

CENTER FOR
INFORMATION
TECHNOLOGY
SOCIETY AND
LAW — ITSL

Volume 12

Nadja Braun Binder / Florent Thouvenin (Hrsg.)

Ein Rechtsrahmen für KI
in der Schweiz:
Perspektive öffentliches Recht

CENTER FOR INFORMATION TECHNOLOGY SOCIETY AND LAW — ITSL

Schriften aus dem ITSL, herausgegeben
von Florent Thouvenin und Rolf H. Weber

Volume 12

Nadja Braun Binder / Florent Thouvenin (Hrsg.)

Ein Rechtsrahmen für KI in der Schweiz: Perspektive öffentliches Recht

EIZ  Publishing



Ein Rechtsrahmen für KI in der Schweiz: Perspektive öffentliches Recht Copyright © by Florent Thouvenin und Rolf H. Weber is licensed under a [Creative Commons Namensnennung-Nicht kommerziell-Keine Bearbeitung 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/), except where otherwise noted.

© 2025 – CC BY-NC-ND

Verlag: EIZ Publishing (eizpublishing.ch), Hirschengraben 56, 8001 Zürich, eizpublishing@europa-institut.ch

Herausgeber: Center for Information Technology Society and Law ITSL, Florent Thouvenin (Hrsg.)

ISBN:

978-3-03805-803-8 (Print – Softcover)

978-3-03805-804-5 (PDF)

978-3-03805-805-2 (ePub)

DOI: <https://doi.org/10.36862/eiz-itsl-12>

Version: 1.00 – 20250425

Dieses Werk ist als gedrucktes Buch sowie als E-Book (open access) in verschiedenen Formaten verfügbar. Weitere Informationen finden Sie unter der URL: <https://eizpublishing.ch/publikationen/ein-rechtsrahmen-fuer-ki-in-der-schweiz-perspektive-oeffentliches-recht/>.

Vorwort

Am 12. Februar 2025 hat der Bundesrat beschlossen, die KI-Konvention des Europarats zu ratifizieren. Mit dieser Entscheidung hat er den Startschuss für die Gesetzgebungsarbeiten zur Regulierung von Künstlicher Intelligenz (KI) in der Schweiz gegeben. Bis 2026 soll eine Vernehmlassungsvorlage vorliegen, welche die notwendigen gesetzlichen Massnahmen zur Umsetzung der KI-Konvention des Europarats festlegt. Dazu gehören insbesondere die Bereiche Transparenz, Datenschutz, Nichtdiskriminierung und Aufsicht.

Bereits vor dieser Entscheidung haben wir uns im Rahmen des Forschungsprojekts «Nachvollziehbare Algorithmen: Ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz» (Juli 2021 – Juni 2024) diesen Themen gewidmet. Entstanden sind mehrere Beiträge, die sich aus öffentlich- und privatrechtlicher Sicht mit zentralen Herausforderungen befassen: fehlende Transparenz und Nachvollziehbarkeit der Entscheidungen von KI-Systemen, Risiken für den Schutz der Privatsphäre, Gefahr von Diskriminierung und Manipulation sowie Fragen des Urheberrechts.

Neben dem Beschluss zur Ratifikation der KI-Konvention des Europarats hat der Bundesrat am 12. Februar 2025 auch eine Auslegeordnung zu möglichen Regulierungsansätzen für KI publiziert. Dieser Auslegeordnung lagen drei Analysen zugrunde: eine rechtliche Basisanalyse, eine sektorielle Analyse und eine Länderanalyse. Wir freuen uns, dass Erkenntnisse aus unserem Forschungsprojekt in diese Analysen eingeflossen sind.

Der vorliegende Sammelband bündelt einen Teil der im Rahmen des Forschungsprojekts bislang veröffentlichten Beiträge. Diese befassen sich mit dem Einsatz von KI-Systemen in der öffentlichen Verwaltung. Die privatrechtlichen Erkenntnisse aus dem Forschungsprojekt werden in einem separaten Sammelband zusammengefasst. Da es sich um eine Wiedergabe bereits veröffentlichter Aufsätze handelt, werden Abkürzungen und Literaturangaben wie in der jeweiligen Originalpublikation verwendet. Dies kann zu Abweichungen in der Schreibweise zwischen den einzelnen Beiträgen führen. Zur besseren Übersicht wurde jeder Beitrag um ein einheitlich formatiertes Inhaltsverzeichnis ergänzt.

Die Entwicklungen auf Bundesebene haben gezeigt, dass weiterhin Forschungs- und Vermittlungsbedarf besteht. Das Forschungsprojekt wurde deshalb bis im März 2026 verlängert.

Das Forschungsprojekt wird in Kooperation zwischen der Forschungsstelle für Digitalisierung in Staat und Verwaltung (e-PIAF) der Universität Basel und dem Center for Information Technology, Society, and Law (ITSL) der Universität Zürich durchgeführt. Wir danken unseren Mitarbeitenden – Dr. Stephanie Volz (Uni Zürich), Soraya Weiner, MLaw, und Fabienne Graf, MLaw, LL.M (beide Uni Zürich), sowie Liliane Obrecht, MLaw, und Nina Laukenmann, MLaw, LL.M, (beide Uni Basel) – für ihre wertvolle Unterstützung und ihr Engagement. Ihr Fachwissen und Ihre Ideen haben massgeblich zur Weiterentwicklung unserer Forschung beigetragen. Ein besonderer Dank gilt auch der Stiftung Mercator Schweiz für die Unterstützung des Forschungsprojektes. Herrn Lars Rufli, Hilfsassistent bei e-PIAF, danken wir für die tatkräftige Unterstützung bei der Vorbereitung dieses Sammelbandes.

Basel, April 2025

Prof. Dr. Nadja Braun Binder, Professorin für öffentliches Recht an der Universität Basel

Prof. Dr. Florent Thouvenin, Inhaber des Lehrstuhls für Informations- und Kommunikationsrecht an der Universität Zürich

Inhaltsverzeichnis

<u>Vorwort</u>	V
<u>Autorinnenverzeichnis</u>	IX
<u>Künstliche Intelligenz in der Verwaltung – Möglichkeiten, rechtliche Rahmenbedingungen und Grenzen</u>	1
<i>Nadja Braun Binder / Liliane Obrecht</i>	
<u>KI in der Verwaltung: Entwicklungen und Herausforderungen</u>	27
<i>Nadja Braun Binder / Nina Laukenmann / Liliane Obrecht</i>	
<u>Whitepaper: Transparenz durch Begründung von Verfügungen</u>	57
<i>Nadja Braun Binder / Liliane Obrecht</i>	
<u>Die Begründung von Verfügungen beim Einsatz algorithmischer Systeme</u>	67
<i>Nadja Braun Binder / Liliane Obrecht</i>	
<u>Verfügung und automatisierte Einzelentscheidung – same same but different?</u>	91
<i>Liliane Obrecht</i>	
<u>White Paper: Transparenz durch öffentliches Verzeichnis</u>	107
<i>Nadja Braun Binder / Liliane Obrecht</i>	
<u>Transparenz über den staatlichen Einsatz algorithmischer Entscheidungssysteme</u>	115
<i>Nadja Braun Binder / Liliane Obrecht</i>	
<u>Maschinelle Gesichtserkennung im öffentlichen Raum</u>	141
<i>Nadja Braun Binder / Eliane Kunz / Liliane Obrecht</i>	
<u>Algorithmisch überprüfte Steuererklärungen im ordentlichen gemischten Veranlagungsverfahren</u>	161
<i>Nadja Braun Binder / Liliane Obrecht</i>	
<u>Vertrauensschutz bei fehlerhaften Behördenauskünften durch Chatbots</u>	187
<i>Nadja Braun Binder / Liliane Obrecht / Grace Wittmer</i>	

Autorinnenverzeichnis

Prof. Dr. Nadja Braun Binder, MBA, Professorin für Öffentliches Recht an der Juristischen Fakultät der Universität Basel

Liliane Obrecht, MLaw, Doktorandin an der Juristischen Fakultät der Universität Basel

Nina Laukenmann, MLaw, LL.M., Doktorandin an der Juristischen Fakultät der Universität Basel

Eliane Kunz, MLaw, Kanzleimitarbeiterin am Zivilkreisgericht Basel-Landschaft West

Grace Wittmer, MLaw, Juristische Volontärin

Künstliche Intelligenz in der Verwaltung – Möglichkeiten, rechtliche Rahmenbedingungen und Grenzen

Nadja Braun Binder / Liliane Obrecht

Erstabdruck in: SVVOR (Hrsg.), Jahrbuch 2021/2022, Bern 2023, S. 21–48

Inhalt

I.	Einleitung	2
II.	Begriffsverständnis	2
	1. KI – Allgemeine Definition?	2
	2. KI-System	3
	2.1. Verständnis der Bundesverwaltung	3
	2.2. Verständnis der EU-Kommission	3
	2.3. Vorliegendes Verständnis	4
III.	Charakteristika des KI-Einsatzes	5
	1. Standardisierung von Prozessabläufen	5
	2. Massenverwaltung	6
	3. Intelligente Durchsuchung von Informationen	7
	4. Prognosen	8
IV.	Rechtliche Rahmenbedingungen und Grenzen	8
	1. Vorbemerkung	9
	2. Legalitätsprinzip	9
	3. Verfahrensgarantien und Verfahrensgrundsätze	10
	3.1. Rechtliches Gehör	10
	a) Anspruch auf vorgängige Äusserung und Mitwirkung	11
	b) Begründungspflicht	13
	3.2. Untersuchungsgrundsatz	14
	4. Diskriminierungsverbot	15
	4.1. Verfassungsrechtliches Diskriminierungsverbot	15
	4.2. Diskriminierungsquellen beim KI-Einsatz	16
	a) Präexistierender Bias in den Daten	16
	b) Technischer Bias und fehlende Daten	17
	c) Emergender Bias und statistische Diskriminierung	18
	4.3. Massnahmen zur Verhinderung von Diskriminierung	18

5.	Transparenz	19
5.1.	Individuelle Kontrollmöglichkeiten	19
5.2.	Allgemeine Kontrollmöglichkeiten	20
V.	Künftige Entwicklungen	20
1.	Schweiz	21
2.	International	22
VI.	Fazit	22

I. Einleitung

Aufgrund ihrer grossen Leistungsfähigkeit und der zunehmenden Verfügbarkeit von Daten sind Technologien, die etwas pauschal als Künstliche Intelligenz (KI) oder auch algorithmische Systeme bezeichnet werden, für öffentliche Verwaltungen zunehmend von Interesse. Auf Bundesebene, aber auch in den Kantonen, laufen derzeit verschiedene Projekte zum Einsatz von KI, während gleichzeitig die Entwicklung von Leitlinien und Rechtsnormen insbesondere auch im internationalen Kontext vorangetrieben werden. Der vorliegende Beitrag geht zuerst einem allgemeinen Begriffsverständnis nach (II) und beschreibt grundlegende Charakteristika eines KI-Einsatzes (III). Sodann werden die rechtlichen Rahmenbedingungen und dabei gleichzeitig die Grenzen desselben dargelegt (IV). Der Beitrag schliesst mit einem Blick in die Zukunft (V) und einem Fazit (VI).

II. Begriffsverständnis

Im Folgenden werden verschiedene Annäherungsversuche an eine Begriffsdefinition im nationalen und internationalen Kontext aufgeführt und ein für den vorliegenden Beitrag geltendes Verständnis dargelegt.

1. KI – Allgemeine Definition?

Der Begriff «Künstliche Intelligenz» (KI – «Artificial Intelligence – AI») erlebt seit einigen Jahren eine regelrechte Renaissance. Seine Ursprünge gehen in die 1950er Jahre zurück,¹ dennoch hat sich bisher keine allgemeingültige Definition etabliert. Den verschiedenen Definitionsversuchen liegt immerhin ein gemeinsamer Tenor zugrunde. Aus technischer Perspektive handelt es sich um

¹ KAPLAN, S. 13 f.

einen etablierten Sammelbegriff, der gleich mehrere Technologien umfasst, die automatisierte Entscheidungen fällen, Empfehlungen abgeben, Schlussfolgerungen ziehen oder Vorhersagen treffen.² Dazu gehören wissensbasierte Systeme und statistische Methoden ebenso wie Ansätze des maschinellen Lernens (z. B. unter Einsatz neuronaler Netze).³ Ein Computer wird mit anderen Worten so gebaut oder programmiert, dass er Dinge tun kann, die normalerweise dem Menschen und seinen biologischen Fähigkeiten vorbehalten sind. Dazu zählen bspw. visuelle Wahrnehmungen wie die Bilderkennung, aber auch etwa die Spracherkennung und -übersetzung.⁴

2. KI-System

2.1. Verständnis der Bundesverwaltung

In der Bundesverwaltung kursieren aktuell verschiedene Begriffsdefinitionen.⁵ Das Kompetenznetzwerk für Künstliche Intelligenz (CNAI⁶) versucht, die Terminologie zu vereinheitlichen. Nach dem im Rahmen des CNAI entwickelten Ansatzes ist ein KI-System «ein maschinenbasiertes System, welches für vom Menschen klar definierte Problemstellungen, Vorhersagen, Empfehlungen oder Entscheidungen machen kann, welche reale oder virtuelle Umgebungen beeinflussen. KI-Systeme können mit unterschiedlichem Ausmass an Autonomie ausgestattet werden.»⁷ Es handelt sich dabei allerdings nicht um eine Legaldefinition.

2.2. Verständnis der EU-Kommission

Im Gegensatz zur Schweiz bahnt sich auf unionsrechtlicher Ebene mit dem Verordnungsvorschlag der EU-Kommission zur Festlegung harmonisierter

² ERTEL, S. 1 ff.; OETTINGER, S. 91.

³ THOUVENIN et al., S. 1.

⁴ Terminologie Kompetenznetzwerk CNAI, Version 1.0 (Stand am 15.12.2021), abrufbar unter <https://cnaai.swiss/> (zuletzt besucht am 10.10.2022), S. 6.

⁵ Siehe zu den verschiedenen Ansätzen Bericht IDAG KI 2019, S. 18 ff.

⁶ Das CNAI ist dem Bundesamt für Statistik BFS angegliedert und ist für den Austausch zu KI-Themen innerhalb der Verwaltung zuständig. Es setzt selbst jedoch keine Projekte um, siehe <https://cnaai.swiss/ueber-das-cnai/> (zuletzt besucht am 19.10.2022).

⁷ Terminologie Kompetenznetzwerk CNAI, Version 1.0 (Stand am 15.12.2021), abrufbar unter <https://cnaai.swiss/> (zuletzt besucht am 10.10.2022), S. 6.

Vorschriften für Künstliche Intelligenz (EU-KI-VOV⁸) eine Legaldefinition an.⁹ Gemäss Art. 3 Ziff. 1 EU-KI-VOV handelt es sich bei einem KI-System um «eine Software, die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren.» Techniken und Konzepte der Künstlichen Intelligenz sind sodann gemäss Anhang I EU-KI-VOV¹⁰ «Konzepte des maschinellen Lernens, mit beaufsichtigtem, unbeaufsichtigtem und bestärkendem Lernen unter Verwendung einer breiten Palette von Methoden, einschliesslich des tiefen Lernen (Deep Learning); Logik- und wissensgestützte Konzepte, einschliesslich Wissensrepräsentation, induktiver (logischer) Programmierung, Wissensgrundlagen, Inferenz- und Deduktionsmaschinen, (symbolischer) Schlussfolgerungs- und Expertensysteme; Statistische Ansätze, Bayessche Schätz-, Such- und Optimierungsmethoden.»

2.3. Vorliegendes Verständnis

Rechtliche Herausforderungen entstehen nicht nur beim Einsatz spezifischer Systeme und Konzepte wie sie etwa der Anhang zur EU-KI-VOV auflistet, sondern können ganz allgemein beim Einsatz algorithmenbasierter Systeme auftreten. Wird im Folgenden von KI oder KI-Einsatz gesprochen, geht es somit nicht um bestimmte heutige oder künftige Technologien, sondern (auch) um den sozialen Kontext, in dem sie eingesetzt werden. Der Begriff «algorithmische Systeme» wäre passender.¹¹

⁸ Vorschlag der EU-Kommission für eine Verordnung des europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz, 21.4.2021, COM(2021)206 final, abrufbar unter <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52021PC0206>, Dokument 1 (zuletzt besucht am 10.10.2022).

⁹ Siehe zum Verordnungsvorschlag der EU-Kommission allgemein etwa MÜLLER, S. A1 ff.; HORNUNG, S. 561 ff.

¹⁰ Anhänge des Vorschlags für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz, 21.4.2021, SEC(2021) 167 final, abrufbar unter <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52021PC0206>, Dokument 2 (zuletzt besucht am 10.10.2022).

¹¹ THOUVENIN et al., S. 1, f.

III. Charakteristika des KI-Einsatzes

In verschiedenen Ländern wird KI bereits heute in den öffentlichen Verwaltungen eingesetzt.¹² Die Einsatzbereiche sind zahlreich und divers. Zusammenfassend lassen sich vier Charakteristika der aktuellen und künftigen KI-Einsatzbereiche identifizieren, auf die im Folgenden näher eingegangen wird.

1. Standardisierung von Prozessabläufen

Ein erstes Merkmal ist die Standardisierung von Prozessabläufen. Sich häufig wiederholende, strukturierte Tätigkeiten ohne Ermessensspielraum der Behörden lassen sich gut automatisieren. Dies zeigt sich am bereits hohen Automatisierungsgrad in Steuerverfahren und den verschiedenen dort geplanten Anwendungen.¹³ Auch der Einsatz von Chatbots, d. h. technischen Dialogsystemen, mit denen sprach- oder textbasiert kommuniziert werden kann,¹⁴ bestätigt diese Tendenz, da diese vor allem dort genutzt werden, wo häufig wiederkehrende ähnliche Informationsanfragen auszumachen sind.¹⁵

Bspw. setzt die Stadt St. Gallen seit drei Jahren den Chatbot «Gallus» ein, der Informationen zur Mobilität und Entsorgung, zu Parkplätzen und Veranstaltungen oder zum Umzug beantwortet. Künftig soll er auch Prozesse selbständig abwickeln können.¹⁶ Die Stadt Zürich setzt einen Chatbot ein, der die Hauseigentümerinnen und Hauseigentümer in Sachen Heizungsersatz berät.¹⁷ Im Kanton Basel-Stadt setzt insbesondere der Bereich Bevölkerungsdienste und Migration Chatbots ein – seit dem Frühjahr 2021 bspw. das Zivilstandsamt zur Bestellung von Urkunden.¹⁸

¹² Vertieft BRAUN BINDER et al., Studie KI 2021, S. 28 ff.

¹³ Siehe dazu sogleich die Beispiele unter Abschnitt III.2.

¹⁴ DEMAJ/SÄGESSER, S. 1.

¹⁵ BRAUN BINDER, KI Verwaltung, S. 471.

¹⁶ Stadt St. Gallen, Home News vom 22.12.2020, abrufbar unter https://www.stadt.sg.ch/news/stsg_home/2020/12/chatbot.html (zuletzt besucht am 10.10.2022).

¹⁷ Medienmitteilung der Stadt Zürich vom 30.03.2021, Klimaschutz: Ein neuer Chatbot unterstützt Hauseigentümer/-innen beim Heizungsersatz, abrufbar unter <https://www.stadt-zuerich.ch/gud/de/index/departement/medien/medienmitteilungen/2021/maerz/210330a.html> (zuletzt besucht am 10.10.2022).

¹⁸ Medienmitteilung des Justiz- und Sicherheitsdepartements des Kantons Basel-Stadt vom 12.05.2021, abrufbar unter <https://www.jsd.bs.ch/jsd-medien/public-newsroom/chatbot.html> (zuletzt besucht am 23.10.2022).

2. Massenverwaltung

Ein zweites – mit dem Standardisierungspotenzial verbundenes – Merkmal ist die Massenverwaltung. Dazu gehören Verwaltungsbereiche, in denen eine Behörde in einem bestimmten Aufgabenbereich eine grosse Zahl an einzelnen Verfügungen erlässt.¹⁹ Typische Beispiele der Massenverwaltung sind das steuerrechtliche und das sozialversicherungsrechtliche Verfahren, aber auch das Asylwesen kann bspw. dazu gezählt werden.²⁰ Der KI-Einsatz in der Massenverwaltung soll insbesondere die Effizienz der Verwaltungstätigkeit steigern und gleichzeitig die verfahrensrechtlichen Vorgaben wahren.²¹

Bereits heute werden in den meisten Kantonen die digital eingereichten Steuererklärungen durch automatische Veranlagungsprogramme zumindest teilweise automatisiert bearbeitet. In Zukunft soll KI die Steuerverwaltungen jedoch noch weiterführend unterstützen,²² weshalb derzeit verschiedene Kantone einen solchen KI-Einsatz prüfen.²³

*Im Kanton Bern werden bspw. bereits fast 20 Prozent der steuerpflichtigen natürlichen Personen vollständig automatisiert veranlagt – das Ziel ist ein Automatisierungsgrad durch maschinelles Lernen von 40 Prozent.*²⁴

Im Sozialrecht stehen insbesondere die Automatisierung der Gewährung von Prämienverbilligungsansprüchen und die Betrugsbekämpfung im Zentrum, der Einsatz tatsächlicher KI-Systeme ist derzeit – soweit ersichtlich – jedoch sehr gering.²⁵

¹⁹ KIENER/RÜTSCHÉ/KUHN, N 879.

²⁰ KIENER/RÜTSCHÉ/KUHN, N 886 ff.

²¹ BRAUN BINDER, Dystopie, S. 270, 274. Zum Potenzial von KI in der Verwaltung allgemein, vgl. TA-SWISS KI 2020, S. 209 f.

²² Siehe dazu NUFER, S. 259 ff.

²³ Z. B. im Kanton St. Gallen, vgl. St. Galler Tagblatt vom 08.04.2018, Digitalisierung: Ostschweizer Steuerverwaltungen setzen auf Roboter, abrufbar unter <https://www.tagblatt.ch/ostschweiz/digitalisierung-ostschweizer-steuerverwaltungen-setzen-auf-roboter-ld.1006738> (zuletzt besucht am 10.10.2022) oder im Kanton Obwalden, vgl. Projektpräsentation vom 2. Juni 2020, abrufbar unter https://www.egovernment-wettbewerb.de/presentationen/2020/Data_Analytics_in_der_Steuerveranlagung_Obwalden_Automatische_Veranlagung.pdf (zuletzt besucht am 10.10.2022).

²⁴ FISCHER/DAEPP, S. 329.

²⁵ Die SUVA offenbart immerhin einige Bestrebungen, Automatisierungsprozesse zu fördern, siehe NZZ Content Creation im Auftrag der SUVA, Beitrag vom 07.12.2020, abrufbar unter <https://www.nzz.ch/sponsored-content/machine-learning-ohne-schlaue-koepfe-gehts-nicht-ld.1581399> (zuletzt besucht am 10.10.2022).

*Im Kanton Genf sollen KI-Systeme im Bereich der Sozialarbeit genutzt werden, um u. a. die ungerechtfertigte Auszahlung von Sozialleistungen zu verhindern.*²⁶

Neben den bereits genannten Bereichen sind zahlreiche weitere KI-Anwendungen in verschiedensten Gebieten der öffentlichen Verwaltung möglich.²⁷

*Eine Studie ist bspw. zum Ergebnis gekommen, dass eine KI-unterstützte, arbeitsmarktorientierte Kantonszuteilung von Asylsuchenden deren Beschäftigungsaussichten um rund 75 Prozent erhöhen könnte.*²⁸

3. Intelligente Durchsuchung von Informationen

Durch die Konfrontation mit immer grösser werdenden Datenmengen, deren Auswertung innert nützlicher Frist kaum mehr auf konventionellem Weg erfolgen kann, bietet sich für die intelligente Durchsuchung der Informationen ein KI-Einsatz an. Dies etwa bei der Polizeiarbeit im Bereich der Wirtschaftskriminalitätsbekämpfung, der allgemeinen organisierten Kriminalität oder der Cyberkriminalität.²⁹ Aber auch die intelligente Videoanalyse zählt dazu, da KI besonders gut bei der Bilderkennung funktioniert.³⁰

*In Luzern setzt etwa die Kantonspolizei eine Software ein, die Videoaufnahmen auswertet und dabei Menschen nach Alter und Geschlecht sowie Objekte und Farben unterscheiden kann.*³¹

²⁶ Secrétariat du Grand Conseil, Projet présenté par le Conseil d'Etat Date de dépôt : 29 août 2018, PL 12386 Projet de loi, abrufbar unter <https://ge.ch/grandconseil/data/texte/PL12386.pdf> (zuletzt besucht am 10.10.2022). Das Gesetz, das einen Investitionskredit von CHF 9'000'000 vorsieht, wurde am 25. Januar 2019 vom Grossen Rat des Kantons Genf verabschiedet (Loi 12386).

²⁷ Ausführlich BRAUN BINDER et al., Studie KI 2021, S. 27 f.

²⁸ Siehe die Medienmitteilung der ETH Zürich vom 18.01.2018, Algorithmus verbessert Erwerbschancen von Flüchtlingen, abrufbar unter <https://ethz.ch/de/news-und-veranstaltungen/eth-news/news/2018/01/algorithmus-verbessert-erwerbschancen-von-fluechtlingen.html> (zuletzt besucht am 10.10.2022). Vertieft BANSAK et al., S. 325.

²⁹ Siehe etwa GLAUS CHRISTIAN, Ein Computer soll für die Luzerner Polizei ermitteln, in: Luzerner Zeitung vom 26.10.2019, abrufbar unter <https://www.luzernerzeitung.ch/zentral-schweiz/luzern/kuenstliche-intelligenz-ein-computer-soll-fuer-die-luzerner-polizei-ermitteln-ld.1163004> (zuletzt besucht am 10.10.2022).

³⁰ Siehe etwa GOLDA/CORMIER/BEYERER, S. 1 ff. Eng in Verbindung mit der Bilderkennung steht die durch KI ermöglichte Gesichtserkennung, vgl. zu deren Risiken BRAUN BINDER/KUNZ/OBRECHT, S. 57 ff.

³¹ Siehe GLAUS (Fn. 29).

4. Prognosen

Schliesslich wird KI dort eingesetzt, wo es darum geht, Prognosen zu treffen. So dienen Anwendungen im Rahmen des Sanktionenvollzugs und im Bereich des *Predictive Policing* (vorausschauende Polizeiarbeit³²) der statistikbasierten Vorhersage von Verhaltensweisen oder Ereignissen.³³

*In der Deutschschweiz wird im Strafvollzug bspw. anhand des Programms ROS (Risikoorientierter Sanktionenvollzug) die Möglichkeit der Vollzugslockerung geprüft, wodurch das Rückfallrisiko während und nach dem Vollzug gesenkt werden soll.*³⁴

In diesen Prognosebereich fällt auch die KI-Entwicklung im Bereich der Stadtentwicklung in Richtung *Smart Cities* und der damit einhergehenden Verknüpfung von Daten.³⁵

*Im Kanton Basel-Stadt laufen verschiedene Projekte zur nachhaltigen Stadtentwicklung, bspw. die intelligente Fussgängersteuerung im Strassenverkehr oder das Schaffen eines intelligenten Energienetzes.*³⁶

IV. Rechtliche Rahmenbedingungen und Grenzen

Die folgenden Abschnitte gehen den rechtlichen Rahmenbedingungen und Grenzen eines KI-Einsatzes in der öffentlichen Verwaltung nach. Dabei werden insbesondere das Legalitätsprinzip (2), die Verfahrensgarantien und Verfahrensgrundsätze (3), das Diskriminierungsverbot (4) sowie transparenzrechtliche Anforderungen (5) thematisiert.³⁷

³² Vertieft PULLEN/SCHEFER, S. 103 ff.

³³ SIMMLER/BRUNNER/SCHEDLER, S. 14 ff., 31 ff.

³⁴ SIMMLER/BRUNNER/SCHEDLER, S. 14 ff., 31 f. Eine Übersicht zu den in den Kantonen eingesetzten Tools im Rahmen von Smart Criminal Justice SIMMLER/BRUNNER, S. 16.

³⁵ ZANELLA et al., S. 22 ff.; vertieft NIKITAS et al., S. 1 ff.

³⁶ Siehe dazu Smart City Basel, abrufbar unter <https://www.smartcity.bs.ch/> (zuletzt besucht am 10.10.2022).

³⁷ Die Autorinnen des vorliegenden Beitrages analysieren im Rahmen des Forschungsprojektes «Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz», gefördert durch die Mercator Stiftung, zahlreiche Rechtsfragen des KI-Einsatzes in der öffentlichen Verwaltung. Mehr Informationen dazu unter <https://ius.unibas.ch/de/e-piaf/nachvollziehbare-algorithmen/>.

1. Vorbemerkung

Es ist zu betonen, dass die rechtlichen Herausforderungen sich einerseits aus dem Einsatz von KI selbst, andererseits aber auch aus dem mithilfe von KI hergestellten Automatisierungsgrad ergeben können. Im Folgenden wird deshalb bei Bedarf zwischen teilautomatisierten Verfahren, in denen KI-Systeme unterstützend eingesetzt werden, und vollautomatisierten Verfahren, in denen KI-Systeme genutzt werden und keine menschliche Intervention mehr stattfindet, unterschieden.³⁸ Das hier zugrunde liegende Verständnis von vollautomatisierten Verfahren deckt sich mit dem Ausdruck «automatisierte Einzelentscheidung» nach dem revidierten Datenschutzgesetz.³⁹

2. Legalitätsprinzip

Zunächst ist aufgrund des Legalitätsprinzips gemäss Art. 5 Abs. 1 Bundesverfassung (BV⁴⁰) erforderlich, dass für staatliches Handeln eine sowohl hinsichtlich der Normstufe als auch hinsichtlich der Normdichte ausreichende Rechtsgrundlage existiert.⁴¹ Dies gilt es, hinsichtlich des Einsatzes von KI zu berücksichtigen.⁴² Für den KI-Einsatz zur Entscheidungsfindung im Rahmen von Verwaltungsverfahren ist eine bereichsspezifische und explizite gesetzliche Grundlage zu fordern.⁴³ Ausserdem ist zu bedenken, dass nach den datenschutzrechtlichen Vorgaben für die Bearbeitung von besonders schützenswerten Personendaten oder in Fällen des Profilings eine formell-gesetzliche Grundlage notwendig ist (Art. 34 Abs. 2 lit. a und b revDSG⁴⁴).

Inbesondere im Sozialversicherungsbereich ist davon auszugehen, dass KI-Anwendungen vorwiegend bei der Bearbeitung besonders schützenswerter Per-

³⁸ Zu den unterschiedlichen Automatisierungsgraden siehe WEDER, S. 241.

³⁹ Art. 21 revDSG. Das revidierte Datenschutzgesetz tritt am 1. September 2023 in Kraft. Zur Unterscheidung von Teil- und Vollautomation in diesem Kontext OBRECHT, S. 39.

⁴⁰ Bundesverfassung der Schweizerischen Eidgenossenschaft vom 18. April 1999, SR 101.

⁴¹ SGK BV-SCHINDLER, Art. 5 N 33 ff.

⁴² THURNHERR, S. 150 ff.

⁴³ THURNHERR, S. 151. In Anlehnung an GLASER, Einflüsse, S. 188, lässt sich auch das Argument anführen, dass das Bundesgericht allein schon für die Digitalisierung einzelner Verfahrensabschnitte wie etwa die gültige Einreichung elektronisch signierter Beschwerdeschriften (BGE 143 I 187) oder die Durchführung einer Hauptverhandlung ohne Einverständnis der Parteien im Rahmen einer Videokonferenz (BGE 146 III 194) eine gesetzliche Grundlage verlangt.

⁴⁴ Das revidierte Datenschutzgesetz tritt am 1. September 2023 in Kraft. Im Folgenden wird auf die Literatur zum aktuell geltenden DSG verwiesen.

sonendaten eingesetzt werden und dementsprechend eine formell-gesetzliche Grundlage erforderlich ist. Der Einsatz von Chatbots zu Informationszwecken dürfte hingegen von bestehenden Rechtsgrundlagen abgedeckt sein – sofern diese keine anderweitigen Aufgaben übernehmen und dabei besonders schützenswerte Personendaten bearbeitet werden.⁴⁵

3. Verfahrensgarantien und Verfahrensgrundsätze

Verfahrensgarantien sind die in Art. 29–32 BV verankerten Mindeststandards, die von Behörden in allen Verfahren zu beachten sind.⁴⁶ Im Zusammenhang mit dem Verwaltungshandeln sind dabei insbesondere die sogenannten «allgemeinen Verfahrensgarantien» nach Art. 29 BV von Bedeutung,⁴⁷ in Zusammenhang mit dem Einsatz von KI etwa der Anspruch auf rechtliches Gehör (3.1).⁴⁸ Verfahrensgrundsätze stellen Leitlinien des Prozesses dar und sind positiv-rechtlich im Verfahrensrecht verankert.⁴⁹ Der Einsatz von KI stellt speziell den im Verwaltungsverfahren vorherrschenden Untersuchungsgrundsatz gemäss Art. 12 Verwaltungsverfahrensgesetz (VwVG) vor Herausforderungen (3.2).⁵⁰

3.1. Rechtliches Gehör

Der Anspruch auf rechtliches Gehör gemäss Art. 29 Abs. 2 BV dient einerseits der Sachverhaltsabklärung und stellt andererseits ein persönlichkeitsbezogenes Mitwirkungsrecht dar.⁵¹ Aus dem rechtlichen Gehör leiten sich verschiedene Teilgehalte ab, wobei im Folgenden der Anspruch auf vorgängige Äusserung und Mitwirkung im Verfahren (a) sowie das Recht auf Begründung (b) betrachtet werden.⁵²

⁴⁵ BRAUN BINDER et al., Studie KI 2021, S. 58 f.

⁴⁶ KIENER/KÄLIN/WYTTENBACH, § 40 N 1; RHINOW et al., N 255.

⁴⁷ BSK BV-WALDMANN, Art. 29 N 6.

⁴⁸ BRAUN BINDER, Dystopie, S. 273 f.; THURNHERR, S. 147 ff.; REITER, S. 991 f.; RECHSTEINER, N 19 ff.

⁴⁹ TSCHANNEN/ZIMMERLI/KERN, § 30 N 771; RHINOW et al., N 974 f.

⁵⁰ BRAUN BINDER, Untersuchungsgrundsatz, S. 35 ff.

⁵¹ BSK BV-WALDMANN, Art. 29 N 44; RHINOW et al., N 309.

⁵² Auf das Akteneinsichtsrecht, das ebenfalls Bestandteil des rechtlichen Gehörs bildet, wird im Folgenden nicht eingegangen. Siehe zum Akteneinsichtsrecht im automatisierten Verwaltungsverfahren WEDER, S. 251 ff. und für einen allgemeinen Überblick RHINOW et al., N 309.

a) *Anspruch auf vorgängige Äusserung und Mitwirkung*

Der Anspruch auf vorgängige Äusserung und Mitwirkung in Verwaltungsverfahren wird auf Bundesebene in Art. 30 Abs. 1 VwVG konkretisiert. Die Parteien müssen demnach vor dem Erlass einer Verfügung angehört werden bzw. Gelegenheit zur Stellungnahme erhalten.⁵³ Dies stellt mithin das wichtigste Mittel dar, um der Partei einen Einfluss auf die Ermittlung des rechtserheblichen Sachverhalts zu sichern.⁵⁴ Nur in den in Art. 30 Abs. 2 VwVG vorgesehenen Ausnahmesituationen müssen Parteien im Vorfeld nicht angehört werden. Damit eine Stellungnahme überhaupt möglich ist, müssen den Betroffenen zumindest die wesentlichen Elemente des voraussichtlichen Verfügungsinhaltes bekannt gegeben werden.⁵⁵ Auch wenn die Behörde den Äusserungen der Betroffenen nicht folgen muss, sie muss diese immerhin tatsächlich zur Kenntnis nehmen, überprüfen und sich mit den Einwänden bei der Entscheidungsfindung und -begründung sachgerecht auseinandersetzen.⁵⁶

In weitgehend standardisierten und vollautomatisierten Verfahren kann der Anspruch auf vorgängige Äusserung beeinträchtigt sein. Wird ein Verfahren durch eine natürliche Person eingeleitet (etwa durch einen Antrag) und läuft dieses sodann ohne jegliche menschliche Intervention ab, kann sich diese nur bei der eigentlichen Antragsstellung äussern, nicht aber im weiteren Verlauf des vollautomatisierten Verfahrens.⁵⁷ Wie stark der Anspruch auf vorgängige Äusserung somit eingeschränkt wird, hängt schliesslich von den Äusserungsmöglichkeiten bei der Antragsstellung ab und inwiefern die getätigten Äusserung durch das System berücksichtigt werden kann.⁵⁸

In der Literatur wird deshalb dafür plädiert, dass vollautomatisierte Entscheidungen nur zulässig sein sollen, wenn den Begehren der betroffenen Person vollumfänglich entsprochen wurde. Dagegen soll bei Entscheidungen, die zugunsten der betroffenen Person ausfallen, das rechtliche Gehör durch die sachverständige Person gewährt werden.⁵⁹ Damit würde es sich in diesen Fällen allerdings nicht mehr um vollautomatisierte Entscheidungen handeln.

⁵³ VwVG Komm-SUTTER, Art. 30 N 1.

⁵⁴ Statt vieler RHINOW et al., N 323 ff.

⁵⁵ VwVG Komm-SUTTER, Art. 30 N 3.

⁵⁶ RHINOW et al., N 328 ff.; KIENER/KÄLIN/WYTTENBACH, § 41 N 40 ff.

⁵⁷ BRAUN BINDER, Perspektive Datenschutzrecht, S. 28 f.; BRAUN BINDER, Dystopie, S. 273 f.; RECHSTEINER, N 18 ff.; WEBER, S. 24.

⁵⁸ BRAUN BINDER, KI Verwaltung, S. 476; bereits BRAUN BINDER, Steuerbescheide, S. 526 ff.

⁵⁹ RECHSTEINER, N 23; BRAUN BINDER, KI Verwaltung, S. 476.

Art. 21 revDSG sieht verschiedene Mechanismen vor, die das rechtliche Gehör – zumindest in gewissen Ausprägungen – adressieren sollen. Bspw. besteht für automatisierte Einzelentscheidungen, die für die betroffene Person mit einer Rechtsfolge verbunden sind oder sie erheblich benachteiligen, eine Informationspflicht (Abs. 1).⁶⁰ Ausserdem erhält die betroffene Person die Möglichkeit, ihren Standpunkt darzulegen, wenn sie dies beantragt, und kann verlangen, dass die Entscheidung von einer natürlichen Person überprüft wird (Abs. 2).⁶¹ Bundesorgane haben vollautomatisierte Verfügungen zu kennzeichnen (Abs. 4).⁶² Die Möglichkeit der betroffenen Person, ihren Standpunkt darzulegen, entfällt, wenn nach Art. 30 Abs. 2 VwVG ohnehin kein vorgängiges Recht auf Äusserung besteht (Art. 21 Abs. 4 Satz 2 revDSG).⁶³

Aus datenschutzrechtlicher Sicht sind diese Überlegungen durchaus sinnvoll und nachvollziehbar. Das Datenschutzrecht zielt auf den Persönlichkeits- und Grundrechtsschutz derjenigen Person, deren Personendaten bearbeitet werden (Art. 1 revDSG). Im Verwaltungsverfahren hingegen gehen diese Lösungsansätze nicht weit genug; der Anspruch auf vorgängige Äusserung steht nicht nur der von der Verfügung adressierten Person – und somit der Person, deren Personendaten bearbeitet werden – zu, sondern gemäss dem Wortlaut von Art. 30 Abs. 1 VwVG einem erweiterten Kreis an Parteien, deren Rechte durch die Verfügung betroffen sein können.

Bspw. können die Nachbarn eines Bauprojektes von der Baubewilligung betroffen sein, womit ihnen das vorgängige Äusserungsrecht zusteht.

Im Verwaltungsverfahren sollten deshalb andere Mechanismen greifen. Eine Anknüpfungsmöglichkeit wären die im entsprechenden Verfahrensrecht vorgesehenen Einsprachemöglichkeiten.⁶⁴ Die Einsprache ist ein Rechtsmittel ohne Devolutiveffekt, d. h., die betroffene Person kann sich vor der verfügenden Behörde und nicht erst vor der Rechtsmittelinstanz äussern.⁶⁵ Dadurch würde das Recht auf vorgängige Äusserung gewahrt.

⁶⁰ Siehe zur Informationspflicht BIERI/POWELL, S. 1538 f.

⁶¹ Siehe dazu BRAUN BINDER, Perspektive Datenschutzrecht, S. 32 f.

⁶² Zur automatisierten Einzelentscheidung als Verfügung i.S.v. Art. 5 VwVG siehe OBRECHT, S. 42 ff.

⁶³ Kritisch BRAUN BINDER, Dystopie, S. 259.

⁶⁴ THURNHERR, S. 148 f.

⁶⁵ Statt vieler zur Einsprache KÖLZ/HÄNER/BERTSCHI, N 786 ff.; RHINOW et al., N 638 ff.

b) Begründungspflicht

Der Anspruch auf Begründung einer Verfügung wird in Art. 35 Abs. 1 VwVG konkretisiert. Dadurch wird Transparenz hinsichtlich der behördlichen Entscheidungsfindung im Einzelfall hergestellt.⁶⁶ Die Ausführlichkeit der Begründung ist an den konkreten Umständen auszurichten – ausschlaggebend ist die Möglichkeit einer sachgerechten Anfechtung eines Entscheids.⁶⁷ Beim Einsatz von KI ist es möglich, dass die Behörde ihrer Begründungspflicht nicht vollständig nachkommen kann. Denn Entscheidungen, die von einem KI-basierten System gefällt werden, können aufgrund der Komplexität nicht ohne Weiteres für die Sachbearbeiterin bzw. den Sachbearbeiter nachvollziehbar sein.⁶⁸

Zu denken ist etwa an Systeme, die statistische Auswertungen basierend auf Kriterien vornehmen, die für die Sachbearbeiterin bzw. den Sachbearbeiter unklar sind.

Die verfassungs- und verwaltungsverfahrensrechtlichen Vorschriften verlangen Angemessenheit der Begründung im Einzelfall. Diese muss je nach Komplexität der Rechtslage ausführlicher oder weniger ausführlich sein.⁶⁹ Die gelieferte Begründung muss in jedem Fall für die betroffene Person verständlich sein.⁷⁰ Ungenügend wäre jedoch die allgemeine Erklärung der Funktionsweise des KI-Systems, da diese gerade keine einzelfallweise Begründung des ergangenen Entscheids darstellt.⁷¹

In der Informatik wird derzeit Grundlagenforschung zur Erklärbarkeit von KI-Systemen («Explainable AI») betrieben.⁷² Diese kann aber keine Lösung zur Sicherstellung der Begründungspflicht im Einzelfall sein, solange sie lediglich darauf abzielt, in allgemeiner Weise zu erklären, wie ein KI-System zu Ergebnissen kommt.

⁶⁶ WIEDERKEHR/MEYER/BÖHME, S. 133 f., 139; WIEDERKEHR, S. 536

⁶⁷ KÖLZ/HÄNER/BERTSCHI, N 630 ff.; VwVG Komm-KNEUBÜHLER/PEDRETTI, Art. 35 N 10 ff. Die Anforderungen an die Begründungsdichte sind mithin Ergebnis einer Interessenabwägung im Einzelfall, siehe dazu WIEDERKEHR/MEYER, S. 1099 f.

⁶⁸ WIEDERKEHR/MEYER, S. 1105 f.; WEDER, S. 254 f.

⁶⁹ WIEDERKEHR/MEYER, S. 1099 f.

⁷⁰ THURNHERR, S. 149; ausführlich ROTH-ISIGKEIT, S. 1025.

⁷¹ Vertieft BRAUN BINDER et al., Studie KI 2021, S. 37 f.

⁷² Zum Konzept im Zusammenhang mit Recht etwa WILLIAMS, S. 318 ff.

3.2. Untersuchungsgrundsatz

Der Untersuchungsgrundsatz – auch Untersuchungsmaxime – ist auf Bundesebene in Art. 12 VwVG verankert. Demnach hat die Behörde den Sachverhalt grundsätzlich von Amtes wegen festzustellen, sie beschafft die Entscheidungsgrundlagen selbst.⁷³ Die Mitwirkungspflicht der Parteien gemäss Art. 13 VwVG relativiert den Untersuchungsgrundsatz sodann. Die Parteien sind verpflichtet, an der Abklärung des Sachverhalts mitzuwirken.⁷⁴

Grundvoraussetzung für den KI-Einsatz ist das Vorhandensein von Daten.⁷⁵ Im Rahmen des Verwaltungsverfahrens handelt es sich dabei um Informationen einen Sachverhalt betreffend, den es in diesem Verfahren zu beurteilen gilt.⁷⁶ Da KI in der öffentlichen Verwaltung insbesondere zur Automatisierung von Prozessen eingesetzt wird,⁷⁷ stellt sich die Frage, wie die Sachverhaltsermittlung im Rahmen dieser Automatisierung umgesetzt werden kann. Dabei ist zu bedenken, dass das automatische Sammeln, Auswerten und Verifizieren von Sachverhaltsdaten gerade zu den Wesensmerkmalen automatisierter Verwaltungsverfahren zählt.⁷⁸ Im vollautomatisierten Verwaltungsverfahren kommt hinzu, dass die Mitarbeitenden der Behörde schon per definitionem keinerlei zusätzliche Abklärungen betreffend den Sachverhalt treffen.

Mit der Automatisierung der Sachverhaltsermittlung geht die Gefahr einher, dass nur standardisiert erhobene Daten genutzt und einzelfallspezifische Informationen nicht berücksichtigt werden.⁷⁹ Dies ist vor allem der Fall, wenn sich das Verfahren auf Daten stützt, die der Behörde bereits vorliegen, insbesondere Daten aus bestehenden Datenbanken und Registern. Der Untersuchungsgrundsatz verpflichtet die Behörde aber dafür Sorge zu tragen, dass die genutzten Daten alle rechtserheblichen Tatsachen abbilden. Von zentraler Bedeutung ist demnach, dass die vom KI-System genutzten Daten geeignet, vollständig, korrekt und verfügbar sind.⁸⁰

⁷³ VwVG Komm-AUER/BINDER, Art. 12 N 7; RHINOW et al., N 1002.

⁷⁴ Komm VwVG-AUER/BINDER, Art. 13 N 1; vertieft MEYER, S. 3 ff.

⁷⁵ Statt vieler ERTEL, S. 191 ff.; OETTINGER, S. 101 ff.

⁷⁶ BRAUN BINDER, Untersuchungsgrundsatz, S. 35 ff.

⁷⁷ Z. B. BRAUN BINDER, KI Verwaltung, S. 470 ff.; für internationale Beispiele LISCHKA/KLINGEL, S. 8 ff.

⁷⁸ GLASER, Einflüsse, S. 184.

⁷⁹ BRAUN BINDER, KI Verwaltung, S. 475 f.; BRAUN BINDER, Dystopie, S. 274.

⁸⁰ Dieses Erfordernis erstreckt sich je nach System auf die genutzten Trainingsdaten, wobei sich die rechtliche Begründung dafür nicht aus dem Untersuchungsgrundsatz, sondern aus dem Anspruch auf rechtmässige, insbesondere diskriminierungsfreie Verfahren ergibt.

Anders verhält es sich beim Verfahren auf Erlass einer mitwirkungsbedürftigen Verfügung, bei dem die Sachverhaltsdaten durch Eingabe der antragstellenden Person ermittelt werden.⁸¹ Die antragstellende Person wird ihre Angaben dabei jeweils in standardisierter Form erfassen, damit das KI-System die Daten verarbeiten kann.⁸² Dies kann jedoch dazu führen, dass die für das Verfahren notwendigen einzelfallspezifischen Informationen nicht berücksichtigt werden. Zugleich führt diese Vorgehensweise zu einer Verlagerung der Sachverhaltsabklärung von der Behörde auf die (lediglich zur Mitwirkung verpflichtete) antragstellende Person.⁸³

Bspw. besteht im Falle eines vollautomatisierten Veranlagungsverfahrens im Steuerbereich zwar eine ausgeprägte Mitwirkungspflicht und das Selbstdeklarationsprinzip findet Anwendung.⁸⁴ Trotzdem muss die Behörde sicherstellen, dass dem Veranlagungsverfahren korrekte Sachverhaltsdaten zugrunde gelegt werden. Sicherstellen kann sie dies bspw. mittels Kontrollmechanismen wie einem Risikomanagementsystem.⁸⁵

4. Diskriminierungsverbot

4.1. Verfassungsrechtliches Diskriminierungsverbot

Das Diskriminierungsverbot gemäss Art. 8 Abs. 2 BV stellt einen qualifizierten Schutz vor Ungleichbehandlung dar.⁸⁶ Nach der bundesgerichtlichen Rechtsprechung liegt eine Diskriminierung 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.»⁸⁷ Das Verbot ist mithin die grundrechtliche Antwort auf historische Erfahrungen der Ausgrenzung, Herabwürdigung und Stigmatisierung von Menschen allein wegen eines sensi-

⁸¹ Siehe dazu beispielhaft die Entwicklungen im Zollverfahren, dazu BRAUN BINDER, Dystopie, S. 261 f.

⁸² GUCKELBERGER, S. 396 f.

⁸³ BRAUN BINDER, Dystopie, S. 275.

⁸⁴ MEYER, S. 160 f.

⁸⁵ Wie dies etwa in Deutschland der Fall ist, vgl. BRAUN BINDER, Risikomanagement, S. 162 ff.; DRÜEN, S. 193 ff.

⁸⁶ BSK BV-WALDMANN, Art. 8 N 45 ff.; MÜLLER/SCHEFER, S. 684 ff.

⁸⁷ 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.

blen Persönlichkeitsmerkmals, das so wesentlich ist, dass es den Betroffenen nicht möglich oder nicht zumutbar ist, sich des Merkmals zu entledigen.⁸⁸

Einige dieser sensiblen Merkmale werden in Art. 8 Abs. 2 BV ausdrücklich genannt, wobei die Aufzählung jedoch nicht abschliessend ist: Herkunft, Rasse, Geschlecht, Alter, Sprache, soziale Stellung, Lebensform, religiöse, weltanschauliche oder politische Überzeugung sowie körperliche, geistige oder psychische Behinderung.

Der Diskriminierungsbegriff kann als «Schlüsselvokabel»⁸⁹ der heutigen Zeit bezeichnet werden. Es wird zwischen direkter (unmittelbarer) und indirekter (mittelbarer) Diskriminierung unterschieden.⁹⁰ Eine direkte Diskriminierung liegt vor, wenn ein Rechtsakt ohne Rechtfertigung direkt an ein verpöntes Merkmal anknüpft.⁹¹ Indirekt ergibt sich die Diskriminierung aus der praktischen Auswirkung einer gesetzlichen Regel, die formal neutral ist und keine offensichtliche Benachteiligung einer bestimmten Gruppe enthält.⁹²

4.2. Diskriminierungsquellen beim KI-Einsatz

In automatisierten Verwaltungsverfahren kann eine Diskriminierung aus verschiedenen Gründen resultieren, es gibt also verschiedene Diskriminierungsquellen.

a) Präexistierender Bias in den Daten

Ein besonders hervorzuhebendes Beispiel sind KI-Systeme, deren Entwicklung auf Trainingsdaten angewiesen ist. Diese Trainingsdaten prägen die spätere Funktionsweise des Systems.⁹³ Falsche Trainingsdaten können zu falschen Ergebnissen führen.⁹⁴ Aber auch «korrekte» Daten können Probleme evozieren, denn häufig perpetuieren sich in der Gesellschaft bestehende Vorurteile in den Daten.⁹⁵ Werden diese Daten sodann als Trainingsgrundlage verwendet, übertragen sich gesellschaftliche Vorurteile gewissermassen auf die Technologie.⁹⁶

⁸⁸ MÜLLER/SCHEFER, S. 684; KIENER/KÄLIN/WYTTENBACH, § 36 N 4 ff.

⁸⁹ DIGGELMANN et al., S. 211.

⁹⁰ Statt vieler BSK BV-WALDMANN, Art. 8 N 62 f.

⁹¹ BSK BV-WALDMANN, Art. 8 N 62.

⁹² MÜLLER/SCHEFER, S. 695 ff.; BSK BV-WALDMANN, Art. 8 N 63.

⁹³ ERTEL, S. 194, 230.

⁹⁴ ERTEL, S. 191 ff., aus rechtlicher Sicht ausführlich BRAUN BINDER, KI Verwaltung, S. 469.

⁹⁵ WOLFANGEL, S. 171 ff.

⁹⁶ Siehe zum Begriff «preexisting bias» bereits FRIEDMANN/NISSENBAUM, S. 333 f.

Werden bspw. in ein KI-System, das Bewerbungen für einen Stellenbescrieb aussortiert, die Bewerbungsunterlagen der letzten zwanzig Jahre eingespeist, ist es möglich, dass Frauen systematisch schlechter bewertet und somit aussortiert werden, da in der Vergangenheit Männer häufiger eingestellt wurden als Frauen.

Das österreichische Bundesverwaltungsgericht hat im Dezember 2020 eine Entscheidung bezüglich des österreichischen «Arbeitsmarkt-Chancen-Assisystem» (AMAS) gefällt.⁹⁷ Dieses unterstützt den österreichischen Arbeitsmarktservice (AMS) durch statistische Analysen historischer Daten, welche die Berechnungsgrundlage für die Beurteilung der zukünftigen Chancen von Arbeitssuchenden am Arbeitsmarkt darstellen. Das System führte zu diskriminierenden Handlungsempfehlungen, da insbesondere Frauen und Menschen mit Behinderung systematisch schlechtere Chancen auf dem Arbeitsmarkt zugesprochen wurden.⁹⁸ In der rechtlichen Auseinandersetzung ging es allerdings nicht in erster Linie um das Diskriminierungsverbot, sondern um die Frage der Rechtsgrundlage für den Einsatz von AMAS.

b) Technischer Bias und fehlende Daten

Auch technische Fehleinstellungen können zu Diskriminierungen führen.⁹⁹ Diese Diskriminierungsform hängt eng mit dem Problem fehlender Daten zusammen, insbesondere, wenn eine Bevölkerungsgruppe in den Trainingsdaten untervertreten ist.¹⁰⁰

Setzt der Staat bspw. Gesichtserkennungstechnologien ein, die auf der Basis unvollständiger Datensets trainiert wurden, werden die untervertretenen Gruppen – häufig Frauen, jüngere Menschen oder nicht kaukasische Personen – schlechter erkannt. Die Auswirkungen dieser falschen Identifizierungen zeigen sich umso stärker, wenn an ein Merkmal nach Art. 8 Abs. 2 BV angeknüpft wird.¹⁰¹

⁹⁷ Bundesverwaltungsgericht der Republik Österreich, Entscheid vom 18.12.2020, W256 2235360-1/5E.

⁹⁸ Siehe dazu etwa FRÖHLICH WIEBKE/SPIECKER genannt DÖHMANN INDRA, Können Algorithmen diskriminieren?, Verfassungsblog vom 26. Dezember 2018, abrufbar unter <https://verfassungsblog.de/koennen-algorithmen-diskriminieren/> (zuletzt besucht am 23.10.2022).

⁹⁹ Siehe zum Begriff «technical bias» bereits FRIEDMANN/NISSENBAUM, S. 335 f.

¹⁰⁰ ZWEIG, S. 214 f.

¹⁰¹ BRAUN BINDER/KUNZ/OBRECHT, S. 60.

c) *Emergenter Bias und statistische Diskriminierung*

Von einem emergenten Bias wird dort gesprochen, wo sich die Diskriminierung erst aus der Nutzung des Systems selbst ergibt, z. B. indem Ausgaben des algorithmischen Systems falsch interpretiert werden.¹⁰² Die Diskriminierung kann schliesslich auch aus Korrelationen resultieren, die an sog. Proxies anknüpfen. Bei Proxies handelt es sich um harmlos erscheinende Merkmale, die mit sensiblen Merkmalen jedoch stark korrelieren können.¹⁰³ Bei der Proxy-Diskriminierung («Stellvertreter Diskriminierung») werden somit scheinbar neutrale Merkmale als Grundlage für einen Datensatz beigezogen, wobei jedoch eine bestimmte geschützte Gruppe unverhältnismässig stark von einer Diskriminierung betroffen ist.¹⁰⁴ Aus rechtlicher Sicht ist die Proxy-Diskriminierung Teil der indirekten Diskriminierung.¹⁰⁵

*Bspw. kann ein Algorithmus am (vermeintlich neutralen) Merkmal «Wohnort» anknüpfen. Korreliert dieser überproportional oft mit dem sensiblen Merkmal «Rasse», weil in einem Stadtteil überwiegend Personen einer bestimmten Ethnie wohnen, kann eine Proxy-Diskriminierung resultieren.*¹⁰⁶

4.3. Massnahmen zur Verhinderung von Diskriminierung

Als Massnahme zur Verhinderung von Diskriminierung ist beim KI-Einsatz deshalb darauf zu achten, dass korrekte Daten – unter Berücksichtigung des Datenschutzes¹⁰⁷ – und fehlerfreie Systeme genutzt werden.¹⁰⁸ Damit lässt sich aber nicht jede Diskriminierung verhindern.¹⁰⁹ Unter Umständen ist in sensiblen Bereichen auf vollautomatisierte Entscheidungen zu verzichten und durch geeignete Massnahmen sicherzustellen, dass Sachbearbeiterinnen und Sachbearbeiter über die notwendigen Kenntnisse und Kompetenzen verfügen, um im Einzelfall eine Entscheidung zu treffen, die vom allenfalls diskriminierenden Vorschlag abweicht.

¹⁰² Siehe zum Begriff «emergent bias» bereits FRIEDMANN/NISSENBAUM, S. 336.

¹⁰³ PRINCE/SCHWARZ, S. 1268 ff.

¹⁰⁴ Vertieft BUCHHOLTZ/SCHEFFEL-KAIN, S. 612 ff.

¹⁰⁵ BUCHHOLTZ/SCHEFFEL-KAIN, S. 615.

¹⁰⁶ WEBER/HENSELER, S. 31 f.; BUCHHOLTZ/SCHEFFEL-KAIN, S. 613.

¹⁰⁷ Zu KI und Datenschutz siehe BRAUN BINDER, Perspektive Datenschutzrecht, S. 28 ff.

¹⁰⁸ STEVENS, S. 74 ff.; VOSS, S. 167 ff.; ZWEIG/KRAFFT, S. 211 ff.

