

Nadja Braun Binder / Nina Laukenmann / Liliane Obrecht

KI in der Verwaltung: Entwicklungen und Herausforderungen

Der vorliegende Beitrag widmet sich den rechtlichen Aspekten eines Einsatzes von Künstlicher Intelligenz (KI) in der öffentlichen Verwaltung. Anhand von konkreten Beispielen werden mögliche Einsatzbereiche und verschiedene rechtliche Herausforderungen aufgezeigt. Neben den rechtsetzenden Entwicklungen in der Schweiz wird zudem auf internationale Regulierungsbestrebungen hingewiesen. Die Autorinnen kommen zum Schluss, dass auch in der Schweiz zumindest punktuelle rechtliche Grundlagen sinnvoll und notwendig sind, um den Herausforderungen zu begegnen und den KI-Einsatz rechtsstaatskonform auszugestalten.

Beitragsart: Beiträge

Region: Schweiz

Rechtsgebiete: Artificial Intelligence & Recht, E-Government

Zitiervorschlag: Nadja Braun Binder / Nina Laukenmann / Liliane Obrecht, KI in der Verwaltung: Entwicklungen und Herausforderungen, in: Jusletter IT 4. Juli 2024

Inhaltsübersicht

1. Einleitung
2. Begriffsverständnis
3. Ausgewählte Einsatzbereiche in der öffentlichen Verwaltung
 - 3.1. Standardisierung und Massenverwaltung
 - 3.2. Chatbots und Voicebots
 - 3.3. Intelligente Informationsdurchsuchung
 - 3.4. Prognosen
4. Rechtliche Herausforderungen
 - 4.1. Legalitätsprinzip
 - 4.2. Verfahrensgarantien und Verfahrensgrundsätze
 - 4.2.1. Rechtliches Gehör
 - 4.2.2. Untersuchungsgrundsatz
 - 4.3. Diskriminierungsverbot
 - 4.3.1. Diskriminierungsarten und Erscheinungsformen
 - 4.3.2. Beispiele algorithmischer Diskriminierung
 - 4.4. Informationelle Selbstbestimmung und Datenschutz
 - 4.4.1. Die Verfassungsbestimmung und ihre gesetzliche Konkretisierung
 - 4.4.2. Beispiel 1: Maschinelle Gesichtserkennung
 - 4.4.3. Beispiel 2: Risikomanagementsysteme in der Steuerverwaltung
 - 4.5. Vertrauensschutz
 - 4.6. Transparenz
5. Entwicklungen aus der Perspektive der Bundesverwaltung
6. Fazit

1. Einleitung

[1] Künstliche Intelligenz (KI) hat sich als Schlagwort im Kontext der digitalen Transformation etabliert. Die enormen Datenmengen und der technische Fortschritt im Bereich des maschinellen Lernens führen dazu, dass neben der Privatwirtschaft auch der Staat verschiedene KI-Anwendungen einsetzt. Das Bundesamt für Statistik BFS führt in einer Projektdatenbank des Kompetenznetzwerks für Künstliche Intelligenz (*Competence Network for Artificial Intelligence, CNAI*) über 60 Projekte auf, in denen die Bundesverwaltung KI einsetzt, Pilotprojekte durchführt oder durchgeführt hat. So setzt etwa die Informatikabteilung des Eidgenössischen Departements für auswärtige Angelegenheiten EDA¹ einen Mailbot ein. Dieser analysiert automatisiert eingehende deutschsprachige E-Mails, priorisiert sie und erstellt entsprechende *Incidenttickets* für das Support System. Dadurch konnten die personellen Ressourcen des EDA-Helpdesks entlastet werden.² Bei MeteoSchweiz werden die Daten des Pollenmessnetzes SwissPollen durch ein

¹ Die Informatikabteilung des EDA erbringt rund um die Uhr mit über 100 Mitarbeitenden IT-Dienstleistungen für das EDA auf der ganzen Welt, siehe Informatik EDA, <https://www.eda.admin.ch/eda/de/home/das-eda/arbeiten-eda/berufserfahrene/taetigkeiten-zentrale.html>. Alle in diesem Beitrag angegebenen Internetquellen wurden zuletzt am 6. Juni 2024 abgerufen.

² Siehe zum Ganzen CNAI, Projektdatenbank, Version 8.0 vom 29. Mai 2024, siehe <https://cnaai.swiss/dienstleistungen/projektdatenbank/> sowie Innovation Process Technology ipt, EDA: Schnellerer Kunden-Support durch Helpdesk Mailbot, <https://ipt.ch/de/kundenstories/success-stories/eda-schnellerer-kunden-support-durch-helpdesk-mailbot>.

KI-System ausgewertet, wodurch der Bevölkerung die Pollenmessung und die Pollenprognosen auf der MeteoSchweiz Webseite oder der App in Echtzeit zur Verfügung stehen.³

[2] Der vorliegende Beitrag gibt zunächst einen Überblick über die verschiedenen international und national verbreiteten Begriffsverständnisse von KI-Systemen (Kapitel 2.). Anschliessend werden ausgewählte Einsatzbereiche von KI in der öffentlichen Verwaltung präsentiert (Kapitel 3). Darauf aufbauend werden die rechtlichen Herausforderungen und Rahmenbedingungen skizziert (Kapitel 4). Schliesslich wird ein Blick auf die Rechtsetzung aus der Perspektive der Bundesverwaltung geworfen (Kapitel 5). Der Beitrag schliesst mit einem Fazit (Kapitel 6).

2. Begriffsverständnis

[3] Es existieren verschiedene Begriffsverständnisse zu «KI». Obwohl der Begriff «KI» bereits auf die 1950er Jahre zurückgeht,⁴ kristallisieren sich erst im Jahr 2024 verschiedene Legaldefinitionen heraus.

[4] So enthält Art. 3 Ziff. 1 des europäischen AI Act⁵ eine Legaldefinition, die sich an jener der OECD (Organisation for Economic Co-operation and Development) orientiert.⁶ Demnach handelt es sich bei einem KI-System um «ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist, das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben, wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können».

[5] Am 17. Mai 2024 hat das Ministerkomitee des Europarates die Konvention über künstliche Intelligenz verabschiedet.⁷ Diese definiert KI-Systeme in Art. 2 wie folgt: «a machine-based system that for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment.» Die Definition ist angelehnt an jene des AI Act der EU. Da die Schweiz Mitglied im Europarat ist und die Konvention voraussichtlich ratifizieren wird, wird diese Definition für die Schweiz von Bedeutung sein.

[6] Neben den beiden vorgenannten Legaldefinitionen ist auf die Terminologie des vom Bundesrat 2022 eingesetzten CNAI⁸ hinzuweisen. Das CNAI versteht unter einem KI-System «ein

³ Bundesamt für Meteorologie und Klimatologie MeteoSchweiz, Automatisches Pollenmessnetz SwissPollen, <https://www.meteoschweiz.admin.ch/wetter/messsysteme/bodenstationen/automatisches-pollenmessnetz-swisspollen.html>. Detailliertere Informationen finden sich in der Projektdatenbank des CNAI (Fn. 2), S. 26.

⁴ JERRY KAPLAN, *Artificial Intelligence. What everyone needs to know*, New York 2016, S. 13 f.

⁵ Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828, Version PE-CONS 24/1/24 REV 1 vom 13.06.2024 abrufbar unter <https://data.consilium.europa.eu/doc/document/PE-24-2024-REV-1/DE/pdf>.

⁶ Zur Definition STUART RUSSELL/KARINE PERSET/MARKO GROBELNIK, Updates to the OECD's definition of an AI system, <https://oecd.ai/en/wonk/ai-system-definition-update>.

⁷ Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 17. Mai 2024, CM(2024)52-final, https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680afb11f.

⁸ Siehe Kapitel 1.

maschinenbasiertes System, das für explizite oder implizite Ziele aus den empfangenen Inputs schlussfolgert, wie es Outputs wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erzeugen kann, welche die physische oder virtuelle Umgebung beeinflussen können. KI-Systeme können mit unterschiedlichem Ausmass an Autonomie ausgestattet werden.»⁹ Bei dieser Definition handelt es sich allerdings nicht um eine Legaldefinition.

[7] Die drei vorstehend erwähnten Definitionen beziehen sich auf KI-Systeme. Allerdings können auch bei algorithmenbasierten Systemen, die nicht unter die Definition eines KI-Systems fallen, dieselben rechtlichen Herausforderungen entstehen.¹⁰ Der Einheitlichkeit halber wird im Folgenden dennoch der Begriff des KI-Systems verwendet. Ferner ist darauf hinzuweisen, dass KI-Systeme unterschiedliche Automatisierungsgrade aufweisen können. Häufig wird zwischen Teil- und Vollautomatisierung unterschieden.¹¹ Ein Verfahren ist teilautomatisiert, wenn KI-Systeme unterstützend eingesetzt werden (z.B. Entscheidungsempfehlung). Ein Verfahren ist vollautomatisiert, wenn keinerlei menschliche Intervention stattfindet.¹²

3. Ausgewählte Einsatzbereiche in der öffentlichen Verwaltung

[8] Das folgende Kapitel befasst sich mit der Eignung der öffentlichen Verwaltung für den Einsatz von KI sowie mit den Bereichen, in denen KI in der Schweiz eingesetzt wird. Dabei sind die Standardisierung im Bereich der Massenverwaltung (Abschnitt 3.1), der Einsatz von Chatbots und Voicebots (Abschnitt 3.2), die intelligente Durchsuchung von Informationen (Abschnitt 3.3) sowie das Treffen von Prognosen (Abschnitt 3.4) von zentraler Bedeutung.

3.1. Standardisierung und Massenverwaltung

[9] Der KI-Einsatz erfordert in der Regel grosse, bestenfalls strukturierte Datenmengen.¹³ Die öffentliche Verwaltung verfügt über umfangreiche Datenbestände aus unterschiedlichen Quellen. Zudem laufen viele Tätigkeiten und Aufgaben immer gleich ab (z.B. die Prüfung, ob ein Gesuch mit allen notwendigen Angaben bei der richtigen Stelle eingereicht wurde).¹⁴ Bereiche, in

⁹ Terminologie Kompetenznetzwerk CNAI in der Version 2.0 vom 21. Dezember 2023, S. 7, <https://cnaai.swiss/dienstleistungen/terminologie-2/>.

¹⁰ Zum Vorschlag, die Umschreibung algorithmische bzw. algorithmenbasierte Systeme zu verwenden, siehe etwa FLORENT THOUVENIN et al., Ein Rechtsrahmen für Künstliche Intelligenz, Positionspapier November 2021, S. 3, <https://www.itssl.uzh.ch/dam/jcr:4687cf40-85a8-47f6-88ea-7ba54ad31d34/dsi-strategy-lab-21-de.pdf> (zit. Positionspapier 2021).

¹¹ Z.B. NADJA BRAUN BINDER, Vollautomatisierte Verwaltungsverfahren, vollautomatisiert erlassene Verwaltungsakte und elektronische Aktenführung, in: Margrit Seckelmann (Hrsg.), Digitalisierte Verwaltung – Vernetztes E-Government, 2. Aufl., Berlin 2019, S. 311 ff. Es können jedoch noch weitere Automatisierungsgrade unterschieden werden, siehe z.B. REGINA WEDER, Verfahrensrechtliche Anforderungen an automatisierte Verwaltungsverfahren, in: Monika Simmler (Hrsg.), Smart Criminal Justice, Basel 2021, S. 237 ff., S. 241.

¹² Siehe vertieft zu Teil- und Vollautomatisierung, LILIANE OBRECHT, Verfügung und automatisierte Einzelentscheidung – same same but different?, *ex/ante* 2/2022, S. 38 ff., S. 39 f., <https://ex-ante.ch/index.php/exante/article/view/189>.

¹³ WOLFGANG ERTEL, Grundkurs Künstliche Intelligenz, 4. Aufl., Wiesbaden 2016, S. 195 ff.

