NEWS

AI Act & Responsible AI: Rechtssicherer und verantwortungsvoller Einsatz von KI

13 August 2026

Fairness und Erklärbarkeit sind nicht mehr nur Themen für technische Experten, sondern gewinnen für die Einhaltung gesetzlicher Vorschriften zunehmend an Bedeutung.

Seit Jahren arbeitet die Technologiebranche weitgehend auf der Grundlage freiwilliger Grundsätze, ethischer Leitlinien und interner Sicherheitsvorkehrungen, doch in Europa signalisiert das im Juli 2024 veröffentlichte EU-KI-Gesetz – die Verordnung (EU) 2024/1689 über künstliche Intelligenz – einen entscheidenden Wandel hin zu einer verbindlichen rechtlichen Aufsicht.

Die Gesetzgebung führt risikobasierte Vorschriften für KI-Systeme ein, wobei die Re[illegible]ngen entsprechend dem potenziellen Schaden, den diese verursachen können, a[illegible] sind. Das Gesetz ermächtigt die Aufsichtsbehörden, bei schwerwiegenden V[illegible]n Geldbußen von bis zu 35 Millionen Euro oder 7 % des weltweiten Jahresumsatzes eines Unternehmens zu verhängen.

Von ethischen Grundsätzen zu rechtlichen Anforderungen

Die sieben Grundsätze für vertrauenswürdige KI, die 2019 von der hochrangigen Expertengruppe für künstliche Intelligenz der Europäischen Kommission formuliert wurden, bilden die Grundlage für die Beurteilung, ob KI-Systeme in der Praxis verantwortungsvoll eingesetzt werden können.

Sie umfassen menschliche Mitwirkung und Aufsicht; technische Robustheit und Sicherheit; Datenschutz und Datenverwaltung; Transparenz; Vielfalt, Nichtdiskriminierung und Fairness; gesellschaftliches und ökologisches Wohlergehen sowie Rechenschaftspflicht.

Obwohl diese Grundsätze einflussreich waren, handelte es sich um unverbindliche ethische Leitlinien ohne Rechtskraft.

Das EU-Gesetz setzt viele ihrer Kernanliegen durch einen risikobasierten Ansatz in rechtlich durchsetzbare Verpflichtungen um und unterteilt KI-Systeme dabei in vier große Kategorien.

Auf der höchsten Stufe befinden sich KI-Systeme, von denen angenommen wird, dass sie ein inakzeptables Risiko darstellen. Diese sind verboten, da sie als unvereinbar mit den Grundrechten eingestuft werden; dazu gehören Social Scoring, manipulative Techniken und die biometrische Identifizierung in Echtzeit zu Strafverfolgungszwecken. Hochrisikosysteme sind nicht verboten, unterliegen jedoch sowohl in der Phase vor der Markteinführung als auch nach der Inbetriebnahme sehr strengen Kontrollen. Zu diesen Systemen gehören KI-Anwendungen für Verbraucherkredite, Lebens- und Krankenversicherungen, die Personalauswahl, die Rechtspflege und demokratische Prozesse. Systeme mit begrenztem Risiko, darunter Chatbots und nicht-systemische Open-Source-GPAI-Modelle, unterliegen hauptsächlich Transparenzpflichten gegenüber den Nutzern. Auf der untersten Stufe befinden sich Systeme, die kein nennenswertes Risiko darstellen, wie beispielsweise viele Spamfilter oder KI-Funktionen in Videospielen.