¹⁰⁹ In der Literatur werden weitere Massnahmen zur Verhinderung von Diskriminierung diskutiert, bspw. durch den Einsatz von Kontrollalgorithmen oder Überprüfung durch staatliche Institutionen oder Drittorganisationen. Siehe dazu m.w.H. BRAUN BINDER et al., Studie KI 2021, S. 42.

5. Transparenz

Transparenz ist eine im Kontext von KI häufig geäußerte Forderung, um die Rechtmässigkeit des KI-Einsatzes sicherzustellen.¹¹⁰ Die Herstellung von Transparenz beim staatlichen KI-Einsatz ist nicht nur unter dem Blickwinkel individueller Kontrollmöglichkeiten von Einzelentscheidungen (5.1), sondern auch in Bezug auf eine allgemeine Kontrolle (5.2) – etwa durch die Zivilgesellschaft – zu diskutieren.¹¹¹

5.1. Individuelle Kontrollmöglichkeiten

Da KI-Systeme regelmässig auf grosse Datenmengen angewiesen sind, um sinnvoll eingesetzt werden zu können, und dabei häufig Personendaten bearbeitet werden, ergeben sich aus dem Recht auf informationelle Selbstbestimmung nach Art. 13 Abs. 2 BV bzw. den datenschutzrechtlichen Grundsätzen verschiedene Anforderungen. Gemäss Art. 6 Abs. 3 revDSG dürfen Personendaten nur zu einem bestimmten und für die betroffene Person erkennbaren Zweck bearbeitet werden. Die gesetzliche Regelung der Erkennbarkeit des Zwecks der Datenbearbeitung setzt dabei zwingend voraus, dass auch die Datenbearbeitung als solche erkennbar ist.¹¹² Weiter zeigen zentrale Bestimmungen des revidierten Datenschutzgesetzes, dass das Gesetz auf dem Grundsatz der Transparenz beruht, so namentlich die Regelung der Informationspflichten (Art. 19 ff. revDSG) und des Auskunftsrechts (Art. 25 ff. revDSG).¹¹³ Zu nennen ist insbesondere die Informations- und Kennzeichnungspflicht beim Erlass einer automatisierten Einzelentscheidung durch ein Bundesorgan gemäss Art. 21 Abs. 4 revDSG.¹¹⁴ Diese Kontrollmöglichkeiten bieten mit Blick auf einen KI-Einsatz zumindest eine partielle Überprüfungsmöglichkeit der (Personen-)Datenbearbeitung durch KI und stellen somit zumindest partiell Transparenz her.

¹¹⁰ Z. B. Schweiz: Bericht IDAG KI 2019, S. 37 f. Deutschland: Datenethikkommission Gutachten 2019, S. 169 f. International: OECD Prinzipien 2019, Forderung 1.3.

¹¹¹ In Bezug auf den KI-Einsatz stellen sich weitere transparenzrechtliche Herausforderungen, wie insb. Zusammenhang mit dem rechtlichen Gehör (vgl. die Ausführungen im vorliegenden Beitrag unter 3.3 und vertieft WIEDERKEHR/MEYER, S. 1105 f.; WEDER, S. 248 ff.) oder dem Recht auf Zugang zu Information (vgl. dazu BRAUN BINDER et al., Studie KI 2021, S. 49 f.), auf die vorliegend nicht weiter eingegangen wird.

¹¹² SHK DSG-BAERISWYL, Art. 4 N 47 f.

¹¹³ Zum Verhältnis von Transparenz und Datenschutz THOUVENIN/BRAUN BINDER, S. 13 ff.

¹¹⁴ Dazu BRAUN BINDER, Perspektive Datenschutzrecht, S. 28 f.

5.2. Allgemeine Kontrollmöglichkeiten

Um allgemeine, zivilgesellschaftliche Kontrolle – und somit Transparenz in einem weit verstandenen Sinn – rechtlich zu konkretisieren, sind verschiedene Ansatzpunkte vorstellbar. In der Literatur¹¹⁵, der Politik¹¹⁶ und der Zivilgesellschaft¹¹⁷ findet sich verschiedentlich der Vorschlag bzw. die Forderung, ein öffentlich zugängliches Register zu schaffen. Aus diesem soll ersichtlich werden, in welchen Bereichen die öffentliche Verwaltung KI-Systeme einsetzt und das u. a. über die Art und Herkunft der bearbeiteten Sach- und Personendaten, die Rechtsgrundlage, den Zweck und die Mittel der Bearbeitung, das verantwortliche Organ, die KI-Anwendung und deren Logik sowie diejenigen Akteure, die an der Entwicklung des Systems mitgewirkt haben, Auskunft gibt.¹¹⁸ Eine weitere Methode, Transparenz herzustellen, besteht darin, einen Transparenzbericht zu verfassen, der zeigt, dass die wichtigsten Fragen, z. B. zur Verhinderung von Diskriminierung, sowohl erkannt als auch reflektiert wurden und dass eine angemessene Rechenschaftspflicht für den Prozess sichergestellt wurde.¹¹⁹ Zu diskutieren wäre, wie und wo ein Register bzw. die Erstellung eines Transparenzberichts rechtlich zu verankern und welche Punkte dabei zu berücksichtigen wären.

V. Künftige Entwicklungen

Auf Bundesebene (1), aber auch im internationalen Kontext (2), laufen verschiedene Projekte und Rechtsetzungsvorhaben, um den KI-Einsatz durch Staat und Private zu begleiten bzw. zu regulieren. Diese Entwicklungen sind bei der

¹¹⁵ BRAUN BINDER et al., Studie KI 2021, S. 50; TA-SWISS KI 2020, S. 294; THOUVENIN et al., S. 3.

¹¹⁶ Bund: Motion 21.4508 Schlatter Marionna vom 16. Dezember 2021 betreffend ein öffentliches Verzeichnis der in der Verwaltung eingesetzten Algorithmen. Kanton Zürich: Postulat 9/2022 Yuste Nicola/Moser Arianne/Mäder Gabriel/Wilma Willi/Erika Zahler vom 17. Januar 2022 betreffend Transparenz über den Einsatz von künstlicher Intelligenz in der Verwaltung, das vom Kantonsrat genehmigt und am 25.04.2022 an den Regierungsrat überwiesen wurde. Im Oktober 2022 hat der Regierungsrat des Kantons Zürich sodann vorgeschlagen, ein Verzeichnis für automatisierte Entscheidungssysteme einzurichten, siehe dazu <https://algorithm-watch.ch/de/adm-register-kt-zh/> (zuletzt besucht am 19.10.2022).

¹¹⁷ Etwa die zivilgesellschaftliche Organisation AlgorithmWatch Schweiz, siehe Automatisierte Entscheidungssysteme im öffentlichen Sektor – einige Empfehlungen, 25. Februar 2022, abrufbar unter <https://algorithmwatch.ch/de/adm-offentlichersektor-empfehlungen/> (zuletzt besucht am 19.10.2022).

¹¹⁸ BRAUN BINDER et al., Studie KI 2021, S. 50.

¹¹⁹ BRAUN BINDER et al., Studie KI 2021, S. 65 ff., 76 ff.

rechtlichen Befassung mit dem KI-Einsatz in der öffentlichen Verwaltung im Blick zu behalten. Nachfolgend findet sich deshalb ein kurzer Überblick über entsprechende Vorhaben und Entwicklungen.

1. Schweiz

Der Bund stellte im Jahr 2019 fest, der allgemeine Rechtsrahmen in der Schweiz reiche zum gegenwärtigen Zeitpunkt grundsätzlich aus, um mit Herausforderungen wie KI umzugehen.¹²⁰ Im Jahr darauf wurden sieben Leitlinien erlassen, die den Einsatz von KI in der Bundesverwaltung transparent, verantwortungsvoll und sicher gestalten sollen.¹²¹ Gleichzeitig wurde das Kompetenznetzwerk für künstliche Intelligenz (CNAI) ins Leben gerufen, das Anfang des Jahres 2022 operativ tätig wurde.¹²² Das CNAI soll die Vertrauenswürdigkeit in und von KI oder anderen Technologien innerhalb der Bundesverwaltung und auch darüber hinaus fördern und zur Information der Öffentlichkeit im jeweiligen Bereich beitragen.

Im April 2022 veröffentlichte das Eidgenössische Departement für auswärtige Angelegenheiten (EDA) einen Bericht, in dem es unterschiedliche KI-Regulierungsansätze im entstehenden internationalen Regelwerk beleuchtet.¹²³ Darin wird ausgeführt, dass die Schweiz zu einer massvollen internationalen Regulierung von KI beitragen werde, die Diskussionen in der Schweiz allerdings bis anhin von weniger umfassendem Regulierungsbedarf ausgehen.¹²⁴

Neben dem Bund beschäftigen sich auch verschiedene Kantone mit einem möglichen KI-Einsatz.¹²⁵ Ausserdem engagieren sich zivilgesellschaftliche Organisationen¹²⁶ und die Wissenschaft¹²⁷ für das Schaffen eines rechtlichen Rahmens zum rechtskonformen und nachhaltigen KI-Einsatz. Der Ansatz soll jedoch ein anderer sein, als ihn bspw. die EU-Kommission verfolgt. Wo erforderlich, sollen

¹²⁰ Bericht IDAG KI 2019.

¹²¹ Leitlinien KI Bund 2020, S. 2.

¹²² Zur Website des CNAI siehe <https://cnaai.swiss/> (zuletzt besucht am 11.10.2022).

¹²³ Bericht EDA KI 2022.

¹²⁴ Bericht EDA KI 2022, S. 26.

¹²⁵ Siehe etwa BRAUN BINDER et al., Studie KI 2021; für eine Übersicht zum *status quo* der Automatisierung kantonaler Prozesse GLASER, Automatisierte Verwaltungsverfahren, S. 159 ff.

¹²⁶ So etwa AlgorithmWatch Schweiz (<https://algorithmwatch.ch>) oder die Digitale Gesellschaft (<https://www.digitale-gesellschaft.ch/>).

¹²⁷ BRAUN BINDER et al., Handlungsbedarf, N 1 ff.; THOUVENIN et al., S. 1 ff.

sektorspezifische Regelungen geschaffen und bestehendes Recht angepasst werden – die Schaffung eines separaten KI-Gesetzes wird abgelehnt.¹²⁸

2. International

Auf internationaler Ebene – in einem europäischen Kontext – sind insbesondere zwei geplante Regelwerke zu nennen. Im Jahr 2021 stellte die EU-Kommission einen Verordnungsvorschlag zur Regulierung von Künstlicher Intelligenz vor (EU-KI-VOV¹²⁹) und entfachte damit eine breite Debatte über die rechtlichen Herausforderungen, die mit einem KI-Einsatz einhergehen.¹³⁰ Die EU-Kommission folgt dabei einem vierstufigen risikobasierten Ansatz, wobei die Klassifizierung entsprechend dem Grad der Risiken für Gesundheit, Sicherheit und Grundrechte erfolgt.¹³¹

Die vier Kategorien sind: unzulässiges Risiko, hohes Risiko, begrenztes Risiko, minimales Risiko. Ein unzulässiges Risiko liegt etwa bei der Verwendung von biometrischen Echtzeit-Identifizierungssystemen zu Strafverfolgungszwecken im öffentlichen Raum vor.

Weiter erarbeitet der Ausschuss für Künstliche Intelligenz (*Committee on Artificial Intelligence, CAI*)¹³² des Europarats eine Konvention zur Regulierung von Künstlicher Intelligenz.

VI. Fazit

KI-Systeme werden mittelfristig in verschiedenen Verwaltungsbereichen zum Einsatz kommen. Im Moment existiert jedoch weder auf kantonaler noch auf Bundesebene eine Übersicht über die von den Behörden eingesetzten oder geplanten KI-Systeme. Dies erschwert die rechtliche Befassung mit den Entwicklungen. Aus einer allgemeinen Perspektive ist festzuhalten, dass der KI-

¹²⁸ Insbesondere THOUVENIN et al., S. 1 ff.

¹²⁹ Siehe bereits Fn. 8.

¹³⁰ Etwa BARKANE, S. 147 ff.; SMUHA et al., S. 9 ff.

¹³¹ Siehe Fn. 9.

¹³² Das CAI ist das Folgekomitee des Ad hoc Committee on Artificial Intelligence (CAHAI), das von 2019–2021 eine Auslegeordnung über Elemente eines möglicherweise rechtlich bindenden Instruments erarbeitete, siehe <https://www.coe.int/en/web/artificial-intelligence/cahai> (zuletzt besucht am 11.10.2022). Den Vorsitz des CAI führt der Schweizer Botschafter Thomas Schneider des Bundesamtes für Kommunikation (BAKOM). Weitere Informationen unter <https://www.coe.int/en/web/artificial-intelligence/cai> (zuletzt besucht am 19.10.2022).

Einsatz in der Verwaltung Regulierungsbedarf sowohl in Querschnittsgesetzen (insbesondere im RVOG und im VwVG) als auch in bereichsspezifischen Rechtsgrundlagen (z. B. im Steuer- oder Sozialversicherungsrecht) erfordert.

VII. Literatur

- AUER CHRISTOPH, BINDER ANJA MARTINA, Kommentierung des Art. 12 VwVG, in: Auer Christoph/Müller Markus/Schindler Benjamin (Hrsg.), 2. Aufl., Zürich/St. Gallen 2019 (zit. VwVG Komm)
- BAERISWYL BRUNO, Kommentierung des Art. 4 DSG, in: Baeriswyl Bruno/Pärli Kurt (Hrsg.), Datenschutzgesetz, Bern 2015 (zit. SHK DSG)
- BANSAK KIRK/FERWERDA JEREMY/HAINMUELLER JENS/DILLON ANDREA/HANGARTNER DOMINIK/LAWRENCE DUNCAN/WEINSTEIN JEREMY, Improving refugee integration through data-driven algorithmic assignment, in: Science 2018, S. 325 ff.
- BARKANE IRENA, Questioning the EU proposal for an Artificial Intelligence Act: The need for prohibitions and a stricter approach to biometric surveillance, in: Information Polity 27/2022, S. 147 ff.
- BIERI ADRIAN/POWELL JULIAN, Informationspflicht nach dem totalrevidierten Datenschutzgesetz, in: AJP 2020, S. 1533 ff.
- BRAUN BINDER NADJA, Künstliche Intelligenz und automatisierte Entscheidungen in der öffentlichen Verwaltung, in: SJZ 115/2019, S. 467 ff. (zit. KI Verwaltung)
- BRAUN BINDER NADJA, Automatisierte Einzelentscheidungen: Perspektive Datenschutzrecht und öffentliche Verwaltung, in: SZW 2020, S. 27 ff. (zit. Perspektive Datenschutzrecht)
- BRAUN BINDER NADJA, Als Verfügungen gelten Anordnungen der Maschinen im Einzelfall... – Dystopie oder künftiger Verwaltungsalltag?, in: ZSR 139/2020, S. 253 ff. (zit. Dystopie)
- BRAUN BINDER NADJA, Der Untersuchungsgrundsatz als Herausforderung vollautomatisierter Verfahren, in: zsis) 2/2020, S. 26 ff. (zit. Untersuchungsgrundsatz)
- BRAUN BINDER NADJA, Algorithmisch gesteuertes Risikomanagement im digitalisierten Besteuerungsverfahren, in: Unger Sebastian/von Ungern-Sternberg Antje (Hrsg.), Demokratie und künstliche Intelligenz, Tübingen 2019, S. 161 ff. (zit. Risikomanagement)
- BRAUN BINDER NADJA, Ausschliesslich automationsgestützt erlassene Steuerbescheide und Bekanntgabe durch Bereitstellung zum Datenabruf, in: DStZ 2016, S. 526 ff.
- BRAUN BINDER NADJA/BURRI THOMAS/LOHMANN MELINDA FLORINA/SIMMLER MONIKA/THOUVENIN FLORENT/VOKINGER KERSTIN NOËLLE, Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht, in: Jusletter 28. Juni 2021 (zit. BRAUN BINDER et al., Handlungsbedarf)
- BRAUN BINDER NADJA/KUNZ ELIANE/OBRECHT LILIANE, Maschinelle Gesichtserkennung im öffentlichen Raum, in: sui generis 2022, S. 35 ff.
- BRAUN BINDER NADJA/SPIELKAMP MATTHIAS/EGLI CATHERINE/FREIBURGHaus LAURENT/KUNZ ELIANE/LAUKENMANN NINA/LOI MICHELE/MÄTZENER ANNA/OBRECHT LILIANE/WULF JESSICA, Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28.02.2021 zum Vorprojekt IP6.4 zu Händen des Kantons Zürich (zit. BRAUN BINDER et al., Studie KI 2021), abrufbar unter 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 (zuletzt besucht am 10.10.2022).
- BUCHHOLTZ GABRIELE/SCHEFFEL-KAIN MARTIN, Algorithmen und Proxy Discrimination in der Verwaltung: Vorschläge zur Wahrung digitaler Gleichheit, in: NVwZ 9/2022, S. 612 ff.

- DEMAJ LABINOT/SÄGESSER PATRICK, Chatbots für Organisationen: Anatomie, Potentiale und Anwendungsmöglichkeiten zur direkten Kommunikation mit Anspruchsgruppen, Diskussionspapier, abrufbar unter https://www.byerley.ch/assets/pdf/byerley_2017_Chatbots_Diskussionspapier.pdf (zuletzt besucht am 10.10.2022)
- DIGGELMANN LIVIER/EMERY MATTHIAS/ENZLER LIVIA/RÜFLI DANIEL/STIRNIMANN MARTINA/TSCHUDI LARISSA, Der Diskriminierungsbegriff im öffentlichen und grundrechtlichen Diskurs, in: ZSR 2/2022, S. 211 ff.
- DRÜEN KLAUS-DIETER, Amtsermittlungsgrundsatz und Risikomanagement, in: DSJG 2019, S. 193 ff.
- ERTEL WOLFGANG, Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung, 4. Aufl., Wiesbaden 2016
- FISCHER CLAUDIO/DAEPP ANNEMARIE, Digitalisierung am Beispiel der Steuerverwaltung des Kantons Bern, in: EF 4/2019, S. 327 ff.
- FRIEDMANN BATYA/NISSENBAUM HELEN, Bias in Computer Systems, in: ACM Transactions of Information Systems, 3/1996, S. 330 ff.
- GLASER ANDREAS, Einflüsse der Digitalisierung auf das schweizerische Verwaltungsrecht, in: SJZ 2018, S. 181 ff. (zit. Einflüsse)
- GLASER ANDREAS, Automatisierte Verwaltungsverfahren in den Schweizer Kantonen, in: Braun Binder Nadja/Bussjäger Peter/Eller Matthias (Hrsg.), Auswirkungen der Digitalisierung auf die Erlassung und Zuordnung behördlicher Entscheidungen, Wien/Hamburg 2021, S. 159 ff. (zit. Automatisierte Verwaltungsverfahren)
- GOLDA THOMAS/CORMIER MICKAEL/BEYERER JÜRGEN, Intelligente Bild- und Videoauswertung für die Sicherheit, in: Wehe Dieter (Hrsg.), Handbuch Polizeimanagement – Polizeipolitik – Polizeiwissenschaft – Polizeipraxis, Wiesbaden 2022, S. 1 ff.
- GUCKELBERGER ANNETTE, Verwaltung im Zeitalter der Digitalisierung, Baden-Baden 2019
- HORNUNG GERRIT, KI-Regulierung im Mehrebenensystem, in: Datenschutz und Datensicherheit 9/2022, S. 561 ff.
- KAPLAN JERRY, Artificial Intelligence. What everyone needs to know, New York 2016
- KIENER REGINA/KÄLIN WALTER/WYTTENBACH JUDITH, Grundrechte, 3. Aufl., Bern 2018
- KIENER REGINA/RÜTSCHKE BERNHARD/KUHN Mathias, Öffentliches Verfahrensrecht, 3. Aufl., Zürich/St. Gallen 2021
- KNEUBÜHLER LORENZ/PEDRETTI RAMONA, Kommentierung des Art. 35 VwVG, in: Auer Christoph/Müller Markus/Schindler Benjamin (Hrsg.), VwVG Kommentar, 2. Aufl., Zürich/St. Gallen 2019 (zit. VwVG Komm)
- KÖLZ ALFRED/HÄNER ISABELLE/BERTSCHI MARTIN, Verwaltungsverfahren und Verwaltungsrechtspflege des Bundes, 3. Auflage, Zürich 2013
- LISCHKA KONRAD/KLINGEL ANITA, Wenn Maschinen Menschen bewerten: Internationale Fallbeispiele für Prozesse algorithmischer Entscheidungsfindung, Arbeitspapier der Bertelsmann Stiftung 2017, abrufbar unter https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/ADM_Fallstudien.pdf (zuletzt besucht am 19.10.2022)
- MEYER CHRISTIAN, Die Mitwirkungsmaxime im Verwaltungsverfahren des Bundes: Ein Beitrag zur Sachverhaltsfeststellung als arbeitsteiligem Prozess, Diss. Zürich 2019
- MÜLLER ANGELA, Der Artificial Intelligence Act der EU: Ein risikobasierter Ansatz zur Regulierung von Künstlicher Intelligenz, in: EuZ 1/2022, S. A1 ff.
- MÜLLER JÖRG PAUL/SCHEFER MARKUS, Grundrechte in der Schweiz, 4. Aufl., Bern 2008
- NIKITAS ALEXANDROS/MICHALAKOPOULOU KALLIOPI/TCHOUAMOU NJOYA ERIC/KARAMPATZAKIS DIMITRIS, Artificial Intelligence, Transport and the Smart City: Definitions and Dimensions of a New Mobility Era, in: Sustainability 7/2020, S. 1 ff.
- NUFER MARIANNE, Künstliche Intelligenz in der Steuerveranlagung, in: ASA 88/2019/2020, S. 259 ff.

- OBRECHT LILIANE, Verfügung und automatisierte Einzelentscheidung – *same same but different?*, in: *ex/ante* 2/2022, S. 38 ff.
- OETTINGER MICHAEL, Data Science: Eine praxisorientierte Einführung im Umfeld von Machine Learning, künstlicher Intelligenz und Big Data, 2. Aufl., Hamburg 2020
- PRINCE ANYA E. R./SCHWARCZ DANIEL, Proxy Discrimination in the Age of Artificial Intelligence and Big Data, in: *Iowa Law Review* 2020, S. 1257 ff.
- PULLEN JENNIFER/SCHEFER PATRICIA, Predictive Policing – Grundlagen, Funktionsweisen und Wirkung, in: Simmler Monika (Hrsg.), *Smart Criminal Justice*, Basel 2021, S. 103 ff.
- RECHSTEINER DAVID, Der Algorithmus verfügt, in: *Jusletter* 26. November 2018
- REITER CATHERINE, Künstliche Intelligenz im Verwaltungsverfahren, in: *AJP* 2022, S. 984 ff.
- RHINOW RENÉ/KOLLER HEINRICH/KISS CHRISTINA/THURNHERR DANIELA/BRÜHL-MOSER DENISE, *Öffentliches Prozessrecht*, 4. Aufl., Basel 2021
- ROTH-ISIGKEIT DAVID, Die Begründung des vollständig automatisierten Verwaltungsakts, in: *DÖV* 2020, S. 1018 ff.
- SCHINDLER BENJAMIN, Kommentierung des Art. 5 Rz. 1–56 BV, in: *Ehrenzeller Bernhard/Schindler Benjamin/Schweizer Rainer J./Vallender Klaus A.* (Hrsg.), 3. Aufl., Zürich/Basel7genf 2014 (zit. SGK BV)
- SIMMLER MONIKA/BRUNNER SIMONE, *Smart Criminal Justice in der Schweiz – Die Kantone im Bann der Algorithmen?*, in: Simmler Monika (Hrsg.), *Smart Criminal Justice*, Basel 2021, S. 10 ff.
- SIMMLER MONIKA/BRUNNER SIMONE/SCHEDLER KUNO, *Smart Criminal Justice – Eine empirische Studie zum Einsatz von Algorithmen in der Schweizer Polizeiarbeit und Strafrechtspflege*, St. Gallen 2020, S. 14 ff., 31 ff., abrufbar unter https://www.alexandria.unisg.ch/261666/1/Simmler_et_al._Smart_Criminal_Justice_Forschungsbericht_vom_10.12.2020.pdf (zuletzt besucht am 10.10.2022)
- SMUHA NATHALIE/AHMED-RENGERS EMMA/HARKENS ADAM/LI WENLONG/MCLAREN JAMES/PISELLI RICCARDO/YEUNG KAREN, How the EU can achieve Legaly Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act, Birmingham 2021 (zit. SMUHA et al.), abrufbar unter https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991 (zuletzt besucht am 19.10.2022)
- STEVENS JEREMY, Datenqualität bei algorithmischen Entscheidungen, in: *CR* 2/2020, S. 73 ff.
- SUTTER PATRICK, Kommentierung von Art. 30 VwVG, in: *Auer Christoph/Müller Markus/Schindler Benjamin, VwVG – Bundesgesetz über das Verwaltungsverfahren Kommentar*, 2. Aufl., Zürich 2019 (zit. VwVG Komm)
- THOUVENIN FLORENT/BRAUN BINDER NADJA, Transparenz durch Datenschutzerklärungen, in: *ZSR* 1/2022, S. 5 ff.
- THOUVENIN, FLORENT/CHRISTEN MARKUS/BERNSTEIN ABRAHAM/BRAUN BINDER NADJA/BURRI THOMAS/DONNAY KARSTEN/JÄGER LENA/JAFFÉ, MARIELA/KRAUTHAMMER MICHAEL/LOHMANN MELINDA/MÄTZENER ANNA/MÜTZEL SOPHIE/OBRECHT LILIANE/RITTER NICOLE/SPIELKAMP MATTHIAS/VOLZ STEPHANIE, Positionspapier: Ein Rechtsrahmen für Künstliche Intelligenz, 2021, abrufbar unter <https://dsi.uzh.ch/dam/jcr:3a0cb402-c3b3-4360-9332-f800895fdc58/dsi-strategy-lab-21-de.pdf> (zuletzt besucht am 10.10.2022)
- THURNHERR DANIELA, Automatisierte Verwaltungsverfahren auf Bundesebene (Schweiz), in: *Braun Binder Nadja/Bussjäger Peter/Eller Matthias* (Hrsg.), *Auswirkungen der Digitalisierung auf die Erlassung und Zuordnung behördlicher Entscheidungen*, Wien/Hamburg 2021, S. 133 ff.
- TSCHANNEN PIERRE/MÜLLER MARKUS/KERN MARKUS, *Allgemeines Verwaltungsrecht*, 5. Aufl., Bern 2022
- VOSS JAKOB, Datenqualität als Grundlage qualitativer Inhaltserschliessung, in: *Franke-Maier Michael/Kasparzik Anna/Ledl Andreas/Schürmann Hans*, *Qualität in der Inhaltserschliessung*, Berlin/Boston 2021, S. 167 ff.

- WALDMANN BERNHARD, Kommentierung von Art. 8 BV, in: Waldmann Bernhard, Belser Eva Maria/Epiney Astrid (Hrsg.), Basler Kommentar Bundesverfassung, Basel 2015 (zit. BSK BV)
- WALDMANN BERNHARD, Kommentierung von Art. 29 BV, in: Waldmann Bernhard, Belser Eva Maria/Epiney Astrid (Hrsg.), Basler Kommentar Bundesverfassung, Basel 2015 (zit. BSK BV)
- WEBER ROLF H., Automatisierte Entscheidungen: Perspektive Grundrechte, in: SZW 2020, S. 18 ff.
- WEBER ROLF H./HENSELER SIMON, Regulierung von Algorithmen in der EU und in der Schweiz, in: EuZ 2020, S. 28 ff.
- WEDER REGINA, Verfahrensgrundrechtliche Anforderungen an automatisierte Verwaltungsverfahren, in: Simmler Monika (Hrsg.), Smart Criminal Justice: Der Einsatz von Algorithmen in der Polizeiarbeit und Strafrechtspflege, Basel 2021, S. 237 ff.
- WIEDERKEHR RENÉ, Transparenz als Grundsatz rechtsstaatlichen Handelns (Art. 5 BV), in: ZBl 2007, S. 521 ff.
- WIEDERKEHR RENÉ/MEYER CHRISTIAN, Schranken, Ausnahmen und Relativierungen des rechtlichen Gehörs: Insbesondere mit Blick auf automatisierte Verwaltungsverfahren, in: AJP 2022, S. 1092 ff.
- WIEDERKEHR RENÉ/MEYER CHRISTIAN/BÖHME ANNA, Transparenz als Verfahrensgrundsatz: Gesetzliche Transparenztechniken und praktische Argumentationsfelder, in: recht 2022, S. 131 ff.
- WILLIAMS MARY-ANNE, Explainable artificial intelligence, in: Vogl Roland, Research Handbook on Big Data Law, Cheltenham/Northampton 2021, S. 318 ff.
- WOLFANGEL EVA, Wo hat sie das nur gelernt?, in: Bischoff Manon (Hrsg.), Künstliche Intelligenz: Vom Schachspieler zur Superintelligenz?, Berlin 2022, S. 171 ff.
- ZANELLA ANDREA/BUI NICOLA/CASTELLANI ANGELO/VANGELISTA LORENZO/ZORZI MICHELE, Internet of Things for Smart Cities, in: IEEE Internet of Things Journal, 1/2014, S. 22 ff.
- ZWEIG KATHARINA, Ein Algorithmus hat kein Taktgefühl, München 2019
- ZWEIG KATHARINA A./KRAFFT TOBIAS D., Fairness und Qualität algorithmischer Entscheidungen, in: Mohabbat Kar Resa/Thapa Basanta/Parycek Peter (Hrsg.), (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft, Berlin 2018, S. 204 ff.

VIII. Materialien

- Eidgenössisches Departement für auswärtige Angelegenheiten EDA, Künstliche Intelligenz und internationales Regelwerk, Bericht an den Bundesrat, 2022 (zit. Bericht EDA KI 2022), abrufbar unter <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-88019.html> (zuletzt besucht am 19.10.2022)
- Leitlinien «Künstliche Intelligenz» für den Bund – Orientierungsrahmen für den Umgang mit künstlicher Intelligenz in der Bundesverwaltung, 2020 (zit. Leitlinien KI Bund 2020), abrufbar unter <https://www.sbf.admin.ch/sbfi/de/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html> (zuletzt besucht am 19.10.2022)
- Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat: Herausforderungen der künstlichen Intelligenz, 2019, abrufbar unter <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-77514.html> (zit. Bericht IDAG KI 2019)
- Gutachten der Datenethikkommission der Bundesregierung, Berlin 2019, abrufbar unter https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digital-politik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6 (zit. Datenethikkommission Gutachten 2019)
- Recommendation of the Council on Artificial Intelligence, OECD 2019, abrufbar unter <https://oecd.ai/en/ai-principles> (zit. OECD Prinzipien 2019)

KI in der Verwaltung: Entwicklungen und Herausforderungen

Nadja Braun Binder / Nina Laukenmann / Liliane Obrecht

Erstabdruck in: Jusletter IT vom 4. Juli 2024

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.

Inhalt

I.	Einleitung	28
II.	Begriffsverständnis	29
III.	Ausgewählte Einsatzbereiche in der öffentlichen Verwaltung	31
	1. Standardisierung und Massenverwaltung	31
	2. Chatbots und Voicebots	32
	3. Intelligente Informationsdurchsuchung	33
	4. Prognosen	34
IV.	Rechtliche Herausforderungen	35
	1. Legalitätsprinzip	35
	2. Verfahrensgarantien und Verfahrensgrundsätze	36
	2.1. Rechtliches Gehör	36
	2.2. Untersuchungsgrundsatz	39
	3. Diskriminierungsverbot	40
	3.1. Diskriminierungsarten und Erscheinungsformen	40
	3.2. Beispiele algorithmischer Diskriminierung	43
	4. Informationelle Selbstbestimmung und Datenschutz	44
	4.1. Die Verfassungsbestimmung und ihre gesetzliche Konkretisierung	44
	4.2. Beispiel 1: Maschinelle Gesichtserkennung	45
	4.3. Beispiel 2: Risikomanagementsysteme in der Steuerverwaltung	48

5.	Vertrauensschutz	50
6.	Transparenz	52
V.	Entwicklungen aus der Perspektive der Bundesverwaltung	53
VI.	Fazit	55

I. Einleitung

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

Der vorliegende Beitrag gibt zunächst einen Überblick über die verschiedenen international und national verbreiteten Begriffsverständnisse von KI-Syste-

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 29.05.2024, siehe <https://cna.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>.

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.

men ([Kapitel II](#)). Anschliessend werden ausgewählte Einsatzbereiche von KI in der öffentlichen Verwaltung präsentiert ([Kapitel III](#)). Darauf aufbauend werden die rechtlichen Herausforderungen und Rahmenbedingungen skizziert ([Kapitel IV](#)). Schliesslich wird ein Blick auf die Rechtsetzung aus der Perspektive der Bundesverwaltung geworfen ([Kapitel V](#)). Der Beitrag schliesst mit einem Fazit ([Kapitel VI](#)).

II. Begriffsverständnis

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.

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 und 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».

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 au-

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> (AI Act – AI für Artificial Intelligence).

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, 17.05.2024, CM(2024)52-final, https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680afb11f.

tonomy 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.

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

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

⁸ Siehe I.

⁹ Terminologie Kompetenznetzwerk CNAI in der Version 2.0 vom 21.12.2023, S. 7, <https://cna1.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.itsl.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>.

III. Ausgewählte Einsatzbereiche in der öffentlichen Verwaltung

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 1](#)), der Einsatz von Chatbots und Voicebots ([Abschnitt 2](#)), die intelligente Durchsuchung von Informationen ([Abschnitt 3](#)) sowie das Treffen von Prognosen ([Abschnitt 4](#)) von zentraler Bedeutung.

1. Standardisierung und Massenverwaltung

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 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 wer-

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 28.02.2021 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).

¹⁵ Siehe zu den damit einhergehenden datenschutzrechtlichen Herausforderungen [Abschnitt IV.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 DAEPF, Digitalisierung am Beispiel der Steuerverwaltung des Kantons Bern, in: EF 4/2019, S. 327 ff., S. 329.

den.¹⁷ Im Sozialversicherungsbereich stehen insbesondere die automatisierte Bearbeitung von Anträgen zur Prämienverbilligung sowie die automatisierte Betrugsbekämpfung im Vordergrund.¹⁸

2. Chatbots und Voicebots

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 Gemeindefwebseite 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

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

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

²¹ FRANZISKA OTT, Voicebot: Definition, Anbieter und Vorteile, HubSpot 31.05.2023, <https://blog.hubspot.de/service/voicebot>.

²² OTT (Fn. 21).

²³ Zu den rechtlichen Herausforderungen betreffend potenzielle fehlerhafte Chatbot-Auskünfte siehe unten [Abschnitt IV.5](#).

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

Nutzenden 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 Anrufern zu beantworten.²⁷

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. Intelligente Informationsdurchsuchung

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

²⁵ Bundesamt für Statistik, StatBot.swiss: mit offenen Daten diskutieren, <https://www.bfs.admin.ch/bfs/de/home/dscc/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, 13.09.2023, <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 26.10.2019, <https://www.luzernerzeitung.ch/zentralschweiz/luzern/kuenstliche-intelligenz-ein-computer-soll-fuer-die-luzerner-polizei-ermitteln-ld.1163004>.

intelligente Videoanalyse, da der Einsatz von KI bei der Bilderkennung und -auswertung besonders effizient ist.³¹

4. Prognosen

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

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

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>.

³³ Siehe unten [Abschnitt III.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.

IV. Rechtliche Herausforderungen

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 1](#)), die Verfahrensgarantien und Verfahrensgrundsätze ([Abschnitt 2](#)), das Diskriminierungsverbot ([Abschnitt 3](#)), die informationelle Selbstbestimmung und das Datenschutzrecht ([Abschnitt 4](#)) sowie der Vertrauensschutz ([Abschnitt 5](#)) sind dabei von entscheidender Bedeutung. Schliesslich wird das Schlagwort «Transparenz» im Kontext von KI erörtert ([Abschnitt 6](#)).

1. Legalitätsprinzip

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

³⁷ Bundesverfassung der Schweizerischen Eidgenossenschaft vom 18.04.1999, 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.

⁴³ Siehe dazu vertieft [Abschnitt IV.4](#).

2. Verfahrensgarantien und Verfahrensgrundsätze

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 2.1](#)).

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 2.2](#)).⁴⁸

2.1. Rechtliches Gehör

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

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 NADIA 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.

der Anspruch auf vorgängige Äusserung und Mitwirkung im Verfahren sowie das Recht auf Begründung.⁵¹ Diese beiden Teilgehalten sind im Kontext von KI von besonderer Relevanz.

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

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 na-

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.

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/WYTTEBACH (Fn. 44), § 41 N 40 ff.

⁵⁸ Siehe zu den verschiedenen Automatisierungsgraden [Abschnitt II](#).

türlichen 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.

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 Wei-

⁵⁹ NADIA BRAUN BINDER, *Automatisierte Entscheidungen: Perspektive Datenschutzrecht und öffentliche Verwaltung*, SZW 2020, S. 27 ff., S. 28 f.; NADIA 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*, Jusletter 26.11.2018, Rz. 18 ff.

⁶⁰ NADIA 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.

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.

teres 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

2.2. Untersuchungsgrundsatz

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

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 Schon aufgrund der Definition der Vollautomatisierung unternehmen Behördenmitarbeitende im vollautomatisierten Verfahren keine weiteren Abklärungen zur Sachverhaltsermittlung.⁷³ Dies birgt die Gefahr, dass nur

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.

⁷³ Siehe [Abschnitt II](#).

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

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

3. Diskriminierungsverbot

Dass der KI-Einsatz zu Diskriminierungen führen kann ([Abschnitt 3.1](#)), ist mittlerweile unbestritten. Dies zeigen auch zwei reale Beispiele aus den Niederlanden und Frankreich ([Abschnitt 3.2](#)).

3.1. Diskriminierungsarten und Erscheinungsformen

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 Wirk-

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 NADIA 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.

lichkeit 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 aufgebbaaren Bestandteil der Identität der betreffenden Person ausmacht»⁷⁹.

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.

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.

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.

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/SCHEFFEL-KAIN (Fn. 81), S. 612.

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

⁸⁴ BUCHHOLTZ/SCHEFFEL-KAIN (Fn. 81), S. 613.

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

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 voll-

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>.

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

automatisierte 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.

3.2. Beispiele algorithmischer Diskriminierung

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

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

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 26.10.2021, <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 29.03.2022, <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 29.12.2021, <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 da-

Schülerinnen und Schüler aus den letzten Schuljahren 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. Informationelle Selbstbestimmung und Datenschutz

Im Folgenden wird zunächst auf die allgemeinen Herausforderungen des KI-Einsatzes in Bezug auf die informationelle Selbstbestimmung und den Datenschutz eingegangen ([Abschnitt 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 4.2](#)) und um den Einsatz von Risikomanagementsystemen im Steuerveranlagungsverfahren ([Abschnitt 4.3](#)).

4.1. Die Verfassungsbestimmung und ihre gesetzliche Konkretisierung

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

zu 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* 2023, S. 165 ff., S. 175. Seit dem 21.05.2018 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 andere Faktoren als die Noten eingeführt (siehe WALLY BORDAY, *Parcoursup: quelles sont les nouveautés de la plateforme en 2020?*, *le figaro étudiant* 20.01.2020, https://etudiant.lefigaro.fr/article/parcoursup-queles-sont-les-nouveautes-de-la-plateforme-en-2020_557794fe-3b5f-11ea-81a1-30c9bbb28c82/).

2023: 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>).

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 insbesondere die in Art. 6 DSG 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

4.2. Beispiel 1: Maschinelle Gesichtserkennung

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

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.

¹⁰² Die Ausführungen lehnen sich an die folgenden zwei Publikationen an: NADIA 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>; NADIA BRAUN BINDER/LILIANE OBRECHT, Algorithmisch überprüfte Steuererklärung im ordentlich gemischten 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>.

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

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 verhältnismässig sein.¹¹² Dabei stellt sich insbesondere die Frage, ob die mit der

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 09.02.2019, <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 22.03.2012, 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: 16.08.2023, siehe <https://onlinekommentar.ch/de/kommentare/dsg34>, Rz. 18 f.

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 ver-

maschinellen Gesichtserkennung verfolgten Ziele (z.B. der Schutz der öffentlichen Ordnung und Sicherheit) nicht auch mit anderen Mitteln erreicht werden können.

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, Menschhandel oder sexueller Ausbeutung, vorgesehen sind.

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

nünftiges Verhältnis zwischen der Datenbearbeitung und dem Grundrechtseingriff bestehen.

¹¹³ Art. 6 Abs. 3 DSG.

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 DSG.

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

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.

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

4.3. Beispiel 2: Risikomanagementsysteme in der Steuerverwaltung

Die fortschreitende Digitalisierung der Steuerverwaltungen geht mit einer zunehmenden Automatisierung des Steueranlagungsverfahrens 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,

¹²⁰ Art. 6 Abs. 5 DSGVO.

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

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

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 01.01.2017 eine Rechtsgrundlage in Kraft getreten, die die verpflichtende Einführung von Risikomanagementsystemen enthält (§ 88 Abs. 5 Abgabenordnung in der Fassung der Bekanntmachung vom 01.10.2002 (BGBl. I S. 3866; 2003 I S. 61), die zuletzt durch Artikel 1 des Gesetzes vom 12.07.2022 (BGBl. I S. 1142) geändert worden ist). In Österreich besteht ebenfalls eine entsprechende Rechtsgrundlage (§§ 48a, 114 Abs. 4 Bundesgesetz vom 28.06.1961 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.

da eine Vollautomation ohne entsprechende Kontrollmöglichkeiten das Risiko birgt, dass falsche Angaben in Steuererklärungen nicht erkannt werden.¹²⁶

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 anderen Zweck nicht durch die gesetzliche Grundlage gedeckt, welche die ursprüng-

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 DSG, zum Begriff der Personendaten Art. 5 lit. a DSG.

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

¹²⁹ Siehe zu den datenschutzrechtlichen Grundsätzen [Abschnitt IV.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.

liche Erhebung rechtfertigt.¹³² Unzulässig ist auch eine Datenerhebung «auf Vorrat», um diese später in einem anderen Kontext zu verwenden.¹³³

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 DSG handeln,¹³⁵ das zumindest auf Bundesebene, aber auch in den meisten Kantonen, eine formell-gesetzliche Grundlage voraussetzt.¹³⁶

5. Vertrauensschutz

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

¹³² Eine Zweckänderung kann gesetzlich zwar vorgesehen sein, wenn Steuerdaten von der Steuerverwaltung an andere Behördenstellen «bekannt gegeben» (im Bund: Art. 5 lit. e DSG) 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 DSG.

¹³⁶ Art. 34 Abs. 2 lit. b DSG. 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 DSG) oder der Bearbeitungszweck für die Grundrechte der betroffenen Person keine besonderen Risiken birgt (Art. 34 Abs. 3 lit. b DSG). 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/>.

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

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 Chat-

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) Vorbeihaltlosigkeit 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. 141), Rz. 624 ff.).

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 Survery 2022, S. 1 ff.

¹⁴⁷ Siehe dazu etwa PENGHAO ZHAO et al., Retrieval-Augmented Generation for AI-Generated Content: A Survey, arXiv 29.02.2024, 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, Berkeley 2024, S. 66, ZHAO et al. (Fn. 147), S. 3.

bot-Antworten durch RAG nur verringert oder gar gänzlich ausgeschlossen werden kann.¹⁴⁹

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 Auskunftserteilung geeignet sind.¹⁵¹

6. Transparenz

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

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 zivilgesell-

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/sbf/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 IV.4](#) sowie BRAUN BINDER/OBRECHT, SVVOR Jahrbuch (Fn. 76), S. 39.

¹⁵⁴ Siehe [Abschnitt IV.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 16.12.2021 betreffend ein öffentliches Verzeichnis der in der Verwaltung eingesetzten Algorithmen; im Kanton Basel-

schaftlicher 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

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.

V. Entwicklungen aus der Perspektive der Bundesverwaltung

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 zuneh-

Stadt Anzug 23.5474.01 betreffend Transparenz über den Einsatz von algorithmischen Systemen in der Verwaltung vom 18.10.2023.

¹⁵⁷ Etwa die zivilgesellschaftliche Organisation AlgorithmWatch Schweiz, siehe Automatisierte Entscheidungssysteme im öffentlichen Sektor – einige Empfehlungen, 25.02.2022, <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 I](#). Es handelt sich dabei allerdings nicht um eine tatsächliche Datenbank, sondern um ein PDF-Dokument.

¹⁶¹ Der Bundesrat, Strategie «Digitale Schweiz», verabschiedet am 05.09.2018, <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-72053.html>.

mende 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, 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

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

¹⁶² Bericht IDAG KI 2019 (Fn. 152).

¹⁶³ Siehe dazu Der Bundesrat, Künstliche Intelligenz: Schweiz befindet sich in guter Ausgangslage, Medienmitteilung vom 13.12.2019, <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», 25.11.2020, <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 22.11.2023, <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-98791.html>.

¹⁶⁷ Siehe Medienmitteilung in Fn. 166.

VI. Fazit

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.

Whitepaper: Transparenz durch Begründung von Verfügungen

Nadja Braun Binder / Liliane Obrecht

Erstveröffentlichung: https://ius.unibas.ch/fileadmin/user_upload/ius/03_Forschung/10_e-PIAF/WhitePaper_Begrueendung_Juni24.pdf

Verfügungen sind in der Regel zu begründen. Werden beim Erlass von Verfügungen KI-Systeme oder algorithmische Entscheidungssysteme eingesetzt, kann dies für die Begründung eine Herausforderung sein, da je nach technischer Ausgestaltung des Systems dessen Ergebnisse nicht nachvollziehbar sind. Entsprechend kann der Einsatz solcher Systeme dazu führen, dass Verfügungen nicht oder nur mit zusätzlichen Vorkehrungen begründet werden können. Unter Berücksichtigung der Funktionen der staatlichen Begründungspflicht skizziert das vorliegende White Paper konkrete Herausforderungen und zeigt mögliche Lösungsansätze auf.

Inhalt

I.	Problemstellung	58
II.	Begründungspflicht und Begründungsdichte	58
III.	Herausforderungen bei algorithmenbasierten Verfügungen	60
IV.	Auswirkungen auf die Funktionen der Begründung	61
	1. Rechtsmittelfunktion	61
	2. Akzeptanzfunktion	61
	3. Rechtssicherheitsfunktion	62
	4. Selbstkontrollfunktion	62
	5. Fremdkontrollfunktion	63
V.	Lösungsansätze	63
	1. Kennzeichnungspflicht	63
	2. Verweis auf weiterführende Informationen	63
	3. Wahl des Modells/der Technik	63
	4. Menschliche Verantwortung	64
	5. Reduktion der Begründungsdichte?	64

I. Problemstellung

Von verschiedenen Seiten, namentlich der Politik, von zivilgesellschaftlichen Organisationen, aber auch der Wissenschaft, wird für den Einsatz von Künstlicher Intelligenz (KI) bzw. von algorithmischen Entscheidungssystemen in der öffentlichen Verwaltung Transparenz gefordert. Dazu zählt die Forderung nach einer allgemeinen, gesellschaftlich breit angelegten Transparenz darüber, welche Behörde zu welchen Zwecken derartige Systeme einsetzt. Dies kann (insbesondere) durch ein entsprechendes öffentliches Verzeichnis erfüllt werden (siehe White Paper «Transparenz durch öffentliches Verzeichnis»).

Transparenz wird ferner mit Blick auf konkrete Einzelfälle gefordert, die mithilfe eines algorithmischen bzw. KI-Systems entschieden werden. Der Begriff algorithmische Systeme zielt nicht auf eine bestimmte Technologie oder Methode ab, sondern erfasst alle Anwendungen, die auf (Computer-)Algorithmen basieren. Er ist insofern für die Anliegen des vorliegenden White Papers besser geeignet als der Begriff KI-System, als er auch Technologien umfasst, die die gleichen Auswirkungen auf Einzelne und die Gesellschaft haben wie KI-Systeme.

Die von einer **algorithmenbasierten Entscheidung** betroffene Person soll diese Entscheidung nachvollziehen können. Diese Forderung findet ihre Basis für Verwaltungsverfahren u. a. in den allgemeinen verfassungsrechtlichen Verfahrensgarantien gemäss Art. 29 Bundesverfassung (BV, SR 101). Dazu zählt insbesondere die **Begründungspflicht**, der im vorliegenden Kontext eine zentrale Bedeutung zukommt.

Mit «Entscheidungen» sind im Kontext von Verwaltungsverfahren «Verfügungen» gemeint. Unter **algorithmenbasierten Verfügungen** werden Verfügungen verstanden, die mithilfe von algorithmischen Systemen vorbereitet und/oder erlassen werden. Im Folgenden geht es demnach um die Bedeutung und Umsetzbarkeit der Begründungspflicht bei algorithmenbasierten Verfügungen.

II. Begründungspflicht und Begründungsdichte

Als Teilaspekt des rechtlichen Gehörs (Art. 29 Abs. 2 BV) ist die Begründungspflicht staatlicher Entscheidungen Bestandteil der allgemeinen verfassungsrechtlichen Verfahrensgarantien. Das rechtliche Gehör wird im anwendbaren Verfahrensrecht weiter konkretisiert. Für das erstinstanzliche Verwaltungsverfahren auf Bundesebene sind die Bestimmungen des Verwaltungsverfah-

rensgesetzes des Bundes (VwVG, SR 172.021) massgebend. Dort wird zunächst bestätigt, was bereits aus der Verfassung hervorgeht: Verfügungen sind grundsätzlich zu begründen (Art 35 Abs. 1 VwVG). Dies gilt allerdings nicht, wenn die Behörde den Begehren der Partei(en) vollständig entspricht und die Partei keine Begründung verlangt (Art. 35 Abs. 3 VwVG). Durch die Begründung soll die betroffene Person nachvollziehen können, weshalb sich die Behörde im konkreten Fall so und nicht anders entschieden hat.

Die von der Entscheidung betroffene Person hat Anspruch auf eine **inhaltlich und umfangmässig angemessene Begründung**. Obwohl die Angemessenheit nur anhand des konkreten Einzelfalls beurteilt werden kann, haben Lehre und Rechtsprechung Kriterien entwickelt, die zur Bestimmung der Begründungsdichte herangezogen werden können:

- **Intensität des Eingriffs:** Je stärker die Entscheidung Rechte und Pflichten der betroffenen Person berührt, desto detaillierter muss die Begründung sein.
- **Komplexität der Sach- und Rechtslage:** Je komplexer die Sach- und Rechtslage, desto schwieriger ist es für die betroffene Person, die Subsumtion (Unterordnung) des Sachverhalts unter die anwendbaren Normen und damit die behördliche Entscheidung zu verstehen. Eine gesteigerte Begründungsdichte kann dies kompensieren.
- **Entscheidungsspielraum der Behörde:** Obliegt der entscheidenden Behörde kein Entscheidungsspielraum, genügt regelmässig ein Verweis auf die angewendeten Rechtsnormen. Je grösser allerdings der Entscheidungsspielraum ist, desto detaillierter ist die Entscheidung zu begründen.
- **Abweichung von einer behördlichen Praxis:** Weicht die Behörde in einem Fall von einer konstanten Praxis der Gesetzesanwendung ab, bedarf es einer detaillierten Begründung.
- **Verfügung der Massenverwaltung:** In Verwaltungsbereichen, in denen eine Vielzahl an gleichgelagerten Verfügungen ergeht, sind im Sinne der Prozessökonomie geringere Anforderungen an die Begründungsdichte zulässig.

III. Herausforderungen bei algorithmenbasierten Verfügungen

Die Herausforderungen, die sich bei der Begründung einer algorithmenbasierten Verfügung stellen, können je nach zugrundeliegendem Modell unterschiedlich sein. Dabei kann zwischen regelbasierten und datenbasierten Algorithmen unterschieden werden.

Regelbasierte Algorithmen verarbeiten eine Eingabe nach durch natürliche Personen festgelegten Regeln, um zu einem Ergebnis zu kommen. Sie sind deterministisch, verarbeiten daher eine bestimmte Eingabe immer nach denselben Abläufen. Die Algorithmen sind dabei auch determiniert, eine bestimmte Eingabe erzeugt immer dieselbe Ausgabe. Dies führt dazu, dass die Reihenfolge der einzelnen Schritte zu jedem Zeitpunkt bestimmt ist, was grundsätzlich bewirkt, dass das Zustandekommen des Ergebnisses nachvollziehbar ist. Dies bedeutet allerdings nicht, dass die betroffene Person die Entscheidung ohne Weiteres nachvollziehen kann. Dies trifft insbesondere dann zu, wenn die Grundlage des algorithmischen Entscheidungssystems komplexe Algorithmen bilden, wie etwa vielfach verzweigte Entscheidungsbäume. Die Begründung kann entsprechend herausfordernd sein.

Hingegen verarbeiten **datenbasierte *machine learning* (ML)-Algorithmen** eine Eingabe nicht anhand vorgegebener Regeln, sondern analysieren i. d. R. enorme Datensätze (Trainingsdaten) auf Korrelationen, um selbst eine verallgemeinerbare Regel aufzustellen. Diese kann sodann auf andere (unbekannte) Eingaben angewendet werden. Korrelation bedeutet, dass zwei Phänomene statistisch miteinander zusammenhängen, aber es sagt nichts über die Ursache-Wirkungs-Beziehung (Kausalität) aus. Die eingesetzten Algorithmen führen keine im Vorhinein vom Menschen definierten Regeln aus, sondern finden selbständig anhand der zur Verfügung stehenden Daten einen Lösungsweg. Derartige Algorithmen sind also wahrscheinlichkeitsbasiert. Daneben können auch Zufälle den Prozess prägen. Dies führt dazu, dass die gleiche Eingabe nicht immer dasselbe Resultat erzeugt.

Das Zustandekommen der Ergebnisse von datenbasierten ML-Algorithmen ist somit oft – je nach Komplexität der angewendeten Technologie – nicht nachvollziehbar. Dies betrifft etwa künstliche neuronale Netze (KNN), eine Klasse von ML-Verfahren, die aktuell besonders viel Aufmerksamkeit auf sich zieht und auch in der öffentlichen Verwaltung als vielfältig einsetzbar gehandelt wird. Auch in solchen Fällen stellt die Herstellung der Nachvollziehbarkeit der

algorithmenbasierten Entscheidung im Rahmen ihrer Begründung eine besondere Herausforderung dar.

IV. Auswirkungen auf die Funktionen der Begründung

Die Begründung von Verfügungen dient verschiedenen Funktionen. Eben diese Funktionen sind beim Einsatz von Algorithmen schwieriger zu erfüllen.

1. Rechtsmittelfunktion

Die Begründung muss es der betroffenen Person erlauben, die Verfügung sachgerecht anzufechten. Insbesondere muss in verständlicher Weise aus der Begründung hervorgehen, welche **tatsächlichen und rechtlichen Erwägungen** zur Entscheidung geführt haben (Kausalität zwischen der rechtlichen Handlung und der erlassenen Verfügung).

Datenbasierte ML-Algorithmen: Es ist möglich, dass sich die Ergebnisse von datenbasierten ML-Algorithmen gerade nicht auf kausale Beziehungen, sondern auf Korrelationen zwischen den einzelnen Daten stützen. Es ist diesfalls kaum darzulegen, welche Sachverhaltsdaten und Rechtsgrundlagen ausschlaggebend für das Ergebnis des Systems im Einzelfall waren.

Regelbasierte Algorithmen: Insbesondere bei komplexen regelbasierten Algorithmen reicht es nicht aus, die einzelnen Rechenschritte aufzulisten. Es bedarf zur Gewährleistung der Rechtsmittelfunktion der Begründung einer Herausarbeitung und verständlichen Formulierung der für die Entscheidung ausschlaggebenden tatsächlichen und rechtlichen Erwägungen.

2. Akzeptanzfunktion

Verfügungen werden hoheitlich – d. h. einseitig – durch die zuständige Behörde erlassen. Die betroffene Person soll die Erwägungen der Behörde nachvollziehen können und wissen, weshalb sich die Behörde so und nicht anders entschieden hat. Dadurch zeigt die Behörde, dass das **Anliegen der Person gewürdigt** wurde, was die **innere Zustimmung der betroffenen Person** herbeiführen soll. Dies leistet einen massgeblichen Beitrag dazu, dass die Entscheidung nicht an die nächste Instanz weitergezogen wird, und dient somit dem Rechtsfrieden sowie der Prozessökonomie.

Datenbasierte und regelbasierte Algorithmen: Unabhängig vom eingesetzten Modell ist bei algorithmenbasierten Verfügungen sicherzustellen, dass die be-

troffene Person die Entscheidung versteht und nachvollziehen kann, dass bzw. wie ihre Anliegen dabei gewürdigt wurden. Die Begründung muss für einen Laien bzw. eine Laiin verständlich abgefasst sein.

3. Rechtssicherheitsfunktion

Die Begründung muss ersichtlich machen, welche Tatsachen bzw. Sachverhaltselemente berücksichtigt wurden und unter welche Rechtsnormen die Behörde diese subsumiert. Dies macht **staatliches Handeln vorhersehbar** und fördert eine **einheitliche Praxis**.

Datenbasierte ML-Algorithmen: Da der Output von ML-Algorithmen nicht auf kausalen Beziehungen, sondern auf Korrelationen und ggf. Zufällen beruht, besteht diesbezüglich eine grosse Diskrepanz. Wenn nicht klar ist, dass gleiche Inputs immer zum gleichen Output führen, kann nicht von einer einheitlichen Praxis und damit auch nicht von Vorhersehbarkeit ausgegangen werden.

4. Selbstkontrollfunktion

Behörden haben sich in Erfüllung ihrer Aufgaben an die geltende Rechtsordnung zu halten und einzelfallgerecht zu entscheiden. Die Entscheidungsfindung kann durch subjektive (z. B. emotionale Einflüsse), aber auch objektive Faktoren (z. B. Wetterlage) beeinflusst werden. Das Verfassen der Begründung kann somit zur Selbstreflexion der die Verfügung ausstellenden Person führen, indem sie überprüft, ob sie sich von unsachlichen Motiven hat leiten lassen.

Datenbasierte und regelbasierte Algorithmen: Auch wenn menschliche Bearbeitungsschritte bei algorithmenbasierten Verfügungen ganz oder teilweise wegfallen, können sich gerade aufgrund der eingesetzten Daten oder Algorithmen Fehler ergeben. Über die Begründung soll auch in diesen Fällen – jedenfalls dann, wenn die algorithmenbasierte Verfügung im Rahmen eines teilautomatisierten Verfahrens ergeht – eine Selbstkontrolle vor Erlass der Verfügung ermöglicht werden. Dies bedingt, dass die entscheidungsrelevanten Faktoren des Systems erkennbar sind und durch einen Menschen nachvollzogen werden können. Die Begründung kann in diesem Fall auch als Mittel zum Zweck verstanden werden, indem z. B. allfällige Diskriminierungsprobleme sichtbar gemacht werden. Im vollautomatisierten Verfahren kann die Selbstkontrollfunktion hingegen gar nicht erfüllt werden.

