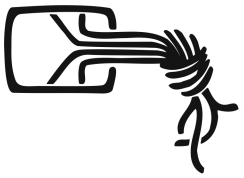

Stellungnahme des Chaos Computer Clubs

an das Bundesverfassungsgericht

zur

automatisierten Datenanalyse

im Hessischen Gesetz über die öffentliche Sicherheit und Ordnung (HSOG)

1 BvR 1557/24

Berlin, 27. März 2026

Constanze Kurz, Stefan Leibfarth, Matthias Marx

Einleitung

Im reformierten Hessischen Gesetz über die öffentliche Sicherheit und Ordnung (HSOG) wird der Einsatz automatisierter Datenanalyse und die Zusammenführung verschiedener Datensammlungen in ganz erheblichem Umfang ermöglicht. Am 13. Dezember 2024 wurden zusätzlich Beschränkungen zum Einsatz von Künstlicher Intelligenz aufgehoben. Als Folge ist die massenhafte Verarbeitung personenbezogener Daten mittels KI-Systemen nun legalisiert, die große Mengen unstrukturierter Daten verarbeiten, aus diesen Daten selbst lernen und neue Verknüpfungen erstellen dürfen.

Aus technischer Sicht wirft das Gesetz grundlegende Fragen zur Bestimmtheit, zu den konkret zulässigen Funktionalitäten der Software sowie zur Kontrolle des Technikeinsatzes, zur Übernahme von Daten aus externen Quellen und zur Begrenzung der Datenverarbeitung auf.

Der Chaos Computer Club (CCC) nimmt Stellung hinsichtlich der Eingriffe in Grundrechte sowie zur Funktionsweise des Softwaresystems. Hierbei werden zentrale technische und strukturelle Risiken betrachtet.

1) Automatisierte Analyse aggregierter Daten

Aus informationstechnischer Perspektive erlaubt der reformierte § 25a HSOG die Erstellung eines dauerhaften, umfänglichen, anwachsenden sowie veränderbaren polizeilichen Datenverarbeitungssystems, das technisch nicht näher spezifiziert ist, aber dessen Inhalt syntaktisch und semantisch erschlossen und analysiert werden darf und soll. Damit sollen vorhandene polizeiliche Datenbestände vernetzt und systematisch zusammengeführt werden.

Nicht nur die Zusammenführung der Datensammlungen ist dauerhaft, auch eine anlasslose automatisierte interne Auswertung wird bei jeder Änderung des Systems permanent vorgenommen. Wie Unbeteiligte aus dieser softwareseitigen Analyse herausgenommen werden könnten, ist nicht spezifiziert.

Zweck des Systems ist es, die aggregierten Daten nach manuellem Befehl zu durchsuchen, um möglichst bisher unbekannte Beziehungen oder Zusammenhänge zwischen Personen, Personengruppierungen, Institutionen, Organisationen, Objekten und Sachen aufzuzeigen. In der Praxis erstellt Software, welche auch KI zur Datenanalyse nutzt, hierzu eine eigene neue Datenbank als Grundlage für die Suchanfragen. Die Art und Weise, mit der die einbezogenen Ursprungsinformationen verknüpft und gewichtet werden, erzeugen hierbei neue Daten. Von einer zweckändernden Verarbeitung der Daten ist dabei auszugehen.

Die Art und der Umfang der neu gespeicherten persönlichen und personenbezogenen Daten und die Verknüpfung zueinander bleibt von außen indes intransparent. Dies hat etwa zur

Folge, dass bei einer manuellen Suchanfrage der Kreis der Betroffenen in der softwareseitigen Verarbeitung nicht mehr sinnvoll eingegrenzt werden kann, so dass von einer Betroffenheit aller in der zusammengeführten Datenbank gespeicherten Personen ausgegangen werden muss.

Praktisch kommt in Hessen eine Softwarelösung des US-Konzerns Palantir mit dem Namen „HessenDATA“ zum Einsatz, die häufig genutzt wird. Nach journalistischen Recherchen wird sie 15.000 Mal im Jahr, also etwa 60 Mal pro Werktag verwendet.[1] Der häufige Gebrauch lässt darauf schließen, dass die Software nicht nur für im § 25a Absatz 2 HSOG vorgesehene Gefahrenlagen oder deren vorbeugende Bekämpfung in schweren Fällen, sondern vielmehr polizeialltäglich verwendet wird.

Nicht ausgeschlossen in der Norm ist der Einsatz eines KI-Systems.[2] Die Gesetzesbegründung bestätigt das explizit: Die „Einsatzmöglichkeiten von hessenDATA [sollen] durch den Einsatz von KI ausgeweitet werden“.[3] Praktisch wirbt auch der derzeitige Vertragspartner Palantir damit, dass seine Software „Gotham“, die „HessenDATA“ zugrundeliegt, mit einer Form von Künstlicher Intelligenz versehen ist („AI-powered“ bzw. „AI-driven“).[4]

Das System soll dabei „verlässliche Tatsachenfeststellungen“ auswerfen und diskriminierende Ergebnisse sicher vermeiden. Um diesen Zielen zu entsprechen, sollen zur Datenauswertung nur Methoden verwendet werden, die zuverlässig arbeiten und dabei weitgehend fehler- und diskriminierungsfrei sind.[5]

2) Methodenoffenheit und unscharfe Normen

Da der Grundrechtseingriff besonders intensiv ist, sollten auch die Anforderungen an die Normenklarheit besonders hoch sein. Aus technischer Sicht sind die Regelungen jedoch so methodenoffen wie nur vorstellbar und wirken zugeschnitten auf den jetzigen Vertragspartner, der Interna des technischen Systems als Geschäftsgeheimnis versteht. Ausreichend klare Absicherungsmechanismen für den Grundrechtsschutz der Betroffenen sind im Gesetz für die Methoden der Datenverknüpfung und -analyse nicht vorgegeben.

