EU AI Act: Stand der Umsetzung

Welche Regelungen gibt es bislang?

Supranational

- **Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD):** OECD-Grundsätze für künstliche Intelligenz (22. Mai 2019)
- **UNESCO:** UNESCO-Empfehlung zur Ethik der künstlichen Intelligenz (23. November 2021)
- **Europarat:** Rahmenübereinkommen über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit (17. Mai 2024), erster rechtverbindlicher internationaler Vertrag
- **EU:** AI Act

Gemeinsamkeit: Sicherstellung, dass AI-Systeme über den gesamten Lebenszyklus (bis zur Entsorgung) ethische Grundsätze und menschenrechtliche Anforderungen erfüllen.

Beispiel: UNESCO-Empfehlung

- Leitfaden für Regierungen
- Regierungen haben regelmäßige Berichtspflichten
- Von 193 Mitgliedstaaten verabschiedet
- Inhalt:
    - Werte und ethische Prinzipien (Folie, s.u.)
    - Konkrete Gestaltungsaufgaben für elf Politikfelder
    - Übergreifende Fragen ethischer Folgenabschätzung und Regulierung

National

USA

- **Bundesprogramme:**
  - seit 2016 Förderung von AI
  - seit 2019 Executive Order: AI keine Hindernisse in den Weg legen, also insbesondere nicht reglementieren
  - 1.1.2021: National Artificial Initiative Act: Koordinierung AI-bezogener Aktivitäten der Regierung
- **Gesetze:**
  - US-Bundesstaaten haben Gesetze zu Einzelfragen im Umgang mit AI-Systemen verabschiedet. Gesetze befassen sich meist mit technischen Fragen bzw. betreffen nur bestimmte Einsatzbereiche von AI-Systemen. In Illinois müssen Arbeitgeber Bewerber vorab darauf hinweisen, wenn sie Bewerbungsvideos mittels AI-Systemen auswerten möchten.
  - Anfang 2022 Entwurf eines Algorithmic Accountability Act 2022. Verpflichtet bestimmte Unternehmen dazu, bestimmte Systeme zur automatisierten Entscheidungsfindung einer Folgenabschätzung zu unterziehen (davor und in regelmäßigen Abständen).
- Verbindliche und unverbindliche **behördliche Regulierungen**
- **Aktuell:** Trump unterzeichnete ein Dekret, das die Ausarbeitung eines "Aktionsplans für Künstliche Intelligenz" innerhalb von 180 Tagen vorsieht. Ziel des Plans ist es, "Amerikas globale KI-Dominanz zu erhalten und auszubauen" wohl im Zusammenhang mit einer umfassenden Deregulierungsagenda, also Zurückfahren von Schutzmaßnahmen.

China

- **Grundkonstellation:**
  - Mehrere nebeneinanderstehende und nicht abgestimmte Regelungsinitiativen
  - Mehrere sektorübergreifende und sektorspezifische Regelungen
- Die in China diskutierten Vorgaben für vertrauenswürdige AI-Systeme gleichen vom Wortlaut her den in westlichen Staaten diskutierten Mindeststandards. Die Inhalte variieren jedoch aufgrund des unterschiedlichen Werteverständnisses.

Zwischenergebnis

- Der von der EU favorisierte **sektorübergreifende und produktrechtliche Ansatz** wird von den USA und China bislang **nicht verfolgt.**
- Die **USA** bevorzugen derzeit einen **ergebnisbezogenen technologieneutralen Ansatz.** Dieser lässt den Entwicklern **weite Freiräume** bei der Gestaltung von AI-Systemen.

Deutschland

- Hier gilt der **AI Act,** der **umgesetzt** werden muss (s.u.).
- **Initiativen** zu einer **eigenen deutschen Regulierung** gab es bislang beim **Beschäftigtendatenschutz** und dem damit zusammenhängenden Einsatz von KI.
    - Derzeit regelt allein § 26 BDSG den Beschäftigtendatenschutz. Dieser dürfte aber nach einem Urteil des EuGH vom 30.03.2023 (Az. C-34/21) **europarechtswidrig** sein und deswegen die Regelung des Beschäftigtendatenschutzes dringlicher denn je machen.
    - Es soll vor allem für **Transparenz** bei der Verwendung von KI gesorgt werden.
    - Außerdem soll KI **in bestimmten Bereichen reglementiert** werden, z.B. im Einstellungsverfahren und bei der Leistungsbeurteilung sowie Beförderung.
    - Jetzt aber Opfer der **Diskontinuität.**

Welche Regelungen sieht der AI Act vor?