Verantwortlichkeiten entlang der Wertschöpfungskette

Wer trägt die Verantwortung für KI-Systeme? Der KI-Gesetzentwurf erkennt an, dass die Verantwortung über eine komplexe Wertschöpfungskette verteilt ist – vom Entwickler über den Importeur und Vertreiber bis hin zum Nutzer. Der Betreiber – also derjenige, der ein KI-System im Rahmen seiner beruflichen Tätigkeit unter seiner Verantwortung einsetzt – ist nicht bloß ein passiver Verbraucher. Er ist verpflichtet, die Konformität des Systems mit den geltenden regulatorischen Anforderungen zu überprüfen und die Verteilung der Pflichten gegenüber seinen Lieferanten vertraglich zu regeln. Durch den sogenannten „Brüssel-Effekt“ könnte die KI-Verordnung ihren Einfluss weit über die Europäische Union hinaus ausdehnen. B[illegible] außerhalb der EU müssen die Vorschriften möglicherweise dennoch einhalten, w[illegible] KI-Systeme auf dem europäischen Markt in Verkehr gebracht oder dort vertrieben werden oder wenn sie Ergebnisse liefern, die in der EU genutzt werden. Infolgedessen

könnte die Verordnung zu einem globalen Maßstab werden.

Verpflichtungen für KI-Systeme mit hohem Risiko

KI-Systeme mit hohem Risiko unterliegen den strengsten Schutzmaßnahmen des KI-Gesetzes. Bevor sie in Verkehr gebracht oder in Betrieb genommen werden, müssen die Anbieter nachweisen, dass sie ein Risikomanagementsystem implementiert, eine umfassende technische Dokumentation erstellt, angemessene Maßnahmen zur Datenverwaltung getroffen, wirksame Mechanismen zur menschlichen Aufsicht eingerichtet und ein Konformitätsbewertungsverfahren abgeschlossen haben. Sobald ein System mit hohem Risiko auf dem Markt ist, bestehen die Verpflichtungen des Anbieters fort. Dies erfordert eine kontinuierliche Überwachung nach dem Inverkehrbringen, die Registrierung in der entsprechenden europäischen Datenbank, die CE-Kennzeichnung sowie die Einrichtung eines systematischen Prozesses zur Meldung und zum Management von Vorfällen.

Fairness und Erklärbarkeit werden zu überprüfbaren technischen Anforderungen

Fairness und Erklärbarkeit sind messbare und rechtlich durchsetzbare technische Anforderungen an KI-Systeme. Sie sind von zentraler Bedeutung für den verantwortungsvollen Einsatz von KI, insbesondere dort, wo Systeme Entscheidungen über die Arbeitsplätze, die Bildung, die Kreditvergabe, die Gesundheitsversorgung oder den Zugang zu öffentlichen Dienstleistungen von Menschen beeinflussen.

In diesem Zusammenhang wird Fairness als algorithmische Gerechtigkeit verstanden. Ein scheinbar neutrales System kann dennoch bestimmte Gruppen benachteiligen, wenn es auf voreingenommene Daten trainiert wurde, oder zu ungleichen sozialen Verhältnissen beitragen. Organisationen müssen anhand dokumentierter technischer Nachweise belegen, dass ihre Systeme keine systematischen Ungleichheiten erzeugen, die sich nachteilig auf Gruppen auswirken, die durch geschützte Merkmale wie Geschlecht, Alter, ethnische Herkunft oder Behinderung definiert sind. Darüber hinaus müssen Systeme auf unterschiedliche Ergebnisse bei verschiedenen Teilgruppen innerhalb der relevanten Bevölkerungsgruppen geprüft werden.

Erklärbarkeit bedeutet, sicherzustellen, dass die Ergebnisse eines KI-Systems für Nutzer, betroffene Personen und Aufsichtsbehörden nachvollziehbar sind. Insbesondere müssen KI-Systeme mit hohem Risiko die Rückverfolgbarkeit und Erklärbarkeit der Ergebnisse in einer zugänglichen Form gewährleisten, beispielsweise die Gründe für die Ablehnung eines Kreditantrags oder die Kriterien, die zum Ausschluss eines Bewerbers aus einem Auswahlverfahren geführt haben. Diese Transparenz ermöglicht es den Nutzern, Grenzen und wesentliche Entscheidungsfaktoren zu erkennen und schädliche oder fehlerhafte E[illegible]se anzufechten. Dies ist in risikoreichen Bereichen wie dem Risikomanagement, der B[illegible]bewertung oder der Personalauswahl von entscheidender Bedeutung, wo Entscheidungen erhebliche Konsequenzen für den Einzelnen haben. Diese Anforderung

stellt eine besondere Herausforderung für KI-Modelle dar, die auf undurchsichtigen, sogenannten „Black-Box"-Architekturen basieren. Sie zwingt technische Teams dazu, sich bereits in den frühesten Entwicklungsphasen mit Transparenz und Erklärbarkeit auseinanderzusetzen, anstatt diese Aspekte als Probleme zu behandeln, die erst nach der Einführung gelöst werden müssen.