¹⁴ NADJA BRAUN BINDER et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4, S. 15, https://www.zh.ch/content/dam/zhweb/bilder-dokumente/themen/politik-staat/kanton/digitale-verwaltung-und-e-government/projekte_digitale_transformation/ki_einsatz_in_der_verwaltung_2021.pdf (zit. Studie KI 2021).

denen grosse Datenbestände vorhanden sind und Prozesse strukturiert ablaufen, sind für den KI-Einsatz besonders geeignet. Dazu zählen in der öffentlichen Verwaltung insbesondere das Steuer- und das Sozialversicherungsverfahren. In beiden Bereichen werden jährlich riesige Datenmengen generiert und viele Prozesse laufen standardisiert ab. Zur Verarbeitung dieser Datenmengen kann KI eingesetzt werden. Im Steuerverfahren können z.B. die Steuererklärungen aufbereitet und für das Veranlagungsverfahren vorbereitet werden.¹⁵ So werden in einigen Kantonen elektronisch eingereichte Steuererklärungen durch automatisierte Veranlagungssysteme (teil-)automatisiert bearbeitet.¹⁶ In Zukunft soll KI auch weiterführend in den Steuerverwaltungen eingesetzt werden.¹⁷ Im Sozialversicherungsbereich stehen insbesondere die automatisierte Bearbeitung von Anträgen zur Prämienverbilligung sowie die automatisierte Betrugsbekämpfung im Vordergrund.¹⁸

3.2. Chatbots und Voicebots

[10] Chatbots sind Onlinedialogsysteme, die in natürlicher Sprache und rund um die Uhr in gleichbleibender Qualität kommunizieren können.¹⁹ Die Interaktion findet dabei in einer digitalen Umgebung statt (z.B. auf der Gemeindefachseite oder einem Social-Media-Kanal).²⁰ Voicebots sind KI-basierte Computerprogramme, die menschliche Sprache erkennen und auf diese reagieren, um mit Nutzenden zu interagieren.²¹ Auch sie sind rund um die Uhr und mit derselben Qualität im Einsatz.²²

[11] In den vergangenen Jahren lässt sich in Schweizer öffentlichen Verwaltungen eine Zunahme der Nutzung von Chatbots und Voicebots für die Kommunikation mit der Bevölkerung beobachten.²³ So hat etwa die Eidgenössische Stiftungsaufsicht ESA im Januar 2024 den Chatbot «Esi» lanciert. Der Chatbot beantwortet verschiedene Fragen rund um das Schweizer Stiftungswesen.²⁴ Der Chatbot «StatBot.swiss» des Bundesamtes für Statistik interagiert mit den Nutzenden

¹⁵ Siehe zu den damit einhergehenden datenschutzrechtlichen Herausforderungen Abschnitt 4.4.3.

¹⁶ Im Kanton Bern werden fast 20 Prozent der steuerpflichtigen natürlichen Personen vollständig automatisiert veranlagt – das Ziel ist ein Automatisierungsgrad durch maschinelles Lernen von 40 Prozent. Siehe dazu CLAUDIO FISCHER/ANNEMARIE DAEPP, Digitalisierung am Beispiel der Steuerverwaltung des Kantons Bern, in: EF 4/2019, S. 327 ff., S. 329.

¹⁷ Siehe dazu MARIANNE NUFER, Künstliche Intelligenz in der Steuerveranlagung, in: ASA 88/2019/2020, S. 259 ff., S. 259 ff.; siehe ferner die Vernehmlassungsvorlage vom 12. März 2024 zu einer Anpassung des Steuergesetzes im Kanton Solothurn, die eine vollautomatische Veranlagung unter Nutzung von maschinellen Lernverfahren ermöglichen soll, https://so.ch/fileadmin/internet/staatskanzlei/stk-regierungsdienste/pdf/Vernehmlassungen/VL_Staats_Gemeindesteuern_2024.pdf.

¹⁸ BRAUN BINDER et al., Studie KI 2021 (Fn. 14), S. 25.

¹⁹ BEATE BRUNS/CÄCILIE KOWALD, Praxisleitfaden Chatbots: Conversation Design für eine bessere User Experience, Wiesbaden 2023, S. 1 f.; FABIA STÖCKLIN, Robot Recruiting: Die Haftung des Arbeitgebers für Persönlichkeitsverletzungen beim Einsatz von Künstlicher Intelligenz im Bewerbungsprozess, *sui generis* 2023, S. 2, <https://sui-generis.ch/article/view/4045>.

²⁰ NADJA BRAUN BINDER/LILIANE OBRECHT/GRACE WITTMER, Vertrauensschutz bei fehlerhaften Behördenauskünften durch Chatbots, *iusnet* 2024, <https://digitalesrecht-datenrecht.iusnet.ch/de/fachbeitraege/vertrauensschutz-bei-fehlerhaften-behoerdenauskueften-durch-chatbots>.

²¹ FRANZISKA OTT, Voicebot: Definition, Anbieter und Vorteile, HubSpot 31. Mai 2023, <https://blog.hubspot.de/service/voicebot>.

²² OTT (Fn. 21).

²³ Zu den rechtlichen Herausforderungen betreffend potenzielle fehlerhafte Chatbot-Auskünfte siehe unten Abschnitt 4.5.

²⁴ Eidgenössische Stiftungsaufsicht, Frag Esi, <https://www.fragesi.ch>.

nach dem Frage-Antwort-Prinzip. Das Ziel dieses Chatbots besteht in der Unterstützung bei der Durchsuchung offener Datensätze.²⁵ Die Stadt St. Gallen implementierte den Chatbot «Gallus», dessen Zielsetzung in der vereinfachten Zugänglichkeit bestimmter Basisdienste für die Bevölkerung besteht.²⁶ Im Kanton Aargau wird seitens des Strassenverkehrsamtes der Voicebot «Vicky» eingesetzt, um Standardfragen (auf Schweizerdeutsch) von Anrufenden zu beantworten.²⁷

[12] Es ist zu betonen, dass nicht alle eingesetzten Chatbots tatsächlich auf KI basieren. Häufig werden auch sog. regelbasierte²⁸ Chatbots eingesetzt, die nicht selbstständig lernen und deren Antworten auch nicht über die vorprogrammierten Antworten hinausgehen können. So setzt etwa das Zivilstandsamt des Kantons Basel-Stadt einen regelbasierten Chatbot u.a. für die Bestellung von Geburtsurkunden oder Heimatscheinen ein.²⁹

3.3. Intelligente Informationsdurchsuchung

[13] Der Einsatz von KI bietet sich zudem an, um die rasant wachsenden Datenmengen, deren Auswertung durch Menschen aufgrund der schieren Menge und Komplexität kaum mehr in angemessener Zeit möglich ist, intelligent zu durchsuchen. Im Bereich der Polizeiarbeit betrifft dies insbesondere die Bekämpfung von Wirtschaftskriminalität, der organisierten Kriminalität im Allgemeinen oder auch der Cyberkriminalität.³⁰ Dazu gehört schliesslich auch die intelligente Videoanalyse, da der Einsatz von KI bei der Bilderkennung und -auswertung besonders effizient ist.³¹

3.4. Prognosen

[14] Die Erstellung von Prognosen bildet einen weiteren Anwendungsbereich der KI. Im Kontext der Stadtplanung und -entwicklung hin zu sog. *Smart Cities* erweist sich der KI-Einsatz als besonders attraktiv. Dabei werden verschiedene Daten in der Stadt erhoben und miteinander verknüpft. Derzeit werden z.B. im Kanton Basel-Stadt verschiedene Projekte zur nachhaltigen Stadtentwicklung durchgeführt. Zu den Anwendungsgebieten zählt u.a. die intelligente Steuerung von Fussgängerinnen und Fussgängern im Strassenverkehr mittels Abstimmung von Lichtsignalanlagen, das Schaffen eines intelligenten Energienetzes oder die Analyse und Verbesserung der Luftqualität durch den Einsatz von Mikrosensoren auf Trams.³²

²⁵ Bundesamt für Statistik, StatBot.swiss: mit offenen Daten diskutieren, <https://www.bfs.admin.ch/bfs/de/home/dssc/blog/2023-02-statbot.html>.

²⁶ Chatbot der Stadtverwaltung St. Gallen, https://www.stadt.sg.ch/news/stsg_home/2020/12/chatbot.html.

²⁷ Kanton Aargau, VoiceBot im Strassenverkehrsamt, 13. September 2023, <https://www.ag.ch/de/verwaltung/dvi/ueber-uns/dossiers-projekte/voicebot>.

²⁸ Regelbasierte Algorithmen funktionieren nach dem «wenn (A) dann (B)»-Prinzip, siehe CHRISTOPH BEIERLE/GABRIELE KERN-ISBERNER, Methoden wissensbasierter Systeme, 3. Aufl., Wiesbaden 2006, S. 71 ff.

²⁹ Kanton Basel-Stadt, Bevölkerungsdienste und Migration, <https://www.bdm.bs.ch/Zivilstand.html>.

³⁰ Siehe etwa CHRISTIAN GLAUS, Ein Computer soll für die Luzerner Polizei ermitteln, Luzerner Zeitung vom 26. Oktober 2019, <https://www.luzernerzeitung.ch/zentralschweiz/luzern/kuenstliche-intelligenz-ein-computer-soll-fuer-die-luzerner-polizei-ermitteln-ld.1163004>.

³¹ Siehe etwa THOMAS GOLDA/MICKAEL CORMIER/JÜRGEN BEYERER, Intelligente Bild- und Videoauswertung für die Sicherheit, in: Dieter Wehe (Hrsg.), Handbuch Polizeimanagement – Polizeipolitik – Polizeiwissenschaft – Polizeipraxis, Wiesbaden 2022, S. 1 ff.

³² Für alle Projekte: Kanton Basel-Stadt, Smart City Basel, siehe <https://www.smartcity.bs.ch/projekte.html>.

[15] In der Polizeiarbeit wird KI neben der Informationsdurchsuchung³³ auch zur Prognose eingesetzt. Im Bereich der vorausschauenden Polizeiarbeit, auch als *Predictive Policing* bezeichnet, können statistikbasierte Vorhersagen über Verhaltensweisen oder Ereignisse getroffen werden.³⁴ In allen Deutschschweizer Kantonen wird im Rahmen des «Risikoorientierten Sanktionenvollzugs» (ROS) ein «Fall Screening Tool» zur Prüfung von Vollzugslockerungen eingesetzt.³⁵ Dadurch soll das Rückfallrisiko während und nach dem Vollzug gesenkt werden.³⁶

4. Rechtliche Herausforderungen

[16] Die folgenden Abschnitte thematisieren verschiedene rechtliche Herausforderungen, die sich im Kontext des KI-Einsatzes in der öffentlichen Verwaltung stellen, und skizzieren die erforderlichen rechtlichen Rahmenbedingungen für den KI-Einsatz in der öffentlichen Verwaltung. Das Legalitätsprinzip (Abschnitt 4.1), die Verfahrensgarantien und Verfahrensgrundsätze (Abschnitt 4.2), das Diskriminierungsverbot (Abschnitt 4.3), die informationelle Selbstbestimmung und das Datenschutzrecht (Abschnitt 4.4) sowie der Vertrauensschutz (Abschnitt 4.5) sind dabei von entscheidender Bedeutung. Schliesslich wird das Schlagwort «Transparenz» im Kontext von KI erörtert (Abschnitt 4.6).

4.1. Legalitätsprinzip

[17] Das in Art. 5 Abs. 1 Bundesverfassung (BV³⁷) verankerte Legalitätsprinzip besagt, dass staatliches Handeln immer einer rechtlichen Grundlage bedarf.³⁸ Die rechtliche Grundlage muss sowohl hinsichtlich ihrer Normstufe als auch hinsichtlich ihrer Normdichte ausreichend bestimmt sein.³⁹ Dieser Grundsatz ist auch im Zusammenhang mit dem KI-Einsatz zu beachten.⁴⁰ So ist für den KI-Einsatz bei der Entscheidungsfindung in Verwaltungsverfahren regelmässig eine spezifische und explizite Gesetzesgrundlage zu fordern.⁴¹ Neben den in der Bundesverfassung festgelegten Anforderungen, sind auch die Bestimmungen des Datenschutzgesetz (DSG⁴²) zu beachten. Demnach ist z.B. für die Bearbeitung von besonders schützenswerten Personendaten durch Bun-

³³ Siehe unten Abschnitt 3.3.