Einführung

- Die **erste umfassende KI-Regulierung weltweit**
- **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 …
    - Kurztitel: Verordnung über künstliche Intelligenz
    - Abkürzungen: KI-Verordnung (KIVO) oder AI Act
- **Werdegang:**
    - 2019 ausgearbeitet
    - am 21. Mai 2024 von den 27 EU-Mitgliedsstaaten endgültig verabschiedet
    - am 12. Juli 2024 verkündet
    - zum 1. August 2024 in Kraft getreten
- **Gestaffelte Anwendung** (s.u.):
    - größtenteils ab 2. August 2026
    - sonst 2. Februar 2025, 2. August 2025 und 2. August 2027
- **Ziele:**
    - Schaffung eines **Schutzniveaus**, das folgendes gewährleisten soll:
        - Sicherung der Grundrechte (einschl. Demokratie, Rechtsstaatlichkeit und Umweltschutz
        - Keine negativen Auswirkungen auf Gesundheit und Sicherheit der Menschen
        - Ethische Verantwortbarkeit von KI-Systemen
    - Funktionieren des **Binnenmarktes** verbessern
    - **KI fördern** (insbesondere Investitionen)

Regulierungsmethode

- **Risikobasierter Ansatz** → sorgt für einheitliche Grundregeln
    - Anforderungen richten sich nach dem Risikoniveau der KI-Anwendung, je höher das Risiko für Grundrechts- oder Sicherheitsgefährdung desto höher die Anforderungen, also abgestufte Verbote bzw. Compliance[1]-, Berichts-, Dokumentations-, Sorgfalts- und Informationspflichten
    - Darüber hinaus spezielle Regeln für multimodale Modelle (GPAI) und solche, die besonders hohe Rechenleistungen haben (bspw. besonders leistungsfähige GenAI-Modelle wie ChatGPT).
- **Sektorübergreifender Ansatz** → sorgt für Verhältnismäßigkeit
    - gilt über alle Wirtschaftssektoren hinweg, es sei denn es gibt für bestimmte Bereiche spezifische Regelunge, z.B. Medizinprodukte
- **Produktrechtlicher Ansatz** → sorgt für praktische Umsetzbarkeit
    - Anwendungen werden nach der Logik des EU-Produktsicherheitsrechts behandelt:
        - Hersteller müssen Konformität (nach harmonisierten Standards und Bewertungsverfahren) nachweisen
        - CE-Kennzeichnung
        - Marktüberwachung durch Behörden.
    - Dadurch
        - klare Verantwortlichkeiten
        - Einbindung in bestehende Strukturen und
        - Vereinfachung des internationalen Handels.
    - Heißt aber auch:
        - Keine Regulierung des Outputs von AI-Systemen.
        - Außerdem sind keine subjektiven Rechte für betroffene Personen vorgesehen. Dafür sollen die bereichsspezifischen Regelungen ausreichen, etwa Datenschutz- oder Antidiskriminierungsregeln.

Regulierungsmittel

- Anwendbarkeitsverbote
- Pflichten für Entwickler, Anbieter, Nutzer etc.
- Marktbeobachtung und Marktüberwachung
- Aufsicht (vor allem durch nationale Behörden)
- Besonderheiten
    - Einrichtung von **KI-Reallaboren:** bieten einen kontrollierten Rahmen, der von einer zuständigen Behörde geschaffen wird und den Anbieter oder zukünftige Anbieter von KI-Systemen nach einem Plan für das Reallabor einen begrenzten Zeitraum und unter regulatorischer Aufsicht nutzen können, um ein innovatives KI-System zu entwickeln, zu trainieren, zu validieren und — gegebenenfalls unter Realbedingungen — zu testen.
    - Einrichtung eines **KI-Büros:** Beratung und Unterstützung der EU-Kommission

[1] Gesamtheit aller betrieblichen Maßnahmen, die das rechtmäßige Verhalten aller Unternehmensangehörigen sicherstellen sollen.

Akteure (Art. 3 KIVO)

Anbieter GAPI-Modell oder KI-System → Einführer, Händler, Bevollmächtigte → Betreiber eines KI-Systems

- **Anbieter:** natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System oder ein KI-Modell mit allgemeinem Verwendungszweck entwickelt oder entwickeln lässt und es unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringt oder das KI-System unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nimmt, sei es entgeltlich oder unentgeltlich;
- **Betreiber:** eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet (Haushaltsausnahme); [nicht gemeint sind Endbenutzer]
- **Bevollmächtigter:** eine in der Union ansässige oder niedergelassene natürliche oder juristische Person, die vom Anbieter eines KI-Systems oder eines KI-Modells mit allgemeinem Verwendungszweck schriftlich dazu bevollmächtigt wurde und sich damit einverstanden erklärt hat, in seinem Namen die in dieser Verordnung festgelegten Pflichten zu erfüllen bzw. Verfahren durchzuführen;
- **Einführer:** eine in der Union ansässige oder niedergelassene natürliche oder juristische Person, die ein KI-System, das den Namen oder die Handelsmarke einer in einem Drittland niedergelassenen natürlichen oder juristischen Person trägt, in Verkehr bringt;
- **Händler:** natürliche oder juristische Person in der Lieferkette, die ein KI-System auf dem Unionsmarkt bereitstellt, mit Ausnahme des Anbieters oder des Einführers;

Zudem sind bestimmte **Produkthersteller indirekt betroffen,** wenn sie ein KI-System in potenziell riskante Produkte wie Maschinen oder Medizingeräte integrieren.

