

Künstliche Intelligenz und automatisierte Entscheidungen in der öffentlichen Verwaltung

Prof. Dr. iur. Nadja Braun Binder (Basel)*

I. Einleitung

Die digitale Transformation ist allgegenwärtig. Fortschritte im Bereich des maschinellen Lernens, eine enorm hohe Datenverfügbarkeit und stark zunehmende Rechenkapazitäten erlauben heute diverse Anwendungen, die verschiedentlich mit dem Schlagwort Künstliche Intelligenz (KI) beschrieben werden. Im Kern geht es dabei häufig um automatisierte bzw. automationsgestützte Entscheidungsfindung (Automated Decision Making – ADM).¹

Nicht nur die Privatwirtschaft setzt zunehmend auf den Einsatz von KI und ADM. Weltweit prüfen und nutzen öffentliche Verwaltungen entsprechende Anwendungen. Inzwischen berührt die Thematik auch die Schweiz. So wird der Bundesrat in einem Postulat aufgefordert zu prüfen, wie die Effizienz in der Bundesverwaltung mithilfe von Prozessautomatisierung und KI optimiert werden kann.² Unabhängig von diesem Postulat hat der Bundesrat im September 2018 eine interdepartementale Arbeitsgruppe unter Federführung des Staatssekretariats für Bildung, Forschung und Innovation eingesetzt, die den Wissens- und Meinungs austausch zum Thema KI sicherstellen und bis Herbst 2019 einen Bericht vorlegen soll.³ Dieser soll sich unter anderem mit KI in der Bundesverwaltung beschäftigen. Auch in der Justiz findet das Thema zunehmend Beachtung. Die Europäische Kommission für die Wirksamkeit der Justiz des Europarates hat im Dezember 2018 eine Charta mit Ethikgrundsätzen für die Anwendung von KI in der Justiz verabschiedet.⁴ Sie soll den verant-

Der Einsatzbereich von Künstlicher Intelligenz (KI) umfasst neben der Privatwirtschaft immer häufiger auch die öffentliche Verwaltung. In der Schweiz soll der Bundesrat prüfen, wie die Effizienz in der Bundesverwaltung mithilfe von Prozessautomatisierung und KI optimiert werden kann. Ausgehend von der Prämisse, dass KI in Zukunft in der schweizerischen öffentlichen Verwaltung eine Rolle spielen wird, nimmt der Beitrag eine Auslegeordnung offener Rechtsfragen vor. Zudem beleuchtet die Autorin anhand konkreter Beispiele aus dem Ausland mögliche Einsatzfelder in der öffentlichen Verwaltung und in der Justiz.

Le recours à l'intelligence artificielle (IA) n'a pas seulement lieu dans l'économie privée, mais de plus en plus aussi dans l'administration publique. En Suisse, il faut que le Conseil fédéral examine comment optimiser l'efficacité de l'administration fédérale à l'aide d'une automatisation des processus et de l'IA. Partant du principe que l'IA jouera un rôle dans l'administration publique suisse à l'avenir, cet article présente une interprétation des questions juridiques encore en suspens. En outre, l'auteure utilise des exemples concrets de l'étranger pour mettre en lumière les domaines d'application possibles dans l'administration publique et le système judiciaire. P.P.

* Die Autorin ist seit dem 1. August 2019 Professorin für Öffentliches Recht an der Universität Basel.

¹ Vgl. Ziff. II.3. unten.

² Postulat FDP-Liberale Fraktion (18.3783) vom 19.9.2018: Effizienzsteigerung beim Bund durch intelligente Prozessautomatisierung in der Verwaltung.

³ Vgl. Aktionsplan Digitale Schweiz vom 5.9.2018, Punkt 7.5, abrufbar unter <<https://bit.ly/2xwR38m>> zuletzt besucht am 27.5.2019.

⁴ European Commission for the Efficiency of Justice (CEPEJ), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, 3./4.12.2018, CEPEJ(2018)14, abrufbar unter <<https://bit.ly/2G18u8x>> zuletzt besucht am 27.5.2019.

wortlichen Gremien in den Mitgliedstaaten, zu denen auch die Schweiz zählt, als Leitfaden im Umgang mit KI in Justizverfahren dienen. Im Bereich der Gesetzgebung hingegen ist der Einsatz von KI derzeit nicht absehbar. Zwar ist der Gesetzgeber für die Festlegung der Rahmenbedingungen eines verantwortungsvollen KI-Einsatzes in der öffentlichen Verwaltung zuständig. Aber eine Einbindung von KI im Rahmen der Tätigkeit der Legislative selbst findet momentan nicht erkennbar statt.⁵

Ausgehend von der Prämisse, dass KI in Zukunft auch in öffentlichen Verwaltungen in der Schweiz eine Rolle spielen wird, soll dieser Beitrag eine – nicht als abschliessend zu verstehende – Auslegeordnung offener Rechtsfragen präsentieren. Zu diesem Zweck wird in einem ersten Schritt erörtert, was hier unter KI bzw. ADM verstanden wird (II.). Anhand konkreter Beispiele aus dem Ausland werden sodann mögliche Einsatzfelder in der öffentlichen Verwaltung und in der Justiz (III.) sowie, darauf aufbauend, offene Rechtsfragen skizziert (IV.). Den Schluss bildet ein Ausblick (V.).

II. Begriffliche Eingrenzung

Der Begriff KI erlebt in letzter Zeit eine regelrechte Renaissance. Seine Ursprünge liegen in den 1950er-Jahren. So wird in der Literatur etwa auf eine Konferenz am Dartmouth College in Hanover (New Hampshire, USA) aus dem Jahr 1956 verwiesen, an welcher der Ausdruck erstmals Erwähnung gefunden haben soll.⁶ Bisweilen wird die Geburtsstunde der KI auch früher angesetzt, nämlich 1947, als Alan Turing auf einem Symposium in Manchester die zentrale Frage der KI formulierte: «Können Maschinen denken?»⁷ Die Diskussionen reichen also mehrere Jahrzehnte zurück, und dennoch gibt es bis heute keine einheitliche Definition. KI, in erster Linie eine Subdisziplin der Informatik, beschreiben Informatiker in ganz

unterschiedlicher Art und Weise.⁸ Es existiert nicht etwa ein Programmiercode, der in Abgrenzung zu anderen Codes als KI bezeichnet werden könnte. Vielmehr existiert eine Vielzahl von Methoden, die mit dem Schlagwort KI versehen werden. Dazu zählen unter anderem maschinelle Lernverfahren und künstliche neuronale Netze. Mithin kann die KI selbst in verschiedene Teildisziplinen unterteilt werden.⁹ Neben der Informatik befassen sich u.a. die Philosophie oder die Neurowissenschaften mit KI und legen jeweils aus der eigenen Forschungsperspektive entsprechende Verständnis zugrunde.¹⁰ Es ist demnach notwendig, die Terminologie für die Zwecke dieses Beitrages zu klären. Dies geschieht im Folgenden in drei Schritten: Zuerst durch eine Abgrenzung zwischen starker und schwacher KI (1.), sodann werden die Begriffe maschinelle Lernverfahren (2.) und ADM (3.) näher erörtert.

