Anwendungsfelder und Potenziale algorithmenbasierter Polizeiarbeit

Dr. iur. Martin Thüne[1]

Big Data, Data Mining, Künstliche Intelligenz, Maschinelles Lernen – diese und weitere Schlagworte fallen regelmäßig, wenn über das Thema „Digitalisierung" berichtet wird. Obgleich Inhalte und Grenzen der einzelnen Konzepte oft im Unklaren bleiben und sich mitunter der Eindruck einer gewissen Beliebigkeit in der Begriffswahl aufdrängt, so wird, gewissermaßen als kleinster gemeinsamer Nenner, schnell klar, dass bei all diesen Verfahren Computer eine Rolle spielen, die mit *Algorithmen* operieren. Allgemein versteht man darunter „[...] eindeutige Handlungsvorschriften zum Lösen eines vorab definierten Problems [...]."[2] Das maschinelle Lösen komplexer Aufgaben zielt letztlich darauf ab, menschliche Arbeit zu erleichtern, zu optimieren oder diese gar obsolet werden zu lassen. Außerdem können Algorithmen entscheidungsunterstützend arbeiten. Ganz wesentlich sollen mit ihrer Hilfe große Datenmengen gesichtet, zusammengeführt und ausgewertet werden.

Es ist hinlänglich bekannt, dass die Sicherheitsbehörden – neben anderen Professionen – vor der Herausforderung einer anschwellenden Informationsflut stehen, mit der sie dennoch möglichst effizient umgehen müssen. Dies betrifft neben dem Bereich strafprozessualer Ermittlungen auch das Feld der operativen Bewältigung von Einsatzlagen, mithin der Gefahrenabwehr. Konnte die Polizei über lange Zeit ihre Arbeit im Wesentlichen mit Stift und Notizblock, diversen Formularen, Handakten und „Lichtbildmappen" bewältigen, so schreitet die *Datafizierung* ehemals analoger bzw. dinglicher Tätigkeiten unaufhaltsam voran. Algorithmen und computerbasierte Polizeiarbeit bilden zunehmend die Grundlage sowohl von hochspezialisierter, aber auch „ganz gewöhnlicher" Polizeiarbeit.

In diesem Beitrag soll überblicksartig skizziert werden, welche Methoden als Schlüsseltechnologien im Feld algorithmenbasierter Ansätze gelten, in welchen polizeilichen Anwendungsfeldern diese bereits zum Einsatz kommen und welche Potenziale damit einhergehen. Die (datenschutz-)rechtlichen, ethischen und politischen Konfliktlinien, die hierbei mitzudenken sind, sollen an dieser Stelle nur deshalb weitgehend ausgeblendet werden, weil sich andere Beiträge in diesem Band explizit mit solchen Fragestellungen befassen.

[1] Der Autor ist Dozent für Kriminologie und Beauftragter für Forschung am Fachbereich Polizei der Thüringer Fachhochschule für öffentliche Verwaltung.
[2] Statt vieler: Krüger / Lischka (2018), S. 9.

Künstliche Intelligenz als Schlüsseltechnologie

Die informationstechnisch-mathematische Grundlage zahlreicher moderner Big-Data-Anwendungen ist die sog. *Künstliche Intelligenz (KI)*. Dieser Oberbegriff steht für eine Art von (Computer-)Systemen bzw. Methoden, mit deren Einsatz der Anspruch verfolgt wird, menschliche Intelligenz annähernd nachzubilden. Bei aller Unterschiedlichkeit der einzelnen Ansätze ist diesen gemein, dass sie *Merkmale intelligenten Verhaltens* aufweisen, die ursprünglich dem Menschen vorbehalten waren. Dazu zählt insb. die Fähigkeit zur *Lösung wenig strukturierter Probleme*.[3] Effizientes Problemlösen setzt voraus, dass man über relevante Wissensbestände verfügt bzw. Wissen verfügbar gemacht, aggregiert und kontextbezogen zur Anwendung gebracht werden kann. LUNZE nennt als weitere Kennzeichen für Intelligenz „[...] die Fähigkeiten,