---

[1] J. Edelhoff, L. Jeric, et. al.: Wird die Palantir-Software unangemessen genutzt? https://www.tagesschau.de/investigativ/ndr-wdr/palantir-einsatz-deutschland-100.html vom 19. Juni 2025, abgerufen am 27. März 2026.

[2] Es handelt sich um eine Hoch-Risiko-KI nach der KI-Verordnung der EU. Siehe dazu Abschnitt 5: KI-System mit besonders hohen Risiken, S. 9.

[3] Vgl. https://starweb.hessen.de/cache/DRS/21/8/01448.pdf, S. 10, abgerufen am 27. März 2026.

[4] Vgl. https://www.palantir.com/platforms/gotham/, abgerufen am 27. März 2026.

[5] Schon wegen der Fehlerquoten und wegen der Mängel bei verlässlicher Reproduzierbarkeit sowie Nachvollziehbarkeit ist damit der Einsatz heutiger Sprachmodelle (Large Language Models) mit diesen Anforderungen nicht vereinbar.

§ 25a HSOG beschreibt ein System, dessen konkrete Funktionsweise und Eingriffstiefe nicht festgelegt sind, sondern sich erst aus einer technischen Umsetzung ergeben. Es bleibt unklar, welche Daten in welcher Form verarbeitet werden dürfen, welche Analyseverfahren zulässig sind und welche Art von Ergebnissen überhaupt erzeugt werden sollen oder dürfen.

Begriffe wie automatisierte Datenanalyse oder Verknüpfung sind unscharf und können in technischer Hinsicht einfache Abfragen, aber auch sehr komplexe, intransparente und selbstlernende Systeme umfassen. Das bedeutet: Zwei Softwaresysteme, die beide durch den § 25a HSOG erlaubt wären, können sich technisch fundamental unterscheiden, ohne dass sich diese Unterschiede aus dem Gesetz ergeben.

Welche zuvor unverbundenen Daten kombiniert werden, wie stark sie gewichtet sind, welche Verknüpfungen als relevant gelten, welche Schwellenwerte für Treffer gelten sollen oder ob es probabilistische Systemkomponenten gibt, wird nicht vorgegeben, sondern erst durch Konfiguration, Training, Updates und Anbieterentscheidung bestimmt. Damit hängt der tatsächliche Grundrechtseingriff nicht vom Gesetz ab, sondern von den technischen Details der Umsetzung, die nicht transparent sind, vom Anbieter abhängen und zudem einer ständigen Veränderung unterliegen.

Um es ganz deutlich zu formulieren: Nach den Vorgaben des § 25a HSOG darf irgendein technisches System auf nicht spezifizierte Weise die einfließenden heterogenen Daten in einer beliebigen Art manipulieren, um irgendetwas herauszufinden, solange das erzielte Ergebnis als zuverlässig und nicht diskriminierend interpretiert wird. Wesentliche Elemente des Systems sind also gar nicht oder unzureichend durch den Gesetzgeber bestimmt.

Ergänzend bleiben auch das Rollen- und Rechtekonzept sowie die Kategorisierung und Kennzeichnung personenbezogener Daten aus technischer Sicht unklar und stellen keine wirksame Begrenzung der technischen Ausgestaltung dar. Im Gesetz werden nur allgemeine Prinzipien formuliert, nach denen

1. mehr Berechtigte Zugriff auf weniger und wenige Berechtigte Zugriff auf mehr der in der Analyseplattform zusammengeführten Daten haben dürfen und

2. eine automatisierte Datenanalyse umso komplexer sein darf, je gewichtiger der Veranlassungszusammenhang ist, und dass sie umso einfacher sein soll, je weniger gewichtig der Veranlassungszusammenhang ist.

Das damit intendierte Schutzkonzept kann in der Praxis leicht unterlaufen und weitreichende Zugriffe aus praktischen Gründen gebündelt werden.

Auch genügt eine bloße Reglementierung des Zugangs nicht, um eine wirksame Kontrolle der Datenverarbeitung und der Verarbeitungsmethoden sicherzustellen. Es sind lediglich Zugriffe auf die Anwendung des Systems zu protokollieren. Nicht erfasst wird die eigentliche Datenverarbeitung innerhalb des Systems, die auch durch selbstlernende Komponenten erfolgen kann.

Daher sollte eine weitergehende Protokollierung erfolgen, mit der erst nachvollzogen werden kann, welche personenbezogenen Daten in welchen Analysen verwendet, verknüpft oder für Trainingszwecke herangezogen wurden. Ohne diese Protokollierung können weder die Polizei selbst noch Kontrollinstanzen wie der Hessische Beauftragte für Datenschutz und Informationsfreiheit oder Gerichte im Nachhinein nachvollziehen, in welchem Umfang und wie personenbezogene Daten verarbeitet wurden.

§ 25a HSOG erlaubt also ein weitgehend unbegrenztes Spektrum an technischen Ausgestaltungen. Die Eingriffstiefe der Datenverarbeitung und Komplexität der automatisierten Datenanalyse ergeben sich erst aus der technischen Ausgestaltung eines Softwaresystems, welches sich über die Zeit, gesteuert durch den Anbieter noch verändern kann. Da das Gesetz keine Begrenzung der Eingriffsintensität vorgibt, kann nur eine maximal negative Abschätzung des Grundrechtseingriffs konstatiert werden.

