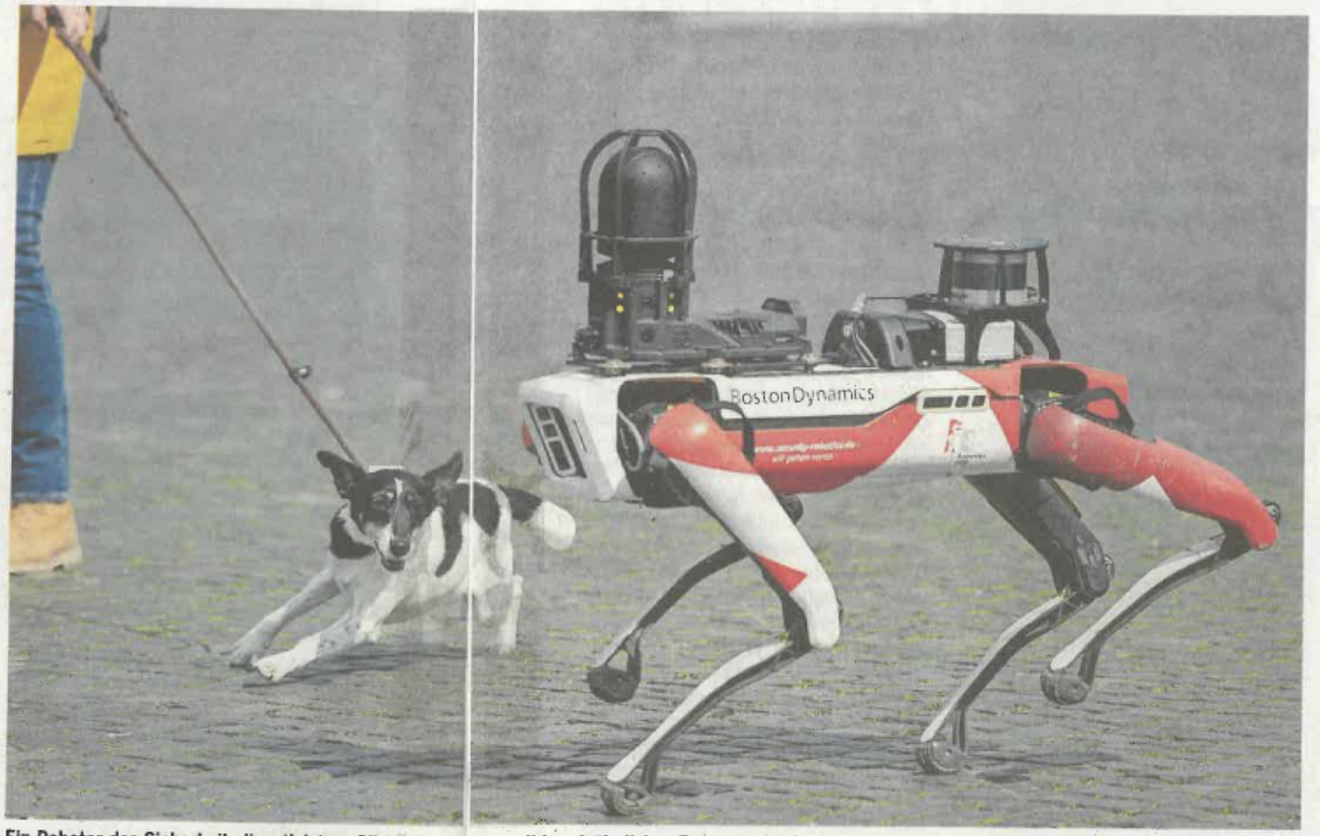
KI: Hohes Risiko nur unter menschlicher Aufsicht

Weltpremiere. EU-Kommission entwarf den weltweit ersten Rechtsrahmen für künstliche Intelligenz (KI): eine Analyse.

VON WOLFGANG ZANKL



Ein Roboter des Sicherheitsdienstleisters Ciborius mit hundeähnlichen Bewegungen, künstlicher Intelligenz und 360-Grad-Kamera. [APA/DPA/M. Reichel]

Wien. Der von der EU-Kommission jüngst vorgeschlagene „Artificial Intelligence Act" erfasst ein breites und durchaus reales Anwendungsspektrum: von Pflege-, Haushalts-, Hunderobotern, Chatbots und Sprachassistenten über Kreditwürdigkeit und Berufsbewerbung bis hin zu heiklen Fragen bei biometrischer Überwachung, Social Scoring oder Triage-Situationen. Durch den bisher nur englisch vorliegenden Entwurf eines „weltweit ersten Rechtsrahmens für künstliche Intelligenz (KI)" soll in Europa „Rechtssicherheit gewährleistet, KI-Verbreitung gefördert sowie Innovation und Investition verstärkt" werden. Ob Letzteres gelingt, könnte bei allem Respekt vor der Regulierungspremiere fraglich sein: Die häufige Verwendung unbestimmter Rechtsbegriffe in einer unmittelbar anwendbaren EU-Verordnung unter gleichzeitiger Androhung exorbitanter Strafen bis 30 Mio. Euro bzw. sechs Prozent des weltweiten Jahresumsatzes erscheint nicht gerade investitionsfördernd.

