

Checkliste

Verbotene Praktik Nr. 4
Prädiktive Polizeiarbeit
i.S.d. Art 5 Abs. 1 d.) KI-VO

Praktische Relevanz

● ● ● ● ●

Technische Relevanz

● ● ○ ○ ○

Rechtliche Relevanz

● ● ● ● ●

Orga.Relevanz

● ○ ○ ○ ○

Verortung im Prüfungsschema

I. Bereichsausnahmen

II. Anwendungsbereiche

III. Risikoeinordnung

a. Risikoeinordnung – Verbotene Praktiken gem. Art. 1 Abs. 2 b.) i.V.m. Art. 5 KI-VO

i. Überblick Verbotene KI-Praktiken

ii. Verbotene KI-Praktiken im Einzelnen

1. Unterschwelligen Beeinflussung und Manipulation Art 5 Abs. 1 a.) KI-VO, 29 ErwGr

2. Ausnutzen von Vulnerabilität und Schutzbedürftigkeit Art. 5 Abs. 1 b.) KI-VO

3. Soziale Bewertung oder Soziale Einstufung Art 5 Abs. 1 c.) KI-VO

→ **4. Prädiktive Polizeiarbeit Art. 5 Abs. 1 d.) KI-VO**

Beachte hierzu auch:

Orientierungshilfe Verbotene KI-Praktiken Überblick (Version 1.1.),

Orientierungshilfe Verbotene KI-Praktik Nr. 4 (Version 1.3.), Übersicht Verbotene KI-Praktiken (Version 1.4)

Einleitung

Im Rahmen der Risikobewertung im Strafverfolgungsbereich werden KI-Systeme verwendet, um die Wahrscheinlichkeit zu berechnen oder vorherzusagen, dass eine Person in Zukunft eine Straftat begehen könnte – und zwar ausschließlich auf der Grundlage von Profiling oder der Bewertung persönlicher Merkmale und Eigenschaften.

Das Verbot sieht vor, dass solche Systeme weder entwickelt noch in Verkehr gebracht oder verwendet werden dürfen, wenn sie Entscheidungen über Personen rein automatisiert treffen oder Vorhersagen treffen, ohne eine menschliche Überprüfung zu ermöglichen. Die Verordnung legt zudem fest, dass nur spezifische Ausnahmen gelten, beispielsweise wenn KI-Systeme unterstützend eingesetzt werden, jedoch keine abschließende Entscheidungsgrundlage darstellen.

Im Folgenden wird dargelegt, welche Merkmale und Bedingungen erfüllt sein müssen, damit eine Praktik unter dieses Verbot fällt, und wie die Regelung dazu beiträgt, den Schutz schutzbedürftiger Gruppen zu gewährleisten und eine ethische Nutzung von KI-Systemen in der EU zu fördern.

Warum gibt es die Regelung

Das Verbot der Risikobewertung im Strafverfolgungsbereich ist notwendig, um zentrale Grundrechte wie die Unschuldsvermutung, das Recht auf ein faires Verfahren und den Schutz vor Diskriminierung zu gewährleisten. KI-Systeme arbeiten oft mit historischen Daten, die bestehende Vorurteile oder Ungleichheiten widerspiegeln. Dies kann dazu führen, dass bestimmte Gruppen unverhältnismäßig überwacht oder benachteiligt werden, etwa aufgrund ihres Alters, ihrer sozialen oder wirtschaftlichen Situation.

Das Verbot schützt außerdem vor der Untergrabung der Unschuldsvermutung, da Entscheidungen im Strafverfolgungsbereich auf überprüfbaren Fakten beruhen müssen, nicht auf spekulativen Vorhersagen. Gleichzeitig ist die mangelnde Transparenz vieler KI-Modelle ein Problem, da ihre Entscheidungen für Betroffene und Behörden oft nicht nachvollziehbar sind.

Schließlich reagiert die EU mit diesem Verbot auf internationale Erfahrungen, wie etwa mit dem COMPAS-System in den USA, das für diskriminierende Vorhersagen und intransparente Entscheidungen kritisiert wurde.

Definition

Art. 5 Abs. 1 lit. d.) KI-VO beschreibt die Praktik folgendermaßen:

(...) das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung eines KI Systems zur Durchführung von Risikobewertungen in Bezug auf natürliche Personen, um das Risiko, dass eine natürliche Person eine Straftat begeht, ausschließlich auf der Grundlage des Profiling einer natürlichen Person oder der Bewertung ihrer persönlichen Merkmale und Eigenschaften zu bewerten oder vorherzusagen; dieses Verbot gilt nicht für KI Systeme, die dazu verwendet werden, die durch Menschen durchgeführte Bewertung der Beteiligung einer Person an einer kriminellen Aktivität, die sich bereits auf objektive und überprüfbare Tatsachen stützt, die in unmittelbarem Zusammenhang mit einer kriminellen Aktivität stehen, zu unterstützen;

Checkliste Verbotene KI-Praktik Nr. 4 i.S.d. Art. 5 Abs. 1 d.) KI-VO

Zur besseren Orientierung und für ein formalisiertes Vorgehen wurde die folgende Checkliste entwickelt, um festzustellen, ob eine verbotene KI-Praktik nach Art. 5 Abs. 1 d) KI-VO vorliegt.