3) Prüfbarkeit und Auditierbarkeit

An dem System dürfen neben der Wartung auch „wesentliche Änderungen“ vorgenommen werden, wenn die Behördenleitung oder ein von ihr Beauftragter dies anordnen. Damit ist bis auf die Zuverlässigkeit und Diskriminierungsfreiheit keine funktionelle Eigenschaft dieses Systems im Gesetz festgelegt. Auch ganz neue Analyseverfahren können implementiert und zusätzliche Datenquellen integriert sowie Auswertungslogiken verändert werden. Was eine „wesentliche“ Eigenschaft ist, kann bei Softwaresystemen durch Austausch weniger Zeilen Quellcode umfunktioniert werden.

Wesentliche Änderungen sind auch Updates. Solche Updates können auch Änderungen oder Aktualisierungen sein, die Funktionalitäten, Analyseverfahren, Datenverknüpfungen oder Gewichtungen grundlegend verändern und erweitern sowie Datenquellen hinzufügen. Der Eingriffsgehalt des Systems in Grundrechte kann sich erheblich verschieben, insbesondere wenn neue Datenarten oder Analyseverfahren integriert werden. Palantir als Vertragspartner der hessischen Polizei bewirbt explizit für die Durchführung von Updates ein „Continuous-Delivery-System“, das den Wartungsaufwand nach eigenen Angaben verringern soll. Der Vertragspartner entsendet im polizeilichen Regelbetrieb eigenes Personal zur Polizei, um das System zu pflegen, zu warten und Updates vorzunehmen.

Die effektive Prüfbarkeit und Kontrolle solcher Änderungen ist insbesondere bei proprietären Systemen stark eingeschränkt und kann nur punktuell erfolgen. Updates bleiben für Kontrollinstanzen wie den Hessischen Beauftragten für Datenschutz und Informationsfreiheit, der vor einer wesentlichen Änderung anzuhören ist, weitgehend intransparent. Zudem ist fraglich, wer mit welchen Kriterien für jedes Update einschätzt, ob es sich dabei um eine „wesentliche Änderung“ handelt oder nicht.

Selbst wenn Dokumentation und Quellcode vorlägen, ist nicht ohne Weiteres prüfbar, wie sich konkrete Änderungen auf ein System im Einzelfall auswirken. Bei manchen KI-Systemen

können Entscheidungsprozesse generell – auch für die polizeilichen Nutzer und erst recht für die Betroffenen – intransparent bleiben. Wie es zu einem bestimmten Resultat gekommen ist, kann nicht bei allen KI-Systemen nachvollzogen werden. Insofern ist auch die Prüfung von vertraglich zugesicherten Eigenschaften der Software erschwert.

Externe Kontrollen stoßen durch fehlenden Zugang zu Quellcode oder durch die Berufung auf Geschäftsgeheimnisse wie im konkreten Fall von „HessenDATA“ an praktische und rechtliche Grenzen. Die 2023 vom Fraunhofer-Institut SIT in Bayern durchgeführte Sichtung der dortigen Software von Palantir zeigt die Grenzen einer solchen Prüfung auf: Trotz Kosten von 430.000 EUR[6] konnte nur eine punktuelle und gerade keine fortlaufende Bewertung der Systemänderungen im Echtbetrieb stattfinden sowie keine Einschätzung über Fehlmeldungsquoten des Softwaresystems erfolgen. Das vor Jahren geprüfte System unterscheidet sich von den heute in Bayern und Hessen eingesetzten Varianten, mit hoher Wahrscheinlichkeit sogar signifikant.

Die Tatsache, dass laut Gesetz „wesentliche Änderungen“ vorgenommen werden dürfen, wirkt sich auch auf die Revisionsfähigkeit aus. Denn Updates der Software oder ein „Weiter-Trainieren“ einer KI können Gewichtungen und Verknüpfungen verändern, so dass frühere Ergebnisse des Systems dann nicht mehr nachvollziehbar sind. Mehrere Updates zusammen können stark vom ursprünglichen System abweichen. Es muss neben den vorgeschriebenen Protokollen der Zugriffe auf das System mindestens dokumentiert werden, welche Version des Systems in welcher Konfiguration zu einem bestimmten Zeitpunkt im Einsatz war.

Es sind keine verbindlichen Verfahren vorgesehen, die sicherstellen würden, dass Updates als gewolltes oder ungewolltes Ergebnis keine diskriminierenden, fehlerhaften oder verzerrten Resultate der Datenanalyse liefern. Letztlich widerspricht sich die Erlaubnis zu „wesentlichen Änderungen“ im Gesetz mit der Vorgabe, dass nur weitgehend fehler- und diskriminierungsfreie Methoden verwendet werden dürfen.

4) Externe Quellen und Daten

Eine „direkte Anbindung“ an Internetdienste darf zwar nicht vorgenommen, jedoch dürfen gespeicherte externe Daten aus dem Internet in das System übernommen werden. Praktisch bedeutet das nur, dass eine Übernahme neuer Daten von externen Quellen sowie etwa das ständige Einspeisen aktualisierter Polizeidaten nicht über Internetdienste, sondern über die internen Netze erfolgen darf. Auch wenn die Software durch die berechtigten Bedarfsträger technisch genutzt wird, darf das demnach nicht über das Internet erfolgen.