³⁴ MONIKA SIMMLER/SIMONE BRUNNER/KUNO SCHEDLER, Smart Criminal Justice – Eine empirische Studie zum Einsatz von Algorithmen in der Schweizer Polizeiarbeit und Strafrechtspflege, St. Gallen 2020, S. 14 ff., S. 31 ff.; JENNIFER PULLEN/PATRICIA SCHEFER, Predictive Policing – Grundlagen, Funktionsweisen und Wirkung, in: Simmler Monika (Hrsg.), Smart Criminal Justice, Basel 2021, S. 103 ff.

³⁵ Zum Ganzen MONIKA SIMMLER/SIMONE BRUNNER, Smart Criminal Justice in der Schweiz – Die Kantone im Bann der Algorithmen?, in: Monika Simmler (Hrsg.), Smart Criminal Justice, Basel 2021, S. 9 ff., S. 16 f. mit einer Übersicht zu weiteren Tools im Rahmen von Smart Criminal Justice.

³⁶ SIMMLER/BRUNNER (Fn. 35), S. 16.

³⁷ Bundesverfassung der Schweizerischen Eidgenossenschaft vom 18. April 1999, SR 101.

³⁸ SGK BV-SCHINDLER, Art. 5 (I. – IV) N 21, in: Bernhard Ehrenzeller et al. (Hrsg.), Die schweizerische Bundesverfassung, St. Galler Kommentar, 4. Aufl., Zürich 2023 (zit. SGK BV).

³⁹ SGK BV-SCHINDLER (Fn. 38), Art. 5 (I. – IV) N 34.

⁴⁰ DANIELA THURNHERR, Automatisierte Verwaltungsverfahren auf Bundesebene (Schweiz), in: Nadja Braun Binder/Peter Bussjäger/Matthias Eller (Hrsg.), Auswirkungen der Digitalisierung auf die Erlassung und Zuordnung behördlicher Entscheidungen, Wien/Hamburg 2021, S. 133 ff., S. 150.

⁴¹ THURNHERR (Fn. 40), S. 151; ausführlich auch ANDREAS GLASER, Einflüsse der Digitalisierung auf das schweizerische Verwaltungsrecht, SJZ 2018, S. 181 ff., S. 188.

⁴² Bundesgesetz über den Datenschutz vom 25. September 2020, SR 235.1.

desorgane oder in Fällen von Profiling eine formell gesetzliche Grundlage notwendig (Art. 34 Abs. 2 lit. a und b DSG).⁴³

4.2. Verfahrensgarantien und Verfahrensgrundsätze

[18] Verfahrensgarantien und Verfahrensgrundsätze gewährleisten die Rechtmässigkeit eines Verfahrens. Die verfassungsrechtlichen Verfahrensgarantien sind in den Artikeln 29 bis 32 BV verankert. Sie richten sich an alle Verfahrensbeteiligten und sind von den Behörden in allen Verfahren anzuwenden.⁴⁴ Es handelt sich dabei um Minimalstandards mit grundrechtlichem Charakter. Die Verfahrensgarantien verfolgen sowohl individualrechtliche als auch rechtsstaatliche Ziele. Der betroffenen Person sollen sie die prozeduralen Rechte garantieren. Aus rechtsstaatlicher Perspektive zielen sie auf die Wahrheitsfindung ab.⁴⁵ Mit Blick auf das Verwaltungshandeln sind die «allgemeinen Verfahrensgarantien» gemäss Art. 29 BV von besonderer Bedeutung.⁴⁶ Im Zusammenhang mit dem Einsatz von KI ist im Folgenden insbesondere der Anspruch auf rechtliches Gehör vertieft zu betrachten (Abschnitt 4.2.1).

[19] Im Gegensatz zu den Verfahrensgarantien, die konkrete Rechte und Garantien gewähren, handelt es sich bei den Verfahrensgrundsätzen um Leitlinien oder Prinzipien, die die Durchführung eines Prozesses steuern. Sie sind positivrechtlich im Verfahrensrecht verankert.⁴⁷ Mit Blick auf den Einsatz von KI wird im Folgenden insbesondere auf den Untersuchungsgrundsatz eingegangen (Abschnitt 4.2.2).⁴⁸

4.2.1. Rechtliches Gehör

[20] Der verfassungsrechtlich garantierte Anspruch auf rechtliches Gehör (Art. 29 Abs. 2 BV) dient der Sachverhaltsabklärung und stellt zugleich ein persönlichkeitsbezogenes Mitwirkungsrecht dar.⁴⁹ Rechtsprechung und Lehre leiten aus dem rechtlichen Gehör verschiedene Teilgehalte ab.⁵⁰ Dazu zählen insbesondere der Anspruch auf vorgängige Äusserung und Mitwirkung im Verfahren sowie das Recht auf Begründung.⁵¹ Diese beiden Teilgehalte sind im Kontext von KI von besonderer Relevanz.

⁴³ Siehe dazu vertieft Abschnitt 4.4.

⁴⁴ REGINA KIENER/WALTER KÄLIN/JUDITH WYTTENBACH, Grundrechte, 3. Aufl., Bern 2018, § 40 N 1; RENÉ RHINOW et al., Öffentliches Prozessrecht, 4. Aufl., Basel 2021, Rz. 255.

⁴⁵ ANJA MARTINA BINDER, Expertenwissen und Verfahrensgarantien, Zürich 2016, S. 21 f.

⁴⁶ BSK BV-WALDMANN, Art. 29 N 6, in: Bernhard Waldmann/Eva Maria Belser/Astrid Epiney (Hrsg.), Basler Kommentar Bundesverfassung, 1. Aufl., Basel 2015 (zit. BSK BV).

⁴⁷ PIERRE TSCHANNEN/MARKUS MÜLLER/MARKUS KERN, Allgemeines Verwaltungsrecht, 5. Aufl., Bern 2022, § 30 Rz. 771; RHINOW et al. (Fn. 44), Rz. 974 f.

⁴⁸ Vertieft dazu NADJA BRAUN BINDER, Der Untersuchungsgrundsatz als Herausforderung vollautomatisierter Verfahren, in: zsis 2/2020, Rz. 25, <https://www.zsis.ch/artikel/der-untersuchungsgrundsatz-als-herausforderung-vollautomatisierter-verfahren> (zit. Untersuchungsgrundsatz).

⁴⁹ BSK BV-WALDMANN (Fn. 46), Art. 29 N 44; RHINOW et al. (Fn. 44), Rz. 309.

⁵⁰ M. w. H. BSK BV-WALDMANN (Fn. 46), Art. 29 N 44.

⁵¹ Auf das Akteneinsichtsrecht als Bestandteil des rechtlichen Gehörs wird im Folgenden nicht eingegangen. Siehe dazu im automatisierten Verwaltungsverfahren WEDER (Fn. 11), S. 251 ff.

[21] Auf Bundesebene konkretisiert Art. 30 Abs. 1 Verwaltungsverfahrensgesetz (VwVG⁵²) den Anspruch auf vorgängige Äusserung und Mitwirkung im Verwaltungsverfahren. Die Parteien sind vor dem Erlass einer Verfügung anzuhören bzw. ihnen ist Gelegenheit zur Stellungnahme zu geben.⁵³ Dieser Anspruch stellt für die betroffenen Parteien das wichtigste Mittel dar, um auf die Ermittlung des rechtserheblichen Sachverhalts Einfluss zu nehmen.⁵⁴ Art. 30 Abs. 2 VwVG zählt abschliessend fünf Ausnahmesituationen auf, in denen die Behörde von einer Anhörung absehen kann.⁵⁵ Damit die Betroffenen überhaupt Stellung nehmen können, müssen sie die wesentlichen Elemente des voraussichtlichen Verfügungsinhalts kennen.⁵⁶ Die Behörde ist hingegen nicht verpflichtet, den Äusserungen der Betroffenen zu folgen. Sie ist allerdings verpflichtet, die Äusserungen zur Kenntnis zu nehmen, zu überprüfen und sich angemessen mit den Einwänden bei der Entscheidungsfindung und -begründung auseinanderzusetzen.⁵⁷

[22] Der Einsatz von KI im Verwaltungsverfahren schränkt den Anspruch auf vorgängige Äusserung nicht per se ein. Vielmehr kommt es auf den Zeitpunkt des KI-Einsatzes und den Automatisierungsgrad des KI-Systems an.⁵⁸ Insbesondere in vollautomatisierten Verfahren kann der Anspruch auf vorgängige Äusserung jedoch beeinträchtigt sein. Wird ein solches Verfahren von einer natürlichen Person eingeleitet (z.B. durch einen Antrag) und anschliessend ohne jegliche menschliche Intervention abgewickelt, kann sich die antragstellende Person nur bei der eigentlichen Antragstellung äussern. Eine Äusserung im weiteren Verlauf des vollautomatisierten Verfahrens ist nicht möglich.⁵⁹ Der Grad der Beeinträchtigung des Anspruchs auf vorgängige Äusserung ist letztlich von zwei Faktoren abhängig: Einerseits von den Möglichkeiten der Äusserung bei der Antragstellung und andererseits davon, inwiefern die getätigten Äusserungen vom System berücksichtigt werden können.⁶⁰ In der Literatur wird daher dafür plädiert, vollautomatisierte Verfahren nur dann einzusetzen, wenn den Begehren der Betroffenen vollumfänglich entsprochen wird. Im Falle einer Entscheidung, die sich nachteilig für die betroffene Person auswirkt, ist das rechtliche Gehör durch eine natürliche Person zu gewährleisten.⁶¹ Dies würde allerdings wiederum dazu führen, dass es sich nicht mehr um eine vollautomatisierte Entscheidung handelt.

⁵² Bundesgesetz über das Verwaltungsverfahren vom 20. Dezember 1968, SR 172.021.

⁵³ BERNHARD WALDMANN/JÜRG BICKEL, Kommentierung Art. 30 Rz. 1, in: Bernhard Waldmann/Patrick L. Krauskopf (Hrsg.), VwVG – Praxiskommentar Verwaltungsverfahrensgesetz, 3. Aufl., Zürich 2023 (zit. Praxiskomm VwVG).

⁵⁴ Statt vieler RHINOW et al. (Fn. 44), Rz. 323 ff.

⁵⁵ Dabei handelt es sich um die folgenden Konstellationen: Zwischenverfügungen, die nicht selbstständig durch Beschwerde anfechtbar sind (Bst. a); Verfügungen, die durch Einsprache anfechtbar sind (Bst. b); Verfügungen, in denen die Behörde den Begehren der Parteien voll entspricht (Bst. c); Vollstreckungsverfügungen (Bst. d); andere Verfügungen in einem erstinstanzlichen Verfahren, wenn Gefahr im Verzuge ist, den Parteien die Beschwerde gegen die Verfügung zusteht und einen keine andere Bestimmung des Bundesrechts einen Anspruch auf vorgängige Anhörung gewährleistet (Bst. e).

⁵⁶ Praxiskomm VwVG-WALDMANN/BICKEL (Fn. 53), Art. 30 Rz. 18.

⁵⁷ RHINOW et al. (Fn. 44), Rz. 328 ff.; KIENER/KÄLIN/WYTTENBACH (Fn. 44), § 41 N 40 ff.

⁵⁸ Siehe zu den verschiedenen Automatisierungsgraden Abschnitt 2.