5. Fremdkontrollfunktion

Die Begründung erlaubt einer allfälligen Rechtsmittelinstanz, die Erwägungen der verfügenden Behörde nachzuvollziehen, um sich ein sachgerechtes Bild von der Angelegenheit zu machen und prüfen zu können, ob die Vorinstanz rechtskonform entschieden hat. Dazu benötigt sie Angaben über das Verhältnis zwischen den angewendeten Rechtsnormen und Tatsachen sowie der erlassenen Verfügung.

Datenbasierte und regelbasierte Algorithmen: Die insbesondere bei der Rechtsmittel- und Akzeptanzfunktion angeführten Herausforderungen stellen sich auch bezüglich der Fremdkontrollfunktion.

V. Lösungsansätze

Die Begründungspflicht verlangt nach einer Erklärung, welche Sachverhaltsinformationen und Rechtsgrundlagen im konkreten Einzelfall berücksichtigt wurden und für das Ergebnis ausschlaggebend waren. Die folgenden Lösungsansätze können dabei helfen, die Begründungspflicht umzusetzen:

1. Kennzeichnungspflicht

Damit die betroffene Person Kenntnis darüber hat, dass ihr Anliegen durch ein algorithmenbasiertes Entscheidungssystem bearbeitet wurde, bedarf es einer entsprechenden Kennzeichnung der Entscheidung, etwa im Sinne der Kennzeichnungspflicht für automatisierte Einzelentscheide von Bundesorganen in Art. 21 Abs. 4 Datenschutzgesetz. Dies stellt zwar einen Baustein zur Erfüllung der Begründungspflicht dar, reicht allerdings für sich allein nicht aus.

2. Verweis auf weiterführende Informationen

Neben der Kennzeichnung können auch weitere Informationen zum Entscheidungssystem selbst geliefert werden. Dazu zählt etwa ein Verweis auf die Ergebnisse einer vorgängig durchgeführten Folgenabschätzung, die in einem öffentlich zugänglichen Verzeichnis über den Einsatz des algorithmischen Systems zu finden sind.

3. Wahl des Modells/der Technik

Es existieren ML-Modelle, die aufgrund ihres Designs (intrinsische Eigenschaft) nachvollziehbare Ergebnisse liefern können. So können Ergebnisse ba-

sierend auf (einfachen) Entscheidungsbäumen eher nachvollziehbar sein als jene basierend auf künstlichen neuronalen Netzen, da erkenntlich gemacht werden kann, welchen «Weg» das System eingeschlagen hat und welche Datenpunkte für die Entscheidung ausschlaggebend waren. Dementsprechend könnten für behördliche Entscheidungen nur jene ML-Techniken/Modelle erlaubt werden, die der Nachvollziehbarkeit zugänglich sind.

4. Menschliche Verantwortung

Von verschiedenen Seiten wird gefordert, dass im automatisierten Entscheidungsverfahren immer ein **Mensch in den Entscheidungsprozess involviert** sein muss: sei es bei der Entscheidungsunterstützung, bei der das System einem Sachbearbeiter bzw. einer Sachbearbeiterin Vorschläge zur Entscheidungsfindung unterbreitet, sei es bei der vollautomatisierten Entscheidungsfindung, bei der ein Mensch die durch das System gefällte Entscheidung – inklusive der Begründung – im Nachhinein überprüft. In beiden Konstellationen ist darauf zu achten, dass die entscheidende Person sich nicht durch Verzerrungen im System zu Fehlentscheidungen verleiten lässt bzw. Fehlentscheidungen billigt. In diesem Zusammenhang wird auch von **«Automation Bias»** gesprochen. Dabei misst der Mensch der Computerantwort mehr Bedeutung bei, als ihr eigentlich zukommen sollte. Um dies zu vermeiden, muss einerseits die Datenkompetenz der Behördenmitarbeitenden gefördert werden. Die sachbearbeitende Person muss grundsätzlich in der Lage sein, den Prozess, der zu einer Entscheidung führt, nachzuvollziehen und eine entsprechende Begründung zu formulieren. Andererseits sollen **retrospektive externe Analysen** durchgeführt werden. Denn ob sich jemand von unsachlichen Motiven oder Datenverzerrungen hat leiten lassen und dadurch (indirekte) Diskriminierungen entstanden sind, zeigt sich oft erst in der Gesamtschau der Entscheidungspraxis und nicht in einer einzelnen Verfügung. Damit erfährt die Selbstkontrollfunktion der Begründungspflicht in automatisierten Entscheidungsverfahren eine gewisse Kompensation.

5. Reduktion der Begründungsdichte?

In Verwaltungsbereichen, in denen eine Vielzahl an Entscheidungen bezüglich gleichartiger Fälle ergehen, stellen Lehre und Rechtsprechung reduzierte Anforderungen an die Begründungsdichte. Dies ist auf die weitgehende Standardisierung des Verfahrens zurückzuführen. Zentral ist dabei, dass gleiche Inputs stets zu gleichem Output führen. Die Kausalität des Entscheidungsprozesses ist gewährleistet. Für regelbasierte Algorithmen könnte diese Praxis

allenfalls analog angewendet werden, da sie ebenfalls standardisierte Abläufe ausführen. Allerdings müsste auf Verlangen der von der Entscheidung betroffenen Person trotzdem eine Begründung geliefert werden können. Die Behördenmitarbeitenden müssten also dennoch in der Lage sein, eine Begründung zu verfassen. Demgegenüber beruhen datenbasierte ML-Algorithmen regelmässig auf Korrelationen. Im Lichte der verfassungsrechtlichen Anforderungen und der bestehenden Praxis im Bereich der Massenverwaltung scheint es deshalb nicht gerechtfertigt, beim Einsatz datenbasierter ML-Algorithmen die Begründungsanforderungen zu reduzieren. Gegebenenfalls ist jedoch ein **Paradigmenwechsel** denkbar. An die Stelle der Nachvollziehbarkeit und Begründung der Entscheidung im Einzelfall könnte eine **retrospektive Systemkontrolle** und Analyse der **Entscheidungspraxis** treten (z.B. durch einfache Zufallsstichproben, «random sampling»). Die Akzeptanzfunktion kann dadurch allerdings nur kompensiert werden, wenn die Bevölkerung den automatisierten Verfahren vertraut. Auch ist ein solches Kontrollsystem ohnehin nur in nicht sensiblen Bereichen denkbar.

Die Begründung von Verfügungen beim Einsatz algorithmischer Systeme

Nadja Braun Binder / Liliane Obrecht

Erstabdruck in: SJZ 15/2024 15.8.2024

Der Einsatz algorithmischer Entscheidungssysteme beim Erlass von Verfügungen kann eine Herausforderung für die Begründungspflicht darstellen, die sich aus dem verfassungsrechtlichen Anspruch auf rechtliches Gehör ergibt. Je nach technischer Ausgestaltung des Systems ist das Zustandekommen der Verfügung nicht nachvollziehbar. Schwierigkeiten können sowohl bei der Verwendung von datenbasierten Machine Learning Algorithmen als auch bei der Verwendung von komplexen regelbasierten Algorithmen auftreten. Entsprechend kann der Einsatz solcher Systeme dazu führen, dass Verfügungen nicht oder nur mit zusätzlichen Vorkehrungen begründet werden können. Der vorliegende Beitrag skizziert unter Berücksichtigung der Funktionen der Begründung von Verfügungen ungeeignete und – zumindest teilweise – geeignete Lösungen.

L'utilisation de systèmes décisionnels algorithmiques lors de la prise de décisions peut représenter un défi pour l'obligation de motiver qui découle du droit constitutionnel d'être entendu. Selon la conception technique du système, l'élaboration de la décision peut ne pas être compréhensible. Des difficultés peuvent survenir lors de l'utilisation d'algorithmes d'apprentissage automatique basés sur des données (machine learning), mais aussi lors de l'utilisation d'algorithmes complexes basés sur des règles. En conséquence, l'utilisation de tels systèmes peut avoir pour conséquence que les décisions ne peuvent pas être motivées ou seulement avec des précautions supplémentaires. Le présent article esquisse des solutions inappropriées et des solutions au moins partiellement appropriées en tenant compte des fonctions de la motivation des décisions.

Inhalt

I.	Einleitung	68
II.	Algorithmenbasierte Verfügung	70
1.	Produkt eines algorithmischen Entscheidungssystems	70
2.	Regelbasierte und datenbasierte Algorithmen	72
2.1.	Regelbasierte Algorithmen	72
2.2.	Datenbasierte (Machine Learning) Algorithmen	73

III.	Funktionen der Begründungspflicht	74
1.	Vorbemerkung	74
2.	Massgeblichkeit der Begründungsdichte	75
3.	Die einzelnen Begründungsfunktionen	77
IV.	Herausforderungen für die Funktionen der Begründung beim Einsatz algorithmischer Entscheidungssysteme	79
1.	Fehlende menschliche Beteiligung im Entscheidungsverfahren	79
2.	Fehlende Vorhersehbarkeit	80
3.	Fehlende Nachvollziehbarkeit der Information	81
V.	Ungeeignete und teilweise geeignete Lösungen	81
1.	Keine Erfüllung der Begründungsfunktionen	82
1.1.	Erklärung der Funktionsweise des algorithmischen Systems	82
1.2.	Erklärung durch das System selbst oder ein externes System	83
1.3.	Offenlegung der Trainings- bzw. Validierungsdaten	84
2.	Teilweise Erfüllung der Begründungsfunktionen	85
2.1.	Kennzeichnungspflicht	85
2.2.	Weiterführende Informationen, insbesondere zur Folgenabschätzung	85
2.3.	Aufführen einzelner Datenpunkte	86
2.4.	Wahl des Modells und der Technik	87
2.5.	Menschliche Verantwortung	87
2.6.	Reduktion der Begründungsdichte	88
VI.	Fazit	89

I. Einleitung

Der Ruf nach der Herstellung von Transparenz über den Einsatz algorithmischer Entscheidungssysteme ist in den letzten Jahren immer lauter geworden.¹

¹ Vorliegend werden nur Schweizer Positionen berücksichtigt. Siehe zur Politik nachfolgend Fn. 3; siehe zur Zivilgesellschaft: *Angela Müller/Estelle Pannatier*, Automatisierte Entscheidungssysteme im öffentlichen Sektor – einige Empfehlungen, AlgorithmWatch CH vom 27.11.2023, <<https://algorithmwatch.ch/de/adm-offentlichersektor-empfehlungen/>> (zuletzt besucht am 25.6.2024); Wissenschaft: *Nadja Braun Binder et al.*, Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht, Jusletter vom 28.6.2021, Rz. 8 ff.; *Florent Thouvenin et al.*, Ein Rechtsrahmen für Künstliche Intelligenz, Positionspapier vom November 2021, 3 ff., <<https://www.dsi.uzh.ch/de/research/projects/strategy-lab/strategy-lab-21.html>> (zuletzt besucht am 25.6.2024); Bund: Siehe etwa *Bundesrat*, Leitlinien «Künstliche Intelligenz» für den Bund vom 25.11.2020, Orientierungsrahmen für den Umgang mit künstlicher Intelligenz in der Bundesverwaltung, <<https://www.sbf.admin.ch/sbf/de/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html>> (zuletzt besucht am 25.6.2024).

Transparenz weist dabei unterschiedliche Facetten auf. Neben allgemeiner, gesellschaftlich breit angelegter Transparenz im Sinne einer Offenlegungspflicht² wird auch die Schaffung von Nachvollziehbarkeit mit Blick auf algorithmenbasierte Einzelentscheidungen gefordert.³ Im Kontext hoheitlicher Tätigkeit ergibt sich der Anspruch auf Nachvollziehbarkeit einer algorithmenbasierten Entscheidung aus dem in Art. 29 Abs. 2 BV⁴ verankerten Anspruch auf rechtliches Gehör bzw. der daraus abgeleiteten Begründungspflicht staatlicher Entscheidungen.⁵ Art. 35 Abs. 1 VwVG⁶ bestätigt bzw. konkretisiert dies für Verfügungen von Bundesverwaltungsbehörden. Zwar kann von einer Begründung abgesehen werden, wenn die Behörde den Begehren der Partei(en) vollständig entspricht und keine Partei eine Begründung verlangt.⁷ Dies ändert aber nichts daran, dass die Behörde die Verfügung im Einzelfall begründen können muss.⁸ Problematisch ist dabei insbesondere, dass die Nachvollziehbarkeit eines Ergebnisses eines algorithmischen Entscheidungssystems abnimmt, je leistungsfähiger dieses ist.⁹

Der vorliegende Beitrag geht einerseits der Frage nach, weshalb die Nachvollziehbarkeit bei algorithmenbasierten Verfügungen erschwert ist und skizziert

- ² Siehe dazu *Nadja Braun Binder/Liliane Obrecht*, Transparenz über den staatlichen Einsatz algorithmischer Entscheidungssysteme – Rechtliche Erwägungen ausgehend von Forderungen nach einem öffentlichen Verzeichnis (im Erscheinen; zit. Öffentliches Verzeichnis).
- ³ Siehe etwa *Nationalrätin Min Li Marti*, Motion (23.3806) «Deklarationspflicht bei Anwendungen der künstlichen Intelligenz und automatisierten Entscheidungssystemen» vom 15.6.2024, <<https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20233806>> (zuletzt besucht am 25.6.2024); *Staatssekretariat für Bildung, Forschung und Innovation (SBFI)*, Herausforderungen der künstlichen Intelligenz, Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat vom 13.12.2019, <<https://www.sbf.admin.ch/sbf/de/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html>> (zuletzt besucht am 25.6.2024) 31.
- ⁴ Bundesverfassung der Schweizerischen Eidgenossenschaft (BV) vom 18. April 1999 (SR 101).
- ⁵ Die betroffene Person muss die Entscheidung verstehen (siehe *Gerold Steinmann/Benjamin Schindler/Damian Wyss*, in: Bernhard Ehrenzeller et al. [Hrsg.], *St. Galler Kommentar, Die schweizerische Bundesverfassung*, 4. A., Zürich/Genf/St. Gallen 2023, Art. 29 BV N 65) weshalb sie diese auch nachvollziehen können muss.
- ⁶ Bundesgesetz über das Verwaltungsverfahren (VwVG) vom 20. Dezember 1968 (SR 172.021).
- ⁷ Art. 35 Abs. 3 VwVG.
- ⁸ Die Behörde muss in der Lage sein, die Verfügung zu begründen, denn die Partei kann in jedem Fall eine Begründung verlangen (siehe *René Wiederkehr/Christian Meyer/Anna Böhme*, *Orell Füssli Kommentar, VwVG Kommentar, Bundesgesetz über das Verwaltungsverfahren und weiteren Erlassen*, Zürich 2022, Art. 35 N 39).
- ⁹ *Thomas Wischmeyer*, *Regierungs- und Verwaltungshandeln durch KI*, in: Martin Ebers et al. (Hrsg.), *Rechtshandbuch Künstliche Intelligenz und Robotik*, München 2020, 612 ff., 629.

andererseits einige, nicht abschliessend zu verstehende Lösungsansätze.¹⁰ Dazu werden zunächst der Begriff der algorithmenbasierten Verfügung (II) und sodann die Begründungspflicht (III) analysiert. Anschliessend werden die Auswirkungen des Einsatzes algorithmischer Entscheidungssysteme auf die Begründungsfunktionen behandelt (IV). Wie schliesslich aus rechtlicher Sicht mit der Herausforderung der fehlenden Nachvollziehbarkeit umgegangen werden könnte, wird mit einem Blick auf ungeeignete und – zumindest teilweise – geeignete Lösungen diskutiert (V).

II. Algorithmenbasierte Verfügung

1. Produkt eines algorithmischen Entscheidungssystems

Im vorliegenden Beitrag wird die Nachvollziehbarkeit von *Verfügungen* – Entscheidungen, die im nichtstreitigen, erstinstanzlichen Verwaltungsverfahren ergehen¹¹ – analysiert. Der Verfügungsbegriff ist in Art. 5 Abs. 1 VwVG geregelt. Es handelt sich demnach um eine hoheitliche, individuell-konkrete Anordnung einer Behörde, mit der im Einzelfall ein Rechtsverhältnis zwischen Behörde und betroffener Person geregelt wird.¹²

Ein *Algorithmus* ist eine eindeutige Handlungsvorschrift an einen Computer zur Lösung eines Problems oder einer Klasse von Problemen.¹³ Werden viele Algorithmen zu dieser Problemlösung zusammengesetzt, liegt ein algorithmisches (Entscheidungs-)System vor.¹⁴ Der Begriff des algorithmischen Ent-

¹⁰ Dass algorithmenbasierte Verwaltungsverfahren künftig vermehrt durchgeführt werden, steht ausser Frage, die konkrete rechtliche Umsetzung erweist sich allerdings noch als lückenhaft (siehe René Wiederkehr/Christian Meyer, Schranken, Ausnahmen und Relativierungen des rechtlichen Gehörs, Insbesondere mit Blick auf automatisierte Verwaltungsverfahren, AJP 2022, 1092 ff., 1102).

¹¹ Regina Kiener/Bernhard Rütscbe/Mathias Kuhn, Öffentliches Verfahrensrecht, 3. A., Zürich/St. Gallen 2021, Rz. 313.

¹² Felix Uhlmann/Matthias Kradolfer, in: Bernhard Waldmann/Patrick L. Krauskopf (Hrsg.), Praxiskommentar Verwaltungsverfahrensgesetz (VwVG), unter Einschluss des Reglements über die Kosten und Entschädigungen vor dem Bundesverwaltungsgericht (VGKE), 3. A., Zürich/Genf 2023, Art. 5 VwVG N 16 ff.

¹³ Hartmut Ernst/Jochen Schmidt/Gerd Beneken, Grundkurs Informatik, Grundlagen und Konzepte für die erfolgreiche IT-Praxis – Eine umfassende, praxisorientierte Einführung, 7. A., Wiesbaden 2020, 501 f.

¹⁴ Siehe Katharina A. Zweig, Algorithmische Entscheidungen: Transparenz und Kontrolle, Januar 2019, 4f., <<https://d-nb.info/1190558548/34>> (zuletzt besucht am 25.6.2024), für da-

scheidungssystems zielt nicht auf eine bestimmte Methode oder Technologie ab, sondern erfasst alle Anwendungen, die auf (Computer-)Algorithmen basieren.¹⁵ Für die Anliegen des vorliegenden Beitrags passt der Begriff besser als der häufig verwendete Begriff des KI-Systems (KI steht für «Künstliche Intelligenz»),¹⁶ da er auch weitere Technologien umfasst, die dieselben Auswirkungen auf Individuen und Gesellschaft haben wie KI-Systeme.

Wird eine Verfügung mithilfe eines algorithmischen Entscheidungssystems vorbereitet oder erlassen, wird diese im vorliegenden Beitrag als *algorithmienbasierte Verfügung* bezeichnet.¹⁷ Da es im vorliegenden Beitrag um die Frage der Nachvollziehbarkeit bzw. der Umsetzung der Begründungspflicht geht, werden sowohl teil- als auch vollautomatisiert erlassene Verfügungen betrachtet. Die Schwierigkeiten in der Umsetzung der Begründungspflicht ergeben sich aus dem Umstand, dass ein algorithmisches Entscheidungssystem genutzt wird.¹⁸ Diesem Umstand soll durch die Verwendung des Begriffs

tenbasierte *Machine-Learning*-Algorithmen. Vorliegend geht es auch um regelbasierte Algorithmen. Zur Unterscheidung [II.2](#).

¹⁵ Thoubenin et al. (Fn. 1) 1.

¹⁶ Siehe insbesondere die Definitionen, die der EU-Gesetzgebung zur Regulierung von KI (siehe hinten [V](#)) und der Konvention zur Regulierung von KI des Europarats (Art. 2 des Rahmenübereinkommens des Europarates über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit vom 17. Mai 2024) zugrunde liegen.

¹⁷ Die Terminologie in der Literatur ist nicht einheitlich: Erdem Büyüksagis, *Décision algorithmiques: mieux vaut responsabiliser qu'informer*, HAVE 2020 225 ff., nutzt den Begriff der algorithmischen Entscheidung. Verschiedentlich wird – in Anlehnung an die Terminologie des Datenschutzrechts (siehe Nadja Braun Binder, *Automatisierte Entscheidungen: Perspektive Datenschutzrecht und öffentliche Verwaltung*, SZW 2020 27 ff. [zit. Perspektive Datenschutzrecht]) – der Begriff der automatisierten Einzelentscheidung verwendet (siehe z.B. Nadja Braun Binder, *Künstliche Intelligenz und automatisierte Entscheidungen in der öffentlichen Verwaltung*, SJZ 2019 467 ff. [zit. KI]; David Rechsteiner, *Der Algorithmus verfügt*, Verfassungs- und verwaltungsrechtliche Aspekte automatisierter Einzelentscheidungen, Jusletter vom 26.11.2018, Rz. 5 ff.). Im Kontext des Datenschutzrechts sind damit aktuell lediglich vollautomatisierte Entscheidungen gemeint (siehe Liliane Obrecht, *Verfügung und automatisierte Einzelentscheidung – same same but different?, ex/ante 2/2022 38 ff., 39 f.*). Die Rede ist bisweilen auch vom automatisierten Verwaltungsentscheid (siehe Daniela Thurnherr, *Automatisierte Verwaltungsverfahren auf Bundesebene [Schweiz]*, in: Nadja Braun Binder/Peter Bussjäger/Mathias Eller [Hrsg.], *Auswirkungen der Digitalisierung auf die Erlassung und Zuordnung behördlicher Entscheidungen*, Wien 2021, 133 ff., 152) oder von der automatisierten Verfügung (siehe Nadja Braun Binder, *Als Verfügungen gelten Anordnungen der Maschinen im Einzelfall... – Dystopie oder künftiger Verwaltungsalltag?*, ZSR 2020 253 ff., 256 f. [zit. Dystopie]).

¹⁸ Siehe Nadja Braun Binder, *Staat, Mensch, Algorithmen*, Zur Frage der Entbehrlichkeit menschlicher Bearbeitung im digitalisierten Staat, BJM 2023 2 ff., 19; Rechsteiner (Fn. 17) Rz. 27.

der algorithmenbasierten Verfügung Rechnung getragen werden. Eine algorithmenbasierte Verfügung kann demnach vollständig durch das Entscheidungssystem erlassen werden, ohne oder nur mit teilweiser menschlicher Beteiligung am Entscheidungsprozess, etwa wenn das Entscheidungssystem unterstützend oder vorbereitend eingesetzt wird.

2. Regelbasierte und datenbasierte Algorithmen

Algorithmischen Entscheidungssystemen können verschiedene Arten von Algorithmen zugrunde liegen (II.2.1 und II.2.2). Deren Einsatz hat unterschiedliche Auswirkungen auf die vorliegend interessierende Begründungspflicht von Verfügungen.

2.1. Regelbasierte Algorithmen

Regelbasierte Algorithmen verarbeiten eine Eingabe nach durch natürliche Personen festgelegten Regeln, um zu einem Ergebnis zu gelangen.¹⁹ Sie sind deterministisch, d.h. sie verarbeiten eine bestimmte Eingabe immer nach denselben Abläufen.²⁰ Die Algorithmen sind dabei auch determiniert, eine bestimmte Eingabe erzeugt immer die gleiche Ausgabe.²¹ Dies führt dazu, dass die Reihenfolge der einzelnen Schritte zu jedem Zeitpunkt bestimmt ist, was grundsätzlich bewirkt, dass das Zustandekommen des Ergebnisses nachvollziehbar ist. Dies bedeutet allerdings nicht, dass die betroffene Person die Entscheidung ohne Weiteres nachvollziehen kann. Dies gilt insbesondere dann, wenn dem algorithmischen Entscheidungssystem komplexe Algorithmen zugrunde liegen, wie etwa vielfach verzweigte Entscheidungsbäume.²² Die Begründung solcher algorithmenbasierter Entscheidungen kann entsprechend herausfordernd sein.

¹⁹ Basanta E. P. Thapa/Peter Parycek, Data Analytics in Politik und Verwaltung, in: Resa Mohabbat-Kar/Basanta Thapa/Peter Parycek (Hrsg.), (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft, Berlin 2018, 40 ff., 61 f.

²⁰ Nach klaren «Wenn-Dann-Schemata» (siehe Ernst/Schmidt/Beneken [Fn. 13] 501 f.).

²¹ Bei Stefan Kirn/Claus D. Müller-Hengstberg, Intelligente (Software-)Agenten: Von der Automatisierung zur Autonomie? Verselbständigung technischer Systeme, MultiMedia und Recht (MMR) 2015 226 ff., 226, als «Automatisierung» beschrieben.

²² Jonathan Dollinger, Folgenabschätzungen für Verwaltungs-Algorithmen, Tübingen 2023, 20 f.

2.2. Datenbasierte (Machine Learning) Algorithmen

Algorithmen, die selbständig lernen und nicht wie regelbasierte Algorithmen durch den Menschen vorprogrammiert werden, werfen aus rechtsstaatlicher Perspektive besondere Probleme auf und stehen daher im Mittelpunkt entsprechender Diskussionen.²³ Solche datenbasierten *Machine Learning* (ML) Algorithmen verarbeiten eine Eingabe nicht anhand vorgegebener Regeln, sondern analysieren i.d.R. enorme Datensätze (Trainingsdaten) auf Korrelationen, um selbst eine verallgemeinerbare Regel aufzustellen.²⁴ Korrelation bedeutet dabei, dass zwei Phänomene zwar statistisch miteinander zusammenhängen,²⁵ was allerdings nichts über die Ursache-Wirkungs-Beziehung (Kausalität) aussagt.²⁶ Das Problem verstärkt sich, wenn zusätzlich Zufälle den Prozess prägen.²⁷ Solche Algorithmen sind also nicht deterministisch, sie verarbeiten eine Eingabe nicht immer nach den gleichen Abläufen, sondern sie sind wahrrscheinlichkeitsbasiert.²⁸ Dies führt dazu, dass die gleiche Eingabe nicht immer zuverlässig dasselbe Resultat erzeugt – der Algorithmus ist mithin nicht determiniert.²⁹

Das Zustandekommen der Ergebnisse von datenbasierten ML-Algorithmen ist somit oft – je nach Komplexität der angewendeten Technologie – selbst für Expertinnen und Experten nicht nachvollziehbar.³⁰ Dies betrifft etwa künstli-

²³ So etwa in Deutschland (Auswahl): David Roth-Isigkeit, Die Begründung des vollständig automatisierten Verwaltungsakts, Die Öffentliche Verwaltung (DÖV) 2020 1018 ff., 1019 ff.; Wischmeyer (Fn. 9) 629 ff.

²⁴ Alfred Früh/Dario Haux, Foundations of Artificial Intelligence and Machine Learning, Berlin 2020, 4 f., <<https://www.ssoar.info/ssoar/handle/document/80203>> (zuletzt besucht am 25.6.2024).

²⁵ Definition Duden online zum Begriff «Korrelation» (zweiter mathematischer Begriff), <<https://www.duden.de/rechtschreibung/Korrelation>> (zuletzt besucht am 25.6.2024).

²⁶ Zur Unterscheidung von Kausalität und Korrelation siehe Moreen Heine et al., Künstliche Intelligenz in öffentlichen Verwaltungen, Grundlagen, Chancen, Herausforderungen und Einsatzszenarien, Wiesbaden 2023, 16 f.

²⁷ Nadja Braun Binder et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4, Zürich 2021, 41 (zit. Studie KI 2021), <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> (zuletzt besucht am 25.6.2024).

²⁸ Ernst/Schmidt/Beneken (Fn. 13) 801 ff.

²⁹ Ernst/Schmidt/Beneken, (Fn. 13) 501 f.

³⁰ Markus Christen et al., Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz, in: TA-SWISS, Zürich 2020, 294 (zit. TA-SWISS Studie KI 2020); Ma-

che neuronale Netze (KNN), eine Klasse von ML-Verfahren, die aktuell besonders viel Aufmerksamkeit auf sich zieht und auch in der öffentlichen Verwaltung als vielfältig nutzbar gehandelt wird.³¹ Aus rechtsstaatlicher Perspektive ist jedoch nebensächlich, welche spezifische Technologie eingesetzt wird.³² Entscheidender ist, dass durch die Verwendung von maschinellen Lernmethoden ein Perspektivenwechsel erfolgt: Das den Rechtswissenschaften inhärente Konstrukt der Kausalität wird gewissermassen durch Wahrscheinlichkeiten abgelöst. Dies ist nicht grundsätzlich mit schlechteren Ergebnissen verbunden. Im Gegenteil, die Ergebnisse können äusserst präzise sein. Allerdings stellt auch in solchen Fällen die Herstellung der Nachvollziehbarkeit der algorithmusbasierten Entscheidung im Rahmen ihrer Begründung eine besondere Herausforderung dar.

III. Funktionen der Begründungspflicht

1. Vorbemerkung

Menschliche Entscheidungsträgerinnen und -träger sind an rechtlich klar umschriebene Kriterien gebunden, die beim Erlass einer Verfügung zu berücksichtigen sind.³³ Es ist jedoch nicht möglich, im Einzelnen nachzuvollziehen (oder gar nachzuzeichnen), welche Denkschritte die Menschen bei ihrer Entscheidungsfindung durchlaufen haben.³⁴ Das Zustandekommen der Verfügung unterscheidet sich *insofern* nicht vom Zustandekommen einer datenbasierten ML-Verfügung.³⁵ Bei regelbasierten Algorithmen ist es sogar so, dass diese bei ihrer Entscheidungsfindung nur jene Kriterien berücksichtigen, die durch die vorherige Programmierung zugelassen wurden, sie haben keinen autonomen

rio Martini, *Blackbox Algorithmus, Grundfragen einer Regulierung Künstlicher Intelligenz*, Berlin 2019, 28 ff.

³¹ Zum Begriff Michael Oettinger, *Data Science, Eine praxisorientierte Einführung im Umfeld von Machine Learning, künstlicher Intelligenz und Big Data*, 2. A., Hamburg 2020, 119 ff.; zu Einsatzbereichen siehe Heine et al. (Fn. 26) 3 f. und 27 f.

³² Nadja Braun Binder/Liliane Obrecht, *Algorithmisch überprüfte Steuererklärung im ordentlichen gemischten Veranlagungsverfahren*, zsis 2023, Rz. 26, in Bezug auf Risikomanagementsysteme in der Steuerverwaltung.

³³ Christian Ernst, *Algorithmische Entscheidungsfindung und personenbezogene Daten*, *JuristenZeitung (JZ)* 2017 1026 ff., 1027.

³⁴ Wischmeyer (Fn. 9) 631, etwa beschreibt dieses Phänomen so: «[...] auch Menschen sind für andere Menschen black boxes».

³⁵ Siehe [II.2.2.](#)

Entscheidungsspielraum.³⁶ Solche Entscheidungen könnten diesbezüglich sogar wesentlich nachvollziehbarer sein als menschliche Entscheide.³⁷

Allerdings verfolgt die Begründung von Verfügungen nicht das Ziel, die einzelnen Denk- oder Programmschritte vollständig nachzuzeichnen. Vielmehr liefert die Begründung die Erklärung für einen *bereits gefällten* Entscheid nach.³⁸ Begründungen stellen somit nicht das reale Zustandekommen von Entscheiden – im Sinne einer Erklärung der neuronalen Denkabläufe im menschlichen Gehirn – dar, sondern liefern der betroffenen Person in verständlicher Weise die Gründe für die getroffene Entscheidung. In diesem Punkt bestehen zwischen menschlichen und algorithmischen Entscheiden keine grundlegenden Unterschiede.³⁹

2. Massgeblichkeit der Begründungsdichte

Die von der Entscheidung betroffene Person hat Anspruch auf eine inhaltlich und umfangmässig *angemessene* Begründung. Was angemessen ist, kann allerdings nur teilweise abstrakt geregelt werden.⁴⁰ Die Angemessenheit kann letztlich nur anhand einer Interessenabwägung des konkreten Einzelfalls beurteilt werden.⁴¹ Dennoch haben Lehre und Rechtsprechung Kriterien entwickelt, die zur Bestimmung der Begründungsdichte herangezogen werden können.⁴² Je stärker die Entscheidung Rechte und Pflichten der betroffenen Person berührt, desto detaillierter muss die Begründung sein (*Intensität des*

³⁶ Siehe [II.2.1](#).

³⁷ Ernst (Fn. 33) 1027; Rechsteiner (Fn. 17) Rz. 26.

³⁸ Gemäss Thomas Kürsteiner, Künstliche Intelligenz und Rechtsprechung diskutiert am Beispiel der Begründungspflicht, *Justice – Justiz – Giustizia* 1/2024, Rz. 49, handelt es sich bspw. bei der Begründung von Gerichtsurteilen um eine «ex-post nachgeschobene Begründung eines Resultats».

³⁹ So auch Kürsteiner (Fn. 38) Rz. 49, betreffend Begründungen von Gerichtsurteilen.

⁴⁰ OFK VwVG-Wiederkehr/Meyer/Böhme (Fn. 8) Art. 29 VwVG N 16.

⁴¹ BGE 112 Ia 107 E. 2b; Michele Albertini, Der verfassungsmässige Anspruch auf rechtliches Gehör im Verwaltungsverfahren des modernen Staates, eine Untersuchung über Sinn und Gehalt der Garantie unter besonderer Berücksichtigung der bundesgerichtlichen Rechtsprechung, Bern 2000, 405.

⁴² Angelehnt an René Wiederkehr, Die Begründungspflicht nach Art. 29 Abs. 2 BV und die Heilung bei Verletzung, ZBl 2010 481 ff.; Albertini (Fn. 41) 205 ff.; Lorenz Kneubühler, Die Begründungspflicht, Eine Untersuchung über die Pflicht der Behörden zur Begründung ihrer Entscheide, Bern/Stuttgart/Wien 1998, 178 ff.

Eingriffs).⁴³ Je komplexer die Sach- und Rechtslage, desto schwieriger ist es für die betroffene Person, die Subsumtion (Unterordnung) des Sachverhalts unter die anwendbaren Normen und damit die behördliche Entscheidung zu verstehen (*Komplexität der Sach- und Rechtslage*). Eine gesteigerte Begründungsdichte kann dies kompensieren. Obliegt der entscheidenden Behörde kein Entscheidungsspielraum, genügt regelmässig ein Verweis auf die angewendeten Rechtsnormen. Je grösser allerdings der Entscheidungsspielraum, desto detaillierter ist die Entscheidung zu begründen (*Entscheidungsspielraum der Behörde*).⁴⁴ Weicht die Behörde in einem Fall von einer konstanten Praxis der Gesetzesanwendung ab, bedarf es einer detaillierten Begründung (*Abweichung von einer behördlichen Praxis*). Schliesslich sind in Verwaltungsbereichen, in denen eine Vielzahl an gleichgelagerten Verfügungen ergehen (*Verfügungen der Massenverwaltung*), im Sinne der Prozessökonomie geringere Anforderungen an die Begründungsdichte zulässig.⁴⁵

Fest steht daher bereits, dass die Anforderungen an die Begründungsdichte nicht reduziert werden können, nur weil ein algorithmisches Entscheidungssystem im Einsatz ist.⁴⁶ Zusammenfassend müssen «wenigstens kurz die Überlegungen genannt werden, von denen sich die Behörde hat leiten lassen und auf welche sich ihr Entscheid stützt»⁴⁷. Die für den Entscheid wesentlichen Aspekte des Sachverhalts und deren rechtliche Würdigung müssen stets dargelegt werden.⁴⁸

⁴³ BGE 133 I 270 E. 3.1. Beispielhaft kann das Diskriminierungsverbot gemäss Art. 8 Abs. 2 BV angeführt werden: Ungleiche Behandlungen bedürfen gemäss BGE 129 I 232 E. 3.4 einer besonders qualifizierten Begründung.

⁴⁴ BGE 112 Ia 107 E. 2b; BGE 104 Ia 201 E. 5g; OFK VwVG-Wiederkehr/Meyer/Böhme (Fn. 8) Art. 29 VwVG N 22 und N 25, die auf die fehlende Voraussehbarkeit der Rechtsfolgen und den damit verbundenen Verlust an Rechtssicherheit hinweisen; gemäss BGE 127 V 431 E. 2b/cc fungiert die Begründung in diesem Fall als Kompensationsmechanismus.

⁴⁵ Insbesondere ist die Verwendung von Textbausteinen zulässig und auch üblich (siehe Lorenz Kneubühler/Ramona Pedretti, in: Christoph Auer/Markus Müller/Benjamin Schindler [Hrsg.], VwVG, Bundesgesetz über das Verwaltungsverfahren, Kommentar, Zürich/St. Gallen 2019, Art. 35 VwVG N 18).

⁴⁶ Auch Wiederkehr/Meyer (Fn. 10) 1107; Braun Binder (Fn. 17) KI, 473.

⁴⁷ BGE 133 I 270 E. 3.1.

⁴⁸ In konstanter Rechtsprechung (Auswahl) BGE 146 II 335 E. 5.1; BGE 142 I 135 E. 2.1; BGE 138 I 232 E. 5.1; BGE 133 II 439 E. 3.3. Im Gesuchsverfahren etwa muss sich die Behörde in der Begründung nicht mit allen Parteivorbringen detailliert auseinandersetzen und sie muss nicht zwingend alle Überlegungen bis ins Detail ausführen.

3. Die einzelnen Begründungsfunktionen

Die Begründung einer Verfügung erfüllt mehrere Funktionen.⁴⁹ Die *Rechtsmittelfunktion* verlangt, dass die Begründung der betroffenen Person eine sachgerechte Anfechtung der Verfügung ermöglicht.⁵⁰ Insbesondere muss in verständlicher Weise aus ihr hervorgehen, welche tatsächlichen und rechtlichen Erwägungen zur Entscheidung geführt haben.⁵¹ Es muss somit Kausalität zwischen der rechtlichen Handlung und der erlassenen Verfügung bestehen. Da Verfügungen hoheitlich, d.h. einseitig von der zuständigen Behörde erlassen werden, verlangt die Begründungspflicht im Sinne der *Akzeptanzfunktion*, dass die betroffene Person die Erwägungen der Behörde nachvollziehen kann.⁵² Setzt sich die Behörde mit der Entscheidung auseinander und zeigt sie insbesondere auch, weshalb sie sich nicht anders entschieden hat, lässt sie erkennen, dass das Anliegen der Person gewürdigt wurde und ihre allfälligen Vorbringen in das Ergebnis miteingeflossen sind.⁵³ Dadurch kann die innere Zustimmung der betroffenen Person herbeigeführt werden (*subjektive Komponente* der Akzeptanzfunktion).⁵⁴ Dies leistet einen massgeblichen Beitrag dazu, dass die betroffene Person die Entscheidung nicht an die nächste Instanz weiterzieht und dient somit dem Rechtsfrieden sowie der Prozessökonomie.⁵⁵

⁴⁹ Die folgenden Ausführungen sind angelehnt an *Kneubühler* (Fn. 42) 94 ff. (Selbstkontroll-, Befriedungs-, Rechtsmittel-, Fremdkontroll-, Effizienz- und Rechtssicherheitsfunktion); *Mark Eugen Villiger*, Die Pflicht zur Begründung von Verfügungen, ZBl 1989 137 ff., 159 ff. (Inhaltliche Präzisierung der Verfügung, Bestätigung der Gewährung des rechtlichen Gehörs, Förderung der Akzeptanz, Ermöglichung eines Rechtsmittels, Prüfung der Rechtmässigkeit der Verfügung, Transparenz des Verwaltungshandelns, Ökonomie des Verwaltungshandelns); zwar explizit bzgl. Gerichtsentscheidungen, die aber diesbezüglich mit Verfügungen vergleichbar seien, *Hans-Jakob Mosimann*, *Entscheidbegründung, Begründung und Redaktion von Gerichtsurteilen und Verfügungen*, Zürich/St. Gallen 2013, 1 und 13 ff. (Rechtsfrieden, Richtigkeitsnachweis, Adressaten, Akzeptanz, Orientierung für künftiges Handeln, Recht erfahrbar machen, Selbst- und Fremdkontrolle, Weiterzug, Rechtsstaatlichkeit).

⁵⁰ Diese Funktion ist insbesondere gemäss Bundesgericht in konstanter Rechtsprechung die zentrale Funktion der Begründung (siehe etwa BGE 143 III 65 E. 5.2; BGE 141 III 28 E. 3.2.4). Kritik daran verübt *Villiger* (Fn. 49) 163 f.

⁵¹ OFK VwVG–Wiederkehr/Meyer/Böhme (Fn. 8) Art. 35 VwVG N 8.

⁵² Siehe ausführlich zur menschlichen Beteiligung an Entscheidungen und der damit einhergehenden Förderung der Akzeptanz der Entscheidung durch die betroffene Person *Dollinger* (Fn. 22) 33 f.

⁵³ OFK VwVG–Wiederkehr/Meyer/Böhme (Fn. 8) Art. 35 VwVG N 8.

⁵⁴ *Mosimann* (Fn. 49) 18.

⁵⁵ Die Akzeptanzfunktion kann daher auch als Gegenstück zur Rechtsmittelfunktion betrachtet werden.

Ferner kann die Behörde belegen, dass sie das Recht korrekt angewendet hat und damit darlegen, dass das Verfahren im Sinne des Rechtsstaates abgelaufen ist (*objektive Komponente der Akzeptanzfunktion*).⁵⁶

Um die *Rechtssicherheitsfunktion* zu erfüllen, muss die Begründung darüber hinaus erkennen lassen, auf welche Tatsachen und Rechtsnormen die Behörde den konkreten Sachverhalt stützt und welche sie für die Verfügung als relevant erachtet.⁵⁷ Dies macht staatliches Handeln vorhersehbar und fördert eine einheitliche Praxis.⁵⁸ Schliesslich kommen der Begründung einer Verfügung zwei Kontrollfunktionen zu. Gemäss der *Selbstkontrollfunktion* haben sich Behörden in Erfüllung ihrer verfassungsmässigen oder gesetzlichen Aufgaben an die geltende Rechtsordnung zu halten und darauf basierend rechtskonform und einzelfallgerecht zu entscheiden.⁵⁹ Insbesondere sind Behörden an die Grundrechte gebunden,⁶⁰ etwa das Diskriminierungsverbot nach Art. 8 Abs. 2 BV. Die Auslegungs- und Subsumtionsarbeit ist ferner fehleranfällig und kann von verschiedenen Faktoren abhängig sein.⁶¹ Die Entscheidungsfindung kann etwa durch subjektive Einflüsse (z.B. Emotionen), aber auch objektive Faktoren (z.B. Wetterlage) beeinflusst werden.⁶² Das Verfassen der Begründung kann somit zur Selbstreflexion der die Verfügung ausstellenden Person führen, indem sie überprüft, ob sie sich von unsachlichen Motiven hat leiten lassen.⁶³ Wird eine Verfügung angefochten, erlaubt die Begründung der Rechtsmittelinstanz

⁵⁶ Mosimann (Fn. 49) 18.

⁵⁷ BGE 142 I 135 E. 2.1; BGE 138 I 232 E. 5.1; OFK VwVG-Wiederkehr/Meyer/Böhme (Fn. 8) Art. 35 VwVG N 8, bezeichnen dies als «Transparenzfunktion».

⁵⁸ Durch die Begründung wird das Legalitätsprinzip, das der Berechenbarkeit staatlichen Handelns dient (*Benjamin Schindler*, in: Bernhard Ehrenzeller et al. [Hrsg.], St. Galler Kommentar, Die schweizerische Bundesverfassung, 4. A., Zürich/Genf/St. Gallen 2023, Art. 5 BV N 21 ff.), gewissermassen auf den Einzelfall heruntergebrochen.

⁵⁹ Gemäss *Kneubühler* (Fn. 42) 96, wird durch die Begründung der Rechtsanwendungsvorgang «rationalisiert» bzw. gemäss BGE 112 Ia 107 E. 2b das Verheimlichen sachfremder Motive erschwert.

⁶⁰ Art. 35 Abs. 2 BV.

⁶¹ *Jacques Meylan*, La motivation des actes administratifs en droit Suisse, in: Recueil de travaux suisses présentés au VIIIe Congrès international de droit comparé, Basel 1970, 313 ff., 314.

⁶² Zur menschlichen Entscheidungsfindung *David Nink*, Justiz und Algorithmen, Über die Schwächen menschlicher Entscheidungsfindung und die Möglichkeiten neuer Technologien in der Rechtsprechung, Berlin 2021, 28 ff., insb. 40 ff.

⁶³ Gemäss Bundesgericht geht es bei der Selbstkontrollfunktion um das Verhindern des Bezugs sachfremder Motive (BGE 112 Ia 107 E. 2b). Unsachlich sind etwa Einflüsse wie Sympathie oder Antipathie (siehe dazu *Angelika Bernhard*, Die Entscheidungsbegründung im schweizerischen zivilgerichtlichen Verfahren, Zürich 1983, 47).

schliesslich, die Erwägungen der verfügenden Behörde nachzuvollziehen, um sich ein sachgerechtes Bild von der Angelegenheit zu machen und prüfen zu können, ob die Vorinstanz rechtskonform entschieden hat.⁶⁴ Der Begründung kommt in diesem Sinne eine *Fremdkontrollfunktion* zu.

IV. Herausforderungen für die Funktionen der Begründung beim Einsatz algorithmischer Entscheidungssysteme

Der Einsatz der verschiedenen Algorithmen (regelbasiert/datenbasiert)⁶⁵ wirkt sich unterschiedlich auf die im vorangehenden Kapitel (III) dargelegten Begründungsfunktionen aus.

1. Fehlende menschliche Beteiligung im Entscheidungsverfahren

Die zunehmende Digitalisierung und die damit einhergehende Anonymisierung des Verwaltungshandelns führen zu einem Rückgang der *menschlichen Komponente* im Entscheidungsprozess. Dies wirkt sich in zweierlei Hinsicht auf die Begründungsfunktionen aus. Erstens sind die verschiedenen Automatisierungsstufen zu beleuchten. So kann etwa die *Selbstkontrollfunktion* bei vollständig menschenunabhängigen algorithmenbasierten Verfügungen in keiner Weise erfüllt werden. Bei teilweise menschenunabhängigen Verfahren, etwa wenn das algorithmische Entscheidungssystem dem Menschen eine Empfehlung unterbreitet, ist dies nicht so eindeutig. Es ist jedoch denkbar, dass die Entscheidungsträgerin bzw. der Entscheidungsträger den computergenerierten Vorschlägen ein unangemessen hohes Mass an Vertrauenswürdigkeit oder gar letztgültiger Richtigkeit beimisst.⁶⁶ In diesem Zusammenhang spricht man auch von *Automation Bias*.⁶⁷ Die Funktion der *Selbstkontrolle* stösst also in bei-

⁶⁴ BVGer 2013/46 vom 3.12.2013 E. 6.2.5.

⁶⁵ Zur Unterscheidung siehe [II.2](#).

⁶⁶ Kürsteiner (Fn. 38), Rz. 56, sieht in der menschlichen Kontrolle eine Chance, potenziell fehlerhafte algorithmische Entscheidungen zu korrigieren, ohne die damit einhergehenden Risiken zu nennen.

⁶⁷ Der *Automation Bias* ist ein verhaltenspsychologisches Phänomen, bei dem Menschen wider besseren Wissens auf Vorschläge und Entscheidungen von Maschinen vertrauen. Siehe zum Begriff im Kontext der öffentlichen Verwaltung etwa Hannah Ruschemeier, *The Problems of the Automation Bias in the Public Sector – A Legal Perspective*, Weizenbaum Conference proceedings 2023, 2; Saar Alon-Barkat/Madalina Busuic, *Human-AI Interactions in Public Sector*

den Automatisierungsstufen an Grenzen. Zweitens wird die *Akzeptanzfunktion* durch das Fehlen menschlichen Handelns vor Herausforderungen gestellt. Viele Betroffene akzeptieren eine Verfügung eher, wenn ihr Fall von einer Person bearbeitet wurde, die sie als verständnisvoll und gerecht empfinden.⁶⁸ Zudem kann die digitale Transformation der Verwaltung generell auf Ablehnung stossen.⁶⁹ Trifft das algorithmische Entscheidungssystem fehlerhafte Entscheidungen, wird diese ablehnende Haltung eher noch verstärkt.⁷⁰

2. Fehlende Vorhersehbarkeit

Staatliches Handeln muss für die Bevölkerung vorhersehbar sein und eine einheitliche Praxis gewährleisten. Dies ist ein grundlegendes rechtsstaatliches Gebot. Im konkreten Einzelfall bedeutet dies, dass dargelegt werden können muss, welche tatsächlichen Gegebenheiten und Rechtsgrundlagen für das Ergebnis des algorithmischen Entscheidungssystems im Einzelfall massgeblich waren (Kausalität).⁷¹ Je nach Art des dem System zugrunde liegenden Algorithmus ist dies jedoch nicht möglich.⁷² Einfache Korrelationen, auch wenn sie zu einem plausiblen Ergebnis führen, genügen den Begründungsanforderungen nicht.⁷³ Dies gilt insbesondere für Systeme, die auf datenbasierten ML-Algorithmen beruhen.⁷⁴ Dadurch wird einerseits die *Rechtsmittelfunktion*, andererseits aber auch die *Rechtssicherheitsfunktion* erheblich beeinträchtigt. Denn wenn gleiche Inputs nicht immer zum gleichen Output führen, kann nicht von einer einheitlichen Praxis ausgegangen werden und die *Rechtssicherheitsfunktion* ist nicht gewährleistet.

Decision Making: «Automation Bias» and «Selective Adherence» to Algorithmic Advice, *Journal of Public Administration Research and Theory (JPART)* 2023 153 ff., 155.

⁶⁸ Annette Guckelberger, *Öffentliche Verwaltung im Zeitalter der Digitalisierung, Analysen und Strategien zur Verbesserung des E-Governments aus rechtlicher Sicht*, Baden-Baden 2019, Rz. 539.

⁶⁹ Annette Guckelberger, *Verwaltung im Zeitalter «4.0»*, in: Walter Frenz (Hrsg.), *Handbuch Industrie 4.0: Recht, Technik, Gesellschaft*, Berlin/Heidelberg 2020, 404 ff., 421.

⁷⁰ Jörn von Luke, *Smart Government auf einem schmalen Grat*, in: Resa Mohabbat Kar/Basanta E. P. Thapa/Peter Parycek (Hrsg.), *(Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft*, Berlin 2018, 97 ff., 116 f.

⁷¹ Siehe [II.2](#) und [III.3](#).

⁷² Siehe [II.2](#).

⁷³ Catherine Reiter, *Künstliche Intelligenz im Verwaltungsverfahren, Ermessen als Stolperstein?*, *AJP* 2022 984 ff., 992; *Wiederkehr/Meyer* (Fn. 10) 1107; *Rechsteiner* (Fn. 17) Rz. 26, mit illustrativem Beispiel.

⁷⁴ Siehe [II.2.2](#).

Bei regelbasierten Algorithmen ergibt sich die Problematik zwar nicht aus der fehlenden Kausalität zwischen Input und Output. Werden jedoch lediglich die kausalen Ketten im Sinne einer blossen Darstellung der einzelnen Rechenschritte aufgeführt, so ist für eine Laiin oder einen Laien ebenfalls nicht vorhersehbar, welcher Sachverhalt aufgrund welcher Rechtsgrundlagen zu einem entsprechenden Ergebnis führt, und der *Rechtsmittelfunktion* ist ebenfalls nicht gedient. Gleiches gilt für die *Fremdkontrollfunktion*.

3. Fehlende Nachvollziehbarkeit der Information

Die Prozesse, die zu einer algorithmenbasierten Entscheidung führen, sind komplex und werden vom System in einer für Menschen kaum verständlichen Sprache ausgeführt.⁷⁵ Die entscheidungsrelevanten Faktoren sind somit insbesondere beim Einsatz von datenbasierten ML-Algorithmen sowie komplexen regelbasierten Algorithmen nur sehr schwer oder gar nicht detektierbar.⁷⁶ Zudem reicht die Darstellung der verwendeten Daten bei datenbasierten ML-Algorithmen – sofern technisch überhaupt möglich – oder die Auflistung der einzelnen Rechenschritte bei regelbasierten Algorithmen nicht aus, um eine nachvollziehbare Information im Sinne des Begründungserfordernisses zu liefern. Dass die entscheidungsrelevanten Faktoren im Kontext algorithmischer Entscheidungssysteme nicht erkannt werden können, bewirkt daher insbesondere, dass die *Selbstkontrollfunktion* nicht zum Tragen kommt. Überdies wird ein algorithmisches System kaum in der Lage sein, relevante von irrelevanten Parteivorbringen zu unterscheiden. Begründungen würden dadurch mit für die Entscheidung irrelevanten Informationen angereichert und entsprechend umfangreicher.⁷⁷ Die Parteien könnten nur schwer erkennen, ob eine Information tatsächlich entscheidungserheblich war.⁷⁸ Dies stellt die *Rechtsmittel-, die Akzeptanz- und die Fremdkontrollfunktion* vor Herausforderungen.

V. Ungeeignete und teilweise geeignete Lösungen

Die Diskussion rund um die Regulierung von algorithmischen Systemen ist spätestens seit der Vorlage des Gesetzgebungsvorhabens der Europäischen Union über eine Verordnung harmonisierter Vorschriften für Künstliche In-

⁷⁵ Braun Binder et al. (Fn. 27) Studie KI 2021, 30.

⁷⁶ Siehe [II.2.1](#) und [II.2.2](#).

⁷⁷ Kürsteiner (Fn. 38) Rz. 59.

⁷⁸ Kürsteiner (Fn. 38) Rz. 59.

telligenz (AI Act) im Jahr 2021 entfacht.⁷⁹ Seither kursieren verschiedene Vorschläge, wie die Nachvollziehbarkeit algorithmenbasierter Entscheidungen hergestellt werden kann. Die inzwischen verabschiedete Fassung des AI Act sieht in Art. 86 ein Recht auf Erläuterung der Entscheidungsfindung im Einzelfall vor, wenn die Entscheidung auf der Grundlage bestimmter Hochrisiko-KI-Systeme getroffen wurde und Rechtswirkungen entfaltet oder die Person in ähnlicher Weise beeinträchtigt.⁸⁰ Art. 86 Abs. 1 AI Act präzisiert, dass die Erläuterung sich insbesondere auf die wichtigsten Elemente der getroffenen Entscheidung beziehen sowie klar und aussagekräftig sein muss. Es bleibt abzuwarten, wie diese Pflicht in Zukunft umgesetzt wird.

Auch in der Wissenschaft werden verschiedene Vorschläge zur Schaffung von Transparenz über algorithmische Systeme diskutiert. Einige dieser Vorschläge sind zur Umsetzung der Begründungspflicht bei Verfügungen allerdings nicht geeignet (1). Gleichwohl gibt es Lösungsansätze, mit denen die Begründungsfunktionen beim Einsatz algorithmischer Entscheidungssysteme – zumindest teilweise – erfüllt werden können (2).

1. Keine Erfüllung der Begründungsfunktionen

1.1. Erklärung der Funktionsweise des algorithmischen Systems

Die Begründungspflicht verlangt eine Darlegung der berücksichtigten Sachverhaltsinformationen und Rechtsgrundlagen, die für das Ergebnis massgeblich waren.⁸¹ Dabei muss nicht der gesamte Ablauf des (menschlichen oder

⁷⁹ Siehe den Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union vom 21.4.2021, COM/201/2026 final, <<https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52021PC0206>> (zuletzt besucht am 25.6.2024).

⁸⁰ 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> (zit. AI Act); siehe auch ErwGr 171 AI Act.

⁸¹ Siehe III.3.

algorithmischen) Entscheidungsprozesses nachvollziehbar sein.⁸² Im Gegenteil: Komplexe, deskriptive Beschreibungen der technischen Aspekte in Verbindung mit der Offenlegung der dem System zugrundeliegenden Algorithmen⁸³ können gerade keine Begründung i.S.v. Art. 29 Abs. 2 BV liefern, da sie keinen wesentlichen Bezug zum konkreten Einzelfall aufweisen. Forschungszweige der Informatik wie *Explainable Artificial Intelligence* (XAI), die versuchen, hochkomplexe algorithmische Systeme zu erklären, sind daher als Ersatz der Begründung wenig hilfreich.⁸⁴

1.2. Erklärung durch das System selbst oder ein externes System

In der Literatur wird ferner diskutiert, nicht nachvollziehbare Blackbox-Modelle nachvollziehbar zu machen, indem ein weiteres Modell «aufgesetzt» wird, das speziell für die Analyse der Zusammenhänge zwischen Input und Output verwendet wird.⁸⁵ Da dieses Modell eine möglichst einfache Erklärung liefern soll, würden die im Einzelfall massgeblichen Kriterien vermutlich stark vereinfacht, sodass es mit hoher Wahrscheinlichkeit nicht alle massgeblichen Tatsachen, Rechtsnormen und Sachverhaltsinformationen berücksichtigen

⁸² Was in der Realität sowieso kaum möglich ist (siehe III.1). «Transparenz» des algorithmischen Entscheidungssystems richtet sich höchstens an Expertinnen und Experten (Peter Parycek/Verena Schmid/Anna-Sophie Novak, *Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations, and Framework Conditions*, *Journal of the Knowledge Economy* 2023, <<https://doi.org/10.1007/s13132-023-01433-3>> [zuletzt besucht am 25.6.2024]).

⁸³ Zur Forderung nach Offenlegung der Algorithmen siehe Christen et al. (Fn. 30) TA-SWISS Studie KI 2020, 134 Fn. 52; Jan Etscheid, *Automatisierungspotenziale in der Verwaltung*, in: Resa Mohabbat-Kar/Basanta E. P. Thapa/Peter Parycek (Hrsg.), (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft, Berlin 2018, 126 ff., 148 f.

⁸⁴ XAI beschreibt einen computerwissenschaftlichen Forschungszweig, der sich mit der Erklärbarkeit der Funktionsweise algorithmischer Systeme beschäftigt (siehe dazu etwa Filip Karlo Dosiilovic/Mario Brcic/Nikica Hlupic, *Explainable Artificial Intelligence: A Survey*, 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIpRo), Opatija 2018, <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8400040>> [zuletzt besucht am 25.6.2024]). Gemäss Wischmeyer (Fn. 9) 629, vermitteln solche Erklärungen höchstens Expertinnen und Experten ein «besseres Verständnis» über das System selbst.

