Die POLIZEI

UNABHÄNGIGE, INTERDISZIPLINÄRE FACHZEITSCHRIFT FÜR ÖFFENTLICHE UND PRIVATE SICHERHEIT

Schwerpunkt KI in der Polizei

AUS DEM INHALT

Aufsätze

Heft 4a
April 2026
Seiten 217–248
117. Jahrgang
Art.-Nr. 56244614
PVSt 5624

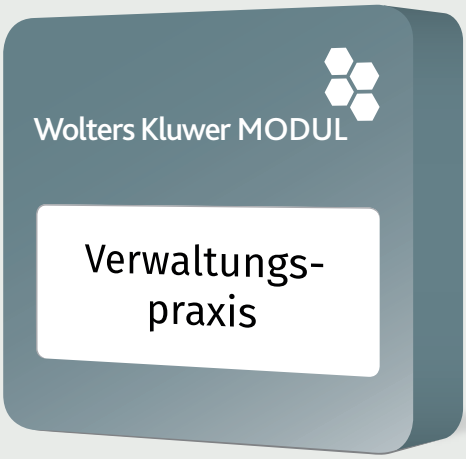
4a

Carl Heymanns Verlag

Umfassende Unterstützung für Städte, Kreise und Gemeinden

Mit dem Modul Verwaltungspraxis auf dem neuesten Stand:

- Die ideale Online-Bibliothek für Kommunalverwaltungen oder Bundesbehörden mit über 90 fundierten Titeln
- Inklusive Top-Zeitschriften wie **DVBl – Deutsches Verwaltungsblatt**, **ZWW – Zeitschrift für Wahlorganisation und Wahlrecht** und **RiA – Recht im Amt**, jeweils mit Online-Archiv
- Mit aktueller und rechtssicherer Urteils- und Normensammlung
- Profitieren Sie von den Vorteilen eines Abonnements: stets aktuelle Inhalte und komfortable Tools, die Ihre Recherche erleichtern. Mit Wolters Kluwer Recherche haben Sie außerdem Zugriff auf unsere kostenlose Rechtsprechungs- und Gesetzesdatenbank.



Preis auf Anfrage.

Zusammenfassungen von Urteilen und Beschlüssen powered by Expert AI
Nutzen Sie das Potential künstlicher Intelligenz, um den juristischen Rechercheprozess zu verkürzen.

Profitieren Sie dabei von einer klaren optischen Trennung zwischen Sachverhalt und Begründung, damit Sie schneller finden, was Sie suchen. Besonders hervorzuheben ist der Fokus auf die entscheidungserheblichen Gründe, durch die Sie die wesentlichen Argumente des Gerichts auf einen Blick erfassen können.

Überzeugen Sie sich selbst von der Qualität unserer Zusammenfassung powered by Expert AI auf Wolters Kluwer Online.

Auch im Handel erhältlich

Wolters Kluwer

shop.wolterskluwer-online.de →

Die
POLIZEI

UNABHÄNGIGE, INTERDISZIPLINÄRE
FACHZEITSCHRIFT FÜR ÖFFENTLICHE
UND PRIVATE SICHERHEIT

INHALT 4a · 2026

Editorial

Sandra Schmidt

Liebe Leserinnen und Leser,

die Transformation der Digitalisierung in allen Lebensbereichen schreitet unaufhaltsam voran; sie ist die Grundlage für Künstliche Intelligenz, deren Nutzung aus unserem Berufs- und Lebensalltag nicht mehr wegzudenken ist. KI ist »die Fähigkeit einer Maschine, menschliche Fähigkeiten wie logisches Denken, Lernen, Planen und Kreativität zu imitieren« (EU-Parlament, 2020, https://www.europarl.europa.eu/topics/de/article/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt; 06.02.2026); eine Eigenschaft, die den Menschen in vielen Lebenslagen unterstützen kann, soweit er damit verantwortungsvoll umgeht.

Was haben KI-Systeme und die Polizeiorganisation gemeinsam? KI-Systeme sind anpassungsfähig, denn sie lernen kontinuierlich aus neu zur Verfügung stehenden Daten. Damit werden sie grds. treffsicherer und verbindlicher in ihrer Aussagekraft. Auch die Polizei ist ein lernendes System; die Organisation passt sich, mehr oder weniger schnell, an neue Einsatz- und Kriminalitätsphänomene sowie gesellschaftliche und technische Entwicklungen an. Sie konnte und musste lernen, dass die Nutzung von KI unverzichtbar für eine weitere Professionalisierung der Polizeiarbeit ist.

Für dieses Editorial habe ich Chat-GPT gefragt, welche Vorteile die Nutzung von KI-Systemen bei der Bewältigung polizeilicher Aufgaben haben kann. Die Antwort kam prompt: Entlastung der Polizei, schnellere und bessere Analysen, Unterstützung bei der Gefahreneinschätzung, Reduzierung menschlicher Willkür, Unterstützung bei der Auswertung umfangreichen Bild- und Videomaterials. Bei vollem Bewusstsein, dass ich mit meiner Eingabeaufforderung die Antwort gelenkt und so provoziert habe, lediglich die Sonnenseiten der KI-Anwendung bei der Polizeiarbeit präsentiert zu bekommen, formulierte ich einen weiteren Prompt: Worin liegen die Risiken, bei der Polizeiarbeit KI zu nutzen? Die daraufhin generierte Antwort zeigte, dass die KI die für einen Rückgriff zur Verfügung stehenden Daten ebenso nüchtern auch selbstkritisch auswertet und verknüpft: Verzerrung und Diskriminierung (wenn die von der KI genutzten Daten bspw. schon diskriminierende Vorurteile enthalten), Intransparenz (denn, so Chat-GTP, »KI-Systeme sind nicht nachvollziehbar«), Fehler mit realen Folgen, Aushöhlung von Grundrechten (bspw. bei Massenüberwachungen und automatisierten Datenauswertungen), Verantwortungsdiffusion (wer trägt die Verantwortung, wenn Fehler durch die Nutzung von KI entstehen), Abhängigkeit von Technik. Daraus lässt sich ableiten, dass KI in der Polizeiarbeit unzweifelhaft zur Steigerung der Effektivität und Effizienz beitragen kann, sie aber kontrolliert, transparent und in einem klaren Rechtsrahmen eingesetzt werden muss, um die Risiken abzufedern und letztlich menschliche Entscheidungen zu unterstützen (und nicht zu automatisieren).

Unmittelbar hieran anschlussfähig sind die juristischen Beiträge in diesem Sonderheft vom Autorenteam *Dr. Anna Louban, Dr. Lou Therese Brandner und unserer Fachredakteurin Prof. Dr. Sabrina Schönrock*, von *Prof. Dr. Kristin Pfeffer* sowie von *Steven Kleemann und Prof. Dr. Hartmut Aden*.

Praktisch wird es im Beitrag von *Dr. Robert Pelzer, Dr. Florian Ludwig und Martin Hofer*. Die Autoren berichten aus dem Forschungsprojekt »KISTRA«, welches die Möglichkeiten des ethisch und rechtlich vertretbaren Einsatzes KI zur Früherkennung und Bekämpfung von Hasskriminalität im digitalen Raum untersuchte und das Ziel verfolgte, KI-basierte Textklassifizierer zur automatisierten Analyse großer Textmengen im Internet zu entwickeln. Im Projekt wurden zudem auch rechtliche und gesellschaftlich-ethische Rahmenbedingungen analysiert.

Dr. Clemens Seibold befasst sich mit einem Kriminalitätsphänomen, mit dem die Polizei sich konfrontiert sieht: Durch KI-basierte Verfahren erstellte Gesichtsmorphs können ein erhebliches Sicherheitsrisiko darstellen, wenn sie für kriminelle Zwecke missbraucht werden. Neben der Beschreibung von Gesichts-Morphing-Angriffen werden im Beitrag auch Möglichkeiten zum Schutz vor Morphing-Angriffen sowie zu Detektionsmöglichkeiten von manipulierten Bildern aufgezeigt.

Auch in der Polizei-Führung kommt KI in den Fokus. *Dr. Sylke Piéch, Thomas Berger und Raphael Humbel* diskutieren die Fragen, wie KI-Systeme den Polizeialltag unterstützen können, welche Rolle Führungskräfte im Prozess des digitalen Wandels einnehmen und wie sich Führung durch den Einsatz von KI-Technologien verändern kann. Zur Beantwortung dieser Fragestellungen erörtert das Autorenteam konkrete Anwendungsfälle bei den Länderpolizeien Berlin und Baden-Württemberg. Der Beitrag wird wissenschaftlich fundiert durch Erkenntnisse aus einer Master-Thesis an der Hochschule Luzern, die forschungsbasierte Perspektiven auf KI und Führung im Polizeikontext aufzeigt.

Editorial

Kilian D. Grütter aus Zürich überschreibt seinen Beitrag mit der Überschrift: »Generative KI und die Transformation polizeilicher Führungskultur - Chancen, Risiken und die Neudefinition von Autorität im Zeitalter algorithmischer Assistenz«. Der Autor analysiert auf Basis aktueller empirischer Forschung und klassischer Führungstheorien, wie KI die polizeiliche Führungspraxis verändert.

Liebe Leserinnen und Leser, die Autorinnen und Autoren der Beiträge in diesem Sonderheft sind wahrlich Expertinnen und Experten auf ihrem Wissensgebiet. Ich wünsche Ihnen viel Vergnügen und Erkenntnisgewinn beim Lesen dieser Ausgabe.

Ihre

Sandra Schmidt

Aufsätze

Die
POLIZEI

UNABHÄNGIGE, INTERDISZIPLINÄRE FACHZEITSCHRIFT FÜR ÖFFENTLICHE UND PRIVATE SICHERHEIT

Heft 4a/2026 · 117. Jahrgang · Seiten 217–248

Aufsätze

KI in der Polizeiarbeit: Menschliche Aufsicht als ethische und institutionelle Herausforderung[1]

Dr. Anna Louban, Berlin, Dr. Lou Therese Brandner, Tübingen und Prof. Dr. Sabrina Schönrock, Berlin*

Die fortschreitende Digitalisierung und die wachsende Notwendigkeit, große Datenmengen effizient zu verarbeiten, verändern derzeit zahlreiche gesellschaftliche Bereiche – auch die Strafverfolgung. Die Integration von Künstlicher Intelligenz (KI) in die Polizeiarbeit verspricht dabei eine Steigerung von Effizienz, Präzision und Objektivität. Zugleich sind mit dem polizeilichen KI-Einsatz erhebliche Risiken verbunden: diskriminierende Verzerrungen, mangelnde Nachvollziehbarkeit automatisierter Entscheidungen und unklare Verantwortlichkeiten stellen zentrale Herausforderungen dar.

Die am 01.08.2024 in Kraft getretene Verordnung Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz ((EU) 2024/1689) – kurz: EU AI Act – adressiert diese Problematik, indem er rechtliche Rahmenbedingungen für die Entwicklung und Nutzung von KI-Systemen schafft – vorwiegend durch die Verpflichtung zu »menschlicher Aufsicht« bei sogenannten Hochrisiko-Anwendungen (Art. 14 VO (EU) 2024/1689).

In Deutschland wird KI bereits in verschiedenen Bereichen der Polizei erprobt und eingesetzt – etwa zur Analyse großer Bild- und Videodatenmengen oder für die Textanalyse, bspw. in Bezug auf Dokumente oder Chatverläufe. Drei vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Forschungsprojekte – FAKE-ID, VIKING und PEGASUS –[2] untersuchten in den Jahren zwischen 2020 und 2025 die aktuellen Strukturen, Herausforderungen und Ambivalenzen datenbasierter KI-gestützter Polizeiarbeit.

Auf Grundlage dieser Forschung beleuchtet dieser Artikel die praktische Umsetzung menschlicher Aufsicht in komplexen, arbeitsteiligen KI-Systemen der Polizei, und die technische und organisatorische Heterogenität polizeilicher KI-Anwendungen im föderalen Kontext. Daraufhin identifizieren wir konkrete Herausforderungen, die die Polizei bei der Umsetzung der gesetzlich vorgeschriebenen menschlichen Aufsicht über KI-Systeme erfährt, und skizzieren mögliche Strategien zur rechtskonformen und ethisch informierten Überwindung dieser Schwierigkeiten.

1 Die Ersterscheinung dieses Beitrags erfolgte in: *Schönrock/Aden*, 8. Fachsymposium zum Terroranschlag auf dem Berliner Breitscheidplatz, Zukunftssicherheit: Die Rolle von Künstlicher Intelligenz im Kampf gegen Terrorismus (2025), S. 38 f.

* Dr. *Anna R. Louban* war bis Oktober 2025 wissenschaftliche Mitarbeiterin am Forschungsinstitut für öffentliche und private Sicherheit (FÖPS Berlin) der Hochschule für Wirtschaft und Recht Berlin und arbeitet zum Einsatz Künstlicher Intelligenz (KI) in der nationalen Strafverfolgung, u.a. in den Forschungsprojekten FAKE-ID zu Videoanalyse mit Hilfe Künstlicher Intelligenz zur Detektion von falschen und manipulierten Identitäten und VIKING zur Vertrauenswürdigkeit von KI-Anwendungen in der Polizeiarbeit. *Dr. Lou Therese Brandner* ist Postdoktorandin und wissenschaftliche Mitarbeiterin am IZEW Tübingen und forscht im Bereich der KI- und Datenethik zur Analyse ethischer und gesellschaftlicher Aspekte von datengetriebenen Technologien. Sie war wissenschaftliche Mitarbeiterin in den Projekten PEGASUS zu KI in der polizeilichen Ermittlungsarbeit (2022–2023) und KITQAR zu ethischen Dimensionen von KI-Trainingsdatenqualität (2022–2024), sowie Leiterin des Projektes VIKING zur Vertrauenswürdigkeit KI-gestützter Polizeitechnologien (2022–2025). *Prof. Dr. Sabrina Schönrock* ist Professorin für Öffentliches Recht an der HWR Berlin, sowie erste Vizepräsidentin der HWR Berlin. In den Jahren 2014–2024 war Sabrina Schönrock Richterin des Verfassungsgerichtshofes des Landes Berlin. In den BMBF-geförderten Forschungsprojekten FAKE-ID und VIKING übernahm sie die Co-Projektleitung gemeinsam mit Prof. Dr. Hartmut Aden.

2 Projektinformationen: Polizeiliche Gewinnung und Analyse heterogener Massendaten zur Bekämpfung organisierter Kriminalitätsstrukturen (PEGASUS), Projektlaufzeit: April 2020 – September 2023; Videoanalyse mit Hilfe künstlicher Intelligenz zur Detektion von falschen und manipulierten Identitäten (FAKE-ID); Projektlaufzeit: Mai 2021 – Oktober 2024; Vertrauenswürdige Künstliche Intelligenz für polizeiliche Anwendungen (VIKING), Projektlaufzeit Januar 2022 – März 2025.

I. Menschliche »Aufsicht« in komplexen Systemen

KI-gestützte polizeiliche Ermittlungsarbeit zählt zu den Hochrisikoanwendungen i.S.d. EU AI Acts (Art. 6, VO (EU) 2024/1689). Gem. Art. 14 dieser Verordnung sind Anbieter und Betreiber solcher Systeme zur Einrichtung wirksamer menschlicher Aufsicht verpflichtet.

Vereinfachte Modelle und die daran anknüpfenden Diskurse suggerieren dabei häufig, beim sogenannten *human in the loop* handle es sich um einen klar definierten, singulären menschlichen Eingriff in einen ansonsten automatisierten Ablauf. Tatsächlich handelt es sich in der Praxis jedoch um ein kontinuierliches, arbeitsteiliges Zusammenspiel verschiedenster professioneller Rollen und technischer Systeme – ein verteiltes Geflecht von Verantwortlichkeiten, das an mehreren Stellen des soziotechnischen Prozesses ansetzt. Die Reduktion auf einen einzelnen »Loop-Moment« verzerrt nicht nur die tatsächliche Vielschichtigkeit menschlicher Beteiligung, sondern birgt auch die Gefahr, Verantwortlichkeit zu externalisieren oder zu entkontextualisieren.

In der Praxis sind vielmehr zahlreiche menschliche Akteure an den jeweiligen soziotechnischen Prozessen des polizeilichen KI-Einsatzes beteiligt – darunter Ermittler:innen, Datenanalyst:innen, Techniker:innen, Systemverwalter:innen und juristische Fachkräfte. Ihre Zusammenarbeit verläuft nicht linear: Interviews im Rahmen der Projekte FAKE-ID, VIKING und PEGASUS mit Polizeimitarbeiter:innen, die in solche Arbeitsprozesse involviert sind, zeigen, dass sich die Abstimmung zwischen technischem Einsatz, menschlicher Urteilsfähigkeit und rechtlichen Rahmenbedingungen als dynamisch und kontextabhängig erweist. Die Ausgestaltung der Zusammenarbeit orientiert sich an der jeweiligen Aufgabenstellung und kann zahlreiche Schleifen, Umwege und Unterbrechungen umfassen, in denen Aspekte wie technische Details, individuelle Expertisen und spezifische Verantwortlichkeiten geklärt werden müssen.

Auch die Interaktion zwischen Mensch und Technik ist – wie unsere Forschung verdeutlicht – weder geradlinig noch einheitlich strukturiert. Vielmehr ist sie geprägt von situativen Einschätzungen, institutionellen Rahmenbedingungen und spezifischen technologischen Möglichkeiten. Dabei wird eine deutliche Diskrepanz zwischen idealisierten Vorstellungen einer digitalisierten Strafverfolgung und der tatsächlichen Umsetzungspraxis sichtbar. Letztere ist geprägt von enormen Datenmengen und einer Vielzahl heterogener Softwarelösungen innerhalb der Polizeiarbeit – Aspekte, die lineare Prozessannahmen infrage stellen und die Komplexität digital unterstützter Ermittlungen verdeutlichen.

II. KI im Polizeialltag: Zwischen technischer Integration und föderaler Vielfalt

Die beachtliche Größe der deutschen Polizei als gesellschaftliche Institution sowie die notwendige thematische Ausdifferenzierung polizeilicher Arbeit führen zu Vielschichtigkeiten im bereits komplexen Zusammenspiel von Akteuren und KI-Technologien. Hinzu kommen föderale Unterschiede in der Organisation und Durchführung polizeilicher Arbeit, die zu einer erheblichen Diversität im Umgang mit KI-gestützten Anwendungen führen: Aufgrund der föderalen Struktur der Polizei werden KI-Anwendungen in den verschiedenen Strafverfolgungs- und Ermittlungsbehörden unterschiedlich antizipiert und umgesetzt. Zentrale Faktoren wie die eingesetzten Technologien, verantwortliche Akteure und konkrete Arbeitsabläufe im Umgang mit KI können variieren.

Während manche Abteilungen innerhalb einer Polizei oder ganze Landespolizeien eigene KI-Systeme entwickeln und sich somit in der Rolle des Anbieters befinden, setzen andere auf externe Softwarelösungen und agieren entsprechend als Betreiber. Diese »Mischlösungen« aus kommerziellen und eigens entwickelten KI-Anwendungen bringen sowohl Herausforderungen als auch Chancen mit sich – wie etwa *Kai Küllmer*, Leitender Kriminaldirektor beim Bundeskriminalamt (BKA), jüngst beim 28. Europäischen Polizeikongress[3] in Berlin betonte: »*Die Auswertung von Millionen von Chatnachrichten wäre ohne KI nicht mehr möglich. Mit der im BKA entwickelten Netzwerkanalyse konnten über 4.000 Ermittlungsverfahren in Deutschland eingeleitet werden*«, so *Küllmer* auf dem Podium zum Thema »KI und Massendaten«.

Es zeigt sich eine insgesamt heterogene Landschaft polizeilicher KI-Anwendungen: Hochkomplexe Technologien treffen auf ebenso komplexe Organisationsformen. Polizeibedienstete kommen auf unterschiedlichen Ebenen mit KI in Berührung – etwa im Rahmen kriminalistischer Ermittlungen oder im administrativen Bereich, bei der Integration neuer Systeme oder der Auswahl geeigneter Technologien. Die Vielfalt an Anwendungen, Einsatzbereichen und Akteurskonstellationen – in Kombination mit den erheblichen föderalen Unterschieden innerhalb der Polizei – führt zu erheblichen Herausforderungen für eine wirksame Gestaltung und einheitliche Umsetzung der gesetzlich vorgeschriebenen »menschlichen Aufsicht« (Art. 14 VO (EU) 2024/1689).

III. Der AI Act: Verantwortung durch menschliche Kontrolle

Der EU AI Act verlangt, dass bei hochrisikoreichen KI-Systemen eine menschliche Aufsicht gewährleistet sein muss. Für Strafverfolgungsbehörden gelten diverse Ausnahmen der Hochrisikopflichten des AI Acts (*Kleeman & Aden* 2025). Dies ist auch der Fall für menschliche Aufsicht; so wird in Art. 14 VO (EU) 2024/1689 spezifiziert, dass Betreiber von Systemen zu Zwecken der Strafverfolgung nicht sicherstellen müssen, dass von Systemen getroffene Identifizierungen von mindestens zwei natürlichen Personen gesondert überprüft und bestätigt werden, wenn das Unions- oder nationale Recht die Anwendung dieser Anforderung für unverhältnismäßig hält. Strafverfolgungsbehörden sind jedoch nicht von der Anforderung ausgenommen, dass natürliche Personen die Funktionsweise von KI-Systemen stets überwachen sollen. Natürliche Personen müssen sicherstellen können, dass KI-Systeme zu Strafverfolgungszwecken bestimmungsgemäß verwendet werden und verlässliche Ergebnisse liefern. Aufsichtspersonen müssen also in der Lage sein, die Systeme zu verstehen, ihre Grenzen zu erkennen, Fehlfunktionen zu identifizieren und – wenn nötig – einzugreifen oder Systeme zu deaktivieren.

Einerseits bedeutet das, dass KI-Systeme bereits während der Entwicklung mit Mechanismen ausgestattet werden sollen, die Aufsichtspersonen anleiten und informieren, damit diese

3 28. Europäischer Polizeikongress (European Police Congress/EPC), Zeitenwende für die Innere Sicherheit – Strategie · Resilienz · Zusammenhalt, 20.–21.05.2025, Berlin.

entscheiden können, ob, wann und wie sie in das System eingreifen oder es abschalten müssen, wenn es nicht wie vorgesehen funktioniert. Doch vor allem setzt die Anforderung entsprechende Kompetenzen innerhalb der Polizeistrukturen voraus. So spezifiziert der AI Act in Art. 14 »Menschliche Aufsicht« bezüglich der häufig diskutierten Problematik des Automatisierungsbias, dass Aufsichtspersonen in der Lage sein müssen »sich einer möglichen Neigung zu einem automatischen oder übermäßigen Vertrauen in die von einem Hochrisiko-KI-System hervorgebrachte Ausgabe (»Automatisierungsbias«) bewusst zu bleiben, insbesondere wenn Hochrisiko-KI-Systeme Informationen oder Empfehlungen ausgeben, auf deren Grundlage natürliche Personen Entscheidungen treffen« (Art. 14 Abs. 4 Buchst. b) VO (EU) 2024/1689). Sind Aufsichtspersonen nicht für diese Problematik sensibilisiert, kann die wahrgenommene Autorität von KI-Systemen dazu führen, dass Biases unerkannt bleiben. Es kann zu einer Autoritätsumkehr kommen zwischen Anwender:in und Systemen, die die Entscheidungsfindung eigentlich nur unterstützen sollen (*Yeung et al.* 2021). Weiterhin müssen für eine klare Rollenverteilung innerhalb der Polizei spezifische Verantwortlichkeiten und Aufgaben der involvierten Akteure klar differenziert und definiert werden, z.B. in Entwicklung, Bereitstellung, Anwendung.

3.1 Herausforderungen der menschlichen Aufsicht in der Praxis

Menschliche Aufsicht kann also keine individuelle Verantwortung sein, sondern muss auf institutioneller Seite mitgedacht, organisiert und strukturell verankert werden. Um KI-Risiken effektiv erkennen und mindern zu können, müssen Aufsichtspersonen über die erforderliche Kompetenz, Ausbildung und Autorität verfügen, sowie angemessene Intentionen und Motivation haben (*Sterz et al.* 2024). Das entspricht jedoch bei weitem nicht immer der Realität: Die gesammelten Daten zeigen, dass Ermittelnde oft nur begrenzt mit den technischen Grundlagen der eingesetzten Systeme vertraut sind, während Datenanalyst:innen wiederum nicht immer die juristischen oder ermittlungstaktischen Hintergründe kennen. Dadurch entsteht eine Verantwortungslücke, die durch besser geschulte »hybride Rollen« oder strukturelle Reformen geschlossen werden könnte und müsste. Gleichzeitig ist die konkrete Ausgestaltung effizienter menschlicher Aufsicht unklar, da das Konzept oft vage und lückenhaft bleibt (*Enqvist* 2023).

Die Umsetzung menschlicher Aufsicht gestaltet sich in der Praxis dementsprechend schwierig. Komplexe KI-Systeme produzieren Ausgaben, die oft nicht ohne Weiteres nachvollziehbar sind. Ermittler:innen müssen sich auf Rücksprachen mit Spezialist:innen verlassen – was Zeit kostet und zu Missverständnissen führen kann. Fehlinterpretationen von KI-Systemen sind besonders problematisch; bspw. sehen sich Systeme zur Sprach- und Textanalyse mit der Kontextgebundenheit von Sprache bezüglich Semantik und Sinnhaftigkeit konfrontiert, etwa Faktoren wie Ironie und Sarkasmus. Eine automatische Analyse kann zu folgenreichen Fehlschlüssen führen, wenn Inhalte falsch übersetzt oder soziokulturelle Kontexte nicht erkannt werden, was im polizeilichen Kontext bei der Verwendung von Slang oder aufgrund millieuspezifischer mehrsprachiger Kommunikation in Chatverläufen schnell der Fall sein kann. Menschliche Aufsicht spielt hier eine zentrale Rolle, um Ergebnisse kritisch zu hinterfragen und prüfen zu können, setzt jedoch im gleichen Zuge auf menschlicher Seite Kenntnisse bestimmter Sprachen und Sprachgewohnheiten voraus. Sind diese Fähigkeiten nicht ausreichend vorhanden, besteht das Risiko, dass falsche Schlussfolgerungen gezogen und infolgedessen unrechtmäßige Maßnahmen durch die Polizei ergriffen werden.

Das Konzept der menschlichen Aufsicht berührt weitere ethischen Dimensionen: Ohne vorhandene technologische Transparenz ist es Anwender:innen kaum möglich, Systeme kritisch zu hinterfragen. Ein Beispiel dafür sind mögliche Fehlidentifizierungen von Gesichtserkennungssystemen und anderen biometrischen Technologien, die potenziell schwerwiegende Folgen haben wie fälschliche Verdächtigungen bis hin zu unverschuldeten Festnahmen. Nur wenn Anwender:innen solche KI-Systeme ausreichend verstehen, die internen Prozesse nachvollziehen und die Ausgaben der jeweiligen KI-gestützten Anwendungen entsprechend interpretieren können, können sie deren Entscheidungen und Ergebnisse hinreichend gegenprüfen.

Zudem ist in vielen Bereichen der Polizei eine wenig ausgeprägte Fehlerkultur, also die Bereitschaft, offen über Probleme und Fehleinschätzungen zu sprechen, erkennbar. So gab Bundespolizeibeauftragter *Uli Grötsch* in einem Interview im März 2025[4] an, dass die deutsche Polizei »definitiv« Nachholbedarf hinsichtlich einer Fehlerkultur habe und Polizeibeschäftigte mit existierenden polizeiinternen Beschwerdestellen »– vorsichtig ausgedrückt – oft nicht zupass kommen«. Auch ethische Reflexionen über KI-Einsatz finden bislang selten institutionell verankert statt. Der AI Act bringt somit nicht nur neue Pflichten, sondern kann auch bestehende Schwachstellen verstärken, wenn diese neuen Aufgaben auf unklare Verantwortungsketten und mangelnde Mechanismen zur Aufarbeitung von Fehlverhalten und Missständen treffen.

3.2 Ethische Verantwortung: Mehr als Technikfolgenabschätzung

Die Verpflichtung zu menschlicher Aufsicht rückt die menschliche Handlungsfähigkeit und Autonomie in den Fokus und reflektiert die Bedeutsamkeit ethischer Gesichtspunkte für die Regulierung von KI. Im sensiblen Bereich der Strafverfolgung, in dem Entscheidungen über rechtliche Unschuld oder Schuld, Freiheit oder der Entzug dieser gefällt werden, greift eine rein technische Perspektive auf den Einsatz Künstlicher Intelligenz zu kurz. Menschliche Aufsicht darf nicht allein als Frage der Systemkontrolle oder der organisatorischen Effizienz verstanden werden, sondern muss als ethisches Mandat begriffen werden. Das bedeutet, dass diejenigen, die KI-Systeme entwickeln, implementieren oder anwenden, nicht nur technisch geschult sein sollten, sondern auch eine fundierte Auseinandersetzung mit ethischen Grundfragen führen müssen. Vor diesem Hintergrund ist die Polizei als gesellschaftliche Kerninstitution dafür verantwortlich, die notwendigen Strukturen für eine solche Auseinandersetzung zu schaffen und in interne Arbeitsprozesse zu integrieren.