1. Starke KI und schwache KI

Eine erste terminologische Eingrenzung kann aus der Differenzierung zwischen schwacher und starker KI gewonnen werden. *Schwache KI* umfasst Anwendungen, die den Menschen intelligent beim Erreichen seiner Ziele unterstützen sollen.¹¹ Dazu zählen etwa die Zeichen-, Text-, Bild- und Spracherkennung, wie sie heute beispielsweise in Smartphones eingesetzt wird. Auch die individuelle Anzeige von Werbung in Social Media, automatisierte Übersetzungsprogramme oder die Funktionen der Autovervollständigung und Korrekturvorschläge bei Suchvorgängen können genannt werden. Die *starke KI* umschreibt hingegen Phänomene, denen aktuell keine praktische Relevanz zukommt. Eine starke KI soll die gleichen intellektuellen Fertigkeiten wie Menschen erlangen oder diese sogar übertreffen.¹² Sie agiert nicht mehr nur reaktiv, sondern soll die menschliche Intelligenz nachbilden und entsprechend auch aus eigenem Antrieb handeln

⁵ Siehe *Nadja Braun Binder*, Algorithmic Regulation – Der Einsatz algorithmischer Verfahren im staatlichen Steuerungskontext, in: Hermann Hill/Joachim Wieland (Hrsg.), *Zukunft der Parlamente*, Berlin 2018, 107–120; vgl. aber die Überlegungen bei *Wolfgang Hoffmann-Riem*, Verhaltenssteuerung durch Algorithmen – Eine Herausforderung für das Recht, *AöR* 2017 1–42, zur «Konkurrenz» von algorithmischer und rechtsnormativer Steuerung.

⁶ *Jerry Kaplan*, *Artificial Intelligence. What everyone needs to know*, New York 2016, 13 f.

⁷ *Erhard Konrad*, Zur Geschichte der Künstlichen Intelligenz in der Bundesrepublik Deutschland, in: Dirk Siefkes/Peter Eulenhöfer/Heike Stach/Klaus Städtler (Hrsg.), *Sozialgeschichte der Informatik*, Wiesbaden 1998, 287–296 (287).

⁸ *Wolfgang Ertel*, *Grundkurs Künstliche Intelligenz*, 4. A., Wiesbaden 2016, 1 ff.

⁹ Vgl. *Günther Görz/Ute Schmid/Ipke Wachsmuth*, Einleitung, in: Görz/Schneeberger/Schmid (Hrsg.), *Handbuch der Künstlichen Intelligenz*, 5. A., München 2014, 1–18 (13 ff.).

¹⁰ Vgl. nur etwa *Karl Leidlmair*, KI – Eine Teildisziplin der Philosophie?, in: Johannes Rettig/Karl Leidlmair (Hrsg.), 5. Österreichische Artificial-Intelligence-Tagung, *Informatik-Fachberichte* Nr. 208, Berlin u.a. 1989, 313–324.

¹¹ Bitkom e.V./Deutsches Forschungszentrum für Künstliche Intelligenz (Hrsg.), *Entscheidungsunterstützung mit Künstlicher Intelligenz*, 2017, 29, abrufbar unter <<https://bit.ly/2eZFIdU>> zuletzt besucht am 27.5.2019.

¹² Bitkom e.V./Deutsches Forschungszentrum für Künstliche Intelligenz (Fn. 11) 29.

können.¹³ Diese Form der KI ist nicht Gegenstand des vorliegenden Beitrages. Es geht im Weiteren also ausschliesslich um Anwendungen schwacher KI.

2. Maschinelle Lernverfahren

Eine zweite Eingrenzung ergibt sich aus den potenziellen Anwendungen, die für die öffentlichen Verwaltungen aus heutiger Sicht von Interesse sind.¹⁴ Diese basieren meist auf maschinellen Lernverfahren. Es gibt wiederum verschiedene Methoden maschinellen Lernens, die sich beispielsweise in überwachtes und unüberwachtes Lernen unterteilen lassen.¹⁵ Während bei ersterem der Algorithmus aus einer Reihe von vorgegebenen Beispielen eine allgemeine Funktion erlernt, liegen den Algorithmen bei letzterem keine klassifizierte Beispiele zugrunde.¹⁶ Die Algorithmen ermitteln vielmehr selbst interessante Strukturen in unklassifizierte Daten. Der informationstechnischen Differenzierung zwischen «überwacht» und «unüberwacht» lässt sich übrigens noch keine rechtlich relevante Unterteilung, etwa in «von Menschen überwacht» und «nicht von Menschen überwacht», entnehmen.

All diese maschinellen Lernverfahren benötigen eine grosse Menge an Daten, aus denen sie lernen können, ohne dass im Voraus jeder Einzelfall explizit programmiert worden wäre.¹⁷ Auch sog. künstliche neuronale Netze gehören zu den maschinellen Lernverfahren.¹⁸ Dabei geht es, grob gesagt, darum, die im Gehirn von Menschen und Tieren existierenden Netzwerke aus Nervenzellen (natürliche neuronale Netze) zu simulieren oder in Hardware nachzubauen.¹⁹ Maschinelle Lernverfahren, die auf grossen neuronalen Netzen basieren, werden auch als Deep Learning bezeichnet.²⁰

Neben einer ausreichenden Datengrundlage ist den hier erwähnten Verfahren auch gemeinsam, dass der Lösungsweg, den diese lernenden Algorithmen einschlagen, nicht determiniert ist.²¹

Die Algorithmen werden vielmehr so programmiert, dass sich das System das Problem selbst erschliesst, Muster erkennt und – gestützt auf die vorhandenen Daten – selbsttätig zu einer Lösung kommt.²² Damit geht die Schwierigkeit einher, dass der Weg, auf dem solche Systeme zu einem Ergebnis kommen, nur schwer bzw. gar nicht erklärt werden kann.²³ Dies gilt sowohl für die Ex-ante- als auch für die Ex-post-Perspektive. Es lassen sich also unter Umständen weder akkurat die Entscheidungen eines lernenden Systems vorher sagen noch im Nachhinein der Entscheidungsweg rekonstruieren.²⁴

3. ADM

Zwar werden die nachstehend (III.) skizzierten Anwendungen in der Politik,²⁵ aber auch in der Literatur²⁶ häufig mit dem Schlagwort KI umschrieben, allerdings ist dies nicht sehr aussagekräftig. KI kann, wie eben dargelegt, Verschiedenes bedeuten.²⁷ Aus rechtlicher Sicht ist bei einem staatlichen Einsatz von maschinellen Lernverfahren von besonderer Relevanz, dass dabei grosse Datenmengen nicht durch Menschen, sondern durch Maschinen in unter Umständen nicht nachvollziehbarer Art und Weise ausgewertet werden und das Resultat dieses Verarbeitungsprozesses entweder als Entscheidungsgrundlage dient oder die menschliche Entscheidung substituiert. Vereinzelt ersetzt deshalb die Umschreibung Automated Decision Making (ADM) den Begriff KI.²⁸ Jedoch er-

¹³ Werner Sesink, *Menschliche und künstliche Intelligenz. Der kleine Unterschied*, Stuttgart 1993.

¹⁴ Vgl. Ziff. III. unten.

¹⁵ Vgl. Stefan Wrobel/Thorsten Joachims/Katharina Morik, *Maschinelles Lernen und Data Mining*, in: *Handbuch der Künstlichen Intelligenz* (Fn. 9) 405–471.

¹⁶ Wrobel/Joachims/Morik (Fn. 15) 405 f.

¹⁷ Bitkom e.V./Deutsches Forschungszentrum für Künstliche Intelligenz (Fn. 11) 13.

¹⁸ Wrobel/Joachims/Morik (Fn. 15) 406.

¹⁹ Ertel (Fn. 8) 265.