Erläuterung von Entscheidungen und Messung von Verzerrungen

Um Fairness und Erklärbarkeit zu stärken, hat CRIF spezifische Ansätze und Rahmenwerke entwickelt.

Um eine entscheidende Lücke bei den derzeit auf dem Markt verfügbaren Tools zur Erklärbarkeit zu schließen, hat CRIF einen Algorithmus entwickelt und veröffentlicht, der darauf ausgelegt ist, bestehende Marktstandards wie LIME (Local Interpretable Model-agnostic Explanations) zu verbessern.

Der Algorithmus soll stabilere Erklärungen für komplexe, nichtlineare Modelle liefern, die häufig in Risikomanagementprozessen der Finanzdienstleistungsbranche zum Einsatz kommen. Er ermöglicht die Generierung zuverlässiger, fallspezifischer Erklärungen, selbst in hochdimensionalen Kontexten und in Fällen, in denen die Daten keinen statistischen Standardverteilungen folgen.

Im Hinblick auf Fairness hat CRIF ein strukturiertes analytisches Framework entwickelt und veröffentlicht, der sowohl die Algorithmen als auch die zu deren Training verwendeten Daten abdeckt. Dieses Framework verwendet spezielle Bewertungsskalen, um zu beurteilen, inwieweit sensible Variablen Verzerrungen in den Datensätzen verursachen können, sowie die Tendenz von Algorithmen, diese Verzerrungen im Entscheidungsprozess zu reproduzieren oder zu verstärken. Es kann während der Modellentwicklung angewendet werden, indem es Maßnahmen zur Verbesserung der Datenqualität und -zusammensetzung anleitet, sowie während der Validierung, indem es objektive und vergleichbare Kennzahlen für interne Audits und die Berichterstattung an Aufsichtsbehörden bereitstellt.

Nachhaltiges Wachstum durch Compliance

Für Organisationen, die KI entwickeln oder einsetzen, sollte Compliance nicht nur als regulatorische Belastung betrachtet werden. Sie ist auch eine Chance, die interne Governance zu stärken, die Qualität von KI-Systemen zu verbessern und Vertrauen bei Kunden, Partnern und Behörden aufzubauen. In diesem Sinne sind Fairness, Erklärbarkeit und menschliche Aufsicht keine Hindernisse für Innovation, sondern Voraussetzungen dafür, dass diese nachhaltig und legitim ist.

WEITERE SPANNENDE BERICHTE ENTDECKEN

NEWS

6 August 2026

Zunehmende Insolvenzgefahr im Kfz-Gewerbe: Knapp Jeder zehnte Betrieb betroffen

Mehr lesen →

NEWS

29 Juli 2026

Insolvenzwelle 2026: Welche Branchen jetzt besonders gefährdet sind

Mehr lesen →

NEWS

24 Juli 2026

Die stille Gefahr in der Lieferkette: Warum finanzielle Risiken oft zu spät erkannt werden

Mehr lesen →

Folgen Sie uns

LinkedIn · YouTube

FÜR UNTERNEHMEN

BRANCHEN

FÜR PRIVATPERSONEN

Impressum

Datenschutz

Code Of Conduct

AGB Für Leistungen Im Risiko- Und Chancenmanagement

AGB Für Data And Marketing Solutions

Business Ethics Policy

Victor-Gollancz-Straße 5 | 76137 Karlsruhe

Company with Management System Certified by DNV - ISO 9001, ISO 45001, ISO/IEC 27001, ISO 14001