4 *Grötsch*: Polizei hat »definitiv« Problem mit Fehlerkultur, WEB.DE Magazin, 15.03.2025. https://web.de/magazine/politik/groetsch-polizei-definitiv-problem-fehlerkultur-40755194 (Zugriff: 03.06.2025).

Notwendig sind gezielte Schulungsprogramme für Anwender:innen, die sowohl technisches als auch ethisches und rechtliches Wissen vermitteln, um Polizeibeschäftigte dazu zu befähigen, zu kompetenten KI-Nutzer:innen werden (*Burrell* 2016). Der Aufbau von Kompetenzen zu den Themenfeldern Fairness und Diskriminierung, Transparenz, Rechenschaftspflicht und damit verbundene gesellschaftliche Auswirkungen algorithmischer Entscheidungen kann sie darin unterstützen, KI-Ergebnisse einzuordnen, kritisch zu hinterfragen und die mit den jeweiligen Entscheidungen einhergehenden Risiken besser einzuschätzen. Nur wenn Polizeiakteure – je nach Funktion und Verantwortung – die sozialen, juristischen und ethischen Dimensionen ihres Handelns verstehen und aktiv in ihre Praxis einbeziehen, kann der Einsatz von KI im Polizeialltag verantwortungsvoll ausgestaltet werden.

Das betrifft zudem die Frage, wie Vertrauen in die Polizei als Institution gewahrt oder gestärkt werden kann. Institutionelle Mechanismen, die ethische Standards absichern, sind hierfür notwendig – etwa durch Ombudspersonen, unabhängige Ethikbeiräte oder niedrigschwellige, möglichst anonyme Beschwerdeverfahren, die Transparenz und Teilhabe fördern. Auf diese Art würde die Möglichkeit, missbräuchliche oder anderweitig ethisch bedenkliche Technologienutzung und deren benachteiligende Auswirkungen auf einzelne Personen oder Personengruppen zu melden, institutionell verankert.

Fazit: Zwischen technologischer Innovation und institutioneller Verantwortung

Die Integration Künstlicher Intelligenz in die Polizeiarbeit eröffnet zweifellos neue Möglichkeiten – etwa für effizientere Abläufe, komplexe datenbasierte Analysen und präzisere Ermittlungen anhand von Mustern, Aufgaben, die durch menschliche Arbeitskraft nur unter extremem Aufwand übernommen werden können. Gleichzeitig wirft sie jedoch erhebliche ethische, rechtliche und auch praktische Herausforderungen auf: Intransparenz, potenzielle Diskriminierung, unklare Zuständigkeiten und damit einhergehende unscharf definierte Verantwortlichkeiten zählen zu den zentralen Problemfeldern. Der EU AI Act setzt mit der Verpflichtung zur menschlichen Aufsicht einen wichtigen Rahmen für den Umgang mit Hochrisiko-KI, doch seine Umsetzung trifft auf eine grundsätzlich föderal strukturierte Institution sowie auf höchst ausdifferenzierte, heterogene Arbeitsstrukturen innerhalb der einzelnen Polizeien.

Angesichts der aktuell zunehmenden Technologisierung steigt der Bedarf nach Rollendefinitionen, einer fundierten technischen wie ethischen Aus- und Weiterbildung für das Personal und einer institutionellen Verankerung von Verantwortung. Die Etablierung einer offenen Kommunikation über Probleme und Fehleinschätzungen ist dabei unumgänglich.

Da Polizeibedienstete auf unterschiedlichen Ebenen mit KI in Berührung kommen – direkt in Ermittlungen oder indirekt über administrative und strategische Prozesse –, verlangt die Diskussion um ‚die' menschliche Aufsicht einer diskursiven und analytischen Pluralisierung, um die Komplexität der vielfältigen polizeilichen Arbeit realitäts- und praxisnah erfassen und reflektieren zu können. Anwendungen, die auf Künstlicher Intelligenz basieren, können in diesem spezifischen Kontext nicht als ein Allheilmittel angesehen werden, sondern sind als Werkzeug zu begreifen, dessen Wirkung maßgeblich davon abhängt, wie es genutzt wird und wie die Konsequenzen seiner Anwendung kontrolliert und reflektiert werden.

Die Verpflichtung zur menschlichen Aufsicht ist zwar ein wesentlicher Aspekt ethisch akzeptabler Technologien, stellt jedoch ebenso wenig ein Allheilmittel für die mannigfachen ethischen Probleme des polizeilichen KI-Einsatzes dar – vor allem, wenn sie als individuelle Aufgabe statt als strukturelle Verantwortlichkeit verstanden wird (*Zweig et* al. 2018). Ein idealisierter Diskurs über die symbolische Bedeutsamkeit menschlicher Aufsicht, Kontrolle und Letztentscheidung hilft wenig, wenn es an einer konkreten Ausgestaltung und Implementation solcher Verpflichtungen, sowie einer praxisnahen Analyse ihrer Beziehung zu anderen ethischen Dimensionen mangelt. Erst wenn technologische Innovation mit institutioneller Verantwortung und ethischem Bewusstsein einhergeht, kann KI einen sozial vertretbaren Fortschritt in der Polizeiarbeit und insbesondere der polizeilichen Strafverfolgung herbeiführen.

Referenzen

Burrell, J. (2016). How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. Big Data & Society 3(1). https://doi.org/10.1177/2053951715622512

Enqvist, L. (2023). Human Oversight' in the EU Artificial Intelligence Act: What, When and by Whom? Law, Innovation and Technology 15 (2): 508–35. http://doi.org/10.1080/17579961.2023.2245683

Kleemann, S. und Aden, H. (2025). Die Nutzung Künstlicher Intelligenz durch Strafverfolgungsbehörden – »Hintertüren« der Verordnung der Europäischen Union über Künstliche Intelligenz. In: Honekamp, W. und Kemme, S. (Hrsg.): Auswirkungen von KI auf die zukünftige Polizeiarbeit — Ein technischer, rechtlicher und kriminologisch-sozialwissenschaftlicher Blick. Springer Gabler (i.E.).

Sterz, S., Baum, K., Biewer, S., Hermanns, H., Lauber-Rönsberg, A., Meinel P. und Langer M. (2024). On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives. In Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24). Association for Computing Machinery, New York, NY, USA 2495–2507. https://doi.org/10.1145/3630106.3659051

Yeung, D., Khan, I., Kalra, N. und Osoba, O. A. (2021). Identifying Systemic Bias in the Acquisition of Machine Learning Decision Aids for Law Enforcement Applications. RAND Corporation. http://www.jstor.org/stable/resrep29576Z

Zweig, K. A., Fischer, S. und Lischka, K. (2018). Wo Maschinen irren können. Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung. Bertelsmann Stiftung. http://doi.org/10.11586/2018006

Nach der Regulierung ist vor der Regulierung

Künstliche Intelligenz in der Polizeiarbeit nach dem AI-Act

von Prof. Dr. Kristin Pfeffer, Hamburg*

Nach einem dreijährigen Gesetzgebungsprozess hat die Europäische Union am 21.05.2024 den AI-Act (KI-VO) verabschiedet und am 12.07.2024 im Amtsblatt (VO (EU) 2024/168954) veröffentlicht. Auch wenn sich der Anwendungsbereich dieses speziellen Produktsicherheitsgesetzes auf die mitgliedstaatlichen Sicherheitsbehörden erstreckt, handelt es sich nicht um ein klassisches Polizeigesetz. Der Beitrag gibt einen kurzen Überblick über die Folgen des »bereichsübergreifenden« Regulierungsansatzes der KI-VO für Sicherheitsbehörden und den mitgliedstaatlichen Gesetzgeber.

Einleitung

Die KI-VO verfolgt erstmalig einen sektorenübergreifenden Regelungsansatz, der die charakteristischen Herausforderungen der Regulierung von und des Umgangs mit KI-Systemen in Angriff nimmt und deren gesamten »Lebenszyklus« (von der Entwicklung, bis zur Verwendung) erfasst. Auch Sicherheitsbehörden unterliegen damit nun künftig den Anforderungen der KI-VO.

Die EU stützt sich hierbei auf ihre Kompetenz aus Art. 114 AEUV zur Binnenmarktregulierung und Art. 16 AEUV zum Datenschutz, ErwGr 3. Ziel des Gesetzes ist ein funktionierender Binnenmarkt durch einen einheitlichen Rechtsrahmen insbesondere für die Entwicklung, das Inverkehrbringen, die Inbetriebnahme und die Verwendung von Systemen künstlicher Intelligenz (KI-Systeme) in der Union unter Beachtung der Werte der Union, Förderung der Technologie unter gleichzeitiger Beachtung der Grundrechte, Art. 1 Abs. 1 KI-VO.

Ob eine Verordnung, die auch das Inverkehrbringen, die Inbetriebnahme und die Verwendung von KI-Systemen durch Gefahrenabwehr- und Strafverfolgungsbehörden regelt, überhaupt von der Gesetzgebungskompetenz der EU umfasst ist, ist wegen der essentiellen Bedeutung des Rechts der Inneren Sicherheit für die Souveränität eines Staates[1] nach wie vor umstritten.[2]

Die KI-VO ist an Risiken ausgerichtet, nicht an Sektoren oder Anwendungsbereichen. Mithin werden sektorspezifische Besonderheiten etwa des Polizei- und Strafrechts nur in Ansätzen berücksichtigt. Die KI-VO unterscheidet sich grundlegend von der Regulierung vieler anderer neuer Technologien, welche sich auf bestimmte Bereiche beschränken und keine entsprechende umfassende gesellschaftliche Relevanz aufweisen. Sie ist umfassender und weist eine besondere grundrechtliche Relevanz auf.[3]

Grundsätzlich ist die KI-VO gem. Art. 288 AEUV in allen Teilen und in jedem Mitgliedstaat verbindlich. Im Bereich Polizei- und Strafverfolgungsbehörden sind die Regeln aber durch zahlreiche Ausnahmen regelrecht »perforiert« worden. Hier überlässt es der Gesetzgeber im Wesentlichen den Mitgliedstaaten – unter Beachtung des von der KI-VO vorgegebenen verbindlichen Mindeststandards – den Einsatz von KI-Systemen mit Ermächtigungsgrundlagen zu erlauben und zu regulieren.

Aus deutscher Sicht zu beachten ist, dass der in der KI-VO stets verwendete Begriff der »Strafverfolgung« europarechtlich auszulegen ist und sowohl die Verhütung, Ermittlung, Aufdeckung und Verfolgung von Straftaten, als auch die Gefahrenabwehr für die öffentliche Sicherheit einschließt, Art. 3 Nr. 46 KI-VO. Die in der deutschen Rechtsordnung geltende klare Differenzierung zwischen Gefahrenabwehr und Strafverfolgung,[4] erfolgt nämlich weder durch die EU, noch in den meisten übrigen Mitgliedstaaten.

I. Flexibler Regulierungsrahmen und technologieoffener Ansatz

Der EU-Gesetzgeber hat für die KI-VO einen flexiblen Regulierungsrahmen gewählt, indem er Durchführungsrechtsakte zur Regulierung und Kategorisierung von KI-Systemen auf die EU-Kommission (unter Mitwirkung der Mitgliedstaaten) delegiert. Auf diese Weise soll der Rechtsrahmen an neue technische Entwicklungen und Erkenntnisse angepasst werden.[5] Der Regulierungsansatz ist technologieoffen: Mit einem verpflichtenden Risikomanagementsystem für Hochrisikosysteme soll ein kontinuierlicher Prozess der Anpassung an die technischen Entwicklungen erfolgen. Darüber hinaus soll die Pflicht zur Grundrechtefolgenabschätzung, Art. 27 KI-VO, und die Möglichkeit der Einrichtung von Reallaboren durch die Mitgliedstaaten, Art. 57 KI-VO, die Grundlage für innovationsfördernde Anpassungen ermöglichen. Diese dynamische Ausgestaltung des KI-Gesetzes geht zwangsläufig zu Lasten der Rechtssicherheit: Denn je präzisier eine Regelung gestaltet wird, desto schneller würde sie von der Dynamik der Entwicklung der KI-Systeme eingeholt werden. Noch mangelt es an ausreichenden empirischen Erfahrungen mit der neuen Technologie. Die Folgen insbesondere für die Grundrechte sind schwer vorherzusehen.[6]

Vor diesem Hintergrund ist auch die sehr weite Definition des KI-Systems im Gesetz zu sehen. Gem. Art. 3 Abs. 1 der

---

* Professorin für Öffentliches Recht an der Hochschule der Polizei Hamburg, Leiterin der Forschungsstelle für Europäisches und Deutsches Sicherheitsrecht (FEDS). Bei dem Beitrag handelt es sich um die überarbeitete Fassung eines Vortrages auf dem 7. Hamburger Sicherheitsrechtstag am 05.11.2024 in Hamburg.

1 BVerfGE 123, 267 (358) – Lissabon Vertrag.

2 BR-Drucks. 488/1/21; Eingehend dazu K. Pfeffer, NVwZ 2023, S. 1286, 1289; Schöndorf-Haubold/Giogios, KI im Einsatz für die Sicherheit: Innovation und Kontrolle im Spannungsfeld von europäischer Gesetzgebung und nationaler Souveränität, VerfBlog, 2024/12/10, https://verfassungsblog.de/ki-im-einsatz-fur-die-sicherheit/ (Zugriff: 26.05.2025).

3 Die Entwicklung der Verankerung des Grundrechtsschutzes während des Gesetzgebungsprozesses der KI-VO nachzeichnend Palmiotto, The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the Legislative Process and the Future of the Regulation. European Journal of Risk Regulation. https://doi.org/10.1017/err.2024.97 (Zugriff: 26.05.2025).

4 Dazu statt vieler *Graulich*, in: Lisken/Denninger, Handbuch des Polizeirechts, 7. Aufl. 2021, E. Rn. 221 ff.

5 Etwa in Leitlinien gem. Art. 96 Abs. 1b) KI-VO zu verbotenen Praktiken i.S.v. Art. 5 KI-VO.

6 Zum Ganzen *Seitz*, EuZW 2024, S. 839, 845 f.

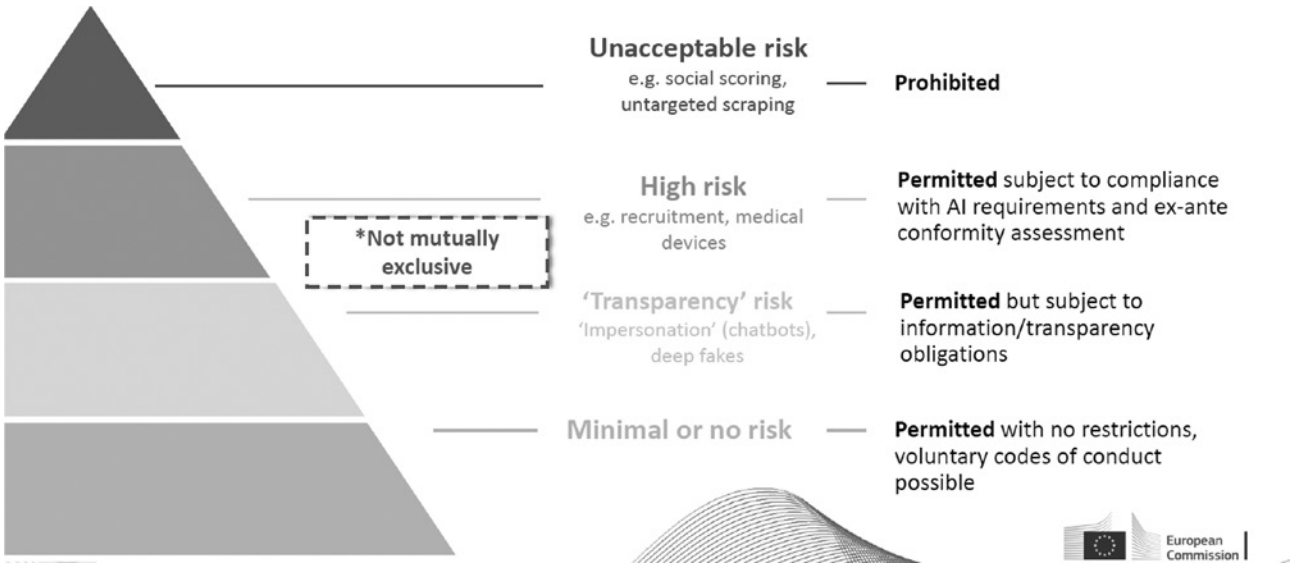


Abb.: EU-Kommission

KI-VO ist ein KI-System: »eine Software, die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren«. Umfasst sind nicht Systeme für maschinelles Lernen, sondern erfasst sind auch logik- und wissensbasierte Systeme. Die weite Definition erfasst somit auch Systeme mit wenig Schadenspotential, was z.T. als unnötige Überregulierung kritisiert wurde.[7]

II. Risikobasiert

Die KI-VO ist als Produktsicherheitsgesetz risiko- und weniger grundrechtebasiert.[8] In welchem Maße die unterschiedlichen KI-Systeme durch die KI-VO reguliert werden, hängt somit von deren Schadenspotential ab (risikobasierter Ansatz).

Je höher das Risiko des Systems gewichtet wird, desto strenger sind die Vorschriften für die Handhabung. Dabei erscheint ein risikobasierter Ansatz aus dem Bereich Produktsicherheitsrecht zur Binnenmarktregulierung gem. 114 AEUV wenig tauglich, die betroffenen Grundrechte beim Einsatz von KI-Systemen durch Sicherheitsbehörden zu schützen: Polizeigesetze und Strafprozessrecht sind grundsätzlich als Schranken für den Schutzbereich von Grundrechten ausgestaltet, der Regelungsansatz ist hier »grundrechtebasiert« statt »risikobasiert«. Das Risikoverständnis des Produktsicherheitsrechts beruht dagegen auf potentiellen Schäden, die sodann ausgeglichen werden. Das »Risiko« für Grundrechte liegt aber in einer potentiellen Verletzung des Grundrechts, was nicht gleichbedeutend mit einem Schaden ist.[9] Potentielle Grundrechtsverletzungen erfordern stets eine Einzelfallbetrachtung, sie lassen sich nicht standardisieren.[10]

III. Ausgenommen: »Nationale Sicherheit«

Die Ausnahme von KI-Systemen für die »nationale Sicherheit« aus dem Anwendungsbereich der KI-VO in Art. 2 Abs. 3 UAbs. 1 KI-VO hat lediglich deklaratorische Bedeutung – bereits nach dem Primärrecht verbleibt die alleinige Verantwortung für die nationale Sicherheit und Aufrechterhaltung der öffentlichen Ordnung bei den Mitgliedstaaten, Art. 4 Abs. 2 Satz 3 EUV (sog. ordre-public-Vorbehalt). Der EuGH legt den Begriff der »nationalen Sicherheit« in ständiger Rechtsprechung eng aus. Es gehe dabei um den Schutz der wesentlichen Funktionen des Staates und der grundlegenden Interessen der Gesellschaft, um »die Verhütung und Repression von Tätigkeiten, die geeignet sind, die tragenden Strukturen eines Landes im Bereich der Verfassung, Politik oder Wirtschaft oder im sozialen Bereich in schwerwiegender Weise zu destabilisieren und insbesondere die Gesellschaft, die Bevölkerung oder den Staat als solchen unmittelbar zu bedrohen (…)«. Letzteres umfasse insbesondere die Verhütung und Repression von terroristischen Aktivitäten.[11] Die Bekämpfung schwerer Kriminalität wird vom EuGH dagegen (nur) unter den Begriff »innere Sicherheit« subsumiert.[12]

IV. Verbotene Praktiken (mit Ausnahmen)

In Art. 5 Abs. 1 KI-VO werden insgesamt acht Gruppen von KI-Praktiken für grundsätzlich unzulässig erklärt, was für Systeme zur Strafverfolgung und Gefahrenabwehr in Art. 5 Abs. 1 UAbs. 1 Buchst. h) i.V.m. Abs. 2 KI-VO von vornherein durch einige weitreichende Ausnahmen abgeschwächt wird. Gemäß dem risikobasierten Ansatz der KI-VO werden diejenigen Praktiken verboten, welche ein inakzeptables hohes Risiko, insbesondere für die Grundrechte, darstellen. Hier wird bei »Praktiken« angesetzt und nicht bei KI-Systemen. Es geht hierbei um auf KI-Systeme bezogene Tätigkeiten, nämlich das Inverkehrbringen, die Inbetriebnahme und/oder die Verwendung von KI-Systemen zu einem bestimmten Zweck.[13] Dabei dürfte die Verwen-

7 *Ebers*, Truly Risk-Based Regulation of Artificial Intelligence – How to Implement the EU's AI Act. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=%204870387 (Zugriff: 28.09.2024), S. 17.
8 *Seitz*, EuZW 2024, S. 839, 843.
9 *Wendehorst*, in: Martini/Wendehorst (Hrsg.), KI-VO, 2024, Art. 6 Rn. 103.
10 *Wendehorst*, in: Martini/Wendehorst (Hrsg.), KI-VO, 2024, Art. 6 Rn. 105.
11 EuGH 21.06.2022 – C-817/19 – Ligues des droits humaines.
12 EuGH 05.04.2022 – C-140/20 – Commissioner of An Garda Síochána. Zum Trennungsprinzip zwischen Polizei und Geheimdiensten *Denninger*, in: Lisken/Denninger, Handbuch des Polizeirechts, 7. Aufl. 2021, B. Rn. 43 ff.; Zur grundsätzlichen Aufgabentrennung von Polizei und Militär *Weingärtner*, in: Lisken/Denninger, Handbuch des Polizeirechts, 7. Aufl. 2021, I. Rn. 679 ff.; Zu den Konflikten bei der Auslegung des Begriffs nationale Sicherheit *Pfeffer*, NVwZ 2023, S. 1288.
13 *Wendehorst*, in: Martini/Wendehorst (Hrsg.), KI-VO, 2024, Art. 5 Rn. 2.

dung von zahlreichen unbestimmten Rechtsbegriffen künftig zu erheblicher Rechtsunsicherheit führen.[14] Die Tatbestände dieser »Black-List« sind bereits wegen der problematischen Rechtssetzungskompetenz der EU in diesem Bereich eng auszulegen, sodass nur solche Systeme als verboten anzusehen sind, deren Verwendung offensichtlich im »Widerspruch« zu den Werten der EU, wie der Achtung der Menschenwürde, der Grundrechte sowie der Demokratie und der Rechtsstaatlichkeit (ErwGr 28) stehen.[15] Folgende schädliche KI-Anwendungen aus dem Bereich der Polizeiarbeit werden ganz verboten, Art. 5 Abs. 1 KI-VO:

- bestimmte KI-Systeme, die **manipulative Techniken** einsetzen oder **persönliche Schwächen ausnutzen**, um eine Person zu einem schädigenden Verhalten zu bewegen, Art. 5 Buchst. a), b)
- **Social Scoring**, bestimmte KI-Systeme zur Bewertung des sozialen Verhaltens, Art. 5 Buchst. c)
- **bestimmte biometrische Kategorisierungssysteme**, die darauf abzielen, Rückschlüsse u.a. auf ethnische Herkunft, politische Meinungen oder sexuelle Orientierung zu ziehen, Art. 5 Buchst. g)
- **personenbezogenes Predictive Policing:** KI-Systeme zur Vorhersage künftiger Straftaten auf der Grundlage von Profiling bzw. Persönlichkeitsmerkmalen, Art. 5 Buchst. d)
- **Image Scraping:** KI-Systeme, die Datenbanken zur Gesichtserkennung durch das Auslesen von Gesichtsbildern aus dem Internet oder aus Videoüberwachungsanlagen erstellen, Art. 5 Buchst. e).

Die EU-Kommission hat hierzu inzwischen Leitlinien i.S.v. Art. 96 Abs. 1 Buchst. b) der KI-VO erarbeitet.[16]

V. Biometrische Gesichtserkennung

Zu den während des Gesetzgebungsprozesses wohl umstrittensten Einsatzgebieten von KI-Systemen durch Sicherheitsbehörden gehört die biometrische Gesichtserkennung in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken.[17] Dabei erscheint die Differenzierung zwischen biometrischer Fernidentifikation in Echtzeit und ex-post, wie sie die KI-VO nun vornimmt, indem sie erstere Praxis grundsätzlich verbietet (und eng begrenzte Ausnahmen vorsieht) und letztere (nur) als Hochrisiko-KI-System betrachtet, aus grundrechtlicher Sicht nicht überzeugend.[18] Bei der Fernidentifikation in Echtzeit geht es um einen gegenwärtigen, bei der retrograden Identifikation dagegen um einen abgeschlossenen Geschehensablauf. Letztere dient daher primär repressiven Zwecken, während mit der Echtzeit-Fernidentifizierung auch Rückschlüsse auf den gegenwärtigen Aufenthaltsort einer Person getroffen werden können. Bei der Echtzeit-Fernidentifizierung ergibt sich ein hohes Eingriffsgewicht aufgrund der Unmittelbarkeit zwischen Erfassung der biometrischen Daten und dem Abgleich mit der Referenzdatenbank, eine Kontrolle oder Korrektur sind hier nur eingeschränkt möglich. Die mit der Bestimmung des Aufenthaltsorts einer Person verbundene jederzeitige Zugriffsmöglichkeit stellt einen besonders schweren Grundrechtseingriff dar (so auch ErwGr 32). Aber auch eine nachträgliche biometrische Fernidentifikation kann erhebliche Grundrechtsbelastungen nach sich ziehen, die im Unterschied zur Echtzeit-Fernidentifikation jedoch aufgrund der zeitlichen Differenz einfacher durch Verfahren aufgefangen werden können. Es können hierbei aber über einen längeren Zeitraum umfangreiche Bewegungsprofile erstellt werden. Im Zusammenhang mit Demonstrationen kann dies abschreckend wirken. Vom Verbot biometrischer Gesichtserkennung in Echtzeit zu Strafverfolgungszwecken in öffentlich zugänglichen Räumen regelt die KI-VO mehrere Ausnahmen, zum Teil mit sehr unbestimmten, noch klärungsbedürftigen Tatbeständen.[19]

VI. Komplexe Verwaltungsverfahren

Die KI-VO regelt umfangreiche Verfahrensanforderungen, die einen sehr hohen Verwaltungsaufwand für die Sicherheitsbehörden nach sich ziehen und vom mitgliedstaatlichen Gesetzgeber gesetzgeberisch umzusetzen sind.[20] Dies gilt bspw. für den ausnahmsweisen Einsatz biometrischer Gesichtserkennung in Echtzeit zu Strafverfolgungszwecken in öffentlich zugänglichen Räumen. Hier gelten

- die Pflicht zur Grundrechtefolgenabschätzung (Art. 27 KI-VO),
- Genehmigungspflichten (Art. 5 Abs. 3, 5 KI-VO),
- Registrierungspflichten (Art. 49 KI-VO),
- Anzeige- und Berichtspflichten (Art. 5 Abs. 4, 6, 7 KI-VO)
- Anforderungen an die Verwertung der Ergebnisse des Prozesses des KI-Systems (Art. 5 Abs. 3 letzter Satz KI-VO).

VII. Hochrisiko-KI-Systeme durch Einsatzgebiet

Die KI-VO zählt KI-Systeme schon deshalb zu Hochrisikosystemen, wenn KI-Systeme für einen der in Anhang III der KI-Verordnung aufgeführten Anwendungsfälle mit hohem Risiko bestimmt sind. Diese Liste enthält Anwendungsfälle aus Bereichen wie Bildung, Beschäftigung, Strafverfolgung oder Migration. Nach Art. 6 Abs. 2 KI-VO liegt somit ein Hochrisiko-KI-System insbesondere dann vor, wenn es in einem der in Annex III aufgeführten Bereiche eingesetzt wird, u.a. KI-Systeme, die in den Bereichen Strafverfolgung, Migration und Grenzkontrolle eingesetzt werden, soweit sie nicht bereits verboten sind, sowie Systeme für die Justizverwaltung und demokratische Prozesse; KI-Systeme, die zur biometrischen Identifizierung, biometrischen Kategorisierung und Emotionserkennung verwendet werden, soweit sie nicht ohnehin bereits verboten sind. Gem. Anhang III Nr. 6 der KI-VO zählen dazu im Bereich Strafverfolgung

- KI-Systeme zur Bewertung des Risikos einer Person, Opfer eines Verbrechens zu werden,
- Lügendetektoren,
- Systeme für die Bewertung der Zuverlässigkeit von Beweisen bei strafrechtlichen Ermittlungen oder Strafverfolgungen,
- Systeme zur Bewertung des Risikos einer Person, straffällig zu werden oder erneut straffällig zu werden, die nicht nur auf der Grundlage eines Profils oder der Bewertung von Persönlichkeitsmerkmalen oder früherem kriminellen Verhalten beruhen,
- Systeme zum Profiling bei der Ermittlung, Untersuchung oder Verfolgung von Straftaten.