⁸⁵ P. Jonathon Phillips et al., *Four Principles of Explainable Artificial Intelligence*, Draft NISTIR 8312, September 2020, 10, <<https://doi.org/10.6028/NIST.IR.8312-draft>> (zuletzt besucht am 25.6.2024).

würde. Damit wäre den Begründungsfunktionen, insbesondere der *Rechtsmittelfunktion*, nicht gedient.

1.3. Offenlegung der Trainings- bzw. Validierungsdaten

Zur Herstellung von Transparenz algorithmischer Entscheidungssysteme wird teilweise die Offenlegung der dem datenbasierten ML-System zugrunde liegenden Trainings- bzw. Validierungsdaten gefordert.⁸⁶ Fraglich ist jedoch, ob damit der Begründungspflicht – als Teilaspekt der Transparenzforderung – gedient ist. Die Trainings- und Validierungsdaten bilden die Grundlage für die Regeln, die ein ML-System aufstellt, um sie auf unbekannte Eingaben anzuwenden.⁸⁷ Sie enthalten Informationen über vergangene Sachverhalte und insbesondere darüber, wie diese bewertet wurden.⁸⁸ Im Kontext des Behördeneinsatzes sollten die Trainings- und Validierungsdaten daher vergangene rechtliche Sachverhalte und deren rechtliche Beurteilung umfassen. Dazu gehört auch die Subsumtion, die letztlich die Grundlage für die zu erstellenden Regeln bildet. Allerdings ist es bereits aus praktischen Gründen nicht möglich, sämtliche Trainings- und Validierungsdaten in der Begründung selbst abzubilden, da es sich regelmässig um sehr umfangreiche und komplexe Datensätze handelt.⁸⁹ Enthielte die Begründung daher sämtliche Trainings- und Validierungsdaten, die dem algorithmischen System zugrunde liegen, würden aufgrund des erforderlichen Datenumfangs viele für die Einzelfallentscheidung irrelevante Informationen aufgeführt.⁹⁰ Dies würde weder eine sachgerechte Anfechtung im Sinne der *Rechtsmittelfunktion* noch eine nachvollziehbare Erklärung des Einzelfalls im Sinne der *Akzeptanzfunktion* ermöglichen. Die Gesamtheit der Trainings- und Validierungsdaten hilft der betroffenen Person daher regelmässig nicht weiter.⁹¹

⁸⁶ Etwa Christen et al. (Fn. 30) TA-SWISS Studie KI 2020, 310; siehe zu den Anforderungen an Trainings-, Validierungs- und Testdatensätze Art. 10 Abs. 2–5 AI Act.

⁸⁷ Philipp Hacker, Ein Rechtsrahmen für KI-Trainingsdaten, Zeitschrift für geistiges Eigentum (ZGE) 2020 239 ff., 241 f.

⁸⁸ Siehe dazu Felix Neutatz/Ziawasch Abedjan, What is «Good» Training Data?, in: Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz/Frauke Rostalski (Hrsg.), Künstliche Intelligenz, Tübingen 2022, 1 ff., 2.

⁸⁹ Siehe [II.2.2](#).

⁹⁰ Auch wenn gemäss Wischmeyer (Fn. 9) 631, Begründungen algorithmenbasierter Entscheidungen über die Basisinformationen hinausgehen können, würde das Aufführen jeglicher Trainingsdaten den Umfang klar sprengen.

⁹¹ Sie können dagegen ein wichtiger Anknüpfungspunkt für die allgemeine Transparenz sein (siehe Braun Binder/Obrecht [Fn. 2] Öffentliches Verzeichnis [im Erscheinen]).

2. Teilweise Erfüllung der Begründungsfunktionen

Im Folgenden werden verschiedene, nicht abschliessende Lösungsansätze für die Begründung algorithmenbasierter Verfügungen vorgestellt, die auch kombiniert werden können.⁹²

2.1. Kennzeichnungspflicht

Damit die betroffene Person Kenntnis davon erlangt, dass ihr Anliegen durch ein algorithmenbasiertes Entscheidungssystem bearbeitet wurde, bedarf es einer entsprechenden Kennzeichnung der Verfügung.⁹³ Eine solche Kennzeichnungspflicht sieht Art. 21 Abs. 4 DSGVO⁹⁴ bereits für automatisierte Einzelentscheidungen von Bundesorganen vor. Sie stellt ein Mittel zur Gewährleistung des rechtlichen Gehörs der betroffenen Person dar.⁹⁵ Die (nachträgliche) Kennzeichnungspflicht ist zwar ein wichtiges Element einer algorithmenbasierten Verfügung,⁹⁶ genügt für sich allein aber noch nicht der Begründungspflicht.⁹⁷

2.2. Weiterführende Informationen, insbesondere zur Folgenabschätzung

Neben der Kennzeichnungspflicht kann der Begründungspflicht auch durch einen Verweis auf weitere Informationen über das algorithmische Entscheidungssystem selbst gedient werden.⁹⁸ Solche Informationen könnten etwa in

⁹² Die Lösungsansätze für die Sicherstellung der Begründungspflicht können auch für Verwaltungshandeln, das nicht in Form einer Verfügung stattfindet, aber dennoch nach Transparenz im Einzelfall verlangt, herangezogen werden. Zu denken ist etwa an den Einsatz von datenbasierten Chatbots zur Erteilung von behördlichen Auskünften (siehe dazu *Nadja Braun Binder/Liliane Obrecht/Grace Wittmer*, Vertrauensschutz bei fehlerhaften Behördenauskünften durch Chatbots, *iusNet DigR* vom 30.5.2024).

⁹³ *Verbraucherzentrale Bundesverband*, Thesenpapier: Algorithmische Entscheidungsfindung, in: *Resa Mohabbat-Kar/Basanta E. P. Thapa/Peter Parycek* (Hrsg.), (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft, Berlin 2018, 516 ff., 522.

⁹⁴ Bundesgesetz über den Datenschutz (Datenschutzgesetz, DSGVO) vom 25. September 2020 (SR 235.1).

⁹⁵ *Braun Binder* (Fn. 17) Perspektive Datenschutzrecht, 32.

⁹⁶ Gemäss *Rechsteiner* (Fn. 17) Rz. 24, stellt diese Kennzeichnungspflicht sogar einen Ausfluss des verfassungsrechtlichen rechtlichen Gehörs dar.

⁹⁷ *Wiederkehr/Meyer* (Fn. 10) 1106.

⁹⁸ Gemäss *Wischmeyer* (Fn. 9) 631, können Begründungen algorithmenbasierter Entscheidungen auch über die Basisinformationen hinausgehen, was Ausdruck der hohen Qualität des

einem öffentlichen Verzeichnis über den Einsatz algorithmischer Entscheidungssysteme enthalten sein.⁹⁹ Hierzu gehört insbesondere auch ein Verweis auf die Ergebnisse einer vorab durchgeführten Folgenabschätzung des algorithmischen Entscheidungssystems.¹⁰⁰ Denn die Folgenabschätzung soll erstens die Behörde, die ein solches System einsetzt, in die Lage versetzen, frühzeitig die Auswirkungen des Systems auf die Grundrechte der betroffenen Personen darzulegen.¹⁰¹ Zweitens können allgemeine Ausführungen über das System und seine Auswirkungen die fehlende Nachvollziehbarkeit einer Einzelentscheidung zumindest insoweit relativieren, als der betroffenen Person vor Augen geführt wird, inwieweit die Veränderung einer bestimmten Input-Variable das Ergebnis verändern kann.¹⁰² Auch wenn sich die Behörde damit nicht mit jeder Einzelentscheidung auseinandersetzt, so kann damit doch die *Selbstkontrollfunktion* einerseits und die *Akzeptanzfunktion* andererseits zumindest teilweise erfüllt werden. Insofern kommt der Folgenabschätzung – aber auch den anderen im öffentlichen Verzeichnis enthaltenen Informationen – eine gewisse Kompensationsfunktion für verfahrensrechtliche Defizite beim Einsatz algorithmischer Systeme zu.¹⁰³

2.3. Aufführen einzelner Datenpunkte

Auch wenn es sich für die Erfüllung der Begründungspflicht als nicht sinnvoll erwiesen hat, sämtliche Trainings- und Validierungsdaten vorzulegen,¹⁰⁴ kann es hilfreich sein, diejenigen Datenpunkte aufzuführen, die für den Einzelfall von besonderer Bedeutung waren. Dies gilt umso mehr, je grösser der behördliche Ermessensspielraum im Einzelfall ist. Insbesondere bei negativen Entscheidungen im Rahmen eines Gesuchverfahrens dürfte es die Akzeptanz der Verfügung erheblich fördern, wenn für die betroffene Person erkennbar ist,

Rechtsschutzes bildet und nicht etwa die betroffene Person mit hoher Informationsdichte überfordern soll.

⁹⁹ Siehe dazu *Braun Binder/Obrecht* (Fn. 2) Öffentliches Verzeichnis (im Erscheinen).

¹⁰⁰ Der Begriff der Folgenabschätzung steht insbesondere auch im Zusammenhang mit Art. 22 DSGVO, wonach der Verantwortliche vorgängig zu einer Datenbearbeitung eine Datenschutz-Folgenabschätzung erstellt, wenn eine Bearbeitung ein hohes Risiko etwa für die Grundrechte darstellt (Abs. 1). Sie enthält eine Beschreibung der geplanten Bearbeitung, eine Bewertung der Risiken für (etwa) die Grundrechte der betroffenen Person sowie Massnahmen zu deren Schutz (Abs. 3).

¹⁰¹ Siehe dazu *Müller/Pannatier* (Fn. 1).

¹⁰² *Dollinger* (Fn. 22) 46 f. m.w.H.

¹⁰³ *Dollinger* (Fn. 22) 46.

¹⁰⁴ Siehe [V.1.3](#).

in welchen Punkten ihr Gesuch von den positiven Entscheidungen abweicht. Hinweise in der Begründung, wie ein Gesuch zu ändern wäre, um eine bestätigende Verfügung zu erhalten, könnten die Akzeptanz der Entscheidung zusätzlich fördern.

2.4. Wahl des Modells und der Technik

Der Einsatz algorithmischer Entscheidungssysteme darf nicht zu einer Verringerung der Anforderungen an die Begründungsdichte führen.¹⁰⁵ Daher muss bei der Wahl des Modells und der Technik selbst angesetzt werden. So existieren etwa datenbasierte ML-Modelle, die aufgrund ihrer intrinsischen Eigenschaften nachvollziehbare Ergebnisse liefern können. So sind zum Beispiel Ergebnisse auf Basis von Entscheidungsbäumen besser nachvollziehbar als Ergebnisse auf Basis von künstlichen neuronalen Netzen, da erkennbar ist, welchen «Weg» das System genommen hat und welche Datenpunkte für die Entscheidung ausschlaggebend waren.¹⁰⁶ Dementsprechend könnten nur solche ML-Techniken/Modelle für behördliche Entscheidungen zugelassen werden, die der Nachvollziehbarkeit zugänglich sind.¹⁰⁷

2.5. Menschliche Verantwortung

Der Mensch kann in unterschiedlichem Masse in den algorithmenbasierten Entscheidungsprozess eingebunden sein. Sei es als endgültiger Entscheidungsträger, dem das System lediglich Entscheidungsalternativen offeriert, oder (im Nachgang) als Entscheidungsprüfer bei der vollständig vom Menschen unabhängigen Entscheidungsfindung.¹⁰⁸ In beiden Szenarien obliegt es der verantwortlichen Entität, sich nicht durch systemimmanente Verzerrungen zu fehlerhaften Entscheidungen verleiten zu lassen bzw. solche Entscheidungen nicht abzusegnen (Stichwort: Automation Bias¹⁰⁹). Um solche Fehl-

¹⁰⁵ Siehe [III.2](#).

¹⁰⁶ Wolfgang Ertel, Grundkurs Künstliche Intelligenz, Eine praxisorientierte Einführung, 4. A., Wiesbaden 2016, 217.

¹⁰⁷ *Wiederkehr/Meyer* (Fn. 10) 1106 f., knüpfen beim Automatisierungsgrad an und propagieren, diesen zu reduzieren, sofern keine angemessene Begründung durch ein «künstliches Entscheidungssystem» erreicht werden kann. Dabei wird nach v.A. jedoch zu wenig berücksichtigt, dass es gerade nicht allein auf den Automatisierungsgrad des Verfahrens ankommt, sondern dass eine hinreichende Begründung auch im teilweise algorithmenbasierten Verfahren nicht erfolgen kann, wenn nicht nachvollziehbare Modelle verwendet werden.

¹⁰⁸ Zu den verschiedenen Stufen an Automatisierung siehe [II.1](#).

¹⁰⁹ Siehe dazu bereits [IV.1](#).

allokationen von Vertrauen präventiv zu vermeiden, muss zum einen die Datenkompetenz des Verwaltungspersonals gefördert werden.¹¹⁰ Die entscheidungsbefugte Person muss in der Lage sein, den Entscheidungsprozess zu rekonstruieren, um eine nachvollziehbare Begründung verfassen zu können.

Zum anderen ist die Durchführung retrospektiver, externer Analysen zu befürworten.¹¹¹ Damit können insbesondere unsachliche Einflüsse oder Verzerrungen in der Entscheidungsfindung identifiziert und bewertet werden. Diese Massnahmen dienen als kompensatorische Elemente zur Stärkung der Selbstkontroll-, Fremdkontroll- und Akzeptanzfunktion der Begründungspflicht im algorithmenbasierten Entscheidungsverfahren.

2.6. Reduktion der Begründungsdichte

In Verwaltungsbereichen, in denen eine Vielzahl von Entscheidungen über gleichartige Fälle getroffen werden, stellen Lehre und Rechtsprechung geringere Anforderungen an die Begründungsdichte (Massenverwaltung).¹¹² Dies ist auf die weitgehende Standardisierung des Verfahrens und die damit verbundenen prozessökonomischen Überlegungen zurückzuführen.¹¹³ Zentral ist dabei, dass gleiche Eingaben immer zum gleichen Ergebnis führen, es also keinen Ermessensspielraum der Behörde im Einzelfall gibt.¹¹⁴ Damit ist die Kausalität des Entscheidungsprozesses gewährleistet. Für regelbasierte Algorithmen könnte diese bestehende Praxis hinsichtlich der Begründungsdichte allenfalls analog angewendet werden, wenn sie ebenfalls standardisierte Prozesse ausführen. Für solche Entscheidungen könnten entsprechend geringere Anforderungen an die Begründungsdichte gestellt werden. Auf Verlangen der von der Entscheidung betroffenen Person müsste allerdings trotzdem eine Begrün-

¹¹⁰ Mirco Schoenfeld/Paloma Mendes Saldanha, Bias is not just in the Data: A holistic view on AI systems, the need for data literacy, and regulatory intervention, 13th Congresso de Ciências Criminais do PPGCCRIM PUCRS vom 14.9.2022, Abschn. 2.1, <<https://ebooks.pucrs.br/edipucrs/acessolivre/anais/congresso-internacional-de-ciencias-criminais/assets/edicoes/2022/arquivos/44.pdf>> (zuletzt besucht am 25.6.2024), auch unter «AI literacy» (für: KI Kompetenz) gehandelt.

¹¹¹ Braun Binder (Fn. 17) KI, 474 f.; Wischmeyer (Fn. 9) 631, schlägt in diesem Zusammenhang den Aufbau einer «arbeitsteiligen Begründungs- und Kontrollarchitektur» vor.

¹¹² OFK VwVG-Wiederkehr/Meyer/Böhme (Fn. 8) Art. 35 VwVG N 28.

¹¹³ René Wiederkehr, Öffentliches Verfahrensrecht, 2. A., Bern 2022, Rz. 112; René Wiederkehr/Kaspar Plüss, Praxis des öffentlichen Verfahrensrechts, Eine systematische Analyse der Rechtsprechung, Bern 2020, Rz. 598.

¹¹⁴ OFK VwVG-Wiederkehr/Meyer/Böhme (Fn. 8) Art. 35 VwVG N 28.

ung geliefert werden können.¹¹⁵ Die Behördenmitarbeitenden müssten also trotzdem in der Lage sein, eine Begründung zu verfassen.

Bei datenbasierten Algorithmen stellt sich die mögliche Reduzierung der Begründungsdichte komplexer dar, da gerade keine kausalen Zusammenhänge vorliegen, sondern auf Wahrscheinlichkeitsberechnungen abgestellt wird.¹¹⁶ Im Lichte der verfassungsrechtlichen Anforderungen und der bestehenden Praxis im Bereich der Massenverwaltung erscheint es daher nicht gerechtfertigt, Entscheidungen, die auf datenbasierten Algorithmen beruhen, reduzierten Begründungsanforderungen zu unterwerfen. Gegebenenfalls wäre für gewisse Verwaltungsbereiche ein Paradigmenwechsel zu prüfen. An die Stelle der Nachvollziehbarkeit und Begründung der Verfügung im Einzelfall könnten eine retrospektive Kontrolle des Systems im Sinne einer Analyse der Entscheidungspraxis¹¹⁷ sowie eine Kontrolle des Systems vor dessen Einsatz treten (etwa durch einfache Zufallsstichproben, *Random Sampling*¹¹⁸). Die *Akzeptanzfunktion* kann dadurch allerdings nur kompensiert werden, wenn das Vertrauen der Bevölkerung in automatisierte Verfahren steigt.¹¹⁹ Auch wäre ein solches Kontrollsystem ohnehin nur in nicht sensiblen Bereichen denkbar.

VI. Fazit

Behörden können den rechtsstaatlichen Anforderungen der Begründungspflicht beim Einsatz algorithmischer Entscheidungssysteme im Rahmen des Erlasses von Verfügungen nicht ohne Weiteres genügen. Die Funktionen der Begründungspflicht müssen allerdings auch beim Einsatz neuer Technologien gewahrt bleiben. Um diese erfüllen zu können, sind verschiedene Lösungsansätze in Betracht zu ziehen. Insbesondere bietet sich auch eine Kombination der verschiedenen Möglichkeiten an. Die Begründungspflicht hat neben ihrem rechtsstaatlichen, individuell-konkreten Gehalt im Zusammenhang mit algorithmenbasierten Entscheidungen noch einen weiteren – *generelleren* – Gehalt. Unter der Voraussetzung, dass die Begründung die Funktionsweise eines

¹¹⁵ Siehe [I](#).

¹¹⁶ Siehe [II.2.2](#).

¹¹⁷ Siehe bereits den Lösungsansatz in [V.2.5](#).

¹¹⁸ Steven Vethman et al., Random Sample as a Pre-pilot Evaluation of Benefits and Risks for AI in Public Sector, in: Slawomir Nowaczyk et al. (Hrsg.), ECAI 2023 International Workshops, Basel 2024, 115 ff., schlagen vor, *Random Sampling* bereits in der Evaluationsphase des Systems einzusetzen.

¹¹⁹ Siehe [IV.1](#).

Algorithmus auf den Einzelfall herunterbricht,¹²⁰ werden dem algorithmischen Entscheidungssystem gewissermassen insgesamt Transparenzpflichten auferlegt.

¹²⁰ Siehe [III.3](#); Martini (Fn. 30) 190.

Verfügung und automatisierte Einzelentscheidung – same same but different?

Liliane Obrecht

Erstabdruck in: *ex/ante* 2/2022, S. 38–45

Das Verwaltungsverfahren ist massgeblich vom Verfügungsbegriff bestimmt. Im Zuge der Revision des Datenschutzgesetzes wird der Begriff der automatisierten Einzelentscheidung eingeführt – und als «gewöhnliche Verfügung» bezeichnet. Der Frage, ob diese beiden Begriffe identisch sind oder sich nicht doch unterscheiden, geht der vorliegende Aufsatz nach.

The administrative procedure is decisively determined by the concept of rulings. In the course of the revision of the Data Protection Act, the automated individual decision is being introduced – and is called an «ordinary ruling». The question of whether these two terms are not different after all is explored in this article.

La procédure administrative est définie de manière déterminante par la notion de décision. Dans le cadre de la révision de la loi sur la protection des données, la décision individuelle automatisée a été introduite et qualifiée de «décision ordinaire». Le présent article se penche sur la question de savoir si ces deux termes ne sont pas différents.

Inhalt

I.	Einleitung	92
II.	Automatisierte Einzelentscheidung	94
1.	Der datenschutzrechtliche Begriff	94
1.1.	Vorbemerkung: Teil- und Vollautomation	94
1.2.	Ausgangspunkt: Art. 21 nDSG	95
2.	Der verwaltungsrechtliche Begriff	97
2.1.	Einsatzszenarien in der Verwaltung	97
2.2.	Art. 21 nDSG als allgemeine Zulässigkeitsnorm im Verwaltungsalltag?	99
III.	Automatisierte Einzelentscheidung als Verfügung im Sinne von Art. 5 VwVG?	100
1.	Traditioneller Verfügungsbegriff	100
2.	Diskrepanz: behördlicher Willensbildungsprozess...	101

3. ... bei teilautomatisierter Einzelentscheidung?	102
4. ... bei vollautomatisierter Einzelentscheidung?	102
IV. Fazit	104

I. Einleitung

In den letzten Jahren ist im Zuge der allgemeinen digitalen Transformation auch die Digitalisierung des öffentlichen Sektors in den Blick gerückt. Dabei ist die Idee einer automatisierten Verwaltung keineswegs neu – die rechtswissenschaftliche Auseinandersetzung damit begann bereits in den 1970er Jahren.¹ Innerhalb des öffentlichen Sektors stellt sich dabei insbesondere die Frage, inwiefern das Verwaltungsverfahren digitalisiert werden kann. Letzteres ist sowohl im Bund als auch in den Kantonen vom Verfügungsbegriff² geprägt – die Verfügung ist das Ziel des erstinstanzlichen Verwaltungsverfahrens sowie elementares Handlungsinstrument der Verwaltungsbehörden.³ Ihr automatisierter Erlass bildet im Rahmen der Verwaltungsdigitalisierung ein zentrales Thema.⁴ In der Schweiz wird diese «automatisierte Einzelentscheidung»

¹ Insbesondere im deutschsprachigen Raum NIKLAS LUHMANN, *Recht und Automation in der öffentlichen Verwaltung*, Berlin 1966; RALF-MICHAEL POLOMSKI, *Der automatisierte Verwaltungsakt*, Diss. Göttingen 1992, Berlin 1993; BENNO DEGRANDI, *Die automatisierte Verwaltungsverfügung*, Diss. Zürich 1977, Zürich 1997.

² Im vorliegenden Beitrag geht es um den bundesrechtlichen Verfügungsbegriff, weshalb nicht weiter auf die in den Kantonen unterschiedlichen Begrifflichkeiten – z. B. *Entscheid* oder *Anordnung* – eingegangen wird.

³ Statt vieler REGINA KIENER/BERNHARD RÜTSCHÉ/MATHIAS KUHN, *Öffentliches Verfahrensrecht*, 3. Aufl., Zürich/St. Gallen 2021, N 341.

⁴ NADJA BRAUN BINDER, *Automatisierte Entscheidungen: Perspektive Datenschutzrecht und öffentliche Verwaltung*, SZW 2020, 27 ff. (zit. *Perspektive*); NADJA BRAUN BINDER, *Künstliche Intelligenz und automatisierte Entscheidungen in der öffentlichen Verwaltung*, SJZ 2019, 467 ff. (zit. *KI*); FLORENT THOUVENIN/ALFRED FRÜH, *Automatisierte Entscheidungen: Grundfragen aus der Perspektive des Privatrechts*, SZW 2020, 3 ff.; ROLF WEBER, *Automatisierte Einzelentscheidungen: Perspektive Grundrechte*, SZW 2020, 18 ff.; DAVID RECHSTEINER, *Der Algorithmus verfügt, Verfassungs- und verwaltungsrechtliche Aspekte automatisierter Einzelentscheidungen*, Jusletter 26. November 2018; SIMON ROTH, *Die automatisierte Einzelentscheidung*, digma 2017, 104 ff.; FLORENT THOUVENIN/ALFRED FRÜH/DAMIAN GEORGE, *Datenschutz und automatisierte Entscheidungen*, Jusletter 26. November 2018; MARIA-URANIA DOVAS, *Automatisierte Einzelentscheidungen*, digma 2017, 98 ff. In Deutschland wurde diese Diskussion bereits im Jahr 2016 entfacht und läuft seither intensiv, siehe (Auswahl) NADJA BRAUN BINDER, *Vollautomatisierte Verwaltungsverfahren im allgemeinen Verwaltungsrecht?*

insbesondere im Zusammenhang mit dem neuen Datenschutzgesetz (nDSG⁵) diskutiert, das den Begriff in Art. 21 nDSG einführt. Demnach liegt eine automatisierte Einzelentscheidung immer dann vor, wenn eine Entscheidung ausschliesslich auf einer automatisierten (Daten-)Bearbeitung beruht und für die betroffene Person mit einer Rechtsfolge verbunden ist oder sie erheblich beeinträchtigt. Im Zuge der Revision des Datenschutzgesetzes werden auch weitere Erlasse um den Begriff der automatisierten Einzelentscheidung und einen Verweis auf Art. 21 nDSG ergänzt.⁶ Dabei geht der Bundesgesetzgeber davon aus, dass es sich bei solchen automatisierten Einzelentscheidungen um «gewöhnliche Verfügungen» handelt, die allerdings ohne menschliches Zutun erlassen werden.⁷ Gar nicht beziehungsweise nicht ausreichend debattiert wird jedoch die Frage nach der Qualifikation dieser automatisierten Einzelentscheidungen (handelt es sich wirklich um Verfügungen i.S.v. Art. 5 VwVG?). Dem möchte der vorliegende Beitrag im Folgenden nachgehen. Insbesondere die Bedeutung der Verfügung als Handlungsinstrument für die öffentliche Verwaltung sowie die damit einhergehende Frage nach der verwaltungsrechtlichen Verankerung eines solchen Instruments sind unklar.⁸ Denn wird die Verwaltungstätigkeit vollständig menschenunabhängig durch Algo-

Der Gesetzesentwurf zur Modernisierung des Besteuerungsverfahrens als Vorbild für vollautomatisierte Verwaltungsverfahren nach dem VwVfG, NVwZ 2016, 960 ff. (zit. Gesetzesentwurf); HANS PETER BULL, Der «vollständig automatisiert erlassene» Verwaltungsakt: Zur Begriffsbildung und rechtlichen Einhegung von «E-Government», DVBl 2017, 409 ff.; ARIANE BERGER, Der automatisierte Verwaltungsakt, NVwZ 2018, 1260 ff.; ANNETTE GUCKELBERGER, Automatisierte Verwaltungsentscheidungen: Stand und Perspektiven, DÖV 2021, 566 ff.; BETTINA SPILKER, E-Government: Anforderungen an das Verwaltungsverfahren, NVwZ 2022, 665 ff.

⁵ Das neue Gesetz tritt am 1. September 2023 in Kraft.

⁶ Die Veranlagung nach dem Zollgesetz vom 16. März 2005 (ZG; SR 631.00), die Festsetzung der Steuerbeträge nach dem Bundesgesetz über die Tabakbesteuerung vom 21. März 1969 (Tabaksteuergesetz, TStG; SR 641.31), dem Mineralölsteuergesetz vom 21. Juni 1996 (MinöStG; SR 641.411) und dem Bundesgesetz über die Biersteuer (Biersteuergesetz, BStG; SR 641.411) sowie zur Festsetzung der Schwerverkehrsabgabe nach dem Bundesgesetz über eine leistungsabhängige Schwerverkehrsabgabe (Schwerverkehrsabgabengesetz, SVAG; SR 641.81).

⁷ Botschaft über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz vom 15. September 2017, BBl 2017 6941 ff., 7057.

⁸ Klärungsbedarf beim Verfügungsbegriff in Zeiten automatisierter Einzelentscheide identifizieren NADIA BRAUN BINDER, Als Verfügungen gelten Anordnungen der Maschinen im Einzelfall... – Dystopie oder künftiger Verwaltungsalltag?, ZSR 2020 I, 253 ff., 272 (zit. Dystopie); BRAUN BINDER, Perspektive (Fn. 4), 32 sowie ANDREAS GLASER, Einflüsse der Digitalisierung auf das schweizerische Verwaltungsrecht, SJZ 2018, 181 ff. (zit. Einflüsse), 187.

rithmeneinsatz ausgeführt, bedeutet dies im Vergleich zum bisherigen Einsatz von Informationstechniken einen Perspektivenwechsel im Verwaltungsalltag. Diese Veränderungen rechtfertigen es, der Verfügung als «Dreh- und Angelpunkt des Verwaltungsverfahrens- und Verwaltungsprozessrechts»⁹ im Kontext der Verwaltungsdigitalisierung besondere Aufmerksamkeit zu schenken.

Im Folgenden wird zuerst dargelegt, was eine automatisierte Einzelentscheidung ist, wobei der datenschutzrechtliche Begriff beleuchtet und der massgebende verwaltungsrechtliche Begriff herausgearbeitet wird (II.). Sodann wird der traditionelle Verfügungsbegriff der automatisierten Einzelentscheidung gegenübergestellt (III.). Im Fazit werden die wichtigsten Aspekte nochmals aufgegriffen und ein Blick in die Zukunft geworfen (IV.).

II. Automatisierte Einzelentscheidung

1. Der datenschutzrechtliche Begriff

1.1. Vorbemerkung: Teil- und Vollautomation

Der Begriff «automatisierte Einzelentscheidung» ist nicht ohne weitere Präzisierung, insbesondere hinsichtlich der konkreten Automatisierungsgrade¹⁰, verständlich.¹¹ Häufig wird in Voll- und Teilautomation unterteilt.¹² Bei der Teil-

⁹ MÜLLER MARKUS, Art. 5 VwVG N 6, Kommentar zum Bundesgesetz über das Verwaltungsverfahren (VwVG), Christoph Auer/Markus Müller/Benjamin Schindler (Hrsg.), 2. Aufl., Zürich 2019.

¹⁰ Siehe CLARA-ANN GORDON/TANJA LUTZ, Haftung für automatisierte Entscheidungen – Herausforderungen in der Praxis, SZW 2020, 53 ff., 55; REGINA WEDER, Verfahrensgrundrechtliche Anforderungen an automatisierte Verwaltungsverfahren, in: Monika Simmler (Hrsg.), Smart Criminal Justice, Basel 2021, 238 ff., 241 f. Im Kontext der automatisierten Entscheidungsfindung wird gelegentlich von «Automated Decision Making – ADM» gesprochen. Der Begriff ersetzt in diesen Fällen häufig den Begriff der KI, weshalb er der vorliegenden Problematik nicht gerecht wird, siehe BRAUN BINDER, KI (Fn. 4), 469.

¹¹ Siehe allgemein zu den Begriffen «Automation» und «Automatisierung» im rechtlichen Kontext CHRISTIAN MEYER, Digitale Formulare als Angelpunkt automatisierter Verwaltungsverfahren, ZSR 2022 I, 365 ff., 366 ff. Vorliegend werden die Begriffe jedoch synonym verwendet, da die Unterscheidung aus Sicht des vorliegenden Beitrags nicht von weiterer Bedeutung ist.

¹² Siehe NADJA BRAUN BINDER, Vollautomatisierte Verwaltungsverfahren, vollautomatisiert erlassene Verwaltungsakte und elektronische Aktenführung, in: Margrit Seckelmann (Hrsg.), Digitalisierte Verwaltung: Vernetztes E-Government, 2. Aufl., Berlin 2019, 311 ff., N 5 m.w.H.

automation können zwei Konstellationen unterschieden werden. Erstens kann es sich um die sogenannte Entscheidungsunterstützung handeln, deren Potenzial für die Entwicklung der öffentlichen Verwaltung während der nächsten Jahre als besonders hoch eingeschätzt wird.¹³ Dabei bereitet ein System Aufgaben für den menschlichen Entscheidungsträger beziehungsweise die -trägerin vor, ermittelt Informationen und stellt diese übersichtlich zusammen oder hilft bei deren Auswertung.¹⁴ Die erzeugten Ergebnisse können, müssen aber nicht von der natürlichen Person bei der Entscheidungsfindung berücksichtigt werden.¹⁵ Die finale Entscheidung liegt weiterhin in menschlicher Hand.¹⁶ Zweitens kann es sich um die Konstellation handeln, in der die verantwortliche Behörde die Sachverhaltsdaten eruiert oder zumindest prüft und die Automatisierung erst zur Überprüfung des Entscheids eingesetzt wird.¹⁷ Vollautomation bedeutet schliesslich, dass das System jeden einzelnen Schritt bis hin zur Entscheidungsfällung selbst, also menschenunabhängig, ausführt.¹⁸ Dabei muss selbst der Anstoss zur selbständigen Entscheidungsfindung menschenunabhängig erfolgen.¹⁹

1.2. Ausgangspunkt: Art. 21 nDSG

Um den Begriff der automatisierten Einzelentscheidung im Schweizer Rechtssystem zu qualifizieren, bietet es sich an, Art. 21 nDSG als Ausgangspunkt zu nehmen. Gemäss Abs. 1 handelt es sich dabei um eine Entscheidung, die ausschliesslich auf einer automatisierten Bearbeitung beruht und für die betroffene Person mit einer Rechtsfolge verbunden ist oder sie erheblich beeinträchtigt. Eine automatisiert erlassene Einzelentscheidung i.S.v. Art. 21 Abs. 1 nDSG liegt somit immer dann vor, wenn sie «ohne Dazutun einer natürlichen Per-

¹³ MARKUS CHRISTEN et al., Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz, TA-SWISS Studie, April 2020, 278 ff.

¹⁴ Vgl. ERNST/SCHMIDT/BENEKEN (Fn. 21), 748 f.

¹⁵ Vgl. NADJA BRAUN BINDER, Algorithmic Regulation – Der Einsatz algorithmischer Verfahren im staatlichen Steuerungskontext, in: Hermann Hill/Joachim Wieland (Hrsg.), Zukunft der Parlamente: Speyer Konvent in Berlin, Berlin 2018, 107 ff. (zit. Algorithmic Regulation).

¹⁶ ROGER PLATTNER, Digitales Verwaltungshandeln, Diss. Zürich 2021, Zürich 2021, N 423 ff., insb. N 431 und N 615 f.

¹⁷ NADJA BRAUN BINDER et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4 (zit. Vorprojekt IP6.4), 24.

¹⁸ BRAUN BINDER, Algorithmic Regulation (Fn. 15), 112.

¹⁹ BULL (Fn. 4), 410.

son» erfolgt.²⁰ Die Sachverhaltsbeurteilung sowie die darauf beruhende Entscheidung ergehen beide losgelöst von einer natürlichen Person, mithin geht es um eine Vollautomation.²¹ Diese ist auch dann gegeben, wenn die Einzelentscheidung schliesslich durch eine natürliche Person mitgeteilt wird, sofern die automatisierte Entscheidung in diesem Zeitpunkt nicht mehr beeinflusst werden kann.²² In der Literatur wird zudem die Ansicht vertreten, eine automatisierte Einzelentscheidung liege auch dann noch vor, wenn der Mensch im Rahmen der Programmierung miteinbezogen werde.²³ Gänzlich menschenunabhängig muss die automatisierte Einzelentscheidung somit wohl nicht ergehen. In der Botschaft zum nDSG knüpft der Bundesrat allerdings nicht ausschliesslich an das Merkmal «Mensch» an, sondern zusätzlich an dasjenige der «Komplexität» und der besonderen «Betroffenheit» der adressierten Person: Um als automatisierte Einzelentscheidung qualifiziert zu werden, muss die Automatisierung eine gewisse Komplexität aufweisen und für die betroffene Person eine Rechtsfolge herbeiführen beziehungsweise sie besonders beeinträchtigen.²⁴ Einfache «Wenn-Dann-Entscheidungen», wie beispielsweise die Geldausgabe bei einem Bankomaten, fallen gemäss Bundesrat nicht unter den Begriff der automatisierten Einzelentscheidung.²⁵ Wann die erforderliche Komplexität vorliegt, bleibt, abgesehen von diesem Beispiel, in der Botschaft offen.²⁶ Die Hürde, diese zu bejahen, dürfte in der Praxis jedoch nicht

²⁰ Botschaft DSG (Fn. 7), BBl 2017 7056 f.

²¹ Botschaft DSG (Fn. 7), BBl 2017 7056.

²² Botschaft DSG (Fn. 7), BBl 2017 7057.

²³ RECHSTEINER (Fn. 4), N 3, m.w.H. zur deutschen Rechtslage.

²⁴ Botschaft DSG (Fn. 7), BBl 2017 7057. Dies wäre z. B. bei einer vollautomatisierten Steuerveranlagung der Fall, beispielhaft (zwar für Deutschland) NADJA BRAUN BINDER, *Artificial Intelligence and Taxation: Risk Management in Fully Automated Taxation Procedures*, in: Thomas Wischmeyer/Timo Rademacher (Hrsg.), *Regulating Artificial Intelligence*, Cham 2020, 295 ff., N 3 ff. Beispiele insb. im privatrechtlichen Kontext bei SENIDA MEHMEDOVIC, *Le droit d'accès à une décision individuelle automatisée*, Jusletter 19. September 2022, N 25 ff.

²⁵ Botschaft DSG (Fn. 7), BBl 2017 7057.

²⁶ Auch die Verordnung zum Bundesgesetz über den Datenschutz (DSV) sieht keine Präzisierung vor, wann eine automatisierte Einzelentscheidung vorliegt. In einer im September publizierten Übersicht zu den wichtigsten Änderungen für die Erarbeitung der Rechtsgrundlagen betreffend Datenbearbeitungen durch Bundesorgane durch das Bundesamt für Justiz wird zur Feststellung der erforderlichen Komplexität an die Nachvollziehbarkeit des Entscheidungsprozesses angeknüpft. Allerdings führt dies dazu, dass eben die im Verwaltungsalltag eingesetzten Systeme wohl kaum unter Art. 21 nDSG subsumiert werden, siehe Bundesamt für Justiz BJ, *Totalrevision des Datenschutzgesetzes (DSG)*, September 2022, <https://www.bj.admin.ch/bj/de/home/staat/gesetzgebung/datenschutzstaerkung.html>, abgerufen am 6. September 2022.

all zu hoch sein, da die meisten in der Verwaltung eingesetzten Algorithmen auf «Wenn-Dann-Schemata» beruhen²⁷ und die datenschutzrechtliche Regelung ansonsten ins Leere laufen würde.

Art. 21 Abs. 1 nDSG nimmt sodann den Verantwortlichen beziehungsweise die Verantwortliche in die Pflicht,²⁸ die betroffene Person bei einer automatisierten Einzelentscheidung zu informieren – sei dies im privaten oder im öffentlichen Sektor. Ergeht die automatisierte Einzelentscheidung durch ein Bundesorgan, besteht zudem eine Kennzeichnungspflicht (Art. 21 Abs. 4 nDSG). Gleichzeitig geht der Bundesrat davon aus, dass es sich bei solchen einem Bundesorgane zuzuordnenden automatisierten Einzelentscheidungen um Verfügungen im Sinne von Art. 5 Abs. 1 VwVG handelt,²⁹ weshalb diese auch als «automatisiert erlassene Verfügung(en)» bezeichnet werden könnten.³⁰

2. Der verwaltungsrechtliche Begriff

2.1. Einsatzszenarien in der Verwaltung

Verwaltungsaufgaben werden bereits seit mehreren Jahrzehnten automatisiert.³¹ Automatisierung bedeutet ganz generell, dass *Aufgaben*, die bisher durch eine natürliche Person bewältigt wurden, einer Maschine übertragen werden.³² Eine Tätigkeit ist automatisiert, wenn sie zumindest in einem Teilbereich menschenunabhängig ausgeführt wird.³³ Die automatisierte Datenverarbeitung hat sich in den letzten Jahren stark weiterentwickelt. Technologische Anwendungen können heute durch vorgegebene Handlungsanweisungen (Algorithmen) dahingehend programmiert werden, dass Berechnungen, der Einsatz von Textbausteinen oder der Abgleich verschiedener Informationen durch ein maschinen-

²⁷ Vertieft unten [II.2.1](#).

²⁸ Gemäss Art. 5 lit. j nDSG ist der Verantwortliche bzw. die Verantwortliche eine private Person oder ein Bundesorgan, die oder das allein oder zusammen mit anderen über den Zweck und die Mittel der Bearbeitung entscheidet.

²⁹ Botschaft DSG (Fn. 7), BBl 2017 7059.

³⁰ BRAUN BINDER, Perspektive (Fn. 4), 29.

³¹ Quellennachweis in Fn. 1.

³² Vgl. THOMAS RAMGE, Mensch und Maschine, 2. Aufl., Ditzingen 2018, 17 f.

³³ M.w.H. DANIELA THURNHERR, Automatisierte Verwaltungsverfahren auf Bundesebene, in: Nadja Braun Binder/Peter Bussjäger/Mathias Eller (Hrsg.), Auswirkungen der Digitalisierung auf die Erlassung und Zuordnung behördlicher Entscheidungen, Wien/Hamburg 2021, 133 ff., 135 f. Zu den Ursprüngen der Automation vgl. LUHMANN (Fn. 1), 9 f.; POLOMSKI (Fn. 1), 24.

gesteuertes System vorgenommen werden.³⁴ Verschiedentlich werden diese Entwicklungen unter dem Schlagwort «Künstliche Intelligenz» (KI) zusammengefasst.³⁵ Für die Verwaltung ist jedoch – gemäss heutigem Stand der Dinge und Technik – nicht jede maschinengestützte Anwendung von Interesse. Grosses Potenzial wird insbesondere im Bereich der automatisierten Steuerveranlagung identifiziert oder im Sozialversicherungsbereich hinsichtlich der Ausrichtung der kantonalen Prämienverbilligung.³⁶ In der Verwaltungsdigitalisierung steht somit die automatisierte Entscheidungsfindung im Vordergrund.³⁷ Um dieser bestmöglich die Türen zu öffnen, bedarf sie grosser Datenmengen und strukturierter Prozesse.³⁸ Daher zeichnet sich ab, dass sich die für die öffentliche Verwaltung realistischen Einsatzszenarien automatisierter Einzelentscheidungen vorerst auf die sogenannte Massenverwaltung konzentrieren.³⁹ Denn im Massenverfahren werden viele gleichartige Fälle auf dieselbe Art und Weise abgehandelt und es besteht kein Ermessensspielraum der Behörde.⁴⁰

³⁴ Siehe nur HARTMUT ERNST/JOCHEN SCHMIDT/GERD BENEKEN, Grundkurs Informatik, 7. Aufl., Wiesbaden 2020, 501.

³⁵ Aus der Literatur z. B. MARIO MARTINI, Black Box Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, Wiesbaden 2019, 43 f.; LEONID GUGGENBERGER, Einsatz künstlicher Intelligenz in der Verwaltung, NVwZ 2019, 844 ff., 845; RAINER KESSLER/JUTTA SONJA OBERLIN, Künstliche Intelligenz: Quo Vadis?, CB 2020, 89 ff., 92. Aus den Medien nur FLORENT THOUVENIN/NADIA BRAUN BINDER/MELINDA LOHMANN, Die Schweiz braucht Regeln für den Einsatz von künstlicher Intelligenz, NZZ vom 27.01.2021, <https://www.nzz.ch/meinung/die-schweiz-braucht-regeln-fuer-den-einsatz-von-kuenstlicher-intelligenz-ld.1595509>, abgerufen am 24. August 2022.

³⁶ Vertieft BRAUN BINDER, Vorprojekt IP6.4 (Fn. 17), 24 ff.

³⁷ Z. B. Eidgenössisches Departement für auswärtige Angelegenheiten EDA, Künstliche Intelligenz und internationales Regelwerk, Bericht an den Bundesrat, Bern, 13. April 2022; Der Bundesrat, Leitlinien «Künstliche Intelligenz» für den Bund, Bern, 25. November 2020; Eidgenössisches Departement für Wirtschaft, Bildung und Forschung WBF/Staatssekretariat für Bildung Forschung und Innovation SBFI, Herausforderungen der künstlichen Intelligenz, Bern, 13. Dezember 2019, <https://www.sbf.admin.ch/sbfi/de/home/bfi-politik/bfi-2021-2024/transversale-themen/digitalisierung-bfi/kuenstliche-intelligenz.html>, abgerufen am 6. September 2022. LUHMANN (Fn. 1), 18, stellt dies bereits für die Anfänge der Verwaltungsdigitalisierung mittels elektronischer Datenverarbeitung fest.

³⁸ JAN ETSCHIED/JÖRN VON LUKE/FELIX STROH, Künstliche Intelligenz in der öffentlichen Verwaltung, Stuttgart 2020, 27, 37, 45.

³⁹ Diese Annahme wird durch das Einführen des Begriffs der automatisierten Einzelentscheidung im Rahmen der DSGVO-Revision in verschiedenen abgabe- und versicherungsrechtlichen Erlassen verstärkt. Siehe dazu BRAUN BINDER, Dystopie (Fn. 8), 270, 274.

⁴⁰ KIENER/RÜTSCHKE/KUHN (Fn. 3), N 878.

2.2. Art. 21 nDSG als allgemeine Zulässigkeitsnorm im Verwaltungsalltag?

Fraglich ist, ob der Gesetzgeber mit Art. 21 Abs. 1 nDSG eine allgemeine Zulässigkeitsnorm für automatisierte Einzelentscheidungen in der Schweiz schaffen wollte. Denn im Rahmen der DSGVO-Revision werden weitere Gesetze angepasst und um den Begriff der automatisierten Einzelentscheidung ergänzt. Im Bereich des Zollveranlagungsverfahrens sowie der Festsetzung des Tabaksteuer-, Biersteuer-, Mineralölsteuer- sowie Schwerverkehrsabgabebetrags werden automatisierte Einzelentscheidungen vorgesehen, indem auf Art. 21 Abs. 1 nDSG verwiesen wird. Das nDSG dient mithin als Ausgangsnorm für die Verweise in den genannten (Verfahrens-)Gesetzen. Im Gegensatz zur europäischen Regelung, die mit Art. 22 Abs. 1 Datenschutzgrundverordnung (DSGVO) ein allgemeines Automationsverbot – mit Erlaubnisvorbehalt⁴¹ – geschaffen hat und somit deren generelle Zulässigkeit verneint,⁴² scheint das Schweizer Recht nun von der weitgehend voraussetzungslosen Zulässigkeit solcher Entscheidungen auszugehen. Eine solche Zulässigkeitsabsicht lässt sich aber zumindest der Botschaft zum nDSG nicht entnehmen. Dagegen spricht auch das Legalitätsprinzip gemäss Art. 5 Abs. 1 Bundesverfassung (BV⁴³), das neben dem Erfordernis der formell-gesetzlichen Grundlage (Normstufe) für wichtige rechtsetzende Bestimmungen (Art. 164 Abs. 1 BV) auch genügend Bestimmtheit (Normdichte) derselben fordert.⁴⁴ Auch wenn das nDSG ein Gesetz im formellen Sinn darstellt; dessen Bestimmtheit ist mit Blick auf den Erlass vollautomatisierter Verfügungen durch den Staat wohl ungenügend.⁴⁵ Denn es äussert sich nicht zu den verschiedenen Anwendungsfeldern automatisierter Verfahren, sondern geht ganz einfach von deren Existenz aus.⁴⁶ Im Zuge der Datenschutzgesetzrevision wird die automatisierte Einzelentscheidung als *vollständige* Verlagerung der menschlichen Entscheidungsfindung auf die Maschine verstanden. Fälle der Teilautomation, insbesondere der automatisierten Ent-

⁴¹ Dazu ROMY DAEDELLOW, Wenn Algorithmen (unfair) über Menschen entscheiden..., Jusletter 26. November 2018, N 23 ff.

⁴² Vertieft statt vieler CHRISTIAN ERNST, Algorithmische Entscheidungsfindung und personenbezogene Daten, JZ 2017, 1026 ff.

⁴³ Bundesverfassung der Schweizerischen Eidgenossenschaft vom 18. April 1999 (SR 101).

⁴⁴ SGK BV-SCHINDLER/TSCHUMI, Art. 5 N 33 ff.; BSK BV-WYTENBACH/WYSS, Art. 164 N 11.

⁴⁵ BRAUN BINDER, Dystopie (Fn. 8), 260 f. Allgemein zu den Anforderungen aus dem Legalitätsprinzip bezüglich der Normierung automatisierter Verwaltungsverfahren, THURNHERR (Fn. 33), 150 ff.

⁴⁶ BRAUN BINDER, Dystopie (Fn. 8), 261; THURNHERR (Fn. 33), 152.

scheidungsunterstützung, werden vom nDSG nicht erfasst.⁴⁷ Dies wird indes dem Verwaltungsalltag nicht gerecht.⁴⁸ Die Vollautomation dürfte in aktuellen Digitalisierungsprojekten nachrangig sein; die Teilautomation hingegen häufiger vorkommen. Schliesslich ist anzumerken, dass der Begriff der automatisierten Einzelentscheidung nach Art. 21 Abs. 1 nDSG nur im Zusammenhang mit der Bearbeitung von Personendaten von Relevanz ist (Art. 2 Abs. 1 i.V.m. Art. 5 lit. a nDSG). Indes liegen nicht allen Verfügungen immer zwingend Personendaten zugrunde. Das betrifft vor allem Allgemeinverfügungen, die grundsätzlich jedoch gleich behandelt werden wie eine Individualverfügung nach Art. 5 Abs. 1 VwVG,⁴⁹ sie somit theoretisch auch der automatisierten Entscheidungsfindung zugänglich sind. Alle diese Argumente führen zum Schluss, dass mit Art. 21 nDSG keine hinreichende gesetzliche Grundlage für den Erlass vollautomatisierter Verfügungen in der Bundesverwaltung vorliegt.

III. Automatisierte Einzelentscheidung als Verfügung im Sinne von Art. 5 VwVG?

Bis hierhin kann festgehalten werden, dass ohne themenspezifische Rechtsgrundlage der vollautomatisierte Erlass einer Verfügung auf Bundesebene nicht zulässig ist. Oder anders formuliert: es bedürfte wohl der Anpassung der einschlägigen fachspezifischen Gesetze. Unabhängig davon stellt sich zudem die Frage, ob im VwVG als allgemeiner Verfahrensordnung Regelungsbedarf besteht. Die Beantwortung dieser Frage hängt u. a. davon ab, ob automatisierte Einzelentscheidungen ohne Weiteres als Verfügungen im Sinne von Art. 5 VwVG qualifiziert werden können. Auf diese Frage soll im Folgenden eingegangen werden. Zusätzlicher Regulierungsbedarf im VwVG – etwa zur Sicherstellung des rechtlichen Gehörs – wird an dieser Stelle nicht thematisiert.

1. Traditioneller Verfügungsbegriff

Mit Erlass des VwVG im Jahr 1968 wurde erstmalig für das Bundesrecht definiert, was eine Verfügung und insbesondere, was deren Rechtsnatur ist. Gemäss der Legaldefinition in Art. 5 Abs. 1 VwVG sind Verfügungen sich auf öffentliches Recht stützende Anordnungen der Behörden im Einzelfall, durch die Rechte und Pflichten rechtsgestaltend oder feststellend einseitig geregelt werden. Die Verfügung ist somit das Ziel des erstinstanzlichen Verwaltungs-

⁴⁷ Bereits BRAUN BINDER, Perspektive (Fn. 4), 33.

⁴⁸ Oben [II.2.1](#).

⁴⁹ BRAUN BINDER, Dystopie (Fn. 8), 270.

verfahrens.⁵⁰ Der Begriff ist im öffentlichen Recht derart zentral, dass er sich vom VwVG selbst abhebt und sich gewissermassen als Rechtsbegriff verselbstständigt hat.⁵¹ Das Bundesgericht präzisiert in seiner Rechtsprechung, dass es sich um «autoritative, einseitige, individuell-konkrete Anordnungen der Behörde, die in Anwendung von Verwaltungsrecht ergangen, auf Rechtswirkungen ausgerichtet sowie verbindlich und erzwingbar sind»⁵² handelt. Es lassen sich schliesslich verschiedene Begriffselemente erkennen:⁵³ Verfügungen sind eine behördliche Anordnung, stützen sich auf öffentliches Recht (des Bundes), betreffen eine bestimmte Person und einen bestimmten Sachverhalt, regeln rechtsgestaltend oder feststellend Rechte und Pflichten und sind verbindlich und erzwingbar, wodurch sie ohne weitere Konkretisierung vollstreckt werden können.

2. Diskrepanz: behördlicher Willensbildungsprozess...

Die automatisierte Einzelentscheidung weist viele Begriffselemente der Verfügung auf. Ihnen ist gemeinsam, dass sie sich auf öffentliches Recht (des Bundes) stützen, individuell-konkret ergehen, Rechte und Pflichten rechtsgestaltend oder feststellend regeln sowie verbindlich und erzwingbar, mithin vollstreckbar sind. Nicht klar ist jedoch die Frage nach dem Urheber beziehungsweise der Urheberin der Verfügung, mithin ob eine vollständig automatisierte Einzelentscheidung einer natürlichen Person, die hinter einer Behörde steht, zugerechnet werden kann.⁵⁴ Bis anhin ging das allgemeine Verwaltungsrecht davon aus, dass beim Erlass einer Verfügung innerhalb der Behörde eine Willensbildung stattfindet, die Verfügung entsprechend eine Willenserklärung ist.⁵⁵ Dieser Willensbil-

⁵⁰ KIENER/RÜTSCHKE/KUHN (Fn. 3), N 5 f.

⁵¹ MÜLLER (Fn. 9), Art. 5 N 3.

⁵² BGE 141 II 233, 235, E. 3.1; 135 II 38, 44 f., E. 4.3.

⁵³ Je nachdem werden fünf bis sieben Begriffselemente genannt, vgl. etwa Botschaft des Bundesrates an die Bundesversammlung über das Verwaltungsverfahren vom 24. September 1965, BBl 1965 1348 ff., 1362 ff.; MÜLLER (Fn. 9), Art. 5 N 28; aus der Basisliteratur nur PIERRE TSCHANNEN/ULRICH ZIMMERLI/MARKUS MÜLLER, Allgemeines Verwaltungsrecht, 4. Aufl., Bern 2014, § 28 N 16 ff.; vertieft ROLF HEINRICH HALTNER, Begriff und Arten der Verfügung im Verwaltungsverfahren des Bundes (Artikel 5 VwVG), Diss Zürich 1979, Zürich 1979, 21 ff.

⁵⁴ BRAUN BINDER, Dystopie (Fn. 8), 272; GLASER, Einflüsse (Fn. 8), 187 f.; THURNHERR (Fn. 33), 135.

⁵⁵ Zur behördlichen Anordnung als Willenserklärung HALTNER (Fn. 53), 23 ff.; KIENER/RÜTSCHKE/KUHN (Fn. 3), N 355; BERNHARD RÜTSCHKE, Was sind öffentliche Rechte und Pflichten? Überlegungen zum Rechtscharakter der Verfügung, in: Markus Rüssli/Julia Hänni/Reto Häggi Furrer (Hrsg.), Staats- und Verwaltungsrecht auf vier Ebenen, Festschrift für Tobias Jaag, Zürich/Basel/Genf 2012, 69 ff., 77; BERNHARD WALDMANN, Vom Umgang mit or-

dungsprozess beurteilt sich für die jeweiligen Automatisierungsstufen – Teilautomation und Vollautomation – unterschiedlich.

3. ... bei teilautomatisierter Einzelentscheidung?

Bei der Teilautomation können, wie bereits erläutert,⁵⁶ zwei Konstellationen differenziert werden. Bei der Entscheidungsunterstützung liegt der finale Entscheid weiterhin in menschlicher Hand. Die entscheidende natürliche Person kann sich zwar an Empfehlungen des Algorithmus orientieren, letztendlich gibt sie jedoch «ihren» Willen – beziehungsweise denjenigen der gesetzgebenden Gewalt⁵⁷ – kund. Die menschliche Komponente schlägt sich in der eigentlichen Entscheidung nieder, weshalb diese relativ problemlos einer Behörde zugerechnet werden kann und unter den Verfügungsbegriff nach Art. 5 VwVG fällt. Eruiert oder prüft die verantwortliche natürliche Person die Sachverhaltsdaten und setzt die Automatisierung erst für das Fällen des finalen Entscheides ein, liegt gewissermassen die umgekehrte Lage zur Entscheidungsunterstützung vor. Allerdings ist auch in dieser Konstellation eine menschliche Komponente enthalten. Denn neben der Programmierung des den Entscheid fallenden Algorithmus kann die Entscheidung durch das Eruiieren der Sachverhaltsdaten gewissermassen präjudiziert werden.⁵⁸

4. ... bei vollautomatisierter Einzelentscheidung?

Der ursprüngliche Verfügungsbegriff basiert auf der Annahme, die Anordnung würde durch eine natürliche Person getroffen.⁵⁹ Zwar schliesst der Gesetzeswortlaut von Art. 5 VwVG nicht aus, dass eine Verfügung auch ohne jegliches menschliche Zutun ergehen kann.⁶⁰ Daraus folgt aber noch nicht, dass eine vollautomatisierte Einzelentscheidung einem Willensbildungsprozess zugänglich ist. Die Willensbildung ist nach einem traditionellen Verständnis natürlichen Personen vorbehalten.⁶¹ Eine Willenserklärung ist demnach von zwei

ganisatorischen, innerdienstlichen und anderen Anordnungen ohne Verfügungscharakter, ZSR 2014 I, 489 ff., 492. Im Umkehrschluss auch Botschaft VwVG (Fn. 53), BBl 1965 1364.

⁵⁶ Oben [II.1.1](#).

⁵⁷ HALTNER (Fn. 53), 29.

⁵⁸ Für Deutschland bereits BRAUN BINDER, Gesetzesentwurf (Fn. 4), 963 f.

⁵⁹ Botschaft VwVG (Fn. 53), BBl 1965 1362.

⁶⁰ BRAUN BINDER, Dystopie (Fn. 8), 272.