⁵⁹ NADJA BRAUN BINDER, Automatisierte Entscheidungen: Perspektive Datenschutzrecht und öffentliche Verwaltung, SZW 2020, S. 27 ff., S. 28 f.; NADJA BRAUN BINDER, Als Verfügungen gelten Anordnungen der Maschinen im Einzelfall... – Dystopie oder künftiger Verwaltungsalltag?, ZSR 2020, S. 253 ff., S. 273 f. (zit. Dystopie); DAVID RECHSTEINER, Der Algorithmus verfügt, in: Jusletter 26. November 2018, Rz. 18 ff.

⁶⁰ NADJA BRAUN BINDER, Künstliche Intelligenz und automatisierte Entscheidungen in der öffentlichen Verwaltung, SJZ 2019, S. 467 ff., S. 476 (zit. KI Verwaltung).

⁶¹ RECHSTEINER (Fn. 59), N 23; BRAUN BINDER, KI Verwaltung (Fn. 60), S. 476.

[23] Die verfassungsrechtlich verankerte Begründungspflicht wird in Art. 35 Abs. 1 VwVG für schriftliche Verfügungen konkretisiert. Die Begründungspflicht stellt die Nachvollziehbarkeit und Rechtsstaatlichkeit der behördlichen Entscheidungsfindung im Einzelfall sicher und erfüllt verschiedene Funktionen.⁶² So soll die Begründung z.B. eine sachgerechte Anfechtung der Verfügung ermöglichen, bei der betroffenen Person Akzeptanz schaffen oder die Behördenmitarbeitenden zur Selbstreflexion anregen.⁶³ Die verschiedenen Funktionen der Begründung kommen grundsätzlich auch zum Tragen, wenn eine Verfügung KI-gestützt erlassen wird. Je nach Art des eingesetzten KI-Systems kann es jedoch sein, dass der Weg, wie eine Entscheidung zustande gekommen ist, für die Behördenmitarbeitenden nicht nachvollziehbar ist.⁶⁴ Sie können daher den Anforderungen an die Begründungspflicht nicht ohne weiteres nachkommen. Es genügt nicht, der betroffenen Person die Funktionsweise des KI-Systems allgemein zu erklären.⁶⁵ Denn eine solche Erklärung stellt gerade keine einzelfallweise Begründung einer Verfügung dar und legt den Entscheidungsvorgang für die betroffene Person nicht offen.⁶⁶ Zusammenfassend ist festzuhalten, dass die Funktionen, die der Begründungspflicht zukommen, auch bei einem KI-Einsatz sichergestellt sein müssen.⁶⁷

4.2.2. Untersuchungsgrundsatz

[24] Gemäss dem Untersuchungsgrundsatz (Art. 12 VwVG) ermitteln die Behörden den Sachverhalt von Amtes wegen. Die Behörde ist als Entscheidungsträgerin für die Beschaffung der Entscheidungsgrundlage verantwortlich.⁶⁸ Sie ermittelt alle rechtserheblichen Informationen.⁶⁹ Der Untersuchungsgrundsatz wird durch die Mitwirkungspflicht gemäss Art. 13 VwVG eingeschränkt, indem die Parteien verpflichtet werden, an der Feststellung des Sachverhalts mitzuwirken.⁷⁰

[25] Im Verwaltungsverfahren müssen alle Informationen ermittelt werden, die den verfahrensspezifischen Sachverhalt betreffen. Die für den Sachverhalt relevanten Informationen müssen vollständig und korrekt sein. Darüber hinaus müssen sie im erforderlichen Umfang verfügbar sein, um damit den rechtserheblichen Sachverhalt ermitteln zu können.⁷¹ Im automatisierten Verfahren werden die Sachverhaltsdaten automatisiert gesammelt, ausgewertet und verifiziert.⁷²

⁶² Nur RENÉ WIEDERKEHR/CHRISTIAN MEYER/ANNA BÖHME, VwVG – Bundesgesetz über das Verwaltungsverfahren mit weiteren Erlassen, Zürich 2022, Kommentierung Art. 35 N 6 ff.

⁶³ Vertieft zu den Funktionen der Begründungspflicht etwa LORENZ KNEUBÜHLER, Die Begründungspflicht, Diss. Bern 1998, Bern/Stuttgart/Wien 1998, S. 94 ff.

⁶⁴ MARIO MARTINI, Blackbox Algorithmus, Berlin 2019, S. 27 ff.

⁶⁵ Zurzeit wird an der Erklärbarkeit von KI-Systemen geforscht (sog. «Explainable Artificial Intelligence, XAI»). Solange jedoch nur in allgemeiner Weise erklärt werden kann, wie ein KI-System zum Ergebnis kommt, genügt dies nicht, um die Begründungspflicht im Einzelfall sicherzustellen. Zum Konzept im Zusammenhang mit Recht etwa MARY-ANNE WILLIAMS, Explainable artificial intelligence, in: Vogl Roland (Hrsg.), Research Handbook on Big Data Law, Cheltenham/Northampton 2021, S. 318 ff.

⁶⁶ Vertieft BRAUN BINDER et al., Studie KI 2021 (Fn. 14), S. 37 f.

⁶⁷ Siehe dazu ausführlich NADJA BRAUN BINDER/LILIANE OBRECHT, Die Begründung von Verfügungen beim Einsatz algorithmischer Systeme, SJZ 2024 (i.E.) (zit. Begründung).

⁶⁸ Praxiskomm VwVG-KRAUSKOPF/WYSSLING (Fn. 53), Art. 12 / I. – II Rz. 16 ff.

⁶⁹ Praxiskomm VwVG-KRAUSKOPF/WYSSLING (Fn. 53), Art. 12 / I. – II Rz. 20 ff.

⁷⁰ Praxiskomm VwVG-KRAUSKOPF/WYSSLING (Fn. 53), Art. 13 Rz. 4 f.

⁷¹ Praxiskomm VwVG-KRAUSKOPF/WYSSLING (Fn. 53), Art. 12 / I. – II Rz. 20 ff.

⁷² GLASER (Fn. 41), S. 184.

Schon aufgrund der Definition der Vollautomatisierung unternehmen Behördenmitarbeitende im vollautomatisierten Verfahren keine weiteren Abklärungen zur Sachverhaltsermittlung.⁷³ Dies birgt die Gefahr, dass nur standardisiert erhobene Daten bzw. in Datenbanken vorhandene Daten genutzt werden und einzelfallspezifische Informationen unbeachtet bleiben.⁷⁴ Die Behörden müssen jedoch auch bei einem KI-Einsatz sicherstellen, dass alle für die Sachverhaltsermittlung rechtserheblichen Tatsachen erhoben werden.⁷⁵ Vor diesem Hintergrund sind die Vollständigkeit, Aktualität und Richtigkeit der genutzten Daten von besonderer Bedeutung.⁷⁶

[26] Auch bei einer Mitwirkungspflicht der Beteiligten besteht die Gefahr, dass einzelfallspezifische Informationen unberücksichtigt bleiben. Damit ein KI-System die Datenbestände nutzen kann, müssen die Daten in standardisierter Form erfasst werden. Im Gesuchverfahren kann dies dazu führen, dass für den Einzelfall notwendige Informationen nicht erfasst werden können, weil z.B. das standardisierte Formular kein Feld für individuelle Erklärungen enthält.⁷⁷

4.3. Diskriminierungsverbot

[27] Dass der KI-Einsatz zu Diskriminierungen führen kann (Abschnitt 4.3.1), ist mittlerweile unbestritten. Dies zeigen auch zwei reale Beispiele aus den Niederlanden und Frankreich (Abschnitt 4.3.2).

4.3.1. Diskriminierungsarten und Erscheinungsformen

[28] Gemäss Art. 8 Abs. 2 BV darf niemand diskriminiert werden, namentlich nicht wegen der Herkunft, der Rasse, des Geschlechts, des Alters, der Sprache, der sozialen Stellung, der Lebensform, der religiösen, weltanschaulichen oder politischen Überzeugung oder wegen einer körperlichen, geistigen oder psychischen Behinderung. Diese Aufzählung der sensiblen Merkmale ist nicht abschliessend.⁷⁸ Das Diskriminierungsverbot stellt einen qualifizierten Schutz vor Ungleichbehandlung dar. So liegt gemäss der bundesgerichtlichen Rechtsprechung eine Diskriminierung i.S.v. Art. 8 Abs. 2 BV vor, «wenn eine Person ungleich behandelt wird allein aufgrund ihrer Zugehörigkeit zu einer bestimmten Gruppe, welche historisch oder in der gegenwärtigen sozialen Wirklichkeit tendenziell ausgegrenzt oder als minderwertig angesehen wird. (...) Die Diskriminierung stellt eine qualifizierte Art der Ungleichbehandlung von Personen in vergleichbaren Situationen dar, indem sie eine Benachteiligung eines Menschen bewirkt, die als Herabwürdigung oder Ausgrenzung einzustufen ist, weil sie an ein Unterscheidungsmerkmal anknüpft, das einen wesentlichen und nicht oder nur schwer aufgebaren Bestandteil der Identität der betreffenden Person ausmacht»⁷⁹.

⁷³ Siehe Abschnitt 2.

⁷⁴ BRAUN BINDER, Untersuchungsgrundsatz (Fn. 48) Rz. 25; BRAUN BINDER, KI Verwaltung (Fn. 60), S. 470 ff.

⁷⁵ BRAUN BINDER, Untersuchungsgrundsatz (Fn. 48), Rz. 28.

⁷⁶ Siehe dazu NADJA BRAUN BINDER/LILIANE OBRECHT, Künstliche Intelligenz in der Verwaltung – Möglichkeiten, rechtliche Rahmenbedingungen und Grenzen, in: SVVOR Jahrbuch 2021/2022, S. 21 ff., Fn. 80 (zit. SVVOR Jahrbuch).

⁷⁷ BRAUN BINDER, Untersuchungsgrundsatz (Fn. 48), Rz. 28

⁷⁸ BSK-BV-Waldmann (Fn. 46), Art. 8 Rz. 65.

⁷⁹ BGE 139 I 169 E. 7.2.1; 139 I 292 E. 8.2.1; 138 I 305 E. 3.3.; 136 I 297 E. 7.1.

[29] Es wird zwischen direkter (unmittelbarer) und indirekter (mittelbarer) Diskriminierung unterschieden.⁸⁰ Direkt diskriminierend sind Rechtsakte, die ohne hinreichende Rechtfertigung an ein sensibles Merkmal anknüpfen. Sie bewirken damit eine Benachteiligung, die nicht durch qualifizierte Gründe gerechtfertigt werden kann. Eine indirekte Diskriminierung liegt demgegenüber dann vor, wenn ein Rechtsakt zwar neutral formuliert ist und keine offensichtliche Benachteiligung einer bestimmten Gruppe enthält, in seiner praktischen Auswirkung jedoch eine diskriminierende Wirkung entfaltet. Beim KI-Einsatz können sowohl direkte als auch indirekte Diskriminierungen vorkommen.

[30] Von besonderer Relevanz im Zusammenhang mit KI ist die sog. Proxy-Diskriminierung.⁸¹ Proxys (sog. Stellvertreter) sind harmlos erscheinende Merkmale, die stark mit sensiblen Merkmalen korrelieren können.⁸² KI-Algorithmen stellen häufig auf Stellvertretervariablen ab, da sie nach Korrelationen suchen und direkt diskriminierende Merkmale aus den Datensätzen gelöscht wurden.⁸³ Knüpft ein KI-System z.B. an ein vermeintlich neutrales Merkmal wie «Wohnort» an, kann dies dennoch zu einer Diskriminierung führen, wenn dieses Merkmal überproportional häufig mit dem sensiblen Merkmal «Rasse» korreliert. Dies ist z.B. der Fall, wenn in einem Stadtteil überwiegend viele Personen einer bestimmten Ethnie wohnen.⁸⁴ Das Merkmal Stadtteil wirkt diesfalls stellvertretend als Variable für die Ethnie.