²⁰ Bitkom e.V./Deutsches Forschungszentrum für Künstliche Intelligenz (Fn. 11) 13.

²¹ Stefan Kirn/Claus D. Müller-Hengstenberg, *Intelligente (Software-) Agenten: Von der Automatisierung zur Autonomie? Vorselektierung technischer Systeme*, MMR 2014 225–232 (226).

²² Stefan Kirn/Claus D. Müller-Hengstenberg, *Rechtliche Risiken autonomer und vernetzter Systeme*, Berlin/Boston 2016, 59 ff.

²³ Peter Bräutigam/Thomas Klindt, *Industrie 4.0, das Internet der Dinge und das Recht*, NJW 2015 1137–1142 (1138).

²⁴ Mario Martini, *Black Box Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz*, Wiesbaden 2019, 43 f.; Thomas Wischmeyer, *Regulierung intelligenter Systeme*, AöR 2018 1–66 (46 ff.).

²⁵ Vgl. Fn. 2 und 3, aber auch etwa den Auftrag für eine Expertenstudie von TA-Swiss zum Thema «Wenn Algorithmen für uns entscheiden: die Herausforderungen der künstlichen Intelligenz», abrufbar unter <<https://bit.ly/2JG29QG>> zuletzt besucht am 27.5.2019.

²⁶ Vgl. nur Martini (Fn. 24).

²⁷ Vgl. Ziff. II.1. und II.2. oben.

²⁸ Vgl. den von AlgorithmWatch in Zusammenarbeit mit der Bertelsmann Stiftung herausgegebenen Bericht *Automating Society*, Januar 2019, abrufbar unter <<https://bit.ly/2YnDOCz>> zuletzt besucht am 27.5.2019; Konrad Lischka/Annika Klingel, *Wenn Maschinen Menschen bewerten. Internationale Fallbeispiele für Prozesse algorithmischer Entscheidungsfindung*, Arbeitspapier der Bertelsmann Stiftung, 2017, abrufbar unter <<https://bit.ly/2opA5Vd>> zuletzt besucht am 27.5.2019.

fasst auch diese Umschreibung nicht die vollständige Problematik, denn automatisierte Entscheidungen können auch durch deterministische Computerprogramme getroffen bzw. unterstützt werden.²⁹ Darum soll es in diesem Beitrag aber nicht gehen. Im Folgenden wird der Begriff ADM deshalb in einem engeren Sinne verwendet, der lediglich automatisierte Entscheidungen bzw. Entscheidungsunterstützungen, die auf maschinellen Lernverfahren basieren, umfasst.

III. Einsatzfelder von KI und ADM in der öffentlichen Verwaltung

ADM-Verfahren werden in öffentlichen Verwaltungen weltweit eingesetzt oder getestet. Im Folgenden werden einzelne Beispiele mit dem Ziel skizziert, die damit einhergehenden rechtlichen Herausforderungen näher einzugrenzen.³⁰ Eine Analyse der in der Literatur beschriebenen Fälle lässt den Eindruck entstehen, dass ADM in der öffentlichen Verwaltung in der Schweiz nicht bzw. nur sehr zurückhaltend zum Einsatz kommt. Inwiefern dies tatsächlich zutrifft, wird in diesem Beitrag lediglich punktuell besprochen.

1. Vorausschauende Polizeiarbeit («predictive policing»)

International finden sich KI-basierte Systeme sowohl in der *personenbezogenen* vorausschauenden Polizeiarbeit als auch in der *raumbezogenen* vorausschauenden Polizeiarbeit. Einerseits geht es dabei um eine Prognose zur Gefährlichkeit oder Gefährdung einer Person und andererseits um die raumbezogene Prognose in Bezug auf die Wahrscheinlichkeit von Straftaten. Im ersten Fall wird danach gefragt, wer gefährlich

werden bzw. gefährdet sein könnte, im zweiten Fall danach, wo eine bestimmte Gefahr auftreten könnte.³¹

Vor allem in den USA setzt man die *personenbezogene vorausschauende Polizeiarbeit* ein. Zu erwähnen ist etwa die «Strategic Subject List» der Stadt Chicago. Mithilfe von Daten über soziale Kontakte von Personen soll das Risiko errechnet werden, dass eine Person beispielsweise in Bandenkriminalität involviert sein könnte.³²

Computergestützte personenbezogene vorausschauende Polizeiarbeit ist auch in der Schweiz, zumindest ansatzweise, bekannt.³³ Aufmerksamkeit hat 2018 die Software DyRiAS (Dynamisches Risiko-Analyse-System) erlangt, weil damit offenbar zwei von drei Personen falsch verdächtigt wurden.³⁴ Die Schweizer Software setzt aber, soweit ersichtlich, keine KI ein.

Als System von *raumbezogener vorausschauender Polizeiarbeit*, das auf maschinellen Lernverfahren basiert, gilt *Predpol*³⁵, welches die Polizei in US-amerikanischen Grossstädten (z.B. Los Angeles, Chicago, Seattle, Boston) verwendet. 2013 wurde die Software erstmals auch in Europa eingesetzt, und zwar in der Grafschaft Kent (England). Mit diesem System sollen Zeit und Ort von möglichen Gefahren vorhergesagt und entsprechende Polizeipatrouillen bereitgestellt werden.

Das in der Schweiz in einzelnen Kantonen eingesetzte raumbezogene Predictive-Policing-System namens *Precobs* setzt dagegen nicht auf maschinelle Lernverfahren. *Precobs* generiert Prognosen vielmehr auf Basis von vorgängig einge-

²⁹ So im Rahmen automatisierter Besteuerungsverfahren, z.B. in Deutschland, vgl. *Nadja Braun Binder*, Ausschließlich automationsgestützt erlassene Steuerbescheide und Bekanntgabe durch Bereitstellung zum Datenabruf, DSz 2016 526–535. Es ist allerdings davon auszugehen, dass die Risikomanagementsysteme, die solche automatisierten Besteuerungsverfahren kontrollieren sollen, sich auf KI stützen, vgl. *Nadja Braun Binder*, Artificial Intelligence and Taxation: Risk Management in Fully Automated Taxation Procedures, in: Timo Rademacher/Thomas Wischmeyer (Hrsg.), *Regulating Artificial Intelligence*, Wiesbaden 2019 (i.E.), Vorabversion abrufbar unter <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3293577> zuletzt besucht am 27.5.2019.

³⁰ Weitere Fallbeispiele finden sich etwa im Bericht *Automating Society* (Fn. 28) oder bei *Lischka/Klingel* (Fn. 28).

³¹ *Tobias Knobloch*, Vor die Lage kommen. Predictive Policing in Deutschland, Bertelsmann Stiftung, 2018, 13 f., 17, abrufbar unter <<https://bit.ly/2MaSLXg>> zuletzt besucht am 27.5.2019.

³² Offizielle Informationen unter <<https://bit.ly/2JCyRIW>> zuletzt besucht am 27.05.2019. Vgl. *Jessica Saunders/Priscillia Hunt/John S. Hollywood*, Predictions Put Into Practice: A Quasi-experimental Evaluation of Chicago's Predictive Policing Pilot, in: *Journal of Experimental Criminology* 12, Nr. 3, 2016 347–371.

³³ A.A. *Matthias Leese*, Predictive Policing in der Schweiz: Chancen, Herausforderungen, Risiken, in: Christian Nünlist (Hrsg.), *Bulletin 2018 zur schweizerischen Sicherheitspolitik*, 57–71 (59).