Die **meisten Pflichten** der KI-Verordnung treffen die **Anbieter** von KI-Systemen/GPAI-Modellen und die **Betreiber** von KI-Systemen.

Anwendbarkeit

- **Persönlicher Anwendungsbereich** (Art. 2 Abs. 1 KIVO), insbesondere:
    - **Niederlassungsprinzip:** Betreiber von KI-Systemen, die ihren Sitz in der Union haben oder in der Union befinden.
    - **Marktortprinzip:** Anbieter von AI-Systemen, die KI-System in der EU in Verkehr bringen oder in Betrieb nehmen oder GPAI in Verkehr bringen (unabhängig, ob Anbieter in der EU oder einem Drittland niedergelassen ist)
    - **„Outputprinzip“:** Anbieter und Betreiber von KI-Systemen in Drittländern, wenn Ausgabe des KI-Systems in der EU verwendet wird (wirft viele Fragen auf)
    - Erfasst sind **auch:**
        - Einführer und Händler von KI-Systemen
        - Produkthersteller, die KI-Systeme zusammen mit ihrem Produkt unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringen oder in Betrieb nehmen
        - Bevollmächtigte von Anbietern, die nicht in der EU niedergelassen sind.

- **Sachlicher Anwendungsbereich:** Definition von KI (Art. 3 Nr. 1 KIVO)
    - **„KI-System“** [bezeichnet] ein maschinengestütztes System, das für einen in unterschiedlichem Grade **autonomen** Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme **anpassungsfähig** sein kann und 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;
    - **„KI-Modell mit allgemeinem Verwendungszweck“:** KI-Modell — einschließlich der Fälle, in denen ein solches KI-Modell mit einer großen Datenmenge unter umfassender Selbstüberwachung trainiert wird —, das eine erhebliche allgemeine Verwendbarkeit aufweist und in der Lage ist, unabhängig von der Art und Weise seines Inverkehrbringens ein breites Spektrum unterschiedlicher Aufgaben kompetent zu erfüllen, und das in eine Vielzahl nachgelagerter Systeme oder Anwendungen integriert werden kann, ausgenommen KI-Modelle, die vor ihrem Inverkehrbringen für Forschungs- und Entwicklungstätigkeiten oder die Konzipierung von Prototypen eingesetzt werden;
    - **Nicht erfasst**
        - KI-Systeme, die ausschließlich für **militärische Zwecke** entwickelt oder verwendet werden.
        - KI-Systeme oder -Modelle einschließlich ihrer Ausgabe, die eigens für den alleinigen Zweck der **wissenschaftlichen Forschung und Entwicklung** entwickelt und in Betrieb genommen werden.
        - KI-Systeme, die unter **freien und quelloffenen Lizenzen** bereitgestellt werden, es sei denn, sie werden als Hochrisiko-KI-Systeme oder als ein KI-System, das unter Artikel 5 oder 50 fällt, in Verkehr gebracht oder in Betrieb genommen.
    - **Fazit:** Sehr weite Definition mit weitem Anwendungsbereich
        - an sich gewollt, **Technologieoffenheit**,
        - was aufgrund von Abgrenzungsfragen aber auch zu **Rechtsunsicherheit** führen kann
        - letztlich mehr **Regulierung von Software** als von KI?
- **Räumlicher Anwendungsbereich:** EWR

Risikobasierter Ansatz

Abstufung

- **KI-Systeme mit inakzeptablem Risiko** (Kap. II, Art. 5 KIVO):
  Diese KI-Anwendungen sind **verboten** (Ausnahmen für Militär und Geheimdienste)
- **KI-Systeme mit hohem Risiko** (Kap. III Art. 6 – 49 KIVO):
  Strenge Anforderungen und Pflichten (der größte Teil der KIVO befasst sich mit dieser Hochrisiko-Klasse)
- **KI-Systeme mit Transparenzanforderungen** (begrenztes Risiko, Kap. IV Art. 50 KIVO):
  Hier werden lediglich Transparenzanforderungen aufgestellt
- **KI-Systeme mit keinem/niedrigem Risiko** (minimales Risiko):
  Keine spezifischen Pflichten (bzw. nicht reguliert)
- **Sonderrolle GPAI** (Kap. V, Art. 51 – 56 KIVO):
  Müssen spezielle Transparenzanforderungen erfüllen. Wenn sie darüber hinaus systemische Risiken verursachen, gelten strengere Regeln.