Allerdings ist eine Begrenzung von Art und Umfang von Daten aus dem Internet im Gesetz nicht vorgesehen. Damit entsteht praktisch ein Freibrief für Art und Menge von Daten aus

---

[6] Drs 19/7043 des Bayerischen Landtags, S. 11, https://www.bayern.landtag.de/www/ElanTextAblage_WP19/Drucksachen/Schriftliche%20Anfragen/19_0007043.pdf vom 7. Juni 2025, abgerufen am 27. März 2026.

Internetquellen, die zur Analyse auch als Massendaten importiert werden können. Beispielsweise verfügen Social-Media-Plattformen über weitreichende Export-Möglichkeiten, welche etwa nach Standort, Schlüsselwörtern oder Personengruppen vorgefiltert werden können.

Palantir als Vertragspartner der hessischen Polizei wirbt damit, dass in der Software Daten aus dem öffentlichen und dem privaten Sektor kombiniert werden können. In „HessenDATA" ist das nicht in Echtzeit vorgesehen, weil die Verbindung zum Internet nicht erlaubt ist. Doch angesichts der bereits angekündigten bundespolitischen Pläne zur erweiterten Auswertung von biometrischen Daten mit einem automatisierten Abgleich von im Internet verfügbaren Fotos[7] durch die Polizeien des Bundes könnte ein solches Vorgehen in Hessen und anderen Nutzerländern der automatisierten Datenanalyse auf Interesse stoßen. Echtzeit-Abgleiche mit biometrischen Daten aus dem Internet und der Import von Massendaten aus dem Internet sollten daher präventiv ausgeschlossen werden.

Technisch ist neben Abgleichen von biometrischen Daten in Bildern oder Videos auch eine Vielzahl anderer Abgleiche von Mustern möglich, die Menschen kategorisieren oder auch identifizieren können: Zu denken ist beispielsweise an ethnische Merkmale, äußerliche Eigenschaften oder Auffälligkeiten etwa bei Haaren, sichtbaren Tätowierungen oder Narben.

Gar nicht in das hessische System einfließen sollen nur Daten aus Wohnraumüberwachungen und von Staatstrojanereinsätzen, sofern diese die Form der „Online-Durchsuchung" haben, die Schadsoftware also auf das gesamte gehackte informationstechnische System zugreifen konnte. Das sind die einzigen Datenquellen, die aufgrund ihres ursprünglichen Anlasses ausgesondert werden.

Ansonsten ist in § 25a HSOG eine Liste sehr heterogener polizeilicher Datenarten gegeben, die auf der Suche nach Übereinstimmungen und Kreuztreffern ausgewertet werden können: „Vorgangsdaten, Falldaten, Daten aus den polizeilichen Auskunftssystemen, Verkehrsdaten, Telekommunikationsdaten, Daten aus Asservaten und Daten aus dem polizeilichen Informationsaustausch". Jede darin gespeicherte Information in diesen umfangreichen Sammlungen ist für die landesweite hessische Analyse freigegeben. Damit wird ein Großteil der überhaupt polizeilich verfügbaren Datensammlungen dauerhaft zusammengeführt, zentralisiert und fortlaufend ausgewertet und in einem zweiten Schritt dann manipuliert und abgefragt. Es gibt keine Begrenzung der darin gespeicherten Datenarten, Speicherfristen oder Volumina, nur für Verkehrsdaten ist eine „Speicherfrist von regelmäßig zwei Jahren" vorgesehen.

Innerhalb dieser Datenliste können also alle Datenarten jeder Menge vorkommen, die digital speicherbar sind, auch Audio- und Videodaten, biometrische Daten einschließlich genetischer Informationen, höchstpersönliche Daten aus dem Kernbereich privater Lebensgestaltung, etwa aus Asservaten-Daten. Wegen des Einbezugs des polizeilichen Informationsaustauschs

---

[7] Vgl. Online-Bildabgleich und automatisierte Datenanalyse: Bundesjustizministerium schlägt Rechtsgrundlage für neue digitale Ermittlungsmaßnahmen vor, https://www.bmjv.de/SharedDocs/Pressemitteilungen/DE/2026/0312_Digitale_Ermittlungsmassnahmen.html vom 12. März 2025, abgerufen am 27. März 2026.

können auch geheimdienstliche Daten[8] oder Daten aus dem Bundesamt für Migration und Flüchtlinge, die an die Polizei weitergegeben wurden, in das System gelangen.

Nach journalistischen Recherchen[9] reduzierte sich diese Datenfülle in Hessen nach dem Urteil des Bundesverfassungsgerichts zur automatisierten Datenanalyse 2023 auch nicht, sondern wird weiterhin in das KI-System eingespeist und ausgewertet. Demnach wurde lediglich die Sichtbarkeit für die 2.000 Polizeibeamten eingeschränkt (Treffer werden „verborgen"), nicht aber die Verarbeitung innerhalb des Systems verändert.

Zu den in die Analyse einfließenden heterogenen Polizeidaten gehören seit der Novelle des hessischen Psychisch-Kranken-Hilfe-Gesetzes (PsychKHG) auch Patientendaten. Hessische psychiatrische Krankenhäuser, Kliniken und Einrichtungen sind nach § 28 Absatz 4 PsychKHG verpflichtet, Daten von psychisch Kranken an die Polizei weiterzugeben. War eine Person wegen einer Fremdgefährdung untergebracht, sind Polizeibehörden zur Gefahrenabwehr von einer bevorstehenden Entlassung unverzüglich zu unterrichten.

