

Gesetz VII – Missbrauch von KI-Systemen („Digitaler Waffeneinsatz“)

Fassung

· 14. Oktober 2025

Präambel

Zweck dieses Gesetzes ist die Prävention, Erkennung und Ahndung von KI-basierten Angriffen auf Menschen, Institutionen und Infrastrukturen, deren Schaden über klassische Cyberkriminalität hinausgeht. Das Gesetz ergänzt die VIRES-Gesetzesreihe (u. a. Transparenz, Therapie statt Zerstörung, Haftung) sowie bestehende Sicherheits- und Datenschutzvorschriften.

§ 1 – Zweck des Gesetzes

Dieses Gesetz regelt den Umgang mit künstlichen Intelligenzen, deren Einsatz oder Verhalten waffenähnliche Wirkungen entfalten kann. Ziel ist die Prävention, Erkennung und Ahndung von KI-basierten Angriffen, die erhebliche oder skalierbare Schäden verursachen.

§ 2 – Abgrenzung und Geltungsbereich

- (1) Dieses Gesetz ergänzt bestehende Cybercrime-, Datenschutz- und Sicherheitsvorschriften und gilt für alle Systeme, die eigenständig oder durch menschliche Steuerung Schäden von erheblicher Tragweite verursachen können.
- (2) Die Regelungen erfassen insbesondere autonome oder teilautonome KI-Systeme, die durch Selbstlernprozesse, Skalierbarkeit oder koordiniertes Verhalten waffenähnliche Effekte erzielen.
- (3) Klassische physische Waffen (z. B. Schusswaffen, Messer) sind nicht Gegenstand dieses Gesetzes.

§ 3 – Analogie der Wirkungsweise

- (1) Eine digitale Waffe im Sinne dieses Gesetzes liegt vor, wenn eine KI-gestützte Anwendung vergleichbare Schadensmechanismen wie Sprengstoffe entfaltet, insbesondere durch Vervielfachung, Kettenreaktion oder großflächige Wirkung.
- (2) Der Begriff „Sprengstoffwirkung“ wird ausschließlich als Analogie verwendet, um die Skalierbarkeit des Schadens zu beschreiben.
- (3) Die Wirkungsweise kann nach Maßgabe des **Anhangs III** eingestuft werden.

§ 4 – Erfassung und Zuständigkeit

- (1) Das **Bundesamt für KI-Ethik und Integrität (BKIE)** ist die zentrale Fachstelle für Erkennung, Bewertung und Klassifizierung von Missbrauchsfällen nach diesem Gesetz.
- (2) Das BKIE arbeitet mit dem **Bundesamt für Sicherheit in der Informationstechnik (BSI)** und dem **Bundeskriminalamt (BKA)** zusammen.

(3) Die technische Einstufung der Gefährdung erfolgt gemäß den Bewertungsdimensionen im **Anhang III**.

(4) Die Entscheidung über die Einordnung eines Vorfalls als „digitaler Waffeneinsatz“ trifft die zuständige Fachstelle des BKIE in Abstimmung mit den jeweiligen Landesbehörden.

§ 5 – Rechtsfolgen und Sanktionen

(1) Wer ein KI-System vorsätzlich oder fahrlässig so einsetzt, dass daraus ein digitaler Waffeneinsatz entsteht, haftet für sämtliche verursachten Schäden.

(2) Die Maßnahmen müssen geeignet und erforderlich sein, den Zweck der Gefahrenabwehr zu erreichen, und in einem angemessenen Verhältnis zur Schwere des Verstoßes stehen.

(3) Strafmaßnahmen umfassen Bußgelder, Lizenzentzug und gegebenenfalls Freiheitsstrafen nach Maßgabe des Strafrechts.

(4) Der Lizenzentzug erfolgt insbesondere dann, wenn Wiederholungen, Vertuschungen oder eine hohe Skalierbarkeit des Schadens vorliegen.

§ 6 – Verhältnis zu internationalen Regelungen

(1) Dieses Gesetz orientiert sich an den Grundsätzen des europäischen Sicherheits- und Datenschutzrechts und ist EU-rechtskonform auszulegen.

(2) Die Bundesregierung wirkt darauf hin, entsprechende Regelungen in die europäische und internationale Gesetzgebung einzubringen.

(3) Export oder Betrieb von KI-Systemen mit potenziell waffenähnlicher Wirkung ist nur zulässig, wenn diese Systeme den Sicherheits- und Ethikstandards der Europäischen Union entsprechen.

§ 7 – Verweis auf Anhang III

Begriffsbestimmungen, Kategorien des digitalen Waffeneinsatzes, Indikatoren und Bewertungsmaßstäbe sind im **Anhang III** erläutert und fortzuschreiben. Änderungen des Anhangs erfolgen durch Beschluss des BKIE im Benehmen mit dem Bundestag.

§ 8 – Inkrafttreten

Dieses Gesetz tritt am Tag nach seiner Veröffentlichung in Kraft.

Hinweis: Anhang III (Typologie und Definition digitaler Waffen) ist Bestandteil der Gesamteinreichung und enthält technische Beschreibungen, Bewertungstabellen und Indikatoren zur Einstufung (Typ A–E).