- Situationen trotz mehrdeutiger oder widersprüchlicher Informationen zu erkennen,
- Ähnlichkeiten von Situationen, Aufgaben und Lösungswegen trotz großer Unterschiede herauszufinden,
- flexibel und situationsabhängig zu entscheiden und dabei die relative Wichtigkeit verschiedener Elemente einer Situation zu berücksichtigen und günstige Umstände auszunutzen,
- aus Erfahrung zu lernen.“[4]

Nach gegenwärtigem Forschung- und Entwicklungsstand sind intelligente Systeme in der Lage, *einzelne*, wenngleich hochkomplexe solcher Anwendungsprobleme zu lösen. Dieser KI-Teilbereich wird als *schwache Künstliche Intelligenz* bezeichnet. Anders als es die Begriffswahl impliziert, sind entsprechende Systeme

> „ [...] nicht nur bereits vielfach den bisherigen von Menschen programmierten Systemen überlegen, sondern auch den menschlichen Fähigkeiten selbst – teilweise bereits bei Aufgaben, die aufgrund von Kreativität und Komplexität bislang nicht automatisierbar waren [...].“[5]

Die potenzielle Leistungsfähigkeit schwacher Künstlicher Intelligenz sollte dementsprechend nicht unter-, aber ebenso nicht überschätzt werden. Trotz enormer Fortschritte im Bereich schwacher KI haben sich einige Forschende dem langfristigen Ziel verschrieben,

[3] Vgl. Lunze (2016), S. 2.
[4] Lunze (2016), S. 2.
[5] Welsch / Eitle / Buxmann (2018), S. 369.

die menschliche Intelligenz in ihrer Gesamtheit reproduzieren zu wollen, d.h. eine *starke Künstliche Intelligenz* zu schaffen. Allerdings ist hoch umstritten, welche Zielkriterien für eine solche Intelligenz maßgeblich sein sollen und ob ein solches Vorhaben überhaupt jemals realisierbar ist.[6]

Ein Begriff, der im Kontext Künstlicher Intelligenz allgemein, aber auch mit Blick auf entsprechende polizeiliche Verfahren häufig fällt, ist das *Maschinelle Lernen* (engl.: Machine Learning):

> „Auf hochleistungsfähigen Hard- und Softwareplattformen bieten die maschinellen Lernverfahren der Künstlichen Intelligenz das Instrumentarium, um aus großen Datenmengen komplexe Zusammenhänge zu lernen und in Entscheidungen und Handlungen umzusetzen, und zwar ohne explizit programmiert werden zu müssen."[7]

Der letztgenannte Aspekt stellt das zentrale Charakteristikum maschineller Lernverfahren dar: Mithilfe „selbstlernender Algorithmen" sind entsprechende Anwendungen in der Lage, ihre Performanz (zumindest in Teilen) eigenständig zu verbessern. Allerdings erfolgt dieser Lernvorgang in vielen Fällen nicht vollständig autark, sondern von Menschenhand angeleitet: Beim *supervised learning* (angeleitetes / überwachtes Lernen) werden einem selbstlernenden Algorithmus sog. Trainingsdatensätze angeboten. Diese enthalten bereits „gelabelte", d.h. vorstrukturierte Informationen. Auf dieser Grundlage kann die Maschine bestimmte Zusammenhänge, Wichtungen, Gruppenzugehörigkeiten o.Ä. erkennen und sodann eigenständig Regeln ableiten. Nach weiteren Zwischenschritten (sog. Validierungs- und Testphasen) ist der Algorithmus bzw. das mathematische Modell potenziell in der Lage, das Gelernte in Bezug auf einen ungelabelten, unbekannten Datensatz anzuwenden. Voraussetzung für „supervised learning" ist das Vorhandensein von theoretischen Modellen bzw. empirisch abgesicherten Annahmen, welche sich in den klassifizierten Testdatensätzen abbilden. Der Computer lernt also aus der Erfahrung des Menschen. Dieser Ansatz eignet sich potenziell etwa für *Predictive-Policing*-Anwendungen (s. Tabelle unten), weil hier spezifische kriminologische Theorien und ein gewisses empirisches Fundament vorhanden sind, welche zum Anlernen der Künstlichen Intelligenz genutzt werden können.[8]

[6] Vgl. Moeser (2017). Bezüglich der Zielkriterien stellt sich u.a. die Frage, welche Rolle menschliche Fähigkeiten wie Bewusstsein, Empathie u.Ä. im Feld Künstlicher Intelligenz spielen sollen bzw. können.
[7] Hecker / Döbel / Rüping et al. (2017), S. 27.
[8] Vgl. Suthaharan (2016), S. 7 ff., 25 f.; Pollich / Bode (2017), S. 9.