KI-Systeme mit inakzeptablem Risiko

Klassifikation: Folgende Praktiken im KI-Bereich sind verboten

- Einsatz von **unterschwelligen, manipulativen oder täuschenden Techniken,** um das Verhalten zu verzerren und die bewusste Entscheidungsfindung zu beeinträchtigen, wodurch erheblicher Schaden entsteht.
- **Ausnutzung von Schwachstellen** im Zusammenhang mit Alter, Behinderung oder sozio-ökonomischen Verhältnissen, um das Verhalten zu verzerren und erheblichen Schaden anzurichten.
- **biometrische Kategorisierungssysteme, die Rückschlüsse auf sensible Merkmale** (Rasse, politische Meinungen, Gewerkschaftszugehörigkeit, religiöse oder philosophische Überzeugungen, Sexualleben oder sexuelle Ausrichtung) zulassen, außer bei der Kennzeichnung oder Filterung rechtmäßig erworbener biometrischer Datensätze oder bei der Kategorisierung biometrischer Daten durch die Strafverfolgungsbehörden.
- **Soziales Scoring,** d. h. die Bewertung oder Klassifizierung von Personen oder Gruppen aufgrund ihres Sozialverhaltens oder ihrer persönlichen Eigenschaften, was zu einer nachteiligen oder ungünstigen Behandlung dieser Personen führt.
- die **Bewertung des Risikos, dass eine Person Straftaten begeht**, ausschließlich auf der Grundlage von Profilen oder Persönlichkeitsmerkmalen, es sei denn, sie wird zur Ergänzung menschlicher Bewertungen auf der Grundlage objektiver, überprüfbarer Fakten, die in direktem Zusammenhang mit kriminellen Aktivitäten stehen, verwendet.
- Aufbau von **Gesichtserkennungsdatenbanken** durch ungezieltes Auslesen von Gesichtsbildern aus dem Internet oder aus Videoüberwachungsaufnahmen.
- das **Ableiten von Emotionen am Arbeitsplatz oder in Bildungseinrichtungen,** außer aus medizinischen oder Sicherheitsgründen.
- **Biometrische Fernidentifizierung (RBI) in Echtzeit** in öffentlich zugänglichen Räumen für die Strafverfolgung außer wenn:
    - Suche nach vermissten Personen, Entführungsopfern und Menschen, die Opfer von Menschenhandel oder sexueller Ausbeutung geworden sind;
    - Verhinderung einer erheblichen und unmittelbaren Bedrohung des Lebens oder eines vorhersehbaren terroristischen Angriffs; oder
    - Identifizierung von Verdächtigen bei schweren Straftaten (z. B. Mord, Vergewaltigung, bewaffneter Raubüberfall, Drogen- und illegaler Waffenhandel, organisierte Kriminalität, Umweltkriminalität usw.).

  Außerdem weitere Voraussetzungen, wenn ausnahmsweise RBI zulässig.

KI-Systeme mit hohem Risiko

Klassifizierungssystem

- KI-System unterliegt bestimmten **Harmonisierungsvorschriften** (s. Anlage I) und es muss eine **Vorab-Konformitätsbewertung** durchgeführt werden
- Oder: **KI-System aus Anlage III**, außer wenn:
    - das KI-System führt eine enge prozedurale Aufgabe aus;
    - verbessert das Ergebnis einer zuvor durchgeführten menschlichen Tätigkeit;
    - Entscheidungsmuster oder Abweichungen von früheren Entscheidungsmustern aufdeckt und nicht dazu gedacht ist, die zuvor durchgeführte menschliche Bewertung ohne angemessene menschliche Überprüfung zu ersetzen oder zu beeinflussen; oder

    - führt eine vorbereitende Aufgabe für eine Bewertung durch, die für die in Anhang III aufgeführten Anwendungsfälle relevant ist.
- KI-Systeme gelten **immer dann als risikoreich, wenn sie ein Profil von Personen erstellen,** d. h. wenn sie personenbezogene Daten automatisiert verarbeiten, um verschiedene Aspekte des Lebens einer Person zu bewerten, z. B. Arbeitsleistung, wirtschaftliche Lage, Gesundheit, Vorlieben, Interessen, Zuverlässigkeit, Verhalten, Standort oder Bewegung.
- Anbieter, deren KI-System unter die Anwendungsfälle in Anhang III fällt, die aber der Ansicht sind, dass es **kein hohes Risiko** darstellt, müssen eine solche **Bewertung dokumentieren,** bevor sie es in Verkehr bringen oder in Betrieb nehmen.
- **Kommission** hat Möglichkeit, die **Listen zu erweitern** (Technologieoffenheit).
- **Anhang III Anwendungsfälle**
    - **Nicht verbotene biometrische Verfahren:** Biometrische Fernidentifikationssysteme, mit Ausnahme der biometrischen Überprüfung, die bestätigt, dass eine Person diejenige ist, die sie vorgibt zu sein. Biometrische Kategorisierungssysteme, die auf sensible oder geschützte Attribute oder Merkmale schließen. Systeme zur Erkennung von Emotionen.
    - **Kritische Infrastrukturen:** Sicherheitskomponenten bei der Verwaltung und dem Betrieb kritischer digitaler Infrastrukturen, im Straßenverkehr und bei der Versorgung mit Wasser, Gas, Wärme und Strom.
    - **Allgemeine und berufliche Bildung:** KI-Systeme, die den Zugang, die Zulassung oder die Zuweisung zu Bildungs- und Berufsbildungseinrichtungen auf allen Ebenen bestimmen. Bewertung von Lernergebnissen, einschließlich derer, die zur Steuerung des Lernprozesses des Schülers verwendet werden. Bewertung des angemessenen Bildungsniveaus für eine Person. Überwachung und Erkennung von unzulässigem Schülerverhalten bei Prüfungen.
    - **Beschäftigung, Arbeitnehmermanagement und Zugang zur Selbständigkeit:** KI-Systeme für die Einstellung oder Auswahl, insbesondere für gezielte Stellenanzeigen, die Analyse und Filterung von Bewerbungen und die Bewertung von Kandidaten. Beförderung und Beendigung von Verträgen, Zuweisung von Aufgaben auf der Grundlage von Persönlichkeitsmerkmalen oder Eigenschaften und Verhalten sowie Überwachung und Bewertung der Leistung.
    - **Zugang zu und Inanspruchnahme von wesentlichen öffentlichen und privaten Dienstleistungen:** KI-Systeme, die von Behörden zur Beurteilung der Anspruchsberechtigung auf Leistungen und Dienste, einschließlich deren Zuweisung, Kürzung, Entzug oder Rückforderung, verwendet werden. Bewertung der Kreditwürdigkeit, außer bei der Aufdeckung von Finanzbetrug. Bewertung und Klassifizierung von Notrufen, einschließlich der Festlegung von Prioritäten für den Einsatz von Polizei, Feuerwehr, medizinischer Hilfe und Dringlichkeitsdiensten für Patienten. Risikobewertung und Tarifierung in der Kranken- und Lebensversicherung.
    - **Strafverfolgung:** KI-Systeme zur Bewertung des Risikos einer Person, Opfer eines Verbrechens zu werden. Polygraphen. Bewertung der Zuverlässigkeit von Beweisen bei strafrechtlichen Ermittlungen oder Strafverfolgungen. Bewertung des Risikos einer Person, straffällig zu werden oder erneut straffällig zu werden, nicht nur auf der Grundlage eines Profils oder der Bewertung von Persönlichkeitsmerkmalen oder früherem kriminellen Verhalten. Profiling bei der Ermittlung, Untersuchung oder Verfolgung von Straftaten.