[31] Diskriminierungen können beim KI-Einsatz aus verschiedenen Diskriminierungsquellen resultieren.⁸⁵ So kann der Grundstein für spätere Diskriminierungen bereits in der Entwicklung des KI-Systems gelegt werden. KI-Systeme sind bei ihrer Entwicklung auf Trainingsdaten angewiesen. Diese Trainingsdaten wiederum prägen die Funktionsweise des Systems.⁸⁶ Sind bereits die Trainingsdaten fehlerhaft, unvollständig (z.B. wenn bestimmte Bevölkerungsgruppen unterrepräsentiert sind), veraltet oder spiegeln sie in der Gesellschaft existierende Vorurteile wider⁸⁷, können die Ergebnisse diskriminierend sein. An dieser Stelle ist zu betonen, dass der Einsatz von KI zur Entscheidungsfällung jedoch auch dazu führen kann, dass in der Gesellschaft bestehende Diskriminierungen aufgedeckt werden. Darüber hinaus sind die Trainingsdaten durch die Entwickelnden geprägt, da diese i.d.R. die Datensätze bestimmen, mit denen das KI-System trainiert, validiert oder getestet wird.⁸⁸ Neben der Entwicklung können Diskriminierungen auch aus der Anwendung des KI-Systems resultieren. Dies ist z.B. der Fall, wenn in teilautomatisierten Verfah-

⁸⁰ Die folgenden Ausführungen erfolgen gestützt auf BSK-BV-Waldmann (Fn. 46), Art. 8 Rz. 62 ff.; KIENER/KÄLIN/WYTTENBACH (Fn. 44), § 36 Rz. 36 ff.; BGE 138 I 305, E. 3.3.

⁸¹ GABRIELE BUCHHOLTZ/MARTIN SCHEFFEL-KAIN, Algorithmen und Proxy Discrimination in der Verwaltung: Vorschläge zur Wahrung digitaler Gleichheit, in: NVwZ 2022, S. 612 ff.; ANYA E. R. PRINCE/DANIEL SCHWARCZ, Proxy Discrimination in the Age of Artificial Intelligence and Big Data, Iowa Law Review 2020, S. 1257 ff.

⁸² BUCHHOLTZ/SCHWARCZ (Fn. 81), S. 612.

⁸³ PRINCE/SCHWARCZ (Fn. 81), S. 1260 f.

⁸⁴ BUCHHOLTZ/SCHWARCZ (Fn. 81), S. 613.

⁸⁵ Siehe zum Ganzen BRAUN BINDER et al., Studie KI 2021 (Fn. 14), S. 40 ff.

⁸⁶ ERTEL (Fn. 13), S. 194, 230.

⁸⁷ EVA WOLFANGEL, Wo hat sie das nur gelernt?, in: Bischoff Manon (Hrsg.), Künstliche Intelligenz: Vom Schachspieler zur Superintelligenz?, Berlin 2022, S. 171 ff.

⁸⁸ NUR ALMA KOLLECK/CARSTEN Orwat, Mögliche Diskriminierung durch algorithmische Entscheidungssysteme und maschinelles Lernen – ein Überblick, Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, 2020, S. 32, <https://www.tab-beim-bundestag.de/de/pdf/publikationen/berichte/TAB-Hintergrundpapierhp024.pdf>.

ren ein Ergebnis eines KI-Systems von Sachbearbeitenden falsch interpretiert oder dem Ergebnis fälschlicherweise übermässiges Vertrauen zugemessen wird (Stichwort: *Automation Bias*).⁸⁹

[32] Um Diskriminierung beim Einsatz von KI-Systemen zu verhindern, muss sichergestellt werden, dass korrekte, vollständige und repräsentative Daten verwendet werden.⁹⁰ Allerdings lässt sich dadurch nicht jede Diskriminierung verhindern.⁹¹ Unter Umständen ist in sensiblen Bereichen zumindest auf vollautomatisierte Entscheidungen zu verzichten. Zudem ist durch geeignete Massnahmen sicherzustellen, dass die Sachbearbeitenden über die notwendigen Kenntnisse und Kompetenzen verfügen, um eine im teilautomatisierten Verfahren getroffene Entscheidung zu überprüfen und gegebenenfalls eine Entscheidung zu treffen, die von einem diskriminierenden Vorschlag abweicht.

4.3.2. Beispiele algorithmischer Diskriminierung

[33] Die folgenden zwei Beispiele zeigen, dass der Einsatz von KI-Systemen bereits zu Diskriminierungen geführt hat. Im Jahr 2021 sorgte die niederländische Steuerbehörde für medialen Aufruhr.⁹² Von 2010 bis 2019 setzte die niederländische Steuerbehörde KI-basierte Systeme ein, um Betrugsfälle bei der Auszahlung von Kinderbetreuungsbeiträgen aufzudecken. Das System knüpfte dabei an diskriminierende Anhaltspunkte wie ausländisch klingende Namen oder eine doppelte Staatsangehörigkeit an. Dadurch wurden mehr als 20'000 Familien finanziell ruiniert.⁹³

[34] Ein weiteres Beispiel stammt aus Frankreich. Dort wird seit 2018 die online Plattform *Parcoursup* verwendet, um die Studienplätze zentralisiert zu verteilen.⁹⁴ Nebst einer Effizienzsteigerung im Anmeldeverfahren sollte dadurch die Chancengerechtigkeit erhöht werden, indem alle Schulabgängerinnen und -abgänger unabhängig ihrer sozialen Herkunft oder ihres Wohnorts die gleichen Chancen auf einen Studienplatz erhalten.⁹⁵ Das Auswahlverfahren von *Parcoursup* basiert jedoch auf einem Algorithmus, dessen Auswahlkriterien zu Beginn nicht vollständig offen gelegt wurden.⁹⁶ Bekannt ist, dass die Noten der Schülerinnen und Schüler aus den letzten Schul-

⁸⁹ Siehe zum Begriff *Emergent Bias* bereits FRIEDMANN/NISSENBAUM, S. 336; zum *Automation Bias* siehe PETER PARYCEK/VERENA SCHMID/ANNA-SOPHIE NOVAK, Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations, and Framework Conditions, in: Journal of the Knowledge Economy 2023, siehe <https://doi.org/10.1007/s13132-023-01433-3>.

⁹⁰ JEREMY STEVENS, Datenqualität bei algorithmischen Entscheidungen, Computer und Recht 2/2020, S.73 ff., S. 74 ff.

⁹¹ In der Literatur werden weitere Massnahmen zur Verhinderung von Diskriminierung diskutiert, z.B. durch den Einsatz von Kontrollalgorithmen oder Überprüfung durch staatliche Institutionen oder Drittorganisationen. Siehe dazu m.w.H. BRAUN BINDER et al., Studie KI 2021 (Fn. 14), S. 42.

⁹² Amnesty International, Tax office algorithm led to racial profiling, DutchNews 26. Oktober 2021, <https://www.dutchnews.nl/2021/10/tax-office-algorithm-led-to-racial-profiling-amnesty-international/>; MELISSA HEIKKILÄ, Dutch scandal serves as a warning for Europe over risks of using algorithms, politico 29. März 2022, <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>.

⁹³ Siehe dazu INGO DACHWITZ, Niederlande zahlen Millionenstrafe wegen Datendiskriminierung, Netzpolitik 29. Dezember 2021, <https://netzpolitik.org/2021/kindergeldaffaere-niederlande-zahlen-millionenstrafe-wegen-datendiskriminierung>.

⁹⁴ Zur Plattform siehe <https://www.parcoursup.gouv.fr>.

⁹⁵ MARIO MARTINI/JONAS BOTTA/DAVID NINK/MICHAEL KOLAIN, Automatisch erlaubt?, Bertelsmann Stiftung 2020, S. 12, https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Automatisch_erlaubt_final.pdf.

⁹⁶ Die Plattform wird deshalb in der rechtswissenschaftlichen Literatur insbesondere im Zusammenhang mit dem Zugangsrecht zum algorithmischen Quellcode diskutiert. Siehe dazu z.B. LUCIE CLUZEL-MÉTAYER, The Judicial Review of the Automated Administrative Act, European Review of Digital Administration & Law – ERDAL 2020, S. 101 ff., S. 101; NINA LASBLEIZ/LILIANE OBRECHT, Le droit d'accès aux codes sources des systèmes algorithmiques utilisés par l'administration publique, European Review of Digital Administration & Law – ERDAL

jahren berücksichtigt wurden. Schülerinnen und Schüler aus weniger renommierten Schulen (etwa in ländlichen Gebieten) und mit weniger Prüfungsvorbereitungskursen sahen sich dadurch mit schlechteren Chancen auf einen guten Universitätsplatz konfrontiert, da sie aufgrund fehlender Ressourcen und Unterstützung oft schlechtere Noten hatten. Durch eine solche Herangehensweise besteht die Gefahr, dass die sozialen Ungleichheiten im Bildungssystem reproduziert werden. In den Jahren 2020 und 2023 wurde der Algorithmus entsprechend angepasst.⁹⁷

4.4. Informationelle Selbstbestimmung und Datenschutz

[35] Im Folgenden wird zunächst auf die allgemeinen Herausforderungen des KI-Einsatzes in Bezug auf die informationelle Selbstbestimmung und den Datenschutz eingegangen (Abschnitt 4.4.1). Anschliessend werden zwei Beispiele präsentiert, die die zuvor dargelegten Herausforderungen veranschaulichen. Dabei handelt es sich um den Einsatz von maschineller Gesichtserkennung im öffentlichen Raum (Abschnitt 4.4.2) und um den Einsatz von Risikomanagementsystemen im Steuerveranlagungsverfahren (Abschnitt 4.4.3).

4.4.1. Die Verfassungsbestimmung und ihre gesetzliche Konkretisierung

[36] Das Grundrecht auf informationelle Selbstbestimmung gemäss Art. 13 Abs. 2 BV ist weiter gefasst, als es der Wortlaut auf den ersten Blick vermuten lässt. Der verfassungsrechtliche Anspruch erstreckt sich nicht nur auf den Schutz vor Missbrauch, sondern auf jeglichen Umgang mit persönlichen Daten durch den Staat (oder Private⁹⁸). Dazu zählt z.B. die Erfassung, Bearbeitung, Bekanntgabe oder Speicherung von persönlichen Daten.⁹⁹ Das Grundrecht wird in den jeweiligen anwendbaren Datenschutzgesetzen, auf Bundesebene im DSG, konkretisiert.¹⁰⁰ Das DSG schützt die Daten natürlicher Personen, d.h. Angaben, die sich auf eine bestimmte oder bestimmbar natürliche Person beziehen.¹⁰¹ Da das Training, die Validierung oder zumindest das Testen von KI-Systemen auf (grosse) Datenmengen angewiesen ist, ergeben sich aus Art. 13 Abs. 2 BV und den datenschutzrechtlichen Konkretisierungen weitere Anforderungen, sofern sich unter den Daten auch Personendaten befinden. Datenschutzrechtliche Vorgaben sind ferner immer dann zu beachten, wenn beim konkreten Einsatz des KI-Systems Personendaten bearbeitet werden. Demnach müssen sowohl bei der Entwicklung des Systems als auch bei dessen Anwendung

2023, S. 165 ff., S. 175. Seit dem 21. Mai 2018 ist der Code öffentlich zugänglich, <https://www.enseignementsup-recherche.gouv.fr/fr/parcoursup-publication-du-code-informatique-des-algorithmes-46680>.

⁹⁷ 2020: Es wurde eine Quotenregelung für Schülerinnen und Schülern aus benachteiligten Vororten und die stärkere Berücksichtigung anderer Faktoren als die Noten eingeführt (siehe WALLY BORDAY, *Parcoursup: quelles sont les nouveautés de la plateforme en 2020?*, le figaro étudiant 20. Januar 2020, https://etudiant.lefigaro.fr/article/parcoursup-queelles-sont-les-nouveautes-de-la-plateforme-en-2020_557794fe-3b5f-11ea-81a1-30c9bbb28c82/; Die Gewichtung der Noten wurde wiederum weiter reduziert und die Berücksichtigung von Motivation und Talent gestärkt (siehe Ministère de l'éducation nationale et de la jeunesse, *Parcoursup 2023*, <https://www.education.gouv.fr/reussir-au-lycee/parcoursup-2023-une-information-plus-claire-plus-riche-et-plus-transparente-343939>).