Kontrollierter Probebetrieb

Schon eher könnten Anreize von „Regulatory Sandboxes" (für kontrollierten Probebetrieb) ausgehen, speziell bei Bevorzugung von Start-ups und kleineren Anbietern. Die dadurch erleichterte Zulassung wirkt der innovationsfeindlichen Überregulierung entgegen, die Europa oft vorgeworfen wird. Damit gehen freilich Risken einher, weil KI bei regulatorischen Lockerungen weniger beherrschbar ist. Dies gilt zwar auch für ähnliche Ansätze im Finanzbereich, doch entstehen bei Fintechs allenfalls Vermögensschäden, während hier oft Personenschäden drohen. Auch die (von der Kommission als „zukunftssicher" eingestufte) KI-Definition wirft Fragen auf. Ihre Aufzählung einschlägiger Technologien (z. B. Machine Learning) bringt die Gefahr der Programmierung von Systemen mit sich, die sich dieser Kasuistik – unter Umständen sogar beabsichtigt – entziehen. Dies könnte im Hinblick auf den risikobasierten Ansatz der Verordnung zum Problem werden.

Dieser Ansatz besteht darin, dass KI umso strenger reguliert wird, je höher das damit verbundene Risiko ist. Im Einzelnen wird zwischen „unannehmbaren", hohen, geringen und minimalen Risken unterschieden. Erstere beziehen sich auf Anwendungen, die Schwächen „spezieller Gruppen" ausnützen, menschliches Verhalten manipulieren (z. B. Spielzeug mit Sprachassistenz, die Kinder zu gefährlichem Verhalten motiviert) oder Behörden eine Bewertung des Verhaltens von Menschen ermöglichen (Social Scoring) sowie biometrische Fernidentifizierungssysteme (z. B. Gesichtserkennung). Deren behördlicher Echtzeiteinsatz im öffentlichen Raum ist grundsätzlich verboten, mit Ausnahmen bei gerichtlicher Genehmigung und zeitlich-geografischer Beschränkung, z. B. um vermisste Personen zu finden oder terroristische Bedrohungen abzuwenden, aber auch zur Strafverfolgung.

Unfallopfer auswählen?

Diesen Überwachungspraktiken wird besonderes Augenmerk gewidmet. Mehr als zwei Drittel des Art. 5, der die „prohibited artificial intelligence practices" regelt, beschäftigen sich damit. Nicht explizit behandelt sind hingegen andere Anwendungen, die wesentlich massiver in Grundrechte eingreifen als biometrische Überwachungen, etwa solche, die bei selbstfahrenden Autos abwägen, welche Personen bei einem unvermeidbaren Unfall mehr oder weniger schutzwürdig sind (sog. Trolley-Dilemma). Meines Erachtens sind solche Programmierungen schon nach allgemeinen (insbesondere grundrechtlichen) Bestimmungen unzulässig – woran auch straf- und zivilrechtliche Notstandsbestimmungen nichts ändern. Denn diese gelten für die Beurteilung konkreter Situationen im Nachhinein, während es hier um abstrakte Programmierungen im Voraus geht, die über Leben und Tod entscheiden.

Insofern muss auch eine andere Anwendung restriktiv verstanden werden, nämlich jene, die Priorisierungen bei medizinischen Notfällen (Triagen) „etabliert" (Art. 6/2). Solche Programme dürften zur ärztlichen Unterstützung – z. B. Berechnung von Überlebenswahrscheinlichkeiten – eingesetzt werden, nicht aber selbst die Priorisierungsentscheidung treffen. Dies ergibt sich auch aus dem Verordnungsentwurf selbst, weil dieser bei Hochrisikoanwendungen, unter denen die Triage aufgezählt ist, „menschliche Aufsicht" über entsprechende Systeme anordnet.

Bonitätsprüfung für Kredite

Weitere KI-Systeme mit hohem Risiko liegen etwa in folgenden Bereichen vor (mit Beispielen): Infrastrukturen mit Lebens- und Gesundheitsgefahren (Verkehr), Schul- oder Berufsausbildung (Prüfungsbewertung), Sicherheitskomponenten (für roboterassistierte Chirurgie). Personalmanagement (Auswertung von Lebensläufen), „wichtige Dienstleistungen" (Bonitätsprüfung für Kredite), Rechtspflege (automatisierte Bescheide oder Rückfallswahrscheinlichkeiten im Strafrecht). Für Anwendungen in diesen Bereichen müssen auch Nicht-EU-Unternehmen strenge Vorgaben erfüllen, bevor sie entsprechende KI-Systeme auf den Markt bringen bzw. anwenden dürfen: darunter Risikomanagement, Verwendung nicht diskriminierender Datensätze, Dokumentation und, wie erwähnt, menschliche Aufsicht.

Bei geringem Risiko bestehen demgegenüber nur Transparenzpflichten: Nutzer müssen informiert werden, wenn sie es mit einer Maschine (z. B. Chatbot) oder mit manipulierten Videos (Deep Fakes) zu tun haben. Auf KI mit minimalem Risiko (z. B. Videospiele, Spamfilter) soll die Verordnung überhaupt unanwendbar sein. Dasselbe gilt für militärische Anwendungen, aber auch für allgemeinen Datenschutz nach der DSGVO und urheber- oder haftungsrechtliche Themen. Wenngleich damit viele Fragen offenbleiben, setzt der Verordnungsentwurf mit dem Fokus auf „vertrauenswürdige KI" ein starkes Signal. Ob es damit gelingen wird, Europa als „globales Zentrum für Exzellenz in der KI" zu positionieren (Binnenmarkt-Kommissar Thierry Breton), bleibt abzuwarten.

Der Autor ist Professor am Institut für Zivilrecht/Universität Wien, Leiter des E-Center und Internationaler Direktor des Artificial Intelligence Law Institute/Tianjin University.