Dabei sind auch alle für eine Gefährdungseinschätzung notwendigen Informationen an die Polizeibehörde zu übermitteln. Diese Meldepflicht über die Daten der psychisch Kranken gilt schon dann, wenn nicht näher bestimmte Anhaltspunkte dafür bestehen könnten, dass sie möglicherweise in absehbarer Zeit in einer nicht näher bestimmten Form die Gesundheit oder andere bedeutende Rechtsgüter Dritter gefährden könnten. Das Einfließen solcher sensiblen Informationen über Kranke in die automatisierte Datenanalyse kann zur Stigmatisierung, Vorverurteilung und zu ungerechtfertigten Benachteiligungen der Patientinnen und Patienten führen, die nach dem PsychKHG an die Polizei gemeldet wurden. Schon präventiv sollten alle Gesundheitsdaten von der Datenanalyse gesetzlich ausgeschlossen werden.

5) KI-System mit besonders hohen Risiken

Kommen in dem hessischen Datenauswertungssystem auch Anwendungen der Künstlichen Intelligenz zum Einsatz, ist schon wegen der sehr hohen Zahl betroffener Grundrechtsträger von einer Einstufung als hochriskant nach der 2024 verabschiedeten KI-Verordnung der EU auszugehen.[10] Die Verordnung sieht eine Risikoeinstufung vor: Je höher das Risiko ist, das etwa mit der Nutzung eines KI-Systems für die Grundrechte von Menschen einhergeht, desto intensiver und detailreicher soll diese Nutzung reguliert sein.

---

[8] Für die automatisierte Datenanalyse bei Geheimdiensten existieren keine gesetzlichen Vorgaben. Da sie keine Zwangsmaßnahmen einleiten können, übermitteln Geheimdienste bei konkreten Gefahren personenbezogene Daten zur Gefahrenabwehr auch an Polizeibehörden.

[9] J. Edelhoff, L. Jeric, et. al.: Wird die Palantir-Software unangemessen genutzt? https://www.tagesschau.de/investigativ/ndr-wdr/palantir-einsatz-deutschland-100.html vom 19. Juni 2025, abgerufen am 27. März 2026.

[10] Vgl. KI-Verordnung, Anhang III, Nummer 19 x.

Die Auswertung personenbezogener Daten aus polizeilichen Datensammlungen ist zweifellos ein sensibler und risikoreicher Anwendungsbereich. Da mit „HessenDATA“ bereits jahrelang ein solches System betrieben wird, das in der KI-Verordnung als hoch risikoreich aufgeführt ist, müsste die Polizei als Anbieter alternativ eine Bewertung dokumentieren, warum dieses KI-System nicht hochriskant sein sollte.

Die Beschreibung von Palantir für das System „Gotham“, das „HessenDATA“ zugrundeliegt, lautet: „AI-enabled operating system that facilitates and accelerates human decision-making in all functions and task areas“ (KI-gestütztes Betriebssystem, das die Entscheidungsfindung von Menschen in allen Funktionen und Aufgabenbereichen erleichtert und beschleunigt). Soweit öffentlich bekannt, ist der Quellcode dieses „Betriebssystems“ in Hessen nicht geprüft oder auch nur gesichtet worden, obwohl große Datenmengen von vergangenen Straftaten, von zahlreichen Unverdächtigen und auch sensibles polizeiliches Organisationswissen ausgewertet werden.[11]

Das besonders hohe Risiko beim hessischen KI-System der Polizei entsteht dadurch, dass ein hoher Schaden eintreten kann, wenn Personen in der Folge der automatisierten Auswertung fälschlich in Verbindung zu möglichen Straftaten gerückt werden und dadurch polizeiliche Maßnahmen erdulden müssen. Auch fehlerhafte Ausgaben oder Missbrauch können massive Schäden anrichten. Entsprechend wäre zu prüfen, ob und welche Teile der automatisierten Datenanalyse zu den verbotenen KI-Praktiken nach der KI-Verordnung zählen. Dazu würden etwa manuelle Anfragen ins Blaue hinein zählen.

Bei hochriskanten KI-Systemen sind insbesondere bei der Genauigkeit, Robustheit sowie IT-Sicherheit besondere Anforderungen vorgeschrieben, die bei Entwicklung und Betrieb gelten, aber auch bei allen wesentlichen Veränderungen zu beachten sind. Daraus ergibt sich auch unmittelbar die Problematik der Verantwortung: Denn ob die Polizeivertreter, die Behördenleitung oder der Software-Anbieter etwa nach einem Update die Verantwortung für Mängel bei Genauigkeit oder IT-Sicherheit trägt, bleibt bisher offen.

6) Vermeidbare Abhängigkeiten

Da die Polizei in Hessen mit einem außereuropäischen Konzern kooperiert und mit „HessenDATA“ eine modifizierte Variante des Palantir-„Gotham“-Systems einsetzt, ist sie einerseits Anbieter des hochriskanten KI-Systems, andererseits aber auch in hohem Maße abhängig von den Dienstleistungen des Konzerns, um gewünschte Veränderungen, Fehlerkorrekturen, Tests, Aktualisierungen oder IT-Sicherheitsmaßnahmen umzusetzen. Doch auch bei Vertragsschluss mit einem europäischen Anbieter würde sich die Polizei vergleichbare Probleme ins Haus holen. Um eine Verantwortungsdiffusion zu vermeiden, sollte aus dem Gesetz deutlich

[11] Anders in Bayern, wo das Fraunhofer-Institut SIT vor mehreren Jahren einmalig eine Version der Software einsehen durfte. Der Bericht des SIT ist allerdings nicht öffentlich verfügbar. Die bayerischen und hessischen Versionen von „Gotham“ weichen aber ohnehin voneinander ab.

werden, wer als Anbieter des KI-Systems die Pflichten nach der KI-Verordnung zu erfüllen hat.