- **Migrations-, Asyl- und Grenzkontrollmanagement:** Lügendetektoren. Bewertung von irregulärer Migration oder Gesundheitsrisiken. Prüfung von Anträgen auf Asyl, Visa und Aufenthaltsgenehmigungen und damit zusammenhängende Beschwerden in Bezug auf die Anspruchsberechtigung. Aufspüren, Erkennen oder Identifizieren von Personen, mit Ausnahme der Überprüfung von Reisedokumenten.
- **Rechtspflege und demokratische Prozesse:** KI-Systeme, die bei der Erforschung und Auslegung von Fakten und der Anwendung des Rechts auf konkrete Sachverhalte oder bei der alternativen Streitbeilegung eingesetzt werden. Beeinflussung der Ergebnisse von Wahlen und Volksabstimmungen oder des Abstimmungsverhaltens, mit Ausnahme von Ergebnissen, die nicht direkt mit Menschen interagieren, wie z. B. Werkzeuge, die zur Organisation, Optimierung und Strukturierung politischer Kampagnen verwendet werden.

Anforderungen

- Einrichtung eines **Risikomanagementsystems** für den gesamten Lebenszyklus des KI-Systems mit hohem Risiko;
- Durchführung von **Data Governance,** um sicherzustellen, dass die Schulungs-, Validierungs- und Testdatensätze relevant, ausreichend repräsentativ und so weit wie möglich fehlerfrei und vollständig sind.
- Erstellung von **technischen Unterlagen** zum Nachweis der Konformität und Bereitstellung von Informationen für die Behörden, um die Konformität zu bewerten.
- ihr KI-System für hohe Risiken so zu gestalten, dass es Ereignisse, die für die Identifizierung von Risiken auf nationaler Ebene relevant sind, und wesentliche Änderungen während des gesamten Lebenszyklus des Systems automatisch **aufzeichnen** kann.
- Bereitstellung von **Gebrauchsanweisungen** für nachgeschaltete Verteiler, damit diese die Vorschriften einhalten können.
- ihr KI-Hochrisikosystem so zu gestalten, dass die Einsatzkräfte die Möglichkeit haben, eine **menschliche Aufsicht** zu implementieren.
- ihr risikoreiches KI-System so zu gestalten, dass es ein angemessenes Maß an **Genauigkeit, Robustheit und Cybersicherheit** erreicht.
- Einrichtung eines **Qualitätsmanagementsystems** zur Gewährleistung der Einhaltung der Vorschriften.