³⁴ Vgl. *Timo Grossenbacher*, Polizei-Software verdächtigt zwei von drei Personen falsch, SRF, 5.4.2018, abrufbar unter <<https://bit.ly/2WpTkRa>> zuletzt besucht am 27.5.2019.

³⁵ Unternehmenswebsite: <<https://www.predpol.com>> zuletzt besucht am 27.5.2019. Eine Beschreibung des Systems findet sich z.B. bei *George Mohler/Martin B. Short/Sean Malinowski/Mark Johnson/George E. Tita/Andrea L. Bertozzi/P. Jeffrey Brantingham*, Randomized Controlled Field Trials of Predictive Policing, in: *Journal of the American Statistical Association*, 2015, 1399–1411.

gebenen «Wenn-dann-Entscheidungen».³⁶ Das System kommt in den Kantonen Aargau, Basel-Landschaft sowie in der Stadt Zürich regulär zum Einsatz, die Kantone Zug und Zürich testen es.³⁷ Neben dem Verzicht auf KI zeichnet sich Precobs in der Schweiz durch die Beschränkung auf die Prognose von Wohnungseinbruchdiebstahl-Delikten aus. Andere Delikte werden nicht erfasst.

2. Gerichtsverfahren

Weitere prominente Beispiele des KI-Einsatzes durch den Staat finden sich im Rahmen der Beurteilung der Rückfallgefahr von Straftätern. So wird in den USA im Justizsystem etwa eine *KI-basierte Beurteilungssoftware* (COMPAS – Correctional Offender Management Profiling for Alternative Sanctions) verwendet.³⁸ Diese errechnet auf der Basis von 137 Merkmalen eine Wahrscheinlichkeit dafür, ob ein Straftäter rückfällig wird oder nicht. Aufmerksamkeit kam dem System zuteil, weil dieses das Rückfallrisiko von farbigen Straftätern durchwegs höher als dasjenige von weissen Straftätern eingestuft hatte.³⁹ Diese Einstufung war allerdings nachweislich falsch.⁴⁰

In der Schweiz wird im Strafvollzug anhand des Programms ROS (Risikoorientierter Sanktionenvollzug) die Möglichkeit von Vollzugslockerungen geprüft. Ziel ist es, das Rückfallrisiko während und nach dem Vollzug zu senken.⁴¹ ROS wurde zwischen 2010 und 2013 in einem vom Bundesamt für Justiz unterstützten Modellversuch in den Kantonen Luzern, St. Gallen, Thurgau und Zürich getestet und später in der ganzen

Deutschschweiz umgesetzt.⁴² ROS überführt Daten zu einer Person aus dem Strafregisterauszug in das sog. Fall-Screening-Tool (FaST). Es resultiert eine Einteilung in drei Risikokategorien, die als Basis für den Entscheid dient, ob eine weitere Begutachtung notwendig ist. Es gibt allerdings keine Anhaltspunkte dafür, dass die IT-Unterstützung bei ROS auf KI basiert.

3. Chatbots

Einige KI-Anwendungen in der öffentlichen Verwaltung betreffen weit weniger sensible Bereiche. Sprachbasierte Assistenten können z.B. dabei helfen, Bürgeranfragen zu beantworten. So etwa der digitale Sachbearbeiter/die digitale Sachbearbeiterin «Amelia» aus dem Bezirk North London Borough. Die rund 330 000 im Bezirk wohnhaften Personen können ihre Fragen an die Bezirksverwaltung an Amelia richten. Das System beantwortet monatlich 55 000 Anrufe und 100 000 Anfragen über das Internet und unterstützt die Einwohnerinnen und Einwohner dabei, Bescheinigungen zu erhalten, Ausweise und Autokennzeichen zu beantragen etc. Amelia füllt entsprechende Formulare für die Bürgerinnen und Bürger aus und beantwortet Fragen. Das System soll die Qualität für Bürgerdienste zu geringeren Kosten als bisher sicherstellen. Nach drei Monaten Training konnte Amelia 64 Prozent aller Anfragen erfolgreich erledigen. Für Anrufende sank die Wartezeit von 55 auf 2 Sekunden, die Anfragen waren in durchschnittlich 4,5 statt früher 18 Minuten erledigt.⁴³

Auch in der Schweiz überlegt man, die Verwaltungskommunikation durch sog. Chatbots zu unterstützen, etwa im konsularischen Dienst oder – seitens der Parlamentsdienste – im Vorfeld der Nationalratswahlen 2019. Der Chatbot «Parli» soll Fragen rund um die Wahlen und den Wahlvorgang klären.⁴⁴ Allerdings bietet Parli keine Möglichkeit, eigene Fragen zu formulieren. Nutzerinnen und Nutzer können lediglich aus vorgegebenen Optionen entsprechende Themen auswählen. Es handelt sich mithin nicht um eine KI-basierte Anwendung.

³⁶ Vgl. die Beschreibung von Precobs bei *Dominik Gerstner*, Predictive Policing als Instrument zur Prävention von Wohnungseinbruchdiebstahl, *forschung* aktuell Nr. 50, 2017 19 f.

³⁷ Vgl. die Angaben auf der Website des Instituts für musterbasierte Prognosetechnik: <<https://www.ifmpt.de>> zuletzt besucht am 27.5.2019.

³⁸ Unternehmenswebsite: <<http://www.equivant.com>>. COMPAS wurde ursprünglich von der Firma Northpointe entwickelt; im Zuge von Kritik an COMPAS wurde die Firma in Equivant umbenannt.

³⁹ Vgl. *Julia Angwin/Jeff Larson/Surya Mattu/Lauren Kirchner*, Machine Bias – There's software used across the country to predict future criminals. And it's biased against blacks, 2016, abrufbar unter <<https://bit.ly/1XMKh5R>> zuletzt besucht am 27.5.2019. Ausserdem war COMPAS Gegenstand eines Gerichtsverfahrens, vgl. *Wisconsin Supreme Court, State v. Loomis*, Urteil vom 13.7.2016, Az. 2015AP157-CR.

⁴⁰ Vgl. auch Ziff. IV.2.

⁴¹ Vgl. *Anna Vegh*, Künstliche Intelligenz in der Strafzumessung, in: APARIUZ Nr. 20, Zürich/St. Gallen 2019 (i.E.).

⁴² Vgl. <<https://www.rosnet.ch>> zuletzt besucht am 27.05.2019. Vgl. auch *Daniel Treuthardt/Mirjam Loewe-Baur/Melanie Kröger*, Der Risikoorientierte Sanktionenvollzug (ROS) – aktuelle Entwicklungen, SZK 2, 2018 24–32, sowie *Mirjam Loewe-Baur*, Der Risikoorientierte Sanktionenvollzug (ROS), Zürich, 2017.

⁴³ Bitkom e.V./Deutsches Forschungszentrum für Künstliche Intelligenz (Fn. 11) 51.

⁴⁴ <<https://www.parlament.ch/Documents/parli-d.html>> zuletzt besucht am 27.5.2019.

