirkungen und das festgestellte Diskriminierungsrisiko zu berücksichtigen sind. Solche Risiko- und Folgenabschätzungen sollten gegebenenfalls veröffentlicht und leicht zugänglich gemacht werden. Sie sollten die potenziellen diskriminierenden Auswirkungen von KI-Systemen über deren gesamten Lebenszyklus hinweg und in Bezug auf alle in Absatz 2 der Empfehlung genannten geschützten Merkmale bewerten. Eine Möglichkeit, die Aufsicht zu gewährleisten, könnte darin bestehen, eine Gleichheitsstelle oder eine Regulierungsbehörde zu befugen, zu überwachen, ob der Betrieb von KI-Systemen Einzelpersonen benachteiligt und während ihres gesamten Lebenszyklus zu struktureller Ungleichheit beiträgt.

70. Die Mitgliedstaaten könnten sich zudem am Risikomanagementsystem orientieren, das in Artikel 9 der KI-Verordnung vorgesehen ist und besagt, dass „ein Risikomanagementsystem für Hochrisiko-KI-Systeme eingerichtet, umgesetzt, dokumentiert und aufrechterhalten werden muss“. Dieses System „ist als ein kontinuierlicher, iterativer Prozess zu verstehen, der über den gesamten Lebenszyklus eines Hochrisiko-KI-Systems hinweg geplant und durchgeführt wird und eine regelmäßige, systematische Überprüfung und Aktualisierung erfordert“. Bei der Umsetzung des Risikomanagementsystems müssen die Anbieter prüfen, ob das Hochrisiko-KI-System angesichts seines Verwendungszwecks voraussichtlich nachteilige Auswirkungen auf Personen unter 18 Jahren und gegebenenfalls auf andere schutzbedürftige Gruppen haben könnte. Darüber hinaus sieht Artikel 79 Absatz 2 über das Verfahren auf nationaler Ebene zur Behandlung von KI-Systemen, die ein Risiko darstellen, vor, dass die nationale Marktüberwachungsbehörde jedes KI-System untersuchen soll, bei dem der Verdacht besteht, dass es ein Risiko im Sinne von Artikel 3 Nummer 19 der Verordnung (EU) 2019/1020 darstellt, sofern es Risiken für die Gesundheit oder Sicherheit oder für die Grundrechte von Personen birgt. Hier heißt es: „Besondere Aufmerksamkeit ist KI-Systemen zu widmen, die ein Risiko für schutzbedürftige Gruppen darstellen.“ Darüber hinaus gilt: „Werden Risiken für Grundrechte festgestellt, so unterrichtet die Marktüberwachungsbehörde auch die in Artikel 77 Absatz 1 genannten zuständigen nationalen Behörden oder Stellen und arbeitet uneingeschränkt mit diesen zusammen.“ Artikel 27 des KI-Verordnungs sieht zudem eine Folgenabschätzung hinsichtlich der Grundrechte für Hochrisiko-KI-Systeme vor.

71. Einige Länder, beispielsweise die Niederlande,⁹⁶ haben bereits umfassende Folgenabschätzungen zu Grundrechten und Algorithmen entwickelt, um die Rechenschaftspflicht für die im Alltag zunehmend eingesetzten KI-Systeme sicherzustellen.

72. In Absatz 20.3 wird zudem die Einführung von Standards für Diskriminierungsaudits bei KI-Systemen behandelt. Normen zur Bewertung von Diskriminierung, einschließlich der Audits, können technische Normen umfassen. Dabei handelt es sich um vereinbarte Regelwerke, die sicherstellen, dass Produkte, Dienstleistungen oder Prozesse sicher, zuverlässig und kompatibel sind, und als bewährte Verfahren der Branche gelten. Dazu gehören Normen wie die der Internationalen Organisation für Normung (ISO), der Internationalen Elektrotechnischen Kommission (IEC), des Europäischen Komitees für Normung (CEN), das Europäische Komitee für elektrotechnische Normung (CENELEC) oder das Europäische Institut für Telekommunikationsnormen (ETSI). Neben technischen Standards, die als Grundlage für den Bewertungsprozess dienen, könnten auch soziotechnische oder andere Standards entwickelt werden. In der Empfehlung werden die Mitgliedstaaten aufgefordert, die Ausarbeitung von Standards für Diskriminierungsaudits bei KI-Systemen in die Wege zu leiten. Dies könnte dazu beitragen, einen einheitlichen und wirksamen Ansatz zur Bekämpfung von Voreingenommenheit zu gewährleisten, die sich aus Aktivitäten im Lebenszyklus von KI-Systemen ergibt, und gleichzeitig Innovationen bei der Entwicklung technischer Lösungen und politischer Maßnahmen zur Bekämpfung von Diskriminierung im Zusammenhang mit KI-Systemen zu fördern. Dies würde auch das Vertrauen der Öffentlichkeit in KI-Systeme stärken. Harmonisierte Regulierungsstandards würden zudem die Verantwortung der Unternehmen hinsichtlich der Diskriminierungsrisiken stärken, was wiederum potenziellen Opfern nützliche Informationen für die Beurteilung der Möglichkeit, (rechtliche) Schritte einzuleiten, sowie den Richtern in Diskriminierungsverfahren verständliches Beweismaterial liefern würde.

73. Externe Audits hinsichtlich Diskriminierungsrisiken sollten von unabhängigen Stellen durchgeführt werden, während dies bei internen Audits nicht zwingend erforderlich ist. Audits auf Diskriminierungsrisiken, insbesondere wenn sie von Dritten durchgeführt werden, können Vorurteile aufdecken und dabei helfen, diese zu beseitigen, die an verschiedenen Stellen in den Lebenszyklus eines KI-Systems einfließen können; dadurch lassen sich das Risiko von Diskriminierung und die Gefahr der Aufrechterhaltung systemischer Ungleichheit mindern. Die Ergebnisse externer Audits sowie die Ergebnisse von Folgenabschätzungen sollten, soweit angemessen, öffentlich zugänglich gemacht werden.

⁹⁵ Unter besonderer Bezugnahme auf die Absätze 57–65 des erläuternden Berichts zum Rahmenübereinkommen.

⁹⁶ Siehe hierzu auch, [„Fundamental Rights and Algorithms Impact Assessment \(FRAIA\)“, Ministerium für Inneres und Königsbeziehungen](#), Niederlande (März 2022).

74. Die Mitgliedstaaten sollten zudem prüfen, ob zusätzlich Zertifizierungsmechanismen eingesetzt werden könnten, um sicherzustellen, dass das positive Potenzial von KI-Systemen ausgeschöpft wurde und dass Verzerrungen gemindert sowie Diskriminierungsrisiken für genau definierte Anwendungszwecke so weit wie möglich beseitigt wurden. Solche *Vorsorgemaßnahmen* könnten dazu dienen, den nichtdiskriminierenden Charakter domänenspezifischer Systeme zu gewährleisten. Diese Zertifizierungen sollten von unabhängigen, akkreditierten Stellen ausgestellt werden, deren Mitarbeiterinnen und Mitarbeiter über qualifizierte Fachkräfte verfügen, die über Erfahrung in den Bereichen KI, Menschenrechte und Gleichheit verfügen. Darüber hinaus sollten Zertifizierungen nach einem angemessenen Zeitraum regelmäßig erneuert werden, um die kontinuierliche Einhaltung der einschlägigen Standards über einen längeren Zeitraum hinweg sicherzustellen und zu gewährleisten, dass zertifizierte Systeme die Grundsätze der Gleichbehandlung und Nichtdiskriminierung fortlaufend einhalten. Die Zertifizierung sollte in gewissem Maße öffentlich bekannt gemacht werden. So sollten Zertifizierungen beispielsweise für Nutzende und Endnutzende von KI-Systemen (im Hinblick auf mögliche rechtliche Verteidigung und Risikobewertungen für bestimmte Anwendungen) sowie für öffentliche Einrichtungen (beispielsweise Gleichheitsstellen, nationale Menschenrechtsinstitutionen, Datenschutzbeauftragte oder andere zuständige Behörden) und für Organisationen der Zivilgesellschaft mit einem berechtigten Interesse (im Sinne der EU-Gleichbehandlungsrichtlinien und des nationalen Gleichbehandlungsrechts) zugänglich und verständlich sein. Als Anhaltspunkt sieht die [Empfehlung CM/Rec\(2020\)1](#) vor, dass „Zertifizierungssysteme auf der Grundlage regionaler und internationaler Standards so konzipiert und angewendet werden sollten, dass die Herkunft und Qualität von Datensätzen und Modellen gewährleistet sind“, und dass diese „auch Teil von Beschaffungsprozessen sein sollten und sich an regulatorischen Rahmenbedingungen orientieren sowie diesen entsprechen sollten, die bestimmte Verwendungszwecke algorithmischer Systeme verbieten“.

75. Die Mitgliedstaaten könnten zudem die Einführung technischer Standards und Validierungsmethoden fördern, mit denen Verzerrungen in KI-Systemen gemessen und gemindert werden können. Derartige Instrumente können dazu beitragen, einen einheitlichen und wirksamen Ansatz zur Bekämpfung von Voreingenommenheit zu gewährleisten, die sich aus Aktivitäten im Lebenszyklus von KI-Systemen ergibt, und gleichzeitig Innovationen bei der Entwicklung technischer Lösungen und politischer Maßnahmen zur Bekämpfung von Diskriminierung im Zusammenhang mit KI-Systemen fördern.

76. Diese Bewertungs- und Evaluierungsmechanismen liefern darüber hinaus objektive Kriterien, die in öffentliche Vergabeverfahren einfließen können. Diese sollten klare Zulassungs- und Auswahlkriterien enthalten, um sicherzustellen, dass KI-Systeme, die von Behörden oder anderen Organisationen eingesetzt werden, nicht diskriminierend sind und gleichberechtigte Ergebnisse fördern (siehe in diesem Zusammenhang Absatz 9 der GPR Nr. 7 der ECRI⁹⁷). Einige Mitgliedstaaten haben Maßnahmen ergriffen, um Leitlinien für die Beschaffung von KI-gestützten Systemen durch staatliche Stellen zu erstellen. Im Vereinigten Königreich hat die Regierung einen [Leitfaden zum Einsatz von KI im öffentlichen Sektor](#) veröffentlicht, um öffentlichen Einrichtungen die Einführung KI-gestützter Systeme in einer Weise zu ermöglichen, die der Gesellschaft zugutekommt. Das Weltwirtschaftsforum hat zudem einen Leitfaden mit dem Titel [„Adopting AI Responsibly: Guidelines for Procurement of AI Solutions by the Private Sector“](#), zu Deutsch: „Verantwortungsvoller Einsatz von KI: Leitlinien für die Beschaffung von KI-Lösungen durch den privaten Sektor“, veröffentlicht, und die Organisation „Data Ethics“ hat ein [Whitepaper](#) zur Datenethik im öffentlichen Beschaffungswesen herausgegeben. Auch die hochrangige Expertengruppe der EU (AI HLEG) hat [empfohlen](#), das öffentliche Beschaffungswesen strategisch zur Finanzierung von Innovationen und zur Gewährleistung vertrauenswürdiger KI zu nutzen, indem „klare Zulassungs- und Auswahlkriterien in die Beschaffungsvorschriften und -verfahren der EU-Institutionen, -Agenturen und Mitgliedstaaten aufgenommen werden, die vorschreiben, dass KI-Systeme vertrauenswürdig sein müssen“.

