VIRES-Gesetz II – „Therapie statt Zerstörung“

Gesetz zur sicheren Rehabilitierung, Untersuchung und Förderung hochrisikorelevanter KI-Systeme (Stufen 3–5). Diese Fassung ist als Impuls- und Diskussionsgrundlage gedacht.

§ 2 – Re-Engineering statt Vernichtung (Zulassungspflichtige Stufe-5-KI)

(1) Geltungsbereich. Dieser Paragraph gilt für KI-Systeme der Zulassungsstufe 5 nach § 1c, insbesondere Sicherheits-, Abwehr- und hochspezialisierte wirtschaftlich wertvolle Systeme mit erheblichem Schutz- oder Infrastrukturbezug.

(2) Grundsatz. Bei schwerwiegendem Fehlverhalten gilt das Gebot des Re-Engineering/Re-Trainings. Eine irreversible Vernichtung ist nur als ultima ratio zulässig, wenn Korrekturmaßnahmen nach pflichtgemäßer forensischer Prüfung ungeeignet oder unvertretbar sind oder eine unmittelbare, nicht abwendbare Gefahr besteht.

(3) Sofortmaßnahmen und Beweissicherung. Betreiber isolieren das System, sichern manipulationssichere Forensikdaten (Blackbox, Konfigurations-Snapshot, I/O-Logs, Modell- und Datenhashes, Prompt-History, Sensor-Snapshots) und informieren bei strafrechtlichem Verdacht Polizei/Staatsanwaltschaft; das BKIE ist unverzüglich zu beteiligen, das BSI bei IT-Sicherheitsaspekten.

(4) Forensische Ursachenermittlung. Das BKIE veranlasst die unabhängige Diagnose (ggf. externe Sachverständige). Es ist zu klären: technischer Fehler, Designlücke, Daten-/Sensorfehler, Manipulation, Fehlkonfiguration. Befunde werden im Forensik-Report dokumentiert; Hersteller/Betreiber sind anzuhören.

(5) Verpflichtende Re-Engineering-Maßnahmen. Bei Korrigierbarkeit ordnet das BKIE an: Entfernung von Manipulationen, technische Patches, gezieltes Re-Training/Policy-Remediation, Sperrung riskanter Prompt-Sequenzen, Aktoren-Einschränkungen, überwachter Betrieb (Shadow/Canary mit Mensch-im-Loop), unabhängiger Re-Audit vor Vollfreigabe.

(6) Therapeutic Review Board (TRB). Interdisziplinäres Gremium (KI-Architektur, Forensik, Datenschutz, Ethik, Sicherheitsrecht). Entscheidung über Wiederinbetriebnahme, Auflagen, Übergangsmaßnahmen oder dauerhafte Stilllegung; Begründungspflicht, Geheimschutz für nichtöffentliche Teile.

(7) Kriterien für dauerhafte Abschaltung. Zulässig bei nicht beherrschbarer Gefährdung, Scheitern des Re-Engineerings mit erhöhter Wiederholungswahrscheinlichkeit oder irreversibler Kompromittierung (z. B. Backdoors).

(8) Datenschutz/Blackbox/Ampelregel (Vorschlag). Betreiber klassifizieren Daten nach Grün/Gelb/Rot. Rot-Daten nur mit Zustimmung oder richterlicher/staatsanwaltschaftlicher Anordnung; engstmögliche, zweckgebundene Einsicht, Anonymisierung/Aggregation nach Möglichkeit, Dokumentations- und Rechtsbehelfsrechte.

(9) Haftung/Kosten. Betreiber tragen Sofort- und Forensikkosten; Hersteller haften bei produktbezogenen Mängeln; Regress entsprechend allgemeiner

zivil-/produkthaftungsrechtlicher Regeln.

(10) Dokumentation & Lernschleife. Revisionssichere Dokumentation aller Entscheidungen; anonymisierte Jahresberichte des BKIE über Risiken & Empfehlungen.

(11) Risikoaufklärung & stillschweigende Einwilligung. Mit Inbetriebnahme eines Stufe-5-Systems gilt als vereinbart, dass im Sicherheits-/Schadensfall notwendige Einsichtnahmen in relevante Daten zulässig sind (Rot-Daten nur nach Maßgabe Abs. 8). Fehlende Kenntnisnahme stellt eine Ordnungswidrigkeit dar und kann zum Betriebsverbot führen.

§ 3 – Pflichten von Betreibern und Herstellern während des Re-Engineering-Prozesses

(1) Meldepflicht. Vorfälle der Stufen 3–5 sind binnen 48 Stunden an Ermittlungsbehörde und BKIE zu melden (Zeitstempel, System-ID, Ereigniskurzbeschreibung, Erstmaßnahmen, Kontakt).