4. Weitere Beispiele

Weitere Beispiele für KI-Anwendungen in der öffentlichen Verwaltung umfassen den Einsatz von Spracherkennungssoftware, sei es im Rahmen der Büroautomation (Erstellung von Akten und Dokumenten mittels Spracheingabe), sei es in Form einer stimmbiometrischen Lösung zur Erkennung der Herkunftsregion von Flüchtlingen.⁴⁵ Im Bereich des privaten und öffentlichen Verkehrs kann KI zur Optimierung und Steuerung des Verkehrs (sog. intelligente Verkehrsbeeinflussungsanlagen) oder zur Parkraumüberwachung zum Einsatz kommen.⁴⁶ Die deutsche Bundespolizei erprobte vom 1. August 2017 bis 31. Juli 2018 verschiedene, u.a. KI-basierte Systeme zur automatischen Gesichtserkennung am Bahnhof Berlin Südkreuz.⁴⁷ Im Mai 2019 war in den Medien dagegen zu lesen, dass die Stadt San Francisco die bis dahin von der Polizei eingesetzten Gesichtserkennungssysteme als erste US-amerikanische Stadt verboten hatte.⁴⁸ Hintergrund waren der starke Widerstand von Zivilrechtsorganisationen und die Angst vor einem repressiven Überwachungsstaat. Hinweise auf den staatlichen Einsatz von KI finden sich zudem im Zusammenhang mit der Bekämpfung von Wirtschaftskriminalität⁴⁹ oder im Kontext des Risikomanagements in Besteuerungsverfahren.⁵⁰

IV. Offene Rechtsfragen

Der staatliche Einsatz von KI evoziert zahlreiche Rechtsfragen,⁵¹ die je nach Anwendungsbereich unterschiedlich ausfallen. So stellen sich im Kontext staatlich eingesetzter Chatbots andere als im Bereich einer automatisierten Gesichtserkennung. Solche spezifischen Fragen werden hier ausgeklammert. Doch sind KI- bzw. ADM-Systemen gewisse Herausforderungen unabhängig vom jeweiligen Anwendungsbereich gemeinsam. Diese ergeben sich aus dem Umstand, dass die im Rahmen maschineller Lernverfahren eingesetzten Algorithmen nicht mehr ohne Weiteres nachvollziehbar sind (1), dass die maschinellen Lernverfahren vor einem Einsatz trainiert werden müssen (2) und dass grosse Datenmengen verarbeitet werden (3). Eine weitere Herausforderung stellt sodann die Gewährleistung des rechtlichen Gehörs dar, wenn Entscheidungen vollautomatisiert – also ohne jedes menschliche Zutun – von der KI getroffen werden (4).

1. Nachvollziehbarkeit bzw. Begründungserfordernisse

Die mangelnde bzw. eingeschränkte Nachvollziehbarkeit maschineller Lernverfahren ist vor allem mit Blick auf das rechtsstaatliche Begründungserfordernis eines Rechtsaktes bzw. Gerichtsurteils von Bedeutung. Die Begründungspflicht folgt aus dem Anspruch auf rechtliches Gehör (Art. 29 Abs. 2 BV).⁵² Die Begründung dient dazu, Transparenz bezüglich der Entscheidungsgründe herzustellen. Sie soll es den Betroffenen ermöglichen, gegen den behördlichen Entscheid ein sachgerechtes Rechtsmittel einzureichen. Trifft eine KI die Entscheidung oder zumindest wesentliche vorbereitende Teile davon, können die internen Schritte, die innerhalb der Maschine stattfinden, nicht nachvollzogen werden. Daraus wird in der Literatur bisweilen der Schluss gezogen, eine den rechtsstaatlichen Anforderungen genügende Begründung sei nicht mehr möglich.⁵³ Andere Stimmen betonen aber, dass auch bei der Begründung einer von einem Behördenmitarbeitenden getroffenen Entscheidung die Abläufe der einzelnen Gedankengänge der Person nicht wiedergegeben werden müssen. Genauso wenig müssten die einzelnen maschinellen Schritte in der Begründung dargestellt werden. Auch wenn der Herstel-

⁴⁵ Vgl. *Cancom.info* (Redaktion), Dragon Nuance – Spracherkennung in Behörden, 29.1.2019, abrufbar unter <<https://bit.ly/2DLzXrn>> zuletzt besucht am 27.5.2019.

⁴⁶ Vgl. für verschiedene Beispiele den Aktionsplan «Digitalisierung und Künstliche Intelligenz in der Mobilität» des deutschen Bundesministeriums für Verkehr und digitale Infrastruktur (BMVI) vom November 2018, abrufbar unter <<https://bit.ly/2EDilZA>> zuletzt besucht am 27.5.2019, sowie die Liste der KI-Projekte des BMVI unter <<https://bit.ly/2YRVbvv>> zuletzt besucht am 27.05.2019.

⁴⁷ Vgl. Bundespolizeipräsidium Potsdam: Abschlussbericht zum Teilprojekt 1 «Biometrische Gesichtserkennung» vom 18.9.2018, abrufbar unter <<https://bit.ly/2F9QFth>> zuletzt besucht am 27.5.2019.

⁴⁸ *Kate Conger/Richart Fausset/Serge F. Kovaleski*, San Francisco Bans Facial Recognition Technology, *New York Times*, 14.5.2019, abrufbar unter <<https://nyti.ms/2Ec6SFr>> zuletzt besucht am 27.5.2019.

⁴⁹ Vgl. z.B. die verschiedenen Beiträge auf der Herbsttagung des deutschen Bundeskriminalamtes zum Thema «Sicherheit in einer offenen und digitalen Gesellschaft» vom 21./22.11.2018, abrufbar unter <<https://bit.ly/2K7i61G>> zuletzt besucht am 27.5.2019.

⁵⁰ *Braun Binder* (Fn. 29).

⁵¹ Allgemein zu Rechtsfragen im Kontext der Digitalisierung der Verwaltungstätigkeit: *Andreas Glaser*, Einflüsse der Digitalisierung auf das schweizerische Verwaltungsrecht, *SJZ* 2018 181–190; *Ders.*, Der elektronisch handelnde Staat, *ZSR* 2015 II 259–333.

⁵² Zum rechtlichen Gehör auch *Ziff.* IV.4. unten.

⁵³ Zumind. zweifelnd: *David Rechsteiner*, Der Algorithmus verfügt, in: *Jusletter* 26.11.2018, Rz. 26.

lungsprozess einer auf maschinellen Lernverfahren basierenden Entscheidung nicht im Einzelnen nachvollzogen werden könne, müsse dies nicht bedeuten, dass KI keine Entscheidungen treffen dürfe. Es stelle sich aber die Frage, ob bzw. wie KI künftig in der Lage sein werde, die Entscheidung in einer den rechtsstaatlichen Anforderungen an Begründungen genügenden Form darzustellen.⁵⁴ Derzeit entwickelt sich unter dem Schlagwort *Explainable AI* (XAI) ein neues Forschungsgebiet, mit dem Ziel, die in maschinellen Lernverfahren eingesetzten Algorithmen transparent und nachvollziehbar zu machen.⁵⁵ Ob damit inskünftig die rechtsstaatlichen Begründungserfordernisse sichergestellt werden können, ist allerdings noch offen.