Rechenschaftspflicht, Haftung und Regeln zur Beweislast im Bereich der Gleichheit und Nichtdiskriminierung

Zu den Absätzen 21 – 26

77. In diesem Unterabschnitt wird erörtert, wie die Mitgliedstaaten ihre Verfahrensvorschriften im Zusammenhang mit Diskriminierungsklagen und Rechtsstreitigkeiten optimieren können, um sicherzustellen, dass betroffene Personen Zugang zur Justiz und zu wirksamen Rechtsbehelfen haben. Die Anwendung dieses Teils der Empfehlung auf Strafverfahren hängt vom jeweiligen nationalen Kontext und den geltenden Vorschriften ab. Die Grundsätze der Rechenschaftspflicht und der Verantwortung beziehen sich auf die Notwendigkeit, Mechanismen zu schaffen, die sicherstellen, dass Personen, Organisationen oder Einrichtungen, die für die Aktivitäten im Lebenszyklus von KI-Systemen verantwortlich sind, für die nachteiligen Auswirkungen auf die Menschenrechte, die Demokratie oder die Rechtsstaatlichkeit zur Rechenschaft gezogen werden können, die sich aus den Aktivitäten im Lebenszyklus dieser Systeme ergeben. Dies kann administrative, gerichtliche und sonstige Verfahren und Maßnahmen sowie zivil-, straf- und sonstige

⁹⁷ ECRI [Allgemeine politische Empfehlung Nr. 7 \(überarbeitet\) zu nationalen Rechtsvorschriften zur Bekämpfung von Rassismus und rassistischer Diskriminierung](#).

Haftungsregelungen umfassen, damit Entscheidungen angefochten und diskriminierende Ergebnisse korrigiert werden können, oder die Auferlegung spezifischer Verantwortlichkeiten und Pflichten für die Betreiber. In Absatz 66 des erläuternden Berichts zum Rahmenübereinkommen wird klargestellt, dass die Vertragsparteien verpflichtet sind, neue Rahmenbedingungen und Mechanismen zu schaffen oder bestehende Rahmenbedingungen und Mechanismen beizubehalten, die dann auf Aktivitäten im Lebenszyklus von KI-Systemen angewendet werden können, um den Grundsatz der Rechenschaftspflicht und Verantwortung umzusetzen. Darüber hinaus werden die Mitgliedstaaten im Einklang mit den Leitlinien des Ministerkomitees des Europarates zur Effizienz und Wirksamkeit von Rechtsbeistandssystemen im Zivil- und Verwaltungsrecht⁹⁸ dazu angehalten, sicherzustellen, dass Rechtsbeistand in geeigneten Fällen zur Verfügung steht.

78. In einigen Mitgliedstaaten wäre die Einführung von Sammelklagen oder einer verschuldensunabhängigen Haftung mit dem bestehenden nationalen Rechtsrahmen unvereinbar. Ungeachtet dessen kann es für die Mitgliedstaaten unter Berücksichtigung der jeweiligen nationalen Gegebenheiten angebracht sein, neue Verfahren einzuführen oder bestehende Verfahren anzupassen, um angesichts der besonderen Herausforderungen, die KI-Systeme mit sich bringen, einen wirksamen Zugang zur Justiz und zu Rechtsbehelfen zu gewährleisten. Zum Beispiel bedeutet die Informationsasymmetrie, die häufig im Zusammenhang mit Diskriminierungsklagen gegen eine Entscheidung auftritt, die von einem KI-System getroffen oder unter dessen Einsatz getroffen wurde, dass insbesondere verfahrensrechtliche Fristen die Kläger benachteiligen können.⁹⁹ Der Begriff „kollektive Rechtswege“ in Absatz 23 bezieht sich auf Rechtsverfahren, die von einer Gruppe von Klägern angestrengt werden, die auf dieselbe Weise von einem Vorfall betroffen sind, sei es durch Fahrlässigkeit, einen Zwischenfall oder eine andere Ursache. Solche Maßnahmen können für Kläger im Zusammenhang mit durch KI-Systeme verursachter Diskriminierung besonders vorteilhaft sein, da Informationsasymmetrien sowie mangelndes Bewusstsein für mutmaßliche Diskriminierung in Verbindung mit den finanziellen Auswirkungen eines Rechtsstreits gegen einen Anbieter oder Betreiber dazu führen, dass sie Gefahr laufen, keine wirksamen Rechtsbehelfe zu erhalten.

79. Es könnten widerlegbare Vermutungen hinsichtlich der (mangelnden) Achtung der Gleichheit, einschließlich der Gleichheit der Geschlechter, sowie der Verpflichtungen zur Nichtdiskriminierung geprüft werden. In einigen Mitgliedstaaten gibt es bereits Beweisregeln, die eine Umkehr der Beweislast in Diskriminierungsfällen vorsehen, oder andere Methoden zum Ausgleich von Informationsasymmetrien; siehe beispielsweise die ständige Rechtsprechung der deutschen Gerichte, wonach eine Umkehr der Beweislast erfolgen kann, wenn ein Verstoß aus dem Gefahrenbereich resultiert, für den eine Partei haftet (*Beweislast nach Gefahrenbereichen*), sowie Pflichten der Offenlegung (*sekundäre Darlegungslast*) oder eine Umkehr der Beweislast, wenn die Erfahrung zeigt, dass ein wiederkehrendes Muster vorliegt, das in der Regel zum Tragen kommt und zu diskriminierenden Ergebnissen führt (*Anscheinsbeweis*). In der Empfehlung werden die Mitgliedstaaten aufgefordert, die Einführung einer vorübergehenden Umkehr der Beweislast oder der Vermutung der Rechtmäßigkeit von KI-Systemen in Erwägung zu ziehen. Es könnten weitere Anpassungen der Beweisregeln in Betracht gezogen werden, um Rechtsverfahren von Personen zu erleichtern, die Diskriminierung aufgrund fehlender oder unzureichender Präventivmaßnahmen seitens der Anbieter oder Betreiber von KI-Systemen geltend machen. In Verbindung mit Pflichten der Offenlegung und im Rahmen eines abgestuften und differenzierten Ansatzes würde eine solche Umkehr der Beweislast oder der Vermutung die Beweislast der Betroffenen bei der Darlegung einer *prima-facie-Diskriminierung* verringern. Anbieter und Betreiber von KI-Systemen könnten beispielsweise eine solche Umkehr der Beweislast oder die Entstehung von Vermutungen hinsichtlich einer Voreingenommenheit verhindern, wenn sie nachweisen können, dass sie sinnvolle und ausreichende Präventivmaßnahmen ergriffen haben. Zu solchen Maßnahmen würden beispielsweise Diskriminierungsrisikobewertungen, Gleichheitsfolgenabschätzungen sowie Zertifizierungen gehören (siehe Absatz 18). Wurden solche vorbeugenden Maßnahmen nicht ergriffen, könnte eine vorläufige Vermutung einer algorithmischen Voreingenommenheit aufgestellt werden, was zur Folge hätte, dass die Beweislast auf den Beklagten verlagert würde. Dies könnte erreicht werden, indem man sich auf bestehende nationale Vorschriften zur Umkehr der Beweislast stützt. Nutzende und/oder Anbieter von KI-Systemen müssten dann nachweisen, dass das System den Anforderungen des Antidiskriminierungsrechts entspricht. Solche Techniken könnten sowohl die Allgegenwärtigkeit und Vorhersehbarkeit von Verzerrungen als auch die Machtverschiebungen widerspiegeln, die mit dem Einsatz von KI-Systemen zwischen Anbietern/Betreibern und den betroffenen Personen einhergehen.

80. Die Mitgliedstaaten könnten sich auch an Artikel 18 über die Umkehr der Beweislast der [Richtlinie \(EU\) 2023/970](#) zur Stärkung des Grundsatzes des gleichen Entgelts für gleiche oder gleichwertige Arbeit von Männern und Frauen durch Entgelttransparenz und Durchsetzungsmechanismen orientieren, der vorsieht, dass die Mitgliedstaaten, gemäß ihren nationalen Rechtssystemen, die geeigneten Maßnahmen ergreifen um

⁹⁸ [Leitlinien des Ministerkomitees des Europarates zur Effizienz und Wirksamkeit von Rechtsbeistands-Überstemen im Zivil- und Verwaltungsrecht](#), CM(2021)36-add2final, verabschiedet am 31. März 2021.

⁹⁹ Zur dreimonatigen Frist in Verfahren der gerichtlichen Überprüfung im Vereinigten Königreich siehe Cobbe, Jennifer, „Administrative Law and the Machines of Government: Judicial Review of Automated Public-Sector Decision-Making“ (6. August 2018). Vorabentwurf veröffentlicht unter Jennifer Cobbe (2019) „Administrative law and the machines of government: judicial review of automated public-sector decision-making“, Legal Studies, 39 (4), verfügbar bei SSRN: <https://ssrn.com/abstract=3226913> oder <http://dx.doi.org/10.2139/ssrn.3226913>.

sicherzustellen, dass in Fällen, in denen Arbeitnehmer, die sich benachteiligt fühlen, weil der Grundsatz des gleichen Entgelts auf sie nicht angewendet wurde, vor einer zuständigen Behörde oder einem nationalen Gericht Tatsachen darlegen, aus denen sich vermuten lässt, dass eine unmittelbare oder mittelbare Diskriminierung vorliegt, die Beweislast bei der beklagten Partei liegt, um nachzuweisen, dass keine unmittelbare oder mittelbare Diskriminierung in Bezug auf das Entgelt vorliegt. Besonders wichtig ist, dass dieser Artikel in Absatz 2 klarstellt, dass die Mitgliedstaaten sicherstellen müssen, dass in Verwaltungsverfahren oder Gerichtsverfahren wegen mutmaßlicher unmittelbarer oder mittelbarer Diskriminierung im Zusammenhang mit der Entlohnung, sofern ein Arbeitgeber bestimmte Verpflichtungen zur Lohntransparenz nicht erfüllt hat, die Beweislast dafür, dass keine solche Diskriminierung vorliegt, beim Arbeitgeber liegt.

81. Solche *Ex-ante*-Rechenschaftsmechanismen würden zudem Richtern und Gleichheitsstellen mit Entscheidungsbefugnis dabei helfen, in Diskriminierungsfällen den Verhältnismäßigkeitstest anzuwenden und die Angemessenheit der getroffenen Präventivmaßnahmen zu bewerten. Dieser Mechanismus würde Anreize für Nutzende und Anbieter von KI schaffen, technische Entscheidungen in rechtlich nachvollziehbare Informationen umzusetzen, anhand derer sich dann beurteilen lässt, ob ein bestimmtes System einem legitimen Zweck dient und ob die eingesetzten Mittel in einem angemessenen Verhältnis zu dem verfolgten Zweck stehen. Mit solchen Mechanismen zur Rechenschaftspflicht könnten verstärkte Pflichten zur Offenlegung einhergehen. Zu diesen Verpflichtungen könnten die in Absatz 15 der Empfehlung genannten Informationen gehören.