(1) Objektiver Tatbestand

(a) KI-Technologie

(i) KI-System (Art. 3 Nr. 1 KI-VO)

(b) Handlungen

(i) Handlung I: Liegt eine der rollenspezifischen Handlungen vor?

1. Inverkehrbringen (Art. 3 Nr. 9 KI-VO)
2. Inbetriebnahme (Art. 3 Nr. 11 KI-VO)
3. Verwendung (Art. 3 Nr. 4 KI-VO)

(ii) Wesentliche Schlechterstellung oder Benachteiligung

1. Schlechterstellung oder Benachteiligung in sozialen Kontexten, die in keinem Zusammenhang mit den ursprünglichen Umständen stehen:

Auslegungshilfe: Eine Schlechterstellung liegt vor, wenn die sozialen Bewertungen oder Einstufungen auf Daten beruhen, die ursprünglich in einem anderen Kontext erhoben wurden, ohne dass ein rechtmäßiger Bezug zur neuen Nutzung besteht.

2. Ungerechtfertigte oder unverhältnismäßige Benachteiligung:

Auslegungshilfe: Eine Benachteiligung ist ungerechtfertigt oder unverhältnismäßig, wenn sie im Vergleich zur Tragweite des sozialen Verhaltens, auf dem sie basiert, übertrieben ist. Beispiel: Ein geringfügiges Fehlverhalten führt zu einem generellen Ausschluss von Sozialleistungen.

(2) Subjektiver Tatbestand

(a) Absicht der sozialen Bewertung

(i) Liegt die Absicht vor, soziale Bewertungen oder Einstufungen mit diskriminierenden oder benachteiligenden Folgen durchzuführen?

Es ist nicht erforderlich, dass die Benachteiligung ausdrücklich beabsichtigt ist. Es genügt, dass das KI-System objektiv darauf ausgelegt ist, soziale Bewertungen vorzunehmen, die zu diskriminierenden oder benachteiligenden Folgen führen können.¹

(b) Ausnahme von der Absichtsvermutung

(ii) Keine Absichtsvermutung bei Faktoren außerhalb der Kontrolle des Anbieters:

1. Liegt eine externe Ursache vor, die nicht vorhersehbar oder beeinflussbar war?

Auslegungshilfe: Externe Ursachen können z. B. Datenfehler oder Änderungen in rechtlichen Anforderungen sein, die vom Anbieter nicht antizipiert oder kontrolliert werden konnten.

1. 2. Wurden Faktoren verwendet, die nicht durch das KI-System beeinflusst wurden?

Auslegungshilfe: Wenn externe Faktoren wie gesellschaftliche Normen maßgeblich zur Schlechterstellung beitragen und nicht auf die Funktionalität des KI-Systems zurückzuführen sind, greift die Ausnahme.

¹ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 66, Abs.192

(3) Schaden

(a) Erheblicher Schaden

(i) Liegt ein erheblicher Schaden vor oder ist ein solcher mit hinreichender Wahrscheinlichkeit zu erwarten?

1. Physische oder psychische Gesundheitsschäden
2. Reine Vermögensschäden
3. Schäden an Datenschutz oder Privatsphäre
4. Immaterielle Schäden

Auslegungshilfe: Schäden wie kulturelle Nachteile, Verlust der Autonomie oder Zeitverluste durch soziale Ausgrenzung.

5. Kumulativer Schaden

Auslegungshilfe: Schäden, die sich durch wiederholte Benachteiligungen summieren, etwa durch dauerhafte Ausgrenzung in sozialen oder beruflichen Kontexten.²

(ii) Beeinträchtigung demokratischer Teilhabe oder gesellschaftlicher Gleichstellung

Auslegungshilfe: Eine Beeinträchtigung liegt vor, wenn soziale Bewertungen systematisch dazu genutzt werden, den Zugang zu demokratischen Prozessen, Sozialleistungen oder wirtschaftlichen Möglichkeiten einzuschränken. Dazu gehören z.B. Risikobewertungen im Hinblick auf eine erneute Straffälligkeit, die systematisch dazu führen können, dass bestimmte Bevölkerungsgruppen ohne Vorliegen konkreter Fakten stärker überwacht oder in Präventionsprogramme aufgenommen werden.³

(4) Ausnahmen

(a) Unterstützung menschlicher Entscheidungen:

(i) Wurde das KI-System ausschließlich zur Unterstützung menschlicher Bewertungen eingesetzt?

Auslegungshilfe: Die Ausnahme gilt nur, wenn die menschliche Entscheidung auf objektiven und überprüfbaren Tatsachen basiert, die direkt mit der strafrechtlich relevanten Handlung verknüpft sind.⁴

(ii) Stützen sich menschliche Entscheidungen auf überprüfbare Tatsachen und nicht auf automatisierte Bewertungen allein?

Auslegungshilfe: Eine bloße Bestätigung oder Übernahme des KI-Ergebnisses ohne kritische Prüfung genügt nicht. Die Entscheidungsverantwortung muss durch entsprechend geschulte Personen getragen werden, die die Ergebnisse im Kontext bewerten können.⁵

² Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 66, Abs. 190

³ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 67, Abs. 196

⁴ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 69, Abs. 203

⁵ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 70, Abs. 204