Zwischen diesen beiden Extrempositionen – keine KI, da keine nachvollziehbare Begründung möglich sei, bis hin zu einer ausschliesslich durch die KI getroffenen Entscheidung bei entsprechender Begründung durch die KI selbst – liegt die Konstellation, in der KI zur Entscheidungsvorbereitung genutzt wird und die Entscheidung anschliessend durch einen Menschen getroffen wird. Diesfalls muss die entscheidende Person in der Lage sein, die wesentlichen Entscheidungsgründe darzulegen, auch wenn die Funktionsweise der genutzten maschinellen Lernverfahren nicht im Einzelnen erklärt werden kann. Handelt es sich um eine staatliche Entscheidung, die in vergleichbarer Weise auch andere Personen betreffen kann, wären als Mindestanforderung Angaben über die Vergleichsgruppe, in welche der Algorithmus eine Person einordnet, zu verlangen.⁵⁶ Ferner sollte klar sein, welche individuellen Besonderheiten im Einzelfall für den Entscheid ausschlaggebend waren.⁵⁷

Insgesamt ist festzuhalten, dass staatliche Entscheidungen auch dann zu begründen sind, wenn diese auf Basis einer KI-Empfehlung bzw. von der KI selbst getroffen werden. Eine Einschränkung der Begründungserfordernisse, die sich allein

aus der schwierigeren Nachvollziehbarkeit maschineller Lernverfahren ergibt, ist abzulehnen.⁵⁸

2. Trainingsdaten

Lernende Algorithmen müssen in der einen oder anderen Weise trainiert werden.⁵⁹ Die Trainingsdaten, die ein maschinelles Lernverfahren nutzt, sind demnach prägend für die spätere Funktionsweise des Systems. Sie dienen quasi als Referenz für künftige Differenzierungen und unterstellen zu diesem Zweck, dass die Trainingsdaten korrekt und die auf dieser Basis getroffenen Entscheidungen richtig waren.⁶⁰ Falsche Trainingsdaten können demgegenüber dazu führen, dass das System unkorrekte Resultate liefert. Ein anschauliches Beispiel hierfür liefert ein am Massachusetts Institute of Technology durchgeführtes Experiment mit Bilderkennungsprogrammen: Ein System wurde mit «normalen» Daten trainiert, das andere mit Daten aus dem Darknet. Letztgenanntes System «erkannte» daraufhin auf Bildern, die das andere als Vögel oder Blumen interpretierte, ausschliesslich Gewaltszenen.⁶¹

Es besteht ausserdem das Risiko, dass das System historisch gewachsene Vorurteile, die sich in den Trainingsdaten niederschlagen, übernimmt und als «objektives Faktum» behandelt.⁶² Auch wenn die genutzten Daten korrekt und geeignet sind, können sie zu diskriminierenden Resultaten führen.⁶³ Dies zeigt etwa das Beispiel des österreichischen Arbeitsmarktservice. Seit 2019 kategorisiert dort ein Algorithmus Personen nach ihren Chancen auf dem Arbeitsmarkt. Diese werden drei Personengruppen zugeordnet; solchen mit guten, mittleren und schlechten Perspektiven auf dem Arbeitsmarkt.⁶⁴ Ziel ist es, dass der Arbeitsmarktservice Österreich

⁵⁸ So im Resultat wohl auch *Wischmeyer* (Fn. 24) 61 ff.

⁵⁹ Vgl. Ziff. II.2. oben.

⁶⁰ *Martini* (Fn. 24) 50.

⁶¹ Vgl. *Jane Wakefield*, *Are you scared yet? Meet Norman, the psychopathic AI*, BBC News, 2.6.2018, abrufbar unter <<https://www.bbc.com/news/technology-44040008>> zuletzt besucht am 27.5.2019.

⁶² Vgl. Bericht der Expertengruppe zur Zukunft der Datenbearbeitung und Datensicherheit (EZDD) vom 17.8.2018, 38 f., abrufbar unter <<https://bit.ly/2N2qD8N>> zuletzt besucht am 27.5.2019.

⁶³ *Wiebke Fröhlich/Indra Spiecker genannt Döhmann*, *Können Algorithmen diskriminieren?* VerfBlog, 2018/12/26, abrufbar unter <<https://bit.ly/2KLq2X0>> zuletzt besucht am 27.5.2019.

⁶⁴ *Jürgen Holl/Günter Kernbeiß/Michael Wagner-Pinter*, *Das AMS-Arbeitsmarktchancen-Modell*, Dokumentation zur Methode, Oktober 2018, abrufbar unter <<https://bit.ly/2AiFrbI>> zuletzt besucht am 27.5.2019; *Barbara Wimmer*, *Der AMS Algorithmus ist ein «Paradebeispiel für Diskriminierung»*, *futurezone*, 17.10.2018, abrufbar unter <<https://bit.ly/31oY7C1>> zuletzt besucht am 27.5.2019.

⁵⁴ Zuversichtlich: *Stephan Meyer*, *Künstliche Intelligenz und die Rolle des Rechts für Innovation*, ZRP 2018 233–238 (237). Vgl. auch *Finale Doshi-Velez/Mason Kortz*, *Accountability of AI Under the Law: The Role of Explanation*, Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society working paper, 2017, abrufbar unter <<https://dash.harvard.edu/handle/1/34372584>> zuletzt besucht am 27.5.2019.

⁵⁵ Vgl. z.B. den Hinweis in *Laura Delponte*, *European Artificial Intelligence (AI) leadership, the path for an integrated vision*, Study Requested by the ITRE committee, European Parliament, September 2018, abrufbar unter <<https://bit.ly/2wIME28>> zuletzt besucht am 27.5.2019.

⁵⁶ *Martini* (Fn. 24) 190.

⁵⁷ *Martini* (Fn. 24) 190.

(AMS) seine Ressourcen ab 2020 überwiegend auf Personen der mittleren Gruppe konzentriert, da sie dort am effektivsten eingesetzt seien. Problematisch ist, dass der Algorithmus die Arbeitsmarktchancen von Frauen pauschal als schlecht einstuft. Hat eine Frau Kinder, fällt die Beurteilung noch schlechter aus.⁶⁵ Diese Resultate erzielt das System nicht aufgrund unkorrekter oder ungeeigneter Daten. Vielmehr spiegeln die Daten die gesellschaftliche Realität wider.⁶⁶ In der negativen Bewertung des «Frau-Seins» werden die strukturell bedingten Nachteile der Frauen im Erwerbsleben reproduziert. Folgt der AMS solchen Empfehlungen, werden damit bestehende Missstände festgeschrieben bzw. noch verstärkt.

Das ethnische Diskriminierungspotenzial, insbesondere von Anwendungen in der vorausschauenden Polizeiarbeit und im Bereich der Risikobeurteilung von Straftätern, ist in der Literatur denn auch ein zentrales Thema. Im Bereich der vorausschauenden Polizeiarbeit bezieht sich die Diskussion sowohl auf die personen- als auch auf die raumbezogenen Systeme. Bei personenbezogenem KI-Einsatz ist es deshalb von besonderer Bedeutung sicherzustellen, dass keine Vorurteile gegenüber gewissen Bevölkerungsgruppen oder Minderheiten entstehen bzw. reproduziert werden.⁶⁷ Aber auch mit Blick auf rein ortsbezogene Systeme können Bedenken auftreten – dies etwa dann, wenn der Einsatz von Polizeikräften zur Überrepräsentanz in ärmeren Gegenden mit einem hohen Anteil bestimmter Bevölkerungsgruppen führt. Dieses Risiko besteht unter Umständen unabhängig von Daten wie Einkommen oder Hautfarbe. Entsprechende Faktoren können in ethnisch stark segregierten Gebieten mit der Postleitzahl übereinstimmen.⁶⁸

Die bereits erwähnte KI-basierte Beurteilungssoftware COMPAS wertete Antworten aus einem Fragebogen aus, in dem nicht nach der Hautfarbe gefragt worden war.⁶⁹ Dennoch kam das System zu nachweislich falschen Schlussfolgerungen.⁷⁰