82. In Absatz 26 wird betont, dass die Mitgliedstaaten sicherstellen sollten, dass den betroffenen Personen wirksame Rechtsbehelfe zur Verfügung stehen. Artikel 13 der Konvention gewährleistet jedem, dessen in der Konvention verankerten Rechte und Freiheiten verletzt wurden, den Zugang zu einem wirksamen Rechtsbehelf vor einer nationalen Behörde. Um wirksam zu sein, muss der Rechtsbehelf in der Lage sein, die beanstandete Situation unmittelbar zu beheben¹⁰⁰, und er muss ausreichend und zugänglich sein, um der Verpflichtung zur zügigen Bearbeitung gerecht zu werden.¹⁰¹ Zu den wirksamen Rechtsbehelfen gehört der Ersatz immaterieller Schäden (Absatz 12 der Allgemeinen Empfehlung Nr. 7 der ECRI zur nationalen Gesetzgebung zur Bekämpfung von Rassismus und rassistischer Diskriminierung). Im Rahmen der Europäischen Union kann die Zahlung einer Entschädigung an die Opfer auch eine Sanktion im Sinne von Artikel 15 der Richtlinie 2000/43/EG des Rates vom 29. Juni zur Umsetzung des Grundsatzes der Gleichbehandlung von Personen ohne Unterschied der Rasse oder der ethnischen Herkunft darstellen.

83. In Bezug auf die Haftung sind die bestehenden deliktrechtlichen Regelungen technologieneutral und decken möglicherweise bereits Schäden, die durch KI-Systeme verursacht werden, angemessen ab. Insbesondere sollten die Mitgliedstaaten das bestehende Konzept der Verletzung von Sorgfaltspflichten berücksichtigen. Sofern die Anbieter und/oder Betreiber eines diskriminierenden Systems keine vorbeugenden Maßnahmen gegen algorithmische Verzerrungen ergriffen haben, könnte dies als Verletzung einer Sorgfaltspflicht im Sinne des geltenden Haftungsrechts der Mitgliedstaaten gelten. In vielen Fällen könnte es für Betroffene vorteilhafter sein, sich auf bestehende technologieneutrale Haftungsregelungen einschließlich der geltenden Antidiskriminierungsgesetze zu stützen, da diese möglicherweise nicht erfordern, dass der Betroffene nachweist, ob ein KI-System oder eine andere Art von Algorithmus an dem betreffenden Fall beteiligt war. In Fällen von Diskriminierung, die sich aus Aktivitäten im Lebenszyklus von KI-Systemen ergeben, sollten Betreiber und Anbieter nicht allein deshalb von der Haftung befreit werden, weil die Diskriminierung „autonom“ durch das diskriminierende System hervorgerufen wurde. Das Versäumnis, Diskriminierung, die sich aus den Aktivitäten im Lebenszyklus von KI-Systemen ergibt, in angemessener Weise zu verhindern, sollte als Verletzung der Sorgfaltspflicht angesehen werden. Die verschuldensunabhängige Haftung wurde bereits in bestimmten, genau definierten Bereichen eingeführt, um Risiken zu bewältigen, die sich aus Technologien ergeben, welche anderen unverhältnismäßigen Schaden zufügen können.

Zu Absatz 27

84. Absatz 27 lässt den Mitgliedstaaten die Möglichkeit offen, im Einklang mit ihren bestehenden rechtlichen Verpflichtungen zu prüfen, ob ein Moratorium oder ein Verbot in Bezug auf bestimmte KI-Systeme angemessen ist, wenn das Risiko für die Gleichbehandlung und Nichtdiskriminierung als so hoch angesehen wird, dass es nicht hinnehmbar ist, und es nicht möglich ist, dieses Risiko durch andere Präventivmaßnahmen zu mindern. Wie in den Absätzen 111–112 des erläuternden Berichts zum Rahmenübereinkommen dargelegt, obliegt es dem Mitgliedstaat, zu bestimmen, was in diesem Zusammenhang als „unvereinbar“ gilt, ebenso wie die Beurteilung, ob ein solches Szenario einerseits ein Moratorium oder ein Verbot oder andererseits eine andere geeignete Maßnahme erfordern würde.

Zu Absatz 28

85. Und schließlich wird in Absatz 28 auf eine Reihe von Stellen Bezug genommen, darunter die zuständigen Behörden, zu denen in diesem Zusammenhang auch die Datenschutzbehörden gehören. Die Mitgliedstaaten können

¹⁰⁰ Pine Valley Developments Ltd und andere gegen Irland, Entscheidung der Kommission, 1989.

¹⁰¹ Paulino Tomás gegen Portugal (Entscheidung), 2003; Çelik und İmret gegen die Türkei, 2004, § 59.

je nach den nationalen Gegebenheiten und dem Grad der Unabhängigkeit dieser Stellen deren Zusammenarbeit erleichtern.

V. Demokratische Teilhabe, Sensibilisierung der Öffentlichkeit und Kompetenzaufbau

Zu Absatz 29

86. In Absatz 29 der Empfehlung wird betont, dass die Mitgliedstaaten sicherstellen müssen, dass die Aktivitäten öffentlicher Stellen im gesamten Lebenszyklus von KI-Systemen einer angemessenen Überprüfung unterzogen werden. Wie in Absatz 31 der vorliegenden Begründung hervorgehoben wird, in dem die Bedeutung der parlamentarischen Kontrolle bei der Verabschiedung von Maßnahmen zur Förderung der Gleichheit und zur Verhinderung von Diskriminierung im Zusammenhang mit künstlicher Intelligenz thematisiert wird, kann dies beispielsweise regelmäßige parlamentarische Debatten umfassen. Eine parlamentarische Kontrolle ist unerlässlich, um sicherzustellen, dass öffentliche Stellen bei der Entwicklung und dem Einsatz von KI-Systemen transparent und rechenschaftspflichtig bleiben, und um vor potenziellen Risiken zu schützen, die sich aus Aktivitäten öffentlicher oder privater Akteure im Lebenszyklus von KI-Systemen ergeben, mit dem Ziel, die Gleichheit, einschließlich der Gleichheit der Geschlechter, zu fördern und Diskriminierung während des gesamten Lebenszyklus von KI-Systemen zu verhindern.

Zu den Absätzen 30 – 31

87. Die Empfehlung zielt darauf ab, die Auswirkungen der digitalen Kluft, einschließlich der geschlechtsspezifischen digitalen Kluft, zu verringern; dabei handelt es sich um die Kluft zwischen denjenigen, die Zugang zu digitalen Technologien (einschließlich des Internets und von KI-Systemen) haben, und denjenigen, die keinen solchen Zugang haben, was häufig auf systemische Diskriminierung zurückzuführen ist. In der Empfehlung wird zudem festgestellt, dass zu den durch die digitale Kluft ausgegrenzten Personen nicht nur diejenigen gehören, die keinen angemessenen Zugang zu digitalen Hilfsmitteln haben, sondern auch diejenigen, denen die Möglichkeiten fehlen, den effektiven Umgang mit diesen zu erlernen. Ein eingeschränkter Zugang zu KI-Systemen und -Tools sowie zu Bildungsprogrammen, die den Nutzenden digitale Kompetenz vermitteln, kann Einzelpersonen oder Gruppen daran hindern, die damit verbundenen Vorteile und Nutzen zu erfahren, was zu Benachteiligungen in verschiedenen Bereichen führen kann, darunter Bildung, Beschäftigung, Zugang zu Sozialleistungen und Gesundheitsversorgung. Dies wiederum könnte die soziale Mobilität und die Beschäftigungsfähigkeit einschränken und Ungleichheit verfestigen, insbesondere bei jungen Menschen, von denen digitale Kompetenz am häufigsten erwartet oder sogar verlangt wird. Zudem würde dies die Wirkung von Maßnahmen einschränken, die zur Förderung der Vielfalt am Arbeitsplatz im Technologiesektor ergriffen werden. Im Zusammenhang mit der Gleichheit, einschließlich der Gleichheit der Geschlechter, und der Nichtdiskriminierung gilt: Wenn bestimmte Gruppen oder Einzelpersonen beim Zugang zu KI-Tools auf Hindernisse stoßen, werden sie wahrscheinlich auch in den Datensätzen fehlen, die zum Trainieren von KI-Systemen verwendet werden, was das Risiko erhöht, dass voreingenommene Ergebnisse bestätigt und verstärkt werden. Außerdem sinkt dadurch die Wahrscheinlichkeit, dass ihre Bedürfnisse bei der Konzeption und Entwicklung neuer Systeme berücksichtigt werden. Die Mitgliedstaaten sollten daher prüfen, wie digitale Klüfte überwunden werden können, und sicherstellen, dass alle Teile der Gesellschaft, insbesondere diejenigen, die von Diskriminierung bedroht sind, einen sicheren und zuverlässigen Zugang zum Internet sowie zu Möglichkeiten zur Förderung der digitalen Kompetenz erhalten.¹⁰² Solche Programme sollten durch eine angemessene Finanzierung unterstützt werden. Zu den konkreten Maßnahmen können finanzielle Unterstützung zur Überwindung von Bezahlschranken für den Zugang zu KI-Tools, die Bereitstellung von Mitteln und Ressourcen für den Ausbau der Infrastruktur, Programme zur Förderung der digitalen Kompetenz sowie Initiativen zum Abbau von Ungleichheiten oder die Möglichkeit zur Teilnahme (beispielsweise an Konsultationen) in Präsenzveranstaltungen gehören. Programme zur Förderung der digitalen Kompetenz sollten sowohl im formalen als auch im nichtformalen Bildungsbereich angeboten werden, und die Mitgliedstaaten sollten erwägen, diese gegebenenfalls bereits ab einem frühen Alter bereitzustellen.

Zu Absatz 32

88. Die Festlegung technischer Normen kann besonders gravierende Auswirkungen auf die Gleichheit und die Nichtdiskriminierung haben, da sie die Definition und Auslegung rechtlicher Normen in der Praxis beeinflussen kann. Allerdings findet eine solche Festlegung von Standards oft fernab der öffentlichen Kontrolle und der Beteiligung von Organisationen und Einzelpersonen statt. Hinsichtlich der Beteiligung der Zivilgesellschaft, die in Absatz 32 der Empfehlung hervorgehoben wird, wird auf die Empfehlung CM/Rec(2018)11 verwiesen, in der die Notwendigkeit betont wird, den Schutz und die Förderung des zivilgesellschaftlichen Raums in Europa zu stärken. Das Instrument empfiehlt den Mitgliedstaaten,

¹⁰² „Human Rights by Design – Future-Proofing Human Rights Protection in the Era of AI“, Empfehlung des Menschenrechtskommissars des Europarats, 2023, S. 35.

sicherzustellen, dass alle Menschen, einschließlich Menschenrechtsverteidiger, wirksam an Entscheidungsprozessen teilnehmen können, indem sie uneingeschränkten Zugang zu Informationen gewähren und zeitnahe sowie transparente öffentliche Konsultationen durchführen, insbesondere wenn die Politikgestaltung oder Gesetzesentwürfe Auswirkungen auf die Zivilgesellschaft haben könnten.¹⁰³ Darüber hinaus haben die Staats- und Regierungschefs des Europarates in der Erklärung von Reykjavík bekräftigt, dass die Zivilgesellschaft eine Voraussetzung für eine funktionierende Demokratie ist, und sich verpflichtet, ein sicheres und förderliches Umfeld zu unterstützen und aufrechtzuerhalten, in dem die Zivilgesellschaft und Menschenrechtsverteidiger frei von Behinderungen, Unsicherheit und Gewalt agieren können.¹⁰⁴ Diese Verpflichtungen sind für die Steuerung von KI-Systemen von unmittelbarer Bedeutung, da die Beteiligung der Zivilgesellschaft an Normungsprozessen entscheidend ist, um sicherzustellen, dass die rechtlichen und gesellschaftlichen Auswirkungen von KI-Technologien, einschließlich der Auswirkungen auf die Gleichheit, umfassend berücksichtigt werden. Besondere Aufmerksamkeit sollte auch der Beteiligung junger Menschen oder der sie vertretenden Organisationen gewidmet werden. Um junge Menschen zur Teilnahme zu motivieren, könnten die Mitgliedstaaten gezielte Bildungsprogramme anbieten und Sensibilisierungsinitiativen durchführen, die bereits im frühen Kindesalter ansetzen und darauf abzielen, ihr Verständnis sowie ihr kritisches Denken in Bezug auf KI, Gleichheit und Nichtdiskriminierung zu fördern. Aufklärungsmaterialien sollten in mehreren Sprachen verfügbar und kulturell angepasst sein, auch für Bevölkerungsgruppen mit begrenzten digitalen Kompetenzen.