KI-Systeme mit Transparenzanforderungen

Fallgruppen und Anforderungen

- **KI-Systeme zur direkten Nutzerinteraktion** (Art. 50 Abs. 1 KIVO): Pflicht zur Information über KI-Nutzerinteraktion
- **Generative KI-Systeme** (Art. 50 Abs. 2 KIVO): Pflicht zur Kennzeichnung als maschinenlesbar und künstlich erzeugt/manipuliert
- **KI-Systeme zur Emotionserkennung** (Art. 50 Abs. 3 KIVO): Pflicht zur Information natürlicher Personen, und der Verarbeitung personenbezogener Daten
- **KI-System zur Generierung von Deep Fakes** (Art. 50 Abs. 4 KIVO): Offenlegungspflicht

Beispiele

- Chatbots für grundlegende Kundeninteraktion
- Produktempfehlungssysteme
- Automatisierte Preisanpassungsalgorithmen für nicht-essenzielle Güter
- Inhaltsmoderation für nicht sensible Bereiche
- Inventarverwaltung (Vorhersage von Bestandsbedürfnissen und die Optimierung von Lagerbeständen)
- Kundenstimmungsanalyse (um Zufriedenheitsniveaus zu messen)
- Website-Personalisierung (z. B. Ändern von Layouts oder das Hervorheben bestimmter Produkte entsprechend dem Kundenverhalten)

Besonderheiten

- **Informationspflicht** spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung in klarer und eindeutiger Weise (Art. 50 Abs. 5 KIVO).
- KI-Büro soll **Praxisleitfäden** für die Umsetzung der Pflichten zur Feststellung und Kennzeichnung künstlich erzeugter oder manipulierter Inhalte erstellen (Art. 50 Abs. 6 KIVO)

KI-Systeme mit keinem/niedrigem Risiko

- Beispiele
    - KI-Einsatz bei Spamfiltern
    - KI-gestützte Videospiele

Sonderrolle GPAI

- **Unterscheidung** GPAI-Model (GPAIM, Art. 3 Nr. 63 KIVO, Erwägungsgrund 97) und GPAI-System – grob:
    - **KI-System:** funktionsfähige und mit einer Benutzeroberfläche ausgestattete KI-Anwendung;
      → Hierbei handelt es sich um gewöhnliche KI-Systeme, für die die entsprechenden Regelungen gelten (s. Erwägungsgrund 85); können also auch als Hochrisikosystem eingestuft werden
    - **KI-Modell:** das dahinterstehende (technische) Herzstück, also die KI-gestützte Funktionsweise; Algorithmus und seine Gewichtungen.
      → KIVO hält eigenständige Regelung bereit
      Abgrenzung gewöhnliches KI-Modell zu GPAIM (s. Erwägungsgrund 99):
        - Erhebliche Allgemeinheit
        - Kann für breites Spektrum unterschiedlicher Aufgaben eingesetzt werden
        - Ermöglicht flexible Erzeugung von Inhalten, insbesondere von z.B. Text, Audio oder Video

      LLMs sind GPAIM
- Nur **Pflichten für Anbieter** (es gibt keinen Betreiber)
- **Grundsätzlicher Pflichtenkatalog** für Anbieter von GPAIM (Art. 53, 54 KIVO)
    - Technische Dokumentationspflicht
    - Ausarbeitung von Informationen und Unterlagen für nachgelagerte Anbieter, die das GPAI-Modell in ihr eigenes KI-System integrieren wollen, damit diese die Möglichkeiten und Grenzen verstehen und in die Lage versetzt werden, die Anforderungen zu erfüllen.
    - Festlegung einer Regelung zur Einhaltung des Urheberrechts
    - Transparenzpflicht zu Trainingsmodalitäten des KI-Modells

    - Benennungspflicht eines Bevollmächtigten für Anbieter mit Niederlassung in Drittländern
- **Zusätzliche Pflicht bei Anbietern von GPAI mit systemischen Risiken** (Def. in Art. 51 KIVO, Pflichten in Art. 55 KIVO):
    - Mitteilungspflicht gegenüber EU-Kommission (Art. 52 KIVO)
    - Erweiterte technische Dokumentationspflicht
    - Risikomanagement für systemische Risiken (erkennen/bewerten/mindern)
    - Informationspflicht gegenüber KI-Büro bei schwerwiegenden Vorfällen
    - Gewährleistung von Cybersicherheit für das KI-Modell und die physische Infrastruktur
- **Abgrenzungsmerkmal „systemische Risiken“**:
    - Voraussetzungen:
        - Fähigkeiten mit hohem Wirkungsgrad
        - Erfüllen Kriterien aus Anhang XIII
    - Voraussetzungen liegen insbesondere vor, wenn der kumulierte Rechenaufwand für die Ausbildung mehr als 1025 Gleitkommaoperationen (FLOPs) beträgt.
- Alle Anbieter von GPAI-Modellen können die **Einhaltung ihrer Verpflichtungen** nachweisen, indem sie sich freiwillig an einen **Verhaltenskodex** halten, bis harmonisierte europäische Normen veröffentlicht werden, deren Einhaltung zu einer Konformitätsvermutung führt. Anbieter, die sich nicht an einen Verhaltenskodex halten, müssen für die Genehmigung durch die Kommission alternative angemessene Mittel zur Einhaltung der Vorschriften nachweisen.
- **Erleichterungen:**
    - **Praxisleitfäden** zur Erleichterung des Nachweises der erfüllten Pflichten nach Art. 55 Abs. 1 KIVO
    - Art. 52 Abs. 2 und Abs. 5 KIVO bieten Möglichkeiten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck/ GPAI mit systemischem Risiko, **nicht als ein solches Modell mit systemischem Risiko eingestuft zu werden.** Die Anbieter müssen der Kommission innerhalb von 2 Wochen mitteilen, ob ihr Modell dieses Kriterium erfüllt. Der Anbieter kann Argumente dafür vorbringen, dass sein Modell trotz Erfüllung der Kriterien keine systemischen Risiken birgt. Die Kommission kann von sich aus oder durch eine qualifizierte Warnung des wissenschaftlichen Gremiums unabhängiger Experten entscheiden, dass ein Modell hohe Auswirkungen hat und damit systemrelevant ist.
    - Zudem kennt der AI Act **Ausnahmen** für bestimmte Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck/ GPAI ohne systemische Risiken, wenn das KI-Modell im Rahmen einer freien und quelloffenen Lizenz bereitgestellt wird und weitere Voraussetzungen erfüllt werden.