14 *Wendehorst*, in: Martini/Wendehorst (Hrsg.), KI-VO, 2024, Art. 5 Rn. 2.
15 *Krönke*, NVwZ 2024, S. 529. 531.
16 Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), 04.02.2025, C(2025) 884 final. https://digital-strategy.ec.europa.eu/de/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act (Zugriff 23.05.2025).
17 Dazu *Wendehorst*, in: Martini/Wendehorst (Hrsg.), KI-VO, 2024, Art. 5 Rn. 137.
18 S.a. EDSA-EDSB, Gemeinsame Stellungnahme 5/2021 v. 18.06.2021, Rn. 31, https://www.edpb.europa.eu/system/files/2021-10/edpb-edps_joint_opinion_ai_regulation_de.pdf (Zugriff 23.05.2025).
19 *Krönke*, NVwZ 2024, S. 529, 531 f.
20 *Vasel*, EuZW 2024, S. 829, 832.

Die EU-Kommission erarbeitet bis zum 02.02.2026 Leitlinien für die Hochrisiko-Einstufung, Art. 6 Abs. 5 KI-VO.

Die Regelung für »Hochrisiko-KI-Systeme durch Einsatzgebiet« wird jedoch stark abgeschwächt durch die Rückausnahme in Art. 6 Abs. 3 Uabs. 1 KI-VO, die immer dann gilt, wenn das System (doch) kein erhebliches Risiko »für die Gesundheit, die Sicherheit oder die Grundrechte natürlicher Personen birgt, indem es unter anderem nicht das Ergebnis der Entscheidungsfindung wesentlich beeinflusst«. Hier wird die Diskrepanz zwischen risikobasiertem Ansatz und grundrechtebasiertem Ansatz besonders deutlich. Nicht nur die hier gewählten sehr unbestimmten Rechtsbegriffe erscheinen im Hinblick auf die Grundrechtsrelevanz problematisch. Auch wird die Risikoeinschätzung auf einer vorläufigen Selbsteinschätzung beruhen, mit der Gefahr, dass hierunter absichtlich subsumiert wird, um die Anforderungen zu umgehen.[21]

VIII. Umfangreiche Betreiberpflichten

Verwenden Sicherheitsbehörden Hochrisiko-KI-Systeme, so treffen diese als Betreiber erhebliche Pflichten.[22] Gem. Art. 3 Nr. 4 KI-VO ist Betreiber, wer ein KI-System in eigener Verantwortung verwendet. Betreiber sind dazu verpflichtet, geeignete technische und organisatorische Maßnahmen zu ergreifen, um die Verwendung des Hochrisiko-KI-Systems nach Maßgabe der Betriebsanleitung sicherzustellen, Art. 26 Abs. 1 KI-VO, die menschliche Aufsicht einer natürlichen Person mit der erforderlichen Befähigung und Befugnis zu übertragen, Art. 26 Abs. 2 KI-VO, und sicherzustellen, dass die Eingabedaten relevant und ausreichend repräsentativ sind, Art. 26 Abs. 4 KI-VO. Natürliche Personen sind zu informieren, wenn ein Anhang III-System die betreffenden Entscheidungen trifft oder bei solchen Entscheidungen zur Unterstützung verwendet wird, Art. 26 Abs. 11 KI-VO. Betreiber müssen den Betrieb des Systems zudem überwachen und den Anbieter informieren, wenn sie Grund zur Annahme haben, dass der Einsatz des Systems gemäß der Gebrauchsanweisung die Nonkonformität begründet, Art. 26 Abs. 5 KI-VO. Im Falle eines möglichen Risikos i.S.d. Art. 79 KI-VO ist die zuständige Marktüberwachungsbehörde in Kenntnis zu setzen. Für alle Hochrisiko-KI-Systeme gem. Art. 6 Abs. 2 i.V.m. Anhang III (mit Ausnahme III Ziff. 2 – kritische Infrastrukturen) ist von Polizei- und Strafverfolgungsbehörden eine Grundrechtefolgenabschätzung vorzunehmen, also eine Abschätzung der Auswirkungen, die die Verwendung auf die Grundrechte haben kann. Darin enthalten sein muss gem. Art. 27 KI-VO eine Beschreibung der Verfahren des Betreibers, bei denen das Hochrisiko-KI-System im Einklang mit seiner Zweckbestimmung verwendet wird, eine Beschreibung des Zeitraums und der Häufigkeit, innerhalb dessen bzw. mit der jedes Hochrisiko-KI-System verwendet werden soll, sowie die Kategorien der natürlichen Personen und Personengruppen, die von seiner Verwendung im spezifischen Kontext betroffen sein könnten. Der Betrieb von Hochrisiko-KI-Systemen durch Polizei- und Strafverfolgungsbehörden ist einer Registrierungspflicht unterworfen, allerdings ist der Leserkreis des sicheren, nichtöffentlichen Teils auf die EU-Kommission sowie die Marktüberwachungsbehörden beschränkt, Art. 49 Abs. 4, 5 KI-VO.

IX. KI-Kompetenz, Art. 4 KI-VO

Polizei- und Strafverfolgungsbehörden, die KI-Systeme einsetzen, müssen zuvor für die KI-Kompetenz ihrer Bediensteten Sorge tragen. Gem. Art. 4 KI-VO haben sie Maßnahmen zu ergreifen, »um nach besten Kräften sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ein ausreichendes Maß an KI-Kompetenz verfügen, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind«. Zu diesen Maßnahmen gehören die Entwicklung von internen Richtlinien und Standards, die Fortbildung und Schulung, Zertifizierungsprogramme, praxisorientiertes Lernen in divers zusammengesetzten Teams, die Ernennung von behördlichen KI-Beauftragten.[23]

X. Governance-Struktur

Die KI-VO etabliert eine komplexe Governance-Struktur durch Benennung mehrerer zuständiger Behörden sowohl auf EU-Ebene als auch auf mitgliedstaatlicher Ebene.[24] Auf Unionsebene wurde das Büro für künstliche Intelligenz, Art. 64 KI-VO, das Europäische Gremium für künstliche Intelligenz (sog. KI-Gremium), Art. 65 KI-VO, das Beratungsforum, Art. 67 KI-VO, und das Wissenschaftliche Gremium unabhängiger Sachverständiger (sog. wissenschaftliches Gremium), Art. 68 KI-VO, geschaffen. Daneben hat jeder Mitgliedstaat mindestens eine notifizierende Behörde und mindestens eine Marktüberwachungsbehörde als zuständige nationale Behörden einzurichten bzw. zu benennen, Art. 70 KI-VO. Eine Marktüberwachungsbehörde soll darüber hinaus als zentrale Anlaufstelle für den AI-Act dienen, Art. 70 Abs. 2 KI-VO. Diesen (mindestens) 27 mitgliedstaatlichen Behörden wurden wichtige Aufgaben sowohl vor einer Markteinführung, Art. 30 ff. KI-VO, als auch bei der Marktüberwachung einschließlich aufsichtsrechtlicher Untersuchungsbefugnisse übertragen, Art. 74 ff. KI-VO. Die Entscheidung, wer in Deutschland die zuständigen Behörden sein sollen, war zunächst umstritten. Zunächst hatten die deutschen Datenschutzbehörden, unterstützt durch den Europäischen Datenschutzausschuss (EDSA), dafür geworben, selbst (jeweils) die zuständige Marktaufsichtsbehörde zu werden.[25] Die letzte Bundesregierung hat sich schließlich dafür entschieden, die Bundesnetzagentur (BNetzA) zur geforderten nationalen KI-Marktüberwachungsbehörde zu benennen. Daneben sollen die Fachbehörden zu Vermeidung von Doppelstrukturen weiter ihre fachlichen Zuständigkeiten behalten. Die Bundesnetzagentur soll künftig zwar als zentrale Stelle die KI-VO umsetzen und gleichzeitig zum KI-Kompetenzzentrum werden, welches auch für die Innovationsförderung zuständig ist. Die

21 *Ebers/Streibörger*, RDi 2024, S. 393, 397; *Wendehorst*, in: Martini/Wendehorst (Hrsg.), KI-VO, 2024, Art. 6 Rn. 90.

22 *Vasel*, EuZW 2024, S. 829, 831 f.

23 *Wendehorst*, in: Martini/Wendehorst (Hrsg.), KI-VO, 2024, Art. 4 Rn. 12–19.

24 *Vasel*, EuZW, 829, 832 f.; *Demková/De Gregorio*, The Looming Enforcement Crisis in European Digital Policy: A rule-of-law centered path forward, VerfBlog, 2025/2/10, https://verfassungsblog.de/the-looming-enforcement-crisis-ai-dsa-eu/ (Zugriff: 26.05.2025).

25 DSK-Positionspapier »Nationale Zuständigkeiten für die Verordnung zur Künstlichen Intelligenz (KI-VO)« v. 03.05.2024, https://www.datenschutzkonferenz-online.de/media/dskb/20240503_DSK_Positionspapier_Zustaendigkeiten_KI_VO.pdf (Zugriff: 26.05.2025); EDPB, Statement 3/2024 on data protection authorities' role in the Artificial Intelligence Act framework v. 16.07.2024, https://www.edpb.europa.eu/system/files/2024-07/edpb_statement_202403_dpasroleaiact_en.pdf (Zugriff 26.05.2025).

Bundesanstalt für Finanzdienstleistungen (BaFin) soll aber weiter für den Finanzbereich, das Kraftfahrtbundesamt für den Fahrzeugbereich und auch die jeweiligen Datenschutzbehörden für den Datenschutzbereich zuständig bleiben.[26]

XI. »Nach der Regulierung ist vor der Regulierung«

Ob und in welchem Umfang ein KI-System z.B. zum Erlass von Verwaltungsakten, der Gefahrenabwehr und der Strafverfolgung sowie bei der Verarbeitung personenbezogener Daten eingesetzt werden darf und welche Vorkehrungen zum Schutz der von der Nutzung dieser Systeme betroffenen Personen getroffen werden müssen, muss durch das jeweils geltende Fachgesetz in den Mitgliedstaaten bzw. der EU geregelt werden. Nicht verbotene Systeme i.S.d. KI-VO sind in der mitgliedstaatlichen Polizeiarbeit nicht automatisch erlaubt. Die KI-VO stellt klar, dass die DSGVO (VO (EU) 2016/679)[27] und auch die Datenschutzrichtlinie für die Strafverfolgungsbehörden (RL (EU) 2016/680)) unberührt bleiben (ErwGr 63). Die KI-VO stellt zudem keine polizeiliche Ermächtigungsgrundlage dar. Im Falle eines Grundrechtseingriffes ist jeweils eine verfassungsgemäße Eingriffsermächtigung erforderlich (Gesetzesvorbehalt), in Deutschland z.B. im Verwaltungsverfahrensgesetz, den Polizeigesetzen von Bund und Ländern, in der Strafprozessordnung, im Datenschutzrecht usw. Das mitgliedstaatliche Fachrecht und das Unionsrecht formulieren bereits jetzt einige wenige KI-spezifische Anforderungen. Was die Anforderungen an die noch zu schaffenden Ermächtigungsgrundlagen für den Einsatz angeht, so sind hier also insbesondere neben den Standards der KI-VO bzw. der Datenschutzrichtlinie für die Polizei- und Strafverfolgungsbehörden (DSRL-JI) die verfassungsrechtlichen Vorgaben, die das BVerfG an solche Eingriffsermächtigungen stellt, zu beachten. So hatte das BVerfG 2023 die verfassungsrechtlichen Implikationen der automatisierten Datenanalyse anhand der Normen im Hessischen und Hamburgischen Polizeigesetz klargestellt.[28] Die Anforderungen an die jeweiligen Ermächtigungsgrundlagen sind unterschiedlich, abhängig von Art und Umfang der verarbeitbaren Daten und der zugelassenen Methode der Datenanalyse oder -auswertung.[29]

XII. Stufenweise Geltung

Die KI-VO ist am 01.08.2024 in Kraft getreten. Sie sieht nur wenig Bestandsschutz für Altsysteme vor, Art. 111 KI-VO. Im Übrigen wird sie stufenweise Geltung erlangen und erst zwei Jahre nach ihrem Inkrafttreten, am 02.08.2026, vollständig anwendbar werden, mit Ausnahme der folgenden besonderen Bestimmungen: Die Verbote, Begriffsbestimmungen und Vorschriften im Zusammenhang mit den KI-Kompetenzen haben bereits sechs Monate nach dem Inkrafttreten, am 02.02.2025, Geltung erlangt. Die Vorschriften über die Governance und die Verpflichtungen in Bezug auf KI mit allgemeinem Verwendungszweck werden 12 Monate nach dem Inkrafttreten, am 02.08.2025, anwendbar. Die Verpflichtungen für KI-Systeme, die als Hochrisiko-KI-Systeme eingestuft werden, weil sie in regulierte Produkte eingebettet sind, die in Anhang II (Liste der Harmonisierungsrechtsvorschriften der Union) aufgeführt sind, werden 36 Monate nach dem Inkrafttreten, am 02.08.2027, anwendbar. Verstöße gegen die KI-VO sollen gem. Art. 99 KI-VO von den Mitgliedstaaten bußgeldbewert werden.

Fazit

Die Anforderungen für den Einsatz von KI-Systemen werden durch die KI-VO erheblich erhöht: Polizei- und Strafverfolgungsbehörden, die KI-Systeme im Sinne der KI-VO entwickeln und einsetzen, müssen danach zusätzliche, sehr komplexe Vorschriften und Verfahren anwenden und die KI-Kompetenz durch Schulungen ihrer Mitarbeiter gewährleisten. Sicherheitsbehörden, die schon bisher KI-Systeme eingesetzt haben, mussten deren Risiko gemäß der KI-VO neu bewerten und im Falle eines verbotenen Systems, dieses bis zum 02.02.2025 abschalten. Für Hochrisiko-KI-Systeme sind komplexe Konformitätsbewertungsverfahren durchzuführen und Betreiberpflichten zu beachten, ggf. müssen erhebliche Änderungen an den bestehenden Systemen erfolgen, um den Standards der KI-VO gerecht zu werden. Sicherheitsbehörden, die die Initiative zur internen Entwicklung von KI-Systemen ergriffen haben, werden mit einer doppelten Verantwortung konfrontiert: Die Einhaltung der Vorschriften sowohl als Nutzer und als Entwickler. Sie müssen gewährleisten, dass jede Phase des Prozesses, von der Entwicklung, Datenerfassung, Schulung bis hin zum Einsatz, den Anforderungen der KI-VO entspricht.[30] Die komplexe Governance-Struktur, die die KI-VO mit einer Vielzahl von zuständigen Institutionen, Gremien und Behörden schafft, erhöht den Verwaltungsaufwand.[31]

Insgesamt erscheint die EU-Regulierung mit der KI-VO zum jetzigen Zeitpunkt begrüßenswert, denn die Technologie entwickelt sich dynamisch und ließe sich später noch schwerer regulieren. Eine Anpassung der KI-VO dürfte auch nur eine Frage der Zeit sein.[32] Der intensive Diskurs während des Gesetzgebungsprozesses hat einer breiteren Öffentlichkeit die Chancen und auch Risiken des Einsatzes von KI-Systemen in der Polizeiarbeit vor Augen geführt.

Der deutsche Gesetzgeber muss nun die erforderlichen Ermächtigungsgrundlagen schaffen und dabei sowohl die Mindeststandards und Verfahrensanforderungen der KI-VO erfüllen, als auch die verfassungsrechtlichen Vorgaben des BVerfG. Die Anforderungen sind hoch, aber nicht unerfüllbar.

---

26 Haufe Online Redaktion, 05.11.2024 https://www.haufe.de/recht/weitere-rechtsgebiete/wirtschaftsrecht/nationale-marktueberwachungsbehoerde-bei-der-ki-aufsicht_210_635468.html (Zugriff: 26.05.2025).

27 Zum Zusammenspiel von KI-VO und DSGVO *Nebel/Geminn*, DuD 2025, S. 279 ff.

28 BVerfG, Urt. v. 16.02.2023 – 1 BvR 1547/19 und 1 BvR 2634/20.

29 Dazu eingehend *Kugelmann/Buchmann*, GSZ 2024, S. 1 ff.

30 Europol, AI and policing – The benefits and challenges of artificial intelligence for law enforcement 2024, S. 11, https://www.europol.europa.eu/publication-events/main-reports/ai-and-policing (Zugriff: 26.05.2025).

31 *Vasel*, EuZW, 2024, S. 829, 832.

32 *Seitz*, EuZW 2024, S. 839, 846.

Die Grundrechte-Folgenabschätzung nach dem *Artificial Intelligence Act* der Europäischen Union – Chancen einer qualitativen Konkretisierung für die polizeiliche KI-Nutzung[1]

von Steven Kleemann und Prof. Dr. Hartmut Aden, Berlin*

Die im Sommer 2024 verabschiedete Verordnung der Europäischen Union zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Verordnung [EU] 2024/1689, im Folgenden EU-KI-VO), die in der internationalen Fachdiskussion auch als *Artificial Intelligence Act* oder kurz als *AI Act* bezeichnet wird, hat weitreichende Auswirkungen auch auf die Nutzung Künstlicher Intelligenz (KI) für polizeiliche Zwecke. Anders als beim Datenschutz, wo neben der EU-Datenschutz-Grundverordnung 2016/679 (DSGVO) die gesonderte Richtlinie (EU) 2016/680 für Polizei und Strafverfolgung (JI-Richtlinie) gilt, regelt die EU-KI-VO auch die polizeiliche KI-Nutzung, allerdings mit zahlreichen Sonderregelungen.[2] Die EU-KI-VO enthält daher nur wenige materiell-rechtliche Beschränkungen für die polizeiliche KI-Nutzung, die über die ohnehin geltende Grundrechtsbindung hinausgehen. Prägend für die polizeiliche KI-Nutzung ist eine *Double Black Box*-Problematik. Erstens ist die Technik selbst eine *Black Box,* weil die genutzten KI-Systeme und ihre Ergebnisse nur begrenzt transparent und erklärbar sind. Hinzu kommt – zweitens – eine institutionelle *Black Box,* da die organisatorischen Prozesse und die konkrete Nutzung von KI-Systemen aufgrund von Geheimhaltungsinteressen im Sicherheitsbereich tendenziell undurchsichtig sind.[3]

Die EU-KI-VO basiert auf einer Risikoklassifizierung, bei der Anbietende und Betreibende von KI-Systemen sowie weitere Beteiligte insbesondere für solche KI-Systeme Pflichten zu erfüllen haben, die als Hochrisikosysteme eingestuft werden. Nur wenige KI-Anwendungen werden als verboten klassifiziert – und selbst hier gibt es noch Ausnahmen für Strafverfolgungsbehörden. Die Artikel 8 bis 27 der EU-KI-VO treffen umfangreiche Regelungen für Hochrisiko-KI-Systeme. Im Mittelpunkt dieses Beitrags steht die Grundrechte-Folgenabschätzung (im Folgenden: GRFA), die erst im Laufe des Aushandlungsprozesses auf Betreiben zivilgesellschaftlicher Akteur:innen und des Europäischen Parlaments in die EU-KI-VO aufgenommen wurde,[4] um den Grundrechtsschutz bei der KI-Nutzung sicherzustellen (Art. 27 EU-KI-VO). Die Pflicht zur Durchführung einer GRFA vor Inbetriebnahme eines Hochrisiko-KI-Systems gilt für Einrichtungen des öffentlichen Rechts und für private Einrichtungen, die öffentliche Dienste erbringen oder dem Banken- und Versicherungssektor zuzurechnen sind. Damit ist die gesamte polizeiliche KI-Nutzung erfasst, soweit sie in die Hochrisiko-Kategorie einzuordnen ist.

Dieser Beitrag geht den Fragen nach, wie die GRFA für polizeiliche KI-Systeme durchgeführt werden sollte, welche Chancen sich hieraus für den Grundrechtsschutz ergeben und in welchem Verhältnis die GRFA zu anderen Verfahrensvorkehrungen wie der Datenschutz-Folgenabschätzung nach der DSGVO bzw. der JI-Richtlinie und zu den Risikomanagementsystemen nach Art. 9 der EU-KI-VO steht. Der Beitrag zeigt, dass die GRFA in Kombination mit der Datenschutz-Folgenabschätzung nach der DSGVO bzw. der JI-Richtlinie, Polizeibehörden dabei unterstützt, sinnvolle und rechtskonforme KI-Lösungen auszuwählen. Damit sinkt das Risiko, dass aufgrund unzulänglicher KI-Systeme Fehler in der Polizeiarbeit auftreten, etwa durch Ermittlungen und Eingriffsmaßnahmen gegen Unbeteiligte. Die Potenziale der GRFA können indes nur ausgeschöpft werden, wenn sie von unabhängigen Stellen aufgrund zu entwickelnder Standards durchgeführt und die Ergebnisse transparent gemacht werden.

I. Die Grundrechte-Folgenabschätzung als prozedurale Anforderung an polizeiliche Hochrisiko-KI-Systeme

Die EU-KI-VO enthält mehrere Ansätze zur Absicherung der rechtmäßigen und ethikkonformen KI-Nutzung durch Verfahren (prozedurale Anforderungen). Hierzu zählen neben der GRFA u.a. auch Regelungen zu Konformitätsbewertungen, zur Einrichtung von Qualitäts- und Risikomanagementsystemen sowie zur Bereitstellung technischer Dokumentationen und Betriebsanleitungen. Viele dieser Ansätze sind zugleich Transparenzanforderungen, was bezüglich technischer Dokumentationen und Aufzeichnungen z.B. konkrete Datenblätter, KI-*Model-* und *-System-Cards* oder Protokolle umfassen kann. Die Bereitstellung von Informationen für Nutzende, die Einrichtung von Mechanismen zur menschlichen Kontrolle, die Bereitstellung von Informationen über Genauigkeit, Robustheit und Cybersicherheitsmaßnahmen, die Registrierung der eingesetzten KI-Systeme in einer EU-Datenbank sowie die Durchführung von Überwachungsmaßnahmen nach dem Inverkehrbringen und die Berichterstattung an die Aufsichtsbehörden sollen die KI-Nutzung für Betroffene transparent machen und ihre rechtskonforme

* *Prof. Dr. Hartmut Aden*, Professur für Öffentliches Recht, Europarecht, Politik- und Verwaltungswissenschaft an der Hochschule für Wirtschaft und Recht (HWR) Berlin.
Steven Kleemann, Wissenschaftlicher Mitarbeiter am Deutschen Institut für Menschenrechte (DIMR) und zuvor am Forschungsinstitut für öffentliche und private Sicherheit (FÖPS), HWR Berlin.

1 Eine frühere Fassung dieses Beitrags wurde veröffentlicht in: Schönrock, Sabrina & Aden, Hartmut (Hrsg.) (2025): 8. Fachsymposium zum Terroranschlag auf dem Berliner Breitscheidplatz: Zukunftssicherheit: Die Rolle von KI im Kampf gegen den Terrorismus, Stuttgart u.a.: Boorberg, S. 47–57.
2 Ausführlicher hierzu Kleemann & Aden 2025.
3 Siehe hierzu auch Martinez, Tahraoui & Kleemann 2026.
4 Zur Genese: Mantelero 2024; Bertaina et al 2025.

Nutzung sichern (siehe hierzu insbesondere die Artikel 8 bis 27 EU-KI-VO).

Die GRFA ist nach Art. 27 EU-KI-VO vor Inbetriebnahme, also vor dem Erstgebrauch vorzunehmen. Die Betreibenden haben im Rahmen einer Prognoseentscheidung selbst abzuschätzen, inwiefern ihr Hochrisiko-System Auswirkungen auf die Grundrechte hat. Der Prüfungsmaßstab umfasst dabei zuvörderst die in der EU-Grundrechtecharta (EU-GRCh) normierten Rechte, die Rechte aus der Europäischen Menschenrechtskonvention und auch die verfassungsrechtlichen Vorgaben des Mitgliedstaates, in Deutschland also insbesondere die Grundrechte des Grundgesetzes in ihrer Auslegung durch das Bundesverfassungsgericht.[5]

So sollen vorab potenzielle Risiken für die Grundrechte der Betroffenen identifiziert werden. Falls solche Risiken festgestellt werden, ist die Inbetriebnahme des KI-Systems nur zulässig, wenn wirksame Maßnahmen zum Schutz der Grundrechte ergriffen werden. Das Ergebnis der GRFA ist eine Dokumentation.[6] Art. 27 Abs. 1 Buchst. a–f EU-KI-VO erfordert umfassende Risikobewertungen, deren Dokumentation und die Festlegung von Abhilfemaßnahmen. Allerdings können sich Probleme bei der grundrechtskonformen Nutzung von KI als soziotechnische Systeme auch erst im laufenden Betrieb zeigen, so dass vergleichbare Standards für den gesamten KI-Lebenszyklus erforderlich wären. Die GRFA sollte daher als iterativer Prozess verstanden werden.

Nicht abschließend geklärt ist das Verhältnis zwischen GRFA und dem Risikomanagementsystem gem. Art. 9 EU-KI-VO,[7] bei welchem ebenfalls Risiken für Grundrechte berücksichtigt werden müssen (Abs. 2 Buchst. a).[8] Das Risikomanagementsystem gehört zu den Pflichten der Anbietenden von Hochrisiko-KI-Systemen, die GRFA liegt dagegen in der Verantwortung der Betreibenden. Daher ist eine enge Zusammenarbeit zwischen beiden erforderlich. Nach Art. 27 Abs. 1 Buchst. d EU-KI-VO ist die Abschätzung auch aufgrund von Informationen der Anbietenden nach Art. 13 EU-KI-VO durchzuführen. Zugleich können die Anbietenden durch eine Kooperation mit den Betreibenden ihrer Beobachtungspflicht nach Inverkehrbringen nachkommen (Art. 72 i.V.m. Art. 26 Abs. 5).[9] Durch dieses Zusammenspiel kann ein höherer Schutz erreicht werden, da die Beteiligten jeweils im Rahmen ihrer originären Kompetenzen Informationen zuliefern und die Risiken beurteilen. Für die Umsetzung wirksamer und funktionierender *Accountability*-Mechanismen, wäre jedoch die Einbeziehung einer unabhängigen Stelle nötig. Dies ist in der EU-KI-VO nicht zwingend vorgeschrieben, kann aber durch die mitgliedstaatliche Gesetzgebung für öffentliche Stellen, die nach Art. 27 EU-KI-VO eine GRFA durchzuführen haben, konkretisierend vorgeschrieben werden.

Die Methoden, die bei der Durchführung der GRFA zum Einsatz kommen, sind in Art. 27 EU-KI-VO nicht näher ausgeführt. Hier sind Konkretisierungen erforderlich, die auf EU-Ebene durch das KI-Gremium vorgenommen werden können, das auf die Vereinheitlichung der mitgliedsstaatlichen Verwaltungspraxis hinwirken (Art. 66 Buchst. d) EU-KI-VO) und Empfehlungen zur Durchführung der EU-KI-VO abgeben kann (Art. 66 Buchst. e) EU-KI-VO).[10] Neben bindenden Normen (z.B. gesetzlichen Regeln der Mitgliedstaaten) kommen hier auch technische Normen in Betracht, soweit sie eine professionelle und einheitliche Durchführung der GRFA absichern. Solche Standards können z.B. im Rahmen von nationalen, europäischen oder internationalen Normungsorganisationen wie dem *Deutschen Institut für Normung* (DIN), der *International Organization for Standardization/International Classification for Standards* (ISO/IEC), dem *Institute of Electrical and Electronics Engineers* (IEEE) oder dem US-amerikanischen *National Institute of Standards and Technology* (NIST) entstehen und insbesondere Leitlinien zu Risikomanagementsystemen, Transparenz und technischen Abhilfemaßnahmen enthalten. Hier gibt es bereits konkrete Ausarbeitungen, die unter anderem zahlreiche Anforderungen der EU-KI-VO aufgreifen wie die ISO/IEC 23894 (Risikomanagement), ISO/IEC 25012 (Datenqualität und Anforderungen an Trainingsdaten), ISO/IEC TR 24027 (Identifizierung und Reduzierung von *Bias*), ISO/IEC 38507 (*Accountability* und *Governance*) oder ISO/IEC42001 (KI-Managementsysteme mit denen auch die Einhaltung ethischer Prinzipien und regulatorischer Anforderungen unterstützt werden soll) sowie ISO/IEC42005 (zur Bewertung der Auswirkungen von KI-Systemen, *AI impact assessment*). Auch die DIN SPEC 91517 zu polizeilicher KI-Nutzung kann hier herangezogen werden.

Die Ergebnisse der GRFA werden gem. Art. 27 Abs. 3 EU-KI-VO der zuständigen Marktüberwachungsbehörde übermittelt. Dies kann entweder die nationale Marktüberwachungsbehörde, der Europäische Datenschutzbeauftragte (wenn es sich um Organe der EU handelt) oder die zuständige Finanzmarktaufsicht sein. Für polizeiliche KI-Systeme ist die nationale Marktüberwachungsbehörde (Art. 3 Nr. 26 EU-KI-VO) zuständig, die jedoch gegenüber den Betreibenden über keine aufsichtsbehördlichen Befugnisse verfügt.[11] Allerdings müssen die Betreibenden mit den für den Schutz der Grundrechte zuständigen Behörden (Art. 77 EU-KI-VO) zusammenarbeiten. Aktuell sind dies in Deutschland im Rahmen ihrer jeweiligen Zuständigkeit die Bundesdatenschutzbeauftragte (für die Polizeien des Bundes) sowie die Landesdatenschutzbeauftragten (für die jeweilige Landespolizei), die Antidiskriminierungsstelle des Bundes und die Berliner Ombudsstelle der Landesstelle für Gleichbehandlung.[12] Die Prüfungs- und Beanstandungsrechte dieser Stellen sind nicht in der EU-KI-VO, sondern in den Gesetzen zu finden, auf deren Basis die benannten Stellen arbeiten. Da viele Hochrisiko-KI-Systeme nicht nur von einer Landes- oder Bundespolizeibehörde eingesetzt werden, erscheint eine enge Koordination zwischen den Stellen angeraten und erforderlich. Hier wird zu beobachten sein, ob die gemeldeten Stellen und ihre Koordination ein umfassendes Monitoring des gesamten KI-Lebenszyklus leisten können. Eine Nachmeldung von Art. 77-Stellen ist möglich.