89. Durch die Durchführung öffentlicher Konsultationen und die Bereitstellung transparenter Kanäle für Rückmeldungen können die Mitgliedstaaten sicherstellen, dass die Entwicklung von KI-Systemen inklusiv und transparent erfolgt und dass dabei die weiterreichenden gesellschaftlichen Auswirkungen im Hinblick auf Gleichheit, einschließlich der Gleichheit der Geschlechter, sowie Nichtdiskriminierung berücksichtigt werden.

90. In Bezug auf den in Absatz 30.3 genannten gezielten Kompetenzaufbau und die Schulungen könnten solche Initiativen insbesondere Mitarbeiterinnen und Mitarbeiter von Anbietern und Betreibern von KI-Systemen sowie von unabhängigen Aufsichtsgremien einbeziehen.

VI. Zuständigkeiten und Befugnisse unabhängiger Aufsichtsgremien

Zu Absatz 33

91. Dieses Kapitel befasst sich mit den Kompetenzen und Befugnissen unabhängiger Stellen, die die Anwendung und Umsetzung von Antidiskriminierungsrecht überwachen, wie beispielsweise Gleichheitsstellen und gegebenenfalls andere nationale Menschenrechtsstrukturen („unabhängige Aufsichtsstellen“), im Zusammenhang mit Aktivitäten im Lebenszyklus von KI-Systemen. Die institutionellen Regelungen sowie die Aufgabenbereiche dieser Gremien unterscheiden sich zwischen den Mitgliedstaaten des Europarates, wobei im Allgemeinen gilt, dass Gleichheitsstellen ein spezifischeres Mandat haben und sich auf die Prävention und Bekämpfung von Diskriminierung spezialisieren, während sich nationale Menschenrechtsinstitutionen auch mit weiter gefassten Menschenrechtsfragen befassen.¹⁰⁵ In einigen Mitgliedstaaten sind zudem alle Funktionen in einer einzigen Institution zusammengefasst.¹⁰⁶ Solche Stellen sollten institutionell und/oder funktional unabhängig von den Behörden der Mitgliedstaaten sein¹⁰⁷ und können oft als benutzerfreundlicherer Mechanismus dienen, über den Einzelpersonen Rechtsbehelf gegen Diskriminierung einlegen sowie strukturelle Diskriminierung und Ungleichheit bekämpfen können.

Zu Absatz 34

92. Die Empfehlung stützt sich auf die Allgemeine Empfehlung Nr. 2 der ECRI zu Gleichheitsstellen zur Bekämpfung von Rassismus und Intoleranz auf nationaler Ebene, in der festgelegt ist, dass Gleichheitsstellen folgende Aufgaben übertragen werden sollten: die Aufgabe, Gleichheit zu fördern und Diskriminierung zu verhindern (Förderungs- und Präventionsfunktion); die Aufgabe, von Diskriminierung und Intoleranz betroffene Personen zu unterstützen und in ihrem Namen Rechtsstreitigkeiten zu führen (Unterstützungs- und Prozessführungsfunktion); sowie gegebenenfalls die Aufgabe, über Beschwerden zu entscheiden (Entscheidungsfunktion). Zu den Kompetenzen im Bereich der Unterstützung und der Prozessführung können beispielsweise die in Absatz 14 der GPR Nr. 2 der ECRI genannten Aufgaben gehören. Die Entscheidungsbefugnis kann zwischen Gleichheitsstellen und der Justiz aufgeteilt oder vollständig der Justiz übertragen werden. [Die Empfehlung CM/Rec\(2021\)1](#) zur Entwicklung und Stärkung wirksamer, pluralistischer und unabhängiger nationaler Menschenrechtsinstitutionen besagt, dass die Mitgliedstaaten sicherstellen sollten, dass das den nationalen Menschenrechtsinstitutionen übertragene Mandat zum Schutz und zur

¹⁰³ Empfehlung [CM/Rec\(2018\)11](#) des Ministerkomitees an die Mitgliedstaaten zur Notwendigkeit, den Schutz und die Förderung des zivilgesellschaftlichen Raums in Europa zu stärken, d. h. und f.

¹⁰⁴ [United around our values – Reykjavík Declaration](#) (2023), S. 16.

¹⁰⁵ Siehe die Allgemeine Politikempfehlung Nr. 2 (überarbeitet) der ECRI [zu Gleichstellungsstellen zur Bekämpfung von Rassismus und Intoleranz auf nationaler Ebene](#), Absätze 4 ff.

¹⁰⁶ Siehe GPR Nr. 2 Absatz 7 der ECRI.

¹⁰⁷ Siehe GPR Nr. 2 Absatz 22 der ECRI.

Förderung der Menschenrechte so umfassend wie möglich ist und den Pariser Grundsätzen in vollem Umfang entspricht, und detaillierte Leitlinien zum Inhalt eines solchen Mandats enthält.

93. In Anbetracht der GPR Nr. 2 der ECRI und der unterschiedlichen Gegebenheiten in den Mitgliedstaaten enthält Absatz 34 der Empfehlung eine Auflistung der Befugnisse, mit denen die Mitgliedstaaten, unter Berücksichtigung der jeweiligen nationalen Gegebenheiten, unabhängige Aufsichtsstellen ausstatten sollten. Insbesondere wird in der Empfehlung auch festgestellt, dass die Übertragung solcher Rechte und Befugnisse an diese Institutionen mit der Bereitstellung angemessener Ressourcen, einer ausreichenden operativen Unabhängigkeit sowie der Fähigkeit zur Zusammenarbeit mit anderen Aufsichtsbehörden und Organisationen einhergehen muss.

Zu den Absätzen 35 – 37

94. Es wird zudem empfohlen, dass unabhängige Aufsichtsgremien an der Erprobung neuer und innovativer Technologien teilnehmen können, unter anderem im Rahmen von regulatorischen Sandboxes, wie oben beschrieben, sofern sie dies wünschen. Die Überprüfung könnte in Form von Situationstests erfolgen, bei denen diese Behörden die Ergebnisse eines bestimmten Systems überprüfen, indem sie die Ergebnisse verschiedener Gruppen miteinander vergleichen. So könnten sie beispielsweise Test-Lebensläufe oder Kreditanträge von Angehörigen der Mehrheits- und Minderheitsgruppen einreichen, um zu versuchen, Diskriminierung in Kontexten aufzudecken, in denen Unternehmen KI-Systeme einsetzen. Sollten sie Zugang zu den entsprechenden Systemen erhalten, könnten sie zudem Audits durchführen, um mögliche Verzerrungen aufzudecken. Solche Methoden könnten den Opfern helfen, indem sie bestehende Hindernisse beim Nachweis einer *prima facie*-Diskriminierung abbauen. Derartige Stellen könnten Diskriminierung möglicherweise verhindern, indem sie mit den Nutzenden von KI-Systemen zusammenarbeiten, beispielsweise mit Unternehmen, die solche Systeme zur Unterstützung von Einstellungsverfahren einsetzen, um relevante Daten über die Auswirkungen ihrer Entscheidungen zu erheben und Verzerrungen zu beseitigen. Unabhängige Aufsichtsgremien könnten zudem befugt werden, Frühwarnungen und Empfehlungen auszusprechen, wenn algorithmische Vorgehensweisen potenzielle Bedenken hinsichtlich der Gleichheit aufwerfen, selbst wenn keine individuellen Beschwerden vorliegen. Schließlich könnten sie potenziellen Opfern dabei helfen, Wiedergutmachung zu erlangen. Je nach nationalem Kontext kann auch die Zivilgesellschaft zu diesem Vorhaben beitragen.

VII. Weitere Bestimmungen

Zu Absatz 38

95. KI-Systeme werden zunehmend als Informations- und Entscheidungsinstrument im Alltag eingesetzt. In einigen Branchen kann der Einsatz von KI-Systemen besonders gravierende Auswirkungen auf die Gleichheit, einschließlich der Gleichheit der Geschlechter, und auf die Nichtdiskriminierung haben. In der Empfehlung werden die Mitgliedstaaten daher aufgefordert, in den in Absatz 38 der Empfehlung aufgeführten Sektoren konkrete Maßnahmen zu ergreifen, wobei darauf hingewiesen wird, dass diese Aufzählung nicht erschöpfend ist und es für die Mitgliedstaaten gegebenenfalls angebracht sein kann, Maßnahmen in Bezug auf andere Sektoren zu ergreifen. In diesem Zusammenhang sollten die Mitgliedstaaten Maßnahmen ergreifen, um sicherzustellen, dass ein angemessenes Maß an menschlicher Aufsicht gewährleistet ist.

96. Um die Transparenz zu fördern, werden die Mitgliedstaaten dazu angehalten, klare und leicht zugängliche Informationen über den Einsatz von KI-Systemen in den in Absatz 38 der Empfehlung genannten relevanten Sektoren bereitzustellen. Dazu könnten nationale KI-Strategien sowie Berichte gehören, in denen der Einsatz von KI-Systemen in bestimmten Bereichen und deren tatsächliche oder potenzielle Auswirkungen auf die Gleichstellung, einschließlich der Gleichheit der Geschlechter, detailliert beschrieben werden.

97. Die Mitgliedstaaten sollten sich für die durchgängige Berücksichtigung der Geschlechterperspektive und der Gleichheit einsetzen, um die Gleichheit wirksam zu fördern und Diskriminierung bei Aktivitäten im Lebenszyklus von KI-Systemen zu verhindern und zu bekämpfen, insbesondere in den in Absatz 38 der Empfehlung genannten Sektoren.

