

Checkliste

Verbotene Praktik Nr. 1

Unterschwellige Beeinflussung & Manipulation

i.S.d. Art 5 Abs. 1 a.) KI-VO, 29 ErwGr

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**

Beachte hierzu auch:

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

Orientierungshilfe Verbotene KI-Praktik Nr. 1 (Version 1.3.), Übersicht Verbotene KI-Praktiken (Version 1.4)

Einleitung

Bevor wir uns diese verbotenen KI-Praktiken der unterschweligen Beeinflussung oder absichtlich manipulative, täuschende Technik gemäß Art. 5 Abs. 1 lit. a KI-VO genauer anschauen, solltet ihr wissen: In der Europäischen Union existieren bereits verschiedene Rechtsakte, die sich mit unterschweligen Beeinflussungen sowie manipulativen oder täuschenden Techniken befassen. So definiert etwa die **Richtlinie über unlautere Geschäftspraktiken** klare Vorgaben, um Verbraucher vor irreführenden oder aggressiven Methoden zu schützen. Der **Digital Services Act (DSA)** wiederum adressiert die Verantwortung von Online-Plattformen und untersagt Gestaltungen, die Nutzer täuschen oder in ihrer Entscheidungsfreiheit beeinträchtigen. Diese bestehenden Regelungen bieten wichtige Anhaltspunkte für die Auslegung und Präzisierung der verbotenen Praktiken. Nutzt diese!

Warum gibt es die Regelung

Die in Art. 5 Abs. 1 lit. a KI-VO vorgesehenen Verbote sollen Individuen und Gesellschaft vor erheblichen Nachteilen schützen, die durch manipulative KI-Systeme entstehen können. Bislang existieren in der Europäischen Union bereits rechtliche Rahmenwerke, welche sich mit den unterschweligen Beeinflussungen befassen. Diese bestehenden Regelungen decken zwar viele Formen von Manipulation ab, sind aber häufig auf Verbraucher oder wirtschaftliche Aspekte fokussiert und erfassen nicht durchgängig alle Arten von Beeinflussung, insbesondere dann, wenn kein unmittelbarer wirtschaftlicher Bezug oder kein klassischer Verbraucherkontext vorliegt. **Die KI-VO greift hier weiter: Sie schließt Lücken, die andere Vorschriften offenlassen**, indem sie auch Praktiken erfasst, die unabhängig von einem Verbraucherstatus, einer wirtschaftlichen Transaktion oder einem spezifischen Medienkontext wirken.

Definition KI-VO

Art 5 Abs. 1 a.) KI-VO, 29 ErwGr definiert unterschwellige Beeinflussung folgendermaßen:

(...) das Inverkehrbringen, die Inbetriebnahme oder die Verwendung eines KI Systems, das Techniken der unterschweligen Beeinflussung außerhalb des Bewusstseins einer Person (...) mit dem Ziel oder der Wirkung einsetzt, das Verhalten einer Person oder einer Gruppe von Personen wesentlich zu verändern, indem ihre Fähigkeit, eine fundierte Entscheidung zu treffen, deutlich beeinträchtigt wird, wodurch sie veranlasst wird, eine Entscheidung zu treffen, die sie andernfalls nicht getroffen hätte, und zwar in einer Weise, die dieser Person, einer anderen Person oder einer Gruppe von Personen erheblichen Schaden zufügt oder mit hinreichender Wahrscheinlichkeit zufügen wird.

Checkliste Verbotene KI-Praktik Nr. 1 i.S.d. Art. 5 Abs. 1 a.) KI-VO, ErwGr 29

Zur besseren Orientierung und für ein formalisiertes Vorgehen wurde die folgende Checkliste entwickelt, um festzustellen, ob eine eine verbotene KI-Praktik nach Art. 5 Abs. 1 a.) KI-VO, ErwGr 29 vorliegt.

Voraussetzungen

(1) Objektiver Tatbestand

(a) KI-Technologie

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

(b) Handlungen

- (i) Handlung I: Liegt mind. einer der rollenspezifischen Handlungen vor?

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

- (ii) Handlung II: Liegt mind. einer der folgenden Praktiken vor?

1. Unterschwellige Beeinflussung: Hierbei handelt es sich um Methoden, bei denen durch akustische, visuelle oder audiovisuelle Reize Einfluss auf das Verhalten oder die Entscheidungen von Personen genommen wird, ohne dass diese Reize bewusst wahrgenommen werden. Die betroffenen Personen erkennen nicht, dass sie beeinflusst werden, da die Reize unterhalb ihrer Wahrnehmungsschwelle bleiben. Beispiele sind visuelle Botschaften, die zu schnell eingeblendet werden, um bewusst wahrgenommen zu werden; leise oder maskierte Audiosignale; unmerkliche taktile Reize; eingebettete Bilder; „Misdirection“ (gezielte Ablenkung) oder „temporale Manipulation“ (Verzerrung der Zeitwahrnehmung).¹

2. Manipulative oder täuschende Techniken: Diese Praktiken zielen darauf ab, die Autonomie und Entscheidungsfreiheit von Menschen so zu untergraben, dass sie trotz möglicher bewusster Wahrnehmung der Einflussnahme nicht in der Lage sind, sich ihr wirksam zu entziehen. Dabei kann es sich beispielsweise um den Einsatz von Gehirn-Computer-Schnittstellen oder virtuellen Realitäten (Siehe ErwGr 29) handeln, um Verhalten und Entscheidungen gezielt zu lenken. Weitere Beispiele sind personalisierte Manipulation mittels individueller Daten, sensorische Manipulation zur Stimmungsveränderung, sowie Täuschung durch falsche Informationen (z. B. Chatbots, die sich als Bekannte ausgeben). Auch unbeabsichtigte Manipulation durch KI-Systeme ist verboten, sofern keine angemessenen Schutzmaßnahmen ergriffen wurden.²

- (c) Handlung III: Liegen die folgenden Eigenschaften vor?

