KAN Kommission Arbeitsschutz und Normung

Position zum Vorschlag für eine Verordnung (EU) über künstliche Intelligenz

Stand 2022-03-01

Verein zur Förderung der Arbeitssicherheit in Europa e.V.

Die Kommission Arbeitsschutz und Normung (KAN) ist die Stimme des deutschen Arbeitsschutzes in der Normung. Die KAN setzt sich aus Vertretern der Arbeitgeber, der Arbeitnehmer, des Bundes und der Länder, der gesetzlichen Unfallversicherungsträger und des DIN (Deutsches Institut für Normung e. V.) zusammen. Als neutraler Mittler bündelt sie die öffentlichen Interessen im Arbeitsschutz und bringt sie in Normungs- und Gesetzesvorhaben ein. Sie zeigt Defizite aus Sicht des Arbeitsschutzes auf und macht Verbesserungsvorschläge.

Gefördert durch:

Bundesministerium für Arbeit und Soziales

aufgrund eines Beschlusses des Deutschen Bundestages

Das Projekt „Kommission Arbeitsschutz und Normung" wird finanziell durch das Bundesministerium für Arbeit und Soziales (BMAS) gefördert.

Der Eintrag in das EU-Transparenzregister ist unter der Nummer 90520343621-73 erfolgt.

Ansprechpartner: Corrado Mattiuzzo
Kommission Arbeitsschutz und Normung (KAN)
– Geschäftsstelle –
Alte Heerstraße 111, 53757 Sankt Augustin
Telefon (02241) 231–3466
E-Mail: mattiuzzo@kan.de
Internet: http://www.kan.de

Stand: 1. März 2022

1 Zu den Rechtsgrundlagen

Sachverhalt

Der Rechtsakt ist eine Verordnung auf der Grundlage der Artikel 16 und Artikel 114 des Vertrags über die Arbeitsweise der Europäischen Union (AEUV). Artikel 16 AEUV betrifft das Recht auf Schutz personenbezogener Daten. Artikel 114 AEUV regelt die Verwirklichung des Binnenmarktes.

Abweichend von der bisherigen Tradition im New Legislative Framework verbindet der Vorschlag den Abbau von Handelshemmnissen mit unmittelbar an Anwender gerichteten Verpflichtungen (insbesondere Artikel 29 des Vorschlags) sowie weitreichenden Grundrechtsbelangen.

Position der KAN

Aus Sicht der KAN muss daher vor Verabschiedung der Verordnung geklärt werden (beispielsweise durch die Europäische Kommission):

- ob die Vollharmonisierung der Nutzeranforderungen nur die in Art. 29 geregelten Aspekte oder - wie Erwägungsgrund 58 suggeriert - die Nutzung der KI-Systeme insgesamt betrifft;
- inwieweit deswegen die Rechtsgrundlagen der Verordnung ausreichend sind, da die unmittelbar an Anwender gerichteten Verpflichtungen, sofern dabei Arbeitsschutzanforderungen berührt werden, auch in den Geltungsbereich von Art. 153 AEUV hineinreichen können;
- welche Konsequenzen für die Rechtsunterworfenen mit den oben genannten vertragsrechtlichen Abweichungen von der bisherigen Tradition im New Legislative Framework verbunden sind.

2 Hochrisiko-KI-Systeme (Titel III):

2.1 Zur Klassifizierung von Hochrisiko-KI-Systemen (Kapitel 1, Art. 6 (1)):

Sachverhalt

Laut Art. 6 (1) wären laut Vorschlagskonzept als Hochrisiko-KI-Systeme nur KI-basierte Sicherheitskomponenten für Produkte erfasst, die durch europäische Rechtsakte harmonisiert *und* darin einem Konformitätsbewertungsverfahren durch benannte Stellen unterworfen sind.

Das heißt, dass beispielsweise

- alle KI-basierten Sicherheitskomponenten für Produkte im Anwendungsbereich der Niederspannungs-Richtlinie, sowie
- der gesamte nicht-harmonisierte Produkt- und Anlagenbereich

die Anforderungen aus Titel III nicht zu erfüllen hätten – *unabhängig davon, wie hoch die Risiken wären*, die von diesen Produkten oder Anlagen ausgehen. Ausgenommen davon wären nur KI-Systeme nach Anhang III, Nummer 2. a), die bestimmungsgemäß als Sicherheitskomponenten in der Verwaltung und im Betrieb des Straßenverkehrs sowie in der Wasser-, Gas-, Wärme- und Stromversorgung verwendet werden sollen.