⁶¹ Ausführungen zu Willenserklärungen und dem Willensbildungsprozess existieren insb. in der privatrechtlichen Lehre und Literatur zum Rechtsgeschäft, vgl. statt vieler PETER GAUCH et al., Schweizerisches Obligationenrecht, Allgemeiner Teil ohne ausservertragliches Haft-

Elementen geprägt: dem inneren Verpflichtungswillen und dem äusserlich erkennbaren Erklärungswillen.⁶² Im öffentlichen Recht wird der Willensbegriff jedoch bereits durch das Legalitätsprinzip eingeschränkt: Der Wille der natürlichen Personen, die hinter einer Behörde stehen, wird durch die gesetzlichen Grundlagen geleitet.⁶³ Der innere Verpflichtungswille ist also ein objektiv-gesetzlicher Wille. Durch die einzelfallweise Anwendung der Normen sind jedoch auch die natürlichen Personen in der Verwaltung – in beschränkter Weise – einem Willensbildungsprozess unterworfen.⁶⁴ Vollautomatisierte Verwaltungsverfahren sind dagegen von ihrer menschlichen Unabhängigkeit geprägt.⁶⁵ Die interne Willensbildung kann somit naturgemäss nicht durch eine natürliche Person erfolgen. Fraglich bleibt, ob nicht zumindest der äusserlich erkennbare Erklärungswille der Behörde zugerechnet werden könnte. Dies hängt massgeblich davon ab, ob eine automatisierte Einzelentscheidung einer Informations- und/oder Kennzeichnungspflicht im Sinne von Art. 21 Abs. 1 und Abs. 4 nDSG unterliegt. Diese Kundgebungspflichten führen nämlich dazu, dass die Behörde klar deklariert, dass die automatisierte Einzelentscheidung nicht durch eine natürliche Person gefällt wurde und damit zugleich klärt, dass es keinen Willen einer natürlichen Person gibt, der äusserlich erkennbar gemacht werden könnte. Das Argument, dass mit dem Einführen von automatisierten Einzelentscheidungen der Willensbildungsprozess vorverlagert werde – etwa auf den Zeitpunkt der Anschaffung eines Systems,⁶⁶ das automatisierte Einzelentscheidungen fällen kann oder auf den Zeitpunkt der Programmierung desselben⁶⁷ – verkennt, dass für den Willensbildungsprozess eine Abschlussentscheidung notwendig ist. Dies folgt aus der hier zugrunde gelegten Definition der Willenserklärung.

pfllichtrecht, Band I, 10. Aufl., Zürich/Basel/Genf 2014, N 168 ff. Allerdings wird dies bereits in den Anfängen der Verwaltungsrechtswissenschaft auf behördliche Anordnungen übertragen, siehe z. B. BRUNO BEETSCHEN, Die materielle Rechtskraft von Verwaltungsverfügungen, Diss. Zürich 1923, Zürich 1923, 15. Diese Gedanken wurden auch in der neueren verwaltungsrechtlichen Literatur aufgenommen, vgl. JÜRIG BICKEL, Auslegung von Verwaltungsakten, Diss. Freiburg 2013, Bern 2014, § 2 N 7.

⁶² Statt vieler GAUCH et al. (Fn. 61), N 168 ff.

⁶³ BICKEL (Fn. 61), § 2 N 9.

⁶⁴ BULL (Fn. 4), 415.

⁶⁵ Oben [II.1.1](#).

⁶⁶ M.w.H. zur deutschen Literatur BRAUN BINDER, Gesetzesentwurf (Fn. 4), 963. Dies kann auch darauf zurückgeführt werden, dass bereits mit der Beschaffung des Systems «grundlegende Weichenstellungen für einen verantwortungsvollen KI-Einsatz» getroffen werden, CHRISTEN et. al (Fn. 13), 279.

⁶⁷ M.w.H. zur deutschen Literatur BRAUN BINDER, Gesetzesentwurf (Fn. 4), 963.

IV. Fazit

Es bestehen unterschiedliche Möglichkeiten, eine automatisierte Einzelentscheidung dogmatisch als «menschliche» Willenserklärung aufzufassen. Angesichts der fehlenden Willensbildung bleibt jedoch unklar, ob eine automatisierte Einzelentscheidung tatsächlich eine Verfügung sein kann. Der geltende Wortlaut von Art. 5 VwVG schliesst zwar nicht aus, dass teil- und vollautomatisierte Einzelentscheide im Verwaltungsverfahren ergehen können.⁶⁸ Und auch in der Praxis ist es kaum vorstellbar, dass automatisierte Einzelentscheidungen nicht als Verfügung im Sinne von Art. 5 Abs. 1 VwVG bezeichnet würden, da die Verwaltung ansonsten nahezu handlungsunfähig würde⁶⁹ und den betroffenen Personen kein Rechtsschutz gewährt werden könnte⁷⁰. Der Verfügungsbegriff ist im öffentlichen Recht jedoch derart wichtig, dass eine Neuinterpretation kaum angestrebt werden kann.⁷¹ Eine gesetzliche Klarstellung drängt sich geradezu auf, um Rechtssicherheit sowohl für die erlassende Behörde als auch für die betroffene Person sicherzustellen. Ausserdem trägt eine gesetzliche Klärung zur Transparenz des Verwaltungshandelns bei und verhindert ein Zersplittern der Rechtsordnung.⁷² Grundsätzliche Diskussionen über die Natur von teil- und insbesondere vollautomatisierten Einzelentscheidungen würden dadurch vermieden. Zwar wäre auch denkbar, die Klärung jeweils nur in den bereichsspezifischen Gesetzen herbeizuführen, in welchen die Rechtsgrundlage für den Erlass automatisierter Einzelentscheide verankert wird. Gegen diese Herangehensweise spricht – neben der Verhinderung der Zersplitterung der Rechtsordnung – auch die Tatsache, dass im VwVG neben der Klärung des Verfügungsbegriffs weitere Konkretisierungen notwendig wären, etwa zur Sicherstellung des rechtlichen Gehörs.⁷³ Mit Blick auf die ti-

⁶⁸ BRAUN BINDER, *Dystopie* (Fn. 8), 272.

⁶⁹ Die Verfügung ist das wichtigste Handlungsinstrument der Verwaltung, siehe Fn. 3.

⁷⁰ Erst die Verfügung eröffnet den Rechtsweg, da sie Anfechtungsobjekt der Beschwerde bildet, statt vieler KIENER/RÜTSCHKE/KUHN (Fn. 3), N 338.

⁷¹ GLASER, *Einflüsse* (Fn. 8), 188.

⁷² Dem Leitgedanken des VwVG folgend, um das «bunte Mosaik aus zahllosen Bestimmungen des Spezialverwaltungsrechts über das Verwaltungsverfahren» (Botschaft VwVG [Fn. 53], BBl 1965 1351) zu überwinden. Dies bereits für die deutsche Rechtslage NADJA BRAUN BINDER, *Vollständig automatisierter Erlass eines Verwaltungsaktes und Bekanntgabe über Behördenportale*, DÖV 2016, 891 ff., 891.

⁷³ Zur Beeinflussung der Verwaltungsverfahrensprinzipien durch automatisierte Einzelentscheidungen BRAUN BINDER, *Dystopie* (Fn. 8), 273; NADJA BRAUN BINDER, *Der Untersuchungsgrundsatz als Herausforderung vollautomatisierter Verfahren*, ZSIS 2020, 27 ff., 31 ff.; THURNHERR (Fn. 33), 145 ff.; CATHERINE REITER, *Künstliche Intelligenz im Verwaltungsver-*

telgebende Ausgangsfrage kann abschliessend festgehalten werden, dass eine automatisierte Einzelentscheidung dem Begriff der Verfügung sehr nahe ist, sie sind aber nicht deckungsgleich – eben *same same but different*.

fahren, AJP 2022, 984 ff., 989 ff.; spezifisch zum rechtlichen Gehör RENÉ WIEDERKEHR/
CHRISTIAN MEYER, Schranken, Ausnahmen und Relativierungen des rechtlichen Gehörs: Insbesondere mit Blick auf automatisierte Verwaltungsverfahren, AJP 2022, S. 1092 ff., 1102 f.

White Paper: Transparenz durch öffentliches Verzeichnis

Nadja Braun Binder / Liliane Obrecht

Erstveröffentlichung: https://ius.unibas.ch/fileadmin/user_upload/ius/03_Forschung/10_e-PIAF/WhitePaper_Verzeichnis_Juni24.pdf

Der Einsatz algorithmischer Systeme kann dazu führen, dass staatliche Entscheidungen nicht mehr ohne Weiteres nachvollziehbar sind. Zur Erhöhung des Vertrauens der Bevölkerung in solche Entscheidungen und zur Sicherstellung von externen Kontrollmöglichkeiten sollte allgemein ersichtlich sein, wann ein algorithmisches System für bestimmte Entscheidungen eingesetzt wird. Mit einem öffentlichen Verzeichnis über den Einsatz algorithmischer Systeme kann die notwendige Transparenz hergestellt werden.

Inhalt

I. Problemstellung	107
II. Algorithmische Systeme und Künstliche Intelligenz	108
III. Transparenzbegriff	109
IV. Lösungsansatz ADMS-Verzeichnis	110
V. Rechtliche Erfassung	111

I. Problemstellung

In jüngster Zeit hat das Thema der Herstellung von Transparenz über den Algorithmeinsatz in öffentlichen Verwaltungen erhöhte Aufmerksamkeit erfahren. So wurden Ende 2021 und im Mai 2023 im Nationalrat zwei Vorstösse eingereicht, mit welchen der Bundesrat aufgefordert wird, ein öffentliches Verzeichnis aller in der Bundesverwaltung eingesetzten algorithmenbasierten, automatisierten Entscheidungssysteme anzufertigen (Motion 21.4508 Schlatter; Postulat 23.3566 Schlatter). Auch der Zürcher Kantonsrat hat im April 2022 ein Postulat überwiesen, in welchem der Regierungsrat gebeten wird, ein Verzeichnis zur Erfassung der eingesetzten automatisierten Entscheidungssysteme zu erstellen (Postulat 9/2022 Yuste/Moser/Mäder/Willi/Zahler). Mitt-

lerweile liegt ein Gesetzesentwurf im kantonalen Informations- und Datenschutzgesetz (IDG-ZH) vor, der als Grundlage für die Pflicht zur Führung eines solchen Verzeichnisses dient (§ 13 Abs. 3 revidiertes IDG-ZH). Im Kanton Basel-Stadt wurde im Oktober 2023 ein Anzug betreffend Transparenz über den Einsatz von algorithmischen Systemen in der Verwaltung eingereicht (Anzug 23.5474 der Justiz-, Sicherheits- und Sportkommission). In allen genannten parlamentarischen Vorstössen geht es um algorithmische Systeme, die unter dem Gesichtspunkt der Nachvollziehbarkeit der Entscheidungsfindung problematisch erscheinen.

Die Herstellung von Transparenz über den staatlichen Algorithmeinsatz stellt ein Instrument zur Sicherstellung einer **rechtsstaatskonformen und diskriminierungsfreien digitalen Transformation der öffentlichen Verwaltung** dar. Es ist oft unklar, wo algorithmenbasierte Systeme überhaupt eingesetzt werden. Da der Staat seiner Bevölkerung in der Regel hoheitlich gegenübertritt, bedarf es gewisser Mechanismen, um staatliches Handeln kontrollierbar zu machen. Zusätzlich zu allenfalls existierenden bzw. noch zu schaffenden individuellen Kontrollmöglichkeiten im Einzelfall für unmittelbar Betroffene, soll insbesondere die Zivilgesellschaft die Möglichkeit erhalten, den staatlichen Algorithmeinsatz kritisch zu begleiten.

Dadurch kann die **Vertrauensbeziehung** zwischen dem Staat und der Bevölkerung auch bei zunehmender Nutzung digitaler Technologien erhalten und sogar gestärkt werden. Transparenz kann die Vertrauenswürdigkeit fördern, indem die Bevölkerung die Möglichkeit erhält, sich möglichst umfassend zu informieren. Dadurch wird eine zusätzliche Motivation für **rationales Staatshandeln** sowie **behördliche Selbstkontrolle** geschaffen.

Transparenz über den Einsatz algorithmischer Systeme kann ausserdem dabei helfen, allfällige **Diskriminierungsprobleme** sichtbar zu machen, indem durch Vergleich verschiedener mit demselben System bearbeiteter Fälle erkennbar wird, ob eine Entscheidung an ein verpöntes Merkmal (siehe Art. 8 Abs. 2 BV, z. B. Herkunft, Geschlecht) anknüpft.

II. Algorithmische Systeme und Künstliche Intelligenz

Im vorliegenden Kontext liegt der Fokus auf **algorithmischen Systemen**, die aufgrund erschwerter Nachvollziehbarkeit problembehaftet sein können. Dazu gehören Systeme, die **ohne Mitwirkung einer natürlichen Person** funktionieren und z. B. zum (vollautomatisierten) Erlass sogenannter automatisierter Einzelentscheidungen genutzt werden. Auch **KI-Systeme** zählen dazu, die

insbesondere aufgrund der erschwerten Nachvollziehbarkeit des Zustandekommens ihrer Outputs eine Herausforderung darstellen können. Bilden die Ausgaben von KI-Systemen eine Grundlage für menschliche Entscheidungen (Entscheidungsunterstützung), die sich direkt oder indirekt auf die Rechte und Pflichten natürlicher Personen auswirken, oder handelt es sich um ein algorithmisches System, das selbst eine Entscheidung trifft, die sich direkt oder indirekt auf die Rechte und Pflichten von Personen auswirkt, so können sie als Systeme zur algorithmischen Entscheidungsfindung («Algorithmic Decision Making System», ADMS) zusammengefasst werden. Im Folgenden wird daher von einem **ADMS-Verzeichnis** gesprochen.

III. Transparenzbegriff

Zum Begriff der Transparenz findet sich im Bundesrecht **keine Legaldefinition**. Es existiert auch keine andere einheitliche Definition des Begriffs. Dennoch durchzieht der Transparenzbegriff die gesamte **Bundesverfassung** (BV, SR 101) und kommt in unterschiedlichen Bestimmungen zum Ausdruck. So kann er beispielsweise als Element der **Rechtsstaatlichkeit** (Art. 5 BV) verstanden oder als Ausfluss grundrechtlicher **Verfahrensgarantien** (siehe White Paper «Transparenz durch Begründung von Verfügungen».) eingeordnet werden. Ein weiterer verfassungsrechtlicher Anknüpfungspunkt stellt der **Informationsauftrag** des Bundesrates nach Art. 180 Abs. 2 BV dar. Gemeinsam ist diesen Anknüpfungspunkten, dass sie insbesondere der Stärkung der Vertrauensbeziehung zwischen Staat und Bevölkerung sowie der Legitimität staatlichen Handelns dienen.

Auf **Gesetzesstufe** wird Transparenz häufig im Rahmen des Öffentlichkeitsprinzips der öffentlichen Verwaltung diskutiert. Zentrales Anliegen des seit 2004 auf Bundesebene geltenden **Öffentlichkeitsgesetzes** (BGÖ, SR 152.3) ist die Verbesserung der Kontrolle der Bevölkerung über die Verwaltung, wodurch das Vertrauen in diese gestärkt werden soll. Das BGÖ fördert die Transparenz der Verwaltung, indem jede interessierte Person in die Dokumente der Bundesbehörden Einsicht nehmen darf. Allerdings erfasst das BGÖ lediglich die passive Informationstätigkeit der Verwaltung, eine aktive Information durch Behörden ist nicht vorgesehen.

Die aktive Informationstätigkeit findet ihre Grundlage insbesondere in Art. 10 des **Regierungs- und Verwaltungsorganisationsgesetzes** (RVOG, SR 170.010), der den Informationsauftrag gemäss Art. 180 Abs. 2 BV konkretisiert. Der Informationspflicht unterliegen die Regierung sowie die Bundesverwaltung, wo-

bei Umfang und Inhalt der Information vom jeweiligen Ort und der jeweiligen Zeit abhängig sind. Ausschlaggebend sind die Bedürfnisse der Bevölkerung sowie des Staates selbst.

Auch das **Datenschutzgesetz** (DSG, SR 235.1) enthält verschiedene Vorgaben zur Herstellung von Transparenz. Dabei geht es um die Herstellung von Transparenz bei der Bearbeitung von Personendaten. Diesem Zweck dienen etwa der Grundsatz, dass die Beschaffung von Personendaten und der Zweck, zu dem sie beschafft werden, erkennbar sein müssen (Art. 6 Abs. 3 DSG), die Pflicht der Verantwortlichen, die betroffene Person über die Datenbeschaffung zu informieren (Art. 19 DSG), die Kennzeichnungspflicht bei einem automatisierten Einzelentscheid durch ein Bundesorgan (Art. 21 Abs. 4 DSG) oder das Verzeichnis der Datenbearbeitungstätigkeiten von Bundesorganen (Art. 56 DSG).

IV. Lösungsansatz ADMS-Verzeichnis

Von verschiedenen Seiten wird die Schaffung eines öffentlich zugänglichen ADMS-Verzeichnisses der in der öffentlichen Verwaltung eingesetzten algorithmischen Systeme gefordert, um den Transparenzanforderungen nachzukommen und gesellschaftliche Kontrolle zu konkretisieren. Diese Überlegungen sind auch aus rechtswissenschaftlicher Perspektive sinnvoll.

Das Führen öffentlicher Verzeichnisse ist aus rechtlicher Sicht nichts Neues. Seit jeher werden öffentliche Verzeichnisse in der Rechtswissenschaft als **Mittel zur Transparenzherstellung** und somit zur Sicherstellung von Rechtsstaatlichkeit betrachtet. Ein Verzeichnis ist ein geordnetes Informationssystem, das Orientierung über einen bestimmten Sachbereich bietet. Es ist öffentlich, sobald jede Person zu einem unspezifischen Zeitpunkt und ohne spezifisches Interesse in dieses Einsicht nehmen kann.

Ein ADMS-Verzeichnis verfolgt mehrere Ziele. Zentral ist die **Pflege des Vertrauensverhältnisses** zwischen dem Staat und der Bevölkerung. Da der Staat hoheitlich, d. h. einseitig, über Rechte und Pflichten der Bevölkerung entscheiden kann, entsteht ein Machtungleichgewicht, das potenziell zu Machtmissbrauch führen kann. Indem der Staat jedoch offenlegt, wo er welche ADMS einsetzt, schafft er Kontrollmechanismen und fördert die Akzeptanz solcher Technologien in der Bevölkerung. Darüber hinaus kann ein Verzeichnis auch einer möglichen Intransparenz über die Funktionsweise von ADMS entgegenwirken, indem es diese näher erläutert. Von besonderer Bedeutung ist dabei die Schaffung einer **adressatengerechten Transparenz**. So sollten ne-

ben leicht verständlichen Informationen für Laien auch weiterführende Dokumentationen für Expertinnen und Experten zur Verfügung gestellt werden. Schliesslich wird die Verwaltung selbst durch die Dokumentation entlastet, da sich Interessierte direkt selbst über die eingesetzten Systeme informieren können und kein einzelfallweiser verwaltungsinterner Aufwand für die Auskunftserteilung im Einzelfall anfällt.

V. Rechtliche Erfassung

Behörden dürfen aufgrund ihres Informationsauftrages – wie er z. B. auf Bundesebene in Art. 10 RVOG verankert ist – bereits heute ein ADMS-Verzeichnis führen. Die Nachteile von solchen freiwillig geführten ADMS-Verzeichnissen sind, dass sie keinem einheitlichen Aufbau folgen und auch nicht an zentraler Stelle zugänglich gemacht werden müssen. Darüber hinaus ist bei einer rein freiwilligen Verzeichnisführung nicht klar, ob auch wirklich alle Systeme erfasst werden. Deshalb ist die **Pflicht zur Führung eines ADMS-Verzeichnisses** auf **Gesetzesstufe** zu verankern und in den Grundzügen festzulegen, welche Angaben ein solches Verzeichnis enthalten muss, wer das Verzeichnis zu verwalten und wer die Angaben einzutragen hat.

Für die Bundesebene kämen verschiedene bereits bestehende Gesetze als Grundlage infrage. Dazu zählen das DSG oder das am 1. Januar 2024 in Kraft getretene Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben (EMBAG, SR 172.019). Grundsätzlich wäre auch das RVOG ein möglicher Regelungsort. Das EMBAG ist jedoch spezifisch auf Fragen der Digitalisierung ausgerichtet und weist daher einen engeren Sachbezug für die Regelung eines ADMS-Verzeichnisses auf als das RVOG.

Entscheidend für die Frage der **rechtlichen Verankerung** eines ADMS-Verzeichnisses ist neben dem Sachbezug letztlich der Adressatenkreis. Während das **DSG** alle Bundesorgane erfasst, die eine öffentliche Aufgabe wahrnehmen (Art. 2 Abs. 1 lit. b i. V. m. Art. 5 lit. i DSG), weist das **EMBAG** einen eingeschränkten Adressatenkreis auf. Es richtet sich nur an die zentrale Bundesverwaltung. Zwar kann der Bundesrat auf Verordnungsstufe eine Ausweitung des Adressatenkreises vorsehen (Art. 2 Abs. 1, 2 und 3 EMBAG), doch sind damit in erster Linie punktuelle Ausweitungen gemeint und keine generelle Ausweitung auf alle Bundesorgane, auch wenn dies nur die Vorgaben für ein ADMS-Verzeichnis betreffen würde. Ziel eines ADMS-Verzeichnisses ist es, möglichst alle in der öffentlichen Verwaltung eingesetzten ADMS zu erfassen. **Zielfüh-**

rend erscheint daher die **Verankerung im DSG**. Ansprüche nach dem BGÖ würden durch die Regelung des ADMS-Verzeichnisses nicht berührt.

Konkretisierungen zu den einzelnen Elementen des Verzeichnisses können schliesslich auf **Verordnungsebene** vorgenommen werden und müssen nicht auf Gesetzesstufe verankert werden. Dazu zählen unter anderem die folgenden Aspekte:

– **Basisinformationen**

Als grundlegende Information dient erstens eine allgemeine **Übersicht**, die allen Personen in verständlicher Sprache Erklärungen über die Funktionsweise des algorithmischen Systems, die Art und Weise seines Einsatzes und dessen Auswirkungen liefert. Zweitens sind Informationen zum **verantwortlichen Organ** aufzuführen. Dieses stellt sicher, dass das System über seinen gesamten Lebenszyklus weiterentwickelt und gewartet wird. In diesem Rahmen können auch die an der Entwicklung des Systems beteiligten Personen sowie diejenigen, die das System betreiben, genannt werden. Wichtig ist jedoch, dass klar ersichtlich wird, wer die Verantwortung trägt. Die genannten Personen müssen kontaktiert werden können (beispielsweise via E-Mail oder Telefon). Drittens sind gegebenenfalls weitere **Referenzen** zu nennen, die weiterführende Informationen (auch in Bezug auf Details technischer Natur) und allfällige Prüfungsberichte liefern können.

– **Einsatzbereich des ADMS**

Das Verzeichnis muss Informationen über den Bereich, in dem das ADMS eingesetzt wird, enthalten. Dazu gehört auch die **Rechtsgrundlage**, die den Einsatz eines Systems legitimiert.

– **Daten**

Der **Zweck** der Bearbeitung von Personendaten ist anzugeben. Falls der Zweck aus der gesetzlichen Grundlage hinreichend hervorgeht, kann auf diese verwiesen werden. Ebenso soll das Verzeichnis Auskunft zur Art und Herkunft der bearbeiteten Sach- und Personendaten (**Datensets**) geben. Dabei müssen keine Einzeldaten veröffentlicht, sondern die einschlägigen Merkmale oder Kategorien genannt werden. Allfällige Geheimhaltungsinteressen sowie die Anonymisierung von Personendaten sind zu berücksichtigen. Auch die Art und Weise, wie das algorithmische System die Datenverarbeitung durchführt (**Funktionsweise**) und wie die Resultate erzielt werden, bedarf der Konkretisierung auf Verordnungsebene. Die Bearbeitungsmittel und die Logik des Algorithmus sind aufzu-

zeigen. Zudem müssen Informationen zur Interpretation und Messbarkeit des Systems aufgeführt werden. Dies kann etwa durch das Festhalten der Hauptmerkmale realisiert werden. Die Offenlegung des Quellcodes hingegen ist nicht zwingend erforderlich. Dieser könnte allerdings ausgewählten Expertinnen und Experten offenbart bzw. bei Bedarf bereitgestellt werden. Ist ein System im Einsatz, das mithilfe von **Trainings-, Validierungs- oder Testdaten** entwickelt wurde, muss ersichtlich sein, welche Daten für das Training, die Validierung bzw. das Testen genutzt wurden.

- **Algorithmus**
Ferner sind Informationen zu den eingesetzten Algorithmen notwendig, insbesondere darüber, ob es sich um **regelbasierte** oder **datenbasierte** («Machine Learning») Algorithmen oder auch um eine Kombination aus beiden handelt.
- **Auswirkungen auf die Bevölkerung**
Im öffentlichen Verzeichnis sollten Methoden beschrieben werden, die bei der Implementierung und beim Einsatz des algorithmischen Systems angewendet werden, um potenzielle Diskriminierungsrisiken zu adressieren und zu verhindern (**Antidiskriminierung**). Zusätzlich bedarf es bei der Offenlegung der Resultate einer **Folgenabschätzung**, in der die Risiken des jeweiligen algorithmischen Systems genannt werden. Dies bildet die Grundlage für die Festlegung geeigneter Massnahmen zum Schutz der Bevölkerung vor bzw. zur Bewältigung der erkannten Risiken.
- **Menschliche Aufsicht**
Die Behörde muss in der Lage sein, die eingesetzten Systeme durch Menschen zu beaufsichtigen. Die Ausgestaltung der menschlichen Aufsicht hängt vom Grad der Autonomie des algorithmischen Systems ab. Grundsätzlich sollte aus dem ADMS-Verzeichnis hervorgehen, inwieweit eine menschliche Aufsicht möglich und vorhanden ist. Darüber hinaus sind die für die menschliche Aufsicht erforderlichen Fähigkeiten zu beschreiben.
- **Metadaten**
Schliesslich gilt es, Informationen zu den Metadaten – etwa, wann ein ADMS im Verzeichnis erstmalig erfasst wurde oder Informationen überarbeitet wurden – zu nennen.

Rechtliche Erwägungen ausgehend von Forderungen nach öffentlichen Verzeichnissen

Nadja Braun Binder / Liliane Obrecht

Erstabdruck in: AJP/PJA 10/2024, 1069-1081

Im Bestreben, Transparenz über den staatlichen Einsatz algorithmischer Systeme herzustellen, fordern verschiedene Akteurinnen und Akteure die Schaffung von öffentlichen Verzeichnissen. Die Intention ist, das Vertrauen der Bevölkerung in staatliche Entscheidungen, die auf algorithmischen Systemen basieren, zu erhöhen und externe Kontrollmöglichkeiten sicherzustellen. Der vorliegende Beitrag geht aus rechtswissenschaftlicher Perspektive der Frage nach, ob solche Verzeichnisse notwendig sind, wo diese rechtlich verankert werden könnten und welche Inhalte ein derartiges Verzeichnis aufweisen könnte. Für die Schweiz betrachten es die Autorinnen als zielführend, eine formell-gesetzliche Grundlage im Datenschutzgesetz (DSG) zu schaffen und diese auf Verordnungsebene zu konkretisieren. Allerdings sollen nur Systeme in das Verzeichnis aufgenommen werden, die als Basis für Entscheidungen dienen, die sich direkt auf Rechte und Pflichten von natürlichen Personen auswirken (sog. Algorithmic Decision-Making Systems, ADMS).

Dans le but d'instaurer la transparence sur l'utilisation de systèmes algorithmiques par l'Etat, différents acteurs demandent la création de répertoires publics. L'intention est d'augmenter la confiance de la population dans les décisions étatiques basées sur des systèmes algorithmiques et de garantir des possibilités de contrôle externe. Le présent article examine, dans une perspective juridique, la question de savoir si de tels répertoires sont nécessaires, où ils pourraient être ancrés juridiquement et quels pourraient être les contenus d'un tel répertoire. Pour la Suisse, les auteures considèrent qu'il serait judicieux de créer une base légale formelle dans la loi sur la protection des données (LPD) et de la concrétiser au niveau de l'ordonnance. Toutefois, seuls les systèmes qui servent de base aux décisions ayant un impact direct sur les droits et les obligations des personnes physiques (appelés systèmes décisionnels algorithmiques, ADMS) doivent être inclus dans le répertoire.

Inhalt

I.	Einleitung	116
II.	Ziele eines Verzeichnisses und erfasste Systeme	120
1.	Terminologie	120
2.	Ziele eines ADMS-Verzeichnisses	121
2.1.	Schutz des Menschen und der Menschenwürde	121
2.2.	Kontrolle staatlichen Handelns	122
2.3.	Vertrauensbildung	123
2.4.	Dokumentation	124
3.	Erfasste Systeme	124
III.	Rechtliche Anknüpfungspunkte für Transparenzforderungen	125
1.	Verfassungsrechtliche Herleitung	125
2.	Transparenz zur Verhinderung von Diskriminierung	128
3.	Gesetzliche Anknüpfungspunkte	129
3.1.	Öffentlichkeitsprinzip	129
3.2.	Informationspflicht	131
3.3.	Datenschutz	132
4.	Zwischenfazit	133
IV.	Rechtliche Verankerung	134
1.	Regelungsort	135
2.	Regelungsgegenstand	137
2.1.	Grundlegende Informationen im Gesetz im formellen Sinn	137
2.2.	Konkretisierungen auf Verordnungsstufe	138
2.3.	Verständlichkeit und Zugänglichkeit der Informationen	139
V.	Schlussbemerkung	140

I. Einleitung

In den letzten Jahren hat das Thema der Schaffung von Transparenz über den Einsatz von Algorithmen in öffentlichen Verwaltungen in der Schweiz erhöhte Aufmerksamkeit erfahren. So wurden Ende 2021 und im Mai 2023 im Nationalrat zwei Vorstösse eingereicht, die den Bundesrat auffordern, ein öffentliches Verzeichnis aller in der Bundesverwaltung eingesetzten algorithmenbasierten Entscheidungssysteme vorzusehen.¹ Der Zürcher Kantonsrat hat im April 2022

¹ Motion 21.4508 Schlatter Marionna vom 16. Dezember 2021 betreffend ein öffentliches Verzeichnis der in der Verwaltung eingesetzten Algorithmen; Postulat 23.3566 Schlatter Marionna vom 4. Mai 2023 betreffend die Vollständigkeit des Transparenzregisters für den

ein Postulat überwiesen, in welchem der Regierungsrat angehalten wird, ein Verzeichnis zur Erfassung der eingesetzten algorithmischen Entscheidungssysteme zu erstellen.² Mittlerweile liegt der Entwurf für eine entsprechende Rechtsgrundlage im kantonalen Informations- und Datenschutzgesetz (IDG-ZH) vor.³ Im Kanton Basel-Stadt wurde im Oktober 2023 ein Anzug betreffend Transparenz über den Einsatz von algorithmischen Systemen in der Verwaltung eingereicht.⁴ Jüngst hat der Kanton Appenzell Innerrhoden ein Verzeichnis über die im öffentlichen Sektor eingesetzten algorithmischen Entscheidungssysteme veröffentlicht.⁵

Ferner findet sich die Forderung nach obligatorischen öffentlichen Verzeichnissen⁶ der in Schweizer Verwaltungen eingesetzten algorithmischen Systeme in wissenschaftlichen Artikeln⁷ und Studien⁸ sowie Beiträgen zivilgesellschaft-

Einsatz von künstlicher Intelligenz durch die Bundesverwaltung. Jüngst wurde auch die Ausweitung des Verzeichnisses auf bundesnahe Betriebe gefordert, Postulat 23.3565 Glättli Balthasar vom 4. Mai 2023.

² Postulat 9/2022 Yuste Nicola/Moser Arianne/Mäder Gabriel/Wilma Willi/Erika Zahler vom 17. Januar 2022 betreffend Transparenz über den Einsatz von künstlicher Intelligenz in der Verwaltung, das vom Kantonsrat genehmigt und am 25.4.2022 an den Regierungsrat überwiesen wurde.

³ § 13 Abs. 3 revidiertes IDG-ZH: «Es [das öffentliche Organ] führt ein öffentlich zugängliches Verzeichnis der von ihm verwendeten algorithmischen Entscheidungssysteme, die sich auf die Grundrechte von Personen auswirken können. Der Regierungsrat regelt die Einzelheiten für alle öffentlichen Organe in einer Verordnung.»

⁴ Anzug 23.5474 der Justiz-, Sicherheits- und Sportkommission vom 18.10.2023 betreffend Transparenz über den Einsatz von algorithmischen Systemen in der Verwaltung.

⁵ Kanton Appenzell Innerrhoden, Algorithmische Systeme, Internet: <https://www.ai.ch/the-men/staat-und-recht/digitale-verwaltung/verzeichnis-algorithmischer-systeme> (Abruf 7.8.2024). Ende Juli 2024 führte das Verzeichnis Informationen zu vier algorithmischen Systemen auf.

⁶ Verschiedentlich auch als «Register» bezeichnet; siehe etwa FLORENT THOUVENIN et al., Ein Rechtsrahmen für Künstliche Intelligenz, Positionspapier November 2021, 3, Internet: <https://www.itsl.uzh.ch/dam/jcr:4687cf40-85a8-47f6-88ea-7ba54ad31d34/dsi-strategy-lab-21-de.pdf> (Abruf 7.8.2024) (zit. Positionspapier 2021), 3.

⁷ THOUVENIN et al., Positionspapier 2021 (FN 6), 1 ff.; NADJA BRAUN BINDER et al., Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht, Jusletter 28.6.2021, N 8 ff. (zit. Handlungsbedarf).

⁸ NADJA BRAUN BINDER et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4 (zit. Studie KI 2021), 50, Internet: https://www.zh.ch/content/dam/zhweb/bilder-dokumente/the-men/politik-staat/kanton/digitale-verwaltung-und-e-government/projekte_digitale_transformation/ki_einsatz_in_der_verwaltung_2021.pdf (Abruf 7.8.2024); MARKUS

licher Organisationen⁹. Der Bundesrat erachtete es dagegen unter Verweis auf die freiwillige Datenbank des Kompetenznetzwerks Künstliche Intelligenz (KI, *Competence Network for Artificial Intelligence* – CNAI), in der KI-Projekte des Bundes dokumentiert werden, bislang nicht als notwendig, ein solches Verzeichnis verpflichtend vorzusehen.¹⁰

Auch mit Blick auf andere Rechtsordnungen zeigt sich eine gewisse Entwicklung hin zur Schaffung von Transparenz über den Einsatz von Algorithmen in öffentlichen Verwaltungen mithilfe von Verzeichnissen. So führten etwa Amsterdam¹¹ und Helsinki¹² im Jahr 2020 ein gemeinsam entwickeltes, freiwilliges Verzeichnis (sog. «Algorithmenregister») ein, um die Nutzung von Künstlicher Intelligenz in den beiden Städten transparenter zu machen.¹³ Im Jahr 2021 zog die Stadt Rotterdam nach und veröffentlichte ihr eigenes Algorithmenregister.¹⁴ Nach einem

CHRISTEN et al., Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz, TA-SWISS Publikationsreihe (Hrsg.): TA 72/2020, Zürich, 294 (zit. TA-SWISS Studie KI 2020); zudem auch FLORENT THOUVENIN et al., Towards Principles Regulation of Automated Decision-Making, Workshop Report vom 12.–14.9.2019 (zit. Workshop Report 2019), Internet: <https://www.itsl.uzh.ch/dam/jcr:edba006c-8452-4ffc-bbb0-fd8fde62d114/Workshop> (Abruf 7.8.2024).

⁹ Digitale Gesellschaft, Position der Digitalen Gesellschaft zur Regulierung von automatisierten Entscheidungssystemen, 21.2.2022, 19 f., Internet: <https://www.digitale-gesellschaft.ch/uploads/2022/02/Position-der-Digitalen-Gesellschaft-zur-Regulierung-von-automatisierten-Entscheidungssystemen-1.0.pdf> (Abruf 7.8.2024); AlgorithmWatch CH, Automatisierte Entscheidungssysteme im öffentlichen Sektor: einige Empfehlungen, 25.2.2022, Internet: <https://algorithmwatch.ch/de/adm-offentlichersektor-empfehlungen/> (Abruf 7.8.2024) (zit. Empfehlungen). Siehe auch Zivilgesellschaftliche Grundsatzanforderungen zur Regulierung von Künstlicher Intelligenz und automatisierten Entscheidungssystemen, 30.11.2023, 1, Internet: <https://www.digitale-gesellschaft.ch/2023/11/30/kein-abwarten-bei-der-regulierung-von-ki-und-adm-systemen-zivilgesellschaft-fordert/> (Abruf 7.8.2024).

¹⁰ Stellungnahme des Bundesrates vom 16.2.2022 zur Motion 21.4508 (FN 1). In der Projektdatenbank (Version 8, Stand am 29.5.2024) finden sich Informationen zu knapp 70 Systemen, siehe Internet: <https://cna1.swiss/dienstleistungen/projektdatenbank> (Abruf 7.8.2024).

¹¹ Zum Register siehe Internet: <https://algorithmerregister.amsterdam.nl/> (Abruf 7.8.2024).

¹² Zum Register siehe Internet: <https://ai.hel.fi/en/ai-register/> (Abruf 7.8.2024).

¹³ Siehe Team AI Regulation, Amsterdam and Helsinki launch Algorithm and AI Register, AI-Regulation vom 13. Oktober 2020, Internet: <http://ai-regulation.com/amsterdam-and-helsinki-launch-algorithm-and-ai-register/> (Abruf 7.8.2024); DIMITROVA ASENIYA, Helsinki and Amsterdam with first ever AI registries, The Mayor 28.9.2020, Internet: <https://www.themayor.eu/en/a/view/helsinki-and-amsterdam-with-first-ever-ai-registries-5982> (Abruf 7.8.2024).

¹⁴ Zum Register siehe Internet: <https://algorithmerregister.rotterdam.nl/p/Onzealgoritmes> (Abruf 7.8.2024).

Diskriminierungsskandal¹⁵ richtete die *niederländische Regierung 2022* ein nationales Zentralregister ein¹⁶, und ebenfalls auf nationaler Ebene verortet sind Transparenzbestrebungen in *Frankreich*. Seit 2016 müssen bestimmte französische Behörden die wichtigsten algorithmischen Verarbeitungsregeln online veröffentlichen.¹⁷ In *Grossbritannien* wurde 2021 im Rahmen der *National Artificial Intelligence (AI) Strategy*¹⁸ ein Transparenzstandard entwickelt, um Informationen über algorithmische Systeme bereitzustellen.¹⁹ Der Standard besteht aus zwei Stufen: einer nicht-technischen Erklärung für die breite Öffentlichkeit und technischen Details für Expertinnen und Experten.²⁰ Schliesslich ist auf den AI Act der *Europäischen Union* hinzuweisen, der erstmals eine Registrierungspflicht für Hochrisiko-KI-Systeme und bestimmte andere Systeme vorsieht.²¹

Vor dem Hintergrund der verschiedenen Forderungen nach einem Verzeichnis über den staatlichen Algorithmeinsatz in der Schweiz sowie den skizzierten

¹⁵ Siehe dazu INGO DACHWITZ, Niederlande zahlen Millionenstrafe wegen Datendiskriminierung, Netzpolitik 29.12.2021, Internet: <https://netzpolitik.org/2021/kindergeldaffaere-niederlande-zahlen-millionenstrafe-wegen-datendiskriminierung> (Abruf 7.8.2024).

¹⁶ Siehe dazu INA BIERFREUND, Wie die Verwaltung KI-Register einführen konnte, Tagesspiegel Background 23.11.2023, Internet: <https://background.tagesspiegel.de/smart-city/wie-die-verwaltung-ki-register-einfuehren-koennte> (Abruf 7.8.2024); TOMAS RUDL, Wie die Niederlande aus KI-Skandalen lernen, Netzpolitik 01.2.2023, Internet: <https://netzpolitik.org/2023/auf-sicht-und-transparenz-wie-die-niederlande-aus-ki-skandalen-lernen/> (Abruf 7.8.2024). Zum Algorithmenregister siehe Internet: <https://algoritmes.overheid.nl/> (Abruf 7.8.2024).

¹⁷ Die Vorgabe geht zurück auf das Gesetz «Loi No2016-1321 pour une République numérique du 7 octobre 2016»; sie wurde mittlerweile in der Kodifikation über die Beziehungen zwischen der Öffentlichkeit und der Verwaltung (Code des relations entre le Public et l'Administration – CRPA) implementiert. Wortlaut der einschlägigen Bestimmung L312-1-3 CRPA: «Sous réserve des secrets protégés en application du 2° de l'article L. 311-5, les administrations mentionnées au premier alinéa de l'article L. 300-2, à l'exception des personnes morales dont le nombre d'agents ou de salariés est inférieur à un seuil fixé par décret, publient en ligne les règles définissant les principaux traitements algorithmiques utilisés dans l'accomplissement de leurs missions lorsqu'ils fondent des décisions individuelles».

¹⁸ Siehe Internet: <https://www.gov.uk/government/publications/national-ai-strategy> (Abruf 7.8.2024).

¹⁹ Algorithmic Transparency Recording Standard Hub, Internet: <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub> (Abruf 7.8.2024).

²⁰ NIGEL KINGSMAN et al., Public sector AI transparency standard: UK Government seeks to lead by example, Discover Artificial Intelligence, 21.2.2022, Internet: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3986213 (Abruf 7.8.2024).

²¹ Art. 49 Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung weiterer Erlasse, ABl. L, 2024/1689 (im Folgenden AI Act).

Entwicklungen in anderen Rechtsordnungen, werden im vorliegenden Beitrag zunächst die Zielsetzungen eines öffentlichen Verzeichnisses sowie die in einem solchen Verzeichnis zu erfassenden Systeme besprochen (II). Sodann wird untersucht, welche rechtlichen Grundlagen bestehen, um Transparenz in Bezug auf algorithmische Systeme zu fordern (III). Schliesslich wird analysiert, wie bzw. wo die Transparenzanforderungen rechtlich verankert werden könnten (IV). Die aufgeworfenen Fragen werden im Folgenden ausgehend von der Rechtslage auf Bundesebene diskutiert. Wo angezeigt, wird auf kantonale Diskussionen Bezug genommen.

II. Ziele eines Verzeichnisses und erfasste Systeme

Aufgrund unterschiedlicher Begriffsverwendungen im Kontext algorithmischer Systeme ist zunächst eine kurze terminologische Einordnung vorzunehmen (1). Die Transparenzforderungen sind sodann in Bezug auf die damit verfolgten Ziele zu konkretisieren (2). Im Zentrum stehen dabei der Schutz des Menschen (2.1), die Kontrolle staatlichen Handelns (2.2), die Vertrauensbildung (2.3) und die Dokumentation (2.4). Ausgehend von den identifizierten Zielen ist sodann zu präzisieren, welche Systeme in einem Verzeichnis zu erfassen sind (3).

1. Terminologie

In politischen, zivilgesellschaftlichen und wissenschaftlichen Kontexten werden im Zusammenhang mit Transparenzforderungen über den staatlichen Einsatz algorithmischer Entscheidungssysteme unterschiedliche Begriffe verwendet. Dazu zählt prominent der Begriff des «KI-Systems», der auf EU-Ebene inzwischen eine Legaldefinition erfahren hat. Gemäss Art. 3 Ziff. 1 AI Act bezeichnet ein KI-System «[...] ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und 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 [...]». Zentral ist, dass die Systeme mit einem gewissen Mass an Autonomie ausgestattet und anpassungsfähig sind.²²

²² Siehe auch die Definition in Art.2 AI-Convention oder die Definition der OECD, <https://oecd.ai/en/ai-principles>. Die Definition der EU lehnt sich an jene der OECD an, um Rechtssicherheit, Harmonisierung und Akzeptanz zu schaffen und die internationale Zu-

Hinsichtlich der hier diskutierten Verzeichnisse ist zu bedenken, dass Systeme, die Rechte und Pflichten natürlicher Personen berühren können oder in ihrer Entscheidungsfindung nicht nachvollziehbar sind, auch ohne Anpassungs- bzw. Lernfähigkeit auskommen können.

Deshalb wird die Verwendung des Begriffs «Algorithmic Decision-Making System (ADMS) – Algorithmisches Entscheidungssystem» bevorzugt.²³ Dieser ist technologieneutral und erfasst sowohl Systeme, die auf maschinellen Lernverfahren basieren als auch Systeme, deren Algorithmen nach klassischen Wenn-Dann-Schemata aufgebaut sind. Unter ADMS fallen Systeme, die mit oder gänzlich ohne menschliches Mitwirken funktionieren. Der Begriff verdeutlicht zudem, dass es sich um Systeme handelt, die in einem Entscheidungsfindungsprozess genutzt werden.

2. Ziele eines ADMS-Verzeichnisses

2.1. Schutz des Menschen und der Menschenwürde

Im Zusammenhang mit algorithmischen Systemen beziehen sich Transparenzforderungen häufig auf die Funktionsweise des algorithmischen Systems.²⁴ Insbesondere «komplexere» Algorithmen, die z. B. Methoden maschinellen Lernens einsetzen, sind meist nicht einmal mehr für die Programmiererinnen

sammenarbeit zu fördern, siehe DANIEL BECKER/DANIEL FEUERSTACK, Der neue Entwurf des EU-Parlaments für eine KI-Verordnung, MMR 2024, 22 ff., 23.

²³ ADM steht regelmässig auch für «automated decision-making», siehe etwa THOUVENIN et al., Workshop Report 2019 (FN 8), 1. Der Begriff «algorithmic decision-making» wird aber ebenfalls regelmässig verwendet, siehe z. B. Europäisches Parlament, Understanding algorithmic decision-making: Opportunities and challenges, 2019 (zit. ADM 2019), 2 ff., Internet: [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU\(2019\)624261_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf) (Abruf 7.8.2024); KATHARINA A. ZWEIG/GEORG WENZELBURGER/TOBIAS D. KRAFFT, On Chances and Risks of Security Related Algorithmic Decision-Making Systems, European Journal for Security Research, 2018, 181 ff.

²⁴ MARTIN KMENT/SOPHIE BORCHERT, Künstliche Intelligenz und Algorithmen in der Rechtsanwendung, München 2022, N 84 ff.; MARIO MARTINI, Algorithmen als Herausforderung für die Rechtsordnung, JZ 2017, 1017 ff. (zit. Algorithmen), 1018; LISA KÄDE/STEPHANIE MALTZAN, Die Erklärbarkeit von Künstlicher Intelligenz, CR 2020, 66 ff., 67.

oder Programmierer «transparent» (Blackbox-Phänomen²⁵).²⁶ Transparenz kann in diesem Zusammenhang bedeuten, dass die Person, die von einer derartigen Entscheidung betroffen ist, dies erkennt.²⁷ Auch die Begründbarkeit²⁸, die Erklärbarkeit²⁹ oder die Nachvollziehbarkeit³⁰ staatlicher Entscheidungen können damit gemeint sein. Ziel dieser Transparenz ist der Schutz des Menschen vor einer Degradierung zum Objekt³¹ und damit der Schutz der Menschenwürde (Art. 7 BV).³²

2.2. Kontrolle staatlichen Handelns

Da der Staat i. d. R. hoheitlich auftritt, d. h. einseitig Rechte und Pflichten gegenüber Privaten festlegt,³³ bedarf es gewisser Mechanismen, um staatliches Handeln kontrollierbar zu machen und damit das Funktionieren eines demo-

²⁵ MARIO MARTINI, *Blackbox Algorithmus*, Berlin 2019, 28 ff. (zit. *Blackbox*); THOMAS WISCHMEYER, *Artificial Intelligence and Transparency: Opening the Black Box*, in: Thomas Wischmeyer/Timo Rademacher (Hrsg.), *Regulating Artificial Intelligence*, Cham 2020, 75 ff.

²⁶ KMENT/BORCHERT (FN 24), N 89; BERNHARD WALTL, *Erklärbarkeit und Transparenz im Machine Learning*, in: Klaus Mainzer (Hrsg.), *Philosophisches Handbuch Künstliche Intelligenz*, Wiesbaden 2019, 1 ff., 4 ff.

²⁷ Im Sinne einer Kennzeichnungspflicht, siehe etwa NADJA BRAUN BINDER, *Künstliche Intelligenz und automatisierte Entscheidungen in der öffentlichen Verwaltung*, SJZ 115/2019, 467 ff., 475 f. (zit. KI).

²⁸ Siehe dazu vertieft NADJA BRAUN BINDER/LILIANE OBRECHT, *Die Begründung von Verfügungen beim Einsatz algorithmischer Systeme*, SJZ 2024, 707 ff.

²⁹ Z. B. Der Bundesrat, *Leitlinien «Künstliche Intelligenz» für den Bund*, 25.11.2020, 4, Internet: <https://www.sbf.admin.ch/sbf/de/home/bfi-politik/bfi-2021-2024/transversale-the-men/digitalisierung-bfi/kuenstliche-intelligenz.html> (Abruf 7.8.2024) (zit. BR Leitlinien 2020); KMENT/BORCHERT (FN 24), N 207 ff.; WALTL (FN 26), 1 ff.; WISCHMEYER (FN 25), 75 ff.; PAUL VOGEL, *Künstliche Intelligenz und Datenschutz*, Diss. Universität Würzburg 2020, Baden-Baden 2022, 72 ff.

³⁰ BR Leitlinien 2020 (FN 29), 4; VOGEL (FN 29), 72 ff.

³¹ In diesem Sinne wird auch von einem «Instrumentalisierungsverbot» gesprochen, siehe BSK BV-BELSER/MOLINARI, Art. 7 N 11 ff., in: Bernhard Waldmann/Eva Maria Belser/Astrid Epiney (Hrsg.), *Bundesverfassung, Basler Kommentar*, Basel 2015 (zit. BSK BV-Verfasser). Zur Menschenwürde vertieft statt vieler JÖRG PAUL MÜLLER/MARKUS SCHEFER, *Grundrechte in der Schweiz*, 4. A., Bern 2008, 1 ff.

³² BR Leitlinien 2020 (FN 29), 3; SGK BV-SCHWEIZER/SPENLÉ, Art. 7 N 70, in: Bernhard Ehrenzeller et al. (Hrsg.), *Die Schweizerische Bundesverfassung, St. Galler Kommentar*, 4. A., Zürich 2023 (zit. SGK BV-Verfasser).

³³ Zur Hoheitlichkeit staatlichen Handelns TOBIAS JAAG, *Hoheitliche und nicht-hoheitliche staatliche Tätigkeiten*, in: Peter Hänni (Hrsg.), *Mensch und Staat – Festgabe für Thomas Fleiner*, Freiburg 2003, 619 ff., 630 ff.

kratischen Rechtsstaates zu gewährleisten.³⁴ Diese Kontrolle wird z. B. durch individuelle Kontrollmöglichkeiten (etwa im Rahmen von Einsprachen oder Rekursen), die den Bedürfnissen des Einzelfalles Rechnung tragen, sichergestellt. Zusätzlich spielt auch die Kontrolle staatlichen Handelns durch die Öffentlichkeit eine wichtige Rolle.³⁵ Diese umfasst im Kontext staatlich eingesetzter algorithmischer Systeme zum einen eine tatsächliche Kontrolle der Systeme durch Expertinnen und Experten bzw. durch zivilgesellschaftliche Organisationen.³⁶ Zum anderen kann Transparenz einen Beitrag zur Selbstkontrolle leisten, indem sie dazu führen kann, dass diejenigen, denen Transparenzanforderungen auferlegt werden, davon ausgehen, dass zumindest einige andere erfahren werden, was sie tun oder unterlassen.³⁷ Allgemeine gesellschaftliche Kontrollmöglichkeiten fördern somit rationales Handeln sowie die Selbstkontrolle der Verwaltung.³⁸

2.3. Vertrauensbildung

Ein demokratischer Rechtsstaat kann nur funktionieren, wenn ein gegenseitiges Grundvertrauen zwischen Staat und Bevölkerung besteht.³⁹ Indem Behörden hoheitlich über Rechte und Pflichten betroffener Personen entscheiden können, kann ein Machtungleichgewicht zwischen Staat und Bevölkerung entstehen, das potenziell zu einem Machtmissbrauch führen kann. Dadurch kann die Vertrauensbeziehung zwischen Staat und Bevölkerung ins Schwanken geraten. So wird das Veröffentlichen von Informationen durch den Staat denn auch als Mittel zur Förderung der Vertrauensbeziehung zwischen Staat

³⁴ PATRICK SUTTER, Vertrauen durch Informationszugang: Von der Notwendigkeit eines umfassenden Rechts auf Zugang zu amtlichen Informationen, *digma* 2004, 150 ff., 153; FRANZISKA SPRECHER, Transparenz: Ein Grundprinzip des Rechtsstaats und seine Bedeutung im Gesundheitsrecht, insbesondere im Heilmittel- und Humanforschungsrecht, *ZSR* 2016, 139 ff., 151; RENÉ WIEDERKEHR, Transparenz als Grundsatz rechtsstaatlichen Handelns (Art. 5 BV), *ZBI* 2007, 521 ff., 521 ff.

³⁵ Im Resultat für den Bereich des Datenschutzes ebenso BERTIL COTTIER, *Transparence des traitements de données personnelles opérés par les organes fédéraux: un pas en avant, deux en arrière*, *SZW* 2021, 65 ff., 66.

³⁶ Siehe nur etwa die Aktivitäten von AlgorithmWatch Schweiz, Internet: <https://algorithm-watch.ch/de/> (Abruf 7.8.2024).

³⁷ ONORA O'NEILL, *Transparency and the Ethics of Communication*, in: Christopher Hood/David Heald (Hrsg.), *Transparency, The Key to Better Governance?*, New York 2006, 74 ff., 85. SPRECHER (FN 34), 152.

³⁹ SGK BV-SCHINDLER/TSCHUMI (FN 32), Art. 5 N 62; KASPAR SUTTER, *Vertrauen im Recht*, Habil. Bern 2020, Zürich 2020, 1.

und Gesellschaft sowie zur Sicherstellung der Rechtsstaatlichkeit betrachtet.⁴⁰ Dies dient der Legitimation des Staates insgesamt und der damit einhergehenden Akzeptanz von staatlichen Entscheidungen.⁴¹

2.4. Dokumentation

Das ADMS-Verzeichnis ist eine Möglichkeit, in standardisierter und archivierbarer Form Entscheidungen und Annahmen algorithmischer Systeme bei der Entwicklung, Implementierung, Verwaltung und gegebenenfalls Demontage zu dokumentieren.⁴² Einerseits führt das Verzeichnis somit Protokoll über die genutzten algorithmischen Systeme durch die öffentliche Verwaltung, andererseits kann es der interessierten Öffentlichkeit Informationen über das eingesetzte System liefern.⁴³ Dabei ist die Möglichkeit einer online Durchsuchung des Verzeichnisses besonders wichtig. Indem sich interessierte Personen direkt selbst über die in der Verwaltung eingesetzten ADMS erkundigen können, kann die Verwaltung entlastet werden. Bei entsprechenden Gesuchen auf Einsichtnahme müssen keine aufwändigen verwaltungsinternen Ermittlungen angestellt werden.⁴⁴

3. Erfasste Systeme

Um die skizzierten Ziele zu erreichen, müssen nicht alle algorithmischen Systeme in einem öffentlichen Verzeichnis erfasst werden, sondern nur diejenigen, die mit erhöhten Risiken verbunden sind. So sind etwa algorithmische Systeme, die im Rahmen der administrativen Hilfstätigkeit genutzt werden (z. B. KI-basierte Übersetzungsprogramme), aus grundrechtlicher Perspektive

⁴⁰ THOMAS POLEDNA, Staatliche Informationspflichten (Grundversorgung mit elektronischen Daten), in: Thomas Koller/Heinrich Koller (Hrsg.), Recht und Rechtsdaten – Anspruch und Wirklichkeit, Tagung für Informatikrecht, Bern 2003, 69 ff., 87.

⁴¹ POLEDNA (FN 40), 88.

⁴² MAYA MURAD, Beyond the «Black Box»: Enabling Meaningful Transparency of Algorithmic Decision-Making Systems through Public Registers, Juni 2021, 16, Internet: <https://dspace.mit.edu/bitstream/handle/1721.1/139092/murad-mmurad-sm-idm-2021-the-sis.pdf?sequence=1&isAllowed=y> (Abruf 7.8.2024); MEERI HAATAJA/LINDA VAN DE FLIERT/PASI RAUTIO, Public AI Registers, White Paper on realising AI transparency and civic participation in government use of AI, September 2020, 3, Internet: <https://algoritmeregister.amsterdam.nl/wp-content/uploads/White-Paper.pdf> (Abruf 7.8.2024).

⁴³ HAATAJA/VAN DE FLIERT/RAUTIO (FN 42), 3, 6.

⁴⁴ Siehe Beschluss des Zürcher Regierungsrates (Regierungsratsbeschluss) Nr. 878/2023 vom 5. Juli 2023, 39.

mit geringen Risiken verbunden. Es wäre entsprechend nicht zielführend, diese in das Verzeichnis aufzunehmen. Gleichzeitig ist zu bedenken, dass Transparenzherausforderungen nicht nur bei maschinell lernenden Algorithmen bestehen können, sondern auch bei Algorithmen, die nach klassischen Wenn-Dann-Schemata funktionieren.⁴⁵ Ziel ist somit, komplexe bzw. potenziell problembehaftete algorithmische Systeme von unproblematischen Systemen zu unterscheiden.

Zu diesem Zweck kann auf die Auswirkungen des eingesetzten Systems, d. h. auf die Entscheidung, die aus der Anwendung des algorithmischen Systems resultiert, abgestellt werden.⁴⁶ Wirkt sich die Entscheidung auf Rechte und Pflichten von Menschen aus, ist das eingesetzte Entscheidungssystem im Verzeichnis aufzuführen.⁴⁷

III. Rechtliche Anknüpfungspunkte für Transparenzforderungen

Bevor im nächsten Abschnitt analysiert wird, wie bzw. wo ein ADMS-Verzeichnis auf Bundesebene rechtlich verankert werden kann, soll in diesem Abschnitt zunächst geklärt werden, aus welchen bestehenden Rechtsgrundlagen sich Transparenzanforderungen an staatliche algorithmische Entscheidungssysteme ableiten lassen und ob sich daraus ggf. konkrete Vorgaben für verpflichtende ADMS-Verzeichnisse ergeben. Als Grundlagen kommen sowohl die Verfassung (1 und 2) als auch verschiedene Gesetze in Betracht (3).

1. Verfassungsrechtliche Herleitung

Der Begriff «Transparenz» findet im Bundesrecht keine Legaldefinition. Es existiert auch keine andere einheitliche Definition.⁴⁸ In der Literatur bestehen

⁴⁵ JONATHAN DOLLINGER, Folgenabschätzungen für Verwaltungs-Algorithmen, Diss. Freiburg i.B. 2022, Tübingen 2023, 20.

⁴⁶ MURAD (FN 42), 19. Der Kanton Zürich beschränkt die Transparenzpflichten auf Systeme, die Auswirkungen auf die Grundrechte zeitigen, siehe FN 3.

⁴⁷ MURAD (FN 42), 19, erwähnt daneben auch das Kriterium des Einflusses eines algorithmischen Systems auf öffentliche Güter wie demokratische Institutionen, die Umwelt oder die öffentlichen Finanzen.