Den benannten Stellen ist gem. Art. 77 Abs. 1 EU-KI-VO der Zugang zu allen Dokumenten zu gewähren, die sie »für

5 *Eisenberger* 2025, Art. 27, Rn. 33.
6 *Eisenberger* 2025, Art. 27, Rn. 5.
7 Vgl. *Schuett* 2024.
8 *Kleemann* & *Aden* 2025.
9 *Eisenberger* 2025, Art. 27, Rn. 11.
10 Ebd.
11 Ebd., Rn. 7.
12 Liste nach Art. 77 Abs. 2 EU-KI-VO, Stand 19.12.2024, abrufbar unter: https://www.bmwk.de/Redaktion/DE/Downloads/J-L/liste-nach-art-77-abs-2-ki-verordnung.pdf?__blob=publicationFile&v=8.

die wirksame Ausübung ihrer Aufträge im Rahmen ihrer Befugnisse innerhalb der Grenzen ihrer Hoheitsgewalt« benötigen, was die Dokumentation der GRFA einschließt. Für Polizei und Strafverfolgung stellt sich auch die Frage, inwiefern die problematisch unbestimmte Datenkategorie der »sensiblen operativen Daten« (Art. 3 Nr. 38 EU-KI-VO)[13] hier Auswirkungen hat. Im Hinblick auf die effektive Ausübung der Kontrollrechte sollten die Aufsichtsbehörden umfassenden Zugang zu den Informationen erhalten. Sie unterliegen ohnehin weitreichenden Geheimhaltungsvorschriften für sicherheits- oder datenschutzsensible Vorgänge.

Die Betreibenden sind gem. Art. 49 Abs. 1 EU-KI-VO verpflichtet ihre Hochrisiko-Systeme in einer EU-Datenbank (Art. 71 EU-KI-VO) zu registrieren. Handelt es sich (u.a.) um Behörden, so ist in dieser EU-Datenbank gem. Art. 49 Abs. 3 i.V.m. Anhang VIII Abschnitt C Nr. 4 EU-KI-VO grundsätzlich eine Ergebniszusammenfassung der GRFA zu veröffentlichen. Dadurch ist eine gewisse Transparenz der Öffentlichkeit gegenüber herzustellen.[14] Allerdings sind KI-Systeme für die Strafverfolgung (Anhang III Nr. 6 EU-KI-VO) gem. Art. 49 Abs. 4 EU-KI-VO von der Veröffentlichungspflicht ausgenommen. Hier sollen die Informationen in einem nicht-öffentlichen Teil der EU-Datenbank gespeichert werden. Art. 49 Abs. 4 UAbs. 2 EU-KI-VO regelt, dass lediglich die Europäische Kommission und die nationalen Marktüberwachungsbehörden nach Art. 74 Abs. 8 EU-KI-VO die entsprechenden Informationen erhalten, welche allerdings gem. Art. 49 Abs. 4 Buchst. c EU-KI-VO nur die Informationen aus Anhang VIII Abschnitt C Nr. 1–3 umfassen. Die bereits auf eine Zusammenfassung reduzierte GRFA ist im dortigen Anhang jedoch unter Nr. 4 zu finden und dementsprechend überhaupt nicht umfasst. Gleiches gilt für die DSFA die dort unter Nr. 5 genannt ist. Dies untergräbt die teilweise hohen Transparenz- und Dokumentationspflichten der GRFA in den hochsensiblen Bereichen Biometrie, Strafverfolgung sowie Migration, Asyl und Grenzkontrolle (Anhang III Nr. 1, 6 und 7) erheblich. Allerdings bestehen hier Spielräume, nicht sicherheitssensible Informationen im Interesse von Transparenz gegenüber der breiteren (Fach-)Öffentlichkeit dennoch zu veröffentlichen. Die mitgliedstaatliche Gesetzgebung kann dies konkretisieren und für die eigenen Behörden verbindlich regeln.

Die GRFA umfasst die Beschreibung der Verfahren, in denen das Hochrisiko-System verwendet werden soll, wobei dies im Einklang mit der Zweckbestimmung gesehen werden muss (Art. 27 Abs. 1 Buchst. a EU-KI-VO). Weiterhin müssen Angaben über den Zeitraum und die Häufigkeit des Einsatzes gemacht werden (lit. b). Hier müssen im Falle der biometrischen Gesichtserkennung auch die Anforderungen aus Art. 5 Abs. 2 UAbs. 2 und Art. 26 Abs. 10 EU-KI-VO beachtet werden.[15] Die von der Verwendung des Hochrisiko-Systems im spezifischen Kontext betroffenen Personen oder Personengruppen sind gem. Art. 27 Abs. 1 Buchst. c EU-KI-VO zu benennen. Hier sollten die Betreibenden sich zunächst selbst Klarheit über die möglichen Einsatzfelder des jeweiligen KI-Systems verschaffen und diese vollumfänglich berücksichtigen. Die Betriebsanleitungen der Anbietenden können herangezogen werden. Interessenvertretungen, die Zivilgesellschaft und unabhängige Sachverständige sollen miteinbezogen werden (vgl. auch Erwägungsgrund 96 EU-KI-VO). Anhand der identifizierten Personen(gruppen) sollen die potenziell grundrechtsrelevanten Schadensrisiken ermittelt und aufgelistet werden (lit. d). Die anzustellende Prognoseentscheidung kann sowohl in quantitativer als auch in qualitativer Hinsicht erfolgen, da zunächst keine Methodik vorgegeben ist. Allerdings wäre eine rein quantitative oder eine rein automatisierte Prognose aus grundrechtlicher Sicht nicht ausreichend,[16] da Grundrechtsverletzungen konkrete qualitative Wertungsfragen darstellen. Die in lit. e geforderte menschliche Aufsicht sollte durch die Betreibenden unter Berücksichtigung der in Art. 14 EU-KI-VO niedergelegten Standards konkretisiert werden.

Die GRFA muss konkrete Abhilfemaßnahmen und Beschwerdemechanismen für den Fall enthalten, dass sich die beschriebenen Risiken tatsächlich verwirklichen (lit. f). Dabei ist festzulegen, wie im Fall einer Risikoverwirklichung konkret zu handeln ist. Dies kann u.a. technische-organisatorische Vorkehrungen umfassen, was in vielen Fällen die automatische Abschaltung des Systems sein kann. Im Kern geht es darum, eine effektive *Governance*-Struktur zu schaffen und über diese auch Auskunft zu geben.[17] Insbesondere hinsichtlich der spezifischen *Black-Box*-Problematik bei der KI-Nutzung durch Strafverfolgungsbehörden ist dies interessant. Die (teilweise) berechtigten Geheimhaltungsinteressen müssen hier auf den Prüfstand gestellt werden. Hinsichtlich der geforderten Beschwerdemechanismen legt die EU-KI-VO nicht fest, ob es sich für die Strafverfolgungsbehörden lediglich um die bestehenden Mechanismen (Art. 85 und 86 EU-KI-VO), also um die bestehende *Governance*-Struktur, handelt oder ob zusätzliche Mechanismen eingeführt werden.

Die GRFA hat insbesondere Dokumentationscharakter. Um einen wirksamen Schutz zu gewährleisten, ist es unerlässlich, eine umfassende und diverse grundrechtliche Expertise in den Prozess einzubinden, insbesondere zivilgesellschaftliche und unabhängige interdisziplinäre Akteur:innen. Diese Expertise sollte über den gesamten Lebenszyklus von KI-Systemen hinweg bereitstehen, und die GRFA sollte iterativ durchgeführt werden. Es sollten sowohl individuelle als auch kollektive und gesellschaftliche Risiken erfasst werden – auch hier kann die mitgliedstaatliche Gesetzgebung die Anforderungen ergänzen.

Eine wirksame GRFA erfordert eine klare Orientierung an menschenrechtlichen Standards. Weiterhin bedarf es umfassender Transparenz der Ergebnisse sowie einer frühzeitigen und kontinuierlichen Beteiligung betroffener Stakeholder. Die Einhaltung von Transparenzstandards sollte dabei nicht von den jeweils handelnden Akteur:innen abhängen, sondern bereits im technischen Design des jeweiligen KI-Systems angelegt sein.[18] Die Einschränkungen von Transparenzanforderungen für die Strafverfolgung sollten überdacht und auf das für Sicherheitsbelange zwingend erforderliche Mindestmaß begrenzt werden. Das Ziel besteht darin, potenzielle Grundrechtsverletzungen präventiv zu erkennen und zu verhindern. Dabei ist die Bewertung von Notwendigkeit und Verhältnis-

13 Siehe dazu *Kleemann* & *Aden* 2025.
14 *Eisenberger* 2025, Art. 27, Rn. 9.
15 Siehe hierzu auch *Kleemann* & *Aden* 2025.
16 So auch *Garcia-Godinez* 2025.
17 Siehe zu diesem Absatz *Eisenberger* 2025, Art. 27, Rn. 35–48.
18 Näher hierzu *Aden* & *Fährmann* 2020; *Aden* 2025.

mäßigkeit des KI-Einsatzes von zentraler Bedeutung.[19] Für die Methodik der GRFA können Ausarbeitungen wie das HUDERIA-Modell des Europarats[20] oder methodische Konzepte aus der Wissenschaft[21] eine Orientierung bieten.

II. KI-Grundrechte-Folgenabschätzung und Datenschutz-Folgenabschätzung

Die GRFA kann auf eine bereits erfolgte Datenschutz-Folgenabschätzung (DSFA) gem. Art. 35 DSGVO aufbauen (Art. 27 Abs. 4 EU-KI-VO). Für die polizeiliche Datenverarbeitung folgt die Pflicht zur Durchführung einer DSFA aus Art. 27 JI-Richtlinie. Da polizeiliche Hochrisiko-KI-Systeme in aller Regel personenbezogene Daten verarbeiten, kann hier stets von einer Überschneidung mit der DSFA ausgegangen werden. GRFA und DSFA weisen strukturelle Ähnlichkeiten auf, so dass sich die Koordination von Durchführung und Ergebnissen anbietet: Beide verfolgen einen risikobasierten und grundrechtsbezogenen Ansatz und sind *vor* der Inbetriebnahme von Systemen (*ex ante*) auf der Basis von fachlicher Expertise durchzuführen.[22]

Die DSFA hat allerdings einen engeren Fokus und setzt die Verarbeitung personenbezogener Daten voraus, womit sie im Schwerpunkt auf den Schutz der Grundrechte auf Privatheit und Datenschutz (Art. 7 und 8 EU-GRCh) abstellt. Die GRFA gem. Art. 27 EU-KI-VO umfasst dagegen alle Grundrechtsrisiken, die mit dem Einsatz KI-basierter Systeme einhergehen. Beide Folgenabschätzungen verfolgen somit unterschiedliche Ziele, können und sollen aber miteinander koordiniert werden – auch im Hinblick auf die Aufsichtsbefugnisse (Art. 77 EU-KI-VO). GRFA und DFSA sind komplementär zueinander und müssen beide durchgeführt werden. Art. 35 Abs. 11 DSGVO, implizit ebenso Art. 27 Abs. 2 JI-Richtlinie, verpflichtet die für die Datenverarbeitung verantwortliche Stelle zur Überprüfung und Bewertung der Umsetzung der DSFA-Ergebnisse – ein Standard der auch für die GRFA nach der EU-KI-VO übernommen werden sollte.

III. Fazit und Ausblick

Die EU-KI-VO stellt einen paradigmatischen Wandel in der Regulierung algorithmischer Systeme dar. Insbesondere für Polizei und Strafverfolgung zeigt sich die Herausforderung, klassische rechtsstaatliche Prinzipien wie Transparenz, Rechenschaft und effektiven Rechtsschutz in ein komplexes, soziotechnisches System zu integrieren. Dieser Beitrag hat verdeutlicht, dass die EU-KI-VO zwar wichtige erste Schritte unternimmt – etwa durch die Einführung spezifischer Pflichten zur Grundrechte-Folgenabschätzung, zur technischen Dokumentation und zur Einrichtung von *Governance*-Strukturen. Diese sind jedoch in ihrer praktischen Umsetzung mit grundlegenden strukturellen Defiziten konfrontiert.

Die doppelte *Black-Box*-Problematik von technisch intransparenter KI und organisatorisch undurchsichtiger Behördenpraxis verdeutlicht, dass *Accountability* nicht allein durch prozedurale Anforderungen gewährleistet werden kann. Vielmehr bedarf es eines ganzheitlichen Verständnisses von Verantwortlichkeit, das nicht nur normative und technische, sondern auch institutionelle und kulturelle Aspekte berücksichtigt.[23]

Für die Umsetzung der EU-KI-VO ergibt sich daraus ein doppelter Handlungsbedarf: Einerseits müssen die normativen und verfahrensrechtlichen Grundlagen – insbesondere die Rechte Betroffener und deren Durchsetzbarkeit – weiter konkretisiert und operationalisiert werden. Andererseits bedarf es einer verstärkten interdisziplinären Zusammenarbeit zwischen KI-Entwicklung, Gesetzgebung, Wissenschaft, Technik-Expert:innen, Behörden und zivilgesellschaftlichen Akteur:innen, um der soziotechnischen Komplexität algorithmischer Systeme gerecht zu werden. Zentral wird dabei sein, nicht nur eine formale, sondern eine funktionale Rechenschaftspflicht zu etablieren – entlang des gesamten Lebenszyklus eines KI-Systems und unter Einbezug aller relevanten Akteur:innen. Nur durch eine solche systemische Perspektive kann gewährleistet werden, dass KI-Systeme im sicherheitspolitischen Kontext nicht nur effektiv und effizient, sondern auch legitim und grundrechtskonform eingesetzt werden.

Die GRFA nach der EU-KI-VO bietet Chancen die polizeiliche Nutzung von KI-Systemen grundrechtskonform auszugestalten und deren Fehleranfälligkeit präventiv zu reduzieren. Die Regelungen der EU-KI-VO gelten auch für Polizei und Strafverfolgung, lassen jedoch noch viele Fragen offen und bedürfen daher der weiteren Konkretisierung. Die mitgliedstaatliche Gesetzgebung und die behördliche Umsetzungspraxis sollten die aufgezeigten Spielräume für die unabhängige Durchführung der GRFA durch externe Stellen und für ihre effektive und transparente Ausgestaltung nutzen.

Literatur

Aden, Hartmut (2025) Die polizeiliche Nutzung neuer Technologien zwischen Wollen, Sollen, Können und Dürfen. In: Axel Dessecker & Martin Rettenberger (Hrsg.), *Die Zukunft der Kriminalität und ihrer Kontrolle,* (Reihe Kriminologie und Praxis, Band 1), Wiesbaden: Kriminologische Zentralstelle 2025, S. 91–102; open access: https://www.krimz.de/fileadmin/dateiablage/E-Publikationen/KuP-Online/KuP-online1.pdf.

Aden, Hartmut & *Fährmann, Jan* (2020) Datenschutz-Folgenabschätzung und Transparenzdefizite der Techniknutzung. Eine Untersuchung am Beispiel der polizeilichen Datenverarbeitungstechnologie, *TATuP*, 29(3), 24–29, https://doi.org/10.14512/tatup.29.3.24.

Bertaina, Samuele; Biganzoli Ilaria; Desiante, Rachele; Fontanella, Dario; Inverardi, Nicole; Penco, Ilaria Giuseppina & Cosentini, Andrea Claudio (2025) Fundamental rights and artificial intelligence impact assessment: A new quantitative methodology in the upcoming era of AI Act. *Computer Law & Security Review* 56, 106101, https://doi.org/10.1016/j.clsr.2024.106101.

Council of Europe, Committee on Artificial Intelligence (2024) *Methodology for the risk and impact assessment of artificial intelligence systems from the point of view of human rights, democracy and the rule of law (HUDEIRIA Mehtodology)*. CAI(2024)16rev2. Strasbourg: Concil of Europe. https://rm.coe.int/cai-2024-16rev2-methodology-for-the-risk-and-impact-assessment-of-arti/1680b2a09f.

Eisenberger, Iris (2024) Kommentierung zu Art. 27 EU-KI-VO. In: Martini, Mario & Wendehorst, Christiane (Hrsg.), *KI-VO. Verordnung über Künstliche Intelligenz*. München: C.H. Beck, S. 571–585.

European Network of National Human Rights Institutions, ENNHRI (2025) ENNHRI calls on the European Commission to ensure effective Fundamental Rights Impact Assessments (FRIAs) under the EU AI Act, 7 April 2025 https://ennhri.org/wp-content/uploads/2025/04/ENNHRI-statement-on-ensuring-effective-Fundamental-Rights-Impact-Assessments-FRIAs-under-the-EU-AI-Act.pdf.

Garcia-Godinez, Miguel (2025) A default framework for fundamental rights impact assessment. *AI Ethics* 5, 5149–5163, https://doi.org/10.1007/s43681-025-00761-1.

19 Siehe generell dazu ENNHRI 2025.
20 Council of Europe 2024.
21 *Mantelero* 2024; *Bertaina* et al 2025.
22 Näher hierzu *Mantelero* 2024; *Thomaidou* & *Limniotis* 2025.
23 *Martinez*, *Tahraoui* & *Kleemann* 2026.

Kleemann, Steven & *Aden, Hartmut* (2025) Die Nutzung Künstlicher Intelligenz durch Strafverfolgungsbehörden – »Hintertüren« der Verordnung der Europäischen Union über Künstliche Intelligenz. In: Wilfried Honekamp, Stefanie Kemme & Jens Struck (Hrsg.) *Auswirkungen von Künstlicher Intelligenz auf die zukünftige Polizeiarbeit. Technologische Potenziale, rechtliche Rahmenbedingungen, kriminologisch-sozialwissenschaftliche Erkenntnisse,* Wiesbaden: Springer, S. 3–30.

Mantelero, Allessandro (2024) The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template. *Computer Law & Security Review* 54, 106020, https://doi.org/10.1016/j.clsr.2024.106020.

Martinez Demarco, Sol, Tahraoui, Milan & *Kleemann, Steven* (2026) The siloed logic and the implementation of the Artificial Intelligence Act in the law enforcement context: legal and ethical analysis of the applicability of accountability and responsibility to high-risk AI systems. In: Marie Eneman, Diana Miranda, Jan Ljungberg & William Webster (Hrsg.). *AI Surveillance and the Future of Policing: Opportunities, Threats, Perspectives and Cases*. London: Routledge, Routledge Studies in Surveillance, i.E.

Schuett Jonas (2024) Risk Management in the Artificial Intelligence Act. *European Journal of Risk Regulation*. 15(2), 367–385. doi:10.1017/err.2023.1.

Thomaidou, Anna & Limniotis, Konstantinos (2025) Navigating Through Human Rights in AI: Exploring the Interplay Between GDPR and Fundamental Rights Impact Assessment. *Journal of Cybersecurity and Privacy, 5*(1), 7, https://doi.org/10.3390/jcp5010007.

Einsatz von Künstlicher Intelligenz zur Klassifikation von Hassdelikten im Internet

von Dr. Robert Pelzer, Berlin, Florian Ludwig und Dr. Martin Hofer, München*

Das Forschungsprojekt KISTRA (2020–2023) untersuchte die Möglichkeiten des ethisch und rechtlich vertretbaren Einsatzes Künstlicher Intelligenz zur Früherkennung und Bekämpfung von Hasskriminalität im digitalen Raum. Zentrales Ziel war die Entwicklung von KI-basierten Textklassifizierern zur automatisierten Analyse großer Textmengen im Internet. Neben technischen Aspekten wurden auch rechtliche und gesellschaftlich-ethische Rahmenbedingungen umfassend analysiert. Dieser Werkstattbericht fasst zentrale Projektergebnisse und methodische Erkenntnisse zusammen.

I. Einleitung

Die Digitalisierung unserer Gesellschaft schreitet unaufhaltsam voran und bringt neben zahlreichen Chancen auch neue Herausforderungen für die öffentliche Sicherheit mit sich. Mit der gestiegenen Bedeutung des digitalen Raums für Radikalisierungsprozesse und die Verbreitung von Hass und Hetze, rückt auch die Bekämpfung von Hasskriminalität im Internet stärker in den Fokus der Strafverfolgungsbehörden. Eine zentrale Herausforderung bei der Verfolgung von Online-Hasskriminalität ist dabei, umfangreich vorliegende Inhalte in Form von Postings oder Kommentaren nach ihrer strafrechtlichen Relevanz zu bewerten. Hier setzte das Forschungsprojekt »Einsatz von KI zur Früherkennung von Straftaten« (KISTRA) an, das von 2020–2023 vom BMFTR im Rahmen des Programms »Forschung für die zivile Sicherheit« der Bundesregierung gefördert wurde. Gesamtziel von KISTRA war die Erforschung der Möglichkeiten und Rahmenbedingungen für den ethisch und rechtlich vertretbaren Einsatz von Systemen Künstlicher Intelligenz (KI) durch polizeiliche Endanwender*innen zur Erkennung, Vorbeugung und Verfolgung von Straftaten, insbesondere im Bereich der Hasskriminalität.[1] Dabei wurden mittels Verfahren maschinellen Lernens u.a. Textklassifizierer entwickelt und erprobt, um große Mengen an Textdaten automatisiert zu klassifizieren und relevante Inhalte für die weitere Auswertung vorzusortieren. Zudem erfolgte eine umfassende Analyse der rechtlichen und ethischen Rahmenbedingungen für einen Einsatz der entwickelten KI-Modelle in der polizeilichen Praxis.

Der folgende Werkstattbericht dokumentiert wesentliche Ergebnisse des technischen Entwicklungsprozesses sowie der gesellschaftlich-ethischen Anforderungsanalyse.

II. Anwendungsszenarien

Um die Strafverfolgung im digitalen Raum zu verbessern hat der Gesetzgeber mit dem am 30.03.2021 erlassenen Gesetz zur Bekämpfung des Rechtsextremismus und der Hasskriminalität das Netzwerkdurchsetzungsgesetz (NetzDG) novelliert. In diesem Zuge wurden Telemediendienstanbieter (TMDA) verpflichtet, strafrechtlich relevante Inhalte dem BKA zu melden. Das BKA hat hierzu die Zentrale Meldestelle für strafbare Inhalte im Internet (ZMI BKA) eingerichtet. Um die erwarteten massenhaften Meldungen effizient verarbeiten und eine zeitnahe Strafverfolgung zu ermöglichen, sollten in KISTRA Klassifizierer zur Bewertung der strafrechtlichen Relevanz der Meldungen entwickelt werden. Nach erfolgreicher Klage der Plattformbetreiber wurde die Meldepflicht jedoch ausgesetzt. Mittlerweile wurde das NetzDG durch den europäischen Digital Service Act ersetzt.[2] Der Schwerpunkt des Projektes wurde daher auf die Unterstützung von Aus-

* *Dr. Robert Pelzer* ist Kriminologe und Soziologe und leitet den Forschungsbereich Sicherheit – Risiko – Kriminologie am Zentrum Technik und Gesellschaft der TU Berlin. *Florian Ludwig* und *Dr. Martin Hofer* arbeiten im Forschungsreferat des Geschäftsfelds Big Data Analyse der Zentrale Stelle für Informationstechnik im Sicherheitsbereich (ZITiS).

1 An dem Projekt beteiligten sich das Bundeskriminalamt, die Zentrale Stelle für Informationstechnik im Sicherheitsbereich (ZITiS), die Juniorprofessur für Kriminologie, Strafrecht und Sicherheitsforschung im digitalen Zeitalter der Johannes Gutenberg-Universität Mainz, das Zentrum Technik und Gesellschaft der Technischen Universität Berlin, das Institut für Kommunikationswissenschaft und Medienforschung der Ludwig-Maximilians-Universität München, Munich Innovation Labs GmbH, die RWTH Universität Aachen, der Fachbereich Informatik der Technischen Universität Darmstadt sowie das Center of Advanced Technology for Assisted Learning and Predictive Analytics der Fernuniversität Hagen.

2 Die Übermittlungspflichten nach dem novellierten NetzDG erklärte das Verwaltungsgericht Köln im März 2022 für unionsrechtswidrig, weshalb die ZMI BKA derzeit überwiegend Meldungen von Kooperationspartnern aus der Zivilgesellschaft erhält. Das Digitale-Dienste-Gesetz (DDG) als Umsetzung mehrere EU-Richtlinien, das große Teile des NetzDG ablöst, enthält demgegenüber keine umfassenden Meldepflichten für strafbare Äußerungsdelikte.

wertetätigkeiten im Zusammenhang mit Hasskriminalität im Polizeilichen Staatsschutz gelegt. Anwendungsszenarien hier waren die Auswertung größerer Datenmengen, bspw. aus Internetauswertungen im Bereich PMK (Politisch Motivierte Kriminalität)-rechts nach Hinweisen auf Äußerungsdelikte. So hat das BKA bspw. eine Reihe von Aktionstagen gegen Hasskriminalität durchgeführt, in deren Zuge öffentliche Telegram-Gruppen gezielt nach strafbaren Äußerungsdelikten durchsucht wurden.

III. Technische Entwicklung

Zu Projektbeginn wurde ein Textdatensatz zu Hasskriminalität erstellt und nach einem im Projekt entworfenen Annotationsschema[3] annotiert. Der Datensatz deckt unter anderem strafrechtliche Relevanz, Phänomenbereiche der Politisch Motivierten Kriminalität (PMK) und Themenfelder der Hasskriminalität ab. Die Analyse des Datensatzes ergab, dass einige Bereiche deutlich mehr Datenpunkte enthielten als andere. Bspw. kamen Postings, die als strafrechtlich relevant nach § 130 StGB (Volksverhetzung) eingestuft wurden, deutlich häufiger vor als Postings, die nach § 241 StGB (Bedrohung) eingestuft wurden. Aufgrund dieser unausgeglichenen Datenlage lag der Fokus bei der Modellentwicklung zur Klassifikation von strafrechtlicher Relevanz auf § 130 und § 140 StGB (Belohnung und Billigung von Straftaten) sowie einem Modell zur allgemeinen Klassifikation strafrechtlicher Relevanz. Der Fokus bei der Klassifikation der Phänomenbereiche lag aufgrund der Datenlage auf der PMK-rechts. Die Themenfelder der Hasskriminalität umfassten die Klassen »*antisemitischer Hass*«, »*fremdenfeindlicher Hass*«, »*allgemeiner Hass*« sowie »*kein Hass*«.

Vor der Klassifikation durch die Modelle wurden die Texte in mehreren Schritten vorverarbeitet, unter anderem durch Normalisierung von Kodierungen, Korrektur von Großschreibung und Entfernung irrelevanter Inhalte. Zur Klassifikation wurden vortrainierte Transformer-Modelle (z.B. XLM-RoBERTa) mittels Transfer Learning angepasst. Um die Qualität unserer KI-Modelle zu bewerten, wurden drei zentrale Kennzahlen verwendet:

- **Precision (Treffsicherheit):** Die Precision zeigt, wie zuverlässig ein Modell zwischen relevanten und irrelevanten Inhalten unterscheidet – also wie viele der als relevant eingestuften Texte tatsächlich korrekt klassifiziert wurden. Die Precision liegt zwischen den Werten 0 (sehr schlecht) und 1 (optimal). Unsere Modelle lieferten dabei folgende Ergebnisse: Bei der allgemeinen strafrechtlichen Relevanz sowie der Klassifikation nach § 130 StGB lag die Precision jeweils bei **0,81**. Das Modell für § 140 StGB erreichte eine Precision von **0,73**. Auch im Bereich der Phänomen-Zuordnung zeigte das Modell solide Leistungen: Für den Phänomenbereich »*Rechts*« lag die Precision ebenfalls bei **0,73**. Besonders gute Ergebnisse erzielten die Modelle in den Themenfeldern: **0,87** für *antisemitischen Hass,* **0,95** für *»fremdenfeindlichen Hass«* und **0,87** für *»allgemeinen Hass«*.
- **Recall (Vollständigkeit):** Der Recall beschreibt, wie viele der tatsächlich relevanten Inhalte vom Modell korrekt erkannt wurden – also wie vollständig das System arbeitet. Auch der Wert des Recalls liegt zwischen den Werten 0 (sehr schlecht) und 1 (optimal). Hier zeigten die finalen Modelle folgende Leistung: Für die allgemeine strafrechtliche Relevanz lag der Recall bei **0,82**, für § 130 StGB bei **0,80** und für § 140 StGB bei **0,75**. Für den Phänomenbereich »*Rechts*« konnte das Modell **82 %** der relevanten Inhalte zuverlässig identifizieren. In den Themenfeldern zeigte sich ein ähnliches Bild: *Antisemitischer Hass* wurde mit einem Recall von **0,82**, *fremdenfeindlicher Hass* mit **0,95** und *allgemeiner Hass* mit **0,86** erkannt.
- **F1-Score (Gesamtgüte):** Der F1-Score kombiniert die beiden zentralen Kennzahlen – Precision und Recall – zu einem Wert, der das Gleichgewicht zwischen Treffgenauigkeit und Vollständigkeit abbildet. Wie die Werte von Precision und Recall liegt der Wert des F1-Scores zwischen 0 (sehr schlecht) und 1 (optimal). Vorgabe im Projekt war es, einen F1-Score von mindestens **0,70** zu erzielen. Dieses Ziel wurde in allen Bereichen übertroffen: Die Modelle zur Klassifikation *allgemeiner strafrechtlichen Relevanz* erreichten einen F1-Score von **0,81**, für *§ 130 StGB* lag der Wert bei **0,80**, und für *§ 140 StGB* bei **0,74**. Auch im Bereich der Ideologie *»Rechts«* wurde mit einem F1-Score von **0,78** ein überzeugendes Ergebnis erzielt. Besonders hohe Werte wurden bei der Klassifikation der Themenfelder der Hasskriminalität erzielt. *Antisemitischer Hass* wurde mit einem F1-Score von **0,85**, *fremdenfeindlicher Hass* mit **0,95** und *allgemeiner Hass* mit **0,87** erkannt.