Für die Erfüllung dieser Pflichten wäre eine Softwarelösung anzustreben, die bisher gesetzlich nicht vorgeschrieben ist: ein System, das die Polizei in allen Aspekten selbst beherrschen und prüfen und somit auch selbst anpassen oder erweitern kann. Denn wer die rechtliche und praktische Hoheit über ein System zur Datenanalyse hat, stellt eine langfristige strukturelle Entscheidung darüber dar, wer die operativen Möglichkeiten und polizeiinternen Fähigkeiten (auch in Zukunft) kontrolliert. Mit einer solchen Softwarelösung würden Autonomieverlust und Abhängigkeiten von Konzernen vermieden, deren Prioritäten sich nach den Interessen von Venture-Kapitalgebern oder neuen Eigentümern, aber auch von politischen Akteuren in deren Heimatländern zu Ungunsten von hiesigen polizeilichen Interessen ändern können.

Diese Abhängigkeiten werden durch Interoperabilitätsprobleme verstärkt. Proprietäre Datenformate und -schnittstellen wie im Fall von „HessenDATA“ erschweren eine Migration auf oder die Integration mit anderen Systemen. Dies betrifft nicht nur die Rohdaten, sondern auch daraus generierte Verknüpfungen und trainierte KI-Modelle.[12] In der Folge entsteht eine technische Abhängigkeit, die zwar den Interessen des kommerziellen Vertragspartners dient, jedoch die Austauschbarkeit des Systems einschränkt.

Daraus ergibt sich ein sog. Vendor Lock-in, also eine starke Abhängigkeit von dem jeweiligen Anbieter, der die Weiterentwicklung des Softwaresystems maßgeblich bestimmen, aber auch blockieren kann. Die Polizeibehörden bleiben damit in zentralen Fragen der Funktionsweise und Fortentwicklung eines stark in Grundrechte vieler Menschen eingreifenden Systems auf ein privates Unternehmen festgelegt.

Dieser Vendor Lock-in wirkt sich direkt auf die Möglichkeit der Vertragsbeendigung aus. Selbst bei erheblichen rechtlichen, ethischen oder technischen Bedenken ist ein mittelfristiger Ausstieg praktisch kaum realisierbar. Grund hierfür sind starke funktionale Einschränkungen in der Übergangszeit oder hohe Kosten für eine Migration auf ein neues System. Beides steht einem Ausstieg entgegen und führt in der Praxis zum Verharren auf der bisherigen Lösung.

Dies kann dazu führen, dass trotz bekannter Mängel oder bestehenden Rechtswidrigkeiten ein solches System weiterbetrieben wird. Damit verschiebt sich die Hoheit über das System vom Gesetzgeber hin zu vertraglichen Rahmenbedingungen, die durch den Anbieter geprägt werden. Die tatsächliche Kontrolle – auch ganz praktisch – übt letztlich aber der Vertragspartner aus, dessen Prioritäten nur bedingt beeinflussbar sind.[13] Selbst bei offenkundigem

---

[12] KI ist gewissermaßen zu einem Synonym für Sprachmodelle (LLMs) geworden. Es gibt aber KI-Systeme, die auf anderen Konzepten basieren. Die beschriebenen Abhängigkeiten bleiben jedoch.

[13] Diese Abhängigkeiten führten in der Schweiz zu einer Risikobewertung, die nach einem Gutachten der Schweizer Armee eine Zusammenarbeit mit Palantir als zu riskant einschätzte, vgl. Adrienne Fichter, Marguerite Meyer, et. al.: Wie hartnäckig Palantir die Schweiz umwarb, https://www.republik.ch/2025/12/08/wie-hartnaeckig-palantir-die-schweiz-umwarb vom 8. Dezember 2025, abgerufen am 27. März 2026.

Vertragsbruch gibt es wegen des Vendor Lock-ins für die Polizei kaum praktische Handlungsoptionen.[14]

Warum der Gesetzgeber keine unabhängige Softwarelösung vorgeschrieben hat, obgleich eine derart große Fülle personenbezogener, aber auch sensibler polizeiinterner Daten darin verarbeitet wird und mithin Insider-Informationen über fast alle Arten von Polizeikontakten sowie polizeiinternes Organisationswissen aus der Hand gegeben werden, ist nur erklärbar durch die jahrelange Gewöhnung an „HessenDATA“ und mit dem Unwillen, eine vor Jahren getroffene Entscheidung zu revidieren. Dass der Chief Digital Officer in der hessischen Polizei von einer „großen Fähigkeitslücke“[15] spricht, die ohne die Palantir-Hilfe drohe, bestätigt die Gewöhnung an eine permanente starke Abhängigkeit.