⁴⁸ Duden online zum Begriff «Transparenz»: Durchscheinen, Durchsichtigkeit, (Licht-)Durchlässigkeit, Durchschaubarkeit, Nachvollziehbarkeit, Internet: <https://www.duden.de/rechtschreibung/Transparenz> (Abruf 7.8.2024).

dagegen unterschiedliche Begriffsverständnisse, die auf verschiedene Bestimmungen rekurrieren.⁴⁹

So wird der Transparenzbegriff aus der Rechtsstaatlichkeit nach Art. 5 BV hergeleitet,⁵⁰ da diese in der Verfassung nicht abschliessend geregelt sei und grundsätzlich von einem breiten Verständnis des Begriffs der Rechtsstaatlichkeit ausgegangen werden müsse.⁵¹ Dabei wird einerseits das Fehlen eines staatlichen Transparenzgebots bei den verfassungsrechtlichen Grundsätzen rechtsstaatlichen Handelns bemängelt,⁵² andererseits die Transparenz als ungeschriebener rechtsstaatlicher Grundsatz des Verfassungsrechts verstanden⁵³. Als argumentative Anknüpfungspunkte dienen die Entwicklungen im europäischen und internationalen Umfeld⁵⁴ sowie Entwicklungen auf Bundesebene⁵⁵ und in den Kantonen.⁵⁶ Ein weiterer Ansatz bezeichnet Transparenz ebenfalls als ungeschriebenen «Grundsatz rechtsstaatlichen Handelns»⁵⁷. Dieser stellt bei der Herleitung auf die unterschiedlichen Funktionen ab, namentlich eine demokratische, eine rechtsstaatliche und eine individualrechtliche Funktion, die eine Kontrolle und damit eine Begrenzung staatlicher Macht ermöglichen.⁵⁸ Schliesslich wird

⁴⁹ Die kantonalen Verfassungen hingegen sehen zum Teil einen Verfassungsgrundsatz der «Transparenz» vor, siehe etwa Art. 49 Kantonsverfassung des Kantons Zürich (LS 101; SR 131.211). Für Deutschland siehe grundlegend JÜRGEN BRÖHMER, Transparenz als Verfassungsprinzip, Habil. Universität Saarbrücken 2002, Tübingen 2004.

⁵⁰ SPRECHER (FN 34), 139 ff.; WIEDERKEHR (FN 34), 521 ff.; SGK BV-SCHINDLER/TSCHUMI (FN 32), Art. 5 N 11 BV.

⁵¹ Grundlegend zum Rechtsstaatsprinzip BSK BV-EPINEY (FN 31), Art. 5 N 16 ff.; WIEDERKEHR (FN 34), 526 f., schreibt der Transparenz (u. a.) eine rechtsstaatliche Funktion zu.

⁵² SGK BV-SCHINDLER (FN 32), Art. 8 N 8.

⁵³ SPRECHER (FN 34), 182, 191, die daraus aber kein verfassungsmässiges Individualrecht ableitet.

⁵⁴ SPRECHER (FN 34), 157 ff., 183.

⁵⁵ So verweist SPRECHER (FN 34), 158 ff., 183, auf die Verankerung unterschiedlichster Ausflüsse von Transparenz in der Bundesverfassung.

⁵⁶ SPRECHER (FN 34), 174 ff., 184, analysiert verschiedene transparenzrechtliche Bestimmungen in den kantonalen Verfassungen. Nicht zuletzt betont sie, dass die Genfer Verfassung Transparenz sogar als Grundsatz staatlichen Handelns statuiert (Art. 9 Abs. 3 Verfassung der Republik und des Kantons Genf vom 14. Oktober 2012; SR 131.234).

⁵⁷ WIEDERKEHR (FN 34), 543.

⁵⁸ WIEDERKEHR (FN 34), 526 ff.; RENÉ WIEDERKEHR/CHRISTIAN MEYER/ANNA BÖHME, Transparenz als Verfahrensgrundsatz: Gesetzliche Transparenztechniken und praktische Argumentationsfelder, recht 2022, 131 ff., 131 f.

Transparenz als Grundsatz des Verwaltungshandelns⁵⁹ oder als Grundlage rechtsstaatlichen Handelns⁶⁰ verstanden.

Transparenz wird ferner im Zusammenhang mit individualrechtlichen Bestimmungen thematisiert. Transparenz ist dabei das Ziel, das mithilfe der individualrechtlichen Instrumente hergestellt werden soll.⁶¹ Im Bereich der verfassungsrechtlichen Verfahrensgarantien stellen das rechtliche Gehör (Art. 29 Abs. 2 BV), insbesondere die Begründungspflicht⁶² und das Akteneinsichtsrecht⁶³, solche individualrechtlichen Mittel zur Herstellung von Transparenz dar. Der individuelle Anspruch, der die im Verfahren beteiligten Personen durch den grundrechtlichen Gehalt schützt,⁶⁴ ermöglicht dabei gleichzeitig gesellschaftliche Kontrolle.⁶⁵

Ein weiterer verfassungsrechtlicher Anknüpfungspunkt stellt der Informationsauftrag des Bundesrates nach Art. 180 Abs. 2 BV dar, der insbesondere in Art. 10 Regierungs- und Verwaltungsorganisationsgesetz (RVOG)⁶⁶ konkretisiert wird.⁶⁷ Art. 180 Abs. 2 BV statuiert die aktive Informationstätigkeit,⁶⁸ d. h. ein proaktives Tätigwerden des Staates.⁶⁹ Diese Informationspflichten erfassen die Regierung sowie die Verwaltung.⁷⁰ Staatliche Information wird dabei

⁵⁹ BERNHARD EHRENZELLER, Öffentlichkeit der öffentlichen Verwaltung, in: Walter P. Schlupep (Hrsg.), Recht, Staat und Politik am Ende des zweiten Jahrtausends, Festschrift zum 60. Geburtstag von Bundesrat Arnold Koller, Bern/Stuttgart/Wien 1993, 31 ff., 46.

⁶⁰ STEPHAN BRUNNER, Öffentlichkeit der Verwaltung und informationelle Selbstbestimmung: Von Kollisionen und Verkehrsregeln, in: Patrick Sutter (Hrsg.), Selbstbestimmung und Recht, Festgabe für Rainer J. Schweizer zum 69. Geburtstag, Zürich/Basel/Genf 2003, 31 ff., 31.

⁶¹ JEAN-PATRICK VILLENEUVE/PHILOMÈNE MEILLAND/MARTIAL PASQUIER, La transparence de la transparence, in: Martial Pasquier (Hrsg.), Le principe de transparence en Suisse et dans le monde, Lausanne 2013, 57 ff., 57.

⁶² AXEL TSCHENTSCHER/ANDREAS LIENHARD/FRAZISKA SPRECHER, Öffentliches Recht, 2. A., Zürich/St. Gallen 2019, N 343, 553 sowie vertieft RENÉ RHINOW et al., Öffentliches Prozessrecht, 4. A., Basel 2021, N 309 ff.; WIEDERKEHR/MEYER/BÖHME (FN 58), 137.

⁶³ SPRECHER (FN 34), 159; RHINOW et al. (FN 62), N 1230 ff.

⁶⁴ SPRECHER (FN 34), 158 f.

⁶⁵ BSK BV-REICH (FN 31), Art. 30 N 54.

⁶⁶ Regierungs- und Verwaltungsorganisationsgesetz vom 21. März 1997 (RVOG; SR 172.010).

⁶⁷ Siehe auch weiter unten [III.3.2](#).

⁶⁸ Die passive Seite, das Recht auf Zugang zu Informationen, wird durch das Öffentlichkeitsgesetz geregelt, siehe dazu weiter unten [III.3.1](#).

⁶⁹ BSK BV-MÜLLER (FN 31), Art. 180 N 8; SGK BV-SAXER/BRUNNER (FN 32), Art. 180 N 25.

⁷⁰ Botschaft über eine neue Bundesverfassung vom 20. November 1996, BBl 1997 I, 411.

als «Grundversorgung»⁷¹ verstanden. Deren Umfang und Inhalt ist abhängig vom jeweiligen Ort und der jeweiligen Zeit – ausschlaggebend sind die Bedürfnisse der Bevölkerung sowie des Staates selbst.⁷² Den Behörden kommt mit hin ein grosser Ermessensspielraum zu, ob überhaupt informiert werden soll.⁷³ Ziel ist es, die Legitimität staatlichen Handelns zu gewährleisten und zu sichern.⁷⁴ Information ist somit ein weiteres Mittel, um Transparenz staatlichen Handelns herzustellen.⁷⁵

2. Transparenz zur Verhinderung von Diskriminierung

Im Kontext algorithmischer Entscheidungssysteme wird die Schaffung von Transparenz auch als Massnahme zur Verhinderung von Diskriminierung betrachtet.⁷⁶ Das in Art. 8 Abs. 2 BV statuierte allgemeine Diskriminierungsverbot für staatliche Akteurinnen und Akteure gilt auch, wenn diese algorithmische Entscheidungssysteme einsetzen. Insbesondere, wenn diese auf maschinellen Lernmethoden basieren, besteht die Gefahr von diskriminierenden Resultaten.⁷⁷ Deren Ursprung liegt häufig bereits in der Entwicklung der Systeme

⁷¹ POLEDNA (FN 40), 75.

⁷² POLEDNA (FN 40), 76.

⁷³ SHK BGÖ-BRUNNER/MADER, Bundesgesetz über das Öffentlichkeitsprinzip der Verwaltung: Entstehung, Konzept, Kontext, N 87, in: Stephan C. Brunner/Luzius Mader (Hrsg.), Öffentlichkeitsgesetz, Stämpfli Handkommentar Bern 2008 (zit. SHK BGÖ).

⁷⁴ URS SAXER, Öffentlichkeitsinformationen von Behörden im Rechtsstaat, *medialex* 2004, 19 ff., 19.

⁷⁵ SGK BV-SAXER/BRUNNER (FN 32), Art. 180 N 24. Explizit etwa Art. 49 Verfassung des Kantons Zürich (SR 131.211).

⁷⁶ Dazu z. B. BETTINA BERENDT, The AI Act proposal: Towards the next transparency fallacy? Why AI regulation should be based on principles based on how algorithmic discrimination works, in: Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz/Frauke Rostalski (Hrsg.), Künstliche Intelligenz – Wie gelingt eine vertrauenswürdige Verwendung in Deutschland und Europa?, Tübingen 2022, 31 ff.; GABRIELE BUCHHOLTZ/MARTIN SCHEFFEL-KAIN, Algorithmen und Proxy Discrimination in der Verwaltung: Vorschläge zur Wahrung digitaler Gleichheit, *NVwZ* 2022, 612 ff., 617; FREDERIK J. ZUIDERVEEN BORGESIOUS, Strengthening legal protection against discrimination by algorithms and artificial intelligence, *The International Journal of Human Rights* 2020, 1572 ff., 1577; BRAUN BINDER et al., *Studie KI* 2021 (FN 8), 50.

⁷⁷ ANDREAS TINHOFER, Algorithmenbasierte Entscheidungen und Diskriminierung, *DRdA* 2022, 170 ff., Internet: https://www.drda.at/a/399_DRDA_15/Algorithmenbasierte-Entscheidungen-und-Diskriminierung (Abruf 7.8.2024); siehe für Beispiele NADIA BRAUN BINDER/LILIANE OBRECHT, Künstliche Intelligenz in der Verwaltung – Möglichkeiten, rechtliche Rahmenbedingungen und Grenzen, in: *SVVOR Tagungsband 2021/2022*, Bern 2023, 22 ff., 36 ff.

selbst.⁷⁸ Solche Systeme benötigen i. d. R. skalierbare Datensätze, um trainiert zu werden.⁷⁹ Aus unterschiedlichen Gründen kann es möglich sein, dass diese Datensätze nicht ausgewogen bzw. unvollständig sind.⁸⁰ Hängt eine Entscheidung etwa von neutral erscheinenden Daten ab, knüpft jedoch faktisch an geschützte Merkmale an (sog. Proxy- bzw. Stellvertreter-Diskriminierung⁸¹), kann dies zu einem diskriminierenden Resultat führen.⁸² Dabei ist nach aussen meist nicht ersichtlich, dass das Anknüpfen an das Proxy-Merkmal zu einer Diskriminierung geführt hat.⁸³ Das Phänomen der Proxy-Diskriminierung entsteht dadurch, dass algorithmische Systeme basierend auf maschinellen Lernmethoden nach Korrelationen suchen,⁸⁴ wobei sie häufig Stellvertretervariablen finden.⁸⁵ Die Schaffung von Transparenz hinsichtlich des Einsatzes algorithmischer Systeme sowie der verwendeten Daten stellt zwar keine Lösung des Diskriminierungsproblems dar,⁸⁶ ebnet jedoch den Weg für die Erkennung potenzieller Diskriminierungen sowie für die Einleitung entsprechender Gegenmassnahmen.

3. Gesetzliche Anknüpfungspunkte

3.1. Öffentlichkeitsprinzip

Der Begriff der Transparenz wird häufig im Kontext des Öffentlichkeitsprinzips diskutiert. Kurz vor und seit der Einführung des Öffentlichkeitsgesetzes des Bundes (BGÖ⁸⁷) im Jahr 2004 haben sich zahlreiche Stimmen dazu geäu-

⁷⁸ BERENDT (FN 76), 35 ff.; MARTINI, Algorithmen (FN 24), 1018 f.

⁷⁹ ALFRED FRÜH/DARIO HAUX, Foundations of Artificial Intelligence and Machine Learning, Berlin 2022, 8, Internet: <https://www.ssoar.info/ssoar/handle/document/80203> (Abruf 7.8.2024).

⁸⁰ Dies ist ein Problem der Datenqualität, siehe dazu JEREMY STEVENS, Datenqualität bei algorithmischen Entscheidungen, CR 2020, 73 ff.

⁸¹ «Proxy» ist ein Begriff aus der Statistik und wird mit «Platzhalter» übersetzt, siehe dazu LINDA SUPIK, Statistik und Diskriminierung, in: Albert Scherr/Aladin El-Mafaalani/Gökçen Yüksel (Hrsg.), Handbuch Diskriminierung, Wiesbaden 2017, 191 ff., 194 f.

⁸² Siehe dazu BUCHHOLTZ/SCHEFFEL-KAIN (FN 76), 612 f.

⁸³ BUCHHOLTZ/SCHEFFEL-KAIN (FN 76), 616 f.

⁸⁴ WOLFGANG ERTEL, Grundkurs Künstliche Intelligenz, 4. A., Wiesbaden 2016, 191 ff.

⁸⁵ SUPIK (FN 81), 194.

⁸⁶ BERENDT (FN 76), 47 f.

⁸⁷ Bundesgesetz über das Öffentlichkeitsprinzip der Verwaltung vom 17. Dezember 2004 (Öffentlichkeitsgesetz, BGÖ; SR 152.3). «Öffentlichkeitsprinzip» zu Französisch: «le principe de transparence» und zu Italienisch: «il principio di trasparenza dell'amministrazione».

sert.⁸⁸ Zentrales Anliegen des BGÖ ist die Verbesserung der Kontrolle der Verwaltung durch die Bevölkerung.⁸⁹ Es soll die Transparenz der Verwaltung fördern, indem es jeder interessierten Person den Zugang zu Dokumenten der Bundesbehörden ermöglicht.⁹⁰ Dadurch soll das Vertrauen der Bevölkerung in staatliche Institutionen gefördert werden.⁹¹ Kern des Gesetzes ist Art. 6 BGÖ, der das Öffentlichkeitsprinzip statuiert und jeder Person ein Recht auf Einsicht in amtliche Dokumente einräumt.

Allerdings erfasst das Öffentlichkeitsgesetz lediglich die passive Informations-tätigkeit der Verwaltung, eine aktive Information durch die Behörden ist nicht verlangt.⁹² Es handelt sich damit um eine Holschuld der Bevölkerung, also ein aktives Tätigwerden ihrerseits und nicht seitens des Staates.⁹³ Trotzdem spricht das Bundesgericht im Zusammenhang mit dem BGÖ auch von einem «Transparenzgebot».⁹⁴ Freilich bedeutet dies nicht, dass sich daraus eine aktive Pflicht für Behörden herleiten liesse, über die von ihnen eingesetzten algorithmischen Entscheidungssysteme in einem öffentlichen Verzeichnis zu informieren.

Das Einsichtsrecht des BGÖ muss hingegen auch im Kontext von algorithmischen Systemen gelten. Dies bedeutet, dass interessierte Personen einerseits Kenntnis darüber haben müssen, dass solche Systeme zum Einsatz kommen, andererseits jedoch auch die Möglichkeit haben müssen, Informationen über die Systeme einzusehen.⁹⁵ Dazu gehört auch, dass die Informationen so zur

⁸⁸ Siehe z. B. die Beiträge in Thierry Tanquerel/François Bellanger (Hrsg.), *L'administration transparente*, Genf/Basel/München 2002; ANDREAS LIENHARD, Staats- und Verwaltungsrecht, in: Andreas Ladner et al. (Hrsg.), *Handbuch der öffentlichen Verwaltung in der Schweiz*, Zürich 2013, 225 ff., 227. PIERRE TSCHANNEN, Staatsrecht der Schweizerischen Eidgenossenschaft, 5. A., Bern 2021, § 28, 395 ff., setzt den Begriff des Öffentlichkeitsprinzips dem Transparenzprinzip gleich (siehe SPRECHER [FN 34], 158 zur 4. A. des Buchs).

⁸⁹ Botschaft zum Bundesgesetz über die Öffentlichkeit der Verwaltung (Öffentlichkeitsgesetz, BGÖ), BBl 2003 1963 ff., 1976 (zit. Botschaft BGÖ 2003).

⁹⁰ Botschaft BGÖ 2003 (FN 89), 1964, 2001.

⁹¹ BSK BGÖ-BLECHTA/STÖCKLI/GRÜNINGER, Art. 1 N 4 ff., in: Gabor-Paul Blechta/David Vasella (Hrsg.), *Datenschutzgesetz/Öffentlichkeitsgesetz*, Basler Kommentar, 4. A., Basel 2024 (zit. BSK BGÖ-Verfasser).

⁹² Botschaft BGÖ 2003 (FN 89), 1977.

⁹³ BSK BV-MÜLLER (FN 31), Art. 180 N 8, verwendet den Begriff des «Holprinzips».

⁹⁴ BGE 142 II 313 E. 3.1 S. 316.

⁹⁵ Im Kanton Zürich konnte gestützt auf das Zürcher Informations- und Datenschutzgesetz der Algorithmus der Software eingesehen werden, mit dem die Rückfälle bei Straftätern minimiert werden. Siehe dazu Öffentlichkeitsgesetz.ch, Auch Algorithmen müssen grundsätzlich öffentlich sein, Interview vom 2. September 2018, Internet: <https://www.oeffent->

Verfügung gestellt werden, dass die Personen sie verstehen können.⁹⁶ In der praktischen Umsetzung bietet es sich an, dass die Behörden von sich aus Informationen im Internet publizieren, da hierdurch der Anspruch auf Einsicht erfüllt wird.⁹⁷

3.2. Informationspflicht

Die bereits oben skizzierte verfassungsrechtlich statuierte Informationspflicht von Bundesregierung und Bundesverwaltung wird in Art. 10 RVOG konkretisiert.⁹⁸ Gegenstand der in Art. 10 RVOG verankerten Informationspflicht sind die Tätigkeiten des Bundesrates.⁹⁹ Art. 10 Abs. 2 RVOG verlangt eine einheitliche, frühzeitige und kontinuierliche Information über seine Lagebeurteilungen, Planungen, Entscheide und Vorkehren. Die Information gemäss Art. 180 Abs. 2 BV muss «rechtzeitig» – d. h. früh genug – und «umfassend» – d. h. objektiv, wahrheitsgetreu und vollständig – erfolgen.¹⁰⁰

Die Aufzählung im RVOG ist nur beispielhaft, die Regierung hat über alles Wichtige zu informieren.¹⁰¹ Die Modalitäten sind unter Berücksichtigung des Verhältnismässigkeitsgrundsatzes so zu wählen, dass die Informationstätigkeit die ihr zukommende Funktion erfüllen kann.¹⁰²

Der Einsatz algorithmischer Systeme ist von der Informationspflicht nicht ausgenommen. Da die im vorliegenden Beitrag behandelten ADMS Auswirkungen auf die Rechte und Pflichten der betroffenen Personen haben bzw. ein Diskriminierungsrisiko¹⁰³ besteht, kann argumentiert werden, dass der Einsatz

[lichtkeitsgesetz.ch/deutsch/2018/09/auch-algorithmen-müssen-grundsatzlich-offentlich-sein/](https://www.lichtkeitsgesetz.ch/deutsch/2018/09/auch-algorithmen-müssen-grundsatzlich-offentlich-sein/) (Abruf 7.8.2024).

⁹⁶ Behörden sind gemäss dem verfassungsrechtlichen Verhältnismässigkeitsgrundsatz (Art. 5 Abs. 3 BV) nicht verpflichtet, einen unverhältnismässigen Aufwand zu betreiben, das amtliche Dokument in verständliche Sprache zu übersetzen.

⁹⁷ Art. 6 Abs. 3 BGÖ, siehe dazu SHK BGÖ-BRUNNER/MADER (FN 73), Entstehung, Konzept, Kontext N 73.

⁹⁸ Siehe dazu [III.1](#).

⁹⁹ Damit sind nicht nur behördliche Handlungen im Sinne einer Entscheidungstätigkeit gemeint, sondern z. B. auch Informationen über wichtige aktuelle Geschehnisse, die im öffentlichen Interesse liegen, siehe SGK BV-SAXER/BRUNNER (FN 33), Art. 180 N 45.

¹⁰⁰ SGK BV-SAXER/BRUNNER (FN 32), Art. 180 N 58.

¹⁰¹ SHK RVOG-SÄGESSER, Art. 10 N 24, in: Thomas Sägeser (Hrsg.), Regierungs- und Verwaltungsorganisationsgesetz RVOG vom 21. März 1997, Bern 2007 (zit. SHK RVOG).

¹⁰² SHK RVOG-SÄGESSER (FN 101), Art. 10 N 56 f.

¹⁰³ Siehe dazu [III.2](#).

solcher Systeme als wichtige Information gilt. Der Bundesrat müsste entsprechend über deren Einsatz informieren. Aus diesen Erwägungen lässt sich aber höchstens eine grundsätzliche Informationspflicht herleiten. Eine Pflicht zur Veröffentlichung in einem ADMS-Verzeichnis ist damit nicht verbunden.

3.3. Datenschutz

Das Datenschutzgesetz (DSG)¹⁰⁴ enthält verschiedene Normen zur Schaffung von Transparenz bei der Bearbeitung von Personendaten.¹⁰⁵ Bereits die erste Datenschutzrechtsrevision im Jahr 2006 hatte die Verstärkung der Transparenz bei der Bearbeitung von Personendaten zum Ziel.¹⁰⁶ Dabei stand insbesondere die Information der von der Datensammlung betroffenen Person im Zentrum.¹⁰⁷ Auch die jüngste Revision des Datenschutzgesetzes stand im Lichte der Schaffung von Transparenz.¹⁰⁸ Übergeordnetes Ziel dieser Bestrebungen ist der Schutz der Grundrechte und der Persönlichkeit.¹⁰⁹ Betroffene Personen sollen wissen, wer, wann, weshalb, welche Daten über sie bearbeitet.

Transparenz nimmt im Kontext des Datenschutzes unterschiedliche Ausprägungen an. Die Erkennbarkeit der Datenbearbeitung stellt dabei einen zentralen Grundsatz dar.¹¹⁰ Darüber hinaus sind verschiedene Informations- und Kennzeichnungspflichten zu beachten.¹¹¹ Zudem besteht ein datenschutz-

¹⁰⁴ Bundesgesetz über den Datenschutz vom 25. September 2020 (DSG; SR 235.1).

¹⁰⁵ Das Datenschutzgesetz ist anwendbar, wenn Personendaten bearbeitet werden (Art. 2 Abs. 1 DSG).

¹⁰⁶ Botschaft zur Änderung des Bundesgesetzes über den Datenschutz (DSG) und zum Bundesbeschluss betreffend den Beitritt der Schweiz zum Zusatzprotokoll vom 8. November 2001 zum Übereinkommen zum Schutz des Menschen bei der automatischen Verarbeitung personenbezogener Daten bezüglich Aufsichtsbehörden und grenzüberschreitende Datenübermittlung vom 19. Februar 2003, BBl 2003 2101 ff. (zit. Botschaft DSG 2003), 2102.

¹⁰⁷ Botschaft DSG 2003 (FN 106), 2102.

¹⁰⁸ Botschaft zum Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz vom 15. September 2017, BBl 2017 6941 ff., 6943, 6972.

¹⁰⁹ Art. 1 DSG.

¹¹⁰ Art. 6 Abs. 3 DSG. Dieser wird teilweise auch als «Transparenzprinzip» verstanden, siehe SHK DSG-BAERISWYL, Art. 4 N 48, in: Bruno Baeriswyl/Kurt Pärli/Dominika Blonski (Hrsg.), Datenschutzgesetz (DSG), Stämpfli Handkommentar, 2. A., Bern 2023 (zit. SHK DSG).

¹¹¹ Art. 19 DSG (Informationspflicht bei der Datenbeschaffung), Art. 21 Abs. 1 DSG (Informationspflicht bei einer automatisierten Einzelentscheidung) und Art. 21 Abs. 4 DSG (Kennzeichnungspflicht bei einer automatisierten Einzelentscheidung durch ein Bundesorgan).

rechtliches Auskunftsrecht.¹¹² Schliesslich sind Verzeichnisse und Register über die Bearbeitungstätigkeiten zu führen.¹¹³ Das Führen des datenschutzrechtlichen Registers über die Bearbeitungstätigkeiten wird auch als «allgemeines Transparenzelement der Verwaltung»¹¹⁴ bezeichnet.

Die datenschutzrechtlichen Vorgaben sind auch dann zu beachten, wenn Personendaten mithilfe algorithmischer Systeme bearbeitet werden. Das bedeutet allerdings nicht, dass aus der aktuellen Formulierung des DSG eine Pflicht zur Führung öffentlicher Verzeichnisse über die von Bundesorganen eingesetzten algorithmischen Systeme hergeleitet werden könnte.

4. Zwischenfazit

Die Untersuchung der rechtlichen Grundlagen zur Transparenz staatlichen Handelns zeigt eine starke Verankerung dieses Prinzips in der schweizerischen Rechtsordnung. Transparenz ist ein zentraler Aspekt, der sich durch verschiedene Verfassungs- und Gesetzesbestimmungen zieht und gilt als unverzichtbares Element rechtsstaatlichen Handelns, das die demokratische Kontrolle und Begrenzung staatlicher Macht ermöglicht. Individualrechtliche Bestimmungen wie das rechtliche Gehör und das Akteneinsichtsrecht tragen wesentlich zur Herstellung von Transparenz bei, indem sie den Rechtsunterworfenen ermöglichen, staatliches Handeln nachzuvollziehen und zu überprüfen.

Im Zusammenhang mit algorithmischen Entscheidungssystemen kommt der Transparenz besondere Bedeutung im Hinblick auf die Verhinderung von Diskriminierung zu. Transparenz über den Einsatz von ADMS einerseits und Transparenz über die zugrundeliegenden Daten andererseits erleichtern die Erkennung und Bekämpfung von Diskriminierung, indem sie die Nachvollziehbarkeit und die Überprüfbarkeit von Entscheidungen verbessern.

Das Öffentlichkeitsprinzip soll die Kontrolle der Verwaltung durch die Bevölkerung stärken und das Vertrauen in staatliche Institutionen fördern. Obwohl das BGÖ primär eine passive Informationstätigkeit der Verwaltung fordert, bleibt die Transparenz ein zentrales Anliegen, das den Zugang zu amtlichen

¹¹² Art. 25 (Grundsatz) und Art. 26 (Einschränkungen) DSG.

¹¹³ Art. 12 (Verzeichnis über die Bearbeitungstätigkeiten) und Art. 56 (Register der Datenbearbeitungstätigkeiten von Bundesorganen) DSG.

¹¹⁴ HERBERT BURKERT, Datenschutz auf dem Weg zur Transparenzordnung, in: Bruno Baeriswyl/Beat Rudin (Hrsg.), Perspektive Datenschutz: Praxis und Entwicklungen in Recht und Technik, Zürich 2002, 181 ff., 184.

Dokumenten ermöglicht. Dieses Einsichtsrecht muss auch bei staatlich eingesetzten ADMS gewährleistet sein.

Die Informationspflicht gemäss Art. 180 Abs. 2 BV und Art. 10 RVOG verlangt eine rechtzeitige und umfassende Information durch den Bundesrat. Mit dieser proaktiven Informationspflicht soll die Legitimität staatlichen Handelns sichergestellt und die Transparenz als integraler Bestandteil des Verwaltungshandelns gefördert werden. Eine konkrete Verpflichtung zur Führung eines ADMS-Verzeichnisses lässt sich daraus jedoch nicht ableiten.

Transparenz ist schliesslich ein zentrales Element des DSG, insbesondere bei der Bearbeitung von Personendaten. Dies schliesst auch die Datenbearbeitung durch algorithmische Systeme ein, wobei Transparenz die Nachvollziehbarkeit und Überprüfbarkeit der Datenbearbeitung sicherstellt.

Zusammenfassend zeigt sich, dass Transparenz ein wesentliches Prinzip des schweizerischen Rechtsstaates ist, das durch verfassungsrechtliche und gesetzliche Bestimmungen geschützt und gefördert wird. Die bestehenden Rechtsgrundlagen fördern zwar die Transparenz und sehen gewisse Informationspflichten vor, sind jedoch nicht hinreichend spezifisch, um daraus eindeutig eine behördliche Pflicht zur Führung eines öffentlichen ADMS-Verzeichnisses abzuleiten. Eine solche Verpflichtung müsste durch eine spezifische Rechtsvorschrift eingeführt werden. Diese liesse sich gut in das bestehende schweizerische Verständnis von Transparenz staatlichen Handelns einfügen. Sie entspräche dem Grundgedanken der übrigen Transparenzpflichten und würde dazu beitragen, die Nachvollziehbarkeit und Kontrolle staatlichen Handelns weiter zu stärken bzw. auf den Bereich des Einsatzes algorithmischer Entscheidungssysteme auszudehnen.

IV. Rechtliche Verankerung

Die bisherigen Ausführungen machen deutlich, dass die Schaffung eines ADMS-Verzeichnisses von verschiedenen Seiten gefordert wird und dass sich ein solches Verzeichnis nahtlos in den Grundgedanken der Transparenz des staatlichen Handels zur Sicherstellung von Nachvollziehbarkeit und Kontrolle einfügen würde. Vor diesem Hintergrund werden im Folgenden einige grundsätzliche Überlegungen zum Regelungsort (1) sowie zum Regelungsgegenstand (2) angestellt. Dabei wird von zwei Prämissen ausgegangen: Erstens, ein ADMS-Verzeichnis ist von staatlichen Stellen verpflichtend zu führen. Damit soll si-

chergestellt werden, dass das Verzeichnis vollständig ist.¹¹⁵ Zweitens, ein ADMS-Verzeichnis enthält alle zur Sicherstellung von Nachvollziehbarkeit und Kontrolle relevanten Informationen in einheitlicher Form.

1. Regelungsort

Um das Führen eines ADMS-Verzeichnisses rechtsverbindlich sicherzustellen, sollte die entsprechende Verpflichtung in einem Gesetz im formellen Sinn verankert werden.¹¹⁶ Die detaillierten Vorgaben zu den im Verzeichnis anzugebenden Informationen können anschliessend auf Verordnungsstufe konkretisiert werden.

Auf Bundesebene existieren verschiedene Gesetze, in denen die Verpflichtung zur Führung eines ADMS-Verzeichnisses vorgesehen werden könnte. Dabei ist massgeblich auf den Adressatenkreis des Verzeichnisses abzustellen. Eine Möglichkeit wäre die Verankerung im DSG. Dies entspräche dem Vorgehen des Kantons Zürich.¹¹⁷ Damit wären grundsätzlich alle Bundesorgane von der Registrierungspflicht erfasst.¹¹⁸ Als Bundesorgane gelten gemäss DSG alle Behörden und Dienststellen des Bundes sowie Private, die mit öffentlichen Aufgaben des Bundes betraut sind.¹¹⁹ Der Begriff der öffentlichen Aufgabe ist zwar nicht scharf konturiert, wird aber in Lehre und Rechtsprechung weit verstanden.¹²⁰ Der breit angelegte Adressatenkreis des DSG kann wesentlich zur Erreichung der mit dem ADMS-Verzeichnis verfolgten Ziele¹²¹ beitragen. Da die zu erfassenden ADMS insbesondere zur individuell-konkreten Entscheidungsfindung

¹¹⁵ Siehe dazu auch das Postulat 23.3566 Schlatter Marionna betreffend die Sicherstellung der Vollständigkeit des Transparenzregisters für den Einsatz von künstlicher Intelligenz durch die Bundesverwaltung vom 04.5.2023.

¹¹⁶ Gemäss Art. 164 Abs. 1 BV sind alle *wichtigen* rechtsetzenden Bestimmungen in einem Gesetz im formellen Sinn zu beschliessen. Was wichtig ist, ergibt sich aus von Lehre und Rechtsprechung entwickelten Kriterien (siehe dazu nur SGK BV-TSCHANNEN [FN 32], Art. 164 N 7).

¹¹⁷ Siehe [I](#).

¹¹⁸ Art. 2 Abs. 1 lit. b i. V. m. Art. 5 lit. i. DSG. Dies war schliesslich im Kanton Zürich der Grund, weshalb das Verzeichnis im IDG-ZH und nicht im Gesetz über die Organisation des Regierungsrates und der kantonalen Verwaltung Eingang fand, siehe dazu den Beschluss des Zürcher Regierungsrates (Regierungsratsbeschluss) Nr. 878/2023 vom 5. Juli 2023, 28 f.

¹¹⁹ Zum Begriff des Bundesorgans OFK DSG-JOSURAN-BINDER, Art. 5 lit. i N 3, 4, in: Adrian Bieri/ Julian Powell (Hrsg.), DSG, Orell Füssli Kommentar, Zürich 2023 (zit. OFK DSG).

¹²⁰ Siehe dazu ausführlich BERNHARD RÜTSCHKE, Was sind öffentliche Aufgaben?, recht 2013, 153 ff., 154 f.

¹²¹ Siehe zu den Zielen [II.2](#).

(Erlass von Verfügungen) eingesetzt werden,¹²² besteht regelmässig ein Bezug zu einer Person und damit zu einer Bearbeitung von Personendaten.¹²³

Eine andere Möglichkeit besteht darin, das am 1. Januar 2024 in Kraft getretene Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben (EMBAG) um ein entsprechendes Verzeichnis zu ergänzen. Grundsätzlich käme auch das RVOG als Regelungsort in Betracht. Das EMBAG ist jedoch spezifisch auf die Digitalisierung ausgerichtet und weist daher einen engeren Sachbezug für die Regelung eines ADMS-Verzeichnisses auf als das RVOG.¹²⁴ Mit dem EMBAG wurden die Voraussetzungen einerseits für die Zusammenarbeit zwischen Behörden verschiedener Gemeinwesen und mit Dritten beim Einsatz elektronischer Mittel zur Unterstützung der Erfüllung von Behördenaufgaben geschaffen¹²⁵ und andererseits der Ausbau und die Weiterentwicklung des Einsatzes elektronischer Mittel zur Unterstützung der Erfüllung von Behördenaufgaben ermöglicht¹²⁶. Algorithmische Systeme sind elektronische Systeme, weshalb die Pflicht zur Führung eines ADMS-Verzeichnisses im EMBAG verankert werden könnte. In der Botschaft zum EMBAG wird zudem explizit festgehalten, dass «ein Gesamtbild der geltenden Bestimmungen im Bereich der Digitalisierung»¹²⁷ angestrebt wird, wozu auch der Einsatz von algorithmischen Systemen gehört. Zudem sieht das EMBAG auch Rechtsgrundlagen für eine mögliche Kooperation mit den Kantonen vor.¹²⁸ Diese könnten auch für ein ADMS-Verzeichnis beigezogen werden, um eine möglichst schweizweit harmonisierte Lösung zu finden.

Ausschlaggebend für den hier skizzierten Vorschlag ist jedoch der gegenüber dem DSG stark eingeschränkte Adressatenkreis des EMBAG. Das Gesetz gilt grundsätzlich nur für die zentrale Bundesverwaltung.¹²⁹ Die Einführung der

¹²² BRAUN BINDER et al., Studie KI 2021 (FN 8), 24 ff.

¹²³ DAVID RECHSTEINER, Der Algorithmus verfügt, Jusletter 26.11.2018, N 10 f.

¹²⁴ Siehe etwa Botschaft zum Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben vom 4. März 2022, BBl 2022 804 (zit. Botschaft EMBAG 2022), 50.

¹²⁵ Art. 1 lit. a EMBAG.

¹²⁶ Art. 1 lit. b EMBAG.

¹²⁷ Botschaft EMBAG 2022 (FN 124), 116.

¹²⁸ Art. 4 EMBAG.

¹²⁹ Art. 2 Abs. 1 EMBAG. Der Bundesrat kann allerdings auch Verwaltungseinheiten der dezentralen Bundesverwaltung, die Parlamentsdienste, die eidgenössischen Gerichte und die Bundesanwaltschaft dem Gesetz oder Teilen davon unterstellen (Art. 2 Abs. 2 und Abs. 3 EMBAG). Siehe auch den Vorbehalt in Art. 2 Abs. 4 EMBAG, wonach das Gesetz nur gilt, soweit andere Bundeserlasse nichts Abweichendes vorsehen.

Verpflichtung zur Führung eines ADMS-Verzeichnisses im DSG erscheint daher zielführender.

2. Regelungsgegenstand

2.1. Grundlegende Informationen im Gesetz im formellen Sinn

Auf Bundesebene gibt Art. 164 Abs. 1 BV Anhaltspunkte, welche Vorgaben in einem Gesetz im formellen Sinn zu statuieren sind. Insbesondere in technisch komplexen Bereichen – wie im vorliegenden Fall der algorithmischen Systeme – beschränken sich die formell-gesetzlichen Bestimmungen regelmässig auf Grundsätzliches, während die Konkretisierung auf Verordnungsstufe erfolgt.¹³⁰ Im Gesetz selbst wäre daher erstens die Stelle, die ein solches Verzeichnis führt, vorzusehen. Es bietet sich an, die in der Schweiz bestehenden Strukturen zu nutzen und die Aufgaben des Eidgenössischen Datenschutz- und Öffentlichkeitsbeauftragten (EDÖB) um die Führung eines ADMS-Verzeichnisses zu erweitern. Aber auch eine Zuweisung an das CNAI, das bereits die freiwillige Datenbank über KI-Systeme der Bundesverwaltung führt, wäre denkbar.

Zweitens müsste festgehalten werden, welche Systeme unter die Veröffentlichungspflicht fallen. Wie bereits erörtert, sollten nur Systeme erfasst werden, deren Einsatz im Rahmen der Entscheidungsfindung unmittelbare Auswirkungen auf Rechte und Pflichten von Menschen nach sich zieht.¹³¹ Die vom CNAI geführte Projektdatenbank der auf Bundesebene eingesetzten KI-Systeme trifft aktuell keine Unterscheidung zwischen problembehafteten und nicht problembehafteten algorithmischen Systemen und enthält verschiedene Systeme, die für die Rechte und Pflichten von Personen unproblematisch sind.¹³² Dies mag aus Sicht der Zielsetzung des CNAI, das verwaltungsinterne Know-how hinsichtlich KI zu vernetzen, durchaus sinnvoll sein. Die Ziele des öffentlichen ADMS-Verzeichnisses, um die es in diesem Beitrag geht,¹³³ werden durch die Aufnahme nicht relevanter Systeme eher behindert als gefördert.

¹³⁰ SGK BV-TSCHANNEN (FN 32), Art. 164 N 8.

¹³¹ Siehe dazu [II.3](#).

¹³² Etwa führt die Projektdatenbank das KI-System «SwissPollen» auf, das in Echtzeit Luftteilchen identifiziert, um die Pollenmessung zu automatisieren (S. 10 des PDF-Dokuments, Version 8, siehe FN 10).

¹³³ Siehe [II.2](#).

Drittens ist die Pflicht der Behörden, die von ihnen eingesetzten ADMS in das Verzeichnis einzutragen, gesetzlich zu verankern. Die Konkretisierung der einzelnen Transparenzelemente kann sodann grundsätzlich auf Verordnungsstufe erfolgen. Sollten diese Elemente allerdings politisch stark umstritten sein, wäre eine Klärung durch die Legislative bereits im Gesetz sinnvoll. Unabhängig davon, auf welcher Regelungsstufe die Konkretisierung erfolgt, sollte diese zwei grundlegende Ziele berücksichtigen: Zum einen soll aus dem ADMS-Verzeichnis klar hervorgehen, wo algorithmische Systeme in den Behörden eingesetzt werden.¹³⁴ Zum anderen sollen aussagekräftige, verständliche Informationen über die eingesetzten Systeme veröffentlicht werden, die je nach Interessengruppe technisch mehr oder weniger detailliert ausfallen können.¹³⁵

2.2. Konkretisierungen auf Verordnungsstufe

Zu den Elementen, die auf Verordnungsstufe erfasst werden können, zählen Basisinformationen wie eine Übersicht, Informationen zum verantwortlichen Organ sowie andere Referenzen, die weiterführende Informationen oder Prüfungsberichte liefern können.¹³⁶ Ferner sind der spezifische Einsatzbereich und die Rechtsgrundlage, die den Einsatz des ADMS legitimiert, anzugeben. Zudem sind Informationen zur Art und Herkunft der bearbeiteten Daten sowie zu den Metadaten zu liefern. Werden Personendaten bearbeitet, ist insbesondere der Bearbeitungszweck zu erfassen.¹³⁷ Auch die Funktionsweise der Algo-

¹³⁴ BARBARA THIES/STEFAN HEUSER, Akzeptanz neuer Technologien, in: Kuuya Chibanguza/Christian Kuss/Hans Seege (Hrsg.), Künstliche Intelligenz, Baden-Baden 2022, 59 ff., N 17; DOLLINGER (FN 46), 45.

¹³⁵ HAATAJA/VAN DE FLIERT/RAUTIO (FN 42), 3.

¹³⁶ Diese und die folgenden Ausführungen basieren auf einem open source Standard des Städtetzwerks «Eurocities», der helfen soll, einheitliche und vergleichbare schematische Darstellungen zum Einsatz algorithmischer Systeme in öffentlichen Verwaltungen zu schaffen. Zum Netzwerk Internet: <https://eurocities.eu> (Abruf 7.8.2024), zum Standard Internet: <https://www.algorithmregister.org/standard> (Abruf 7.8.2024). Siehe dazu auch PIA SOMBETZKI/MAITTHIAS SPIELKAMP, Ein KI-Transparenz-Register für die öffentliche Verwaltung, Konzeptpapier 2023, Internet: https://algorithmwatch.org/de/wp-content/uploads/2023/03/Konzept_KI-Transparenzregister_AlgorithmWatch_2023.pdf (Abruf 7.8.2024). Eine zusammengefasste Übersicht zu relevanten Angaben für algorithmische Systeme in einem öffentlichen Verzeichnis findet sich schliesslich auch bei MURAD (FN 42), 13 f. und NADIA BRAUN BINDER/LILIANE OBRECHT, Transparenz mittels ADMS-Verzeichnis, White Paper Juni 2024, 5 f., Internet: https://ius.unibas.ch/fileadmin/user_upload/ius/03_Forschung/10_e-PIAF/WhitePaper_Verzeichnis_Juni24.pdf (Abruf 7.8.2024).

¹³⁷ Siehe auch Art. 6 Abs. 3 DSGVO.

rithmen ist zu dokumentieren. Wird ein System verwendet, das mit Trainings-, Validierungs- und oder Testdaten¹³⁸ entwickelt wurde, sollte ersichtlich sein, welche Daten für das Training, die Validierung oder das Testen genutzt wurden. Schliesslich sind auch die Auswirkungen des ADMS auf die Bevölkerung in der Rechtsgrundlage zu beschreiben. Insbesondere sind Methoden anzugeben, die eingesetzt wurden, um das Diskriminierungsrisiko zu verhindern sowie die Ergebnisse einer Folgenabschätzung, in der Risiken für die Grundrechte der betroffenen Personen genannt werden.

2.3. Verständlichkeit und Zugänglichkeit der Informationen

Ein Verzeichnis ist ein geordnetes Informationssystem, das interessierte Personen über einen bestimmten Sachbereich orientiert.¹³⁹ Es ist öffentlich, sobald jede Person zu einem unspezifischen Zeitpunkt und ohne spezifisches Interesse in dieses Einsicht nehmen kann.¹⁴⁰ Im vorliegenden Kontext bietet es sich an, das ADMS-Verzeichnis online zugänglich zu machen,¹⁴¹ idealerweise in maschinenlesbarer und durchsuchbarer Form.¹⁴²

Um den Transparenzanforderungen gerecht zu werden, müssen möglichst alle Faktoren, die dem ADMS zugrunde liegen, offengelegt werden.¹⁴³ Dabei gilt allerdings die mit den meisten Digitalisierungsvorhaben verbundene Erkenntnis, dass regelmässig viele Informationen anfallen.¹⁴⁴ Bei der Informationsvermittlung geht es jedoch nicht nur um die Erhöhung der Informations-

¹³⁸ Die Unterscheidung zwischen Trainings-, Validierungs- und Testdaten folgt der Unterteilung des AI Acts, siehe insbesondere Art. 10 AI Act.

¹³⁹ Siehe Duden zum Begriff «Verzeichnis», Internet: <https://www.duden.de/rechtschreibung/Verzeichnis> (Abruf 7.8.2024).

¹⁴⁰ So die Definition zur Öffentlichkeit des Handelsverzeichnisses, siehe BSK OR II-ECKERT, Art. 930 N 1, in: Heinrich Honsell/Nedim Peter Vogt/Rolf Watter (Hrsg.), Obligationenrecht II, Basler Kommentar, 5. A., Basel 2016 (zit. BSK OR II-Verfasser). Siehe im spezifischen Kontext zu ADMS-Verzeichnissen MURAD (FN 42), 16.

¹⁴¹ Siehe dazu SOMBETZKI/SPIELKAMP (FN 136).

¹⁴² Aus der Politik bemängelt Marionna Schlatter im Postulat 23.3566 (FN 1) u. a., dass die vom CNAI geführte Projektdatenbank über die in der Bundesverwaltung eingesetzten KI-Systeme lediglich als PDF-Liste und nicht in Form einer durchsuchbaren Datenbank verfügbar ist.

¹⁴³ WOLFGANG HOFFMANN-RIEM, *Recht im Sog der digitalen Transformation*, Tübingen 2022, 200 f.

¹⁴⁴ DANIEL J. CARON/ANDREAS KELLERHALS, *Gestion documentaire, archivage et transparence: une relation équivoque?*, in: Martial Pasquier (Hrsg.), *Le principe de transparence en Suisse et dans le monde*, Lausanne 2013, 197 ff., 206.

menge, sondern auch um die Reduktion von Komplexität.¹⁴⁵ Es geht demnach um die Verständlichkeit der im Verzeichnis zugänglich gemachten Informationen. Dies bezieht sich sowohl auf die Sprache¹⁴⁶ als auch auf die Lesbarkeit der Informationen. Letzteres kann z. B. mittels eines standardisierten Protokolls oder strukturierten Verfahrens erreicht werden.¹⁴⁷ Zentral ist dabei stets die Schaffung adressatengerechter Transparenz.¹⁴⁸ Neben leicht verständlichen Informationen für Laien scheint es daher sinnvoll, eine ausführlichere Dokumentation für Expertinnen und Experten vorzusehen.

V. Schlussbemerkung

Ein ADMS-Verzeichnis dient dazu, Transparenz und Kontrolle bei der Nutzung algorithmischer Systeme in der öffentlichen Verwaltung zu gewährleisten, um den Schutz der Grundrechte, die Kontrolle staatlichen Handelns und das Vertrauen der Bevölkerung in den Staat zu fördern. Das Verzeichnis bietet eine dokumentierte, online durchsuchbare Übersicht über die eingesetzten Systeme und ermöglicht der interessierten Öffentlichkeit einen direkten Zugang zu Informationen. Um die Führung eines ADMS-Verzeichnisses rechtlich verbindlich sicherzustellen, sollten diese Verpflichtung in einem formellen Gesetz verankert und die detaillierten Vorgaben auf Verordnungsebene konkretisiert werden. Geeignete Regelungsorte wären das DSG oder das EMBAG, wobei das DSG aufgrund seines breiteren Adressatenkreises bevorzugt wird. Auf Verordnungsebene sollten spezifische Transparenzelemente und die Verständlichkeit der Informationen klar geregelt sein. Ein verbindliches ADMS-Verzeichnis liesse sich gut in die schweizerische Rechtsordnung integrieren, da die Transparenz staatlichen Handelns bereits ein stark verankerter Grundsatz ist, der durch verfassungsrechtliche und gesetzliche Bestimmungen die demokratische Kontrolle und die Begrenzung staatlicher Macht ermöglicht.

¹⁴⁵ Siehe auch die Richtlinien zur Transparenz unter der DSGVO vom 11.4.2018, 6 ff., Internet: https://www.datenschutz-bayern.de/datenschutzreform2018/wp260rev01_de.pdf (Abruf 7.8.2024).

¹⁴⁶ MURAD (FN 42), 20, schlägt Erklärungen in klaren, ausformulierten Sätzen vor.

¹⁴⁷ Siehe AlgorithmWatch CH, Empfehlungen (FN 9); SOMBETZKI/SPIELKAMP (FN 136), 2.

¹⁴⁸ ROLF H. WEBER, Künstliche Intelligenz: Regulatorische Überlegungen zum «Wie» und «Was», EuZ 2022, B1 ff., B12 f.; MARKUS SCHEFER/RAPHAELA CUENI, Öffentlichkeit im Wandel: Überlegungen aus grundrechtlicher Sicht, Studie zuhanden des BAKOM vom 7. Juli 2020, 69 f. zeigen dies im Kontext der Schaffung von Transparenz des Einsatzes von Algorithmen durch Informationsintermediäre auf, Internet: <https://www.bakom.admin.ch/bakom/de/home/elektronische-medien/studien/einzelstudien.html> (Abruf 22.4.2025).

Maschinelle Gesichtserkennung im öffentlichen Raum

Nadja Braun Binder / Eliane Kunz / Liliane Obrecht

Erstabdruck in: *sui generis*, 2022, S. 53–61

Der Einsatz von maschineller Gesichtserkennung im öffentlichen Raum birgt die Gefahr einer gesellschaftlichen Massenüberwachung. Werden dabei biometrische Daten, die ein Individuum eindeutig identifizieren, genutzt, handelt es sich nach dem neuen Datenschutzgesetz um eine Bearbeitung besonders schützenswerter Personendaten. Die maschinelle Gesichtserkennung ist in der Schweiz regulatorisch kaum umrissen. Dabei sind Überlegungen dazu dringend geboten. Dies ergibt sich einerseits aus der Grundrechtsrelevanz der Thematik, andererseits fordern zivilgesellschaftliche Organisationen ein entsprechendes Verbot. Auch die geplante EU-Regelung zur Künstlichen Intelligenz, in der die Gesichtserkennung enthalten ist, sollte in den Erwägungen berücksichtigt werden. Vor diesem Hintergrund gibt der vorliegende Aufsatz einen Überblick über die im Zusammenhang mit maschineller Gesichtserkennung im öffentlichen Raum auftretenden Rechtsfragen und befasst sich mit der Frage der Notwendigkeit eines Verbots oder Moratoriums.

L'utilisation de la reconnaissance faciale automatique dans l'espace public comporte le risque d'une surveillance de masse de la société. Si des données biométriques identifiant clairement un individu sont utilisées, il s'agit, selon la nouvelle loi sur la protection des données, d'un traitement de données personnelles sensibles. En Suisse, la reconnaissance faciale automatique n'est guère encadrée par la réglementation. Il est pourtant urgent de réfléchir à cette question. Cela résulte, d'une part, de l'importance de cette question en termes de droits fondamentaux et, d'autre part, de la volonté de certaines organisations de la société civile d'interdire une telle utilisation de ces données. La réglementation européenne prévue en matière d'intelligence artificielle, qui inclut la reconnaissance faciale, devrait également être prise en compte dans les réflexions. Dans ce contexte, le présent article donne un aperçu des questions juridiques qui se posent s'agissant de la reconnaissance faciale automatique dans l'espace public et se penche sur la question de la nécessité d'une interdiction ou d'un moratoire.

Inhalt

I.	Einleitung	142
II.	Grundlagen	143
1.	Das Gesicht als Teil biometrischer Daten	143
2.	Maschinelle Gesichtserkennung	144
3.	Definition des öffentlichen Raumes	145
4.	Einsatzbereiche	145
4.1.	Echtzeit-Gesichtserkennung	145
4.2.	Nachträgliche Gesichtserkennung	147
4.3.	Exkurs: Verzicht auf maschinelle Gesichtserkennung	149
III.	Rechtliche Würdigung	150
1.	Datenschutzrechtliche Vorgaben	150
2.	Gesichtserkennung im öffentlichen Raum durch Private	152
3.	Gesichtserkennung im öffentlichen Raum durch staatliche Behörden	154
3.1.	Informationelle Selbstbestimmung	154
3.2.	Versammlungs- und Meinungsfreiheit	156
3.3.	Diskriminierungsverbot	157
IV.	Moratorium statt Verbot	158
V.	Fazit	160

I. Einleitung

In annähernd 100 Staaten werden Technologien zur Gesichtserkennung verwendet.¹ Gleichzeitig schlägt die EU-Kommission ein grundsätzliches Verbot biometrischer Fernidentifizierungssysteme im öffentlichen Raum vor – wenn auch mit weitreichenden Ausnahmen.² In der Schweiz steht das Thema Gesichtserkennung seit Kurzem ebenfalls auf der politischen Agenda.³ Zudem setzen sich zivilgesellschaftliche Organisationen für ein Verbot maschineller

¹ Siehe dazu die Übersichtskarte des britischen Unternehmens Surfshark, The Facial Recognition World Map, abrufbar unter <https://surfshark.com/facial-recognition-map>.

² Europäische Kommission, Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union vom 21.04.2021 (EU-KI-Verordnungsentwurf).

³ Bund: Glättli Ip. 21.3580. Stadt Basel: Baumgartner und Konsorten Anzug 22.5022.01. Stadt Zürich: Maggi/Wey Po. 2021/451, Mo. 2021/450. Stadt Lausanne: Gaillard Po. No8/23.11.2021, projets de règlement No8/07.12.2021 und No8/07.12.2021.

Gesichtserkennung und insbesondere biometrischer Massenüberwachung ein.⁴ In der Schweiz ist die juristische Auseinandersetzung mit dem Thema dagegen bislang überschaubar geblieben.⁵ Bei den meisten Beiträgen handelt es sich um Zeitungs- oder Blogartikel.⁶

II. Grundlagen

Die rechtliche Befassung mit maschineller Gesichtserkennung setzt eine Klärung von deren zentralen Elementen voraus (1.-3.). Zudem soll kurz skizziert werden, wo Gesichtserkennung zum Einsatz kommt (4.).

1. Das Gesicht als Teil biometrischer Daten

Biometrische Merkmale sind messbare, körpereigene, physiologische oder verhaltensspezifische Kennzeichen, die einer Person zugeordnet werden können, grundsätzlich einzigartig sind und nur unter grossem Aufwand verändert werden können.⁷ Das Gesichtsbild ist ein physiologisches biometrisches Merkmal, das vorgegeben und grundsätzlich unveränderbar ist.⁸

⁴ Petition von AlgorithmWatch, Amnesty International & Digitale Gesellschaft Schweiz, abrufbar unter <https://algorithmwatch.ch/de/gesichtserkennung-stoppen/>.

⁵ Siehe etwa SIMMLER MONIKA/CANOVA GIULIA, Gesichtserkennungstechnologie: Die «smarte» Polizeiarbeit auf dem rechtlichen Prüfstand, in: Sicherheit & Recht 3/2021, S. 105 ff.; KEIST RAMONA, Gesichtserkennung im zivilrechtlichen Persönlichkeitsschutz, in: Jusletter vom 30.05.2019; MATTER LIVIA, Gesichtserkennung auf dem Vormarsch, in: digma 2019, S. 14 ff.

⁶ Siehe z. B. LUCHETTA SIMONE, So jagen Schweizer Polizisten mit Gesichtserkennung Verbrecher, in: Tagesanzeiger vom 17.04.2021, abrufbar unter <https://www.tagesanzeiger.ch/so-jagen-schweizer-polizisten-mit-gesichtserkennung-verbrecher-608167461846>; STEIGER MARTIN, Gesichtserkennung: Drei populäre datenschutzrechtliche Irrtümer, abrufbar unter <https://steigerlegal.ch/2020/11/21/gesichtserkennung-datenschutz-irrtuemer/>; CICHOCKI MICHAL, Europarat veröffentlicht Leitlinien zu Gesichtserkennung (Guidelines on facial recognition), abrufbar unter <http://www.lawblogswitzerland.ch/2021/04/europa-rat-veroeffentlicht-leitlinien-zu.html>.

⁷ EDÖB, Leitfaden zu biometrischen Erkennungssystemen, 2009, S. 5; BLONSKI DOMINIKA, Biometrische Daten als Gegenstand des informationellen Selbstbestimmungsrechts, Bern 2015, S. 6; privatim, Leitfaden zur datenschutzrechtlichen Beurteilung von biometrischen Verfahren, 2005, S. 3.

⁸ BLONSKI (Fn. 7), S. 8. Auf Veränderungen des Gesichtsbildes mittels chirurgischer Eingriffe, durch Krankheiten, Unfälle oder Wachstum wird hier nicht näher eingegangen.

Das revidierte Datenschutzgesetz⁹ ordnet die biometrischen Daten, die eine natürliche Person eindeutig identifizieren, der Kategorie der besonders schützenswerten Personendaten zu und verankert den Begriff neu explizit in Art. 5 lit. c Ziff. 4 revDSG. Demnach handelt es sich um Personendaten, «die durch ein spezifisches technisches Verfahren zu den physischen, physiologischen oder verhaltenstypischen Merkmalen eines Individuums gewonnen werden und die eine eindeutige Identifizierung der betreffenden Person ermöglichen oder bestätigen»¹⁰. Als Beispiele nennt die Botschaft digitale Fingerabdrücke, Gesichtsbilder, Bilder der Iris oder Aufnahmen der Stimme. Zwingend ist, dass diese Daten auf einem Verfahren beruhen, welches die eindeutige Identifizierung oder Authentifizierung einer Person erlaubt.¹¹ Dies wäre bei unscharfen Bildern oder Teilabdrücken eines Fingers unter Umständen nicht der Fall.

2. Maschinelle Gesichtserkennung

Maschinelle Gesichtserkennung kann definiert werden als «automatische Verarbeitung digitaler Bilder, die Gesichter von natürlichen Personen enthalten, um bei diesen eine Identifizierung, Authentifizierung bzw. Verifizierung oder Kategorisierung durchzuführen»¹².