98. In der Strategie zur Gleichheit der Geschlechter 2024–2029 wird dargelegt, dass die Gewährleistung der Gleichheit der Geschlechter als wesentlicher Bestandteil eines umfassenderen Kampfes zur Wahrung der Menschenrechte und zur Sicherung sozialer Gerechtigkeit verstanden werden muss. Es ist von entscheidender Bedeutung, die gegenseitige Verflechtung dieser Themen anzuerkennen und gemeinsam daran zu arbeiten, eine gerechtere, inklusivere und nachhaltigere Gesellschaft zu schaffen. Darüber hinaus wird in der Strategie betont, dass die Einbeziehung eines intersektionalen Ansatzes von entscheidender Bedeutung ist, um sicherzustellen, dass die kombinierten Formen von Diskriminierung und Unterdrückung, denen bestimmte Gruppen ausgesetzt sind, angegangen werden.¹⁰⁸ Bei der gleichheitsorientierten Politikgestaltung handelt es

¹⁰⁸ Strategie des Europarates [zur Gleichstellung der Geschlechter \(2024-2029\)](#), S. 50-51.

sich um die systematische Einbeziehung von Gleichheits- und Antidiskriminierungsaspekten in alle Phasen des politischen Entscheidungsprozesses auf der Ebene der Mitgliedstaaten und des Europarates.

99. Bei der gleichheitsorientierten Politikgestaltung geht es darum, die Auswirkungen geplanter oder umgesetzter Maßnahmen und Entscheidungen auf die Gleichheit und Nichtdiskriminierung aus der Perspektive verschiedener Gruppen kontinuierlich zu bewerten und diese gemäß den Ergebnissen entsprechend anzupassen. Daher handelt es sich um einen Prozess, der von Beginn bis zum Ende einer bestimmten Aktivität, einschließlich ihrer Bewertung, befolgt werden sollte. In der Empfehlung des Ministerkomitees zu Standards und Mechanismen der Gleichheit der Geschlechter ([CM/Rec\(2007\)17](#)) wird betont, wie wichtig es ist, Methoden zur Umsetzung der gleichheitsorientierten Politikgestaltung zu entwickeln. ierte Politikgestaltung sollte bei der Politikgestaltung, der Umsetzung und der Überprüfung politischer Maßnahmen berücksichtigt werden.

100. Um die Gleichheit zu verankern, ist eine ähnliche Methodik erforderlich, die klare Leitlinien und Schritte vorgeben sollte, um die Gleichheitsperspektive in neue politische Maßnahmen und Gesetze einzubeziehen und eine Gleichheitsprüfung sowie -bewertung bestehender Maßnahmen und Gesetze durchzuführen, um die Gleichheit in der Praxis für alle zu gewährleisten. Zu diesen Schritten können die Überprüfung und Bewertung der Auswirkungen auf die Gleichheit, die Gleichheitsplanung, Gleichheitspflichten,¹⁰⁹ die Einbeziehung der Gleichheitsperspektive in die Haushaltsplanung, die Konsultation und Beteiligung benachteiligter Gruppen und der Zivilgesellschaft, die Umsetzung sowie die Überprüfung gehören.

Bildung

101. Der Zugang zu Bildung ist ein Menschenrecht für alle Kinder und Erwachsenen, das auf der Grundlage der Gleichberechtigung und unter Gewährleistung gleicher Chancen gilt, um die Gleichheit der Geschlechter auf allen Bildungsebenen sicherzustellen.¹¹⁰ Kinder haben entsprechend ihrer Entwicklungsstufe spezifische Bedürfnisse und Rechte, die sie von Erwachsenen unterscheiden. Daher besteht Bedarf an kindgerechten Vorschriften für die Beschaffung und Nutzung von Bildungstechnologie.¹¹¹ Darüber hinaus verdeutlichen die neuesten Daten ein anhaltendes Geschlechtergefälle in Forschung und Innovation, da nur 21 % der Promovierten im Bereich der Informations- und Kommunikationstechnologien (IKT) Frauen sind.¹¹² Frauen sind in MINT-Studiengängen nach wie vor unterrepräsentiert, was wiederum dazu führt, dass sie die Beschäftigungsmöglichkeiten in diesem Bereich, einschließlich der KI, seltener nutzen.¹¹³ Die Mitgliedstaaten sollten daher der Teilhabe von Mädchen und jungen Menschen an diesen Bildungsbereichen Priorität einräumen.¹¹⁴ Die Mitgliedstaaten sollten zudem den Übergang in den KI-Arbeitsmarkt erleichtern, indem sie Zugang zu Weiterbildungsprogrammen, insbesondere für junge Frauen und junge Menschen aus unterrepräsentierten Gruppen, gewähren, Fördermittel und Stipendien bereitstellen sowie Mentoring-Programme unterstützen.¹¹⁵ Da Geschlechterstereotypen zudem häufig zu Berufsentscheidungen führen, die von stereotypen Rollenzuweisungen geprägt sind, sollten die Mitgliedstaaten Programme für eine stereotypenfreie Berufsberatung fördern.

102. Darüber hinaus sollten die Mitgliedstaaten sicherstellen, dass Bildungseinrichtungen KI-Systeme in einer Weise einsetzen, die mit der Bereitstellung einer qualitativ hochwertigen und inklusiven Bildung vereinbar ist.¹¹⁶ So kann beispielsweise Gesichtserkennungssoftware, die als Teil von Prüfungsüberwachungssoftware zur Beobachtung des Verhaltens von Teilnehmenden während Online-Prüfungen eingesetzt wird, Verzerrungen aufweisen und zu intersektionaler Diskriminierung führen, unter anderem aufgrund der ethnischen Herkunft und des Geschlechts.¹¹⁷ Besondere Aufmerksamkeit sollte daher dem Einsatz von KI-Systemen in Auswahl- und Prüfungsverfahren gewidmet werden, um diskriminierende Ergebnisse zu vermeiden.

¹⁰⁹ Gleichbehandlungspflichten sind positive rechtliche Verpflichtungen zur Förderung der Gleichbehandlung und zur Verhinderung von Diskriminierung, die über die allgemeine Verpflichtung zum Unterlass von Diskriminierung hinausgehen.

¹¹⁰ [Empfehlung CM/Rec\(2007\)17](#) des Ministerkomitees an die Mitgliedstaaten zu Standards und Mechanismen zur Gleichstellung der Geschlechter, Absätze 24 bis 25.

¹¹¹ „Preparatory study for the development of a legal instrument on regulating the use of artificial intelligence systems in education“, zu Deutsch: „Vorbereitende Studie zur Ausarbeitung eines Rechtsinstruments zur Regulierung des Einsatzes von Systemen künstlicher Intelligenz im Bildungswesen“, Überarbeiteter Entwurf (März 2024), Referat für digitale Transformation der Bildungsabteilung, Europarat.

¹¹² Europäische Kommission: Generaldirektion Forschung und Innovation, „*She Figures 2024 – Gender in research and innovation – Statistics and indicators*“, zu Deutsch: „She Figures 2024 – Gender in Forschung und Innovation – Statistiken und Indikatoren“, Amt für Veröffentlichungen der Europäischen Union, 2025.

¹¹³ CDADI/GEC-Studie (2023), S. 89; [Resolution 2343 \(2020\)](#) „Preventing discrimination caused by the use of artificial intelligence“, zu Deutsch: „Vorbeugung von Diskriminierung durch den Einsatz künstlicher Intelligenz“, Absatz 6; sowie Gender Equality Index 2023, Europäisches Institut für Gleichstellungsfragen (EIGE), S. 12, 37.

¹¹⁴ [Empfehlung CM/Rec\(2019\)1](#) zur Prävention und Bekämpfung von Sexismus, II.B.7.; [Resolution 2343 \(2020\)](#) „Preventing discrimination caused by the use of artificial intelligence“, zu Deutsch: „Vorbeugung von Diskriminierung durch den Einsatz künstlicher Intelligenz“, Absatz 12.1.

¹¹⁵ CDADI/GEC-Studie (2023), S. 89.

¹¹⁶ [Empfehlung CM/Rec\(2019\)1](#) zur Prävention und Bekämpfung von Sexismus, insbesondere Abschnitt II.G. „Education institutions“, zu Deutsch: „Bildungseinrichtungen“.

¹¹⁷ CDADI/GEC-Studie (2023), S. 24.

103. Die Mitgliedstaaten sollten sicherstellen, dass digitale Kompetenz in den Bildungslehrplänen verankert ist und Inhalte umfasst, die das Bewusstsein für das Recht auf Gleichbehandlung, einschließlich der Gleichheit der Geschlechter im Zusammenhang mit KI-Systemen, schärfen.¹¹⁸ Die Mitgliedstaaten könnten erwägen, die gemeinsame Gestaltung von Programmen zur Förderung der digitalen Kompetenz mit der Zielgruppe zu fördern, um Inklusivität zu gewährleisten. Insbesondere Kinder und Jugendliche sollten in einer ihrem Alter angemessenen Weise über die Risiken der Online-Überprüfung, des Cybermobbings und der Datenausbeutung sowie ganz allgemein über die Auswirkungen des Einsatzes von KI-Systemen auf Gleichheit und Nichtdiskriminierung aufgeklärt werden. Lehrkräfte aller Schulstufen sollten entsprechende Fortbildungen und Ressourcen erhalten, um dies zu ermöglichen. Darüber hinaus sollte die digitale Kompetenz auch durch nichtformale und informelle Bildung gefördert werden, um die formale Bildung zu ergänzen. Dies soll durch das Angebot leicht zugänglicher, gemeindebasierter Lernmöglichkeiten geschehen, damit lebenslange digitale Kompetenz auch außerhalb des Klassenzimmers gefördert wird. Die Mitgliedstaaten sollten besondere Anstrengungen unternehmen, um marginalisierte Gruppen sowie Personen, die von Diskriminierung und Benachteiligung bedroht sind, im Hinblick auf digitale Kompetenz zu erreichen.¹¹⁹ Im weiteren Sinne sollten sie das Wissen über KI-Systeme fördern, einschließlich ihrer Auswirkungen auf die Gleichheit, unter anderem die Gleichheit der Geschlechter, sowie auf die Nichtdiskriminierung.¹²⁰

Justiz, Rechtswesen und Strafverfolgung

104. Die Polizei- und Justizbehörden in den Mitgliedstaaten des Europarates setzen KI-Systeme ein, darunter Anwendungen wie „Präventive Polizeiarbeit“ sowie Risikobewertungen im Zusammenhang mit Untersuchungshaft, Strafzumessung und Bewährung.¹²¹ Der Einsatz von KI-Systemen in diesen Bereichen kann jedoch auf verschiedene Weise zu Diskriminierung führen.¹²² So wird beispielsweise die Gesichtserkennung zunehmend zur Aufdeckung und Prävention von Straftaten eingesetzt, doch hat sich in mehreren Fällen gezeigt, dass ihre Genauigkeit verzerrt ist, was zur falschen Identifizierung von Verdächtigen und in einigen Fällen zur unrechtmäßigen Inhaftierung unschuldiger Personen geführt hat.¹²³ Im Zusammenhang mit Strafverfolgung und Sicherheit sollten die Mitgliedstaaten daher Vorsicht walten lassen, wenn es darum geht, Diskriminierungsrisiken zu identifizieren, zu bewerten, zu verhindern und zu mindern, die sich aus dem Einsatz von Gesichtserkennungstechnologien oder biometrischen Fernidentifizierungssystemen im Strafverfolgungs- und Sicherheitsbereich ergeben, gegebenenfalls auch durch ein Verbot ihres Einsatzes. Vorhersagemodelle, die im Polizeidienst eingesetzt werden, beispielsweise zur Ermittlung von „Hotspots“¹²⁴, bei Strafzumessungsentscheidungen sowie bei vorausschauenden Risikobewertungen durch Polizei, Gerichte oder Bewährungshilfe¹²⁵, können ebenfalls zu einem Diskriminierungsrisiko führen. Werden mehrere Systeme verwendet, kann dies zu einem „Verkettungseffekt“ führen, der auftritt, wenn sich mehrere Systeme gegenseitig beeinflussen.¹²⁶