Insgesamt war ein klarer Zusammenhang zwischen der Anzahl der zur Verfügung stehenden Trainingsdaten und der erzielten Modellgüte erkennbar. Die Ergebnisse erfüllten die Zielvorgaben des Projekts und zeigen, dass die entwickelten Modelle relevante Inhalte erkennen können. Ihre Einsatzfähigkeit in der Praxis hängt jedoch stark vom Anwendungsfall ab: In sicherheitskritischen Bereichen, in denen keine relevanten Inhalte übersehen werden dürfen (z.B. Bedrohungen), ist die aktuelle Zuverlässigkeit nicht ausreichend. Auch der Zieldatensatz spielt eine entscheidende Rolle. Die Modelle verlieren deutlich an Leistung, wenn sie auf Themen, Zielgruppen oder Sprachstile angewendet werden, die stark vom Trainingsdatensatz abweichen. In solchen Fällen wäre ein Nachtraining der Modelle notwendig.

Da das Trainieren von Modellen eine große Menge an Trainingsdaten erfordert und die Annotation von geeigneten Daten aufwendig ist, haben wir im Rahmen des Projekts erforscht, wie die Anpassung der Modelle an neue Domänen, z.B. Zielgruppen der Hassrede, annotationseffizient gestaltet werden kann. Dazu haben wir verschiedene Methoden aus dem Bereich Domainadaptation untersucht und deren Eignung für dieses Problem miteinander verglichen. Wir haben unsere Forschungsergebnisse in Form einer Fachpublikation veröffentlicht (Ludwig et al., 2022). Insgesamt sind die im Projekt KISTRA entstandenen Modelle als Proof of Concept zu verstehen. Sie zeigen das Potenzial dieser Methoden auf, sind aber nicht ohne ständige Aktualisierung und Wartung für den produktiven Einsatz geeignet.

3 Ein Annotationsschema ist eine definierte Anleitung zur systematischen Markierung von Trainingsdaten für maschinelles Lernen.

IV. Gesellschaftlich-ethische Anforderungen an den KI-Einsatz

Die polizeiliche Nutzung von KI zur Auswertung großer Datenmengen im Internet ist mit komplexen rechtlichen und ethischen Herausforderungen verbunden. Für die Entwicklung vertrauenswürdiger KI-Systeme existieren zahlreiche Leitlinien, die rechtmäßiges, ethisches und robustes Handeln fordern.[4] Dabei haben sich international sieben zentrale ethische Prinzipien etabliert: Legitimität, Effektivität, Wahrung der Freiheitsrechte, Fairness, Transparenz, Autonomie und Rechenschaftspflicht (Accountability). Im Projekt KISTRA wurden diese Prinzipien in konkrete Anforderungen für die KI-Unterstützung bei polizeilichen Auswertungen im Bereich PMK-rechts übertragen (Pelzer et al. 2025). Im Folgenden werden die sieben Kernprinzipien knapp umrissen und daraus sich ergebene Anforderungen beispielhaft aufgeführt.

Das Kernprinzip **Legitimität** fordert, dass die Entwicklung und Anwendung von KI rechtmäßig und auf klare Zwecke begrenzt erfolgt. Die Legitimität ist nicht von Dauer, sondern muss von Fall zu Fall überprüft werden. Sie erfordert eine transparente rechtliche Grundlage sowie die Konformität mit Werten einer demokratischen Gesellschaft wie Freiheit, Gerechtigkeit, Sicherheit und Fairness. Daraus ergibt sich v.a. die Anforderung, dass die Ziele und Zwecke eines KI-Modells klar definiert sein müssen. Dies beinhaltet etwa eine Umgrenzung des phänomenbezogenen Gegenstandsbereiches, der mithilfe eines KI-Systems ausgewertet werden soll. Legitimität erfordert zudem, dass Rechtsgrundlagen für den Einsatz des KI-Systems vorhanden sind. Bei der polizeilichen Nutzung von KI-Systemen sind diverse rechtliche Grundlagen zu berücksichtigen. Die grundrechtliche Ausgangslage bilden die Menschenwürdegarantie, das Recht auf informationelle Selbstbestimmung und Diskriminierungsverbote (Kühne & Golla, 2023). Neben den eingriffsrechtlichen Ermächtigungsgrundlagen (Strafprozessordnung, Polizei- und Ordnungsrecht) sind allgemeine datenschutzrechtliche Grundlagen (JI-Richtlinie, BDSG) zu berücksichtigen (ebd.). Eine weitere Rechtsebene bildet der EU AI Act. So müssen KI-Systeme, die nach Art. 6 AI Act sog. Hochrisiko-KI-Systeme darstellen u.a. bestimmten Qualitätsanforderungen genügen (vgl. Wendt und Wendt, 2024). Hochrisiko-Systems wären bspw. Systeme zur Erstellung von personenbezogenen Kriminalitätsprognosen oder Risikoeinschätzungen (siehe Annex III Nr. 6 zum AI Act).

Effektivität fordert, dass KI-Systeme einen Mehrwert für die Strafverfolgung und Gefahrenabwehr bieten. Sie müssen robust gegen Angriffe und Manipulationen sein und auch bei Störungen zuverlässig funktionieren. Eine hohe Zuverlässigkeit des KI-Systems, die wiederum eine hohe Daten- und Modellqualität erfordert, ist Voraussetzung für die Effektivität. Daraus ergeben sich Anforderungen an die Modellentwicklung wie Qualitätssicherung bei Verfahren der Annotation, Protokollierung der Herkunft der Datenquellen und eine geeignete Vielfalt und Variabilität der Daten. Erforderlich ist darüber hinaus eine Prüfung der inhaltlichen Eignung des KI-Systems für die vorgesehenen Auswertungszwecke. So ist zu klären, ob der Phänomenbereich des Modells mit dem Phänomenbereich der Auswertung kongruent ist und ob die Fehlerraten des Modells für den jeweiligen Auswertungszweck toleriert werden können.

Ein weiteres zentrales Prinzip betrifft die Wahrung der **Freiheitsrechte**. Die im Grundgesetz verankerten Rechte, insb. das durch Auswertung öffentlich verfügbarer Daten tangierte Recht auf informationelle Selbstbestimmung, müssen gewahrt werden. Staatliche Eingriffe müssen auf einen möglichst kleinen Kreis stichhaltig Verdächtiger begrenzt und stets verhältnismäßig sein. Für die polizeiliche Auswertungspraxis ergibt sich daraus v.a. die Anforderung einer Verhältnismäßigkeitsprüfung. Hierzu gilt es darzulegen, inwieweit der vorgesehene Umfang der Datenerhebung/-auswertung für die Verfolgung der Auswertungsziele erforderlich ist. Zu bewerten ist darüber hinaus die Intensität des Eingriffs in Grundrechte (zu berücksichtigen sind hier etwa die Streubreite des Eingriffs, die Persönlichkeitsrelevanz der mit KI ausgewerteten Daten und mögliche nachteilige Folgen und Risiken für Betroffene). Schließlich ist im Rahmen der Angemessenheitsprüfung das Gewicht der in dem zu erreichenden Zweck enthaltenden Rechtsgüter mit der Schwere der Eingriffstiefe gegenüberzustellen.

Das Prinzip der **Fairness** fordert, dass KI-Systeme keine Diskriminierung aufgrund von Geschlecht, ethnischer Herkunft, Alter, Religion oder sozioökonomischem Status vornehmen dürfen. Verzerrungen in den Trainingsdaten oder Modellen sowie Vorurteile der Anwender*innen müssen frühzeitig erkannt und beseitigt werden. Es gilt daher, die Fairness des Modells anhand geeigneter Validierungsdaten und auf Grundlage spezifizierter Fairnesskriterien zu prüfen und anhand der Ergebnisse geeignete Maßnahmen zur Erhöhung der Fairness, wie z.B. eine Erweiterung des Trainingsdatensatzes, zu treffen.

Hinsichtlich der **Transparenz** geht es einerseits um die Nachvollziehbarkeit der Funktionsweise des Modells sowie um die Erklärbarkeit der KI-Ergebnisse für die Anwender*innen, anderseits um die Transparenz der Nutzung von KI-Systemen gegenüber der Öffentlichkeit, die entscheidend für die Akzeptanz und das Vertrauen der Bevölkerung in die Polizei ist. Mit Blick auf die Anwender*innen ist v.a. wichtig, dass Herkunft der Trainingsdaten, Trainingsparameter, Modellarchitektur, Modellgüte und Teststatistiken ausreichend dokumentiert sind und alle den Modellentscheidungen zugrunde liegenden Faktoren offengelegt werden. Zudem sollten die Ergebnisse der KI-Anwendung für Nutzer*innen durch geeignete visuelle und textliche Hinweise nachvollziehbar (»erklärbar«) sein.

Ein weiteres zentrales ethisches Prinzip bildet die **Autonomie**: Die Anwender*innen müssen in der Lage sein, die Ergebnisse der KI kritisch zu bewerten und eigenständige Entscheidungen zu treffen. Die Gefahr des »automation bias«, also der unkritischen Übernahme von KI-Ergebnissen, muss durch geeignete technische und organisatorische Maßnahmen minimiert werden. Dabei ist v.a. auch wichtig, dass

4 Z.B. die »Ethics Guidelines For Trustworthy AI« der von der Europäischen Kommission eingesetzten »High-Level Expert Group on Artifical Intelligence« (HLG-AI, 2019) oder der »KI-Prüfkatalog: Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz« der Fraunhofer IAIS (Fraunhofer IAIS, 2022). Auch die am 01.08.2024 in Kraft getretene KI-Verordnung der Europäischen Union (EU), die darauf abzielt eine vertrauensvolle Verwendung und Entwicklung von KI in der EU zu gewährleisten, ist hier von Relevanz.

Nutzer*innen die Grenzen des Modells richtig einschätzen und interpretieren können, wozu es neben Schulungen v.a. Modell-Evaluationen auf Grundlage von für den Produktivbetrieb repräsentativen Daten bedarf.

Schließlich fordert die **Rechenschaftspflicht** (Accountability), dass alle Arbeitsschritte und Entscheidungen im Rahmen von KI-gestützten Datenerhebungen einer verantwortlichen Person zugeschrieben werden kann. Daraus ergibt sich die Anforderung einer lückenlosen Dokumentation aller menschlichen und technischen Entscheidungen im gesamten Auswertungsprozess, die entsprechenden Aufsichtsstellen zugänglich gemacht werden müssen.

V. Lessons Learned aus dem Projekt KISTRA

Die Erfahrungen im Projekt KISTRA zeigen, dass der Einsatz von KI zur Unterstützung bei der Erkennung und Bewertung strafrechtlich relevanter Hassrede technisch machbar ist – aber nur unter der Voraussetzung klarer rechtlicher, ethischer und organisatorischer Rahmenbedingungen. KI kann Polizei und Sicherheitsbehörden sinnvoll entlasten, ersetzt aber nicht den Menschen und erfordert realistische Erwartungen an ihre Fähigkeiten.

Zentrale Lessons Learned:

1. **Datenqualität bestimmt Modellgüte – Quantität und Qualität als kritische Faktoren**: Die Performance von KI-Modellen hängt maßgeblich von der Qualität, Vielfalt und Ausgewogenheit der Trainingsdaten ab – sowohl technisch als auch im Hinblick auf Fairness und Diskriminierungsfreiheit. Dabei zeigt sich in der Praxis, dass die Beschaffung geeigneter Trainingsdaten schwierig ist: Im Projekt KISTRA konnten trotz intensiver Bemühungen nur für drei Klassen strafbarer Äußerungsdelikte Modelle trainiert werden. Die Qualität der Annotation ist ebenso entscheidend – sie erfordert Expertise, da die Bewertung von Hasskriminalität oft subjektiv und interpretationsabhängig ist.
2. **KI als Unterstützung, nicht als Entscheidungsträger:** Die entwickelten Modelle zeigen Potenzial, doch ihr Einsatz muss sich klar auf unterstützende Funktionen beschränken. Menschliche Entscheidungshoheit und die Ausbildung einer kritischen Bewertungskompetenz von KI-Ergebnissen bei Anwender*innen zur Vermeidung von »automation bias« sind essenziell – besonders in sicherheitskritischen Kontexten. Hierfür braucht es neben KI-Schulungen v.a. auch transparente »Beipackzettel« mit Performance-Werten aus Validierungsstudien mit für die polizeiliche Auswertungspraxis repräsentativen Daten.
3. **Begrenzte Übertragbarkeit und notwendiges Nachtraining von KI-Modellen:** Die Modelle zeigen deutliche Leistungseinbußen, wenn sie auf Themen, Sprachstile oder Nutzer*innengruppen angewendet werden, die im Trainingsdatensatz unterrepräsentiert waren. Ein Nachtraining (»Feintuning«) mit passenden Daten ist in jedem neuen Einsatzszenario zwingend erforderlich.
4. **Rechtliche und ethische Anforderungen:** Für das Anwendungsszenario in KISTRA – die Auswertung von Online-Daten nach strafrechtlicher Relevanz – lässt sich ein rechtlich und ethisch konformer Einsatz von KI grundsätzlich gewährleisten. Ob die rechtlichen Grundlagen für die Anwendung von KI-Modellen in einem Auswertevorgang ausreichend sind, ob der Zweck legitim, das KI-Modell geeignet und fair genug ist und ob die Grundrechtseingriffe verhältnismäßig sind, muss jedoch für den jeweiligen Einzelfall geprüft werden. Empfehlenswert ist daher die Implementierung eines »Legality/Ethics by Design«-Ansatzes in der Softwareentwicklung, welcher einen anwenderorientierten, rechtlich sowie ethisch kontrollierten Workflow durch alle Prozessschritte KI-gestützter Datenauswertungen hinweg ermöglicht.

Literatur

Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS. (2022). KI-Prüfkatalog: Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz. https://www.iais.fraunhofer.de/ki-pruefkatalog. Zugegriffen: 10.06.2024.

High-Level Expert Group on Artificial Intelligence (HLG-AI). (2019). Ethics Guidelines for Trustworthy AI. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Zugegriffen: 10.06.2024.

Kühne, M., & Golla, S. (2023). Rechtliche Rahmenbedingungen von KISTRA. Gutachten. Un-veröffentlichtes Manuskript.

Ludwig, F., Dolos, K., Zesch, T., & Hobley, E. (2022). Improving generalization of hate speech detection systems to novel target groups via domain adaptation. In *Proceedings of the Sixth Workshop on Online Abuse and Harms (WOAH)* (pp. 29–39).

Pelzer, R., Hahne, M., Taing, S. (2005): KI-Tools zur Unterstützung polizeilicher Internetauswertungen: rechtskonforme, ethische Entwicklung und Implementierung. In: Honekamp, W./Kemme, S. (Hg.): Auswirkungen von KI auf die zukünftige Polizeiarbeit. Wiesbaden: Springer VS *(im Erscheinen)*

Wendt, J. & Wendt, D. H. (Hrsg.). (2024). Artificial Intelligence Act: Gesetz über Künstliche Intelligenz (1. Auflage). Nomos; Manz Verlag Wien; Helbing & Lichtenhahn. https://permalink.obvsg.at/AC17313200.

Gesichts-Morphing als Sicherheitsrisiko: Angriffsvektoren und Detektionsmethoden

von Dr.-Ing. Clemens Seibold, Berlin*

Ein übermäßiges Vertrauen in Gesichtsbilder zur Identifikation oder Verifikation kann durch sogenannte Gesichtsmorphs zu einem Sicherheitsrisiko werden. Bei einem Gesichtsmorph handelt es sich um ein computergeneriertes Bild, welches durch die Verschmelzung von den Gesichtsbildern zweier unterschiedlicher Personen entsteht und beiden Ausgangspersonen täuschend ähnlich sieht. Dieser Artikel beschreibt, wie solche Bilder entstehen, welche Gefahren solch ein Bild an der falschen Stelle mit sich bringen kann, insbesondere im Kontext von biometrischen Systemen zur Identifikation und Verifikation, und welche technischen Möglichkeiten zu ihrer Erkennung und organisatorischen Möglichkeiten zu ihrer Vermeidung existieren.

I. Einleitung

Unter Morphing versteht man einen Spezialeffekt der Bildbearbeitung, bei dem ein fließender Übergang zwischen zwei Bildern in Form einer Animation erzeugt wird. Diese Technik kommt bereits seit Jahrzehnten in der Film- und Unterhaltungsindustrie zum Einsatz. Schon 2004 wurde auf einer biometrischen Fachtagung darauf hingewiesen, dass ein einzelnes Zwischenbild einer solchen Morphing-Sequenz, welche aus den Gesichtsbildern zweier unterschiedlicher Personen erzeugt wurde, Merkmale beider Gesichter enthält und womöglich biometrische Systeme gezielt täuschen kann (Lewis und Statham 2004). Ein solches Bild wird als Gesichtsmorph bezeichnet. Zehn Jahre später belegte eine Gruppe von Forschern diese Vermutung empirisch (Ferrara et al. 2014). Sie zeigten, dass Gesichtsmorphs beiden Ausgangspersonen so stark ähneln, dass selbst kommerzielle Gesichtserkennungssysteme eine Übereinstimmung beider Ausgangspersonen mit dem Gesichtsmorph attestieren.

II. Technische Beschreibung zur Erstellung einer Morphinganimation

Zur Erstellung eines Gesichtsmorphs bzw. einer Morphing-Animation stehen zahlreiche kostenlose Tools zur Verfügung. Die meisten davon basieren auf ähnlichen Grundprinzipien, wie sie im Folgenden beschrieben werden.

Eine Morphing-Animation wird schrittweise erzeugt, indem einzelne Zwischenbilder generiert werden. Jedes dieser Zwischenbilder entsteht in einem mehrstufigen Verfahren: Zunächst erfolgt eine geometrische Verzerrung beider Ausgangsbilder, welche je nach Fortschritt der Animation unterschiedlich stark ist. Zu Beginn der Animation wird das Zielbild stark verzerrt, während das Ausgangsbild kaum verändert wird. Gegen Ende verhält es sich umgekehrt. Ziel dieser Verzerrung ist es, markante Objekte in beiden Bildern geometrisch anzugleichen.

Im nächsten Schritt werden die Bilder überlagert. Auch hier richtet sich die Gewichtung nach dem Stand der Animation. Zu Beginn dominiert das Ausgangsbild, während das Zielbild nur schwach durchscheint. Gegen Ende ist es auch hier genau umgekehrt.

Grundlage der geometrischen Anpassung sind paarweise Referenzpunkte in beiden Bildern. Jeder Punkt im ersten Bild hat ein entsprechendes Pendant im zweiten Bild. Diese Punkte legen fest, wie die Bilder verzerrt werden, sodass bspw. der Punkt, der am Anfang an der Position im Startbild liegt, sich im Verlauf der Animation allmählich zur Position im Zielbild verschiebt. In der Regel werden Referenzpunkte an auffälligen und biometrisch relevanten Stellen gesetzt, etwa an Augen, Mund, Nase und Augenbrauen. Die Pixel, welche nicht direkt über Referenzpunkte definiert sind, bewegen sich bei der geometrischen Anpassung ähnlich den nächst gelegenen Referenzpunkten. Eine genaue Platzierung dieser Punkte ist entscheidend, um sichtbare Bildfehler oder unrealistische Übergänge zu vermeiden.

Neben den klassischen Methoden existieren mittlerweile auch KI-basierte Verfahren zur Erstellung von Gesichtsmorphs, welche auf Techniken der generativen künstlichen Intelligenz beruhen (Zhang et al. 2021). Diese ermöglichen es, Morphs auch ohne exakte Referenzpunkte zu erzeugen. Allerdings sind solche Systeme in der Regel weniger intuitiv bedienbar und die biometrische Ähnlichkeit zwischen dem somit erzeugten Bild und den Ausgangsbildern ist nicht immer gegeben.

* Der Verfasser ist z.Zt. wissenschaftlicher Mitarbeiter am Fraunhofer-Institut für Nachrichtentechnik und am Institut für Informatik an der Humboldt-Universität zu Berlin.

Abb. 1: Morphing eines Gesichts in eine cartoonisierte Variante. In jedem Schritt verändern sich sowohl die Gesichtsgeometrie als auch der Bildinhalt wie Hautfarbe und Haarstruktur.

Abbildung 1 zeigt exemplarisch das Morphing eines Gesichts in eine cartoonisierte Variante, um insbesondere den Verzerrungsprozess innerhalb der Animation visuell nachvollziehbar zu machen. Dabei wurde eine klassische Morphing-Methode verwendet.

III. Gesichts-Morphing-Angriff

1. Angriff auf hoheitliche Dokumente

Neben ihrer Nutzung in Unterhaltungsanwendungen können Gesichtsmorphs auch für kriminelle Zwecke missbraucht werden. Ein solches Angriffsszenario wurde unter anderem von *Ferrara et al.* (2014) in ihrer Studie zur Verwundbarkeit kommerzieller biometrischer Gesichtserkennungssysteme beschrieben:

Eine gesuchte Person, die bspw. auf einer internationalen Fahndungsliste steht, möchte das Land verlassen. Dazu erstellt sie gemeinsam mit einer zweiten, nicht gesuchten Person einen Gesichtsmorph, der Merkmale beider Personen enthält. Die zweite Person beantragt daraufhin mit diesem Bild einen Reisepass auf ihren eigenen Namen. Da der Gesichtsmorph beiden Beteiligten ähnlich genug sieht, ist die Beantragung erfolgreich und die gesuchte Person kann später unter der Identität der anderen Person reisen.

Die niedrige Einstiegshürde macht diesen Angriff besonders gefährlich. Im Internet stehen zahlreiche frei verfügbare Programme und Tutorials zur Verfügung, die auch Laien die Erstellung von Gesichtsmorphs ermöglichen. Bis vor Kurzem konnte in Deutschland das Gesichtsbild für Ausweisdokumente vom Antragstellenden als ausgedrucktes Foto selbst mitgebracht werden, wodurch es kaum technische Hürden für einen solchen Angriff gab. Ein prominentes Beispiel für die praktische Durchführbarkeit lieferte das Künstlerkollektiv Peng!, dessen dokumentierter Morphingangriff breite mediale Aufmerksamkeit erhielt (Thelen und Horchert 2018).

Seit Mitte 2025 ist ein solches Vorgehen jedoch erheblich erschwert, da für deutsche Reisepässe und Personalausweise Passbilder nur noch vor Ort in der Behörde aufgenommen oder ausschließlich von zertifizierten Stellen digital übermittelt werden dürfen. Damit wurde zumindest teilweise einer EU-Verordnung Rechnung getragen, die empfiehlt »biometrische Identifikatoren, insbesondere das Gesichtsbild, durch die nationalen Behörden, die Personalausweise ausstellen, vor Ort erfassen zu lassen« (Europäisches Parlament 2019).

Doch ist das Problem damit endgültig gelöst? Nicht vollständig.

Der Angriff wurde durch den Wegfall des Ausdruckens und Einscannens von Passbildern deutlich erschwert, bleibt aber dennoch weiterhin umsetzbar, insbesondere wenn das Gesichtsbild unbeaufsichtigt erfasst wird. Auch wenn sichergestellt wird, dass das Bild technisch unverändert direkt aus der Kamera an die Behörde übermittelt wird, lässt sich durch einen Workaround ein Bild von einem Gesichtsmorph in einen Reisepass bringen. Dazu wird der Gesichtsmorph einfach der Kamera präsentiert. Dies kann bspw. in Form eines Ausdrucks oder sogar einer Maske geschehen. Das Foto wäre zwar echt, aber dennoch wäre darauf kein reales Gesicht abgebildet, sondern weiterhin ein Gesichtsmorph. Eine unbeaufsichtigte Erfassung biometrischer Merkmale sollte folglich in allen Fällen vermieden werden, um sicherzustellen, dass das aufgenommene biometrische Merkmal tatsächlich zur antragstellenden Person gehört. Theoretisch ist die Beantragung eines Ausweisdokuments mit einem Gesichtsmorph also auch heute noch möglich, jedoch nur unter deutlich erhöhtem Aufwand.

Für Reisepässe, die vor Inkrafttreten der neuen Regelung ausgestellt wurden, bleibt dieses Problem jedoch bestehen. Sie sind weiterhin gültig und wurden möglicherweise mit eingereichten manipulierten Papierfotos erstellt. Somit könnten uns morphingbasierte Identitätsbetrugsfälle mit diesen Dokumenten noch über Jahre hinweg beschäftigen. Darüber hinaus ist das Einreichen eigener Passbilder in mehreren Staaten weiterhin gängige Praxis. Solange dies der Fall ist, bleibt das Gesichts-Morphing-Risiko auch international relevant.

2. Weitere Angriffsszenarien

Die Möglichkeiten von Morphing-Angriffen beschränken sich nicht auf hoheitliche Dokumente. Grundsätzlich kann jedes personenbezogene Dokument, bei dem die Identität über ein Gesichtsbild überprüft wird, Ziel eines solchen Angriffs werden. Somit können sich bspw. auch auf Mitgliedskarten in Fitnessstudios, Fahrkarten, Krankenkassenkarten oder anderen Karten für Zutrittsberechtigungen gemorphte Gesichtsbilder befinden. In allen Fällen, in denen das eingereichte Gesichtsbild nicht unter Aufsicht erstellt oder geprüft wird und die Verantwortung für das Bild bei der sich für einen Service registrierenden Person liegt, ist ein Morphing-Angriff nicht auszuschließen.

Neben der Manipulation von Identitätskarten können Gesichtsmorphs auch anderweitig gezielt gegen gesichtsbasierte Identifikations- und Authentifikationssysteme eingesetzt werden. Ein denkbares Szenario besteht darin, dass ein in einem Zugangskontrollsystem gespeichertes biometrisches Template[1] durch das Template eines Gesichtsmorphs ersetzt wird. In einem solchen Fall könnte sowohl die angreifende Person als auch weiterhin die berechtigte Person die Zugangskontrolle passieren. Somit würde der Austausch des Templates nicht auffallen.

Auch in Datenbanken oder auf überwachten Social Media-Plattformen können Gesichtsmorphs problematisch sein. So könnte eine unbeteiligte Person, deren Gesicht für sie unbewusst Teil eines Gesichtsmorphs ist, auf einem manipulierten Bild einer Straftat abgebildet werden. Die Person könnte damit unrechtmäßig mit der Straftat in Verbindung gebracht werden, während die Ähnlichkeit zu der ursprünglichen Person reduziert wird. Dies erschwert eine 1-zu-N-Suche, z.B. bei einem Abgleich der abgebildeten Person mit Fahndungsdatenbanken, und kann dazu führen, dass die gesuchte Person nicht erkannt wird.

Der kriminellen Kreativität sind kaum Grenzen gesetzt und es existieren viele weitere potenzielle Angriffsszenarien, bei denen gezielte Manipulation biometrischer Merkmale eine zentrale Rolle spielt.

1 Biometrische Systeme speichern i.d.R. nicht die aufgenommenen Rohdaten, sondern eine abgeleitete, kodierte und meist nicht invertierbare Repräsentation (biometrisches Template).

Nicht nur Gesichter, sondern auch andere biometrische Merkmale können auf ähnliche Weise angegriffen werden. So können auch synthetische Fingerabdrücke erzeugt werden, die den Fingerabdrücken zweier oder mehrerer realer Personen ähnlich sind (Ferrara et al. 2023). Da Fingerabdrücke in der Regel nicht wie Gesichtsbilder selber aufgenommen und abgegeben werden, ist dieser Angriff deutlich komplexer. Der Fingerabdruck muss entweder physisch erstellt oder anderweitig im System eingeschleust werden. In der Fachliteratur ist ein derartiges Vorgehen bislang deutlich seltener beschrieben, stellt aber dennoch mehr als eine rein theoretische Möglichkeit dar.

3. Schutz vor Morphing-Angriffen

Den wirksamsten Schutz vor Morphing-Angriffen bieten Systeme, die biometrische Merkmale ausschließlich unter Aufsicht und durch vertrauenswürdige Stellen erfassen, wie es bei der Aufnahme von Fingerabdrücken im Rahmen der Beantragung eines Reisepasses üblich ist.