7) KI-Training

Die Nutzung und Weiterentwicklung der hessischen datenbasierten Auswertungssysteme sind – soweit öffentlich bekannt – nicht durch Evaluationen oder durch eine wissenschaftliche Forschung begleitet worden. Eine unabhängige Evaluation polizeilich eingesetzter Software von Privatunternehmen ist gesetzlich auch nicht vorgesehen. Schon angesichts der durchschnittlich 60 Nutzungen pro Arbeitstag in Hessen und der massiven Eingriffstiefe sollte der Gesetzgeber regelmäßige Berichtspflichten und eine systematische Evaluation vorschreiben. Anekdoten über Treffer oder die Effizienz der Software ersetzen keine Evaluation, die auch schon das KI-Training begleiten sollte.

Werden polizeiinterne sensible Daten neben der eigentlichen Nutzung auch für das Training der Software verwendet, gehen sie über Operationen des maschinellen Lernens als Trainingssätze in den Corpus der verwendeten KI ein. Dadurch werden auch personenbezogene Daten für das systemimmanente KI-Training verwendet. Wenn man personenbezogene Daten zum Training in ein KI-System aufnimmt, ergeben sich daraus erhebliche Datenschutzprobleme.

Diese weitreichende Datennutzung ist ein eigener Grundrechtseingriff und mit zusätzlichen Risiken für die betroffenen Menschen verbunden. Das Recht auf Korrektur falscher Daten oder auf Löschung von Daten sowie das Auskunftsrecht werden dadurch faktisch entkernt, weil in KI-Systeme eingelernte Daten nicht wieder vollständig „entlernt“ werden können. Weder ein Zurückrufen noch eine Aktualisierung von Daten im eigentlichen Sinne ist mehr sinnvoll machbar, wenn sie in eine KI eingelernt und darin weiterverarbeitet wurden.

---

[14] Vertragsstrafen dürften wegen Milliardenverträgen im Heimatmarkt im Fall von Palantir kaum abschreckende Wirkung entfalten. Im August 2025 schloss Palantir einen Vertrag mit der US-Armee im Umfang von bis zu 10 Milliarden US-Dollar, vgl. https://www.army.mil/article/287506/u_s_army_awards_enterprise_service_agreement_to_enhance_military_readiness_and_drive_operational_efficiency, abgerufen am 27. März 2026.

[15] J. Edelhoff, L. Jeric, et. al.: Wird die Palantir-Software unangemessen genutzt? https://www.tagesschau.de/investigativ/ndr-wdr/palantir-einsatz-deutschland-100.html vom 19. Juni 2025, abgerufen am 27. März 2026.

Vorschriften zur Anonymisierung oder andere technische Schutzmaßnahmen für Betroffene, deren Daten für das KI-Training verwendet werden, sind nicht erkennbar. Auch zeitliche oder andere Begrenzungen sind nicht ersichtlich, so dass in das Training auch ältere Daten aller Art einfließen könnten.

Da kommerzielle Anbieter das Anlernen und Optimieren der KI und die Ausgestaltung ihres Softwaresystems in der Regel als Geschäftsgeheimnisse betrachten und nicht publik machen, erhält die Polizei als Anwender (oder auch als Anbieter im Sinne der KI-Verordnung) keinen ausreichenden Einblick darin, wie der Vertragspartner das KI-Training praktisch vollzieht und wie auch personenbezogene Daten dabei verarbeitet werden.

Die KI-Verordnung schreibt Anbietern von bestimmten KI-Systemen zwar vor, Informationen über das Training zu veröffentlichen, allerdings hat der hessische Gesetzgeber keine Regelung für Transparenz vorgesehen. Auch jenseits von Veröffentlichungspflichten gibt es im Gesetz keine Vorgaben über die Ausgestaltung und Durchführung des KI-Trainings, etwa gegenüber dem Vertragspartner Polizei, auch nicht solche, die Zusicherungen über personenbezogene Daten beinhalten würden.

Palantir als bisheriger Partner der Polizei Hessen vermarktet seine Software als „Software as a Service“ und entsendet nach eigenen Angaben dafür stets Personal,[16] um das System zu verwalten, die IT-Sicherheit zu gewährleisten, Fehler zu analysieren und zu beheben und die Produktweiterentwicklung durchzuführen. Dazu dürfte auch das KI-Training gehören. Für dieses Personal des Vertragspartner, das auf polizeiinterne Daten für das KI-Training Zugriff nimmt, fehlen Vorgaben im hessischen Gesetz.

8) Strukturelle Auswirkungen der Datenanalyse

Der Einsatz von automatisierten Datenanalyse-Werkzeugen setzt die Polizei systematischen Fehlern aus, beispielsweise durch sog. Automation Bias. Das Phänomen beschreibt eine bei Menschen oft beobachtete Voreingenommenheit gegenüber Ergebnissen oder Vorschlägen von Softwaresystemen verglichen mit menschlichen Ergebnissen: Die maschinellen Resultate werden weniger kritisch betrachtet und weniger hinterfragt als beispielsweise Entscheidungen von Kollegen.

Es ist eine negative Auswirkung der Datenanalyse, dass automatisierte Ergebnisse als Ersatz für die eigene sorgfältige Informationsverarbeitung und -bewertung verwendet werden. Die maschinellen Resultate und Vorschläge gelten Menschen tendenziell als zu verlässlich.

Dadurch steigt das Risiko, dass unzutreffende Annahmen des Systems verstärkt werden, da sie durch die scheinbare Objektivität der Software legitimiert erscheinen. So können etwa

---

[16] In der militärnahen Palantir-Terminologie „forward deployed software engineers“, vgl. Palantir-FAQ https://blog.palantir.com/about-palantir-ddddb78aec29, abgerufen am 27. März 2026.

fehlerhafte Verknüpfungen des Systems direkt zu fälschlichen operativen Maßnahmen gegen Personen oder Personengruppen führen.