Überwachung und Sanktionen

- **Überwachung:**
  - Aufsicht über GPAI-Modelle: KI-Büro
  - Aufsicht über KI-Systeme: nationale Aufsichtsbehörden

  Zudem Mechanismen zur **einheitlichen Auslegung** der KIVO

- **Bußgelder** hängen von der **Schwere des Verstoßes** ab:
  - Bis zu 35 Mio. EUR oder 7 % des weltweiten Jahresumsatzes für verbotene KI-Systeme.
  - Bis zu 15 Mio. EUR oder 3 % des Umsatzes für Verstöße gegen die Pflichten für Hochrisiko-Systeme oder die Transparenzpflichten für Anbieter und Betreiber gemäß Artikel 50 KI-Verordnung.
  - Bis zu 7,5 Mio. EUR oder 1 % des Umsatzes für fehlerhafte Informationen gegenüber Behörden.

  Ob sich das Bußgeld an dem festen Geldbetrag orientiert oder an dem Jahresumsatz, hängt davon ab, welcher Betrag höher ist. Bei KMUs und Start-ups gilt jeweils der niedrigere Betrag (vgl. Art. 99 Abs. 6 KIVO).

Chancen und Herausforderungen

- **Chancen:**
  - Einheitlicher Rechtsrahmen schafft **Rechtssicherheit** für Entwicklung und Einsatz von KI
  - **Vertrauensbildung** bei Nutzern (Schutz der Grundrechte, Transparenz)
  - Europäischer Qualitätsstandard als **globaler Maßstab und Wettbewerbsvorteil**
  - **Innovationsförderung durch geschützte Räume,** in denen Unternehmen ihre Neuerungen testen können
- **Herausforderungen:**
  - **Überschneidungen** mit anderen Rechtsakten und widersprüchliche Anforderungen, insbesondere DSA und DSGVO und MPR
  - **Unterschiedliche Anforderungen** in den einzelnen Sektoren (z.B. Finanz-, Medizin- und Automobilbereich, s. Bertelsmann-Studie)
  - **Komplexität der Klassifizierung**
  - **Technische Anforderungen an Hochrisiko-Systeme**
  - Rechtliche Hürden und dadurch verursachte **hemmen die internationale Wettbewerbsfähigkeit** (insbesondere für KMU)
  - **Ressourcenaufwand für Compliance** (z.B. hoher bürokratischer Aufwand und doppelte Regulierung etwa in der Medizin)
  - **Unterschiedliche Anwendung** der Verordnung **in den Mitgliedstaaten**

Umsetzung

Zeitplan

- **2024:**
  - 1.8.2024: Inkrafttreten
- **2025:**
  - 2.2.2025: Das Verbot bestimmter KI-Systeme und die Anforderungen an die KI-Kompetenz beginnen zu gelten (Kapitel 1 und Kapitel 2).
  - 2.5.2025: Verhaltenskodizes für GPAIM müssen fertiggestellt sein
  - 2.8.2025: U.a. die folgenden Regeln beginnen zu gelten:
    - Vorgaben zu GPAI-Modelle (Kapitel V)
    - Governance (Kapitel VII): Einrichtung nat. Aufsichtsbehörden und KI-Büro
    - Sanktionen (Artikel 99 und 100)
- **2026:**
  - 2.8.2026:
    - Die übrigen Bestimmungen des AI-Gesetzes finden Anwendung (allgemeine Anwendbarkeit), insbesondere Anwendbarkeit der Vorschriften zu Hochrisikosystemen (mit Ausnahme der Klassifizierungsregeln)
    - Deadline für Innovationsmaßnahmen wie die Einrichtung mind. Eines nationalen KI-Reallabors
- **2027:**
  - 2.8.2027: Art. 6 Abs. 1 KIVO (Klassifizierungsregeln für Hochrisikosysteme) und die entsprechenden Verpflichtungen beginnen zu gelten.
- Außerdem gibt es **Regelungen zum Bestandsschutz** nach Art. 111 KI-VO für KI-Systeme/GPAI-Modelle, die vor der Anwendbarkeit der sie betreffenden Regelungen der KI-Verordnung in Verkehr gebracht oder in Betrieb genommen wurden.