Dies bedeutet, dass KI-basierte Sicherheitskomponenten für diese Produktgruppen *keine KI-spezifischen Konformitätsbewertungsverfahren* zu durchlaufen hätten. Faktoren wie Risikomanagement, Daten und Daten-Governance, Technische Dokumentation, Aufzeichnungspflichten, Bereitstellung von Informationen für die Nutzer, menschliche Aufsicht, Genauigkeit, Robustheit und Cybersicherheit bräuchten faktisch nicht beachtet zu werden.

Position der KAN

Die KAN setzt sich daher dafür ein, in Art. 84 festzulegen, andere Binnenmarktvorschriften wie etwa die Niederspannungsrichtlinie, die bisher ausschließlich Modul A für das Konformitätsbewertungsverfahren vorsehen, dahingehend zu überprüfen, ob sie deswegen hinsichtlich der Anwendung von Künstlicher Intelligenz ergänzt werden müssten.

2.2 Anforderungen an Hochrisiko-KI-Systeme

2.2.1 Zur Diskriminierung durch Datensammlung (Art. 10 (5))

Sachverhalt

Indirekte Diskriminierung *durch* Datensammlung zu bestimmten Personen ist in der EU, bisher ausnahmslos, illegal. *Nur* für KI-basierte *Hochrisiko*-Systeme macht der Vorschlag nun eine Ausnahme: damit soll erreicht werden, dass eine KI-Anwendung Personengruppen mit selteneren Eigenschaften nicht gerade dadurch gefährdet, dass sie mit für diese Personengruppen ungeeigneten Daten trainiert wurde. Folglich soll erlaubt werden, für diese Personengruppen spezifische Kategorien von Daten zu sammeln.

Position der KAN:

Da dieser Bruch mit dem Verbot der indirekten Diskriminierung *nicht für andere* als Hochrisiko-KI-Systeme gilt (siehe dazu auch Kommentar zu Art. 6), setzt die KAN sich dafür ein, in Art. 84 festzulegen, andere Binnenmarktvorschriften wie etwa die Niederspannungsrichtlinie, die bisher ausschließlich Modul A für das Konformitätsbewertungsverfahren vorsehen, dahingehend zu überprüfen, ob sie deswegen hinsichtlich der Anwendung von Künstlicher Intelligenz ergänzt werden müssten.

2.2.2 Zu Transparenz und Bereitstellung von Informationen für die Nutzer (Art. 13)

Sachverhalt

Nicht nur, aber insbesondere für Systeme, die auf Techniken und Konzepten des maschinellen Lernens basieren, gibt es gegenwärtig erhebliche Forschungsaktivitäten, um deren Verhalten nicht nur transparenter zu machen, sondern auch zu *erklären* und damit Risikobeurteilungen unterstützen zu können. Dieser Aspekt fehlt bisher in Titel III, Kapitel 2, des Verordnungsvorschlags.

ISO/IEC 22989 definiert explainability (Erklärbarkeit) als "property of an AI system that important factors influencing the prediction decision can be expressed in a way that humans would understand". Zu Interpretierbarkeit konnte man sich bisher nicht auf eine Definition einigen.

Position der KAN

Um die Aspekte Transparenz und Erklärbarkeit besser voneinander abzugrenzen, sollte Art. 13 (1) daher wie folgt geändert werden:

> *Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass ihr Betrieb hinreichend transparent ist, damit die Nutzer die Ergebnisse des Systems angemessen ~~interpretieren und~~ verwenden können. Hochrisiko-KI-Systeme werden zudem so konzipiert und entwickelt, dass die Ergebnisse des Systems erklärbar sind, sodass sie vom Nutzer interpretiert werden können. Die Transparenz ~~wird~~ und die Erklärbarkeit werden auf eine geeignete Art und in einem angemessenen Maß gewährleistet, damit die Nutzer und Anbieter ihre in Kapitel 3 dieses Titels festgelegten einschlägigen Pflichten erfüllen können.*

2.2.3 Zur Menschlichen Aufsicht (Art. 14)

Sachverhalt

Es wird für komplexere Systeme kaum möglich sein, dass Personen, denen die menschliche Aufsicht über ein Hochrisiko-KI-System übertragen wurde, dessen Fähigkeiten und Grenzen *vollständig* zu verstehen. Auf der anderen Seite ist es wichtig, dass sie überhaupt dazu in die Lage versetzt werden, sich diese Fähigkeiten und Grenzen auch tatsächlich zu vergegenwärtigen.