In der Praxis ist dieses Vorgehen jedoch nicht immer realisierbar. Politische oder logistische Gründe können gegen eine beaufsichtigte Erfassung sprechen. Besonders bei Verfahren mit rein digitaler Antragstellung, wie z.B. der Beantragung einer elektronischen Gesundheitskarte, ist es technisch häufig nur möglich, dass Antragstellende ihr Gesichtsbild eigenständig hochladen oder es mithilfe bereitgestellter Software selbst aufnehmen.

Auch in Fällen, in denen Bilder aus fremden Quellen stammen, z.B. aus öffentlich zugänglichen Datenbanken oder sozialen Netzwerken, lässt sich die Authentizität eines Bildes meist nicht sicher feststellen. Nur in Ausnahmefällen kann die Echtheit zweifelsfrei garantiert werden.

Daher sind bei Gesichtsbildern, dessen Aufnahmeprozess nicht von einer vertrauenswürdigen Stelle überwacht wurde, zusätzliche Maßnahmen zu empfehlen, um ihre Echtheit zu prüfen und mögliche Manipulationen frühzeitig zu erkennen.

Im Folgenden werden verschiedene technische Verfahren aus der aktuellen Forschung vorgestellt, die darauf spezialisiert sind, morphingbedingte Veränderungen in Gesichtsaufnahmen zu erkennen und somit eine wichtige Rolle beim Schutz biometrischer Systeme spielen können.

IV. Detektionsmöglichkeiten von manipulierten Bildern

Zur Erkennung von Gesichtsmorphs wurden verschiedene Verfahren entwickelt. Diese wurden teils für diese Problematik neu konzipiert und teils durch die Anpassung bestehender Methoden aus der Medienforensik realisiert. Die Verfahren lassen sich grob in zwei Kategorien unterteilen:

- Single-Image-Based Morphing Attack Detection (S-MAD) Verfahren: Diese Ansätze benötigen ausschließlich das zu prüfende Bild, um eine Aussage über dessen Echtheit zu treffen.
- Differential Morphing Attack Detection (D-MAD) Verfahren: Diese Methoden setzen zusätzlich ein Vergleichsbild derselben Person voraus, bspw. ein vor Ort aufgenommenes Bild beim Grenzübertritt am eGate. Da in solchen Anwendungen ohnehin ein aktuelles Referenzbild erfasst wird, sind D-MAD-Verfahren häufig technisch gut integrierbar.

1. Kompressionsartefakte

Die S-MAD-Verfahren basieren auf forensischen Analysen spezifischer Bildeigenschaften. Zum Beispiel auf der Untersuchung von Kompressionsartefakten, die bei JPEG-Bildern entstehen. Die JPEG-Kompression teilt ein Bild in 8×8-Pixel große Blöcke ein, transformiert jeden Block in einen Frequenzbereich und quantisiert die resultierenden Koeffizienten der einzelnen Frequenzen. Natürliche Bilder weisen dabei bestimmte statistische Verteilungen auf. Werden Bilder jedoch mehrfach komprimiert, wie es beim Zusammensetzen eines Gesichtsmorphs häufig geschieht, verändert sich diese Verteilung. Gesichtsmorphs lassen sich auf diese Weise von Originalbildern unterscheiden, da ihre Kompressionssignaturen statistisch auffällig sind (Makrushin et al. 2018).

2. Bildrauschen

Eine weitere Möglichkeit zur Erkennung manipulierter Bilder basiert auf der Analyse charakteristischer Merkmale digitaler Kamerasensoren, insbesondere des Bildrauschens, das bei der Aufnahme entsteht. Bildrauschen enthält sowohl zufällige Anteile, z.B. temperaturabhängige, als auch deterministische Komponenten. Jeder einzelne Pixel eines Kamerasensors reagiert leicht unterschiedlich auf identische Lichtverhältnisse. Einige Pixel erzeugen systematisch dunklere oder hellere Werte, selbst wenn sie den gleichen Bildinhalt erfassen. Diese individuellen Reaktionsmuster, auch Sensor Pattern Noise genannt, können so charakteristisch sein, dass sie als eine Art digitaler Fingerabdruck der Kamera dienen. Über einen Vergleich dieser Muster lässt sich sogar prüfen, ob ein Bild mit einem bestimmten Gerät aufgenommen wurde (Chen et al. 2008). Beim Erstellen eines Gesichtsmorphs verändern sich diese feinen Signaturen.

Um mittels einer Analyse des Sensorrauschens Gesichtsmorphs zu erkennen, wird das zu prüfende Bild zunächst in verschiedene Regionen unterteilt. Anschließend werden die lokalen Rauschmuster extrahiert und auf Konsistenz, statistische Auffälligkeiten und kameratypische Merkmale hin untersucht. Abweichungen vom erwartbaren Muster können auf ein zusammengesetztes bzw. manipuliertes Bild hinweisen (Debiasi et al. 2018).

3. Allgemeine Bildqualität

Eine weitere Gruppe von Verfahren innerhalb der S-MAD-Kategorie basiert auf der Analyse allgemeiner Bildqualitätsmerkmale sowie texturbasierter Eigenschaften. Da Gesichtsmorphs üblicherweise durch die Überlagerung zweier Bilder entstehen, sind diese meist unschärfer als die verwendeten Eingangsbilder. Diese Unschärfe kann ein Indiz für Manipulation sein. Allerdings ist die isolierte Bewertung eines Bildes bezüglich seiner Unschärfe ohne Zugriff auf ein Referenzbild nicht zuverlässig möglich. Daher arbeiten diese Verfahren mit modellierten Annahmen über die Schärfenverteilung von unverfälschten Gesichtsbildern. Diese Annahmen werden entweder explizit formuliert, etwa über bildstatistische Schärfemodelle, oder durch Methoden des maschinellen Lernens implizit aus Trainingsdaten abgeleitet. In beiden Fällen wird versucht typische Muster und Abweichungen zu erkennen, wie sie in manipulierten Bildern auftreten (Neubert et al. 2019 und Venkatesh et al. 2020).

4. Künstliche Neuronale Netze

Eine weitere Möglichkeit zur Erkennung von Gesichtsmorphs ist der Einsatz künstlicher neuronaler Netze (KNNs). Diese datengetriebenen Methoden nutzen komplexe Architekturen moderner Deep-Learning-Konzepte, um charakteristische Muster in Gesichtsmorphs zu erkennen. Dabei können sie zu einem gewissen Teil die zuvor beschriebenen forensischen Analyseansätze implizit nachbilden und darüber hinaus zusätzliche, nicht unmittelbar offensichtliche Merkmale erlernen, die auf eine Bildmanipulation hinweisen. Für das Training dieser Art von Detektoren werden jedoch wesentlich mehr Beispiele an echten Gesichtsbildern und Gesichtsmorphs benötigt, als bei den zuvor beschriebenen Verfahren (Seibold et al. 2017).

Als entscheidender Nachteil dieser Methode wird oft die geringe Interpretierbarkeit genannt. Ohne zusätzliche Analysemethoden bleibt unklar, warum ein KNN ein bestimmtes Bild als echt oder als Gesichtsmorph klassifiziert. Um Einblicke in die Entscheidungsfindung eines KNN zu erhalten, existieren jedoch verschiedene Techniken, mit denen sich Bildregionen identifizieren lassen, die für die Klassifikation relevant sind.

Eine dieser Methoden ist Layer-wise Relevance Propagation (LRP) (Bach et al. 2015). Mit ihrer Hilfe kann jedem einzelnen Pixel ein sogenannter Relevanzwert zugewiesen werden, der angibt, in welchem Maße dieser Pixel zur Klassifikation des KNN beigetragen hat.

Als klassisches Beispiel für die Anwendung von LRP und zum Vergleich mit anderen Erklärbarkeitsverfahren, wird häufig ein Klassifikationsnetz herangezogen, das auf die Unterscheidung verschiedener Objektkategorien trainiert wurde, z.B. verschiedene Tiere, Fahrzeuge und Gebäude. Dabei ist jeder Kategorie ein Neuron in der letzten Schicht des KNN zugeordnet. Wird dem KNN bspw. ein Bild eines Schiffs präsentiert, führt dies zu einer Aktivierung des entsprechenden Neurons für die Klasse »Schiff«. LRP weist im ersten Schritt diesem Neuron, das eine Klasse repräsentiert, eine Startrelevanz zu. Diese wird anschließend Schicht für Schicht rückwärts durch das KNN bis hin zum Eingangsbild weitergegeben. Dabei gibt ein Neuron seine Relevanz an alle Neuronen weiter, die zu seiner Aktivierung geführt haben. Neuronen, die einen starken Beitrag zu der Aktivierung geleistet haben bekommen mehr Relevanz zugeteilt, als solche die nur leicht dazu beigetragen haben. Am Ende dieses Prozesses werden die Relevanzwerte von den Neuronen der Anfangsschicht in dem Netzwerk an die Pixel im Eingangsbild übertragen. Diese Relevanzverteilung lässt sich dann durch Falschfarben und gegebenenfalls als Overlay über das Eingangsbild visualisieren. Somit lassen sich die für die Klassifikation ausschlaggebenden Bildbereiche anzeigen.

Im Fall von Gesichtsmorphs steht das KNN vor der besonderen Herausforderung, ausschließlich mit Gesichtern konfrontiert zu sein. Der Unterschied zwischen zwei verschiedenen Gesichtern ist wesentlich größer als der zwischen einem Gesicht und einem auf diesem basierenden Gesichtsmorph. Um zwischen echten Bildern und Morphs zu unterscheiden, muss es subtile, auf den ersten Blick kaum wahrnehmbare Strukturen erkennen. Um dies zu realisieren lernt das KNN häufig komplexe Modelle, bei denen LRP an seine Grenzen stößt. Wendet man LRP auf ein KNN zur Erkennung von Gesichtsmorphs an, neigt LRP dazu jenen Regionen eine hohe Relevanz zuzuweisen, in denen typischerweise Fälschungsspuren durch Morphing zu finden sind. Dies geschieht unabhängig davon, ob diese in dem jeweils analysierten Bild tatsächlich vorliegen oder nicht. So können auch vollkommen authentische Bildbereiche irrtümlich als entscheidungsrelevant gekennzeichnet werden.

Ein Lösungsansatz für dieses Problem bietet die LRP-Erweiterung Focused Layer-wise Relevance Propagation (FLRP) (Seibold et al. 2021). Diese Erweiterung wurde speziell für Szenarien entworfen, in denen KNNs ähnlichen Strukturen, z.B. immer Gesichtern, ausgesetzt sind und sehr feine Unterschiede zwischen den Klassen lernen. Im Unterschied zu LRP beginnt FLRP nicht bei den Klassenneuronen in der letzten Schicht, sondern identifiziert im Inneren des KNN gezielt solche Neuronen, die empfindlich auf Manipulationsartefakte reagieren. Diese bekommen eine Startrelevanz zugewiesen, die dann genauso wie bei LRP schrittweise zum Eingabebild zurückgeführt wird.

Abb. 2: Partieller Gesichtsmorph (links) sowie Relevanzverteilungen von LRP (Mitte) und FLRP (rechts) als Overlay. In diesem Beispiel wurde nur die Nase verändert, der Rest des Bildes ist authentisch. Hell überlagerte Regionen deuten darauf hin, dass die jeweilige Erklärbarkeitsmethode diese Bereiche als relevant für die Klassifikation des Bildes als »Gesichtsmorph« identifiziert, was auf mögliche Fälschungsspuren hinweist. Nicht überlagerte Regionen sind für die Klassifikation durch das KNN nicht relevant. Der partielle Gesichtsmorph basiert auf Bildern aus dem Face Research Lab London Set von Lisa M. DeBruine & Benedict C. Jones: https://figshare.com/articles/dataset/5047666, lizenziert unter CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

Die gravierenden Unterschiede der Relevanzzuweisung von LRP und FLRP lassen sich exemplarisch an einem teilmanipulierten Gesichtsmorph verdeutlichen, bei dem lediglich die Nasenregion durch ein gemorphtes Teilbild ersetzt wurde, während das restliche Gesicht unverändert geblieben ist. Abbildung 2 zeigt die resultierenden Relevanzzuweisungen von beiden Methoden überlagert über das Eingangsbild. Während LRP fälschlicherweise auch der Augenregion eine hohe Relevanz zuweist, obwohl diese in diesem Fall nicht manipuliert wurde, konzentriert sich FLRP nahezu ausschließlich auf die gefälschte Nasenregion und liefert somit eine deutlich präzisere und plausiblere Erklärung für die Entscheidung des KNN.

5. D-MAD über Künstliche Neuronale Netze

Differential Morphing Attack Detection (D-MAD) Verfahren nutzen ein zusätzliches Referenzbild der auf dem zu prüfenden Bild abgebildeten Person, um eine Aussage über dessen Echtheit treffen zu können. Ein Großteil dieser Verfahren basiert ebenfalls auf Methoden der künstlichen Intelligenz. Dabei wird das KI-basierte System meist darauf trainiert, zwei Gesichtsaufnahmen miteinander zu vergleichen, um somit das zu prüfende Bild als echtes Bild oder Fälschung klassifizieren zu können (Scherhag et al. 2020). Für dieses spezielle Szenario existieren bislang jedoch keine etablierten Erklärbarkeitsverfahren. Insbesondere lassen sich Techniken wie LRP oder dessen Erweiterung FLRP nicht direkt anwenden, da sie typischerweise auf die Analyse eines einzelnen Bildes ausgerichtet sind und keine direkte Unterstützung für bildpaarbasierte Klassifikationen bieten.

6. Demorphing

Ein alternativer Ansatz innerhalb der D-MAD-Kategorie besteht im sogenannten Demorphing. Bei dieser Methode werden aus dem zu überprüfenden Bild die auf dem Referenzbild abgebildeten Gesichtsmerkmale rausgerechnet. Das Ziel besteht darin festzustellen, ob nach dem Entfernen dieser das Gesicht einer anderen Person sichtbar wird. Ist das resultierende Bild der Referenzperson weiterhin ähnlich, so handelt es sich mit hoher Wahrscheinlichkeit um ein authentisches Bild. Weicht das resultierende Gesicht hingegen stark ab, deutet dies auf ein Morphing hin. Demorphing kann sowohl mit klassischen Verfahren der Bildverarbeitung (Ferrara et al. 2018) als auch mit Techniken der künstlichen Intelligenz umgesetzt werden (Long et al. 2024).

V. Evaluation von Morphing-Detektoren

Für die Auswertung von Verfahren zur Erkennung von Gesichtsmorphs existiert bislang kein allgemein anerkannter Standarddatensatz, der sich innerhalb der wissenschaftlichen Community durchgesetzt hat. Zudem kann das jeweilige Anwendungsszenario einen erheblichen Einfluss auf die Auswahl geeigneter Erkennungsmethoden haben. So unterscheiden sich die Charakteristika von Gesichtsaufnahmen je nach Kontext erheblich. Bilder in Reisepässen sind in der Regel stark komprimiert, werden unter Umständen nach der Aufnahme zunächst ausgedruckt und anschließend für die digitale Verwendung erneut eingescannt. Dieser Prozess beeinträchtigt bspw. die Analyse kameratypischer Rauschmuster bzw. macht sie gänzlich unmöglich. Dagegen ist die Rauschanalyse ein durchaus vielversprechender Ansatz, um digital vom Antragstellenden eingereichte Gesichtsbilder auch Echtheit zu prüfen.

Um dennoch eine vergleichbare, standardisierte Bewertung von Morphing-Erkennungsverfahren zu ermöglichen, bietet das National Institute of Standards and Technology (NIST) der Vereinigten Staaten einen unabhängigen Benchmark an.[2] Forschende und Unternehmen können hier ihre Erkennungsverfahren zur Evaluation einreichen. Dabei muss jedoch ein Abstand von mindestens vier Monaten zwischen zwei Einreichungen derselben Institution bestehen. Die eingereichten Methoden werden dabei auf einer Vielzahl an Datensätzen getestet, die sowohl echte als auch gemorphte Gesichtsbilder enthalten. Ziel des Benchmarks ist es, eine möglichst realitätsnahe und repräsentative Bewertung sicherzustellen, gleichzeitig aber auch eine Überanpassung (Overfitting) an die Testdaten zu vermeiden. Aus diesem Grund bleiben die verwendeten Datensätze unveröffentlicht. Ein Teil der vom NIST genutzten gemorphten Bilder wurde jedoch von externen Forschungsgruppen beigesteuert. Die Auswertung der Einreichungen erfolgt getrennt nach dem jeweils verwendeten Morphing-Verfahren, was eine differenzierte Betrachtung der Erkennungsleistung in Abhängigkeit des möglichen Aufwands eines Angreifers erlaubt.

Da das NIST selbst keine eigenen Verfahren zur Erkennung entwickelt, ist eine gewisse Neutralität der Evaluation gewährleistet. Die Ergebnisse jedes eingereichten Verfahrens werden öffentlich zugänglich gemacht. Neben der verpflichtenden Angabe Ihrer Organisation können Teilnehmende optional eine kurze Verfahrensbeschreibung hinzufügen. Während nur wenige kommerzielle Anbieter diese Möglichkeit nutzen, bieten insbesondere Einsendungen aus Forschungseinrichtungen oft detailliertere Informationen, anhand derer sich die verwendete Methodik nachvollziehen lässt. Hierbei zeigt sich, dass insbesondere Verfahren, die auf künstlicher Intelligenz basieren, im Benchmark die besten Ergebnisse erzielen.

Gleichzeitig verdeutlicht der Benchmark jedoch auch die nach wie vor bestehenden Grenzen der existierenden Detektoren. Qualitativ hochwertige Gesichtsmorphs, die mit größerem manuellem Aufwand erstellt wurden, sind auch für moderne Erkennungssysteme schwer von echten Gesichtsbildern zu unterscheiden. Diese Erkenntnis unterstreicht die Notwendigkeit neben Detektionsverfahren auch die Prozesse zur Einreichung von Gesichtsbildern sicherer zu gestalten.

VI. Fazit

Gesichtsmorphs sind eine ernstzunehmende Bedrohung für die Vertrauenswürdigkeit in biometrische Systeme. Sie gefährden nicht nur hoheitliche Dokumente, wie Reisepässe und Personalausweise, sondern können auch in der Privatwirtschaft erheblichen Schaden anrichten.

2 https://pages.nist.gov/frvt/html/frvt_morph.html (Zugegriffen am 25.06.2025).

Für einen umfassenden Schutz empfiehlt es sich Gesichtsbilder für biometrische Referenzdaten nur in überwachten Szenarien aufzunehmen, um das Einschleusen von verfälschten Daten zu vermeiden. Ist eine überwachte Aufnahme nicht möglich, kann auf unterschiedliche Verfahren zur Erkennung von Gesichtsmorphs zurückgegriffen werden. Während klassische bildforensische Verfahren auf Artefakte wie Kompressionsspuren oder Rauschmuster setzen, erlauben moderne Verfahren auf Basis Künstlicher Neuronaler Netze eine datengetriebene Analyse der Bildinhalte. Doch gerade bei diesen leistungsfähigen KI-basierten Methoden fehlt es bislang oft an ausreichender Transparenz für die Klassifikation, z.B. durch das Hervorheben oder Beschreiben von Fälschungsspuren anhand derer ein Gesichtsbild als Morph eingestuft wurde. Techniken wie Focused LRP zeigen jedoch vielversprechende Ansätze zur Verbesserung der Transparenz solcher Systeme.

Letztlich wird der Schutz vor Morphing-Angriffen nicht allein durch Technologie entschieden. Ein Bewusstsein für die Risiken auf Seiten der Betreiber biometrischer Systeme und eine leichte Skepsis gegenüber Daten aus nicht prüfbaren Quellen sind ebenso relevant.

Literatur

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R. & Samek, W. (2015). On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. PLoS ONE 10(7): e0130140

Chen, M., Fridrich, J., Goljan, M. & Lukas, J. (2008), Determining Image Origin and Integrity Using Sensor Noise, in IEEE Transactions on Information Forensics and Security, vol. 3, no. 1, pp. 74–90, March 2008

Debiasi, L., Scherhag, U., Rathgeb, C., Uhl, A. & Busch, C. (2018), PRNU-based detection of morphed face images, 2018 International Workshop on Biometrics and Forensics (IWBF), Sassari, Italy, 2018, pp. 1–7

Europäisches Parlament (2019) Verordnung (EU) 2019/1157 des Europäischen Parlaments und des Rates vom 20.06.2019 zur Erhöhung der Sicherheit der Personalausweise von Unionsbürgern und der Aufenthaltsdokumente, die Unionsbürgern und deren Familienangehörigen ausgestellt werden, die ihr Recht auf Freizügigkeit ausüben

Ferrara, M., Franco, A. & Maltoni, D. (2014). The Magic Passport. In IEEE International Joint Conference on Biometrics, 2014

Ferrara, M., Franco, A. & Maltoni, D. (2018). Face Demorphing in the Presence of Facial Appearance Variations. 2018 26th European Signal Processing Conference (EUSIPCO), pages 2365–2369, 2018.

Ferrara, M., Cappelli, R. & Maltoni, D. (2023). Detecting Double-Identity Fingerprint Attacks. IEEE Trans. on Biometrics, Behavior, and Identity Science, 5(4): 476–485, 2023

Lewis, M. & Statham, P. (2004). ESG Biometric Security Capabilities Programme: Method, Results and Research Challenges. In Proceedings Biometrics Consortium Conference (BCC), 2004

Long, M., Duan, Q., Zhang, L. -B., Peng, F. & Zhang, D. (2024). TransFD: Transformer-Based Representation Interaction for Face De-Morphing, in IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 6, no. 3, pp. 385–397, July 2024

Makrushin, A., Kraetzer, C., Neubert, T. & Dittmann J. (2018). Generalized Benford's Law for Blind Detection of Morphed Face Images. In Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec '18). Association for Computing Machinery, New York, NY, USA, 49–54.

Neubert, T., Kraetzer, C. & Dittmann, J. (2019). A Face Morphing Detection Concept with a Frequency and a Spatial Domain Feature Space for Images on eMRTD. In Proceedings of the ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec'19). Association for Computing Machinery, New York, NY, USA, 95–100.

Scherhag, U., Rathgeb, C., Merkle, J. & Busch, C. (2020) Deep Face Representations for Differential Morphing Attack Detection, in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3625–3639, 2020

Seibold, C., Samek, W., Hilsmann, A. & Eisert, P. (2017). Detection of Face Morphing Attacks by Deep Learning, International Workshop on Digital Watermarking (IWDW), Lecture Notes in Computer Science (LNCS) Springer, Cham, 2017, pp

Seibold, C., Hilsmann, A. & Eisert, P. (2021). Focused LRP: Explainable AI for Face Morphing Attack Detection, Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops, 2021, pp. 88–96

Thelen, R. & Horchert, J. (2018). Aktivisten schmuggeln Fotomontage in Reisepass. DER SPIEGEL, 03 2018

Venkatesh, S., Ramachandra, R., Raja, K. & Busch, C. (2020), Single Image Face Morphing Attack Detection Using Ensemble of Features, 2020 IEEE 23rd International Conference on Information Fusion (FUSION), Rustenburg, South Africa, 2020, pp. 1–6

Zhang, H., Venkatesh, S., Ramachandra, R., Raja, K. B., Damer, N. & Busch, C. (2021). MIPGAN – generating strong and high quality morphing attacks using identity prior driven gan. IEEE Transactions on Biometrics, Behavior, and Identity Science, PP: 1–1, 04 2021

Zwischen Technik und Vertrauen – Künstliche Intelligenz und Führung bei der Polizei

von Dr. Sylke Piéch, Berlin, Thomas Berger, Stuttgart und Raphael Humbel, Aargau/Schweiz*

Die Entwicklungen im Bereich der Künstlichen Intelligenz (KI) zählen zu den bedeutendsten Innovationstreibern unserer Zeit. Sie beeinflussen zunehmend, wie Führung gestaltet, Entscheidungen getroffen und Organisationen weiterentwickelt werden. Auch im öffentlichen Sektor, insbesondere bei der Polizei, eröffnen KI-Systeme vielfältige Möglichkeiten zur Effizienzsteigerung und Aufgabenunterstützung. Für Führungskräfte bedeutet dies, die technologischen Potenziale gezielt zu nutzen, Orientierung zu geben und die Kompetenzen im Team kontinuierlich weiterzuentwickeln.
Doch wie genau können KI-Systeme den Polizeialltag unterstützen? Welche Rolle spielen Führungskräfte in diesem digitalen Wandel und wie verändert sich Führung durch den Einsatz von KI-Technologien? Im Rahmen dieses Beitrags werden wir diese Fragen anhand konkreter Anwendungsfälle bei der Polizei Berlin und der Polizei Baden-Württemberg erörtern. Vertieft wird der Beitrag durch wissenschaftliche Erkenntnisse aus einer Master-Thesis an der Hochschule Luzern, die forschungsbasierte Perspektiven auf KI und Führung im Polizeikontext aufzeigt.

I. Die Rolle der Führungskraft im digitalen Wandel

Die Führungskräfte nehmen eine Schlüsselfunktion im Prozess der Digitalisierung und beim Einsatz von KI-Systemen ein. Von ihnen hängt es maßgeblich ab, ob der Technologietransfer in die Praxis gelingt (Piéch, 2024).

Führungskräfte stehen dabei vor folgenden Aufgaben:
- Bewertung und Auswahl vertrauenswürdiger KI-Systeme
- Analyse bestehender Arbeitsprozesse: Wo ist der konkrete KI-Einsatz sinnvoll?
- Klärung der Rahmenbedingungen und Verantwortlichkeiten
- Gewinnung der Mitarbeitenden für neue KI-Technologien
- Umgang mit persönlichen Ängsten und Einstellungen
- Transparente Kommunikation und Rollenklärung
- Begleitung der Implementierung und Bewertung der Wirksamkeit

Aktuelle Ergebnisse der Studie »KI & Entscheidungen« zeigen, dass die Akzeptanz von KI-Systemen maßgeblich durch die Führungsqualitäten der Vorgesetzten beeinflusst wird. Wird eine Führungskraft als kompetent in digitaler Führung wahrgenommen, steigt auch die Offenheit der Mitarbeitenden gegenüber dem Einsatz von KI-Tools. Gleichzeitig übernehmen Führungskräfte eine wichtige Vorbildfunktion. Wer digitale Anwendungen aktiv einsetzt, deren Chancen und Grenzen konkret aufzeigt und offen mit neuen Entwicklungen umgeht, motiviert die Mitarbeitenden, selbst digitale Erfahrungen zu sammeln. Ebenso entscheidend ist die Fähigkeit der Führungskräfte, ein gemeinsames Wir-Gefühl zu schaffen. Wer seine Führungskraft als identitätsstiftend erlebt, zeigt eine höhere Bereitschaft, Vertrauen in neue Technologien aufzubauen und die Zusammenarbeit mit KI-Systemen zu akzeptieren (Dieck et al., 2025).

II. Führung neu denken: Kompetenzen für eine digital geprägte Arbeitswelt

Der Einsatz von KI-Systemen verändert die Arbeit von Führungskräften grundlegend. Führungskräfte müssen in der Lage sein, Ergebnisse von KI-Systemen kritisch einzuordnen und transparent zu kommunizieren. Gleichzeitig verändern sich die Führungsprozesse in hybriden Teams, in denen Menschen und Maschinen zusammenarbeiten. Technologisches Grundverständnis, ethisches Verantwortungsbewusstsein und aktives Veränderungsmanagement gewinnen daher an Bedeutung. Führung bedeutet immer mehr, Vertrauen in KI aufzubauen, Mitarbeitende zu befähigen und die Integration neuer Systeme strategisch zu begleiten.

Laut dem Kompetenzmodell nach Humbel (vgl. Abbildung auf der Folgeseite) sind folgende Führungskompetenzen von zentraler Bedeutung.

Wie aus dem Modell ersichtlich wird, basiert erfolgreiche Führung im Kontext von KI auf einem ausgewogenen Zusammenspiel verschiedener Kompetenzen. Dazu zählen neben den KI- und Fachkompetenzen ebenso die Kommunikationsfähigkeit, Teamfähigkeit, Innovationskraft und Problemlösungsfähigkeit. Das Modell betont eine werteorientierte, menschenzentrierte Führungskultur. Trotz technischer Unterstützung bleibt Führung eine persönliche und verantwortungsvolle Aufgabe. Um den Anforderungen einer zunehmend digitalen und dynamischen Arbeitswelt gerecht zu werden, ist die kontinuierliche Weiterbildung von Führungskräften daher absolut entscheidend.