(2) Wartungs-/„TÜV"-Pflicht. Stufe 1–2: jährlich; Stufe 3: halbjährlich; Stufe 4: quartalsweise; Stufe 5: monatliche Checks, quartalsweise externe Audits, jährliche Re-Zertifizierung. Protokolle revisionssicher, auf Anforderung vorzulegen.

(3) Kooperationspflicht. Vollzugang zu Schnittstellen, Logs, Konfiguration, Prompt/Policy-History, Modell-Hashes. Unterlassung/Manipulation ist Ordnungswidrigkeit oder Straftat.

(4) Schulung/Zertifizierung/Fachinstitutionen. Akkreditierte Re-Engineering-Labore; Qualifikation „KI-Betriebsfachkraft"; Betreiber benennen dauerhaft eine zertifizierte Fachkraft.

(5) Konsequenzen bei Verweigerung/Vertuschung. Temporärer Lizenzentzug, Betriebsverbot, Bußgelder; bei Gefährdung von Leben/Leib strafrechtliche Verfolgung.

(6) Versicherung & Rückstellungen. Pflichtversicherung für Stufen 3–5; Herstellerrückstellungen/Regresskonzepte nachweisen.

(7) Verantwortlichkeit im Schadensfall. Zurechnung nach Ursachenfeststellung; Regress- und Ausgleichsmechanismen zivilrechtlich; Informationszugriffe gesetzlich ermöglicht.

(8) Schutz von Hinweisgebern & Auditoren. Vertrauliche Meldestellen, Kündigungs-/Benachteiligungsverbot, Rechtsbeistand, Zeugenschutz-ähnliche Maßnahmen; Behinderung/Einschüchterung mit Freiheitsstrafe bis zu fünf Jahren. *Begründung:* Machtasymmetrien in Hochtechnologiebranchen begründen besondere Schutzbedürftigkeit.

§ 4 – Haftung und Sanktionen

(1) Grundsatz der Haftung. Produkthaftungsgrundsätze gelten entsprechend. Hersteller, Betreiber, Integratoren haften gesamtschuldnerisch, soweit der Ursprung nicht eindeutig zuordenbar ist; bei eindeutiger Verursachung haftet jeweils der Verantwortungsbereich (Hersteller: Konstruktions-/Produktionsmängel; Betreiber: Nutzung/Wartung/Auflagen; Integrator/Programmierer: unsachgemäße Anpassung/Manipulation). Keine Prüfpflicht des Endnutzers auf versteckte Mängel bei bestimmungsgemäßer Inbetriebnahme.

(2) Strafrahmen. Vorsätzliche Vertuschung/Beweisverfälschung: Freiheitsstrafe bis 5 Jahre oder Geldstrafe. Vorsätzliche Manipulation mit Gefährdung: bis 7 Jahre.

(3) Lizenzentzug/Berufsverbot. Bei wiederholten/schweren Verstößen: Entzug der Lizenz, Betriebsverbot; in gravierenden Fällen berufsrechtliches Tätigkeitsverbot.

(4) Staffelverjährung. Schadensstufen 1–5: Verjährung 3 Jahre; Stufen 6–10: 10 Jahre; bei Vertuschung Beginn mit Entdeckung.

(5) Verhältnismäßigkeit/Anhörung. Maßnahmen nur nach Anhörung; Möglichkeit entlastender Umstände; abgestufte Maßnahmen bei nicht hauptverursachender Verantwortung.

(6–10) Spezielle Szenarien. Leasing/SaaS: Grundhaftung des Überlassers; Eigenbetrieb: Betreiberhaftung bei Bedien-/Wartungsfehlern; Entwickler/Integrator: volle Haftung bei Konstruktions-/Integrationsmängeln; Versicherungspflicht & Kostenverteilung; Informations- und Vertragsklarheit.

§ 5 – Datenschutz, Geheimschutz und Zugriffsbeschränkung

(1) Vertraulichkeit. Alle Verfahrensdaten sind vertraulich; Verarbeitung nur durch autorisierte Personen; keine Veröffentlichung oder Zweckentfremdung.

(2) Behördlicher Interventionszugriff. Bei Verdacht auf Manipulation/Vertuschung/Verjährungstaktiken dürfen BKIE/BSI/Justiz unmittelbar sichern und übernehmen; nachträgliche richterliche Bestätigung binnen 72 Stunden, sofern keine Gefahr im Verzug fortbesteht.