Bei der Identifizierung wird das Ermitteln der Identität einer Person aus einer Vielzahl anderer Personen heraus angestrebt – mithin die Beantwortung der Frage, wer eine bestimmte Person ist.¹³ Das Authentifizierungs- bzw. Verifizierungsverfahren dient der Identitätsbestätigung einer Person, also der Abklärung, ob eine Person tatsächlich diejenige ist, als die sie sich ausgibt.

Die im Zusammenhang mit maschineller Gesichtserkennung auftretenden Rechtsfragen stellen sich grundsätzlich unabhängig davon, welche konkrete Technologie

⁹ Das Datenschutzgesetz (DSG, SR 235.1) wurde revidiert. Das revidierte Datenschutzgesetz (revDSG) tritt voraussichtlich am 1. September 2023 in Kraft. Die weiteren Ausführungen werden auf die Rechtslage nach dem revDSG gestützt, wobei auch Bezug auf bisherige Literatur genommen wird, soweit die entsprechenden Ausführungen auch auf die Rechtslage nach dem revDSG zutreffen.

¹⁰ Botschaft zum Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz vom 15. September 2017 (BBl 2017 6941 ff.), S. 7020.

¹¹ BBl 2017 6941 ff. (Fn. 10), S. 7020.

¹² Artikel-29-Datenschutzgruppe, Stellungnahme 02/2012 zur Gesichtserkennung bei Online- und Mobilfunkdiensten, 00727/12/DE, WP 192, Brüssel 22.03.2012, S. 2.

¹³ EDÖB (Fn. 7), S. 5; zur Technik der Gesichtserkennung vertieft KARIOTH GÜNTER, Fähigkeiten der Gesichtserkennung, in: digma 2019, S. 6 ff.

eingesetzt wird und ob man in diesem Zusammenhang von Künstlicher Intelligenz (KI) spricht.¹⁴ Zentral sind vielmehr die damit einhergehende Gefahr der Massenüberwachung, die eine breit angelegte, staatliche Überwachung der Bevölkerung meint, sowie der Umstand, dass biometrische Daten bearbeitet werden.

3. Definition des öffentlichen Raumes

In der Schweiz existiert keine gesetzliche Definition des öffentlichen Raumes. In der Literatur wird vorgeschlagen, den Begriff «Raum» als nicht abgeschlossene bzw. nicht fest eingegrenzte Ausdehnung zu verstehen.¹⁵ Öffentlichkeit kann als der Allgemeinheit zugänglich und benutzbar interpretiert werden, unabhängig davon, ob das Eigentum am Raum in öffentlicher oder privater Hand liegt.¹⁶ Die EU-Kommission definiert in Art. 3 Nr. 39 EU-KI-Verordnungsentwurf den öffentlich zugänglichen Raum als «einen der Öffentlichkeit zugänglichen physischen Ort, unabhängig davon, ob dafür bestimmte Zugangsbedingungen gelten».

Für die Zwecke des vorliegenden Aufsatzes wird der öffentliche Raum in Anlehnung an die oben skizzierten Definitionen als für Personen physisch zugänglicher Ort verstanden, unabhängig davon, ob dafür bestimmte Zugangsbedingungen gelten und unabhängig der Eigentumsverhältnisse. Der Cyberspace bzw. das Internet werden nicht zum physisch zugänglichen Raum gezählt und damit Erwägungen zur Online-Überwachung ausgeklammert.

4. Einsatzbereiche

4.1. Echtzeit-Gesichtserkennung

Gesichtserkennungssysteme können in Echtzeit Daten aus Überwachungskameras zur Identifizierung oder Verifizierung bestimmter Personen auswerten. Diese Möglichkeit wird etwa im Rahmen von automatisierten Passkontrollen an Flughäfen genutzt, indem biometrische Erkennungsmerkmale des Gesichts mit den auf dem Pass gespeicherten Daten abgeglichen werden. Die Daten können ausserdem registriert sowie mithilfe von Fahndungs- und Informationssystemen überprüft werden.¹⁷

¹⁴ HORNUNG GERRIT/SCHINDLER STEPHAN, Datenschutz bei der biometrischen Gesichtserkennung, in: *Datenschutz und Datensicherheit*, 8/2021, S. 515 ff.

¹⁵ Siehe ZUMSTEG PATRICE MARTIN, *Demonstrationen in der Stadt Zürich*, Zürich 2020, S. 6 f.

¹⁶ ZUMSTEG (Fn. 15), S. 7.

¹⁷ Dies ist der Fall am Basler und Genfer Flughafen. Es ist anzumerken, dass die französische Grenzpolizei die Betreiberin dieser Geräte ist, was auf einen Staatsvertrag zwischen der

Ein weiterer denkbarer Einsatzbereich in der Schweiz könnte in Zukunft die Zutrittskontrolle zu Sportstadien sein. Mithilfe maschineller Gesichtserkennung sollen registrierte Hooligans direkt beim Eingang erkannt werden.¹⁸ Solche Systeme wurden und werden international bereits eingesetzt.¹⁹ In der Schweiz hat ein Start-up ausserdem im Sommer 2021 die Idee lanciert, das Vorhandensein des Covid-Zertifikats mittels Gesichtserkennung zu prüfen.²⁰

Echtzeit-Gesichtserkennung kommt in Europa vereinzelt zum Einsatz, um bestimmte, geografisch eingegrenzte Gebiete zu überwachen. Die gescannten

Schweiz und Frankreich zurückzuführen ist. Dazu BALLMER DANIEL, Ein Scan fürs Gesicht: Der Basler Euro-Airport beschleunigt die Passkontrolle, in: bz vom 09.02.2019, abrufbar unter <https://www.bzbasel.ch/amp/basel/basel-stadt/ein-scan-furs-gesicht-der-basler-euro-airport-beschleunigt-die-passkontrolle-ld.134381>; Genève aéroport communiqué, Genève aéroport teste le contrôle automatisé à la frontière vom 05.12.2019, abrufbar unter <https://www.gva.ch/fr/Site/Geneve-Aeroport/News/2020-2016/Geneve-Aeroport-teste-contrôle-automatisé>. Am Flughafen Zürich hingegen findet keine solche Registrierung und Abgleichung statt, nur während eines Pilotprojekts im Jahr 2003; vgl. dazu den Regierungsratsbeschluss Nr. 447/2021, Auszug aus dem Protokoll des Regierungsrates des Kantons Zürich, Sitzung vom 5. Mai 2021, KR-Nr. 128/2021, S. 2. Zu einer Übersicht über Europa vgl. LIEBHERR CHARLES, Gesichtserkennung: Einführung durch die Hintertür, Eine Studie im Auftrag der Grünen Fraktion im EU-Parlament zum Einsatz von Gesichtserkennung in Europa, in: Schweizer Radio und Fernsehen (SRF), Echo der Zeit vom 25.10.2021, abrufbar unter <https://www.srf.ch/audio/echo-der-zeit/gesichtserkennung-einfuehrung-durch-die-hintertuer?partId=1207745>.

¹⁸ Siehe zu Bern: Vadiant.net AG, abrufbar unter <http://www.fussball.ch/Pilotprojekt+mit+Gesichtserkennung/227121/detail.htm>; zum Kanton Wallis: CALOZ JULIEN, Gesichtserkennung beim FC Sion? Datenschützer hätte nichts dagegen, in: Watson vom 15.07.2021, abrufbar unter <https://www.watson.ch/sport/fussball/916971444-gesichtserkennung-beim-fc-sion-datenschuetzer-haette-nichts-dagegen>.

¹⁹ Siehe z. B. zu Dänemark: OVERGAARD SIDSEL, A Soccer Team In Denmark Is Using Facial Recognition To Stop Unruly Fans, in: NPR vom 21.10.2019, abrufbar unter <https://www.npr.org/2019/10/21/770280447/a-soccer-team-in-denmark-is-using-facial-recognition-to-stop-unruly-fans?t=1631265914732>.

²⁰ MARON HANS JOERG, Schweizer Startup kombiniert Covid-Zertifikat-Scan mit Gesichtserkennung, in: inside channels vom 03.11.2021, abrufbar unter <https://www.inside-channels.ch/de/post/schweizer-startup-kombiniert-covid-zertifikat-scan-mit-gesichtserkennung-20211103>. Bis zur Aufhebung der COVID-Massnahmen im Februar 2022 wurde der Covid-Checker im Hotel Dolder in Zürich eingesetzt, vgl. dazu HUBER MARTIN, Dolder Grand setzt auf Gesichtserkennung, Zürcher Tagblatt vom 24.11.2021, abrufbar unter <https://www.tagblatt.ch/limmattal/zueroch/zueroch-gesichtserkennung-statt-covid-zertifikat-das-dolder-setzt-auf-eine-etwas-andere-kontrolle-ld.2218608>.

Gesichter werden mit Überwachungslisten der Polizei oder der Gerichte abgeglichen.²¹

Vermehrt setzen auch Private in Europa Gesichtserkennung ein, wie z. B. Einkaufsgeschäfte,²² die mithilfe von Gesichtserkennung Personen, die in der Vergangenheit gestohlen oder sich nicht regelkonform verhalten haben, beim Betreten des Geschäfts erfassen und dem Personal melden.²³ Ein ähnliches Phänomen ist in den USA zu beobachten, wo lokale Geschäfte Überwachungskameras installieren, die ihre Aufnahmen in Echtzeit an die Polizei übermitteln, die diese daraufhin durch Fahndungs- und Führerschein-Datenbanken laufen lässt und so virtuelle Patrouillen durchführt.²⁴ Ob bzw. in welchem Umfang Gesichtserkennung in der Schweiz im Einzelhandel und in Einkaufszentren eingesetzt wird, ist unklar.²⁵

4.2. Nachträgliche Gesichtserkennung

Bei der nachträglichen Gesichtserkennung geht es um den Einsatz einer Gesichtserkennungssoftware im Nachgang zu einer Foto- oder Videoaufnahme – oftmals nach einer Straftat und während der Strafverfolgung.²⁶ In der Schweiz setzt die Polizei vereinzelt Gesichtserkennung zum Abgleich von Fahndungsbildern ein, wobei eine kantonale Polizeidatenbank bekannter Straftäterinnen und Straftäter mit einer anderen, bspw. kantonsübergreifenden Datenbank, abgeglichen wird.²⁷ Die Fahndungsbilder stammen oft aus Überwachungskameras von

²¹ Siehe etwa zu Grossbritannien: ZIADY HANNA, London will use live facial recognition cameras to police the city, in: CNN vom 24.01.2020, abrufbar unter <https://edition.cnn.com/2020/01/24/tech/london-live-facial-recognition/index.html>.

²² Zu den Einsatzmöglichkeiten und Grenzen von Gesichtserkennung im Detailhandel vgl. GLATTHAAR MATTHIAS, Gesichtserkennung im Supermarkt?, in: digma 2019, S. 20 ff.

²³ Siehe z. B. zu Grossbritannien: LEWIS GARETH, Protecting our customers and colleagues, abrufbar unter <https://www.facewatch.co.uk/2020/10/05/facewatch-at-the-southern-co-op/>.

²⁴ Projekt «Green Light» in Detroit, vgl. JONES CHRISTOPHER, Law enforcement use of facial recognition: bias, disparate impacts on people of color, and the need for federal legislation, in: North Carolina Journal of Law & Technology 22/2021, S. 777 ff., S. 790.

²⁵ Studie von AlgorithmWatch Schweiz, abrufbar unter <https://algorithmwatch.ch/de/kennt-der-supermarkt-ihr-gesicht/>; zudem weiter ODEHNAL BERNHARD, Spionagefirma bietet Gesichtserkennung von Shoppern an, in: Tagesanzeiger vom 19.11.2021, abrufbar unter <https://www.tagesanzeiger.ch/spionagefirma-bietet-gesichtserkennung-von-shoppnern-an-46762371263>.

²⁶ SIMMLER/CANOVA (Fn. 5), S. 8 ff.

²⁷ Vgl. für weitere Informationen zu den einzelnen Gesichtserkennungssoftwares LUCHETTA (Fn. 6). Für Aufruhr sorgte auch die Verwendung der Gesichtserkennungssoftware «Clearview AI» durch einige Angehörige der Stadtpolizei Zürich, die interne Testuntersuchungen

Geschädigten, wie z. B. eines bestohlenen Kleiderladens. Die konkrete Identifizierung der verdächtigten Person obliegt nach wie vor einer Polizistin oder einem Polizisten und nicht einer Software.²⁸

Nachträgliche Gesichtserkennung wird auch im internationalen Kontext insbesondere im Polizeibereich eingesetzt. Interpol nutzt seit 2016 Gesichtserkennung mit einer Gesichtsbilderdatenbank von Personen aus über 179 Staaten zur Identifizierung und Verifizierung; die Ergebnisse werden anschliessend von Hand überprüft.²⁹ In Europa wie auch den USA wird in vereinzelt Staaten bzw. Städten nachträgliche Gesichtserkennung ausserdem eingesetzt, um bei Verdacht auf eine vorsätzlich begangene Straftat einen nachträglichen digitalen Abgleich von Bildern von Überwachungskameras oder Handyaufnahmen mit einer Polizeidatenbank durchzuführen.³⁰ Einsatzszenarien sind bspw. eskalierte Demonstrationen.³¹

durchführten, sie allerdings nicht für Strafverfolgungszwecke einsetzten; vgl. dazu LUCHETTA SIMONE, Zürcher Stadtpolizist testete Clearview nach einer Ausbildung, in: Tagesanzeiger vom 06.09.2021, abrufbar unter <https://www.tagesanzeiger.ch/zuercher-stadtpolizist-testete-clearview-nach-einer-ausbildung-696840503794>.

²⁸ GERNY DANIEL, Wie intelligente Kameras in der Schweiz Einzug halten, in: NZZ online vom 22.06.2018, abrufbar unter <https://www.nzz.ch/schweiz/wie-intelligente-kameras-in-der-schweiz-einzug-halten-ld.1396204>; SCHWENDENER THOMAS, Facial Recognition soll bald St. Galler Polizei unterstützen, in: Inside IT vom 18.06.2020, abrufbar unter <https://www.inside-it.ch/de/post/facial-recognition-soll-bald-st-galler-polizei-unterstuetzen-20200618>.

²⁹ Facial Recognition, abrufbar unter <https://www.interpol.int/How-we-work/Forensics/Facial-Recognition>. Darüber hinaus wurde Clearview auch von der Interpol-Einheit «Verbrechen gegen Kinder» mit einem 30-tägigen Testkonto getestet, um die Opfer von sexuellem Kindesmissbrauch im Internet zu identifizieren; vgl. dazu ANZ PHILIPP, Wer alles die Gesichtserkennung von Clearview testete, in: Inside IT vom 30.08.2021, abrufbar unter <https://www.inside-it.ch/de/post/wer-alles-die-gesichtserkennung-von-clearview-testete-20210830>.

³⁰ Siehe zu Österreich: Beantwortung der parlamentarischen Anfrage «Gesichtsbilderdatenbanken der österreichischen Sicherheitsbehörden» 750/AB 1 zu 708/J (XXVII. GP); Amnesty will Ende der Gesichtserkennung in Österreich, in: Salzburger Nachrichten vom 04.05.2021, abrufbar unter <https://www.sn.at/panorama/oesterreich/amnesty-will-ende-der-gesichtserkennung-in-oesterreich-103305802>; zu Grossbritannien: Metropolitan Police, abrufbar unter <https://www.met.police.uk/advice/advice-and-information/fr/facial-recognition-technology/>; zu den USA: NYPD Questions and Answers Facial Recognition, abrufbar unter <https://www1.nyc.gov/site/nypd/about/about-nypd/equipment-tech/facial-recognition.page>.

³¹ SULZBACHER MARKUS, Polizei nutzt neue Gesichtserkennung, um Demonstranten zu identifizieren, in: Der Standard vom 15.09.2020, abrufbar unter <https://www.derstandard.at/story/2000119996329/polizei-nutzt-neue-gesichtserkennung-um-demonstranten-zu-identifizieren>.

4.3. Exkurs: Verzicht auf maschinelle Gesichtserkennung

Die erwähnten Beispiele bedeuten nicht, dass weltweit eine lineare Zunahme des Einsatzes von maschineller Gesichtserkennung zu verzeichnen wäre. Vielmehr gibt es verschiedene Länder und Städte, in denen Gesichtserkennung inzwischen verboten oder Pilotversuche sistiert wurden. So etwa in Deutschland, wo Rechtsgrundlagen für die Durchführung von Gesichtserkennung an über hundert deutschen Bahnhöfen und Flughäfen im Entwurfsstadium verworfen wurden.³² Auch am Brüsseler Flughafen Zaventem wurde über mehrere Jahre im Rahmen eines Pilotprojekts eine Gesichtserkennungstechnologie eingesetzt, was aufgrund eines nun geltenden landesweiten Verbotes allerdings wieder eingestellt wurde.³³ In Grossbritannien wurde Gesichtserkennung an Schulen eingesetzt, um angesichts der Corona-Pandemie eine kontaktlose Bezahlung des Mittagessens zu ermöglichen. Nach breiter Kritik stellten die Schulen dieses System (vorerst) wieder ein.³⁴ Auch in den USA haben einige Städte den Einsatz von Gesichtserkennung, zumindest durch staatliche Behörden, verboten.³⁵

³² Seehofer verzichtet auf Software zur Gesichtserkennung, in: Der Spiegel vom 24.01.2020, abrufbar unter <https://www.spiegel.de/politik/deutschland/bundespolizeigesetz-seehofer-verzichtet-auf-software-zur-gesichtserkennung-a-c207b3c8-eb1a-48e9-80ce-2642b420bd55>. Mittlerweile wurde der gesamte Gesetzesentwurf abgelehnt.

³³ Im Rahmen der Anpassung des Gesetzes über die Installation und Nutzung von Überwachungskameras (Loi du 21 mars 2007 réglant l'installation et l'utilisation de caméras de surveillance) wurde im März 2018 ein Art. 8/1 eingeführt, der jegliche Überwachungskameras gekoppelt mit Registern oder Dateien mit personenbezogenen Daten verbietet, sofern sie nicht ausschliesslich zur automatischen Erkennung von Kfz-Kennzeichen eingesetzt werden und der Datenschutz gewährleistet ist. Siehe dazu Art. 78 Belgisch Staatsblad vom 16.04.2018, S. 33708.

³⁴ MARON HANS JOERG, Britische Schulen: Schon wieder Schluss mit Gesichtserkennung, in: Inside IT vom 26.10.2021, abrufbar unter <https://www.inside-it.ch/de/post/britische-schulen-schon-wieder-schluss-mit-gesichtserkennung-20211026>.

³⁵ Darunter San Francisco, das im Kontext der Einführung einer Verordnung zur Anschaffung von Überwachungstechnik (Ordinance 190110, Administrative Code – Acquisition of Surveillance Technology) den Einsatz von Gesichtserkennungstechnologien (mit begrenzten Ausnahmen) verbietet und eine öffentliche Bekanntgabe aller verwendeten Überwachungstechnologien vorsieht. Die Verordnung definiert Gesichtserkennungstechnologie als ein automatisiertes oder halbautomatisiertes Verfahren, das dabei hilft, die Identifizierung oder Verifizierung einer Person anhand ihres Gesichts vorzunehmen – somit dürfte die Echtzeit-Gesichtserkennung wie auch die nachträgliche Gesichtserkennung darunter gefasst sein. Verboten auch in Oakland, Berkeley, Portland, Boston und Somerville, vgl. METZ

III. Rechtliche Würdigung

1. Datenschutzrechtliche Vorgaben

Das Datenschutzrecht statuiert einige Definitionen, Grundsätze und Pflichten, die für eine Datenbearbeitung sowohl durch Private als auch Behörden gelten.³⁶ Ausgangspunkt einer Datenbearbeitung bildet deren Rechtmässigkeit (Art. 6 Abs. 1 revDSG). Weiter muss eine Datenbearbeitung verhältnismässig sein (Art. 6 Abs. 2 revDSG). Das bedeutet, dass Daten nur bearbeitet werden dürfen, sofern ihre Bearbeitung geeignet ist, den verfolgten Zweck zu erfüllen und wenn kein milderes Mittel zur Zweckerreichung zur Verfügung steht.³⁷ Schliesslich muss die Datenbearbeitung zumutbar sein, dementsprechend muss ein vernünftiges Verhältnis zwischen der Datenbearbeitung und dem Eingriff in die Privatsphäre der betroffenen überwachten Personen bzw. dem Grundrechtseingriff bestehen.³⁸ Gerade im Bereich des Einsatzes von Gesichtserkennung im öffentlichen Raum zum Schutz der öffentlichen Ordnung und Sicherheit vor Störungen – z. B. die Verhinderung von Vandalismus oder Littering – dürfte sich regelmässig ein milderes Mittel finden lassen, wie etwa baulich-technische oder polizeilich-organisatorische Massnahmen.³⁹ Aus dem Grundsatz der Verhältnismässigkeit werden zudem die Prinzipien der Datenvermeidung und der Datensparsamkeit abgeleitet.⁴⁰ Da im Grundsatz davon ausgegangen werden kann, dass maschinelle Gesichtserkennungssysteme sehr viele (Trainings-)Daten benötigen, um treffsichere Outputs generieren

RACHEL, Portland passes broadest facial recognition ban in the US, in: CNN vom 10.09.2020, abrufbar unter <https://edition.cnn.com/2020/09/09/tech/portland-facial-recognition-ban/index.html>.

³⁶ Zu den datenschutzrechtlichen Grundsätzen im Kontext der Videoüberwachung ausführlich MÜLLER LUCIEN, Videoüberwachung in öffentlich zugänglichen Räumen – insbesondere zur Verhütung und Ahndung von Straftaten, Zürich/St. Gallen 2011, S. 60 ff.

³⁷ BGE 138 II 346 E. 9.2; das Verhältnismässigkeitsprinzip im Datenschutzrecht gilt nicht nur für staatliches Handeln (Art. 5 Abs. 2 BV), sondern auch unter Privaten, was zu hohen Anforderungen an eine Datenbearbeitung führt, vgl. SHK DSG-BAERISWYL, Art. 4 N 25, in: BAERISWYL BRUNO/PÄRLI KURT (Hrsg.), Datenschutzgesetz, Stämpfli Kommentar, Bern 2015.

³⁸ Zum datenschutzrechtlichen Verhältnismässigkeitsprinzip EPINEY ASTRID, § 9 Allgemeine Grundsätze, in: Belsler Eva Maria/Epiney Astrid/Waldmann Bernard (Hrsg.), Datenschutzrecht, Bern 2011, S. 528 ff.

³⁹ Vgl. z. B. den Bericht der Kantons- und Stadtentwicklung Basel-Stadt zur Schaffung von Sicherheit öffentlicher Räume vom September 2018, S. 56, abrufbar unter <https://www.bs.ch/medienmitteilungen/2018-die-sicherheit-im-oeffentlichen-raum-wird-ganzheitlich-gestaerkt>.

⁴⁰ SHK DSG-BAERISWYL (Fn. 37), Art. 4 N 23.

zu können,⁴¹ steht die Technologie in einem Spannungsverhältnis zu den soeben genannten Prinzipien. Zudem stellt sich dabei grundsätzlich die Frage, ob und inwieweit Trainingsdaten überhaupt datenschutzkonform erhoben und genutzt werden können.

Mit der Verhältnismässigkeit hängen ferner die Grundsätze der Zweckbindung und der Erkennbarkeit zusammen (Art. 6 Abs. 3 revDSG). Bei einer Datenbearbeitung muss stets der Bearbeitungszweck festgelegt und eingehalten werden. Insbesondere der Einsatz maschineller Gesichtserkennung zur allgemeinen Gefahrenabwehr kann dazu führen, dass die Festlegung des Bearbeitungszwecks untergraben wird, da dadurch vielseitige und breite Analysemöglichkeiten geschaffen werden, die vom Bearbeitungszweck abweichen.⁴² Der Zweck der Bearbeitung weist auch eine zeitliche Komponente auf. Mithin müssen Daten vernichtet werden, sobald sie zum Bearbeitungszweck nicht mehr erforderlich sind (Art. 6 Abs. 4 revDSG). Somit ist die Speicherdauer festzulegen.⁴³

Die skizzierten Anforderungen gelten bereits für das Erheben der Daten im öffentlichen Raum (z. B. CCTV-Kameraaufnahmen von Passantinnen und Passanten) und nicht erst für die Überprüfung ebendieser mittels Gesichtserkennung. Die Datenbeschaffung sowie insbesondere der Bearbeitungszweck müssen ausserdem für die betroffene Person erkennbar sein. Das revDSG führt mit Art. 19 ausgeweitete Informationspflichten bei der Beschaffung von Personendaten ein, die für Private und Behörden gelten. Unter bestimmten Voraussetzungen, etwa wenn die Person bereits über die entsprechenden Informationen verfügt oder die Bearbeitung gesetzlich vorgesehen ist, entfällt die Informationspflicht (Art. 20 Abs. 1 lit. a und b revDSG). Diese Bestimmung ist allerdings explizit als Ausnahme konstituiert.⁴⁴ Die Erkennbarkeit wird subsidiär, sofern Informationspflichten bestehen.⁴⁵ Die Information muss angemessen sein und diejenigen Informationen enthalten, die die betroffene Person benötigt, um ihre Rechte gemäss Datenschutzrecht geltend machen zu können (Art. 19 Abs. 2

⁴¹ KARIOTH (Fn. 13), S. 11.

⁴² SIMMLER/CANOVA (Fn. 5), S. 109, betiteln dies als eine «anlasslose und verdachtsunabhängige Massnahme mit grosser Streubreite».

⁴³ MÜLLER (Fn. 36), S. 230 f.

⁴⁴ Für Behörden wird die Ausnahme zur Regel, da eine Datenbearbeitung gestützt auf das Legalitätsprinzip stets eine gesetzliche Grundlage voraussetzt. Diese Lösung ist allerdings insbesondere mit Blick auf die Adressatengerechtigkeit der Information kritisch zu sehen, vgl. THOUVENIN FLORENT/BRAUN BINDER NADJA, Transparenz durch Datenschutzerklärungen von Behörden, in: ZSR 141/2022 I, S. 5 ff.

⁴⁵ SHK DSG-BAERISWYL (Fn. 37), Art. 4 N 47.

lit. a–c revDSG). Die Botschaft zum revDSG hält fest, dass bei der Wahl einer allgemeinen Informationsform der Zugang zur Information einfach, die Information vollständig und der Hinweis augenfällig sein muss.⁴⁶ Ein allgemeines Hinweisschild, dass Gesichtserkennung eingesetzt wird, dürfte vor diesem Hintergrund als unvollständig und somit als ungenügend zu bewerten sein.

Zu erwähnen ist ferner der Grundsatz der Datenrichtigkeit (Art. 6 Abs. 5 revDSG), der die Vollständigkeit und die Aktualität der Daten umfasst.⁴⁷ Die Pflicht zur Richtigkeit ist allerdings nicht absolut zu verstehen, sie muss vielmehr in einem angemessenen Verhältnis zum Verarbeitungszweck stehen (relative Datenrichtigkeit).⁴⁸ Gefordert wird, dass mit Rücksicht auf den Zweck alle geeigneten Massnahmen ergriffen werden, die eine Berichtigung, Löschung oder Vernichtung unrichtiger oder unvollständiger Daten ermöglichen.⁴⁹ Im Kontext der Gesichtserkennung ist eine sehr hohe Wahrscheinlichkeit der richtigen Identifizierung bzw. Verifizierung der Betroffenen für die Zulässigkeit eines biometrischen Verfahrens zu fordern.⁵⁰ Um dies zu erreichen, müssen die Trainingsdatensätze insbesondere eine hinreichende Diversität bezüglich Alter, Geschlecht und Hautfarbe aufweisen.⁵¹

2. Gesichtserkennung im öffentlichen Raum durch Private

Die Bearbeitung von Personendaten durch Private fällt grundsätzlich unter das Datenschutzgesetz (Art. 2 Abs. 1 lit. a revDSG), sofern die Bearbeitung nicht ausschliesslich dem persönlichen Gebrauch dient (Art. 2 Abs. 2 lit. a revDSG). Ein rein persönlicher Gebrauch ist beim Einsatz von maschineller Gesichtserkennung im öffentlichen Raum regelmässig auszuschliessen, da das gesammelte Bildmaterial in der Regel Überwachungszwecken und damit potenziell auch als Beweismittel im Rahmen eines späteren Straf- oder Zivilverfahrens dient.⁵²

⁴⁶ BBl 2017 6941 ff. (Fn. 10), S. 7050 f.

⁴⁷ SHK DSG-BAERISWYL/BLONSKI (Fn. 37), Art. 5 N 1.

⁴⁸ BSK DSG-MAURER-LAMBROU/SCHÖNBÄCHLER, Art. 5 N 6, in: MAURER-LAMBROU URS/BLECHTA GABOR P. (Hrsg.), Datenschutzgesetz, Öffentlichkeitsgesetz, Basler Kommentar, 3. A., Basel 2014.

⁴⁹ MÉTILLE SYLVAIN, Le traitement de données personnelles sous l'angle de la nouvelle loi fédérale sur la protection des données du 25 septembre 2020, in: SJ II/2021, S. 1 ff., S. 11.

⁵⁰ privatim (Fn. 7), S. 12.

⁵¹ KARJOTH (Fn. 13), S. 11. Dies ist insb. auch wichtig im Hinblick auf das Diskriminierungsverbot, dazu [III.3.3](#).

⁵² MÜLLER LUCIEN, Private Videoüberwachung in öffentlich zugänglichen Räumen – Datenschutzrechtliche Aspekte, in: Sicherheit & Recht, 2/2012, S. 63 ff., S. 65.

Die Datenbearbeitung durch private Personen darf keine widerrechtliche Persönlichkeitsverletzung darstellen (Art. 6 Abs. 1 i. V. m. Art. 30 Abs. 1 revDSG). Folglich ist eine Datenbearbeitung nur zulässig, sofern die Grundsätze des Datenschutzrechts (Art. 30 Abs. 2 lit. a revDSG) eingehalten werden,⁵³ maschinelle Gesichtserkennung nicht gegen den ausdrücklichen Willen der Betroffenen eingesetzt wird (Art. 30 Abs. 2 lit. b revDSG) oder ein Rechtfertigungsgrund (Art. 31 Abs. 1 revDSG) vorliegt. Für die Bearbeitung biometrischer Daten i. S. v. Art. 5 lit. c Ziff. 4 revDSG ist grundsätzlich eine ausdrückliche Einwilligung der betroffenen Person notwendig (Art. 6 Abs. 7 lit. a revDSG). Die Einwilligung muss freiwillig erfolgen und es muss eine vorgängige Information über die Datenbearbeitung stattfinden (Art. 6 Abs. 6 und 7 i. V. m. Art. 19 Abs. 1 revDSG). Obwohl die Ausdrücklichkeit der Einwilligung keine Schriftlichkeit voraussetzt, muss der Wille der betroffenen Person eindeutig aus den Umständen des Einzelfalles hervorgehen.⁵⁴ Sowohl die Informationspflicht als auch die ausdrückliche und freiwillige Einwilligung sind bei der maschinellen Gesichtserkennung im öffentlichen Raum schwer umzusetzen.⁵⁵ Würde bspw. das Betreten öffentlicher Räume jeweils von einer Einwilligung in die Gesichtserkennung abhängig gemacht, wäre die Freiwilligkeit nicht gegeben.⁵⁶

Als Rechtfertigungsgründe für eine widerrechtliche Persönlichkeitsverletzung – z. B. die Bearbeitung biometrischer Daten ohne ausdrückliche Zustimmung der betroffenen Person – kommen ein überwiegendes privates oder öffentliches Interesse oder das Vorliegen einer gesetzlichen Grundlage in Frage (Art. 31 Abs. 1 revDSG). Analog zur Einschätzung des Eidgenössischen Datenschutz- und Öffentlichkeitsbeauftragten (EDÖB) hinsichtlich der Videoüberwachung dürfte allerdings auch mit Blick auf den Einsatz maschineller Gesichtserkennung davon auszugehen sein, dass die Überwachung des öffentlichen Raums durch Private zur Wahrung privater Interessen regelmässig eine Persönlichkeitsverletzung darstellt, da kein überwiegendes Interesse vorhanden ist.⁵⁷ Es können zwar keine verallgemeinernden Aussagen bezüglich der Interessenabwägung gemacht werden, da diese immer konkret und einzelfallspezifisch erfolgt.⁵⁸ Jedoch sind Konstellationen, in denen das private Interesse an der Überwachung des öffentlichen

⁵³ Vgl. dazu [III.1](#), wobei insb. der Grundsatz der Verhältnismässigkeit zu betonen ist.

⁵⁴ SHK DSG-BAERISWYL (Fn. 37), Art. 4 N 69 ff.

⁵⁵ Zur Informationspflicht im Rahmen der Gesichtserkennung siehe [III.1](#).

⁵⁶ MÜLLER (Fn. 52), S. 73 f.; SHK DSG-BAERISWYL (Fn. 37), Art. 13 N 65 ff.

⁵⁷ EDÖB, Videoüberwachung des öffentlichen Raums durch Privatpersonen, abrufbar unter <https://www.edoeb.admin.ch/de/videoeberwachung-des-oeffentlichen-raums-durch-privatpersonen>.

⁵⁸ SHK DSG-WERMELINGER (Fn. 37), Art. 13 N 12.

Raumes das Interesse der betroffenen Personen am Schutz ihrer biometrischen Daten i. S. v. Art. 5 lit. c Ziff. 4 revDSG und damit am Schutz besonders schützenswerter Daten überwiegt, praktisch nicht vorstellbar. Ferner kann auch die Wahrung von Sicherheit und Ordnung (als öffentliches Interesse) im öffentlichen Raum nicht als Rechtfertigungsgrund für den Einsatz von maschineller Gesichtserkennung durch Private herangezogen werden. Denn die Wahrung von Sicherheit und Ordnung im öffentlichen Raum ist Sache der Polizei und weiterer Behörden.

Bezüglich der allgemeinen Videoüberwachung werden zwei Ausnahmen statuiert:⁵⁹ Erstens bei Vorliegen einer an sich rechtmässigen Überwachung von privatem Grund, bei der eine geringfügige Miterfassung des öffentlichen Raumes erfolgt und die Überwachung des privaten Grundes nicht anders durchführbar ist. Zweitens bei einer privaten Überwachung des öffentlichen Raumes aus Sicherheitsgründen, die mit dem zuständigen Gemeinwesen vereinbart wurde. Soweit es sich bei der maschinellen Gesichtserkennung um eine Bearbeitung von besonders schützenswerten Personendaten handelt, ist fraglich, ob die mit Blick auf die allgemeine Videoüberwachung statuierten Ausnahmen analog anwendbar wären. Davon ist jedenfalls nicht ohne Weiteres auszugehen.

3. Gesichtserkennung im öffentlichen Raum durch staatliche Behörden

Setzen Behörden⁶⁰ Gesichtserkennungssoftware ein, müssen sie die Grundrechte berücksichtigen (Art. 35 Abs. 1 BV). Im Kontext maschineller Gesichtserkennung sind dabei insbesondere das Recht auf informationelle Selbstbestimmung (Art. 13 Abs. 2 BV), das Diskriminierungsverbot (Art. 8 Abs. 2 BV) und die Meinungs- und Versammlungsfreiheit (Art. 16 und 22 BV) zu berücksichtigen.

3.1. Informationelle Selbstbestimmung

Art. 13 Abs. 1 BV schützt unter anderem in der Öffentlichkeit vorgenommene Handlungen oder Äusserungen, die der Pflege persönlicher Kontakte dienen

⁵⁹ Zum Ganzen EDÖB (Fn. 57). Siehe auch FISCHLI-GIESSER LIZ, Private Videoüberwachung im kommunalen öffentlichen Raum, in: KPG Bulletin 03/2016, S. 82 ff.

⁶⁰ Das Datenschutzrecht des Bundes entfaltet keine Wirkung für kantonale oder kommunale Behörden Art. 2 Abs 1 revDSG *e contrario*. Deshalb wird hier (auch) auf verfassungsrechtliche Vorgaben Bezug genommen.

und mithin Ausdruck der Persönlichkeit sind, ganz allgemein vor Videoüberwachung im öffentlichen und privaten Raum sowie dauernder Beobachtung und Dokumentation in der Öffentlichkeit.⁶¹ Der Schutzbereich des sog. Rechts auf informationelle Selbstbestimmung gemäss Art. 13 Abs. 2 BV umfasst u. a. den Schutz vor bzw. die Entscheidungsfreiheit bzgl. jeglichen staatlichen Umgangs mit den eigenen Personendaten, ein Recht am eigenen Bild, sowie die Autonomie betreffend die eigene gesellschaftliche und soziale Rolle.⁶²

Das Recht auf informationelle Selbstbestimmung wird durch das Datenschutzrecht konkretisiert. Ausgangspunkt bildet wiederum die Rechtmässigkeit der Datenbearbeitung nach Art. 6 Abs. 1 revDSG. Für Bundesbehörden setzt die Rechtmässigkeit eine formell gesetzliche Grundlage nach Art. 34 Abs. 2 lit. a revDSG voraus, da es um besonders schützenswerte Personendaten geht. Dadurch werden zugleich die verfassungsrechtlichen Anforderungen gem. Art. 36 Abs. 1 BV bestätigt.⁶³

Neben einer genügenden Normstufe ist auch eine ausreichende Normdichte, mithin eine hinreichend bestimmte gesetzliche Regelung, notwendig.⁶⁴ Erforderlich sind die Definition des konkreten Bearbeitungszwecks, die beteiligten Organe, die Kategorien der bearbeiteten Daten und die grobe Umschreibung des Bearbeitungsumfangs.⁶⁵ Der pauschale Verweis auf die Gewährleistung

⁶¹ SGK BV-BREITENMOSER, Art. 13 N 15 ff. Ausführungen zu internationalen Normen, insb. Art. 8 EMRK (Europäische Menschenrechtskonvention), werden in diesem Aufsatz ausgeklammert.

⁶² MÜLLER (Fn. 36), S. 122 f.

⁶³ Auch gem. Art. 164 Abs. 1 BV ergibt sich die Notwendigkeit einer formell gesetzlichen Grundlage, da die Bearbeitung besonders schützenswerter Personendaten Gefahren für die Persönlichkeit birgt, siehe BSK DSG-UDVARY/BALLENEGGER (Fn. 48), Art. 17 N 25. Zudem ist die Notwendigkeit der demokratischen Legitimation von maschineller Gesichtserkennung im öffentlichen Raum zu betonen, siehe MÜLLER (Fn. 36), S. 203 ff., siehe auch BGE 146 I 11 E. 2.3, in dem das Bundesgericht die automatische Fahrzeugfahndung und Verkehrsüberwachung vor dem Hintergrund der seriellen und simultanen Verarbeitung grosser und komplexer Datensätze innert Sekundenschnelle sowie der Möglichkeit der Erstellung von Persönlichkeits- und Bewegungsprofilen als schweren Grundrechtseingriff qualifiziert hat. Analoges gilt für die Gesichtserkennung im öffentlichen Raum. Ein schwerer Grundrechtseingriff durch Gesichtserkennung wird auch in GJON DAVID, Gesichtserkennung: «Ende der Privatsphäre», in: plädoyer 6/2021, S. 26 f., S. 27, angenommen.

⁶⁴ Dazu ausführlich MÜLLER (Fn. 36), S. 203 ff. Vgl. auch ZYSSET ESTHER, Brauche ich für meine KI-Anwendung eine gesetzliche Grundlage und wenn ja, welcher Art?, abrufbar unter <https://publicsector.ch/brauche-ich-fur-meine-ki-anwendung-eine-gesetzliche-grundlage-und-wenn-ja-welcher-art/>.

⁶⁵ BSK DSG-UDVARY/BALLENEGGER (Fn. 48), Art. 17 N 19.

der öffentlichen Sicherheit, den Schutz von Personen oder die polizeiliche Aufgabenerfüllung ist nicht ausreichend.⁶⁶ Nicht zu vergessen ist zudem die Notwendigkeit einer ausreichenden Rechtsgrundlage für jene Daten, die in den Überwachungsdatenbanken enthalten sind, also die Bilder derjenigen Personen, nach denen bspw. gefahndet wird.⁶⁷

Bis dato ist keine gesetzliche Grundlage für die Gesichtserkennung in der Strafprozessordnung⁶⁸ oder im kantonalen Polizeirecht auszumachen.⁶⁹ Eine gesetzliche Grundlage in einem anderen Rechtsbereich ist z. B. die frühere Verordnung über den Einsatz eines biometrischen Gesichtserkennungssystems am Flughafen Zürich⁷⁰, die durch Art. 102 und 103 AIG⁷¹ abgelöst wurde. Als Konkretisierung ist ausserdem Art. 45 VEV⁷² zur automatisierten Grenzkontrolle zu nennen und Art. 54–62 VEV als Grundlage und Rahmenbedingungen für die Zulässigkeit des Einsatzes eines Gesichtserkennungssystems durch die Grenzbehörden.

3.2. Versammlungs- und Meinungsfreiheit

Der Einsatz von maschineller Gesichtserkennung im öffentlichen Raum kann weitere Grundrechte berühren.⁷³ Zu bedenken ist insbesondere die potenzielle Abschreckungswirkung (sog. *chilling effect*) auf die Grundrechtsausübung.⁷⁴ Würde der öffentliche Raum z. B. während einer Demonstration mittels Gesichtserkennung überwacht, könnte dies dazu führen, dass Personen von der Teilnahme an einer Kundgebung zurückschrecken, weil sie negative Konse-

⁶⁶ FLEISCHMANN FLORIAN SAMUEL, Polizeirechtliche Massnahmen zur Bekämpfung der Gewalt anlässlich von Sportveranstaltungen, Zürich 2019, S. 282 f.

⁶⁷ Zu dieser Unterscheidung und den daraus resultierenden Anforderungen an die gesetzliche Grundlage siehe SIMMLER/CANOVA (Fn. 5), S. 15 ff.; ferner FRA – European Union Agency for Fundamental Rights, Facial recognition technology: fundamental rights considerations in the context of law enforcement, Wien 2019, S. 24.

⁶⁸ Schweizerische Strafprozessordnung vom 05.10.2007 (StPO), SR 312.0.

⁶⁹ Dazu ausführlich SIMMLER/CANOVA (Fn. 5), S. 18 ff. Immerhin erwähnt § 32c PolG-ZH die offene oder verdeckte Videoüberwachung dergestalt, dass Personen damit identifiziert werden können, jedoch stellt diese Norm (gem. SIMMLER/CANOVA) keine ausreichende gesetzliche Grundlage dar; ebenso auch Art. 45b PolG-BL nicht.

⁷⁰ SR 551.113.

⁷¹ Bundesgesetz über die Ausländerinnen und Ausländer und über die Integration vom 16.12.2005, SR 142.20.

⁷² Verordnung über die Einreise und die Visumerteilung vom 15.09.2018, SR 142.204.

⁷³ Siehe eingehend MÜLLER (Fn. 36), S. 97 ff.

⁷⁴ Siehe z. B. BGE 143 I 147 E. 3.3.

quenzen – allein aufgrund einer Teilnahme an einer Demonstration – befürchten. Damit würde die Ausübung der Meinungs- und Versammlungsfreiheit (Art. 16 Abs. 1 bzw. Art. 22 Abs. 1 BV) tangiert. Der Einsatz von maschineller Gesichtserkennung im öffentlichen Raum könnte mithin zu einer mittelbaren Grundrechtsbeeinträchtigung führen.⁷⁵

3.3. Diskriminierungsverbot

Beim Einsatz von Gesichtserkennung gibt es sog. *false positives*, also Resultate, die fälschlicherweise eine Übereinstimmung anzeigen, wo keine vorliegt, und sog. *false negatives*, indem keine Übereinstimmung angezeigt wird, obwohl eine vorliegt. Bei einer hohen überprüften Personenzahl führt auch eine kleine *false positive*-Rate zur falschen Meldung einer Übereinstimmung bei zahlreichen Personen.⁷⁶ Eine Konsequenz von *false positives* könnte die vermehrte Kontrolle bzw. Überprüfung dieser Personen sein oder gar eine ungerechtfertigte Festnahme; eine Konsequenz von *false negatives* das Entstehen von Sicherheitsproblemen.⁷⁷

Die Auswirkungen falscher Identifizierungen zeigen sich umso eindringlicher, wenn die Fehlerrate bei bestimmten Personengruppen höher ausfällt als bei anderen, insbesondere, wenn an ein verpöntes Merkmal im Sinne des Diskriminierungsverbots nach Art. 8 Abs. 2 BV angeknüpft wird. Dies kann eintreffen, wenn Personen, die in den Datenbanken, mit denen die Gesichtserkennungssoftware trainiert wird, untervertreten sind, was vor allem auf dunkelhäutige/nicht kaukasische Personen, Frauen und jüngere Menschen zutrifft.⁷⁸ Durch die geringere Vertretung in den Trainingsdaten erkennt die Software diese Personen schlechter.⁷⁹ Es stellt sich mithin die Frage, ob dies-

⁷⁵ FRA (Fn. 67), S. 29 f.; MÜLLER (Fn. 36), S. 152.

⁷⁶ Bei einer Personenzahl von 100'000 führt eine *false positive*-Rate von bloss 1% dennoch zu 1'000 *false positives*.

⁷⁷ Wobei zu betonen ist, dass in der Schweiz die Ergebnisse der Software noch durch Menschen überprüft werden – hier ist insb. wichtig, die zuständigen Behörden für diese Thematik zu sensibilisieren, vgl. BRAUN BINDER et al., Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht, in: Jusletter vom 28.06.2021; ferner GROTHÉ PATRICK/NGAN MEI/HANAOKA KAYEE, Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects, in: NISTIR 8280 vom 12/2019.

⁷⁸ JACQUET MAËLIG/GROSSRIEDER LIONEL, Enjeux et perspectives de la reconnaissance faciale en sciences criminelles, in: Criminologie 54/2021 Nr. 1, S. 135 ff., Rn. 47; FRA (Fn. 67), S. 27.

⁷⁹ KLARE BRENDAN F. et al., Face recognition performance: role of demographic information, in: IEEE Transactions on Information Forensics and Security 7/2012, S. 1789 ff., S. 1789 und 1800. Ausführlich zur Diskriminierung vgl. auch KOLLECK ALMA/ORWAT CARSTEN, Mögliche

falls der Einsatz maschineller Gesichtserkennung überhaupt als geeignet (i. S. d. Verhältnismässigkeitsgrundsatzes) eingestuft werden kann.

IV. Moratorium statt Verbot

Der EU-KI-Verordnungsentwurf sieht ein Verbot jeglicher KI-Software zur Erfassung biometrischer Daten vor, allerdings mit Erlaubnisvorbehalt.⁸⁰ Die EU-Kommission verfolgt damit einen risikobasierten Ansatz, der für unterschiedliche KI-Systeme unterschiedliche Abstufungen enthält.⁸¹ Gesichtserkennungssoftware wird gem. Art. 6 Abs. 2 EU-KI-Verordnungsentwurf i. V. m. Nr. 1 des Anhangs III EU-KI-Verordnungsvorschlag als Hoch-Risiko-KI kategorisiert, da mit ihr besondere Gefahren für die Sicherheit, die Grundrechte und die Gesellschaft als Ganzes einhergehen. Im öffentlichen Raum soll das Erfassen biometrischer Daten in Echtzeit deshalb grundsätzlich verboten werden, wobei gewisse Ausnahmen vorgesehen sind. Solche sollen bspw. im Polizeirecht greifen, wenn es um das Auffinden vermisster Kinder oder von Schwerverbrecherinnen und Schwerverbrechern geht (Art. 5 Abs. 1 lit. d Ziff. i-iii EU-KI-Verordnungsvorschlag).⁸² Dies allerdings nur, wenn strikte Sicherungsaufgaben eingehalten werden, eine vorgängige Genehmigung erteilt wurde und eine unabhängige Kontrollinstanz existiert (Art. 5 Abs. 2 und 3 EU-KI-Verordnungsvorschlag). Fraglich ist, ob mit einem solchen Verbot mit

Diskriminierung durch algorithmische Entscheidungssysteme und maschinelles Lernen – ein Überblick, 10/2020, S. 53 ff.; GROTHNER/NGAN/HANAOKA (Fn. 77).

⁸⁰ Siehe m. w. H. SPINDLER GERALD, Der Vorschlag der EU-Kommission für eine Verordnung zur Regulierung der Künstlichen Intelligenz, in: IT und Software – Aufsätze 6/2021, S. 361 ff., S. 365; KUGELMANN DIETER, EU-Vorschlag zu Künstlicher Intelligenz muss aus Datenschutz-Perspektive durchleuchtet werden, in: Datenschutz und Datensicherheit 7/2021, S. 433; MARIO MARTINI, Gesichtserkennung im Spannungsfeld zwischen Freiheit und Sicherheit, in: Kristin Pfeffer (Hrsg.), Schriftenreihe der Forschungsstelle Europäische und Deutsches Sicherheitsrecht (FEDS), Band 4, Göttingen 2022, S. 65 ff.; Einzelne EU-Parlaments-Abgeordnete fordern dagegen ein Moratorium, siehe FANTA ALEXANDER, EU-Abgeordnete fordern Auszeit für Gesichtserkennung, abrufbar unter <https://netzpolitik.org/2021/biometrische-ueberwachung-eu-abgeordnete-fordern-auszeit-fuer-gesichtserkennung/>.

⁸¹ Ausführungen zur EU-Regulierung z. B. von SCHALLBRUCH MARTIN, EU-Regulierung der Künstlichen Intelligenz, in: Datenschutz und Datensicherheit 7/2021, S. 438 ff., S. 441.

⁸² Siehe dazu bspw. HERWARTZ CHRISTOPH/KOCH MORITZ, Soll Gesichtserkennung zur Verbrecherjagd erlaubt sein? Streit um die neuen KI-Regeln der EU, in: Handelsblatt vom 22.04.2021, abrufbar unter <https://www.handelsblatt.com/politik/deutschland/kuenstliche-intelligenz-soll-gesichtserkennung-zur-verbrecherjagd-erlaubt-sein-streit-um-die-neuen-ki-regeln-der-eu/27120280.html?ticket=ST-3211797-sg2kkyPNzFoyxbflkIw-ap4>.

Erlaubnisvorbehalt Massenüberwachung tatsächlich verhindert werden kann.⁸³

Die Herangehensweise der EU-Kommission stellt für die Schweiz ohnehin kein taugliches Vorbild dar. Zum einen steht in der Schweiz die Schaffung eines allgemeinen KI-Gesetzes – zurecht – nicht auf der politischen Agenda.⁸⁴ Zum anderen liesse sich eine solche zentrale Regelung nur schwer mit der Kompetenzaufteilung zwischen Bund und Kantonen vereinbaren. Dem Bund kommt insbesondere im Bereich der Polizeiorganisation keine umfassende Kompetenz zu.

Gleichwohl ist festzuhalten, dass in der Schweiz bislang eine gesetzliche Grundlage für den Einsatz von maschineller Gesichtserkennung im öffentlichen Raum fehlt,⁸⁵ was angesichts der Grundrechtsrelevanz des Themas durchaus kritisch zu beurteilen ist. Vor diesem Hintergrund fordern denn auch verschiedene Akteurinnen und Akteure ein grundsätzliches Verbot bzw. ein Moratorium, vor allem für den Einsatz von Gesichtserkennungstechnologie zwecks Massenüberwachung.⁸⁶

Aus rechtlicher Sicht ist ein explizites, durch den Gesetzgeber zu verankern- des Verbot, zumindest für den staatlichen Einsatz von maschineller Gesichtserkennung, nicht notwendig, weil ein solcher Einsatz ohnehin eine entsprechende formell gesetzliche Grundlage und damit ein Tätigwerden des Gesetzgebers voraussetzt. Mit Blick auf Private erscheint ein generelles Verbot ebenfalls nicht unbedingt zielführend, da damit die grundrechtskonforme Entwicklung von Gesichtserkennungssystemen verhindert werden könnte. Ein Moratorium erscheint dagegen sinnvoll, da damit eine gesellschaftliche und politische Debatte angestossen und darauf hingewirkt werden kann, dass der Einsatz von Gesichtserkennungssoftware mit einem genügenden Grundrechtsschutz einhergeht.

⁸³ M. w. H. SPINDLER (Fn. 80), S. 365; KUGELMANN (Fn. 80), S. 433; FANTA (Fn. 80); REICHE MATTHIAS, EU will Künstliche Intelligenz zähmen, Tagesschau vom 21.04.2021, abrufbar unter <https://www.tagesschau.de/wirtschaft/technologie/eu-gesetzentwurf-kuenstliche-intelligenz-ki-101.html>.

⁸⁴ Siehe etwa THOUVENIN FLORENT et al., Positionspapier vom 10. November 2021 «Ein Rechtsrahmen für Künstliche Intelligenz», abrufbar unter <https://www.dsi.uzh.ch/dam/jcr:3a0cb402-c3b3-4360-9332-f800895fdc58/dsi-strategy-lab-21-de.pdf>.

⁸⁵ Bezüglich der Polizeiarbeit haben dies SIMMLER/CANOVA (Fn. 5), S. 17 ff., bereits ausführlich erörtert. Gleichzeitig halten sie jedoch fest, dass die Polizei trotz Fehlen einer gesetzlichen Grundlage gewisse Technologien bereits anwendet.

⁸⁶ Siehe THOUVENIN FLORENT et al. (Fn. 84) sowie im vorliegenden Beitrag Fn. 3 und 4.

V. Fazit

Die Schaffung eines einzelfallweise zulässigen Einsatzes maschineller Gesichtserkennung, wie im EU-KI-Verordnungsentwurf vorgesehen, kann zum Entstehen juristischer Graubereiche führen. Für die Schweiz wird als sinnvoller erachtet, mithilfe eines Moratoriums für maschinelle Gesichtserkennung im öffentlichen Raum einen breiten gesetzgeberischen und gesellschaftspolitischen Diskurs anzustossen und damit nach Möglichkeiten grundrechtskonformer Lösungen zu suchen.

Algorithmisch überprüfte Steuererklärungen im ordentlichen gemischten Veranlagungsverfahren

Nadja Braun Binder / Liliane Obrecht

Erstabdruck in: zsis) 1/2023, Artikel 1

Als Bestandteil des Untersuchungsgrundsatzes prüfen die kantonalen Steuerbehörden von Amtes wegen die Angaben in den Steuererklärungen auf Korrektheit und Vollständigkeit. Im Rahmen der Automatisierung des Veranlagungsverfahrens bildet die Automatisierung dieser Prüfung einen wesentlichen Schritt. In Deutschland werden zu diesem Zweck Risikomanagementsysteme eingesetzt. Diese sondern risikobehaftete Steuererklärungen automatisiert aus, um sie der Prüfung durch eine natürliche Person zuzuführen.

Auch in der Schweiz sind solche Systeme grundsätzlich vorstellbar. Allerdings bestehen diesbezüglich grundrechtliche und datenschutzrechtliche Herausforderungen. So bedeutet etwa der Abgleich von Steuererklärungen mit anderen Steuererklärungen eine Weiterverwendung und somit aus datenschutzrechtlicher Sicht allenfalls unzulässige Zweckänderung der Daten. Eine Risikoanalyse, die der Bewertung der wahrheitsgetreuen Angaben in der Steuererklärung dient, kann ausserdem ein Profiling im Sinne des revidierten Datenschutzgesetzes darstellen und bedarf deshalb einer gesetzlichen Grundlage.

Inhalt

I.	Einleitung	162
II.	Schrittweise Automatisierung	164
III.	Datenerhebung	166
	1. Eingeschränkter Untersuchungsgrundsatz und Mitwirkungspflicht	166
	2. Datenerhebung als Datenbearbeitung	168
	3. Datenquellen	170
IV.	Überprüfung der Steuererklärung	173
	1. Automatisierung und Untersuchungsgrundsatz	173
	2. Einsatz von Risikomanagementsystemen	174

3.	Zweckänderung der Daten im Verfahren	179
4.	Profiling im Veranlagungsverfahren	180
V.	Fazit	185

I. Einleitung

Die Digitalisierung in den Steuerverwaltungen und insbesondere die Automatisierung des Steuerverfahrens schreiten voran.¹ Der Bundesrat hat im Mai 2020 vorgeschlagen, im Bereich der Verrechnungssteuer, der Mehrwertsteuer und der Stempelabgaben eine Verpflichtung zum elektronischen Verkehr mit der eidgenössischen Steuerverwaltung einzuführen.² Im Juni 2021 wurden die Kantone verpflichtet, elektronische Steuerverfahren anzubieten.³ Dies betrifft insbesondere die Möglichkeit des (vollständig) elektronischen Einreichens der Steuererklärung sowie die elektronische Eröffnung der Veranlagungsverfügung durch die Steuerbehörde. Mit Letzterem werden die kantonalen Veranlagungsverfahren im Bereich der direkten Bundessteuern weitgehend digitalisiert. Dies bildet die Basis für weitere Digitalisierungsschritte bis hin zu vollständig automatisierten Veranlagungsverfahren, in denen die Veranlagung durch den Einsatz von Algorithmen und nicht mehr durch Menschen vorgenommen wird.⁴

¹ Dieser Aufsatz entstand im Rahmen des von der Mercator Stiftung geförderten Projekts «Nachvollziehbare Algorithmen: Ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz», online aufgerufen am 17. Februar 2023 unter: <https://ius.unibas.ch/de/e-piaf/nachvollziehbare-algorithmen/>. Die Autorinnen danken Prof. Dr. Luzius Cavelti LL.M. RA, eidg. dipl. Steuerexperte, Prof. Dr. Beat Rudin, Datenschutzbeauftragter Basel-Stadt und Dr. Christian Ulbrich, Geschäftsführer e-PIAF, für ihre inhaltlichen Rückmeldungen zum Text sowie B.A. Anna-Lia Käslin-Tanduo für die sprachliche Durchsicht.

² Siehe die Medienmitteilung des Bundesrates vom 20.05.2020, online aufgerufen am 17. Februar 2023 unter: <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-79192.html>. Dies soll schrittweise vonstattengehen, siehe die Medienmitteilung des Bundesrates vom 03.11.2021, online aufgerufen am 17. Februar 2023 unter: <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-85713.html>.

³ Siehe den Schlussabstimmungstext zum Bundesgesetz über elektronische Verfahren im Steuerbereich vom 18. Juni 2021 (Art. 104a Bundesgesetz vom 14. Dezember 1990 über die direkte Bundessteuer), online aufgerufen am 17. Februar 2023 unter: <https://www.parlament.ch/centers/eparl/curia/2020/20200051/Schlussabstimmungstext.pdf>.