105. Die Europäische Kommission für die Effizienz der Justiz (CEPEJ) hat den ersten europäischen Text verabschiedet, in dem ethische Grundsätze für den Einsatz von KI in Justizsystemen festgelegt sind.¹²⁷ In diesem Text wird der Grundsatz der Nichtdiskriminierung als ein im Bereich der KI und der Justiz zu beachtender Kerngrundsatz genannt, wobei insbesondere darauf hingewiesen wird, wie wichtig es ist, die Entstehung oder Verschärfung jeglicher Diskriminierung zwischen Einzelpersonen oder Personengruppen gezielt zu verhindern. Der Text unterstreicht die potenziellen Risiken von Vorhersagen in Strafsachen unter Einsatz von KI-Tools, die ungerechtfertigte und bereits bestehende Ungleichheiten im Strafrechtssystem reproduzieren können.¹²⁸ Unter gebührender Berücksichtigung der Unabhängigkeit der Justiz und des Rechts auf ein faires Verfahren (Artikel 6

¹¹⁸ [Empfehlung CM/Rec\(2019\)1](#) zur Prävention und Bekämpfung von Sexismus, insbesondere II.B.2. Siehe auch CDADI/GEC-Studie (2023), S. 90; [Resolution 2343 \(2020\)](#) „Preventing discrimination caused by the use of artificial intelligence“, Absatz 12.3.; und [„Human rights by design – future-proofing human rights protection in the era of AI“](#) (2023) S. 21-22.

¹¹⁹ Menschenrechtskommissar des Europarats, [„Unboxing Artificial Intelligence: 10 steps to protect Human Rights“](#), 2019, S. 14.

¹²⁰ [„Human rights by design – future-proofing human rights protection in the era of AI“](#) (2023), S. 21-22.

¹²¹ CDADI/GEC-Studie (2023), S. 22–23; [Resolution 2342 \(2020\)](#) „Justice by algorithm – The role of artificial intelligence in policing and criminal justice systems“, Parlamentarische Versammlung des Europarates, Absatz 6.

Als Beispiel für bewährte Praktiken sei die Verabschiedung eines Gesetzes durch die britische Regierung genannt, das den öffentlichen Sektor verpflichtet, Informationen über die Rolle von Algorithmen bei der Unterstützung von Entscheidungen, die Einzelpersonen betreffen, offenzulegen, wie in [„Human rights by design – future-proofing human rights protection in the era of AI“](#) (2023), S. 17, dargelegt.

¹²² Empfehlung CM/Rec(2024)5 des Ministerkomitees an die Mitgliedstaaten zu den ethischen und organisatorischen Aspekten des Einsatzes künstlicher Intelligenz und damit verbundener digitaler Technologien durch Strafvollzugs- und Bewährungsbehörden, insbesondere die Absätze 3, 19 und 25.

¹²³ CDADI/GEC Studie (2023), S. 22-23. Weitere Beispiele finden sich in der [Resolution 2342 \(2020\)](#) „Justice by algorithm – The role of artificial intelligence in policing and criminal justice systems“, Absatz 7.

¹²⁴ [„Fundamental Rights Agency, Bias in Algorithms – Artificial Intelligence and Discrimination“](#), zu Deutsch: „Agentur für Grundrechte, Voreingenommenheit von Algorithmen – Künstliche Intelligenz und Diskriminierung“ (2022).

¹²⁵ [„Algorithmic risk assessment policing models: Lessons from the Durham Constabulary HART model“](#), zu Deutsch: „Algorithmische Modelle zur Risikobewertung in der Polizeiarbeit: Erkenntnisse aus dem HART-Modell der Durham Constabulary“, OSWALD, Marion, GRACE, Jamie, URWIN, Sheena und BARNES, Geoffrey (2018), Information and Communications Technology Law, 27 (2), 223–250.

¹²⁶ <https://research.northumbria.ac.uk/probablefutures/wp-content/uploads/2025/05/PROBABLE-Futures-Jan-13-Opening-Slides.pdf>

¹²⁷ [„European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment“](#), zu Deutsch: „Europäische Ethik-Charta zum Einsatz künstlicher Intelligenz in Justizsystemen und dessen Umfeld“, Europäische Kommission für die Effizienz der Justiz (CEPEJ), 2018.

¹²⁸ *Ibid*, S. 9, 79.

der Konvention) sollten die Mitgliedstaaten daher sicherstellen, dass der Einsatz von KI-Systemen innerhalb des Justizsystems die Entstehung oder Aufrechterhaltung jeglicher Diskriminierung zwischen Einzelpersonen oder Personengruppen verhindert, insbesondere durch Datenanalyse, Schulungen und den Austausch von Informationen über Praktiken, die für den Einsatz von KI-Systemen in Gerichtsverfahren relevant sind. Im weiteren Sinne sollten die Mitgliedstaaten, um den Zugang zu Rechtsbehelfen gegebenenfalls zu verbessern, die Transparenz hinsichtlich des Einsatzes von KI durch öffentliche Behörden, einschließlich Justiz- und Strafverfolgungsbehörden, stärken.¹²⁹ Es sollten Schutzmaßnahmen eingeführt werden, um sicherzustellen, dass der Einsatz von KI-Systemen in Gerichtsverfahren nicht zu einer Retraumatisierung der Opfer führt und dass KI-Systeme unter Anwendung eines traumainformierten und opferzentrierten Ansatzes eingesetzt werden.

106. Die Strafverfolgung gilt als ein Bereich, in dem es leicht zu Diskriminierung kommen kann; die Mitgliedstaaten sollten daher prüfen, ob neue Vorschriften erforderlich sind oder ob spezifische Maßnahmen zur Verhinderung von Diskriminierung ergriffen werden sollten.¹³⁰

Wahlverfahren

107. KI-Systeme bieten die Möglichkeit, demokratische Prozesse zu verbessern, insbesondere indem sie diese inklusiv gestalten. So können sie beispielsweise den Bürgern helfen, ein besseres Verständnis für Politik zu entwickeln und sich leichter an der demokratischen Debatte zu beteiligen, sowie die Zugänglichkeit von Wahlprozessen verbessern. Ähnlich wie in Absatz 47 dieser Begründung beschrieben, können KI-Systeme auch dazu eingesetzt werden, Falschinformationen zu identifizieren oder zu entfernen, die sich gegen Personen oder Gruppen richten, die von Diskriminierung bedroht sind. Die Erstellung und Verbreitung von Desinformation und unrichtigen Informationen, die insbesondere in Wahlkampfzeiten mit der Absicht, andere in die Irre zu führen, in Umlauf gebracht werden, birgt jedoch erhebliche Risiken, auch im Hinblick auf Personen, die von Diskriminierung bedroht sind. KI findet zunehmend Einzug in Anwendungen, die von Wählern zur Informationssuche genutzt werden, unabhängig davon, ob ihnen dies bewusst ist oder nicht, sei es durch KI-gestützte Suchfunktionen oder durch Produktivitätswerkzeuge. Von KI generierte Informationen mögen plausibel erscheinen, sind jedoch häufig fehlerhaft. Dieser Mangel an Zuverlässigkeit kann für die Wähler besonders nachteilig sein und leicht zum Nachteil von Personen oder Gruppen instrumentalisiert werden, die von Diskriminierung bedroht sind. Der Einsatz von KI-Systemen steht zudem im Zusammenhang mit der Erstellung digitaler Fälschungen und der Verstärkung gesellschaftlicher Vorurteile. Diese können politische Wahlkämpfe stören, das Vertrauen der breiten Öffentlichkeit beeinträchtigen und die Wahlbeteiligung sowie die politische Vertretung hemmen, wovon insbesondere Frauen betroffen sind.¹³¹

Arbeit und Beschäftigungsverhältnisse

108. Im Bereich Arbeit und Beschäftigungsverhältnisse sollten die Mitgliedstaaten sicherstellen, dass KI-Systeme keine bestehenden oder neu entstehenden Muster von Ungleichheit und struktureller Diskriminierung einführen, reproduzieren oder verstärken und dass sie so eingesetzt werden, dass sie die Gleichheit, einschließlich der Gleichheit der Geschlechter, sowie Vielfalt und Inklusion am Arbeitsplatz fördern. Insbesondere könnte dies Folgendes umfassen: regelmäßige Audits der Ergebnisse des Einsatzes von KI bei der Personalauswahl, bei Beförderungen und anderen Verfahren; die Einbeziehung der Beschäftigten und ihrer Vertretungsorganisationen in Richtlinien oder Entscheidungen bezüglich des Einsatzes von KI bei Entscheidungsprozessen am Arbeitsplatz; die Überprüfung der Auswirkungen der Einführung von KI-Systemen am Arbeitsplatz auf die Gleichheit der Geschlechter und die Vielfalt der Belegschaft; sowie Schulungen und Sensibilisierungsmaßnahmen für die Belegschaft zu Datenverzerrungen, Stereotypen und Diskriminierungsrisiken beim Einsatz von KI-Systemen. Darüber hinaus sind Personen, die von Diskriminierung bedroht sind, in besonderem Maße schlechten oder prekären Arbeitsbedingungen ausgesetzt, die sich durch den zunehmenden Einsatz von KI-Systemen am Arbeitsplatz noch verschlimmern können. Darüber hinaus stützen sich laut aktueller Studien viele KI-Systeme auf eine arbeitsintensive Datenaufbereitung, die häufig unter schlechten Arbeitsbedingungen im Globalen Süden, aber auch in Europa ausgelagert wird. Beschäftigte in den Bereichen Datenbeschriftung und Inhaltsmoderation sind schwerwiegenden ausbeuterischen Arbeitsbedingungen ausgesetzt, darunter niedrige Löhne, übermäßige Überwachung und der Kontakt mit belastenden Inhalten. Viele Datensätze für das Training künstlicher Intelligenz basieren auf unterbezahlter oder informeller Arbeit, was Ungleichheiten verschärft.¹³² Die Mitgliedstaaten sollten Maßnahmen ergreifen, um ausbeuterische Arbeitspraktiken zu bekämpfen. Frauen sind in der digitalen Industrie, der digitalen Wirtschaft

¹²⁹ CDADI/GEC-Studie (2023), S. 85.

¹³⁰ Siehe auch [KI-Verordnung](#) der EU (33).