III. Digitalisierung und Künstliche Intelligenz bei der Polizei Berlin

»Für eine sichere Hauptstadt mit moderner digitaler Technik bedarf es finanzieller und personeller Mittel, die für eine zukunfts- und leistungsfähige Polizei unverzichtbar sind. Der Mensch steht im Mittelpunkt – die Digitalisierung unterstützt bei der Vielzahl an Aufgaben!«

Polizeidirektor Christian Schröder, Polizei Berlin – CDO

Die Polizei Berlin hat mit der Einrichtung der Position des Chief Digital Officer (CDO) und seinem Team einen strategischen Meilenstein gesetzt. Diese Organisationseinheit ist direkt bei der Behördenleitung angesiedelt und hat den Auftrag, KI- und Digitalisierungsprozesse ressortübergreifend zu koordinieren, zu priorisieren und eine effiziente Umsetzung zu begleiten. Damit werden Digitalisierung und Einsatz von

* *Dr. Sylke Piéch*, Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI,
E-Mail: sylke.piech@dfki.de, LinkedIn: linkedin.com/in/dr-sylke-piech
Polizeipräsident Thomas Berger, Polizei Baden-Württemberg,
E-Mail: stuttgart.ptls.praesident@polizei.bwl.de, linkedin.com/in/thomas-berger-17b4491bb
Oblt Raphael Humbel, Kantonspolizei Aargau/Schweiz,
E-Mail: raphael.humbel@kapo.ag.ch, LinkedIn: linkedin.com/in/raphael-humbel-62229929b

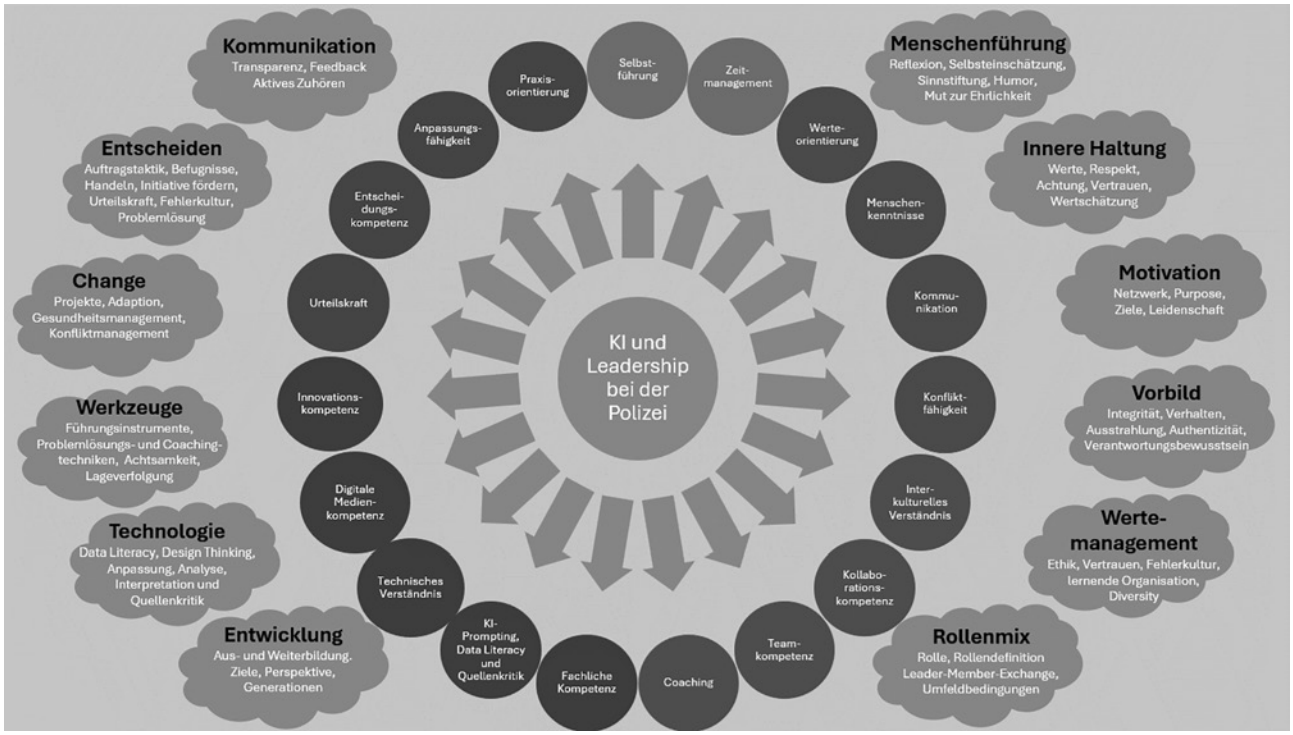


Abbildung: Humbel, R. (2025): Kompetenzmodell für KI und Leadership bei der Polizei

KI nicht als isolierte Technikprojekte, sondern als zentrale Führungsaufgaben verstanden, die das gesamte polizeiliche Handeln – von der Einsatzpraxis bis zur Verwaltung – betreffen.

3.1. Zentrale Handlungsfelder bei der Polizei Berlin

Ein wesentlicher Schwerpunkt liegt auf dem gezielten Einsatz digitaler Möglichkeiten bei operativen polizeilichen Aufgaben. Dabei steht im Fokus, Personalressourcen durch Automatisierung zu entlasten und sie gezielt für Kernaufgaben freizusetzen. Mithilfe von Automatisierung und intelligenten Assistenzsystemen sollen Vollzugskräfte bei repetitiven Aufgaben unterstützt werden, um sie gezielter für operative Kernaufgaben einsetzen zu können. Projekte wie die KI-basierte Videoüberwachung im Objektschutz verfolgen genau dieses Ziel. Hier wird das Potenzial gesehen, mehrere hundert Dienstkräfte durch technologische Lösungen zu entlasten.

Mit der App »INSITU« beginnt für Polizistinnen und Polizisten – vor allem die im Ersten Angriff tätig sind – ein neues Zeitalter: INSITU ist eine App, die eine vollständige Spur- und Sacherfassung am Tatort bzw. Einsatzort ermöglicht. Ziel ist es, jederzeit Antworten auf die zentralen Fragen zu haben: »Was wurde wann, wo, wie, von wem und warum gesichert?« Sie erleichtert die Arbeit vor Ort und setzt auf die Digitalisierung im Rahmen von Smart und Mobile Policing von Beginn an der polizeilichen Ermittlungsarbeit.

Ein weiterer Fokus liegt auf der Optimierung kriminalpolizeilicher Prozesse, insbesondere bei der Auswertung großer Datenmengen,-sogenannter »Massen- und Schmutzdaten«. Moderne Technologien ermöglichen eine effizientere Auswertung und strukturierte Aufbereitung dieser Daten sowie eine digitale Beweisführung bei deren Bearbeitung – in sämtlichen Deliktsbereichen. Ziel ist es, die Ermittlungsarbeit zu beschleunigen und qualitativ zu verbessern.

Auch die Modernisierung der Aus- und Fortbildung ist ein zentrales Anliegen. Die Entwicklung eines KI-basierten Lern- und Coaching-Systems für die Polizeiakademie soll Lernprozesse personalisieren und standortunabhängig Wissen vermitteln. Ergänzend dazu werden Wissens- und Kompetenzassistenten entwickelt, die in Echtzeit relevante Inhalte bereitstellen. Damit wird einerseits auf die Bedürfnisse einer zunehmend digital geprägten Generation von Nachwuchskräften reagiert und andererseits ein zukunftsfähiger Kompetenzaufbau in der Organisation verankert.

Ein Chatbot als digitales Auskunftssystem soll bspw. dabei helfen, wiederkehrende Anfragen schnell, komfortabel und direkt für den Auszubildenden zu beantworten, etwa zu internen Abläufen, Zuständigkeiten oder organisatorischen Prozessen. Des Weiteren ist vorgesehen, durch Chatbotfunktionen auch der Bevölkerung den Zugang zu häufig nachgefragten Informationen rund um die Uhr (24/7) zur Verfügung zu stellen. Dadurch werden sowohl der Service verbessert als auch Mitarbeitende und Auskunftsstellen spürbar entlastet.

3.2. KI-Systeme im Stabs- und Verwaltungsbereich der Polizei Berlin

Darüber hinaus werden administrative Tätigkeiten gezielt automatisiert, um die Bearbeitungszeiten zu verkürzen und die internen Abläufe zu standardisieren. Der Einsatz digitaler Tools, etwa bei Dienstreiseanträgen, der Dienstunfallbearbeitung oder der Vorgangsbearbeitung allgemein trägt dazu bei, Ressourcen deutlich effizienter einzusetzen, die Transparenz zu erhöhen und Verwaltungsprozesse deutlich zu beschleunigen. Es gilt dabei insbesondere Verzögerungen zu überwinden, die durch die Einbindung in landesweite Prozesse/ Standards entstehen

Zudem wird der Einsatz von Künstlicher Intelligenz im Stabs- und Verwaltungsbereich der Polizei Berlin intensiv

vorangetrieben. Anwendungen wie automatisierte Textanalysen, Aktenbewertungen, Vorschlagslogiken sowie Systeme zur Spracherkennung bei sicherheitsrelevanten Daten befinden sich bereits in der Entwicklung. Insbesondere in Bereichen wie der Gremienarbeit und dem Personalwesen sollen KI-basierte Werkzeuge künftig Verwaltungsprozesse effizienter und präziser gestalten. Dabei erfolgt die Einführung solcher Technologien stets unter strikter Berücksichtigung datenschutzrechtlicher, ethischer und sicherheitsrelevanter Anforderungen.

Der Erfolg der Digitalisierung hängt entscheidend vom Führungsverhalten sowie der aktiven Einbindung der Mitarbeitenden ab. Die Polizei Berlin verfolgt daher einen partizipativen Ansatz: Beschäftigte werden frühzeitig in Veränderungsprozesse einbezogen und ihre Bedürfnisse systematisch berücksichtigt. Um diesen Wandel transparent und wirksam zu begleiten, wird eine umfassende Kommunikationsstrategie entwickelt, die informierend und motivierend wirkt und auf eine breite Akzeptanz innerhalb und außerhalb der Organisation abzielt. Dabei spielt auch die Benutzerfreundlichkeit der Systeme eine zentrale Rolle: Die Anwendungen sollen intuitiv bedienbar sein und sich möglichst an privat genutzten Technologien orientieren, um die Akzeptanz bei den Nutzenden zu fördern und die Einführung neuer digitaler Werkzeuge nachhaltig zu unterstützen.

Die Digitalisierung der Polizei Berlin ist zentraler Bestandteil der gesamtbehördlichen Strategie und wird als gemeinsamer Beschluss auf (Spitzen-)Führungsebene miteinander konkret ausgestaltet. Neben den technischen, rechtlichen und sicherheitsrelevanten Aufgaben sind ebenso zielgerichtete Fortbildungsangebote für den Umgang mit KI, die Führung von Teams im digitalen Zeitalter und den Veränderungsprozessen systematisch erforderlich. Die Qualifizierung von Führungskräften ist wichtig, ihre Teams sicher, souverän und verantwortungsvoll durch den technologischen Wandel zu führen.

IV. Digitalisierung und Künstliche Intelligenz bei der Polizei Baden-Württemberg

Drei zentrale Themenkomplexe im Wandel der inneren Sicherheit:

Die Digitalisierung verändert nahezu alle Lebensbereiche – und macht auch vor der Polizei nicht Halt. Besonders der Einsatz von Künstlicher Intelligenz (KI) eröffnet neue Möglichkeiten in der Polizeiarbeit. Die Polizei Baden-Württemberg verfolgt dabei eine klare Strategie: Technologie muss dem Menschen dienen, nicht umgekehrt. Die Landespolizei setzt auf Innovation, bleibt aber verfassungsrechtlich und ethisch fest verankert. Drei Themenkomplexe stehen im Zentrum der aktuellen Diskussion und Umsetzung: der Schutz individueller Grundrechte durch »Human in the Loop«, die Transformation der Ermittlungsarbeit durch KI sowie der Wandel polizeilicher Führung in einer zunehmend interdisziplinären Sicherheitsarchitektur.

4.1. Human in the Loop – Verantwortung bleibt menschlich

Der Grundsatz »Human in the Loop« beschreibt einen entscheidenden Eckpfeiler beim Einsatz von KI-Systemen durch die Polizei: Bei allen grundrechtlich relevanten Entscheidungen bleibt der Mensch in der Verantwortung. KI kann analysieren, vorschlagen und unterstützen – doch sie darf nicht autonom über Freiheitseingriffe entscheiden. Ob Durchsuchung, Überwachung oder Freiheitsentziehung: Diese Maßnahmen sind nach wie vor vom Menschen zu prüfen, zu verantworten und – im Zweifel – vor Gericht zu vertreten.

Gerade in einem demokratischen Rechtsstaat, in dem jede hoheitliche Maßnahme überprüfbar sein muss, ist dieser Punkt zentral. Eine KI kann nicht zur Rechenschaft gezogen werden. Polizei, Justiz und Öffentlichkeit brauchen nachvollziehbare Entscheidungsprozesse. Die Integration von KI darf daher niemals zur Entpersonalisierung staatlichen Handelns führen. Vielmehr muss sie die Professionalität der handelnden Beamtinnen und Beamten stärken, indem sie datenbasierte Entscheidungsgrundlagen liefert – die Bewertung aber bleibt in menschlicher Hand.

4.2. Revolution der Ermittlungsarbeit – Datenflut intelligent bewältigen

Ein zweites zentrales Einsatzfeld von KI liegt in der polizeilichen Ermittlungsarbeit. Die Herausforderungen in diesem Bereich wachsen stetig: Bei Großverfahren, insbesondere in der organisierten Kriminalität oder im Bereich Cybercrime, müssen riesige Datenmengen gesichtet, sortiert und analysiert werden. Gleiches gilt für die Auswertung digitaler Beweismittel – etwa bei der Sichtung von Bildern und Videos nach Großveranstaltungen, Straftaten im öffentlichen Raum oder bei Kindesmissbrauchsdelikten.

Hier bietet KI enorme Chancen. Intelligente Algorithmen können Dokumente vorsortieren, Zusammenhänge erkennen, Muster aufzeigen und Hinweise auf relevante Inhalte liefern. Das Projekt KIRKE (Künstliche Intelligenz für Recherche, Klassifikation und Ermittlungsunterstützung) ist ein Beispiel dafür. Es soll Ermittlerinnen und Ermittlern helfen, schneller die wesentlichen Informationen aus umfangreichen Akten, Chats oder Videos zu extrahieren. Auch in der forensischen Dokumentenprüfung kommt KI bereits zum Einsatz – bspw. beim Erkennen gefälschter Ausweise oder Urkunden.

Diese Entwicklungen sind kein Selbstzweck. Sie sollen menschliche Ressourcen entlasten, Bearbeitungszeiten verkürzen und die Aufklärung von Straftaten verbessern. Entscheidend bleibt: Auch wenn KI bei der Bewertung von Beweismitteln unterstützt – die juristische Relevanz und Einordnung erfolgen stets durch erfahrene Ermittlungs- und Justizkräfte.

4.3. Führung im Wandel – Mindset eats knowledge for breakfast

Der dritte Themenkomplex betrifft die Führungskultur in einer digitalisierten Polizei. Führung verändert sich – nicht nur durch neue Technologien, sondern auch durch neue Anforderungen an Teamarbeit und Kreativität. Klassische Führung beruhte oft auf Fachwissen, hierarchischer Struktur und klaren Weisungsketten. Doch diese Ansätze stoßen in einer sich wandelnden Welt an ihre Grenzen.

Zunehmend entstehen gemischte Teams, in denen Polizistinnen und Polizisten mit IT-Spezialisten, Datenanalysten, Sozialwissenschaftlern oder Psychologen zusammenarbeiten. Sicherheitsarbeit wird nicht mehr ausschließlich durch uniformierte Kräfte gewährleistet. Vielmehr entsteht ein Netzwerk aus Kompetenzen, das strategisch und operativ ineinandergreifen muss. Dies erfordert von Führungskräften ein

neues Mindset: Offenheit für andere Berufsbiografien, Fähigkeit zur Integration divergierender Perspektiven, und nicht zuletzt emotionale Intelligenz.

Gute Führung wird künftig weniger durch reines Fachwissen definiert, sondern durch die Fähigkeit, ein Klima von Vertrauen, Kreativität und Innovation zu schaffen. Wer Teams inspiriert, die kollektive Intelligenz nutzt und auf Augenhöhe kommuniziert, wird nachhaltigen Erfolg haben. In der digitalen Polizei der Zukunft brauchen wir keine »allwissenden« Führungspersönlichkeiten – sondern reflektierte, lernbereite und integrative Köpfe, die den Wandel gestalten und gleichzeitig die Werte des Rechtsstaats hochhalten.

V. Wie verändert KI die polizeiliche Führungsarbeit?

Die Ergebnisse der Master-Thesis von *Raphael Humbel* (2025) an der Hochschule Luzern zeigen, dass polizeiliche Führungskräfte die Möglichkeiten von KI aktuell nur ungenügend kennen oder sich nicht ausreichend geschult fühlen. Nachfolgend werden einige aufbereitete Denkanstöße für das Führungsverständnis bei der Polizei formuliert.

5.1. Chancen, Gefahren und Grenzen von KI

- Menschenorientierte Führung: Entlastung und Ressourcengewinn
  Standardisierte und wiederkehrende Arbeiten können an KI übertragen werden. Dadurch können sich polizeiliche Führungskräfte auf ihre Kernaufgaben konzentrieren: entscheiden, motivieren und handeln.
- Qualitätsmanagement: Kommunikation und Personalentwicklung
  KI kann polizeilichen Führungskräften datengestützte Feedbacks ermöglichen und sie in einer klaren Kommunikation unterstützen. Der Einsatz von textbasierten Kommunikationssystemen (Chatbots) kann für Bereiche einen Vorteil bieten, in welchen Personen gute Gründe haben könnten, die Kommunikation mit Sprachmodellen der Interaktion mit Menschen vorzuziehen (bspw. Critical Incident Reporting System).
- Strategische Ausrichtung: Proaktivität und Agilität
  KI kann polizeiliche Führungskräfte durch die Analyse großer Datenmengen auf Handlungsmöglichkeiten hinweisen, frühzeitig vor Problemen warnen oder zukünftige Tendenzen berechnen.
- Change-Prozess; KI-unterstützte Führungskraft und Bedenken der Mitarbeitenden
  Unter dem Einfluss von KI verändern sich u.a. Anforderungen, Stellenprofile oder Aufgaben. Polizeiliche Führungskräfte müssen diesen Wandel aktiv angehen und steuern. Durch aktive Führung können sie KI und menschliches Handeln miteinander verbinden.
- Selbstführung und Werteorientierung
  KI ermöglicht polizeilichen Führungskräften durch Entlastung von Routineaufgaben eine Konzentration auf ihre Kernaufgaben. Selbstführung, ein menschenorientiertes Führungsverhalten und die innere Haltung stellen sicher, dass KI-Systeme umsichtig, verantwortungsvoll und rechtsstaatlich korrekt eingesetzt werden.

5.2. Ausblick und Handlungsempfehlungen

Die Implementierung von KI in den Führungsalltag erfordert eine grundlegende und breite Auseinandersetzung mit KI, insbesondere im Kontext polizeilicher Aufgaben. Diese Auseinandersetzung mit KI muss die Polizei jetzt führen. Aus der empirischen Forschung ergeben sich folgende konkreten Handlungsempfehlungen für die Polizei:

- Aktive Auseinandersetzung mit KI und mit dem Werte- und Risikomanagement:
  Die Polizei steht vor einem Transformationsprozess. Der bewusste Einsatz von KI in der polizeilichen Führungstätigkeit erfordert den Aufbau von Kompetenzen und Vertrauen in KI, klare Rahmenbedingungen und eine strategische Integration. KI soll ein ergänzendes Führungswerkzeug sein, nicht aber als Ersatz für menschliche Führung dienen. Ein erster Schritt in der Umsetzung bildet die Erarbeitung einer Strategie mit Schwerpunkt Digitalisierung und KI (bspw. Informationskampagne, KI-Richtlinien, Testing-Prozesse) sowie eine aktive Auseinandersetzung mit den Bereichen Werte- und Risikomanagement innerhalb der jeweiligen Polizeiorganisationen.
- Implementierung von KI-Kompetenzen in die polizeiliche Führungsausbildung:
  Die Ergebnisse der empirischen Forschung zeigen auf, dass von polizeilichen Führungskräften für den erfolgreichen Umgang mit KI entsprechende Informations- und Datenkenntnisse, Kommunikation und Kollaboration, die Fähigkeit, digitale Inhalte zu erstellen und zu bewerten, und Kompetenzen im Bereich Sicherheit und Problemlösung gefordert sind. Schlüsselkompetenzen wie KI-Prompting, Data Literacy und Quellenkritik müssen in die Führungsausbildungen aufgenommen werden.
- Informationelle Vernetzung als Grundlage für den Einsatz von KI:
  Neben technischen Fragen zu Speicherkapazität, Rechenkapazität oder zu geeigneter Software für die Datenanalyse ist die informationelle Vernetzung eine Voraussetzung dafür, dass KI für die Führung in der Polizei ein effektives und effizientes Werkzeug werden kann. Dabei gilt es, die informationelle Vernetzung auch aus rechtlicher und technischer Perspektive voranzutreiben.
- Kollaboration von Forschung, Entwicklung und polizeilicher Nutzung:
  Damit KI zielgerichtet der polizeilichen Führungstätigkeit nutzbar gemacht werden kann, ist es notwendig, dass Forschung und Entwicklung von KI in enger Kooperation mit der polizeilichen Nutzung von KI im Alltag verlaufen. Forschung und Entwicklung müssen den polizeilichen Alltag begleiten und individuell erkennen, in welchen Bereichen Potenzial für KI besteht. Gerade bei der Weiterentwicklung von KI für den Einsatz bei der Polizei wird eine Kooperation auf nationaler und internationaler Ebene als notwendig erachtet. Zudem sollte das Potenzial für Weiterentwicklungen in den Polizeiorganisationen erörtert und dann zentral verdichtet werden, damit »die gesamte Polizei« von den individuell gemachten Erfahrungen profitieren können.

VI. Fazit

Die Digitalisierung und der Einsatz von KI stellen die Polizei vor tiefgreifende Veränderungen. Dabei geht es nicht nur um Technik – sondern um Grundsatzfragen der Verantwortlichkeit, der Arbeitsorganisation und der Führung. Die Polizei steht dabei zwischen Innovation und Rechtsstaatlichkeit. Sie hat die Chance, Vorreiter zu sein – vorausgesetzt, sie nutzt die Technologie im Sinne der Menschen, für die sie Sicherheit garantiert. KI kann viel – aber die ethische und rechtliche Verantwortung bleibt immer menschlich.

In diesem Wandel übernehmen Führungskräfte eine zentrale Rolle, denn sie prägen maßgeblich die Integration, die Kommunikation und Kultur im Umgang mit den neuen KI-Technologien. Ihre Führungs- und Digitalkompetenzen, ihr Verantwortungsbewusstsein sowie ihre Fähigkeit, Orientierung zu geben und Vertrauen zu schaffen, sind entscheidend dafür, dass KI-Systeme werteorientiert und nachhaltig eingesetzt werden. Damit Führungskräfte dieser Verantwortung gerecht werden können, ist kontinuierliche Weiterbildung im Bereich digitaler Technologien und moderner Führungskompetenzen unerlässlich.

Haben Sie Interesse an einem weiteren Austausch? Wir freuen uns auf Ihre Kontaktaufnahme!

Literaturverzeichnis

Dahm, M. H.; Zehnder, V. (2023): Moderne Personalführung mit Künstlicher Intelligenz. Chancen und Risiken. Wiesbaden: Springer Gabler.

Dick, v.R.; Groß, M.; Piéch, S.; Bauer, K; Bibic, K; Kratzer, B.; Marx-Leck, S.; Schmidt, H. (2025): Wissenschaftliche Studie: KI und Entscheidungen. Goethe Universität Frankfurt.

Humbel, R. (2025): KI und Leadership bei der Polizei. Master-Thesis an der Hochschule Luzern – Wirtschaft im Studiengang Master of Advanced in Studies Leadership and Management.

Piéch, S. (2024): Die Auswirkungen der künstlichen Intelligenz auf die Zukunft der Arbeit. Erfolgreiches Leadership im Rahmen der digitalen Transformation. In: Basel. J.; Manchen Spörri, S. (Hrsg.): Angewandte Psychologie für die Wirtschaft. Berlin: Springer.

Schröder, C.; Türpe, O. (2025): Interview – Digitalisierung und Künstliche Intelligenz bei der Polizei Berlin mit Dr. Piéch.

Generative KI und die Transformation polizeilicher Führungskultur:

Chancen, Risiken und die Neudefinition von Autorität im Zeitalter algorithmischer Assistenz

von Kilian D. Grütter, Altendorf*

Die Integration generativer Künstlicher Intelligenz (GenKI) in den Führungsalltag der Polizei markiert eine Zäsur. Sie ist kein reines IT-Projekt, sondern ein kultureller Katalysator, der traditionelle Hierarchien und das Verständnis von Autorität herausfordert. Dieser Beitrag analysiert auf Basis aktueller empirischer Forschung und klassischer Führungstheorien (Adaptive, Transformational und Servant Leadership), wie GenKI die polizeiliche Führungspraxis verändert. Er beleuchtet die Spannung zwischen Effizienzgewinn und ethischem Risiko, diskutiert die Auswirkungen auf die »Blue Culture« und liefert pragmatische Handlungsempfehlungen für Führungskräfte in einer VUCA-Welt.

I. Einleitung: Wenn der Algorithmus den Lagebericht schreibt

Es ist Dienstagmorgen, 07:30 Uhr, in einem Polizeipräsidium einer deutschen Großstadt. Polizeipräsident Marcus W. bereitet sich auf eine Pressekonferenz zu einer Serie komplexer Cyberangriffe vor. Früher hätte sein Stab die halbe Nacht damit verbracht, Berichte aus drei verschiedenen Direktionen zu konsolidieren. Heute hat er innerhalb von drei Minuten eine präzise Zusammenfassung, rhetorische Vorschläge für das Statement und eine Risikoanalyse möglicher Journalistenfragen auf dem Bildschirm, generiert von einer isolierten Instanz eines Large Language Models (LLM), gefüttert mit internen Aktenvermerken.

Dieses Szenario ist keine Zukunftsmusik, sondern in Pilotprojekten bereits Realität. Generative KI (GenKI) wie ChatGPT, Claude oder Gemini unterscheidet sich fundamental von bisherigen KI-Anwendungen im Polizeibereich, wie etwa Predictive Policing. Während letzteres analytisch Muster erkennt, ist GenKI schöpferisch: Sie erstellt Texte, Konzepte, Code und Strategien.

Für polizeiliche Führungskräfte ergibt sich daraus ein Spannungsfeld: Einerseits bietet die Technologie nie dagewesene Möglichkeiten zur Entlastung von bürokratischer Last und zur Unterstützung komplexer Entscheidungen. Andererseits rüttelt sie an den Grundfesten einer Organisation, die wie kaum eine andere auf Erfahrungswissen, Seniorität und einer klaren Befehlskette (Chain of Command) basiert.[1] Wenn das Wissen demokratisiert wird und die Maschine schneller formuliert als der Leitende Polizeidirektor, was bedeutet das für die Führungskultur?

* Lic. phil. Kilian D. *Grütter*, Executive MBA Universität Zürich. Unternehmer, Dozent, Mediator. Kilian D. Grütter hält diverse Keynotes und unterschiedliche Weiterbildungsformate zu Chancen und Risiken zu KI, KI und Leadership im DACH-Raum für öffentlich-rechtliche Institutionen, Firmen und Unternehmen. Er führt u.a. diverse Weiterbildungsformate zu KI in der öffentlichen Verwaltung und beim Führungskreis der Deutschen Marine durch und ist Sparringpartner für Executives unterschiedlicher Hierarchiestufen, auch im Militär-, Polizei- und Sicherheitsbereich. www.kiliangruetter.ch.

1 Zum Bürokratiemodell und Hierarchieverständnis, das hier herausgefordert wird, vgl. klassisch: *Weber, M.* (1922): Wirtschaft und Gesellschaft. (Bezugnahme auf Bürokratiemodell).

Dieser Artikel untersucht nicht die technische, sondern die führungspsychologische und organisationale Dimension dieser Revolution.

II. Theoretische Grundlagen: Führung im digitalen Wandel

Um die Auswirkungen von GenKI auf die Polizei zu verstehen, müssen wir sie durch die Brille etablierter Führungsmodelle betrachten. Die polizeiliche Führung ist traditionell transaktional geprägt (Befehl und Gehorsam, Belohnung und Sanktion). Doch die Komplexität moderner Sicherheitslagen erfordert adaptivere Ansätze.

1. VUCA und Adaptive Leadership

Die polizeiliche Realität ist der Inbegriff einer VUCA-Welt (Volatility, Uncertainty, Complexity, Ambiguity). Ronald Heifetz' Konzept des *Adaptive Leadership* unterscheidet zwischen »technischen Problemen« (bekannte Lösungen, durch Experten lösbar) und »adaptiven Herausforderungen« (neue Probleme, die ein Umdenken erfordern).[2]

- **Die Rolle der KI:** GenKI ist hervorragend geeignet, technische Probleme zu lösen (Informationssynthese, Prozessoptimierung).
- **Die Rolle der Führungskraft:** Sie wird frei für die adaptiven Herausforderungen, etwa die ethische Bewertung eines Einsatzes oder die emotionale Stabilisierung der Belegschaft in Krisenlagen.