Aktuell wird in der wissenschaftlichen Literatur diskutiert, ob diese falschen Schlussfolgerungen als Fehler des Systems einzuordnen sind oder vielmehr eine unvermeidbare Folge der zugrunde liegenden Daten darstellen. Letzteres wird etwa damit erklärt, dass an der Rasse orientierte unterschiedliche Rechtsdurchsetzungs- oder Strafverfolgungspraxen, die sich in den Daten der beschuldigten Person widerspiegeln, vom Algorithmus übernommen werden.⁷¹

Die Betrachtung der genannten Beispiele zeigt, dass bereits in der Trainingsphase die Qualität der verwendeten Daten sicherzustellen ist. Die Daten müssen richtig, vollständig, aktuell und ausreichend differenziert sein.⁷² Dies geht in zweierlei Hinsicht über die datenschutzrechtliche Vorgabe des Art. 5 DSGVO⁷³ hinaus, wonach, wer Personendaten bearbeitet, sich über deren Richtigkeit vergewissern muss. Zum einen geht es neben der Richtigkeit der Daten auch um deren Vollständigkeit, Aktualität etc. Zum anderen erstrecken sich die Anforderungen auf nicht personenbezogene Daten. In der Literatur findet sich die Empfehlung, dass im Idealfall sowohl beim Betrieb als auch im Rahmen der Trainingsphase auf Daten der öffentlichen Hand zurückgegriffen werden soll.⁷⁴ Dies schafft allerdings nur dort Abhilfe, wo die öffentliche Hand überhaupt über einschlägige Daten verfügt und wo diese Daten korrekt sind. Setzt der Staat KI ein, hat er über die Nutzung eigener Daten hinaus sicherzustellen, dass das Trainingsmaterial geeignet ist. Dies wird besonders dort relevant, wo der Staat keine eigenen Systeme entwickelt, sondern diese einkauft. Im Rahmen des Beschaffungsvorganges ist demnach unter anderem darauf zu achten, dass der Hersteller eine geeignete Trainingsumgebung nutzt.

Aber auch wenn die Daten richtig, vollständig, aktuell und ausreichend differenziert sind, kann ein in den Daten vorhandener «Bias» im Rahmen von KI bzw. ADM reproduziert oder sogar noch verstärkt werden. Es besteht also auch bei korrekten Daten Bedarf an der Kontrolle der Entscheidung bzw. Ent-

⁶⁵ Fröhlich/Spiecker gen. Döhmman (Fn. 63); Wimmer (Fn. 64).

⁶⁶ Fröhlich/Spiecker gen. Döhmman (Fn. 63); Wimmer (Fn. 64).

⁶⁷ Kritische Perspektiven: Mark Andrejevic, To Preempt a Thief, in: International Journal of Communication 11, 2017 879–896; Peter Mantello, The Machine That Ate Bad People: The Ontopolitics of the Pre-crime Assemblage, in: Big Data & Society 2016 1–11; Mireille Hildebrandt, New Animism in Policing: Re-animating the Rule of Law?, in: Bradford et al. (Hrsg.), The SAGE Handbook of Global Policing, London 2016, 406–428. Kritisch zur Gesichtserkennung im öffentlichen Raum: Big Brother Watch, Defending Civil Liberties, Protecting Privacy, 2018, abrufbar unter <<https://bit.ly/2wPNry3>> zuletzt besucht am 27.5.2019.

⁶⁸ Knobloch (Fn. 31) 12.

⁶⁹ Ein Muster des Fragebogens mit 137 Fragen ist online verfügbar unter <<https://bit.ly/2mqoxBw>> zuletzt besucht am 27.5.2019.

⁷⁰ Angwin/Larson/Mattu/Kirchner (Fn. 39); Frank Pasquale, Secret Algorithms Threaten the Rule of Law, MIT Technology Review 2017, abrufbar unter <<https://bit.ly/2lemv1H>> zuletzt besucht am 27.5.2019. Kritisch zur Studie von Angwin/Larson/Mattu/Kirchner: Anthony Flores/Kristin Bechtel/Christopher T. Lowenkamp, False Positives, False Negatives, and False Analyses, 2016, abrufbar unter <<https://bit.ly/2wSH39r>> zuletzt besucht am 27.5.2019.

⁷¹ Vgl. Megan Stevenson, Assessing Risk Assessment in Action, George Mason University Legal Studies Research Paper Series, LS 17-25, 2017, 26, abrufbar unter <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3016088> zuletzt besucht am 27.5.2019.

⁷² Vgl. Bericht EZDD (Fn. 62) 38.

⁷³ Bundesgesetz über den Datenschutz (DSG) vom 19.6.1992, SR 235.1.

⁷⁴ So z.B. die Empfehlung von Knobloch (Fn. 31) 23.

scheidungsempfehlung. In der Literatur finden sich die Vorschläge, sog. «dummy data» zu nutzen, Kontrollalgorithmen einzusetzen oder regelmässige Audits durchzuführen.⁷⁵

3. Datenbearbeitung

Maschinelle Lernverfahren sind auf Daten angewiesen, die analysiert werden und auf deren Basis sodann eine Prognose oder Entscheidung getroffen wird. Es gilt die allgemeine Erkenntnis, dass die Ergebnisse von Datenverarbeitungsverfahren immer nur so gut sein können, wie es die Qualität, Aktualität und Verfügbarkeit der Daten ermöglichen.⁷⁶ Dies trifft nicht nur auf die Trainingsdaten zu, sondern auch auf jene Daten, die im Laufe des ordentlichen Betriebs in ein System eingespeist und verarbeitet werden. Diesbezüglich kann auf die Ausführungen unter Ziff. IV.2. verwiesen werden.

Im Zusammenhang mit dem Einsatz von KI, insbesondere im Rahmen automatisierter Entscheide, werden häufig personenbezogene Daten verarbeitet. So kann KI z.B. im Rahmen von personenbezogener vorausschauender Polizeiarbeit oder bei der Beurteilung der Rückfallgefahr von Straftätern eine Prognose über die Gefährlichkeit einer Person liefern. In diesem Zusammenhang sind der verfassungsrechtlich im Rahmen des Schutzes der Privatsphäre (Art. 13 BV) verbürgte Schutz vor dem Missbrauch persönlicher Daten (Art. 13 Abs. 2 BV) sowie weitere Persönlichkeitsrechte relevant. In der Praxis steht der Umsetzung der Datenschutzrechte allerdings häufig die Schwierigkeit im Weg, dass die betroffene Person gar nicht weiss, welche Daten über sie gespeichert bzw. im Rahmen einer KI-gestützten Verarbeitung genutzt werden. Abhilfe könnte eine Überarbeitung des Datenschutzgesetzes schaffen. In seinem Entwurf für ein neues Datenschutzgesetz vom September 2017⁷⁷ sieht der Bundesrat in Art. 19 Abs. 1 E-DSG eine Informationspflicht im Falle von automatisierten Einzelentscheidungen vor. Abs. 4 enthält eine Sonderregelung

für Bundesorgane: Diese sollen automatisierte Einzelentscheidungen – gemeint sind Verfügungen⁷⁸ – entsprechend kennzeichnen.⁷⁹ Die Kennzeichnungspflicht von KI-gestützten Entscheidungen ist unbedingt zu befürworten. Nur so erhält die von einer Entscheidung betroffene Person die Möglichkeit, die Korrektheit der Entscheidung unter besonderer Berücksichtigung der Eigenheiten maschineller Lernverfahren zu hinterfragen.