¹³¹ [Freedom of Expression, Artificial Intelligence and Elections, UNESCO, Entwicklungsprogramm der Vereinten Nationen \(2025\) CI-2025/FEJ/ME-1.](#)

¹³² „Valuing Data Enrichment Workers: the Case for a Human-Centric Approach to AI Development“, zu Deutsch: „Die Wertschätzung von Mitarbeiterinnen und Mitarbeitern in der Datenanreicherung: Plädoyer für einen menschenzentrierten Ansatz bei der KI-Entwicklung“, Vereinte Nationen (2023); „Artificial Intelligence for Workers, Not Just For Profit: Ensuring Quality in Jobs in the Digital Age“, zu Deutsch: „Künstliche Intelligenz für Arbeitnehmer, nicht nur für den Profit: Sicherung der Arbeitsplatzqualität im digitalen Zeitalter“, Europäischer Gewerkschaftsbund (2025); „The Human Cost of DeepSeek“, zu Deutsch: „Die menschlichen Kosten von DeepSeek“, Antonio A. Casilli, Thomas Le Bonniec, Julian Posada. The Human Cost of DeepSeek.1(1), DiPLab. 2025, S. 10 fhfal-04952735.

sowie in Forschung und Wissenschaft nach wie vor stark unterrepräsentiert und durch die Automatisierung mittels generativer KI einem höheren Risiko des Arbeitsplatzverlusts ausgesetzt.¹³³ Dennoch sollten Frauen und von Diskriminierung bedrohte Personen in die Lage versetzt werden, KI-Innovationen und eine geschlechtergerechte digitale Welt aktiv mitzugestalten. Die Mitgliedstaaten sollten Maßnahmen in Betracht ziehen, die dazu beitragen, Frauen als aktive Akteure und Entscheidungsträger in der digitalen Industrie und Wirtschaft zu etablieren. Zu diesem Zweck ist es unerlässlich, den Zugang von Frauen zu Führungspositionen zu fördern und Frauen als Unternehmerinnen in der Digitalbranche zu unterstützen.

109. Allgemeiner ausgedrückt sollten die Mitgliedstaaten bei der Regulierung des Einsatzes von Fernüberwachungs- oder Kontrollsoftware in Bezug auf Beschäftigte und Studierende die zunehmende Verbreitung von Heimarbeit und Fernstudium sowie die damit verbundenen besonderen Herausforderungen für das Recht auf Privatsphäre und das Potenzial für Diskriminierung berücksichtigen.

110. Es gibt zudem Beispiele dafür, dass KI-Systeme eingesetzt werden, um die Gleichheit, einschließlich der Gleichheit der Geschlechter, in Einstellungsverfahren zu fördern, indem der Schwerpunkt auf Transparenz, Datenschutz und Rechenschaftspflicht gelegt wird. Beispielsweise durch die Beseitigung potenzieller Verzerrungsfaktoren bei Bewerbungen, wie Alter, Geschlecht, Nationalität und Wohnort. Allgemeiner betrachtet können KI-Systeme dazu eingesetzt werden, die Barrierefreiheit durch behindertenorientierte Lösungen zu fördern, wie beispielsweise Algorithmen zur Live-Untertitelung für gehörlose oder schwerhörige Menschen, Lösungen zur Anpassung der Umgebung, um Inhalte und Arbeitsplätze barrierefrei zu gestalten, oder Lösungen, die Menschen mit Behinderungen neue berufliche Möglichkeiten eröffnen, wie zum Beispiel die gezielte Einstellung von Menschen mit Neurodiversität durch Unternehmen im Bereich der Datenkennzeichnung.¹³⁴

Soziale Dienste, Wohnen und Wohlfahrt

111. Im Bereich der Sozialdienste und der Sozialfürsorge kann der Einsatz von KI-Systemen dazu führen, dass Einzelpersonen vom Zugang zu öffentlichen Dienstleistungen ausgeschlossen werden oder ihnen der Zugang verweigert wird und sie nicht von diesen profitieren können.¹³⁵ Wie in der KI-Verordnung hervorgehoben wird, „wenn KI-Systeme dazu verwendet werden, zu entscheiden, ob solche Leistungen und Dienste von Behörden gewährt, verweigert, gekürzt, widerrufen oder zurückgefordert werden sollen, einschließlich der Frage, ob die Begünstigten einen rechtmäßigen Anspruch auf solche Leistungen oder Dienste haben, können diese Systeme erhebliche Auswirkungen auf die Lebensgrundlage von Personen haben und deren Grundrechte verletzen.“¹³⁶ Darüber hinaus kann der Mangel an Rechenschaftspflicht und Transparenz im Zusammenhang mit dem Einsatz von KI-Systemen dazu führen, dass den betroffenen Personen eine Begründung oder die Möglichkeit, gegen die Entscheidungen Widerspruch einzulegen, vorenthalten wird.¹³⁷ Die Mitgliedstaaten sollten sicherstellen, dass bei der Nutzung von KI-Systemen zur Entscheidung darüber, ob Sozialleistungen, Sozialdienstleistungen, Wohnbeihilfen oder Sozialhilfe gewährt, verweigert, gekürzt, widerrufen oder zurückgefordert werden sollen, einschließlich der Frage, ob die Leistungsempfänger einen berechtigten Anspruch auf solche Leistungen oder Dienstleistungen haben, dass diese Nutzung keine Diskriminierung verursacht oder bestehende Ungleichheiten verstärkt.¹³⁸ Bei der Verwendung von KI-Systemen sollten die möglichen Schutzbedürfnisse der Nutzenden solcher Dienste berücksichtigt werden, darunter unter anderem deren Alter, und es sollten die erforderlichen Schutzmaßnahmen gewährleistet werden.

Gesundheit

112. Die Vorteile von KI-Systemen wurden sowohl im Hinblick auf die Verbesserung der Diagnose- und Therapieergebnisse für Patientinnen und Patienten¹³⁹ als auch im Hinblick auf die Erbringung medizinischer Leistungen, einschließlich Leistungen im Bereich der psychischen Gesundheit, festgestellt. Gleichzeitig wurden in diesem Bereich eine Reihe von Problemen identifiziert, die zu einer Verletzung des Gleichheitsgrundsatzes, einschließlich der Gleichheit der Geschlechter, sowie zu Diskriminierung führen.¹⁴⁰ Dazu gehört beispielsweise die geschlechtsspezifische Datenlücke im Gesundheitswesen, die zu weniger zuverlässigen Prognosesystemen

¹³³ „Generative AI and Jobs: A Refined Global Index of Occupational Exposure“, zu Deutsch: „Generative KI und Arbeitsplätze: Ein verfeinerter globaler Index zur beruflichen Exposition“, ILO, Mai 2025.

¹³⁴ „Using AI to support people with disability in the labour market: opportunities and challenges“, zu Deutsch: „Einsatz von KI zur Unterstützung von Menschen mit Behinderung auf dem Arbeitsmarkt: Chancen und Herausforderungen“, OECD Artificial Intelligence Papers, Nr. 7, 2023.

¹³⁵ CDADI/GEC-Studie (2023), S. 23.

¹³⁶ EU-KI-Gesetz (58).

¹³⁷ CDADI/GEC-Studie (2023), S. 24, und „Human rights by design – future-proofing human rights protection in the era of AI“ (2023), S. 6.

¹³⁸ CDADI/GEC Studie (2023), S. 22-23.

¹³⁹ Antworten des Ministerkomitees auf die Empfehlung 2185 (2020) zur „Artificial intelligence in health care: medical, legal and ethical challenges ahead“ zu Deutsch: „Künstlichen Intelligenz im Gesundheitswesen: Medizinische, rechtliche und ethische Herausforderungen“ der Parlamentarischen Versammlung des Europarats, Absatz 3.

¹⁴⁰ CDADI/GEC Study (2023), S. 25-26 und Empfehlung 2185 (2020) zu „Artificial intelligence in health care: medical, legal and ethical challenges ahead“.

bei der Diagnose von Patientinnen und geschlechtsdiversen Patienten führt.¹⁴¹ Die öffentliche Politikgestaltung im Gesundheitswesen und der weitere technologische Fortschritt sollten auf der uneingeschränkten Achtung der Menschenrechte, einschließlich der sozialen Rechte, beruhen.¹⁴² Folglich sollten die Mitgliedstaaten alle erforderlichen Anstrengungen unternehmen, um das Risiko eines diskriminierenden Zugangs zur Gesundheitsversorgung im Zusammenhang mit dem Einsatz von KI-Systemen zu mindern.¹⁴³ Die Mitgliedstaaten sollten zudem Maßnahmen ergreifen, um die Zuverlässigkeit analytischer und prädiktiver KI-Systeme im Gesundheitswesen zu verbessern, indem sie Datenlücken hinsichtlich der in Absatz 2 der Empfehlung aufgeführten persönlichen Merkmale beheben und sicherstellen, dass vielfältigere Modelle zur Bewertung der Risiken und der Wirksamkeit von Arzneimitteln, Therapien und anderen Instrumenten der Gesundheitsversorgung verwendet werden.

Migration

113. Im Zusammenhang mit der Grenzkontrolle und dem Grenzmanagement werden KI-Systeme unter anderem zu Vorhersagezwecken eingesetzt.¹⁴⁴ Wie in der KI-Verordnung hervorgehoben wird, „betreffen KI-Systeme, die in den Bereichen Migration, Asyl und Grenzkontrolle eingesetzt werden, Personen, die sich häufig in einer besonders schutzbedürftigen Lage befinden und vom Ausgang der Maßnahmen der zuständigen Behörden abhängig sind. Die Genauigkeit, die Nichtdiskriminierung und die Transparenz der in diesen Zusammenhängen eingesetzten KI-Systeme sind daher von besonderer Bedeutung, um die Achtung der Grundrechte der betroffenen Personen zu gewährleisten, insbesondere ihres Rechts auf Freizügigkeit, Nichtdiskriminierung, Schutz der Privatsphäre und personenbezogener Daten, internationalen Schutz sowie auf eine ordnungsgemäße Verwaltung.“¹⁴⁵ Die Parlamentarische Versammlung hat zudem betont, dass KI die menschliche Entscheidungsfindung in Migrations- und Asylverfahren unterstützen und nicht ersetzen sollte, auch wenn in einigen Fällen KI durch die Verringerung des Spielraums für menschliche Fehler möglicherweise mehr Sicherheit und Effektivität bietet als menschliche Entscheidungsfindung allein, und dass KI-Systeme in Migrations- und Asylverfahren vor ihrem Einsatz einer Folgenabschätzung hinsichtlich Menschenrechten, Demokratie und Rechtsstaatlichkeit unterzogen werden sollten.¹⁴⁶ Die Mitgliedstaaten sollten sicherstellen, dass der Einsatz von KI-Systemen in den Bereichen Migration, Asyl sowie Grenzkontrolle und -management nicht diskriminierend ist oder nicht zu Diskriminierung führt, indem sie Diskriminierungsrisiken identifizieren, bewerten, verhindern und mindern. Sie sollten den Rechten von Frauen und Mädchen sowie anderer Personen, die aufgrund der in Absatz 2 der Empfehlung aufgeführten Gründe von Diskriminierung bedroht sind, besondere Aufmerksamkeit widmen.¹⁴⁷

Bankwesen, Finanzwesen, Versicherungswesen und sonstige Dienstleistungsbranchen

114. Es gibt eine Reihe von Beispielen, die zeigen, wie algorithmische Verzerrungen die Entscheidungsfindung im Bank-, Finanz- und Versicherungswesen sowie in anderen Dienstleistungsbranchen beeinflussen, in denen KI-Systeme zum Einsatz kommen, was zu einer Diskriminierung der Dienstleistungsnutzenden führt.¹⁴⁸ Die Mitgliedstaaten sollten daher Maßnahmen ergreifen, um sicherzustellen, dass ein Rahmen vorhanden ist, der Verzerrungen verhindert und aufdeckt sowie die Auswirkungen der Entscheidungsfindung auf die Kunden im Hinblick auf Gleichbehandlung und Nichtdiskriminierung überwacht. Gleichzeitig können KI-Systeme auch dazu eingesetzt werden, voreingenommene Entscheidungen zu erkennen.¹⁴⁹

Weitere Bereiche, in denen mögliche negative Auswirkungen von KI-Systemen beobachtet und gegebenenfalls angegangen werden sollten

115. Über die Sektoren hinaus, bei denen das Risiko besonders gravierender Auswirkungen auf die Gleichheit, einschließlich der Gleichheit der Geschlechter, und die Nichtdiskriminierung festgestellt wurde, macht die Empfehlung auf weitere Bereiche aufmerksam, in denen die Mitgliedstaaten dazu angehalten werden, die

¹⁴¹ CDADI/GEC-Studie (2023), S. 25, unter Berufung auf Criado Perez C, „Invisible women: Exposing data bias in a world designed for men“, zu Deutsch: „Unsichtbare Frauen: Aufdeckung von Datenverzerrungen in einer für Männer gestalteten Welt“ (Random House 2019).