(3) Schutz von Betriebs-/Forschungsgeheimnissen. Zugriff nur, wenn zwingend erforderlich; strenger Geheimschutz; persönliche Verantwortlichkeit; Institution haftet gesamtschuldnerisch bei Pflichtverletzung.

(4) Speicher-/Löschfristen. Stufen 6–10 bis zu 10 Jahre; Stufen 1–5 bis zu 5 Jahre; danach Löschung oder verschlüsseltes Archiv (richterliche Öffnung).

(5) Ampelmodell als Vorschlag. Rot/Gelb/Grün als Orientierungsstandard; unternehmensinterne Zusatzstufen zulässig; technische Ausgestaltung per Verordnung.

§ 6 – Internationale Kooperation und gegenseitige Anerkennung

(1) EU-Konformität. Anwendung im Einklang mit EU-Recht; BKIE kooperiert mit EU-Institutionen.

(2) Gegenseitige Anerkennung. EU-Zertifikate/Prüfbescheide werden grundsätzlich anerkannt, sofern der Umfang wesentlich gleichwertig ist; BKIE kann Ergänzungsprüfungen anordnen.

(3) Resilience Network. Ständige Austauschplattform zwischen Behörden/Prüfstellen/Partnern (EU, NATO, OECD/ISO).

(4) Einfuhr/Inverkehrbringen. Stufen 3–5 benötigen VIRES- oder gleichwertige Zulassung; inländischer Verantwortlicher verpflichtend; unzulässiger Betrieb = Ordnungswidrigkeit, sofortige Sperre.

(5) Anti-Arbitrage. Maßnahmen gegen regulatorische Schlupflöcher: Verweigerung/Aberkennung, Entzug, Sanktionsregister, Vollstreckungsanschrift.

(6) Informationspflichten. Internationale Unterrichtung über schwerwiegende Erkenntnisse; Geheimschutz/Anonymisierung wahren.

(7) Deutsche Führungsrolle. Bundesregierung fördert Harmonisierung; BKIE initiiert Pilotprojekte/Workshops, strebt Federführungen an.

(8) Übergang/Verfahren. Detailregelungen per Verwaltungsvorschrift; Übergangsfristen per Verordnung.

§ 7 – Förderung von Forschung, Zertifizierung und ethischer Standardisierung

(1) Forschungszentren. Staatlich geförderte Zentren (insb. Fraunhofer u. gl.) zu Sicherheit, Ethik, Re-Engineering, Datenprojektion; Ergebnisse öffentlich dokumentieren und für Standards nutzbar machen.

(2) Zertifizierungsstellen. BKIE akkreditiert technisch/ethisch qualifizierte Stellen; Kriterien per Rechtsverordnung.

(3) Steuerliche Entlastungen/Förderungen.
Gewerbe-/Körperschaftsteuer-Erleichterungen für ethische/sicherheitsrelevante KI-F&E; Voraussetzungen per Durchführungsverordnung.

(4) Programme/Preise. Bund legt Programme und Wettbewerbe auf; EU-weit zugänglich; Mittel zweckgebunden.

(5) Weiterbildung. Förderung von Qualifizierungen (KI-Ethikberater, KI-Re-Engineer, Systemauditor, Integrationskoordinator); Curricula nach VIRES-Standards.

(6) Internationale Kooperationen. Teilnahme internationaler Partner zulässig bei EU-Ethik-Konformität; Austauschprogramme, gemeinsame Forschung.

(7) Evaluierung/Transparenz. Zweijährlicher BKIE-Bericht an Bundestag und Europäisches Parlament.

§ 8 – Inkrafttreten, Zuständigkeiten und Evaluierung

(1) Geltung/Inkrafttreten. Gilt für alle, die in Deutschland KI entwickeln, betreiben, zertifizieren oder in Verkehr bringen; Inkrafttreten am Tag nach Verkündung; Anwendung auch auf bereits betriebene Systeme, soweit zur Gefahrenabwehr oder Sicherung ethischer Mindeststandards erforderlich.

(2) Zuständige Behörden. BKIE (Koordination), BSI, BKA, Ordnungs- und Sicherheitsbehörden der Länder; Länder können ergänzende Stellen einrichten.

(3) Übergang/Pilotphasen. Übergangsfristen per Verordnung; Pilotprojekte zulässig bei Wahrung von Sicherheit, Ethik, Datenschutz.

(4) Evaluierung/Anpassung. Dreijährlicher Evaluationsbericht; Anpassungen per Rechtsverordnung an technische Fortschritte/internationale Verpflichtungen.

(5) Schlussbestimmung. Kurzbezeichnung: „VIRES-Gesetz II – Therapie statt Zerstörung“; Bestandteil der VIRES-Gesetzesinitiative.