Position der KAN

Artikel 14 (4) a), sollte daher wie folgt geändert werden:

> *Die in Absatz 3 genannten Maßnahmen müssen den Personen, denen die menschliche Aufsicht übertragen wurde, je nach den Umständen Folgendes ermöglichen:*
>
> *a) sich die Fähigkeiten und Grenzen des Hochrisiko-KI-Systems vergegenwärtigen zu können, ~~vollständig~~ sie zu verstehen und seinen Betrieb ordnungsgemäß zu überwachen, damit Anzeichen von Anomalien, Fehlfunktionen und unerwarteter Leistung so bald wie möglich erkannt und behoben werden können*
>
> *a) ~~fully~~ be aware of and understand the capacities and limitations of the high-risk AI system and be able to duly monitor its operation, so that signs of anomalies, dysfunctions and unexpected performance can be detected and addressed as soon as possible;*

2.2.4 Genauigkeit, Robustheit und Cybersicherheit (Art. 15)

2.2.4.1 Zur Rolle der Umgebungsbedingungen im Zusammenhang mit der Zweckbestimmung von Hochrisiko-KI-Systemen (Art. 15 (1))

Sachverhalt

Bei der Risikobewertung von Hochrisiko-KI-Systemen darf im Zusammenhang mit ihrer Zweckbestimmung nicht vergessen werden, die vorhersehbaren Umgebungsbedingungen mit zu berücksichtigen.

Position der KAN

Art. 15 (1) sollte daher wie folgt ergänzt werden:

> *(1) Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass sie im Hinblick auf ihre Zweckbestimmung - einschließlich der dafür vorhergesehenen Umgebungsbedingungen - ein angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit erreichen und in dieser Hinsicht während ihres gesamten Lebenszyklus beständig funktionieren.*
>
> *1. High-risk AI systems shall be designed and developed in such a way that they achieve, in the light of their intended purpose – including the intended environment - , an appropriate level of accuracy, robustness and cybersecurity, and perform consistently in those respects throughout their lifecycle.*

2.2.4.2 Zu Genauigkeitsgraden und Genauigkeitskennzahlen von Hochrisiko-KI-Systemen (Art. 15 (2))

Sachverhalt

Es ist wichtig, festzulegen, dass Anwendern von Hochrisiko-KI-Systemen in der beigefügten Gebrauchsanleitung nur solche Genauigkeitsgrade und Genauigkeitskennzahlen angegeben werden dürfen, die auf einer verlässlichen Basis ermittelt wurden.

Position der KAN

Art. 15 (2) sollte daher wie folgt ergänzt werden:

> *(2) Die Genauigkeitsgrade und die relevanten Genauigkeitskennzahlen von Hochrisiko-KI-Systemen, die empirisch belegt und praxiserprobt sein müssen, werden in der ihnen beigefügten Gebrauchsanweisung angegeben.*
>
> *2. The levels of accuracy and the relevant accuracy metrics, which shall be empirically valid and proven in practice, of high-risk AI systems shall be declared in the accompanying instructions of use.*

2.2.4.3 Zur technischen Redundanz im Zusammenhang mit der Robustheit von Hochrisiko-KI-Systemen (Art. 15 (3) Satz 2)

Sachverhalt

Um die Widerstandsfähigkeit von Hochrisiko-KI-Systeme gegenüber Fehlern, Störungen oder Unstimmigkeiten durch technische Redundanz zu erreichen, muss sichergestellt sein, dass diese technische Redundanz nicht "dumm", sondern ausreichend diversifiziert konzipiert ist.

Position der KAN

Der 2. Satz von Art. 15 (3) sollte daher wie folgt ergänzt werden:

> *Die Robustheit von Hochrisiko-KI-Systemen kann durch diversifizierte technische Redundanz erreicht werden, was auch Sicherungs- oder Störungssicherheitspläne umfassen kann.*
>
> *The robustness of high-risk AI systems may be achieved through divers technical redundancy solutions, which may include backup or fail-safe plans.*