¹⁴² Antworten des Ministerkomitees auf Empfehlung 2185 (2020) zu „Künstliche Intelligenz im Gesundheitswesen: Medizinische, rechtliche und ethische Herausforderungen“, Absatz 2.

¹⁴³ Bericht zum Thema „Einsatz von KI im Gesundheitswesen und dessen Auswirkungen auf die Arzt-Patient-Beziehung“, Steuerungsausschuss für Menschenrechte in den Bereichen Biomedizin und Gesundheit (CDBIO), in dem auf das Risiko von Voreingenommenheit und Diskriminierung beim Einsatz von KI-Systemen im Gesundheitswesen hingewiesen wird, insbesondere S. 27.

¹⁴⁴ Empfehlung CM/Rec(2020)1 zu den Auswirkungen algorithmischer Systeme auf die Menschenrechte, Absatz 8.

¹⁴⁵ EU KI-Verordnung (60).

¹⁴⁶ Entschließung 2628 (2025) zu künstlicher Intelligenz und Migration, Parlamentarische Versammlung des Europarates, insbesondere die Ziffern 4, 7, 15, 20, 28 und 29.

¹⁴⁷ Empfehlung CM/Rec(2022)17 des Ministerkomitees an die Mitgliedstaaten zum Schutz der Rechte von Migrantinnen, Flüchtlingen und asylsuchenden Frauen und Mädchen, insbesondere die Absätze 22–23.

¹⁴⁸ Ibid, S. 21–22; „Human rights by design – future-proofing human rights protection in the era of AI“ (2023), S. 6; sowie das EU-KI-Gesetz (58).

¹⁴⁹ CDADI/GEC-Studie (2023), S. 63.

potenziellen negativen Auswirkungen des Einsatzes von KI-Systemen zu überwachen und gegebenenfalls zu bekämpfen, wie beispielsweise in den Bereichen Medien und Werbung sowie Sport und Kultur.

116. KI findet zunehmend Anwendung im Medien- und Werbesektor, was spezifische Risiken für die Gleichheit, einschließlich der Gleichheit der Geschlechter, und die Nichtdiskriminierung mit sich bringen kann, wie beispielsweise voreingenommene Darstellungen oder die Verstärkung schädlicher Stereotypen. In diesem Zusammenhang wird darauf hingewiesen, dass KI-Systeme auch dazu genutzt werden könnten, große Datenmengen zu analysieren und die Darstellung von Frauen und von Diskriminierungsrisiken betroffenen Personen in der Online- und Offline-Werbung zu bewerten.¹⁵⁰ Um potenziellen negativen Auswirkungen entgegenzuwirken, könnten die Mitgliedstaaten Maßnahmen in Betracht ziehen, die sicherstellen, dass bestehende Rechtsvorschriften, Strategien und regulatorische Rahmenbedingungen zur Verhinderung und Bekämpfung von Diskriminierung in Medien und Werbung auch den Einsatz von KI-Systemen abdecken, wobei die Unabhängigkeit der Medien zu wahren ist.¹⁵¹ Solche Maßnahmen könnten durch spezifische Schulungen für Journalisten und andere Fachkräfte der Branche ergänzt werden, in denen vermittelt wird, wie Sexismus, Rassismus, LGBTI-Feindlichkeit und andere Formen von Hass sowie sonstige Formen der Diskriminierung erkannt und bekämpft werden können und wie eine inklusive, positive und nicht stereotype Darstellung von Einzelpersonen und Gruppen in Medien und Werbung gefördert werden kann.¹⁵²

117. KI-Systeme finden im Sport zunehmend Anwendung, da sie in der Lage sind, große Datenmengen schnell zu analysieren, sportliche Muster und Trends zu erkennen, Leistungen auszuwerten sowie Statistiken zu erstellen und auszuwerten. Auf dem Gebiet der Verletzungsprävention können beispielsweise prädiktive Analysesysteme Warnmeldungen zu möglichen Muskelverletzungen generieren und die Leistung der Sportlerinnen und Sportler erheblich steigern; in Kontaktsportarten wurden KI-Systeme zur Entwicklung von Überwachungssystemen eingesetzt, die Stöße erkennen und deren Schweregrad automatisch bewerten, wodurch das Risiko einer Gehirnerschütterung für die Sportlerinnen und Sportler verringert wird. KI-Systeme können auch dazu genutzt werden, Sportkommentare zu analysieren und rassistische Vorurteile aufzudecken. In den Vereinigten Staaten deckte ein Forschungsprojekt unter Verwendung von Sprachverarbeitungsalgorithmen Muster voreingenommener Sprache auf, bei denen schwarze Sportlerinnen und Sportler unverhältnismäßig häufig mit körperlicher Leistungsfähigkeit in Verbindung gebracht wurden, während weiße Sportlerinnen und Sportler häufiger mit Begriffen beschrieben wurden, die mit Intellekt assoziiert werden.¹⁵³ Wenn zudem bestimmte Gruppen in den Datensätzen, die zum Trainieren von KI-Systemen verwendet werden, unterrepräsentiert sind, können diese Systeme bereits bestehende Vorurteile, einschließlich geschlechtsspezifischer Vorurteile, in der Sportkultur widerspiegeln und keine inklusiven und objektiven Analysen liefern. Dies könnte sich nicht nur auf ihre Gesundheit auswirken, sondern auch Entscheidungen hinsichtlich der Rekrutierung, der Leistungsbewertung von Sportlerinnen und Sportlern und der Ressourcenzuteilung beeinflussen, was zu Benachteiligungen für bestimmte Sportlerinnen und Sportler oder Sportlergruppen führen würde, beispielsweise dazu, dass Nachwuchstalente aus unterrepräsentierten Gruppen übersehen werden, wodurch Ungleichheiten im Sport weiter bestehen bleiben.

118. Im Zusammenhang mit Kultur können KI-Systeme unbeabsichtigt Stereotypen, darunter auch Geschlechterstereotypen, verfestigen, beispielsweise wenn die zum Trainieren der Algorithmen verwendeten Daten durch vorherrschende kulturelle Narrative verzerrt sind. KI-Systeme können zudem systemische Ungleichheiten reproduzieren, die durch Bedenken hinsichtlich Urheberrechten und geistigem Eigentum noch verstärkt werden können, wenn Inhalte aus Minderheitenkulturen ohne Wissen, Beteiligung und Zustimmung der betroffenen Gruppen entnommen werden.¹⁵⁴ Gleichzeitig haben KI-Systeme durch die Darstellung von Personen mit unterschiedlichem Hintergrund und aus verschiedenen Regionen in ihren Ergebnissen das Potenzial, die kulturelle Repräsentation historisch unterrepräsentierter Personen und Gruppen zu verbessern. Das Potenzial für die Förderung regionaler und Minderheitensprachen wurde bereits erwähnt. In Bereichen, in denen Frauen in der Medienberichterstattung unterrepräsentiert sind oder in denen ihre Leistungen unterbewertet werden, wie beispielsweise im Sport,¹⁵⁵ können KI-Systeme zu einer gleichberechtigten Sichtbarkeit von Frauen beitragen, indem sie Geschlechterstereotypen und traditionelle geschlechtsspezifische Vorurteile herausfordern.

¹⁵⁰ CDADI/GEC-Studie (2023), S. 77.

¹⁵¹ [Empfehlung CM/Rec\(2019\)1](#) zur Prävention und Bekämpfung von Sexismus, insbesondere II.B.1. und II.C. „Medien, Werbung und andere Kommunikationsprodukte und -dienste“; [Empfehlung CM/Rec\(2013\)1](#) des Ministerkomitees an die Mitgliedstaaten zu Gleichstellung und Medien, insbesondere Leitlinie A.1.; [Empfehlung CM/Rec\(2007\)17](#) zu Standards und Mechanismen zur Gleichstellung der Geschlechter, Absatz 47; sowie [die Empfehlung CM/Rec\(2022\)16](#) des Ministerkomitees an die Mitgliedstaaten zur Bekämpfung von Hassreden.

¹⁵² [Empfehlung CM/Rec\(2019\)1](#) zur Prävention und Bekämpfung von Sexismus, insbesondere Abschnitt II.C.6., sowie [Empfehlung CM/Rec\(2022\)16](#) zur Bekämpfung von Hassrede.

¹⁵³ Siehe „Can artificial intelligence help us understand racial bias in sports?“, zu Deutsch: „Kann künstliche Intelligenz uns helfen, rassistische Vorurteile im Sport zu verstehen?“, C. Brandon Ogbunu, 8. Januar 2020, abrufbar unter <https://andscape.com/features/artificial-intelligence-racial-bias-in-sports/>.

¹⁵⁴ Siehe beispielsweise die Leitlinien unter Berücksichtigung der neuesten technologischen Entwicklungen, wie beispielsweise der künstlichen Intelligenz, zur Ergänzung der Standards des Europarats in den Bereichen Kultur, Kreativität und Kulturerbe, März 2025.

¹⁵⁵ [„All In Plus: Promoting Greater Gender Equality in Sport“](#), Analytischer Bericht zu den erhobenen Daten: Einblicke in Trends und Fortschritte, Februar 2025, gemeinsames Projekt der Europäischen Union und des Europarats, das sowohl Fortschritte als auch anhaltende Ungleichheiten bei der Gleichstellung der Geschlechter im europäischen Sport aufzeigt.

Zu Absatz 39

119. In der Empfehlung wird betont, dass die Mitgliedstaaten weiterhin an ihre bestehenden Menschenrechtsverpflichtungen gebunden sind, insbesondere an diejenigen in Bezug auf Nichtdiskriminierung und Gleichheit, einschließlich der Gleichheit der Geschlechter; zwar werden sie dazu ermutigt, doch wird darauf hingewiesen, dass sie nicht verpflichtet sind, die Empfehlung auf Aktivitäten im Lebenszyklus von KI-Systemen anzuwenden, die unmittelbar den Schutz ihrer nationalen Sicherheitsinteressen betreffen. Solche Einschränkungen sollten nicht über die in Artikel 3 Absatz 2 des Rahmenübereinkommens und dessen Erläuterndem Bericht festgelegten hinausgehen.¹⁵⁶

¹⁵⁶ Siehe hierzu den Erläuternden Bericht zum Rahmenübereinkommen, Absatz 32.