Diese Regelung stellt allerdings eine Minimalvorgabe dar. Sie würde noch keine Klarheit über die genutzten Daten schaffen, sondern lediglich einen ersten Anhaltspunkt für betroffene Personen bieten, weitere Informationen einzuholen. Es ist deshalb gerade mit Blick auf den Einsatz von KI durch die Verwaltung der Grundsatz zu betonen, dass der Staat Personendaten nur auf Basis einer gesetzlichen Grundlage bearbeiten darf.⁸⁰ Indirekt ist damit die Transparenz zwar sichergestellt. Diese könnte – im Sinne einer bürgerfreundlichen Handhabung – allerdings insofern ausgeweitet werden, als eine Kennzeichnungspflicht KI-basierter Entscheidungen auch Angaben über die Daten enthalten sollte, die im Rahmen der maschinellen Lernverfahren genutzt wurden.

Ferner ist daran zu erinnern, dass Personendaten nur zu dem Zweck bearbeitet werden dürfen, der bei der Beschaffung angegeben wurde bzw. aus den Umständen ersichtlich oder gesetzlich vorgesehen ist.⁸¹ Dies bedeutet mit Blick auf den staatlichen KI-Einsatz, dass bestehende Datenbanken nicht ohne Weiteres für maschinelle Lernverfahren miteinander verknüpft werden dürfen.

Die bereits erwähnte Regelung in Art. 19 Abs. 4 E-DSG würde sich lediglich auf den engen Anwendungskreis von Entscheidungen beziehen, die ausschliesslich auf einer automatisierten Bearbeitung beruhen. Dies ist der Fall, «wenn keine inhaltliche Bewertung und darauf gestützte Entscheidung durch eine natürliche Person stattgefunden hat. Das heisst, die inhaltliche Beurteilung des Sachverhalts, auf dem die Entscheidung beruht, erfolgte ohne Zutun einer natürlichen Person. Darüber hinaus wird auch der Entscheid, der auf der Basis dieser Sachverhaltsbeurteilung ergeht, nicht von einer natürlichen Person getroffen.»⁸² Mit anderen Wor-

⁷⁵ Martini (Fn. 24) 225 ff., 249 ff., 274 ff.; Wischmeyer (Fn. 24) 61 ff. Vgl. auch Mario Martini/David Nink, Wenn Maschinen entscheiden ... – vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz, NVwZ-extra, 2017/10 1–14, 9 ff., abrufbar unter <https://rsw.beck.de/rsw/upload/NVwZ/NVwZ-Extra_2017_10.pdf> zuletzt besucht am 27.05.2019.

⁷⁶ So mit Blick auf Predictive Policing Knobloch (Fn. 31) 21; Leese (Fn. 33) 67. Vgl. zu den rechtlichen Herausforderungen von «Big Data», u.a. im prädiktiven Bereich, auch die verschiedenen Beiträge in: Wolfgang Hoffmann-Riem (Hrsg.), Big Data – Regulative Herausforderungen, Baden-Baden 2018.

⁷⁷ Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz, BBl 2017 7193 (7215).

⁷⁸ BBl 2017 6941 (7059).

⁷⁹ Art. 19 Abs. 4 E-DSG lautet: «Ergeht die automatisierte Einzelentscheidung durch ein Bundesorgan, so muss es die Entscheidung entsprechend kennzeichnen. Absatz 2 gilt nicht, wenn der betroffenen Person gegen die Entscheidung ein Rechtsmittel zur Verfügung steht.»

⁸⁰ Für Organe des Bundes vgl. Art. 17 Abs. 1 DSG.

⁸¹ Für Bundesorgane vgl. Art. 4 Abs. 3 DSG.

⁸² BBl 2017 6941 (7056 f.).

ten, die vorgeschlagene Regelung in Art. 19 Abs. 4 E-DSG würde im Falle des KI-Einsatzes nur in Situationen einer von der KI autonom getroffenen Entscheidung greifen. In der Praxis sind allerdings Verfahren wahrscheinlicher, in welchen die KI keine vollautomatisierten Entscheidungen trifft, sondern diese lediglich vorbereitet (Entscheidungsempfehlung). Das Bedürfnis, KI-basierte Entscheidungen hinterfragen zu können, besteht aber auch im Falle von KI-basierten Entscheidungsempfehlungen. Die oben unter Ziff. IV.2. angeführten Beispiele des österreichischen Arbeitsmarktservice und des Systems COMPAS zeigen deutlich, dass auch KI-basierte Entscheidungsempfehlungen problematisch sein können. Die Kennzeichnungspflicht sollte sich mithin auf den KI-Einsatz nicht vollständig automatisierter Verfahren erstrecken.

4. Anspruch auf rechtliches Gehör bei vollautomatisierten Entscheidungen

Wird ein KI-Algorithmus so eingesetzt, dass dieser unmittelbar selbst eine verbindliche Entscheidung (Verfügung) trifft, dann besteht die Gefahr, dass der Anspruch auf rechtliches Gehör beeinträchtigt wird. Der Anspruch auf rechtliches Gehör ist Bestandteil der allgemeinen Verfahrensgarantien (Art. 29 Abs. 2 BV) und bedeutet insbesondere, dass eine Person das Recht hat, sich zu äussern, bevor eine sie nachteilig treffende Entscheidung gefällt wird. In der Literatur wird deshalb vorgeschlagen, vollautomatisierte Verfügungen grundsätzlich nur dort zuzulassen, wo den Begehren der betroffenen Person entsprochen wird.⁸³ Eine Ausnahme stellen Verfahren dar, für welche das Verfahrensrecht die Möglich-

keit einer Einsprache vorsieht. Alternativ könnte die Äusserungsmöglichkeit bei Begehren, die elektronisch gestellt werden, in Form eines sog. Freitextfeldes vorgesehen werden. Sobald ein Eintrag in diesem Freitextfeld gemacht wird, würde dies dazu führen, dass das Begehren nicht mehr vom Algorithmus, sondern von einem Menschen, unter Berücksichtigung der getätigten Äusserung, zu entscheiden ist.⁸⁴

V. Ausblick

Die Entwicklungen rund um den Einsatz von KI und ADM in öffentlichen Verwaltungen weltweit lassen es als wahrscheinlich erscheinen, dass künftig auch in schweizerischen öffentlichen Verwaltungen vermehrt auf maschinelle Lernverfahren zurückgegriffen wird. Die erwähnten Beispiele und skizzierten Rechtsfragen belegen, dass dabei neue Herausforderungen auf den Staat zukommen. So spielt insbesondere die Auswahl des Trainingsmaterials für die KI eine zentrale Rolle. Damit können wichtige Weichen für einen sachgerechten und rechtmässigen KI-Einsatz gestellt werden. Der Gesetz- bzw. Verordnungsgeber wird sich bei der Schaffung des rechtlichen Rahmens nicht nur mit Vorgaben für die Programmierung der Algorithmen oder datenschutzrechtlichen Rahmenbedingungen zu befassen haben, sondern auch mit Fragen rund um die Sicherstellung der Datenqualität, die Kontrolle KI-basierter Entscheidungen bzw. Entscheidungsempfehlungen sowie die Umsetzung der rechtsstaatlichen Begründungserfordernisse.

⁸³ Rechsteiner (Fn. 53) Rz. 23.

⁸⁴ Vgl. Braun Binder (Fn. 29); Nadja Braun Binder, Vollständig automatisierter Erlass eines Verwaltungsaktes und Bekanntgabe über Behördenportale, DÖV 2016 891–898.