Der Gesetzgeber in Hessen fordert in § 25a Absatz 6 HSOG von den Polizeibehörden, dass sie sicherstellen, „dass diskriminierende Algorithmen weder herausgebildet noch verwendet werden“. Hierbei handelt es sich um einen begrüßenswerten Vorsatz. In der Praxis wird es, besonders bei vorbeugender Bekämpfung von Straftaten, allerdings beim bloßen Vorsatz bleiben.

Die zur automatisierten Datenanalyse verwendete Software basiert auf bestehenden Daten aus der polizeilichen Arbeit. Diese Daten tragen bereits Voreingenommenheiten in sich, da beispielsweise Minderheiten etwa bei Kontrollen besonders häufig betroffen sind, was dazu beiträgt, deren Vorkommen in entsprechenden Datenbanken zu erhöhen.

Diskriminierende Effekte verstärken sich durch den Einsatz von selbstlernenden Systemen weiter. Wird zur automatisierten Datenanalyse eine intransparente, auch mit externen Daten vortrainierte Software verwendet, stellt dies eine Hürde für die Umsetzung der gesetzlichen Forderung der Diskriminierungsfreiheit dar.

Generell verfestigen sich in der automatisierten Analyse von großen Datenmengen bereits bestehende Verzerrungen im Sinne struktureller und institutioneller Vorurteile, die in den Analyseergebnissen in mindestens genauso starkem, oft auch stärkerem Maße vorkommen wie in den Ausgangsdaten. Dadurch, dass mit „HessenDATA“ ein kommerzieller Anbieter als Vertragspartner für die Auswertung der Polizeidaten gewählt wurde, lassen sich derartige Verzerrungen kaum aufdecken. Denn die Innereien des Softwaresystems sind ein Geschäftsgeheimnis, mithin sind solche systemischen Risiken kaum adressierbar.

9) Überwachungsgesamtrechnung

Es fehlt an einer Erfassung der Gesamtheit aller staatlichen Überwachungsmaßnahmen im Sinne einer Überwachungsgesamtrechnung. Die automatisierte Datenanalyse und die dauerhafte Zusammenführung von großen Datenmengen aus unterschiedlichsten Quellen führen zu einer neuen Eingriffsdimension und sind letztlich ein paradigmatischer Wechsel. Denn zum einen führt die Kombination der Daten zu einer höheren Aussagekraft über Menschen und Personengruppen und ihren Beziehungen. Zum anderen wird im Hintergrund persistent und unabhängig von den manuellen Abfragen eine beständige Zusammenführung und Analyse einer kaum überschaubaren Datenmenge durchgeführt.

Eine kumulative Betrachtung der neuen automatisierten Eingriffsbefugnisse und der Wirkung des Einsatzes von KI-Systemen fehlt bisher. Welche Gesamtbelastung dieser Überwachung für einen Großteil der Bevölkerung als Betroffene in Hessen und darüberhinaus entsteht, nur weil sie aus beliebigem Anlass in Polizeidatensammlungen auftauchen, ist noch zu ermitteln.

Mit der in Hessen durch das Gesetz kaum beschränkten Datenanalyse droht ein schleichender und unverhältnismäßiger Überwachungsausbau.

Fazit

Die automatisierte Datenanalyse und dauerhafte Zusammenführung verschiedener Datensammlungen nach § 25a HSOG ermöglichen einen besonders intensiven Grundrechtseingriff. Die Auswertung großer, heterogener Datenbestände führt zu einer Qualität der Datenverarbeitung, die weit über bisherige polizeiliche Maßnahmen hinausgeht.

Eine Datenanalyse sollte nur bei einem konkreten Anlass erfolgen dürfen, nicht jedoch in permanenter Form im Hintergrund. Massenhaft Daten aus öffentlichen und privaten Quellen in Echtzeit zu kombinieren, muss vom Gesetzgeber schon präventiv untersagt werden.

Der Gesetzgeber hat im vorliegenden Gesetz keine parlamentarische und keine öffentliche Kontrolle hinsichtlich der Nutzung der automatisierten Datenanalyse vorgeschrieben. Dies ist angesichts der häufigen und intensiven Grundrechtseingriffe durch ein solches System aber zwingend erforderlich. Entsprechende Berichtspflichten sind vorzuschreiben, um überhaupt Evaluierungsmöglichkeiten zu eröffnen und auch einen gesellschaftlichen Diskurs über den Einsatz von KI und automatisierter Datenrasterung in der Polizei am Leben zu erhalten.

In der derzeitigen Form ist der intransparente Einsatz von KI in einem so methodenoffenen System, das auch zu systematischen Fehlern neigt, nicht zu rechtfertigen, auch nicht für das bisher unbeschränkte KI-Training.

Die Hoheit über ein eingesetztes System und die Definition der Funktionalitäten muss an den Gesetzgeber übergehen. Falls überhaupt eine automatisierte Datenanalyse in Hessen – und in allen Bundesländern – in grundrechtskonformer Weise stattfinden kann, sollte verpflichtend eine unabhängige und quelloffene Softwarelösung zum Einsatz kommen. Die einfließenden Daten sollten in Art und Umfang klar bestimmt und stärker begrenzt werden. Daten über psychisch Kranke sollten von jeder Form automatisierter Datenanalysen ausgenommen werden. Bereits trainierte KI-Systeme sind zu löschen und die Betroffenen zu benachrichtigen.