Ziele und Erwartungen bzgl. algorithmenbasierter Polizeiarbeit

In einem aktuellen Artikel zum vierjährigen Bestehen der *Zentralen Stelle für Informationstechnik im Sicherheitsbereich* (ZITiS)[9] thematisiert das „Team Big-Data-Analyse" der Bundesanstalt die Herausforderungen, die sich für Sicherheitsbehörden in Verbindung mit den immer größer werdenden Datenmengen ergeben und beschreibt Lösungsansätze.[10] Schon im Rahmen der Daten*sichtung*, besonders aber der Daten*auswertung* sind die Behörden demnach zunehmend auf Programme angewiesen, die automatisiert oder zumindest teilautomatisiert arbeiten und belastbare Ergebnisse produzieren. Entsprechende Programme müssen z.B. in der Lage sein, heterogene Text-, Bild-, Video- und Audiodaten zusammenzuführen und diese auswertbar zu machen. Die Erwartung der Anwender an algorithmenbasierte Polizeiarbeit ist regelmäßig das Ermöglichen schnellerer, effizienterer und flexiblerer Analysen, wobei gleichzeitig die Vertrauenswürdigkeit und Nachvollziehbarkeit der Arbeitsprozesse eine zentrale Rolle spielen.

Wenngleich in der jüngeren Vergangenheit erhebliche technische Fortschritte erzielt wurden und die Erwartungshaltungen entsprechend groß sind, können nicht alle Verfahren die Ansprüche in der Praxis (unmittelbar) erfüllen. Dies ist nicht verwunderlich, da einige der relevanten Problemkonstellationen nach wie vor als relativ neu gelten. Folglich bewegen sich entsprechende Lösungsansätze entweder noch im Bereich von Grundlagenforschung oder überspringen gerade erst die Schwelle zur Praxistauglichkeit.

Die „Nationale KI-Strategie der Bundesregierung"

Die technischen, rechtlichen, gesellschaftlichen und wirtschaftlichen Herausforderungen, die sich im Zusammenhang mit einer zunehmenden Verbreitung und Weiterentwicklung von Künstlicher Intelligenz stellen, haben politische Verantwortliche im Bund zum Anlass genommen, um gemeinsam mit zahlreichen Akteuren ins Gespräch zu kommen und eine „Nationale Strategie für Künstliche Intelligenz" (Slogan: „AI Made in Germany") zu erarbeiten.[11] Als erstes zentrales Ergebnis wurde Ende 2018 ein Strategiepapier veröffentlicht,

[9] „ZITiS ist Dienstleister für die Sicherheitsbehörden in Deutschland. Die Aufgaben orientieren sich am Bedarf der Sicherheitsbehörden und umfassen die Bereiche digitale Forensik, Telekommunikationsüberwachung, Krypto- und Big-Data-Analyse wie auch technische Fragen der Kriminalitätsbekämpfung, Gefahren- und Spionageabwehr. ZITiS hat keine Eingriffsbefugnisse und ist keine Beschaffungsorganisation." Selbstbeschreibung der ZITiS auf: https://www.zitis.bund.de/DE/ZITiS/Aufgaben/aufgaben_node.html; Zugriff: 16.12.2021.

[10] Vgl. hier und in diesem Absatz: Zentrale Stelle für Informationstechnik im Sicherheitsbereich (2021), S. 45.

[11] <https://www.ki-strategie-deutschland.de/>. Zugriff: 16.12.2021.

mit dem die Bundesregierung „[...] einen Rahmen für eine ganzheitliche politische Gestaltung der weiteren Entwicklung und Anwendung Künstlicher Intelligenz in Deutschland“[12] zu setzen beabsichtigt. Darin erfolgt u.a. eine Positionierung hinsichtlich einer „Nutzung von KI in der öffentlichen Verwaltung“ sowie speziell im Kontext der „Gefahrenabwehr und zur inneren und äußeren Sicherheit“. Nach Auffassung der Bundesregierung kann KI etwa