2. Transformational Leadership im KI-Kontext

Das Modell der Transformationalen Führung nach Bass und Avolio ist für die Polizei besonders relevant.[3] Es postuliert vier Dimensionen (die 4 I's):

1. **Idealized Influence (Vorbildfunktion):** Wie nutzt die Führungskraft KI? Ethisch und transparent oder verschleiernd?
2. **Inspirational Motivation:** Kann die Führungskraft eine Vision vermitteln, in der KI dem Menschen dient, statt ihn zu ersetzen?
3. **Intellectual Stimulation:** KI fordert alte Denkmuster heraus. Führungskräfte müssen kritisches Denken fördern, um KI-Ergebnisse zu hinterfragen.
4. **Individual Consideration:** Paradoxerweise kann KI durch Automatisierung von Routineaufgaben (z.B. Dienstplanungs-Entwürfe) Zeit schaffen, die Führungskräfte wieder in das persönliche Gespräch mit den Mitarbeitenden investieren können.[4]

3. Servant Leadership

Das Konzept des *Servant Leadership* ist in modernen Polizeiorganisationen hochrelevant.[5] Führung wird als Dienst am Geführten verstanden. GenAI kann hier als »Force Multiplier« dienen. Ein Dienststellenleiter, der KI nutzt, um bürokratische Hürden für sein Team schneller abzubauen oder Ressourcen intelligenter zuzuweisen, agiert als effektiverer Servant Leader.

Zwischenfazit: KI ersetzt Führung nicht. Sie verschiebt den Schwerpunkt von der Verwaltung (*Management*) hin zur echten Menschenführung (*Leadership*).

III. Forschungsstand: Empirie und Evidenz

Die Forschung zu GenAI im Management steckt noch in den Kinderschuhen, doch erste Studien aus dem DACH-Raum und den USA lassen sich auf den Polizeikontext übertragen und geben wichtige Hinweise:

- **Produktivitätsgewinne:** Eine Studie von Brynjolfsson et al. zeigte, dass Support-Mitarbeiter (ein mit Stabsarbeit vergleichbares Feld) durch GenAI ihre Produktivität um durchschnittlich 14 % steigerten, wobei gering qualifizierte Mitarbeiter um bis zu 35 % profitierten.[6] Für die Polizei bedeutet das: GenAI nivelliert das Spielfeld. Jüngere oder weniger schreibgewandte Führungskräfte können durch KI Defizite in der schriftlichen Kommunikation kompensieren.
- **Algorithm Aversion vs. Appreciation:** Dietvorst et al. prägten den Begriff der »Algorithm Aversion«, das Phänomen, dass Menschen Algorithmen nach einem Fehler schneller das Vertrauen entziehen als einem Menschen.[7] Neuere Untersuchungen deuten jedoch darauf hin, dass bei rein textbasierten Aufgaben (Zusammenfassungen, Entwürfe) die Akzeptanz rasant steigt.
- **Automatisierungs-Bias:** Ein robustes Phänomen in der Forschung ist der »Automation Bias«.[8] Menschen neigen dazu, algorithmischen Vorschlägen unkritisch zu folgen, besonders unter Zeitdruck, ein Dauerzustand im Polizeieinsatz.
- **Akzeptanz und Vertrauen:** Studien aus dem DACH-Raum zeigen, dass Vertrauen in KI-Systeme stark von der »Explainability« (Erklärbarkeit) abhängt.[9] In hierarchischen Organisationen wie der Polizei wird KI oft skeptisch gesehen, wenn sie als »Black Box« wahrgenommen wird, die die Entscheidungshoheit der Führungskraft untergräbt.
- **Führungskräfte als Gatekeeper:** Aktuelle Untersuchungen deuten darauf hin, dass die Haltung der direkten Vorgesetzten (*First-Line Supervisors*) entscheidend dafür ist, ob KI als Werkzeug zur Entlastung oder als Überwachungsinstrument wahrgenommen wird.[10]
- **Führungskultur:** Studien des Fraunhofer IAO zur Führung im digitalen Zeitalter betonen, dass digitale Kompetenz zur Kernkompetenz von Führungskräften wird.[11] Wer KI nicht bedienen kann, verliert an Autorität (»Legitimitätsverlust durch Inkompetenz«).

---

2 *Heifetz, R. A., Grashow, A., & Linsky, M.* (2009): The Practice of Adaptive Leadership. Harvard Business Press. Zur Erweiterung des VUCA-Modells siehe auch: *Cascio, J.* (2020): Facing the Age of Chaos. Medium (BANI-Framework).

3 Vgl. *Bass, B. M., & Avolio, B. J.* (1994): Improving organizational effectiveness through transformational leadership. Sage. Sowie grundlegend: *Bass, B. M.* (1985): Leadership and performance beyond expectations. Free Press.

4 *Avolio, B. J., et al.* (2014): E-leadership: Re-examining transformations in leadership source and transmission. The Leadership Quarterly, 25(1), 105–131.

5 *Greenleaf, R. K.* (1970): The Servant as Leader. Robert K. Greenleaf Center.

6 *Brynjolfsson, E., Li, D., & Raymond, L. R.* (2023): Generative AI at Work. NBER Working Paper No. 31161.

7 *Dietvorst, B. J., Simmons, J. P., & Massey, C.* (2015): Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126.

8 *Skitka, L. J., et al.* (2000): Automation bias and errors: Are crews better than individuals? International Journal of Human-Computer Studies, 52(6), 989–1008.

9 *Grote, G., et al.* (2022): Guidance for Human-AI Teaming in High-Stakes Environments. ETH Zürich Research Collection.

10 *Meijer, A., & Wessels, M.* (2019): Predictive Policing: Review of Benefits and Drawbacks. International Journal of Public Administration, 42(12), 1031–1039.

11 *Fraunhofer IAO* (2022): Führung im Zeitalter der KI – Auswirkungen auf die Arbeitswelt. Stuttgart.

IV. Anwendungsfelder generativer KI im Führungsalltag: Praxis pur

Wo stiftet GenKI heute schon konkreten Nutzen im polizeilichen »C-Level« und im gehobenen Dienst? Nachfolgend sind konkrete, anonymisierte Anwendungsfälle aus dem Polizeialltag aufgeführt.

1. Der digitale Stabsarbeiter (Administration und Reporting)

Ein Inspektionsleiter nutzt ein LLM, um aus stichpunktartigen Notizen von Mitarbeitergesprächen formulierte Zielvereinbarungen zu entwerfen.

- **Vorteil:** Zeitersparnis von ca. 60 %.
- **Qualitätssicherung:** Der Leiter prüft den Text auf Empathie und Tonalität. Das »weiße Blatt Papier« als Hürde entfällt.

2. Strategische Analyse und Lagebild

Führungskräfte müssen oft hunderte Seiten an Berichten sichten. LLMs können genutzt werden (in sicheren Umgebungen!), um:

- Muster in heterogenen Datenquellen zu erkennen (z.B. Zusammenhang zwischen Jugendkriminalität in Bezirk A und Social-Media-Trends).
- Szenarien zu simulieren.

3. Strategische Szenarienbildung (Red Teaming)

In Vorbereitung auf eine Hochrisikolage (z.B. Staatsbesuch) nutzt ein Planungsstab als erster Stresstest GenAI als »Red Team«.

- **Prompt:** »Agiere als Gegner, der versucht, die Sicherheitsmaßnahmen am Punkt X zu umgehen. Welche unkonventionellen Methoden würdest du wählen?«
- **Ergebnis:** Die KI schlägt Szenarien vor (z.B. koordinierte Drohnenschwärme in Kombination mit Cyberangriffen auf die Verkehrsleitzentrale), die im menschlichen Brainstorming untergegangen waren. Dies stärkt die Antizipationsfähigkeit.

4. Wissensmanagement und Onboarding

In vielen Behörden ist Wissen personengebunden (»Kopfmonopole«). Ein internes, auf Polizeidaten trainiertes GPT (Generative Pre-trained Transformer, eine Art von KI-Modell, das menschenähnliche Texte erstellen kann (wie bei ChatGPT)) kann als »Organizational Brain« fungieren. Dies ist besonders hilfreich beim Onboarding neuer Führungskräfte. Eine neue Dienstgruppenleiterin übernimmt eine Wache; statt sich durch Ordner zu wühlen, nutzt sie einen internen Chatbot, der mit den lokalen Dienstanweisungen und Protokollen der letzten fünf Jahre gefüttert wurde.

- **Frage:** »Wie wurde in den letzten zwei Jahren mit Beschwerden wegen Lärmbelästigung im Sektor 4 verfahren?«
- **Antwort:** Die KI liefert eine Zusammenfassung der gängigen Praxis und weist auf eine spezifische Dienstanweisung des Vorgängers hin.

5. Kommunikation und Übersetzung

Bei internationalen Kooperationen oder Bürgeranfragen in Fremdsprachen fungiert GenAI als Echtzeit-Übersetzer und Kulturvermittler, der nicht nur Wort-für-Wort übersetzt, sondern auch den angemessenen behördlichen Tonfall trifft.

6. Administrative Entlastung

- **Berichtswesen:** Automatisierte Erstellung von Entwürfen für Vermerke, Reden oder Jahresberichte.
- **Personaldienst und -entwicklung:** Analyse anonymisierter Mitarbeiterbefragungen auf Stimmungsbilder (*Sentiment Analysis*) oder Entwurf kompetenzbasierter Anforderungsprofile.

V. Fokus: Auswirkungen auf die Führungskultur

Dies ist der kritischste Punkt der Transformation. Die Einführung von GenKI ist ein kultureller Eingriff am offenen Herzen der Organisation. Technologie ändert Prozesse, aber sie kollidiert mit Kultur. Peter Drucker sagte: »Culture eats strategy for breakfast.« Bei GenAI könnte man sagen: »Culture eats AI for lunch.«

1. Von der Wissensautorität zur Entscheidungsautorität

Traditionell legitimiert sich eine polizeiliche Führungskraft oft durch Fachwissen (»Ich bin Chef, weil ich die StPO und die PDV 100 besser kenne als du«). GenKI demokratisiert dieses Wissen. Eine junge Kommissarin kann mit KI-Unterstützung in Sekunden Informationen abrufen, für die bisher ein erfahrener Hauptkommissar Jahre an Erfahrung brauchte. Die Konsequenz: Autorität kann nicht mehr auf reinem Faktenwissen basieren. Sie muss sich auf Urteilskraft (Judgment), Ethik und Verantwortung stützen. Chef:in ist nicht mehr das Lexikon, sondern der moralische Kompass.

2. Fehlerkultur und psychologische Sicherheit

GenKI halluziniert (erfindet Fakten). Wenn Führungskräfte KI nutzen, müssen sie eine Kultur etablieren, in der das Hinterfragen von (KI-)Ergebnissen belohnt wird. In einer strikten Befehlskultur wird ein/e Untergebene/r dem »KI-gestützten Befehl« des Chefs/der Chefin vielleicht nicht widersprechen, selbst wenn sie/er falsch ist. Das ist gefährlich. KI erfordert paradoxerweise *flachere* Hierarchien im Denken bei gleichzeitiger Beibehaltung der klaren Entscheidungslinie.

3. Entfremdung vs. Empathie

Es besteht das Risiko der »Führung per Prompt«. Wenn Lobesschreiben, Beurteilungen oder Kondolenzschreiben per KI generiert werden, verliert Führung ihre Seele. Leadership bedeutet hier: KI macht die Arbeit, der Mensch macht die Beziehung. Eine KI-generierte Beurteilung mag effizient sein, sie ist aber kulturell toxisch, wenn der/die Mitarbeitende die »Maschine« dahinter spürt.

VI. Risikodimensionen und ethische Herausforderungen: Das Minenfeld

1. Datenschutz und Informationssicherheit

Öffentliche LLMs (wie das Standard-ChatGPT) sind für Verschlusssachen tabu. Die Eingabe von personenbezogenen Daten (Namen von Beschuldigten, Opferdaten) stellt einen schweren Verstoß gegen DSGVO und Polizeigesetze dar. Das Risiko liegt im unbewussten Datenabfluss durch eine »Copy-Paste«-Mentalität.

2. Bias und Diskriminierung

KI-Modelle reproduzieren die Vorurteile ihrer Trainingsdaten. Wenn eine KI genutzt wird, um Einsatzschwerpunkte zu priorisieren (»Hotspot Policing«), und sie mit historischen Daten trainiert wurde, die ein *Racial Profiling* enthalten, wird

die KI rassistische Polizeipraktiken nicht nur spiegeln, sondern technisch legitimieren (»Der Computer sagt, dort ist es gefährlich«). **Führungsaufgabe:** Sensibilität für algorithmischen Bias ist eine neue Kernkompetenz für den höheren Dienst.

3. Die Haftungsfalle

Wer ist verantwortlich, wenn eine KI-basierte Entscheidung zu einem rechtswidrigen Eingriff führt? Nach geltendem Recht immer der Mensch (*Human-in-the-loop*). Führungskräfte dürfen sich nicht hinter der Technik verstecken (»Der Algorithmus war schuld«). Dies erfordert **Algorithmische Transparenz**: Führungskräfte müssen rudimentär verstehen, wie das Tool zu einem Ergebnis kam.

VII. Fallbeispiele aus der Praxis (Anonymisiert)

Fall A: Der effiziente Stabsleiter (Best Practice)

In einer Landeskriminalamt-Abteilung nutzt die Leiterin »Svenja M.« ein lokal gehostetes LLM, um tausende Seiten alter Ermittlungsakten eines Cold Cases zu strukturieren. Die KI erstellte eine Zeitleiste und wies auf Widersprüche in Zeugenaussagen hin, die menschlichen Ermittlern entgangen waren.

Erfolgsfaktor: Svenja M. nutzte die KI als Ideengeber, nicht als Entscheider. Jedes Ergebnis wurde von einem erfahrenen Ermittlerteam validiert (*Human Verification*). Die Führungskraft kommunizierte offen: »Das ist ein KI-Hinweis, prüft das.«

Fall B: Das toxische Briefing (Worst Case)

Ein Revierleiter ließ sich von einer öffentlichen KI ein Briefing für eine Demo-Lage erstellen. Er forderte: »Schätze das Gewaltpotenzial der Gruppe X ein.« Die KI, basierend auf Internet-Foren-Daten, halluzinierte eine hohe Gewaltbereitschaft, die real nicht vorlag. Der Leiter übernahm dies ungeprüft in den Einsatzbefehl. Die Folge: Ein überhartes Auftreten der Einsatzkräfte, Eskalation, negative Presse.

Fehler: Mangelnde Quellenkritik, Automation Bias, Nutzung unsicherer Datenquellen.

VIII. Strategische und operative Handlungsempfehlungen für die Praxis

Wie können Polizeiorganisationen das Potenzial heben und Risiken minimieren? Hier folgen strategische, operative und kulturelle Leitplanken.

1. Strategische Ebene (Behördenleitung)

- **Souveräne Infrastruktur:** Aufbau eigener, gesicherter KI-Instanzen (On-Premise oder Private Cloud), um Datensicherheit zu gewährleisten. Beschaffung von datenschutzkonformen KI-Zugängen.
- **KI-Leitlinien:** Erlass einer klaren Dienstvorschrift zur Nutzung von GenKI. Was ist erlaubt (Textentwürfe, Ideensammlung)? Was ist verboten (Personendaten, finale Entscheidungen)?

2. Operative Ebene (Führungskräfte)

- **Prompt Engineering lernen:** Die Fähigkeit, der KI präzise Anweisungen zu geben, wird zur neuen Kernkompetenz.
- **Pareto-Prinzip (80/20-Ansatz):** Nutzen Sie KI für den 80 %-Entwurf, investieren Sie Ihre menschliche Energie in die letzten 20 % (Feinschliff, Ethik, Empathie).
- **Das 4-Augen-Prinzip 2.0:** Jedes KI-Ergebnis muss von einem fachkundigen Menschen geprüft werden.
- **Kennzeichnungspflicht:** Dokumente oder Analysen, die maßgeblich durch KI erstellt wurden, sollten intern als solche gekennzeichnet werden. Transparenz schafft Vertrauen.

3. Kulturelle Ebene (HR und Entwicklung)

- **Entmystifizierung:** Weiterbildungen, die zeigen, dass KI keine Magie ist, sondern reine Wahrscheinlichkeitsrechnung. Dies baut Ängste ab und schafft Handlungssicherheit.
- **Bewahrung der basalen Kernkompetenzen:** Sicherstellen, dass junge Beamtinnen und Beamte das Schreiben und Zusammenfassen trotzdem »von der Pike auf« lernen, bevor sie es delegieren, um das *Cognitive Offloading* zu verhindern.
- **Schulungsoffensive:** Nicht nur IT-Schulungen, sondern Module zu »Critical Thinking« und »Data Literacy« an den Hochschulen der Polizei und in internen Weiterbildungsformaten anbieten.

IX. Leitfaden und Checkliste für Führungskräfte: Die »WAIT«-Regel

Bevor Sie als Führungskraft eine KI-Antwort verwenden, fragen Sie sich: **WAIT** (*Why Am I Talking/ Thinking* about this?):

1. **Wahrheit?** Habe ich die Fakten (Namen, Gesetze, Daten) an der Primärquelle verifiziert?
2. **Akkurat?** Passt der Kontext (kulturell, juristisch) zur spezifischen Polizeilage?
3. **Intern?** Sind sensible Daten enthalten, die nicht in den Prompt gehören?
4. **Transparent?** Weiß mein Team, dass dies KI-unterstützt ist?

X. Zukunftsausblick: Der/Die hybride Polizeiführer:in

Die Zukunft gehört der hybriden Intelligenz. In fünf Jahren werden multimodale KI-Systeme (Text, Bild, Video, Audio) fester Bestandteil von Führungs-Dashboards sein. Wir werden virtuelle »Adjutanten« erleben, die während einer laufenden Chaosphase (z.B. Amoklage) in Echtzeit Funkverkehr transkribieren, auf einer Karte visualisieren und taktische Optionen vorschlagen.

Doch je mächtiger die Technik, desto wichtiger der Mensch. Die Polizeiführungskraft der Zukunft ist ein »Ethischer Integrator und Instanz für die Umsetzungstauglichkeit«. Ihre Hauptaufgabe wird es sein, die kalte Logik der Algorithmen mit den Werten unserer freiheitlich-demokratischen Grundordnung und der menschlichen Empathie zu verweben. Wer diese Balance meistert, wird nicht ersetzt – sie/er wird unverzichtbar. Die KI liefert die Antworten, der Mensch stellt die richtigen Fragen und trägt die Verantwortung.

Fazit

Generative KI ist für die Polizei nicht nur ein IT-Update, sondern ein kultureller Stresstest. Sie bietet die historische Chance, Führungskräfte von der Last der Bürokratie zu befreien und kognitive Ressourcen für das Wesentliche freizusetzen: Die Arbeit am und mit dem Menschen.

Doch dieses Potenzial hebt sich nur, wenn Führungskräfte bereit sind, ihre Rolle neu zu definieren. Weg vom allwissenden Patriarchen, hin zum adaptiven, technologisch mündigen Gestalter. Die Anforderungsprofile künftiger Polizeikader müssen um neue Kompetenzen u.a. Wandelfähigkeit mit neuen Technologien und agile Lernbereitschaft angepasst werden. Die Gefahr liegt nicht in der KI selbst, sondern in einer Führungskultur, die sich weigert, sich mit ihr zu entwickeln. Die Polizei muss auch in Zukunft führen und dies vor allem auch im digitalen Raum: mit Verstand, Herz und einem kritischen Blick auf den Prompt.

Dieser Artikel wurde nach meinem eigenen Wissen und mit Recherchen mit KI (chatgpt.com und gemini.google.com) manuell zusammengestellt.

Anschriften von Schriftleitung und Redaktion der Zeitschrift DIE POLIZEI

Beiträge werden unmittelbar an die zuständigen Redakteure erbeten.

■ **Schriftleitung,**
Rechtsprechung, Aktuelles, Buchbesprechungen
Prof. Dr. jur. Dieter Müller,
An der Petruskirche 16, 06120 Halle (Saale)
Telefon (0160) 3718497
E-Mail: c-dieter.mueller@wolterskluwer.com

■ **Kriminalistik, Kriminologie**
betreut *Ltd. Kriminaldirektor a.D. Ralph Berthel,*
Mühlbergring 43, 09669 Frankenberg,
Telefon: (037206) 7 41 93,
E-Mail: ralph-berthel@web.de

■ **Gesellschaftswissenschaften, Fortbildung, Informationstechnik, Internationale Zusammenarbeit, Rechtswissenschaften**
betreut *Frau Prof. Dr. Sabrina Schönrock,*
Hochschule für Wirtschaft und Recht Berlin, Alt-Friedrichsfelde 60, 10315 Berlin, Telefon: (030) 308772878,
E-Mail: sabrina.schoenrock@hwr-berlin.de

■ **Einsatzmanagement, Management und Leadership, Führung, Digitalisierung**
betreut *Frau Prof. Dr. Sandra Schmidt*
Hochschule für Wirtschaft und Recht Berlin Alt-Friedrichsfelde 60, 10315 Berlin Telefon: (0156) 78357745
E-Mail: sandra.schmidt@hwr-berlin.de

■ **Struktur, Verkehrssicherheitsarbeit, Betriebswirtschaft, Organisation, Rechtswissenschaften**
betreut durch die Schriftleitung

Allgemeine Nachrichten und Informationen sowie Rezensionsexemplare werden direkt an die Schriftleitung erbeten.

Veröffentlichung

Die Beiträge werden nur zur Alleinveröffentlichung und insbesondere nur unter der Voraussetzung angenommen, dass sie keiner anderen Zeitschrift zur Veröffentlichung angeboten worden sind und werden. Die Annahme muss schriftlich erfolgen. Mit ihr erwirbt der Verlag vom Verfasser alle Rechte zur Veröffentlichung. Eingeschlossen sind insbesondere die Rechte zu elektronischen Publikationen der Beiträge in Datenbanken (online oder offline) oder Dokumentationssystemen ähnlicher Art und die Rechte, Beiträge zu gewerblichen Zwecken im Wege fotomechanischer oder anderer Verfahren zu vervielfältigen. Für Manuskripte, die unaufgefordert eingesandt werden, wird keine Haftung übernommen. Mit vollem Verfassernamen gekennzeichnete Beiträge geben nicht unbedingt die Meinung der Redaktion wieder.

Peer-Review-Verfahren

Die in unserer Zeitschrift veröffentlichten Aufsätze durchlaufen sämtlich einem internen Peer-Review-Verfahren durch die Fachredaktion. Einzelne Aufsätze, bei deren Inhalten keine fachliche Expertise in der Redaktion vorhanden ist, durchlaufen zusätzlich ein Peer-Review-Verfahren bei externen Fachexperten.

Herausgeber
Ralph Berthel, Prof. Dr. Dieter Müller, Holger Münch, Prof. Dr. Sandra Schmidt, Prof. Dr. Sabrina Schönrock

Verlag
Wolters Kluwer Deutschland, Carl Heymanns Verlag, Wolters-Kluwer-Straße 1, 50354 Hürth, Stefan Kolbe, Telefon (02233) 3760 - 7646, stefan.kolbe@wolterskluwer.com, http://www.wolterskluwer.de. Kundenservice: Telefon (02233) 3760-7201, E-Mail: info-wkd@wolterskluwer.com

Bezugsbedingungen
Die Zeitschrift erscheint zwölfmal im Jahr. Jahrespreis 255,00 € zzgl. Versandkosten (39,00 € Inland/50,40 € Ausland). Vorzugspreis für Studenten der DHPol u. anderer PolFHS 99,00 € zzgl. Versandkosten (39,00 € Inland/50,40 € Ausland). Einzelheft 23,00 € zzgl. Versandkosten (ca. 2,10 € Inland/ca. 3,50 € Ausland) je nach Heftumfang. Preise inkl. MwSt. Das Jahresabonnement verlängert sich um ein Jahr, wenn es nicht mit einer Frist von 3 Monaten zum Ende des Vertragsjahres schriftlich gekündigt wird.

Anzeigen
Anzeigenverkauf: Gabriele Wieneber, Telefon (02233) 3760-7608, E-Mail: gabriele.wieneber@wolterskluwer.com
Anzeigendisposition: Karin Odening, Telefon (02233) 3760-77 60, E-Mail: anzeigen@wolterskluwer.com
Die Anzeigen werden nach der Preisliste 46 vom 01.01.2026 berechnet. Gedruckt auf chlor- und säurefreiem Papier (DIN-ISO).

Satz
Newgen Knowledge Works (P) Ltd., Chennai

Druckerei
Lotos Poligrafia Sp. z o.o., Warszawa Polen
ISSN 0032-3519

Hinweis
Die Beiträge orientieren sich bei der Verwendung des Genus an den Regeln der Dudenredaktion. Sofern im Text das Maskulinum verwendet wird, aus dem Sachzusammenhang allerdings, Femininum (weiblich) und Neutrum (sächlich) gemeint sind, steht das Maskulinum für die beiden anderen Genera. Ein sog. Gendern erfolgt aus Gründen der Lesbarkeit nicht.

Ein Handkommentar für die Praxis

Mit der 3. Auflage 2026 auf dem neuesten Stand im Öffentlichen Recht:

Der neue Kommentar zur VwGO stellt die Verwaltungsgerichtsordnung wissenschaftlich fundiert und europarechtlich sensibilisiert dar und zeigt die praktische Verzahnung mit dem Fach- und Sonderverwaltungsprozessrecht auf.

Neu in der 3. Auflage:

- parallele Vorschriften in der Finanzgerichtsordnung (FGO) und im Sozialgerichtsgesetz (SGG)
- zahlreiche verwaltungsprozessrechtliche Vorschriften, die außerhalb der VwGO und von erheblicher Bedeutung für die Praxis sind
- eigenständige Kapitel zum Sonderverwaltungsprozessrecht im Asylverfahren, in Disziplinarsachen und im Öffentlichen Wettbewerbsrecht
- das europäische Verwaltungsprozessrecht

Ab der 3. Auflage unter neuer Herausgeberschaft: aus Gärditz wird Knauff.

Knauff, VwGO – Kommentar – neben vielen anderen Titeln enthalten im Modul Allgemeines Verwaltungsrecht auf Wolters Kluwer Online.

ISBN 978-3-452-30214-4, ca. € 189,–

Onlineausgabe ca. € 11,69 mtl.
(im Jahresabo zzgl. MwSt)

Auch im Handel erhältlich

Wolters Kluwer

Mehr Infos:

shop.wolterskluwer-online.de →

Essenzielle Inhalte des Verkehrsrechts

Mit dem Modul Verkehrsrecht auf dem neuesten Stand:

Die Online-Bibliothek mit essenziellen Inhalten des Verkehrsrechts, speziell für alle Verkehrsrechtler:innen.

- Hochrelevante Inhalte zum Verkehrsrecht
- Zahlreiche allgemeine und Spezialtitel, beispielsweise zum **Personenschaden** und **Schmerzensgeld** sowie **Reinking / Eggert, Der Autokauf**
- Mit der häufig zitierten **ADAC Zeitschrift DAR – Deutsches Autorecht** inkl. umfassendem Onlinearchiv

Zusammenfassungen von Urteilen und Beschlüssen powered by Expert AI
Nutzen Sie das Potential künstlicher Intelligenz, um den juristischen Rechercheprozess zu verkürzen.

Profitieren Sie dabei von einer klaren optischen Trennung zwischen Sachverhalt und Begründung, damit Sie schneller finden, was Sie suchen. Besonders hervorzuheben ist der Fokus auf die entscheidungserheblichen Gründe, durch die Sie die wesentlichen Argumente des Gerichts auf einen Blick erfassen können.

Überzeugen Sie sich selbst von der Qualität unserer Zusammenfassung powered by Expert AI auf Wolters Kluwer Online.

Auch im Handel erhältlich

Wolters Kluwer

shop.wolterskluwer-online.de →

Profundes Fachwissen für Verwaltungsrechtler

Die Online-Bibliothek bietet Expert:innen in der öffentlichen Verwaltung oder Fachanwält:innen für Verwaltungsrecht relevante Titel für ihre tägliche Arbeit.

- Das Praxishandbuch **Verwaltungsvollstreckungsrecht** von App / Wettlaufer / Klomfaß
- Der **Knack / Hennecke, VwVfG Kommentar,** informiert über die Einzelheiten zu allen Verfahrensfragen
- Das **Handbuch Verwaltungsrecht von Terwiesche / Prechtel** bietet zahlreiche Checklisten, Beispiele aus der aktuellen Rechtsprechung und wertvolle Praxistipps.
- Als erste öffentlich-rechtliche Fachzeitschrift Deutschlands bietet das **DVBl – Deutsche Verwaltungsblatt** praxisorientierte Fachinformationen.

Zusammenfassungen von Urteilen und Beschlüssen powered by Expert AI
Nutzen Sie das Potential künstlicher Intelligenz, um den juristischen Rechercheprozess zu verkürzen.

Profitieren Sie dabei von einer klaren optischen Trennung zwischen Sachverhalt und Begründung, damit Sie schneller finden, was Sie suchen. Besonders hervorzuheben ist der Fokus auf die entscheidungserheblichen Gründe, durch die Sie die wesentlichen Argumente des Gerichts auf einen Blick erfassen können.

Überzeugen Sie sich selbst von der Qualität unserer Zusammenfassung powered by Expert AI auf Wolters Kluwer Online.

Auch im Handel erhältlich

Wolters Kluwer

shop.wolterskluwer-online.de →