⁹⁸ Auf die Problematik der Drittwirkung von Grundrechten wird im vorliegenden Beitrag nicht eingegangen.

⁹⁹ Siehe nur PHILIPP GLASS, *Kommentierung Art. 5 DSG N 4 ff.*, in: Adrian Bieri/Julian Powell (Hrsg.), *Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlassen*, Zürich 2023 (zit. OFK DSG).

¹⁰⁰ Das DSG ist für Bundesorgane und Private anwendbar (Art. 2 Abs. 1 lit. a und b DSG). Die kantonalen Behörden unterliegen eigenen Datenschutzgesetzen, im Kanton Basel-Stadt z.B. das Gesetz über die Information und den Datenschutz (IDG, siehe dort § 2 IDG).

¹⁰¹ Art. 2 Abs. 1 i. V. m. Art. 5 lit. a DSG.

insbesondere die in Art. 6 DSGVO verankerten Grundsätze beachtet werden. Im Folgenden wird anhand zweier Beispiele veranschaulicht, dass das informationelle Selbstbestimmungsrecht und die datenschutzrechtlichen Grundsätze beim Einsatz von KI vor Herausforderungen gestellt werden.¹⁰²

4.4.2. Beispiel 1: Maschinelle Gesichtserkennung

[37] Es ist längst Realität geworden, dass Staaten Technologien zur Gesichtserkennung einsetzen.¹⁰³ Manche Gesichtserkennungssysteme können in Echtzeit Daten aus Überwachungskameras auswerten, um eine bestimmte Person zu identifizieren.¹⁰⁴ Gesichtserkennungstechnologien können aber auch im Anschluss an eine Foto- oder Videoaufnahme eingesetzt werden.¹⁰⁵ Maschinelle Gesichtserkennung kann grundsätzlich als eine «automatische Verarbeitung digitaler Bilder, die Gesichter von natürlichen Personen enthalten, um bei diesen eine Identifizierung, Authentifizierung bzw. Verifizierung oder Kategorisierung durchzuführen»¹⁰⁶ definiert werden. Unabhängig von der spezifischen Gesichtserkennungs-Technologie, stellen sich rechtliche Fragen.¹⁰⁷ Von entscheidender Bedeutung ist, dass beim Einsatz von maschineller Gesichtserkennung das Gesicht der betroffenen Person vermessen wird, was bedeutet, dass biometrische und damit besonders schützenswerte Personendaten bearbeitet werden. An diese Bearbeitung werden aus grundrechtlicher und datenschutzrechtlicher Sicht besonders hohe Anforderungen gestellt.¹⁰⁸

[38] Ausgangspunkt einer Datenbearbeitung bildet deren Rechtmässigkeit.¹⁰⁹ Für die Bearbeitung besonders schützenswerter Personendaten durch Bundesorgane ist eine formell-gesetzliche Grundlage erforderlich.¹¹⁰ Diese Rechtsgrundlage muss den konkreten Bearbeitungszweck definieren, die beteiligten Organe nennen, die Kategorien der bearbeiteten Daten bezeichnen und den Bearbeitungsumfang eingrenzen.¹¹¹ Weiter muss eine Datenbearbeitung auch verhältnismäs-

¹⁰² Die Ausführungen lehnen sich an die folgenden zwei Publikationen an: NADJA BRAUN BINDER/ELIANE KUNZ/LILIANE OBRECHT, Maschinelle Gesichtserkennung im öffentlichen Raum, *sui generis* 2022, S. 54 ff., siehe <https://sui-generis.ch/article/view/sg.204/2576>; NADJA BRAUN BINDER/LILIANE OBRECHT, Algorithmisch überprüfte Steuererklärung im ordentlich gemischtem Veranlagungsverfahren, *zsis* 2023, <https://www.zsis.ch/artikel/algorithmisch-ueberpruefte-steuererklaerung-im-ordentlichen-verfahren> (zit. Steuererklärung).

¹⁰³ Siehe dazu die Übersichtskarte des britischen Unternehmens Surfshark, <https://surfshark.com/facial-recognition-map>.

¹⁰⁴ Z.B. im Rahmen der automatisierten Passkontrollen an Flughäfen, wie dies am Basler und Genfer Flughafen der Fall ist. Allerdings ist die französische Grenzpolizei Betreiberin dieser Geräte, siehe dazu DANIEL BALLMER, Ein Scan fürs Gesicht: Der Basler Euro-Airport beschleunigt die Passkontrolle, *bz* vom 9. Februar 2019, <https://www.bzbasel.ch/basel/basel-stadt/ein-scan-furs-gesicht-der-basler-euro-airport-beschleunigt-die-passkontrolle-ld.1343811>.

¹⁰⁵ MONIKA SIMMLER/GIULIA CANOVA, Gesichtserkennungstechnologie: Die «smarte» Polizeiarbeit auf dem rechtlichen Prüfstand, *Sicherheit & Recht* 2021, S. 105 ff.

¹⁰⁶ Artikel-29-Datenschutzgruppe, Stellungnahme 02/2012 zur Gesichtserkennung bei Online- und Mobilfunkdiensten vom 22. März 2012, 00727/12/DE, WP 192, Brüssel, S. 2

¹⁰⁷ BRAUN BINDER/KUNZ/OBRECHT (Fn. 102), Rz. 7.

¹⁰⁸ Art. 5 lit. c Ziff. 4 DSGVO bezeichnet «biometrische Daten» als besonders schützenswert.

¹⁰⁹ Art. 6 Abs. 1 DSGVO.

¹¹⁰ Art. 34 Abs. 2 lit. a DSGVO.

¹¹¹ In Anlehnung an ESTHER ZYSSET, Kommentierung zu Art. 34 DSGVO, in: Thomas Steiner/Anne-Sophie Morand/Daniel Hürlimann (Hrsg.), *Onlinekommentar zum Bundesgesetz über den Datenschutz – Version: 16. August 2023*, siehe <https://onlinekommentar.ch/de/kommentare/dsg34>, Rz. 18 f.

sig sein.¹¹² Dabei stellt sich insbesondere die Frage, ob die mit der maschinellen Gesichtserkennung verfolgten Ziele (z.B. der Schutz der öffentlichen Ordnung und Sicherheit) nicht auch mit anderen Mitteln erreicht werden können.

[39] Ferner ist zu berücksichtigen, dass KI-Systeme zur Gesichtserkennung eine signifikante Menge an Trainingsdaten benötigen, um verlässliche Ergebnisse zu generieren. Dies steht u.a. im Widerspruch zu den Prinzipien der Datenvermeidung und -sparsamkeit, die als Teilaspekte des Verhältnismässigkeitsprinzips zu betrachten sind. Erwähnenswert ist in diesem Zusammenhang, dass gemäss Art. 5 Abs. 1 Bst. e AI Act KI-Systeme, «die Datenbanken zur Gesichtserkennung durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder von Überwachungskameras erstellen oder erweitern», zu den in der EU verbotenen Praktiken im KI-Bereich zählen. Der AI Act verbietet grundsätzlich auch die Verwendung sog. biometrischer Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken (Art. 5 Abs. 1 Bst. h AI Act), wobei diesbezüglich verschiedene Ausnahmen, etwa für die gezielte Suche nach bestimmten Opfern von Entführung, Menschenhandel oder sexueller Ausbeutung, vorgesehen sind.

[40] Auch die datenschutzrechtlichen Grundsätze der Zweckbindung und Erkennbarkeit der Datenbearbeitung sind zu erwähnen.¹¹³ Der Zweck der Datenbearbeitung muss dabei genügend bestimmt dargelegt werden.¹¹⁴ Die allgemeine Gefahrenabwehr als Zweck ist bei maschineller Gesichtserkennung zu wenig bestimmt.¹¹⁵ Datenbearbeitungen müssen ferner erkennbar sein.¹¹⁶ In gewissen Situationen besteht eine erweiterte Informationspflicht. In diesen Fällen wird die Erkennbarkeit der Datenbeschaffung subsidiär.¹¹⁷ Die Information muss einfach, vollständig und der Hinweis auf die Information augenfällig sein.¹¹⁸ Allgemeine Hinweisschilder über den Einsatz von Gesichtserkennung im öffentlichen Raum reichen angesichts dieser Anforderungen nicht aus, um der Informationspflicht nachzukommen.¹¹⁹

[41] Schliesslich postuliert das Datenschutzrecht den Grundsatz der Datenrichtigkeit,¹²⁰ der die Aktualität und Vollständigkeit der Daten umfasst.¹²¹ Da der Einsatz maschineller Gesichtserkennung eine hohe Wahrscheinlichkeit der richtigen Identifizierung oder Verifizierung der betroffenen Person fordert, sind sehr hohe Anforderungen an die Qualität der Trainings-, Validierungs- und Testdatensätze für solche KI-Systeme zu stellen.¹²²

¹¹² Art. 6 Abs. 2 DSGVO. Daten dürfen folglich nur bearbeitet werden, wenn die Datenbearbeitung geeignet ist, den verfolgten Zweck zu erfüllen und kein milderes Mittel dafür zur Verfügung steht. Schliesslich muss die Datenbearbeitung auch zumutbar sein, d. h. es muss ein vernünftiges Verhältnis zwischen der Datenbearbeitung und dem Grundrechtseingriff bestehen.

¹¹³ Art. 6 Abs. 3 DSGVO.

¹¹⁴ OFK DSG-FANGER (Fn. 99), Art. 6 N 7 f.

¹¹⁵ BRAUN BINDER/KUNZ/OBRECHT (Fn. 102), Rz. 18.

¹¹⁶ Art. 6 Abs. 3 DSGVO.

¹¹⁷ Art. 19 DSGVO, siehe auch die Ausnahmen in Art. 20 DSGVO.

¹¹⁸ Botschaft zum Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz (BBl 2017 6941 ff.), S. 7050 f.

¹¹⁹ BRAUN BINDER/KUNZ/OBRECHT (Fn. 102), Rz. 19.

¹²⁰ Art. 6 Abs. 5 DSGVO.

¹²¹ OFK DSG-FANGER (Fn. 99), Art. 6 N 11.

¹²² BRAUN BINDER/KUNZ/OBRECHT (Fn. 102), Rz. 20.

4.4.3. Beispiel 2: Risikomanagementsysteme in der Steuerverwaltung

[42] Die fortschreitende Digitalisierung der Steuerverwaltungen geht mit einer zunehmenden Automatisierung des Steuerveranlagungsverfahrens einher.¹²³ Die vollständige Automatisierung von Steuerverfahren erfordert regelmässig den Einsatz sogenannter Risikomanagementsysteme.¹²⁴ Diese überprüfen automatisiert die durch die steuerpflichtige Person eingereichte Steuererklärung, um abschätzen zu können, ob eine steuerpflichtige Person die Steuererklärung wahrheitsgetreu ausgefüllt hat oder nicht. Die heiklen Fälle werden aussortiert und einer natürlichen Person zur Kontrolle vorgelegt.¹²⁵ Grundsätzlich ist der Einsatz von Risikomanagementsystemen auch in der Schweiz vorstellbar, da eine Vollautomation ohne entsprechende Kontrollmöglichkeiten das Risiko birgt, dass falsche Angaben in Steuerklärungen nicht erkannt werden.¹²⁶

[43] Die (kantonalen) Steuerbehörden erheben im Steuerveranlagungsverfahren regelmässig Personendaten zur Festlegung des massgeblichen Sachverhalts. Allein aus diesem Grund ist das Datenschutzrecht relevant.¹²⁷ Zu betonen ist, dass die Kantone für die Datenbearbeitung durch kantonale Organe ihre eigenen Datenschutzgesetze haben, die allerdings mit dem Bundesrecht vergleichbar sind.¹²⁸ Bei einem Einsatz von Risikomanagementsystemen stehen insbesondere die datenschutzrechtlichen Grundsätze der Zweckbindung und der Verhältnismässigkeit vor Herausforderungen.¹²⁹ Die Steuerbehörden erheben die Daten für eine bestimmte Steuerart in einer bestimmten Steuerperiode, um die steuerpflichtigen Personen gleichmässig und verhältnismässig zu veranlagern.¹³⁰ Sie dürfen die Steuerdaten ferner nicht über die gesetzlich vorgesehene Dauer aufbewahren.¹³¹ Im automatisierten Steuerverfahren könnten hingegen Daten verschiedener steuerpflichtiger Personen genutzt werden, um die Steuererklärung einer weiteren Person auf Plausibilität zu prüfen. Darin würde eine Änderung des ursprünglichen Zwecks der Datenerhebung liegen. Der Zweck würde nicht mehr die Bearbeitung der Daten einer bestimmten steuerpflichtigen Person in einer bestimmten Steuerperiode umfassen, sondern die Bearbeitung der Daten verschiedener steuerpflichtiger Personen bzw. die Bearbeitung der Daten einer Person zur Nutzung für eine andere Person. Grundsätzlich ist eine Bearbeitung von Personendaten zu einem

¹²³ BRAUN BINDER/OBRECHT, Steuererklärung (Fn. 102), Rz. 1; siehe ferner die abgeschlossene Vernehmlassung im Kanton Solothurn zur Schaffung der Rechtsgrundlagen für eine vollautomatische Veranlagung unter Nutzung von maschinellen Lernverfahren (Fn 17).