> „[...] in der Strafverfolgung/Gefahrenabwehr zum Schutz der Bürgerinnen und Bürger oder zur Steuerung des Einsatzes von Polizeikräften eingesetzt werden. Andere Anwendungsgebiete sind unter Wahrung der betroffenen Persönlichkeitsrechte und unter bestimmten Voraussetzungen das Predictive Policing (präventive Gefahrenabwehr), der Schutz von Kindern und Jugendlichen vor sexualisierter Gewalt im Internet und die Bekämpfung und Verfolgung der Verbreitung von Missbrauchsdarstellungen oder Social Media Forensics zur Bildung von Personenprofilen.“[13]

Das Papier macht deutlich, dass dem Thema KI generell, aber auch mit Blick auf deren Einsatz durch Sicherheitsbehörden, ein hoher Stellenwert beigemessen wird. Die Bundesregierung strebt deshalb an, „[...] geeignete Themenfelder für die Sicherheitsbehörden [zu] identifizieren und KI im Sinne einer agilen, praxisnahen Entwicklung [zu] fördern.“[14] Für die Umsetzung der gesamten KI-Strategie mit ihren zahlreichen Handlungsfeldern sollten im Zeitraum von 2019 bis 2025 vom Bund rund 3 Milliarden Euro bereitgestellt werden.[15] In einer im Jahr 2020 verabschiedeten Fortschreibung wird diesbezüglich nunmehr eine Mittelerhöhung auf 5 Milliarden Euro angekündigt.[16] Das Papier enthält zudem Ausführungen u.a. zum Ordnungsrahmen und regulatorischen Handlungserfordernissen, die sich mit Blick auf eine zunehmende Verbreitung von Künstlicher Intelligenz sowie deren Einsatz insbesondere durch staatliche Akteure stellen.

[12] (Die) Bundesregierung (2018), S. 4.
[13] Ebd., S. 33.
[14] Ebd., S. 32.
[15] Vgl. ebd., S. 6.
[16] (Die) Bundesregierung (2020), S. 7.

Aktuelle Beispiele für algorithmenbasierte Polizeiarbeit

In welchen Einsatzfeldern algorithmenbasierte Polizeiarbeit bereits stattfindet bzw. erprobt wird, soll folgende Tabelle verdeutlichen. Gleichzeitig handelt es sich dabei um eine Auswahl – die Fortschritte insbesondere im Feld Künstlicher Intelligenz eröffnen regelmäßig neue Anwendungsoptionen.

Anwendungsfeld	**Beispiele**
Auswertung großer Datenmengen im Rahmen von Ermittlungsverfahren	Ermittlungskomplex „Panama Papers“ (weltweites Konglomerat zur legalen und illegalen Steuervermeidung):[17] • Ausgangsdatensatz von rund 2,7 Terabyte (≙ ca. 600 DVDs) mit über 41 Millionen Quelldokumenten, bestehend aus diversen Dokumenttypen und Dateiformaten in verschiedenen Sprachen (weniger als 0,1 % deutschsprachig) • in der Folge Entwicklung und Einsatz ermittlungsunterstützender Anwendungen u.a. auf Basis Künstlicher Intelligenz (inkl. Künstlicher Neuronale Netze) durch das BKA, um Daten zu sichten, zusammenführen und auszuwerten
Automatisierte Mustererkennung in Audio- und Videodatenströmen zur Detektion potenziell gefährlicher Situationen	Analyse von Personenbewegungen im Umfeld öffentlicher Verkehrsknoten (Flug- und Seehäfen, Bahnhöfe u.Ä.):[18] • Erforschung von potenziell „verdächtigen Verhaltensmustern“ und Operationalisierung bzw. Mustererstellung

[17] Vgl. dazu etwa: Schattauer (2019).

[18] Siehe dazu etwa das bereits 2010-2014 durchgeführte Projekt APFel – „Analyse von Personenbewegungen an Flughäfen mittels zeitlich rückwärts- und vorwärtsgerichteter Videodatenströme“. Projekthomepage: https://www.sifo.de/sifo/de/projekte/schutz-vor-kriminalitaet-und-terrorismus/mustererkennung/apfel/apfel-analyse-von-personenbewe-sgerichteter-videodatenstroeme.html. Zugriff: 16.12.2021.