- (i) **Für den Betroffenen nicht erkennbar:** Die Einflussnahme bleibt unterhalb oder außerhalb des bewussten Wahrnehmungsbereichs und ist für die Person nicht eindeutig erkennbar.

- (ii) **Nicht im Interesse des Betroffenen:** Beachte, dass übliche und allgemein akzeptierte Geschäftspraktiken, die im Einklang mit allgemeinen Rechtsvorschriften stehen, nicht als schädliche manipulative KI-Praktiken gelten. Auch rechtmäßige medizinische Behandlungen, etwa psychologische oder physische Rehabilitationsmaßnahmen, fallen nicht unter das Verbot, sofern sie geltendem Recht und medizinischen Standards entsprechen und eine informierte Einwilligung vorliegt (ErwGr. 29).³

- (iii) **Bei dem Betroffenen findet eine wesentliche Verhaltensänderung statt.** Eine wesentliche Verhaltensänderung liegt vor, wenn die Fähigkeit einer Person, eine informierte Entscheidung zu treffen, deutlich beeinträchtigt wird. Die betroffene Person wird so zu einer Entscheidung veranlasst, die sie ohne diese Beeinflussung⁴ nicht getroffen hätte. Es genügt bereits eine „spürbare Beeinträchtigung“; eine wirtschaftliche Entscheidung ist nicht erforderlich.⁵

¹ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 20, Abs. 64, 65

² Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 22, Abs. 69

³ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 28, Abs. 88

⁴ Hierbei kann sich der Anwender sich an Art. 2 lit. e der Richtlinie 2005/29/EG (UGP-RL) orientieren. Dort wird „wesentliche Beeinflussung“ als Anwendung einer Geschäftspraxis verstanden, die die Fähigkeit des Verbrauchers, eine informierte Entscheidung zu treffen, spürbar beeinträchtigt und ihn damit zu einer Entscheidung veranlasst, die er sonst nicht getroffen hätte.

⁵ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 27, Abs. 83

(2) Kausalität: Hinreichende Wahrscheinlichkeit

Auslegungshilfe: Es genügt, wenn die Praxis die objektive Wirkung hat oder haben kann, einen entsprechenden Schaden zu verursachen (ErwGr. 29).

Erheblichkeitsschwelle: Es muss eine gewisse Erheblichkeitsschwelle überschritten werden. Es ist keine geschäftliche Entscheidung erforderlich; jede Art von Entscheidung genügt. Dabei handelt es sich nicht um eine strenge Kausalität, sondern um eine **letztlich fiktive Bewertung des Zusammenhangs („Kausalitätspotenzial“)**. Ein tatsächlicher Kausalitätsnachweis ist nicht erforderlich und wäre aufgrund der potenziell mangelnden Nachvollziehbarkeit von KI-Systemen auch schwierig.⁶

(3) Subjektiver Tatbestand

(a) Absichtsvermutung i.S.d. ErwGr 29

- (i) Die Absicht des Anbieters oder Betreibers muss sich nicht auf die Verursachung des Schadens beziehen.
- (ii) Es genügt eine objektive Wirkung, auch wenn die Manipulationsabsicht schwer nachweisbar ist. Auch Systeme, die unbeabsichtigt manipulativ handeln (z. B. durch fehlerhafte Trainingsdaten), unterliegen dem Verbot, wenn der Anbieter keine angemessenen Schutzmaßnahmen getroffen hat.⁷

(b) Ausnahme: Keine Absichtsvermutung i.S.d. ErwGr 29, wenn vorliegende Faktoren vorliegen:

1. Keine Absichtsvermutung-Faktoren

- a. Faktoren, welche das Verhalten wesentlich beeinflussen
- b. Faktoren, die nicht Teil des KI-Systems sind
- c. Faktoren liegen außerhalb der Kontrolle des Akteurs, weil:
 - i. Faktoren konnten vom Akteur vernünftigerweise nicht vorhergesehen werden, oder
 - ii. Faktoren konnten vom Akteur vernünftigerweise nicht gemindert werden

(4) Schaden

(a) Es muss ein erheblicher Schaden vorliegen oder mit hinreichender Wahrscheinlichkeit eintreten können, der der betroffenen Person, einer anderen Person oder einer Personengruppe zugefügt wird

Auslegungshilfe: „Erheblicher Schaden“ umfasst Schäden jedweder Art, einschließlich:

- Physische und psychische Gesundheitsschäden
- Reine Vermögensschäden
- Schäden in Bezug auf den Schutz personenbezogener Daten oder der Privatsphäre
- Immaterielle Schäden wie kulturelle Schäden, Schäden an Anerkennung und Autonomie, Zeitverluste.
- **Kumulativer Schaden:** Auch Schäden, die sich im Laufe der Zeit anhäufen, können erheblich sein.
- **Schaden bei Dritten oder Gruppen:** Das Verbot erfasst auch Schäden, die anderen Personen oder Personengruppen zugefügt werden. Beispiele: Beeinträchtigung der demokratischen Teilhabe, Manipulation des Wahlverhaltens

Diese Schäden müssen erheblich, aber nicht irreversibel sein.⁸

⁶ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 31, Abs. 94

⁷ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 26, Abs. 82

⁸ Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), S. 30, Abs. 92