¹²⁴ In Deutschland ist am 1. Januar 2017 eine Rechtsgrundlage in Kraft getreten, die die verpflichtende Einführung von Risikomanagementsystemen enthält (§ 88 Abs. 5 Abgabenordnung in der Fassung der Bekanntmachung vom 1. Oktober 2002 (BGBl. I S. 3866; 2003 I S. 61), die zuletzt durch Artikel 1 des Gesetzes vom 12. Juli 2022 (BGBl. I S. 1142) geändert worden ist). In Österreich besteht ebenfalls eine entsprechende Rechtsgrundlage (§§ 48a, 114 Abs. 4 Bundesgesetz vom 28. Juni 1961 betreffend allgemeine Bestimmungen und das Verfahren für die von den Abgabenbehörden des Bundes verwalteten Abgaben, StF: BGBl. Nr. 194/1961).

¹²⁵ TINA EHRKE-RABEL, Die Automatisierung des Verwaltungsverfahrens am Beispiel des österreichischen Abgabenrechts, in: Braun Binder Nadja/Bussjäger Peter/Eller Matthias (Hrsg.), Auswirkungen der Digitalisierung auf die Erlassung und Zuordnung behördlicher Entscheidungen, Wien/Hamburg 2021, S. 21 ff.; Europäische Kommission, Risk Management Guide For Tax Administrations, Version 1.02, Februar 2006, S. 7, https://taxation-customs.ec.europa.eu/system/files/2016-09/risk_management_guide_for_tax_administrations_en.pdf.

¹²⁶ Auf die Fehleranfälligkeit solcher Systeme wird im Folgenden nicht weiter eingegangen, siehe dazu BRAUN BINDER/OBRECHT, Steuererklärung (Fn. 102), Rz. 27.

¹²⁷ Art. 2 Abs. 1 DSGVO, zum Begriff der Personendaten Art. 5 lit. a DSGVO.

¹²⁸ Siehe z.B. das Informations- und Datenschutzgesetz des Kantons Basel-Stadt IDG.

¹²⁹ Siehe zu den datenschutzrechtlichen Grundsätzen Abschnitt 4.4.2. Die kantonalen Datenschutzgesetze sehen ähnliche Regelungen vor; das Verhältnismässigkeitsprinzip geniesst nach Art. 5 Abs. 2 BV in jedem Fall Verfassungsrang.

¹³⁰ Vertieft BSK BV-BEHNISCH (Fn. 46), Art. 127 N 24, 27.

¹³¹ BSK BV-BEHNISCH (Fn. 46), Art. 127 N 24.

anderen Zweck nicht durch die gesetzliche Grundlage gedeckt, welche die ursprüngliche Erhebung rechtfertigt.¹³² Unzulässig ist auch eine Datenerhebung «auf Vorrat», um diese später in einem anderen Kontext zu verwenden.¹³³

[44] Ferner kann der Einsatz von Risikomanagementsystemen dazu führen, dass Vorhersagen über persönlichkeitsrelevante Aspekte wie Ehrlichkeit und Zuverlässigkeit einer steuerpflichtigen Person gemacht werden.¹³⁴ Damit würde es sich um ein Profiling im Sinne des DSGVO handeln,¹³⁵ das zumindest auf Bundesebene, aber auch in den meisten Kantonen, eine formell-gesetzliche Grundlage voraussetzt.¹³⁶

4.5. Vertrauensschutz

[45] Im Kontext des staatlichen KI-Einsatzes etablieren sich vermehrt Chatbots zur behördlichen Informationsvermittlung und Auskunftserteilung.¹³⁷ Seit der Einführung von ChatGPT im November 2022 sind insbesondere sogenannte Large Language Models (LLMs) beliebt.¹³⁸ Aus rechtlicher Sicht ist dabei zentral, wie mit fehlerhaften Chatbot-Auskünften umzugehen ist. Zu betonen ist, dass das im Zusammenhang mit menschlichen Auskünften entwickelte Konzept des Vertrauensschutzes, als Teilaspekt von Treu und Glauben nach Art. 9 BV,¹³⁹ auch beim Einsatz von Chatbots Anwendung findet.¹⁴⁰ Demnach müssen Private¹⁴¹ unter bestimmten Voraussetzungen in ihrem Vertrauen in behördliche Auskünfte geschützt werden.¹⁴²

¹³² Eine Zweckänderung kann gesetzlich zwar vorgesehen sein, wenn Steuerdaten von der Steuerverwaltung an andere Behördenstellen «bekannt gegeben» (im Bund: Art. 5 lit. e DSGVO) werden dürfen, z.B. zum Zweck der Beurteilung, ob jemand bedürftig ist (z.B. bei der Bekanntgabe an die Sozialhilfe). Allerdings fällt die Bearbeitung von Daten der Steuerverwaltung, die zum Zweck der Steuerveranlagung erhoben wurden, durch dieselbe Steuerverwaltung zu einem anderen Zweck – etwa der Überprüfung von Steuererklärungen anderer Steuerpflichtiger – nicht unter die Bekanntgabe.

¹³³ ASTRID EPINEY, Datenschutzrechtliche Grundsätze und Garantien, in: Eva Maria Belser/Astrid Epiney/Bernhard Waldmann/Jürg Bickel (Hrsg.), *Datenschutzrecht: Grundlagen und öffentliches Recht*, Bern 2011, S. 509 ff., § 9 N 33. Dies gilt auch auf kantonaler Ebene, siehe etwa BEAT RUDIN, Kommentierung § 3 IDG-BS N 17 ff., in: Beat Rudin/Bruno Baeriswyl (Hrsg.), *Praxiskommentar zum Informations- und Datenschutzgesetz des Kantons Basel-Stadt*, Zürich 2014.

¹³⁴ BRAUN BINDER/OBRECHT, *Steuererklärung* (Fn. 102), Rz. 30 ff.

¹³⁵ Für den Bund: Art. 5 lit. f DSGVO.

¹³⁶ Art. 34 Abs. 2 lit. b DSGVO. Eine Grundlage im materiellen Sinn genügt, sofern die Bearbeitung für eine in einem Gesetz im formellen Sinn festgelegte Aufgabe unentbehrlich ist (Art. 34 Abs. 3 lit. a DSGVO) oder der Bearbeitungszweck für die Grundrechte der betroffenen Person keine besonderen Risiken birgt (Art. 34 Abs. 3 lit. b DSGVO). Die Ausnahmen werden aber kaum greifen, siehe dazu BRAUN BINDER/OBRECHT, *Steuererklärung* (Fn. 102), Rz. 36.

¹³⁷ BRAUN BINDER/OBRECHT/WITTMER (Fn. 20).

¹³⁸ So setzt z.B. die Eidgenössische Stiftungsaufsicht den Chatbot «Esi» ein, um Stiftungen und Revisionsorganen zu helfen, rasch und unkompliziert Antworten auf Fragen rund um das Schweizer Stiftungswesen zu finden, siehe Eidgenössische Stiftungsaufsicht, <https://www.fragesi.ch/>.

¹³⁹ ELISABETH CHIARIELLO, *Treu und Glauben als Grundrecht nach Art. 9 der schweizerischen Bundesverfassung*, Diss. Bern 2004, S. 27; René Rhinow/Markus Schefer/Peter Uebersax, *Schweizerisches Verfassungsrecht*, 3. Aufl., Basel 2016, Rz. 1996.

¹⁴⁰ Siehe zum Ganzen BRAUN BINDER/OBRECHT/WITTMER (Fn. 20).

¹⁴¹ Terminologie angelehnt an ULRICH HÄFELIN/GEORG MÜLLER/FELIX UHLMANN, *Allgemeines Verwaltungsrecht*, 8. Aufl., Zürich 2020, Rz. 624 ff.

¹⁴² Die Voraussetzungen lauten: 1) Vorliegen einer Vertrauensgrundlage; 2) Vertrauenswürdigkeit der Auskunft; 3) Zuständigkeit der Behörde; 4) Vorbehaltslosigkeit der Auskunft; 5) Vertrauensbetätigung; 6) Kausalität zwischen Vertrauen und Vertrauensbetätigung; 7) Keine Änderung des Sachverhaltes oder der Gesetzgebung; 8) Interessenabwägung (zum Ganzen HÄFELIN/MÜLLER/UHLMANN (Fn. 143), Rz. 624 ff.).

[46] Zu hinterfragen ist der Einsatz von KI-Chatbots¹⁴³ zur behördlichen Auskunft deshalb, weil die Auskünfte aufgrund der technischen Natur der Chatbots fehlerhaft sein können.¹⁴⁴ Dies trifft im besonderen Masse auf Chatbots zu, die LLMs nutzen.¹⁴⁵ In diesem Kontext wird insbesondere von *Halluzinationen* gesprochen.¹⁴⁶ Aktuell werden Methoden entwickelt, um die Fehlerhaftigkeit solcher Modelle zu verringern. Eine solche Methode ist die *Retrieval Augmented Generation* (RAG).¹⁴⁷ Dabei wird eine Anfrage an einen Chatbot zunächst sprachlich analysiert und die relevanten Antwortphrasen werden mittels semantischer Suche auf einem lokalen Textkorpus bestimmt. Das LLM beantwortet die Anfrage daraufhin basierend auf den ihm zur Verfügung gestellten Textsegmenten sowie auf Basis der ursprünglichen Eingabe. Es wird insbesondere dazu verwendet, die Anfrage sprachlich zu gestalten und die Antwort zu strukturieren.¹⁴⁸ Bislang ist allerdings unklar, ob die Fehlerhaftigkeit der Chatbot-Antworten durch RAG nur verringert oder gar gänzlich ausgeschlossen werden kann.¹⁴⁹

[47] Da Chatbots als ebenbürtige Alternative zu menschlichen Auskünften eingesetzt werden sollen, können sich Behörden auch nicht durch den Einsatz von allgemeinen Vorbehalten (sog. Disclaimer) des Vertrauensschutzes entziehen.¹⁵⁰ Vielmehr gilt es zu hinterfragen, ob tatsächlich alle Chatbot-Systeme zur behördlichen Auskunfterteilung geeignet sind.¹⁵¹

4.6. Transparenz

[48] In Bezug auf den behördlichen Einsatz von KI wird häufig die Forderung nach Transparenz geäussert. Damit soll die Rechtmässigkeit des KI-Einsatzes sichergestellt werden.¹⁵² Dabei ist zwischen zwei Dimensionen zu unterscheiden. Einerseits ist die Herstellung von individueller Transparenz beim Einsatz von KI-Systemen zu gewährleisten. Diese Forderung lässt sich zunächst auf die datenschutzrechtlichen Grundsätze und Informationspflichten stützen.¹⁵³ Sodann

¹⁴³ Gemäss BRAUN BINDER/OBRECHT/WITTMER (Fn. 20), handelt es sich um datenbasierte *Machine Learning* Algorithmen, die den Chatbots zugrunde liegen.