Aufbau der Behörden auf nationaler Ebene

- **„zuständige nationale Behörde“:**
  - **Marktüberwachungsbehörde:** Aufsicht ab Inbetriebnahme/Inverkehrbringen des KI-System

    (Art. 79 KI-VO regelt die Zusammenarbeit der nationalen Marktüberwachungsbehörden. Bei Uneinigkeit unter ihnen prüft und entscheidet die Kommission im Rahmen eines sog. Schutzklauselverfahrens, Art. 81 KIVO)
    In Deutschland wohl die Bundesnetzagentur. Es gibt jedoch sektorspezifische Ausnahmen:
    - Bei KI-Systemen in regulierten Produkten nach Anhang 1 Abschnitt A sind die nationalen Behörde zuständig, die auch für die Überwachung des Produkts zuständig sind (Abs. 3).
    - Für KI-Systeme im Finanzbereich ist die BaFin zuständig (Abs. 6).
    - Bei KI-Systemen für die Strafverfolgung, Wahlen, Grenzkontrollen und die Justizverwaltung sind die Datenschutzaufsichtsbehörden die nationalen Marktüberwachungsbehörden (vgl. Art. 74 Abs. 8 KI-VO).
  - **Notifizierende Behörde (Meldebehörde):** Behörde, die für die Einrichtung und Durchführung des Verfahrens zur Bewertung, Benennung und Notifizierung von Konformitätsbewertungsstellen sowie für deren Überwachung zuständig ist.

    In Deutschland wohl die Bundesakkreditierungsstelle.

- **Nationale Behörden,** die die Verpflichtung zur **Achtung der Grundrechte** in den Mitgliedstaaten in Bezug auf die in Anhang III genannten AI-Systeme mit hohem Risiko durchsetzen.

  In Deutschland unklar.

Maßnahmen auf EU-Ebene

Einrichtung des EU AI Office (KI-Büro, Teil der EU-Kommission)

- **Organisation:** Fünf Referate und zwei Beratern
  - Referat „Exzellenz in KI und Robotik“
  - Referat „Regulierung und Einhaltung“
  - Referat „KI-Sicherheit“
  - Referat „KI-Innovation und Politikkoordinierung“
  - Referat „KI für das Gemeinwohl“
  - Der leitende wissenschaftliche Berater
  - Berater für internationale Angelegenheiten
- Das KI-Büro wird **über 140 Mitarbeiter** beschäftigen, darunter Technologiespezialisten, Verwaltungsassistenten, Rechtsanwälte, Politikspezialisten und Ökonomen.
- **Aufgaben:**
  - Umsetzung der KIVO
    - Beitrag zur **kohärenten Anwendung des KI-Gesetzes** in den Mitgliedstaaten, einschließlich der Einrichtung von Beratungsgremien auf EU-Ebene, Erleichterung der Unterstützung und des Informationsaustauschs
    - Entwicklung von **Tools, Methoden und Benchmarks zur Bewertung von Fähigkeiten** und Reichweite von KI-Modellen für allgemeine Zwecke und **Klassifizierung von Modellen** mit systemischen Risiken
    - Erstellung **modernster Verhaltenskodizes** zur Ausgestaltung von Regeln in Zusammenarbeit mit führenden KI-Entwicklern, der wissenschaftlichen Gemeinschaft und anderen Experten
    - **Untersuchung möglicher Verstöße** gegen Vorschriften, einschließlich **Bewertungen zur Bewertung der Modellfähigkeiten**, und Aufforderung der Anbieter, Korrekturmaßnahmen zu ergreifen
    - Ausarbeitung **von Leitlinien und Leitlinien, Durchführungsrechtsakten und delegierten Rechtsakten** sowie anderer Instrumente zur Unterstützung der wirksamen Umsetzung des KI-Gesetzes und zur Überwachung der Einhaltung der Verordnung
  - Stärkung der Entwicklung und Nutzung vertrauenswürdiger KI
  - Förderung der internationalen Zusammenarbeit
  - Zusammenarbeit mit Institutionen, Experten und Stakeholdern

Aufbau weiterer Strukturen

- Das **KI-Gremium/Ausschuss** soll die Arbeit der nationalen Aufsichtsbehörden auf EU-Ebene koordinieren (vgl. Art. 65, 66).
- Das **Beratungsforum und das wiss. Gremium unabhängiger Sachverständiger** soll sowohl das KI-Büro als auch die nationalen Aufsichtsbehörden bei ihren Aufgaben beraten und unterstützen (Art. 67-69).
- Der **Europäischer Datenschutzbeauftragte** ist die Marktüberwachungsbehörde für Organe, Einrichtungen und sonstige Stellen der EU (Art. 74 Abs. 9).

Ausblick

- **Kurz- und mittelfristig**
  - Entwicklung technischer Standards
  - Aufbau von Aufsichtsstrukturen und Testzentren
  - erste Praxiserfahrungen
- **langfristig**
  - Globale Auswirkungen ("Brussels Effect")
  - Anpassungen basierend auf Praxiserfahrungen
  - Integration neuer KI-Entwicklungen