¹⁴⁴ Zum Ganzen BRAUN BINDER/OBRECHT/WITTMER (Fn. 20).

¹⁴⁵ Zum Ganzen BRAUN BINDER/OBRECHT/WITTMER (Fn. 20).

¹⁴⁶ Siehe dazu etwa ZIWEI JI et al., Survey of Hallucination in Natural Language Generation, ACM Computing Survey 2022, S. 1 ff.

¹⁴⁷ Siehe dazu etwa PENGHAO ZHAO et al., Retrieval-Augmented Generation for AI-Generated Content: A Survey, arXiv 29. Februar 2024, S. 1, <https://arxiv.org/abs/2402.19473>.

¹⁴⁸ Zum Ganzen etwa MICHAEL MCTEAR/MARINA ASHURKINA, Transforming Conversational AI: Exploring the Power of Large Language Models in Interactive Conversational Agents, Berkley 2024, S. 66, ZHAO et al. (Fn. 147), S. 3.

¹⁴⁹ Gemäss ZHAO et al. (Fn. 147), S. 1, kann RAG die Halluzinationen zwar verringern, allerdings nicht vollständig ausschliessen.

¹⁵⁰ Zum Ganzen BRAUN BINDER/OBRECHT/WITTMER (Fn. 20).

¹⁵¹ Zum Ganzen BRAUN BINDER/OBRECHT/WITTMER (Fn. 20).

¹⁵² Z.B. Schweiz: Eidgenössisches Departement für Wirtschaft, Bildung und Forschung WBF, Herausforderungen der künstlichen Intelligenz Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat, 2019 (Bericht IDAG KI 2019), <https://www.sbf.admin.ch/sbfi/de/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html>, S. 37 f. Deutschland: Gutachten der Datenethikkommission der Bundesregierung, Berlin 2019, S. 169 f, https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6.

¹⁵³ Siehe Abschnitt 4.4 sowie BRAUN BINDER/OBRECHT, SVVOR Jahrbuch (Fn. 76), S. 39.

präsentiert sie sich auch als wichtiger Aspekt der verfassungsrechtlich gewährleisteten Begründungspflicht von Verfügungen.¹⁵⁴

[49] Andererseits ist für die Schaffung von Transparenz im Kontext des staatlichen KI-Einsatzes die Etablierung allgemeiner gesellschaftlicher Kontrollmöglichkeiten erforderlich. In der Literatur¹⁵⁵, in der Politik¹⁵⁶ sowie seitens zivilgesellschaftlicher Organisationen¹⁵⁷ wird deshalb die Schaffung eines öffentlichen Verzeichnisses (auch: KI-Register) gefordert. Im Kanton Zürich liegt ein Gesetzesentwurf für die Pflicht zur Führung eines entsprechenden öffentlichen Verzeichnisses vor.¹⁵⁸ In einem solchen KI-Verzeichnis sollen die in der öffentlichen Verwaltung eingesetzten KI-Systeme aufgeführt werden. Des Weiteren soll dargelegt werden, in welchen Bereichen der öffentlichen Verwaltung KI-Systeme zum Einsatz kommen, welche Behörde dafür verantwortlich ist, welche Art von Daten bearbeitet wird, welche Metadaten verwendet werden, welche Aufsichtsstelle involviert ist und welche Ergebnisse eine etwaige Folgenabschätzung für die Grundrechte der betroffenen Personen hervorgebracht hat.¹⁵⁹

[50] Obgleich das CNAI für die Bundesverwaltung bereits über 60 Projekte in einer öffentlich zugänglichen «Datenbank» führt,¹⁶⁰ ist davon auszugehen, dass diese nicht vollständig ist. Dies liegt daran, dass das Verzeichnis für die Bundesstellen nicht verpflichtend ist. Ausserdem sind die darin enthaltenen Angaben von unterschiedlicher Ausführlichkeit und Detailtiefe.

5. Entwicklungen aus der Perspektive der Bundesverwaltung

[51] In Anbetracht der fortschreitenden Entwicklungen hat der Bundesrat im Jahr 2018 KI als ein zentrales Thema der Strategie «Digitale Schweiz» definiert und eine interdepartementale Arbeitsgruppe (IDAG) ins Leben gerufen.¹⁶¹ In einem Bericht vom Dezember 2019 präsentierte die Arbeitsgruppe eine Auslegeordnung der relevanten Rahmenbedingungen im Hinblick auf die zunehmende Nutzung von KI. Diese beleuchtete die spezifischen Herausforderungen in verschiedenen Anwendungsfeldern über Politikbereiche der gesamten Bundesverwaltung hinweg und diskutierte möglichen Anpassungsbedarf auf Bundesebene.¹⁶² Die IDAG gelangte zum Schluss, dass die Schweiz insgesamt gut aufgestellt sei, um die Herausforderungen der KI zu bewältigen. Der allgemeine Rechtsrahmen sei zum damaligen Zeitpunkt grundsätzlich geeignet und ausreichend,

¹⁵⁴ Siehe Abschnitt 4.2.1. sowie BRAUN BINDER/OBRECHT, Begründung (Fn. 67), i.E.

¹⁵⁵ BRAUN BINDER et al., Studie KI 2021 (Fn. 14), S. 50; THOUVENIN et al., Positionspapier 2021 (Fn. 10), S. 3.

¹⁵⁶ Z.B. auf Bundesebene Motion 21.4508 Schlatter Marionna vom 16. Dezember 2021 betreffend ein öffentliches Verzeichnis der in der Verwaltung eingesetzten Algorithmen; im Kanton Basel-Stadt Anzug 23.5474.01 betreffend Transparenz über den Einsatz von algorithmischen Systemen in der Verwaltung vom 18. Oktober 2023.

¹⁵⁷ Etwa die zivilgesellschaftliche Organisation AlgorithmWatch Schweiz, siehe Automatisierte Entscheidungssysteme im öffentlichen Sektor – einige Empfehlungen, 25. Februar 2022, <https://algorithmwatch.ch/de/adm-offentlichersektor-empfehlungen/>.

¹⁵⁸ Im kantonalen Information- und Datenschutzgesetz IDG-ZH (§13 Abs. 3 revidiertes IDG-ZH), siehe Medienmitteilung vom 4. August 2023, <https://www.zh.ch/de/news-uebersicht/medienmitteilungen/2023/08/kanton-zuerich-modernisiert-gesetz-ueber-information-und-datenschutz.html>.

¹⁵⁹ BRAUN BINDER et al., Studie KI 2021 (Fn. 14), S. 50.

¹⁶⁰ Siehe Kapitel 1. Es handelt sich dabei allerdings nicht um eine tatsächliche Datenbank, sondern um ein PDF-Dokument.

¹⁶¹ Der Bundesrat, Strategie «Digitale Schweiz», verabschiedet am 5. September 2018, <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-72053.html>.

¹⁶² Bericht IDAG KI 2019 (Fn. 152).

um mit neuartigen Herausforderungen der KI umzugehen.¹⁶³ Im Jahr 2020 wurden sodann sieben Leitlinien zum Umgang mit KI in der Bundesverwaltung präsentiert. Diese dienen der Gewährleistung der Kohärenz der KI-relevanten Politik über die zahlreichen Arbeitsbereiche hinweg.¹⁶⁴ Seither bieten diese Leitlinien der Bundesverwaltung und den verschiedenen Trägerinnen und Trägern von Verwaltungsaufgaben des Bundes einen allgemeinen Orientierungsrahmen.¹⁶⁵

[52] Im November 2023 hat der Bundesrat bekannt gegeben, bis Ende 2024 in einer Auslegeordnung mögliche Ansätze zur Regulierung von KI für die Schweiz aufzuzeigen.¹⁶⁶ Mit der Analyse will der Bundesrat die Basis schaffen, um 2025 einen konkreten Auftrag für eine KI-Regulierungsvorlage erteilen und die Zuständigkeiten regeln zu können. Die Analyse wird auf bestehendem Schweizer Recht aufbauen und mögliche Regulierungsansätze für die Schweiz aufzeigen, die mit dem AI Act der EU und der KI-Konvention des Europarats kompatibel sind. Der Bundesrat hält fest, dass beide internationalen Regelwerke für die Schweiz von Relevanz sind.¹⁶⁷

6. Fazit

[53] Kantonale und öffentliche Verwaltungen setzen KI bereits an verschiedenen Stellen ein. Der Einsatz von KI in der öffentlichen Verwaltung ist zwar mit verschiedenen Chancen verbunden. Der vorliegende Beitrag zeigt jedoch, dass aus rechtlicher Sicht die Risiken solcher Technologien für den Einzelnen und die Gesellschaft als Ganzes von zentraler Bedeutung sind. Es ist daher zu begrüssen, dass der Bundesrat bis Ende 2024 in einer Auslegeordnung mögliche Ansätze für die Schaffung von Rechtsgrundlagen zur Nutzung von KI in der Schweiz aufzeigen wird. Mit Blick auf den staatlichen KI-Einsatz sind Rechtsgrundlagen notwendig, um sicherzustellen, dass dieser rechtmässig, transparent und verantwortungsvoll erfolgt. Darüber hinaus müssen insbesondere die Verhinderung von Diskriminierung, die Beachtung der Verfahrensgrundrechte und die Wahrung des Datenschutzes sichergestellt werden. Mit der Schaffung der entsprechenden Grundlagen sollte nach Vorlage der Auslegeordnung zügig begonnen werden.

Prof. Dr. NADJA BRAUN BINDER, MBA, ist Professorin für öffentliches Recht an der Universität Basel und Co-Leiterin der Forschungsstelle für Digitalisierung in Staat und Verwaltung (e-PIAF), nadja.braunbinder@unibas.ch.

¹⁶³ Siehe dazu Der Bundesrat, Künstliche Intelligenz: Schweiz befindet sich in guter Ausgangslage, Medienmitteilung vom 13. Dezember 2019, <https://www.sbf.admin.ch/sbfi/de/home/aktuell/medienmitteilungen/news-anzeige-nsb.msg-id-77514.html>.

¹⁶⁴ Der Bundesrat, Leitlinien «Künstliche Intelligenz für den Bund», 25. November 2020, <https://www.sbf.admin.ch/sbfi/de/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html>.

¹⁶⁵ Die Leitlinien (siehe Fn. 164) lauten: 1) Der Mensch steht im Mittelpunkt; 2) Der Bund gewährleistet bestmögliche Rahmenbedingungen für die Entwicklung und den Einsatz von KI; 3) Auf KI gestützte Entscheidungsprozesse sind transparent, nachvollziehbar und erklärbar; 4) Beim Einsatz von KI muss die Verantwortlichkeit klar definiert sein; 5) KI-Systeme müssen sicher, robust und resilient konzipiert sein; 6) Die Schweiz gestaltet die globale KI-Gouvernanz aktiv mit; 7) Die Schweiz setzt sich dafür ein, dass alle betroffenen nationalen und internationalen Akteure in die Debatte um KI-Gouvernanz miteinbezogen werden.

¹⁶⁶ Der Bundesrat, Bundesrat prüft Regulierungsansätze für Künstliche Intelligenz, Medienmitteilung vom 22. November 2023, <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-98791.html>.

¹⁶⁷ Siehe Medienmitteilung in Fn.166.

NINA LAUKENMANN, MLaw LL.M., ist wissenschaftliche Assistentin und Doktorandin an der Juristischen Fakultät der Universität Basel, nina.laukenmann@unibas.ch.

LILIANE OBRECHT, MLaw, ist wissenschaftliche Mitarbeiterin und Doktorandin an der Juristischen Fakultät der Universität Basel, liliane.obrecht@unibas.ch.