⁴ Siehe zur Terminologie vollautomatisierter Verfahren Braun Binder Nadja, Vollautomatisierte Verwaltungsverfahren im allgemeinen Verwaltungsverfahrenrecht? Der Gesetzes-

Bei einer solchen Vollautomation des Veranlagungsverfahrens muss – genauso wie im regulären Veranlagungsverfahren – sichergestellt sein, dass die durch die steuerpflichtige Person eingereichte Steuererklärung geprüft wird. Die vollständig automatisierte Veranlagung darf nicht dazu führen, dass fehlerhafte Sachverhaltsangaben die Grundlage für die Veranlagungsverfügung bilden. Auch diese Prüfung wird vermehrt mithilfe von Algorithmen durchgeführt. Internationale Entwicklungen zeigen, dass Steuerbehörden maschinelle Lernverfahren einsetzen, um Steuererklärungen zu überprüfen und Steuerbetrug aufzudecken.⁵ Dazu werden nicht nur Daten genutzt, die auf konventionellem Weg – beispielsweise durch die Mitwirkung der steuerpflichtigen Personen – erhoben werden. So erlaubt etwa ein im Jahr 2021 erlassenes Dekret den französischen Steuerbehörden, verschiedene Online-Plattformen zu durchsuchen, um Beweise für Steuerbetrugsfälle zu ermitteln.⁶ Die Vermutung liegt nahe, dass auch in der Schweiz mit zunehmender Automatisierung des Steuerverfahrens eine zunehmende Automatisierung der Überprüfung der Steuererklärungen stattfindet.⁷

Die digitalisierten bzw. automatisierten Steuerverfahren, insbesondere die dabei notwendige Überprüfung der zugrunde gelegten Daten, haben den Anforderungen der Bundesverfassung (BV) und den einschlägigen gesetzlichen Vorgaben zu genügen.

Der vorliegende Beitrag zeigt in einem ersten Schritt kursorisch die Entwicklung der Automatisierung des Steuerverfahrens auf (II.). Daraufhin wird die automatisierte Datenerhebung im Steuerverfahren genauer betrachtet (III.). Schliesslich

entwurf zur Modernisierung des Besteuerungsverfahrens als Vorbild für vollautomatisierte Verwaltungsverfahren nach dem VwVfG, NVwZ 2016, S. 960 ff., S. 960; zum Begriff des Algorithmus Zweig Katharina, Ein Algorithmus hat kein Taktgefühl: Wo künstliche Intelligenz sich irrt, warum uns das betrifft und was wir dagegen tun können, München 2019, Kapitel 3.

⁵ Siehe die Übersicht der Universität Antwerpen zum Einsatz von Künstlicher Intelligenz durch die europäischen Steuerbehörden, online aufgerufen am 17. Februar 2023 unter: <https://www.uantwerpen.be/en/projects/aitax/country-reports/france/>. Ferner Ulbrich Christian et al., PwC Tax Disruption Report 2021/2022 (zit. PwC 2021/2022), S. 32 ff., online aufgerufen am 17. Februar 2023 unter: <https://www.pwc.ch/en/insights/tax/tax-disruption-report-2021-2022.html>.

⁶ Siehe unten III.3.

⁷ Siehe etwa Ferber Michael, «Der Daten-Hunger der Steuerbehörden ist gestiegen»: Steuererwerbungen prüfen Social-Media-Profile von Steuerpflichtigen – auch in der Schweiz, NZZ, 22.07.2022, online aufgerufen am 17. Februar 2023 unter <https://www.nzz.ch/finanzen/private-finanzen/steueraemter-pruefen-social-media-profile-von-steuerpflichtigen-ld.1694704>.

werden die Überprüfung der Steuererklärung auf Plausibilität im automatisierten Veranlagungsverfahren und die damit einhergehenden grundrechtlichen und datenschutzrechtlichen Konsequenzen untersucht (IV). Abschliessend wird ein Fazit gezogen (V). Im vorliegenden Beitrag werden beispielhaft die Rechtsgrundlagen der Kantone Basel-Stadt (BS), Bern (BE), Obwalden (OW) und Zürich (ZH) herangezogen.

II. Schrittweise Automatisierung

Automatisierungsbestrebungen im Steuerverfahren sind kein Phänomen des 21. Jahrhunderts. Eine gewisse Standardisierung des Steuerprozesses geht bereits auf die Ursprünge des heute geltenden Abgaberechts zurück. Dazu zählen insbesondere verpflichtende Steuererklärungsformulare,⁸ um Effizienz zu steigern und Verständlichkeit zu fördern (Form- bzw. Formularzwang⁹).

Mitte des 20. Jahrhunderts wurde im Rahmen des beginnenden rechtswissenschaftlichen Diskurses deutlich, dass die computergestützte Automatisierung insbesondere in jenen Bereichen sinnvoll ist, in denen es um Kalkulationen geht.¹⁰ Die Prozessautomatisierung bildete bereits damals einen zentralen Bestandteil der Diskussion um die allgemeine Verwaltungsdigitalisierung.¹¹ Dazu waren insbesondere jene Bereiche der öffentlichen Verwaltung geeignet, in denen weitestgehend standardisiert eine grosse Menge an gleichartigen Fällen bearbeitet werden (sogenannte Massenverwaltung).¹² Der Abbau von Automatisierungshemmnissen, wie beispielsweise der Verzicht auf das Unterschriftserfordernis beim Erlass von Massenverfügungen¹³, trug zur weiteren Förderung der allgemeinen Digitalisierung in diesen Bereichen bei. Während in den

⁸ Siehe beispielhaft § 3 Gesetz betreffend die direkten Steuern im Kanton Basel-Stadt vom 6. April 1922, wonach die schriftlichen Steuererklärungen über Einkommen und Vermögen in vom Regierungsrat festgelegten Formularen abgegeben werden mussten.

⁹ Siehe Beusch Michael/Rohner Roger, Möglichkeiten und Grenzen der elektronischen Einreichung von Steuererklärungen bei den direkten Steuern, in: Beusch Michael (Hrsg.), Steuerrecht 2007, Best of zsis, Zürich/Basel/Genf 2007, S. 179 ff., S. 201.

¹⁰ Etwa Luhmann Niklas, Recht und Automation in der öffentlichen Verwaltung: eine verwaltungswissenschaftliche Untersuchung, Diss. Speyer 1966, Berlin 1966, S. 68.

¹¹ Degrandi Benno, Die automatisierte Verfügungsverfügung, Diss. Zürich 1977, S. 18 ff.; Polomski Ralf-Michael, Der automatisierte Verwaltungsakt, Diss. Göttingen 1992, Berlin 1993, S. 53 ff.

¹² Polomski (Fn. 11), S. 25. Zum Begriff der Massenverwaltung Kiener Regina/Rütsche Bernhard/Kuhn Mathias, Öffentliches Verfahrensrecht, 3. Aufl., Zürich/St. Gallen 2021, N 878 ff.

¹³ Siehe BGE 96 V 13 E. 4b S. 21; BGE 97 IV 205 E. 1 S. 207.

Anfängen hauptsächlich die Teilautomation der Verfahren im Zentrum stand¹⁴ – etwa die Entscheidungsunterstützung im Bereich der direkten und indirekten Bundessteuern und Zölle sowie der Kantons- und Gemeindesteuern¹⁵ – geht es nunmehr um den vollautomatisierten Prozess.¹⁶ Beide Automationsgrade können heute durch den Einsatz von Künstlicher Intelligenz (KI) unterstützt werden.¹⁷ Automatisierung und der Einsatz von KI sind insbesondere dann sinnvoll und effizienzbringend, wenn viele Daten (in strukturierter oder unstrukturierter Weise) vorhanden sind (Big Data).¹⁸

Im Rahmen der Totalrevision des Datenschutzgesetzes¹⁹ wurden erste Schritte in Richtung vollautomatisierte Steuerverfahren getätigt, insbesondere indem Grundlagen für eine künftige Vollautomatisierung der Zollveranlagung vorgesehen wurden.²⁰ Das föderalistisch geprägte Steuersystem der Schweiz führt dazu, dass die Voraussetzungen und Bedürfnisse der kantonalen Steuerverwaltungen jedoch ganz unterschiedlich sind.²¹ Für alle Kantone gilt jedoch, dass eine ge-

¹⁴ Luhmann (Fn. 10), S. 49 ff.

¹⁵ Degrandi (Fn. 11), S. 27.

¹⁶ Zur Unterscheidung von Teil- und Vollautomation im Rahmen automatisierter Verfahren Obrecht Liliane, Verfügung und automatisierte Einzelentscheidung – *same same but different?*, ex/ante 2022, S. 38 ff., S. 39 f.

¹⁷ Kment Martin/Borchert Sophie, Künstliche Intelligenz und Algorithmen in der Rechtsanwendung, München 2022, N 45. Siehe dazu unten IV.2. Für Deutschland Schmidt Christoph, Die ausschliesslich automationsgestützte Steuerfestsetzung des § 155 Abs. 4 Satz 1 AO als neues Leitbild des deutschen Steuervollzugs: Ein frommer Wunsch oder tatsächliche Realität?, Tagungsband IRIS 2023, Bern 2023, S. 127 ff., S. 133 ff.

¹⁸ Kment/Borchert (Fn. 17), N 36; Russell Stuart/Norvig Peter, Artificial Intelligence – A Modern Approach, 4. Aufl., Hoboken 2021, S. 707 f.

¹⁹ Botschaft zum Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse vom 15.09.2017, BBl 6941-7192 (zit. Botschaft DSG 2017). Das revidierte Datenschutzgesetz (revDSG; BBl 2020 7639) tritt am 1. September 2023 in Kraft.

²⁰ Art. 21 revDSG führt den Begriff der automatisierten Einzelentscheidung ein, auf die sich das Zollgesetz (SR 631.0) im neuen Art. 38 Abs. 2 abstützt, siehe Botschaft DSG 2017 (Fn. 19), S. 7703. Siehe dazu vertieft Braun Binder Nadja, Der Untersuchungsgrundsatz als Herausforderung vollautomatisierter Verfahren, zsis) 2/2020, A5, N 17 ff. (abrufbar unter publ.zsis.ch/A5-2020). Im Zuge der laufenden Totalrevision des Zollgesetzes und der Schaffung eines neuen Rahmengesetzes sind weitere Schritte und Anpassungen zu erwarten; siehe Botschaft des Bundesrates vom 24. August 2022 zum Bundesgesetz über den Allgemeinen Teil der Abgabenerhebung und die Kontrolle des grenzüberschreitenden Waren- und Personenverkehrs durch das Bundesamt für Zoll und Grenzsicherheit sowie zur Totalrevision des Zollgesetzes zum neuen Zollabgabengesetz, BBl 2022 2724.

²¹ Fischer Claudio, Die digitale Steuerverwaltung, zsis) 2/2020, A6, N 7 ff. (abrufbar unter publ.zsis.ch/A6-2020). Beispielsweise sieht sich der Kanton Zürich nur schon aus demo-

wisse Automatisierung – ob mit KI oder ohne – insbesondere in zwei Bereichen vorstellbar ist: erstens bei der Überprüfung der Steuererklärung und zweitens bei der Steuerfestsetzung.²² Die Umsetzung sieht in den Kantonen allerdings unterschiedlich aus. Je nach Kanton werden einfache Fälle heute vollautomatisiert veranlagt.²³ So sieht es auch mit Blick auf die Überprüfung der Steuererklärung auf ihre Plausibilität aus. Im Kanton Bern beispielsweise erfolgt die Überprüfung zuerst automatisiert durch das Veranlagungssystem. Dieses basiert auf Algorithmen, deren Prüffregeln von Menschen aufgestellt wurden.²⁴ Damit die Überprüfung aber überhaupt automatisiert erfolgen kann, müssen die Steuerklärungsdaten elektronisch vorhanden sein.²⁵

III. Datenerhebung

1. Eingeschränkter Untersuchungsgrundsatz und Mitwirkungspflicht

Das Schweizer Steuersystem ist föderalistisch ausgestaltet.²⁶ Es wird durch die Gesamtheit der verschiedenen Steuerarten sowie die verschiedenen Steuerhoheitsträgerinnen und Steuerhoheitsträger, die zur Steuererhebung be-

graphischen Gründen mit einer viel grösseren Zahl zu veranlagender Personen konfrontiert als die restlichen Kantone.

²² Braun Binder Nadja et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4, Zürich 2021, S. 52 (zit. Braun Binder et al., Schlussbericht).

²³ Für den Kanton Bern ca. 20 % der Steuererklärungen, siehe Fischer Claudio/Daepf Annemarie, Digitalisierung am Beispiel der Steuerverwaltung des Kantons Bern, EF 2019, S. 327 ff., S. 329; für den Kanton St. Gallen ca. 5 %, siehe St. Galler Tagblatt vom 08.04.2018, Digitalisierung: Ostschweizer Steuerverwaltungen setzen auf Roboter, online aufgerufen am 17 Februar 2023 unter: <https://www.tagblatt.ch/ostschweiz/digitalisierung-ostschweizer-steuerverwaltungen-setzen-auf-roboter-ld.1006738>; für Obwalden ebenfalls ca. 20 %, siehe Projektpräsentation «Data Analytics in der Steuerverwaltung», online aufgerufen am 17. Februar 2023 unter: https://www.egovernment-wettbewerb.de/presentation/2020/Data_Analytics_in_der_Steuerveranlagung_Obwalden_Automatische_Veranlagung.pdf.

²⁴ Siehe Kanton Bern, Wie läuft der Steuerprozess ab? Von der Steuererklärung über die Veranlagung zur Steuerrechnung, online aufgerufen am 17. Februar 2023 unter: <https://www.sv.fin.be.ch/de/start/publikationen/steuerwissen/wie.html>.

²⁵ Fischer (Fn. 21), N 18.

²⁶ Siehe etwa Meier-Mazzucato Giorgio, Steuern Schweiz, 2. Aufl., Bern 2022, S. 37 ff.

rechtigt sind, geprägt.²⁷ Die Steuererhebung durch Bund, Kantone und Gemeinden setzt die Feststellung der individuell-konkreten Steuerschuld der steuerpflichtigen Personen voraus.²⁸ Das Steuerveranlagungsverfahren wird grundsätzlich vom Untersuchungsgrundsatz beherrscht; die Sachverhaltsabklärung obliegt der Veranlagungsbehörde.²⁹ Diese erforscht die rechtserheblichen Tatsachen von Amtes wegen. Die grosse Anzahl zu fällender Entscheide führt vor dem Hintergrund der Verfahrensökonomie jedoch dazu, dass die Steuerverwaltungen auf die Mitwirkung der steuerpflichtigen Personen angewiesen sind.³⁰ Nur sie verfügen regelmässig über die erforderlichen Tatsachenkenntnisse und haben Zugang zu den notwendigen Beweismitteln.³¹ Der Verwaltungsaufwand, der aufgebracht werden müsste, um jede einzelne Angabe jeder steuerpflichtigen Person zu überprüfen, stünde – jedenfalls heute noch – in keinem vernünftigen Verhältnis zum Ziel der Veranlagung.³² Die Steuerbehörden stellen den tatsächlichen und rechtlich massgebenden Sachverhalt also unter Mitwirkung der Steuerpflichtigen fest.³³ Diese haben die Pflicht zur Selbstdeklaration, während die Steuerbehörden die massgebenden Tatsachen überprüfen und den Steuerbetrag in der Veranlagungsverfügung festlegen. Mit diesem Schritt ist das Verfahren abgeschlossen.³⁴ Beide nehmen eigentliche Veranlagungshandlungen vor, weshalb von einem gemischten Veranlagungsverfahren die Rede ist.³⁵ Die Mitwirkungspflicht relativiert daher

²⁷ Meier-Mazzucato (Fn. 26), S. 37.

²⁸ Siehe Stauber Stephan, *Steuerverfahren, Verfahrensrechte der Betroffenen sowie Rechtsmittel*, insbesondere auch im Rahmen der schweizerischen Amtshilfe in Steuersachen gegenüber Drittstaaten, in: DACH, Europäische Anwaltsvereinigung e.V., *Verwaltungsverfahren und Verwaltungsgerichtsbarkeit*, 53. Tagung der DACH in Leipzig vom 10. bis 12. September 2015, S. 72 ff., S. 73; Schärrel Erwin, *Verfahrensrecht zur Steuereinschätzung: ein Einblick in das Verfahrensrecht der Steuerveranlagung nach Zürcher Recht unter Berücksichtigung des Wehrsteuerrechts*, Diss. Zürich 1981, S. 36 ff.

²⁹ Mäusli-Allenspach Peter/Oertli Mathias, *Das schweizerische Steuerrecht*, 10. Aufl., Bern 2020, S. 327 f.; Zweifel Martin, *Die Sachverhaltsermittlung im Steuerveranlagungsverfahren unter besonderer Berücksichtigung der Mitwirkung des Steuerpflichtigen und der Ermessensveranlagung bei der direkten Bundessteuer und der zürcherischen Einkommens- und Vermögenssteuer bzw. Ertrags- und Kapitalsteuer*, Habil. Zürich 1989, S. 10 f.

³⁰ Zweifel (Fn. 29), S. 124; Schärrel (Fn. 28), S. 36.

³¹ Zweifel (Fn. 29), S. 7, 45 ff.

³² Schärrel (Fn. 28), S. 36.

³³ § 150 Abs. 1 StG-BS; Art. 166 Abs. 2 StG-BE; Art. 189 Abs. 1 StG-OW; § 132 Abs. 1 StG-ZH.

³⁴ Eine Übersicht zum Steuerverfahren (beispielhaft zum Bund und den Kantonen Zürich und Aargau), in Meier-Mazzucato (Fn. 26), S. 1063 ff.; ferner Schärrel (Fn. 28), S. 36 ff., 71 ff.

³⁵ Mäusli-Allenspach/Oertli (Fn. 29), S. 321 f.

den Untersuchungsgrundsatz.³⁶ Hauptanwendungsfälle des ordentlichen gemischten Veranlagungsverfahrens sind die direkten Steuern auf Einkommen, Vermögen, Gewinn und Kapital.³⁷

Die Steuerbehörden haben die Steuern nach Massgabe der Bundesverfassung gleichmässig und verhältnismässig zu erheben.³⁸ Gleichmässigkeit bedeutet, dass die steuerpflichtigen Personen nach der individuellen, objektiv gemessenen wirtschaftlichen Leistungsfähigkeit veranlagt werden.³⁹ Gleichzeitig müssen steuerpflichtige Personen, deren Sachverhalte ähnlich sind und die sich in gleichen wirtschaftlichen Verhältnissen befinden, gleich behandelt werden (steuerrechtliches Diskriminierungsverbot).⁴⁰ Die Verhältnismässigkeit knüpft an die Mitwirkungspflicht der steuerpflichtigen Person an; sie setzt den Mitwirkungshandlungen die Schranken der Geeignetheit, Notwendigkeit und Zumutbarkeit.⁴¹

2. Datenerhebung als Datenbearbeitung

Zur Feststellung des massgeblichen Sachverhalts sind regelmässig Personendaten, also Angaben, die sich auf eine bestimmte oder bestimmbare natürliche Person beziehen, zu erheben.⁴² Da grundsätzlich jeder Umgang mit Personendaten als Datenbearbeitung i. S. d. Datenschutzes zu qualifizieren ist, gilt dies auch für das Erheben («Beschaffen») von Daten.⁴³ Ob dabei das Bundes-Datenschutzgesetz oder kantonale Datenschutzgesetze zur Anwendung kommen, hängt von der datenbearbeitenden Person ab. Bearbeiten Private oder Bun-

³⁶ Zweifel (Fn. 29), S. 14 f.

³⁷ Meier-Mazzucato (Fn. 26), S. 1063 ff.

³⁸ Vertieft Behnisch Urs, Kommentierung von Art. 127 N 24, 27, in: Waldmann Bernhard/Belser Eva-Maria/Epiney Astrid, Basler Kommentar Bundesverfassung, 3. Aufl., Basel 2015 (zit. BSK BV-VerfasserIn).

³⁹ BSK BV-Behnisch (Fn. 38), Art. 127 N 27.

⁴⁰ BSK BV-Waldmann (Fn. 38), Art. 8 N 13.

⁴¹ Zweifel (Fn. 29), S. 16 f.

⁴² Zum Begriff der Personendaten siehe Art. 5 lit. a revDSG; die Kantone folgen dieser Begriffsdefinition, siehe § 3 Abs. 3 IDG-BS; Art. 2 Abs. 1 KDSG-BE; Art. 2 Abs. 1 i. V. m. Art. 5 lit. a revDSG; § 3 Abs. 3 IDG-ZH; grundlegend Rudin Beat, Kommentierung § 3 IDG-BS N 17 ff., in: Rudin Beat/Baeriswyl Bruno, Praxiskommentar zum Informations- und Datenschutzgesetz des Kantons Basel-Stadt, Zürich 2014 (zit. Praxiskomm IDG-BS-VerfasserIn).

⁴³ Art. 5 lit. d revDSG; § 3 Abs. 5 IDG-BS; Art. 2 Abs. 4 KDSG-BE; Art. 2 Abs. 1 kDSG-OW i. V. m. Art. 5 lit. d revDSG; § 3 Abs. 5 IDG-ZH. Siehe dazu etwa Praxiskomm IDG-BS-Rudin (Fn. 42), § 3 N 50.

desorgane Daten, ist das revidierte Datenschutzgesetz (revDSG) einschlägig,⁴⁴ bearbeiten kantonale oder kommunale öffentliche Organe Daten, sind die kantonalen Datenschutzgesetze anwendbar.⁴⁵ Dabei spielt es – etwa bei der Veranlagung von Einkommens- oder Vermögenssteuern – keine Rolle, ob die Aufgabe im Bundesrecht oder im kantonalen Recht verankert ist. Neben dem in den Datenschutzgesetzen enthaltenen «allgemeinen Datenschutzrecht» sind Vorgaben des «besonderen Datenschutzrechts», hier insbesondere allfällige datenschutzrechtliche Vorgaben der kantonalen Steuergesetze, zu beachten.⁴⁶

Grundsätzlich gilt, dass ein öffentliches Organ – somit auch eine Steuerbehörde – Personendaten bearbeiten darf, wenn dies

(a) eine gesetzliche Grundlage zulässt oder

(b) zur Erfüllung einer gesetzlichen Aufgabe erforderlich ist.⁴⁷

Allerdings dürfen die Daten nur für einen festgelegten Zweck bearbeitet werden.⁴⁸ Der Zweck kann sich aus dem Regelungskontext, für besonders schützenswerte bzw. besondere Personendaten explizit aus einer formell-gesetzlichen Grundlage ergeben.⁴⁹ Die Datenerhebung wird zudem in zeitlicher Hinsicht durch das Verhältnismässigkeitsprinzip begrenzt.⁵⁰ Die Steuerbehörden dürfen die Steuerdaten nicht über die gesetzlich vorgesehene Dauer aufbewahren.⁵¹ Dies ergibt sich entweder aus der gesetzlichen Regelung oder aus dem Verhältnismässigkeitsprinzip, das verlangt, dass die Daten nur so lange aufbewahrt werden dürfen, wie dies zur Aufgabenerfüllung erforderlich ist. Die Steuerbehörden erheben die Daten für eine bestimmte Steuerart in einer

⁴⁴ Art. 2 Abs. 1 revDSG. Das revDSG spielt nur dort eine Rolle, wo die Eidgenössische Steuerverwaltung Steuern veranlagt, wie z. B. die Mehrwertsteuer. Zum revDSG siehe Fn. 19.

⁴⁵ Z. B. § 1 Abs. 1 IDG-BS.

⁴⁶ Zur Abgrenzung zwischen allgemeinem und besonderem Datenschutzrecht siehe Rudin Beat, Kommentierung Art. 2 DSG N 3 ff., in: Baeriswyl Bruno/Pärli Kurt/Blonski Dominika (Hrsg.), Kommentar Datenschutzgesetz, 2. Aufl., Bern 2023.

⁴⁷ § 9 Abs. 1 lit a und b IDG-BS; Art. 5 KDSG-BE; Art. 2 Abs. 1 kDSG-OW i. V. m. Art. 34 Abs. 1 und Abs. 3 lit. a revDSG; § 8 Abs. 1 IDG-ZH.

⁴⁸ § 12 Abs. 1 IDG-BS; Art. 5 Abs. 2 KDSG-BE; Art. 2 Abs. 1 kDSG-OW i. V. m. Art. 6 Abs. 3 revDSG; § 9 Abs. 1 IDG-ZH.

⁴⁹ Siehe zu den Anforderungen an die gesetzliche Grundlage [IV.4.](#)

⁵⁰ Siehe explizit § 9 Abs. 3 IDG-BS sowie Art. 6 Abs. 2 revDSG. Das Verhältnismässigkeitsprinzip genießt in jedem Fall Verfassungsrang (Art. 5 Abs. 2 BV).

⁵¹ Siehe nur Maurer-Lambrou Urs/Steiner Andrea, Kommentierung des Art. 4 N 11a DSG, in: Maurer-Lambrou Urs/Blechte Gabor P. (Hrsg.), Basler Kommentar Datenschutzgesetz Öffentlichkeitsgesetz, 3. Aufl., Basel 2014.

bestimmten Steuerperiode, um die steuerpflichtigen Personen gleichmässig und verhältnismässig zu veranlagern.⁵² Allerdings bleibt mit der Bestimmung des Zwecks unklar, welche Daten durch die Steuerbehörden tatsächlich erhoben werden dürfen. Und auch die zeitliche Perspektive ist gesetzlich nicht umrissen. Mithin wird den kantonalen Steuerbehörden erheblicher Ermessensspielraum eingeräumt.

3. Datenquellen

Neben den Daten, welche die steuerpflichtige Person aufgrund des Mitwirkungsgrundsatzes selbst deklariert, sind verschiedene Quellen vorstellbar, derer sich die Steuerverwaltungen bedienen können, um die Korrektheit von Angaben in der Steuererklärung zu prüfen. Die vorliegend betrachteten kantonalen Steuergesetze sehen beispielsweise vor, dass gewisse Dritte Auskunft- oder Bescheinigungspflichten unterliegen und bestimmte Informationen zu steuerpflichtigen Personen bekanntgeben müssen.⁵³ Es handelt sich dabei um spezifische Befragungen im Rechtsverhältnis zu einer steuerpflichtigen Person, nicht jedoch um eine standardisierte (automatisierte) Datenerhebung.⁵⁴ Weiter kann sie im Rahmen der Amtshilfe von anderen Behörden Informationen herausverlangen.⁵⁵

Das Internet eröffnet darüber hinaus weitere Möglichkeiten. Es kann gezielt nach steuerrechtlich relevanten Daten betreffend eine bestimmte Person

⁵² § 40 Abs. 1 StG-BS; Art. 67 Abs. 1 StG-BE; Art. 60 Abs. 1 StG-OW; § 49 Abs. 1 StG-ZH. Ferner dazu Meier-Mazzucato (Fn. 26), S. 1053 f.

⁵³ § 155 Abs. 1 und 156 Abs. 1 StG-BS; Art. 168 Abs. 1 und Art. 169 Abs. 1 StG-BE; Art. 193 Abs. 1 und Art. 194 Abs. 1 StG-OW; § 136 Abs. 1 und § 137 StG-ZH. Weitere Auskunfts- bzw. Meldepflichten, etwa für Intermediäre wie z. B. Anwältinnen und Anwälte, Steuerberaterinnen und Steuerberater etc., aber auch für Unternehmen dürften insbesondere für grenzüberschreitende Konstellationen von zunehmender Bedeutung werden, siehe die jüngste Änderung der Richtlinie (EU) 2011/16/EU des Rates vom 15. Februar 2011 über die Zusammenarbeit der Verwaltungsbehörden im Bereich der Besteuerung und zur Aufhebung der Richtlinie (DAC6), ABl. L 64 vom 11.03.2011, S. 1, zuletzt geändert durch die Richtlinie (EU) 2021/514 des Rates vom 22. März 2021 zur Änderung der Richtlinie 2011/16/EU über die Zusammenarbeit der Verwaltungsbehörden im Bereich der Besteuerung, ABl. L 104 vom 25.03.2021, S. 1.

⁵⁴ Siehe etwa Regli Florian, Kommentierung § 157 StG-BS N 2, in: Tarolli Schmidt Nadia/Villard Alain/Bienz Andreas/Jaussi Thomas (Hrsg.), Kommentar zum Basler Steuergesetz, Basel 2019; ferner Richner Felix/Frei Walter/Kaufmann Stefan/Rohner Tobias F., Kommentar zum Zürcher Steuergesetz, 4. Aufl., Zürich 2021, § 137 N 15.

⁵⁵ § 139-141 StG-BS; Art. 154 f. StG-BE; Art. 178-179a StG-OW; § 121 StG-ZH.

durchsucht werden.⁵⁶ In Frankreich wurde beispielsweise am 11. Februar 2021 ein Dekret⁵⁷ erlassen, das es den französischen Steuerbehörden erlaubt, verschiedene Online-Plattformen zu durchforsten, um Beweise für Steuerbetrug ausfindig zu machen (z. B. nicht deklarierte Swimmingpools). Dazu können die Generaldirektion für öffentliche Finanzen (DGFiP) und die französischen Zollbehörden während einer Experimentierphase von drei Jahren öffentlich zugängliche, personenbezogene Daten aus sozialen Medien, aber auch von Plattformen wie Airbnb oder der Kleinanzeigen-Webseite Le Bon Coin, erheben.⁵⁸ Die französische Datenschutzbehörde erklärte in einer unverbindlichen Empfehlung⁵⁹, es sei unerlässlich, die Umriss des geplanten Systems zu präzisieren und die Einhaltung des Verhältnismäßigkeitsgrundsatzes in allen Phasen der Datenbearbeitung zu prüfen. Schliesslich sei ein Gleichgewicht zu finden zwischen der Bekämpfung von Steuerbetrug und der Wahrung von Grundrechten der betroffenen Personen.

Eine weitere Entwicklung im Kontext der Digitalisierung stellt der zivile Drohneinsatz dar.⁶⁰ Steuerbehörden werden insbesondere dann an einem Drohneinsatz interessiert sein, wenn es um die bildliche Lageerfassung geht.⁶¹ So wurden etwa in Griechenland experimentell Drohnen eingesetzt, um in touris-

⁵⁶ Siehe auch zur Internetrecherche bei der Sachverhaltsermittlung durch Steuergerichte Jutzi Marc, Der Steuerrichter im Internet – Zulässigkeit und Validität von Internetrecherchen bei der Sachverhaltsermittlung, Justice – Justiz – Giustizia 2/2021, N 17 ff.

⁵⁷ Décret no 2021-148 du 11 février 2021 portant modalités de mise en œuvre par la direction générale des finances publiques et la direction générale des douanes et droits indirect de traitements informatisés et automatisés permettant la collecte et l'exploitation de données rendues publiques sur les sites internet des opérateurs de plateforme en ligne.

⁵⁸ In den Medien kursierten darauffhin international verschiedene Artikel, siehe beispielsweise Darmanin Jules, France starts scouring social media to catch tax fraudsters, POLITICO 19.02.2021, online aufgerufen am 17. Februar 2023 unter: <https://www.politico.eu/article/france-starts-scrapping-social-media-to-catch-tax-fraudsters/>; Breeden Aurelien, French Tax Collectors Use A.I. to Spot Thousands of Undeclared Pools, New York Times 30.08.2022, online aufgerufen am 17. Februar 2023 unter: <https://www.nytimes.com/2022/08/30/world/europe/france-taxes-pools-artificial-intelligence.html>.

⁵⁹ Commission Nationale Informatique & Libertés, Délibération no 2019-114 du 12 septembre 2019 portant avis sur le projet d'article 9 du projet de loi de finances pour 2020, online aufgerufen am 17. Februar 2023 unter: <https://www.legifrance.gouv.fr/cnil/id/CNIL-TEXT000039167079/>.

⁶⁰ Eine Übersicht zur Drohnenutzung durch den Staat für nicht-militärische Zwecke in Straub Sebastian, Beobachtungen aus der Luft – Grenzen des staatlichen und zivilen Drohneinsatzes, 2. Aufl., Frankfurt am Main 2019, S. 2 f.

⁶¹ Straub (Fn. 60), S. 3.

tischen Regionen Angaben in Steuererklärungen zu überprüfen.⁶² Anhand der durch die Drohnen erhobenen Daten konnten die Behörden feststellen, wie viele Passagiere sich an Bord von Ausflugsbooten befanden und diese Daten mit den deklarierten Quittungen der Steuerpflichtigen in der Steuererklärung abgleichen.⁶³

Werden hingegen sogenannte Smart Devices (intelligente Geräte) eingesetzt, die (steuerrechtlich relevante) Daten Dritter in Echtzeit an eine Steuerbehörde übermitteln, wird der Steuererhebungsprozess gewissermassen umgekehrt: Es würde nicht mehr um eine Überprüfung der Angaben in der Steuererklärung gehen, sondern bereits vorausschauend («beobachtend») festgesetzt, welche Angaben beizuziehen sind und welche nicht.⁶⁴

So nutzt beispielsweise die Steuerverwaltung Grossbritanniens ein Datenerfassungs- und Analyseprogramm (His Majesty's Revenue and Customs Connect Computer System⁶⁵), das aus mindestens 30 verschiedenen Datenbanken unterschiedliche Informationen sammelt, analysiert und für weitere Zwecke speichert.⁶⁶ Darunter finden sich Daten von online Zahlungsdienstleistern wie PayPal oder Verkaufswebseiten wie Amazon und eBay, Social Media-Daten oder Informationen zu Flugpassagieren und Flügen.⁶⁷

⁶² De Hoon Iven, Greece uses drones to find tax fraud, No More Tax, online aufgerufen am 17. Februar 2023 unter: <https://nomoretax.eu/greece-drones-tax-fraud/>.

⁶³ Papadimas Lefteris, Greece uses high-tech drones to fight tax evasion in holiday hot spots, reuters 24.09.2018, online aufgerufen am 17. Februar 2023 unter: <https://www.reuters.com/article/us-greece-taxes-drones-idUSKCN1M41SJ>.

⁶⁴ Ehrke-Rabel Tina, Profiling im Steuervollzug, FR 2019, S. 45 ff. (zit. Ehrke-Rabel, Profiling), S. 50.

⁶⁵ Siehe etwa Andersen Chris, What is HMRC's Connect Computer System and how is it Being Used for Tax Compliance?, CompanyDebt 28.11.2022, online aufgerufen am 17. Februar 2023 unter: <https://www.companydebt.com/hmrc-tax-problems/hmrcs-connect-computer-system-used-tax-compliance/>.

⁶⁶ Rigney Paul, The all seeing eye – an HMRC success story?, HMRC Administration November/Dezember 2016, S. 8 f.

⁶⁷ Sanghrajka Jay, HMRC's Connect computer and investigations, TAXATION 14.07.2020, online aufgerufen am 17. Februar 2023 unter: <https://www.taxation.co.uk/articles/hmrc-s-connect-computer-and-investigations>; Ulbrich 2021/2022 (Fn. 5), S. 43 f.

IV. Überprüfung der Steuererklärung

1. Automatisierung und Untersuchungsgrundsatz

Während der Untersuchungsgrundsatz durch die Mitwirkungspflicht der steuerpflichtigen Personen im Veranlagungsverfahren eingeschränkt wird, obliegt es letztlich trotz allem den Steuerbehörden, den massgeblichen Sachverhalt festzulegen.⁶⁸ Dazu zählt auch die Überprüfung der Plausibilität der Angaben in den Steuererklärungen. Das heisst, dass sich eine Behörde von Amtes wegen von der Wahrheit und Richtigkeit des Sachverhaltes überzeugen muss. Der Untersuchungsgrundsatz bedeutet somit zweierlei: Die Behörden sind einerseits verpflichtet und berechtigt, die Sachverhalte abzuklären. Andererseits dürfen sie nur diejenigen Tatsachen übernehmen, von deren Richtigkeit sie sich überzeugt haben. Deshalb prüfen die Steuerverwaltungen die Steuererklärungen von Amtes wegen auf deren Richtigkeit und nehmen die erforderlichen Untersuchungen vor.⁶⁹ Dazu kann sich eine Behörde verschiedener gesetzlich vorgesehener Untersuchungsmittel bedienen.⁷⁰ Diese berechtigen sie, von einer steuerpflichtigen Person wie auch – in einem thematisch eng begrenzten Rahmen – von Dritten gewisse Auskünfte zu verlangen oder sich auf Informationen anderer Behörden im Rahmen der Amtshilfe abzustützen.⁷¹ Damit können Angaben, die Steuerpflichtige über sich selbst machen anhand anderer Informationsquellen über die Steuerpflichtigen überprüft werden. Die im Zuge der Digitalisierung hinzugetretenen Datenquellen – wie etwa Informationen, die Steuerpflichtige selbst auf den sozialen Medien teilen – könnten eine weitere Möglichkeit eröffnen, um Angaben in Steuererklärungen mit anderen Informationsquellen über die betroffene Person zu vergleichen.⁷²

Die von steuerpflichtigen Personen gemachten Angaben könnten zudem mit Angaben anderer steuerpflichtiger Personen verglichen werden. Die jährliche Steuererhebung führt dazu, dass eine grosse Menge historischer und vergleichbarer Daten vorliegt. Da die Ermittlung des steuerrechtlich relevanten Sachverhalts bereits weitgehend in standardisierter Form erfolgt⁷³ und somit eine strukturierte Datengrundlage existiert, bietet sich diese als Ausgangsbasis zur Automatisierung des Steuerverfahrens und der Überprüfung von An-

⁶⁸ Siehe oben [III.1](#).

⁶⁹ § 158 Abs. 1 StG-BS; § 138 Abs. 1 StG-ZH; Art. 196 Abs. 1 StG-OW; Art. 166 Abs. 2 StG-BE.

⁷⁰ Grundlegend Zweifel (Fn. 29), S. 29.

⁷¹ Siehe oben [III.3](#).

⁷² Siehe oben [III.3](#).

⁷³ Siehe oben [II](#).

gaben in Steuererklärungen geradezu an.⁷⁴ Die Ergebnisse der Überprüfung der Plausibilität der Steuererklärung dienen dazu, abzuschätzen, ob eine steuerpflichtige Person die Steuererklärung wahrheitsgetreu ausgefüllt hat oder nicht.⁷⁵ Beispiele im internationalen Kontext zeigen, dass die manuelle Überprüfung der Steuererklärungen durch eine solche automationsgestützte Risikobeurteilung ausgelöst werden kann.⁷⁶

2. Einsatz von Risikomanagementsystemen

Aufgrund des Charakters als Massenverfahren⁷⁷ sehen sich die Steuerbehörden im Veranlagungsverfahren mit der Frage konfrontiert, wie weitgehend sie die Angaben der steuerpflichtigen Personen in den Steuererklärungen tatsächlich prüfen können.⁷⁸ Zu beachten sind dabei stets die Grundsätze der Verhältnismässigkeit und der rechtsgleichen Behandlung.⁷⁹ Angesichts der demographischen Entwicklung, der ansteigenden Datenverarbeitung und der knappen Personalressourcen steht die Einhaltung der verfassungsrechtlichen Grundsätze bei der manuellen Überprüfung der Steuererklärungen vor grossen Herausforderungen.⁸⁰ Die automatisierte Risikoprüfung von in der Steu-

⁷⁴ Denn Automatisierung erfordert Strukturiertheit der Datensets, siehe dazu Esterer Fritz/Just Vanessa/Fettke Peter, Sind Steuerfunktionen bereit für Künstliche Intelligenz?, EF 2019, S. 16 f., S. 16.

⁷⁵ Die steuerpflichtige Person muss die Steuererklärung wahrheitsgemäss ausfüllen, siehe § 151 Abs. 2 StG-BS; Art. 170 Abs. 2 StG-BE; Art. 190 Abs. 3 StG-OW; § 133 Abs. 2 StG-ZH.

⁷⁶ Dazu sogleich IV.2.

⁷⁷ Siehe oben III.1 und spezifisch Kiener/Rütsche/Kuhn (Fn. 12), N 904 ff.

⁷⁸ Schärler (Fn. 28), S. 37.

⁷⁹ BSK BV-Epiney (Fn. 38), Art. 5 N 67 ff.; BSK BV-Waldmann (Fn. 38), Art. 8 N 13; vertieft Schärler (Fn. 28), S. 22 f.

⁸⁰ Die ständige Wohnbevölkerung der Schweiz wird durch das Bundesamt für Statistik BFS auf knapp 10.5 Millionen im Jahr 2050 geschätzt, siehe Bundesamt für Statistik, Zukünftige Entwicklung der Bevölkerung, 25.05.2021, online aufgerufen am 17. Februar 2023 unter: <https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/zukuenftige-entwicklung.html>. Siehe zu den damit einhergehenden Herausforderungen für das Steuerwesen Fraunhofer prognos, Herausforderungen für das Steuerrecht durch die demografische Entwicklung in Deutschland – Analyse einer Problemstellung, Endbericht vom 26. August 2016 im Auftrag des Bundesministeriums der Finanzen, online aufgerufen am 17. Februar 2023 unter: <https://www.bundesfinanzministerium.de/Monatsberichte/2017/01/Inhalte/Kapitel-3-Analysen/3-4-Herausforderungen-Steuerrecht-demografische-Entwicklung.html>.

ererklärung gemachten Angaben kann deshalb Unterstützung leisten.⁸¹ Insbesondere Risikomanagementsysteme (RMS) können dabei zum Einsatz kommen.⁸² Risiko bedeutet im steuerrechtlichen Kontext, dass eine gewisse Wahrscheinlichkeit besteht, dass die Angaben in einer Steuererklärung unwahrheitsgetreu sind,⁸³ mithin die gesetzesmäßige Steuerfestsetzung verfehlt wird.⁸⁴

In Deutschland existiert seit dem 1. Januar 2017 eine Rechtsgrundlage, welche die verpflichtende Einführung von RMS gesetzlich verankert (§ 88 Abs. 5 Abgabenordnung, AO⁸⁵).⁸⁶ Diese wurde im Rahmen der Zulassung vollautomatisierter Steuerverfahren geschaffen.⁸⁷ Die verpflichtende Einführung von RMS soll die mit der Vollautomation einhergehenden Abstriche am Untersuchungsgrundsatz kompensieren.⁸⁸ Die Einzelfallprüfung wurde damit durch eine automationsbasierte Risikoprüfung ersetzt. Gemäss erläuterndem Bericht besteht ein Risikomanagementsystem aus der systematischen Erfassung und Bewertung von Risikopotenzialen sowie der Steuerung von Reaktionen in Abhängigkeit der festgestellten Risikopotenziale.⁸⁹ Durch das Aussondern der risikobehafteten Fälle sollen die personellen Ressourcen (natürlicher Personen)

⁸¹ Koch Stefanie, Das Risikomanagementsystem der Finanzverwaltung, Diss. Münster 2019, S. 1.

⁸² RMS werden in der Privatwirtschaft seit längerer Zeit eingesetzt. Siehe etwa Pechtl Andreas, Ein Rückblick: Risikomanagement von der Antike bis heute, in: Romeike Frank/Finke Robert B. (Hrsg.), Erfolgsfaktor Risiko-Management, Wiesbaden 2003, S. 15 ff.

⁸³ Zum Risikobegriff im Steuerrecht Koch (Fn. 81), S. 74 f.

⁸⁴ Drüen Klaus-Dieter, Amtsermittlungsgrundsatz und Risikomanagement, in: Hey Johanna, Digitalisierung im Steuerrecht, Köln 2019, S. 193 ff., S. 207.

⁸⁵ 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.

⁸⁶ Siehe dazu Schmidt (Fn. 17), S. 128 ff.; Spilker Bettina, E-Government – Anforderungen an das Verwaltungsverfahren: Verfahrensrechte und Mindestanforderungen am Beispiel der AO?, NVwZ 2022, S. 680 ff., S. 682 f.

⁸⁷ Zur historischen Übersicht des Eingangs von Risikomanagementsystemen in die Deutsche Gesetzgebung siehe Koch (Fn. 81), S. 5 ff.

⁸⁸ Braun Binder Nadja, Algorithmisch gesteuertes Risikomanagement in digitalisierten Besteuerungsverfahren, in: Unger Sebastian/von Ungern-Sternberg Antje (Hrsg.), Demokratie und künstliche Intelligenz, Tübingen 2019, S. 161 ff. (zit. Braun Binder, Risikomanagement), S. 162.

⁸⁹ Deutscher Bundestag, Drucksache 18/7457, Gesetzentwurf der Bundesregierung: Entwurf eines Gesetzes zur Modernisierung des Besteuerungsverfahrens, S. 69.

auf jene Fälle beschränkt werden, die tatsächlich prüfungsbedürftig sind.⁹⁰ Risikoreich sind beispielsweise Fälle, die Änderungen in den Sachverhaltsangaben im Vergleich zum Vorjahr aufweisen. Auch in Österreich sind RMS im Einsatz (siehe §§ 48a, 114 Abs. 4 Bundesabgabenordnung, BAO⁹¹). Diese sind in den Rechtsgrundlagen allerdings nicht wie in Deutschland legaldefiniert, sondern werden lediglich erwähnt.⁹²

Zusammengefasst kann Risikomanagement als ein Prozess beschrieben werden, der Risiken identifiziert, evaluiert und bewältigt.⁹³ Bei entsprechender Ausgestaltung des Systems kann dessen Einsatz dazu führen, den Grundsatz der Gleichmässigkeit der Besteuerung⁹⁴ besser umzusetzen, als dies durch natürliche Personen möglich wäre.⁹⁵ Fehler in der Ausgestaltung hätten jedoch weitreichende Konsequenzen.⁹⁶

Der Einsatz von RMS ist prinzipiell auch in der Schweiz vorstellbar. Hierzu bedarf es allerdings einiger grundlegender Voraussetzungen. Dazu zählt, dass die Datensätze so aufbereitet sein müssen, dass sie der technischen Ausgestaltung des Systems entsprechend verarbeitet werden können.⁹⁷ Die elektronische Übermittlung der steuerrelevanten Angaben bildet mit anderen Worten den notwendigen Ausgangspunkt für den Einsatz von RMS. Zwar besteht in den Kantonen keine flächendeckende Pflicht zur elektronischen Einreichung der Steuererklärung. Gemäss Bundesgesetz über elektronische Verfah-

⁹⁰ Ehrke-Rabel Tina, 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. (zit. Ehrke-Rabel, Automatisierung), S. 49.

⁹¹ 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.

⁹² Dazu vertieft Ehrke-Rabel, Automatisierung (Fn. 90), S. 49 ff. Für einen eigentlichen Vergleich siehe Schmidt Christoph, Das modernisierte Besteuerungsverfahren in Deutschland im Vergleich zu Österreich, Diss. Potsdam 2019, Berlin 2021, S. 142 ff.

⁹³ Europäische Kommission, Risk Management Guide For Tax Administrations, Version 1.02, Februar 2006, S. 7, online aufgerufen am 17. Februar 2023 unter: https://taxation-customs.ec.europa.eu/system/files/2016-09/risk_management_guide_for_tax_administrations_en.pdf.

⁹⁴ Siehe oben III.1.

⁹⁵ Ehrke-Rabel, Automatisierung (Fn. 90), S. 49 f.; Schmidt (Fn. 17), S. 130.

⁹⁶ Allgemein beim Einsatz neuer Technologien Martini Mario/Nink David, Wenn Maschinen entscheiden ... – vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz, NVwZ-Extra 2017, S. 1 ff., S. 1.

⁹⁷ Siehe dazu Koch (Fn. 81), S. 66 f.

ren im Steuerbereich⁹⁸ und der damit einhergehenden Anpassung des Steuerharmonisierungsgesetzes⁹⁹ sind die Kantone aber verpflichtet, bis zum 1.1.2024 die Möglichkeit des elektronischen Einreichens der Steuererklärung vorzusehen.¹⁰⁰

Eine weitere Voraussetzung bildet die Schaffung einer hinreichenden Rechtsgrundlage für den Einsatz von RMS. Die Anforderungen an die Rechtsgrundlage lassen sich allerdings nicht in allgemeingültiger Weise formulieren. Sie hängen vielmehr von der konkreten Ausgestaltung des RMS und den genutzten Daten ab. Die genaue Funktionsweise des Systems und insbesondere vordefinierte Kriterien, die der Risikoanalyse als Grundlage dienen, werden wohl aber nicht vollständig öffentlich kommuniziert werden, da dies zu Umgehungsstrategien der Steuerpflichtigen führen könnte.¹⁰¹

Es ist ferner denkbar, dass RMS mit Methoden maschinellen Lernens verbunden werden.¹⁰² Auch wenn es aufgrund der grossen Datenmengen offensichtlich erscheint, RMS mit maschinellen Lernverfahren zu verbinden,¹⁰³ hängt der Einsatz einer spezifischen Technologie massgeblich von der praktischen Umsetzbarkeit und deren potenzieller Auswirkungen auf die betroffenen Personen ab.¹⁰⁴ In Deutschland ist beispielsweise aufgrund einer Ausnahmeregelung in § 32a Abs. 2 Ziff. 2 AO unklar, welche spezifische Technologie für RMS heu-

⁹⁸ Bundesgesetz vom 18. Juni 2021 über elektronische Verfahren im Steuerbereich (SR 641.10).

⁹⁹ Bundesgesetz über die Harmonisierung der direkten Steuern der Kantone und Gemeinden (Steuerharmonisierungsgesetz, StHG) vom 14. Dezember 1990 (SR 642.14).

¹⁰⁰ Art. 38b StHG, siehe die Chronologie der Bundeskanzlei zum Bundesgesetz über elektronische Verfahren im Steuerbereich, online aufgerufen am 17. Februar 2023 unter: <https://www.bk.admin.ch/ch/d/pore/rf/cr/2021/20212111.html>.

¹⁰¹ Für die deutsche Rechtslage Braun Binder, Risikomanagement (Fn. 88), S. 174 f.

¹⁰² Maschinelle Lernverfahren werden als Teilbereich von KI verstanden. Dabei ist allerdings zu berücksichtigen, dass für KI kein einheitliches Begriffsverständnis existiert, siehe etwa Kaplan Jerry, Artificial Intelligence: What everyone needs to know, New York 2016, S. 1 ff.; Ertel Wolfgang, Grundkurs Künstliche Intelligenz, 4. Aufl., Wiesbaden 2016, S. 1 ff. Siehe exemplarisch zur Evaluierung des Kreditvergaberisikos etwa Korn Mona, Potenziale künstlicher Intelligenz und maschinellen Lernens im Risikomanagement von Banken, Controlling – Zeitschrift für erfolgsorientierte Unternehmenssteuerung, 2019, S. 42 ff., S. 43 f.

¹⁰³ Braun Binder Nadja, Artificial Intelligence and Taxation: Risk Management in Fully Automated Taxation Procedures, in: Wischmeyer Thomas/Rademacher Timo (Hrsg.), Regulating Artificial Intelligence, Cham 2020, S. 295 ff. (zit. Braun Binder, Taxation), S. 300 N 16.

¹⁰⁴ Drüen (Fn. 84), S. 214 f. Grundrechts- und verfahrensrechtsrelevante Probleme stellen sich insbesondere dann, wenn die RMS nicht mehr manuell programmiert werden, sondern selbstlernende Systeme (Stichwort *machine learning*) eingesetzt werden, dazu Koch (Fn. 81), S. 71.

te beigezogen wird. Eine Auskunftspflicht besteht nämlich nicht, wenn die Erteilung der Auskunft die ordnungsgemässe Erfüllung der in die Zuständigkeit der Steuerbehörden fallenden Aufgaben und Pflichten gefährden würde. Dies ist beispielsweise der Fall, wenn durch die Offenlegung der Informationen Rückschlüsse auf die Ausgestaltung von RMS möglich sind und dadurch die Aufdeckung steuerlich bedeutsamer Sachverhalte wesentlich erschwert wird.¹⁰⁵

Um aus rechtsstaatlicher Perspektive die Anforderungen zu konkretisieren, ist es jedoch zweitrangig, welche spezifische Technologie zur Anwendung gelangt.¹⁰⁶ Wichtiger ist vielmehr, dass durch den Einsatz maschineller Lernverfahren gewissermassen ein Perspektivenwechsel stattfindet: Anstelle von Kausalitäten basieren diese Systeme – z. B. künstliche neuronale Netze¹⁰⁷ – auf Wahrscheinlichkeiten.¹⁰⁸ Die Daten werden nicht anhand von vorgegebenen Prüfregeln verarbeitet, sondern stellen durch das Erkennen von Korrelationen eigene Regeln auf, die auf die Datensätze angewendet werden.¹⁰⁹ Dabei bleibt regelmässig unklar, wie die Systeme zu ihren Ergebnissen gelangen und welche Daten schliesslich ausschlaggebend sind.¹¹⁰ Für den Einsatz von RMS bedeutet dies, dass Steuererklärungen nicht aufgrund vorgegebener, standardisierter Kriterien beurteilt werden, sondern das System selbst entscheidet, welche Kriterien im Einzelfall für die Ergebnisse massgebend sind. Die Wahl der Kriterien sowie die darauf basierenden Ergebnisse sind dadurch weder vorhersehbar noch überprüfbar.

Erste Praxiserfahrungen im Ausland zeigen, dass solche Systeme fehleranfällig sind. So wurde im Jahr 2019 etwa bekannt, dass die niederländischen Steuerbehörden einen selbstlernenden Algorithmus zur Erstellung von Risikoprofilen verwendet hatten, um Betrug bei Kinderbetreuungsleistungen zu erkennen. Auf der Grundlage der Risikoindikatoren des Systems wurden Familien

¹⁰⁵ Braun Binder, *Taxation* (Fn. 103), S. 299 N 12 f., S. 304 N 4.

¹⁰⁶ Braun Binder, *Risikomanagement* (Fn. 88), S. 174.

¹⁰⁷ Zum Begriff etwa Oettinger Michael, *Data Science*, 2. Aufl., Hamburg 2020, S. 119 ff.

¹⁰⁸ Hartmut Ernst/Schmidt Jochen/Beneken Gerd, *Grundkurs Informatik*, 7. Aufl., Wiesbaden 2020, S. 801 ff.

¹⁰⁹ Ertel (Fn. 102), S. 265 ff. Zu datenschutzrechtlichen Fragen und dem Einsatz maschineller Lernverfahren Glass Philip, *Datenschutzrecht für künstliche Intelligenz in der öffentlichen Verwaltung*, in: Widmer Micheal (Hrsg.), *Datenschutz: Rechtliche Schnittstellen*, Zürich 2023, S. 179 ff., S. 202 ff.

¹¹⁰ Vogel Paul, *Künstliche Intelligenz und Datenschutz*, Diss. Würzburg 2021, Baden-Baden 2022, S. 106 ff.; Wischmeyer Thomas, *Artificial Intelligence and Transparency: Opening the Black Box*, in: Wischmeyer Thomas/Rademacher Timo (Hrsg.), *Regulating Artificial Intelligence*, Cham 2020, S. 75 ff., S. 7 f.; Martini Mario, *Blackbox Algorithmus*, Berlin 2019, S. 28 ff.

schon bei einem blossen Verdacht auf Betrug mit Bussen belangt. Zehntausende von Familien – oft mit geringem Einkommen oder Angehörige ethnischer Minderheiten – waren von den Folgen betroffen.¹¹¹ Die Fehleranfälligkeit maschineller Lernverfahren kann u. a. auf mangelnde Qualität von Trainingsdaten zurückgeführt werden.¹¹² In Bezug auf den Einsatz von RMS besteht das Risiko, dass risikobehaftete Fälle nicht erkannt werden. Dadurch würden der Untersuchungsgrundsatz und der Grundsatz der Gleichmässigkeit der Besteuerung beeinträchtigt.¹¹³ Hinzu kommt, dass bei einem Einsatz (komplexer) maschineller Lernverfahren die Kriterien, die der Risikoanalyse zugrunde gelegt werden, kaum vorhersehbar sind, da diese vom lernenden System selbst entwickelt werden. Damit ist der Prozess kaum mehr nachvollziehbar.

3. Zweckänderung der Daten im Verfahren

Im Rahmen der Veranlagungsverfahren werden Daten der steuerpflichtigen Personen für eine bestimmte Steuerart in einer bestimmten Steuerperiode erhoben (Grundsatz der Zweckbindung).¹¹⁴ Im analogen Verfahren prüfen Steuerbehörden die Erklärungen gestützt auf die durch die steuerpflichtigen Personen bereitgestellten Beweismittel sowie gesetzlich definierte Untersuchungsmittel.¹¹⁵ Im automatisierten Steuerverfahren eröffnen sich hingegen weitere Möglichkeiten (z. B. *Data Mining*¹¹⁶). Vorstellbar ist etwa, dass Daten verschiedener steuerpflichtiger Personen genutzt werden, um die Steuererklärung einer weiteren

¹¹¹ Siehe dazu Heikkilä Melissa, Dutch Scandal serves as a warning for Europe over risks of using algorithms, POLITICO 29.03.2022, online aufgerufen am 17. Februar 2023 unter: <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>. Zur allgemeinen Fehleranfälligkeit algorithmischer Systeme siehe Braun Binder et al., Schlussbericht (Fn. 22), S. 40 ff.

¹¹² Voss Jakob, Datenqualität als Grundlage qualitativer Inhaltserschliessung, in: Franke-Maier Michael/Kasprzik Anna/Ledl Andreas et al. (Hrsg.), Qualität der Inhaltserschliessung, Berlin/Boston 2021, S. 167 ff., S. 168 ff.; Stevens Jeremy, Datenqualität bei algorithmischen Entscheidungen, CR 2020, S. 73 ff., S. 74; Zweig Katharina/Krafft Tobias, Fairness und Qualität algorithmischer Entscheidungen, in: Mohabbat Kar Resa/Thapa Basanta/Parycek Peter (Hrsg.), (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft, Berlin 2018, S. 204 ff., S. 211 ff.

¹¹³ Zu den Grundsätzen oben [III.1](#).

¹¹⁴ Siehe oben [III.2](#).

¹¹⁵ Siehe oben [IV.1](#).

¹¹⁶ Zu Data Mining und datenschutzrechtlicher Zweckbindung siehe Epiney Astrid, Datenschutzrechtliche Grundsätze und Garantien, in: Belser Eva Maria/Epiney Astrid/Waldmann Bernhard/Bickel Jürg (Hrsg.), Datenschutzrecht: Grundlagen und öffentliches Recht, Bern 2011, S. 509 ff., § 9 N 34.

Person auf Plausibilität zu prüfen.¹¹⁷ Darin würde allerdings eine Änderung des Zwecks der erhobenen Daten liegen.¹¹⁸ Der Zweck würde nicht mehr die Bearbeitung der Daten einer bestimmten steuerpflichtigen Person in einer bestimmten Steuerperiode, sondern die Bearbeitung der Daten verschiedener steuerpflichtiger Personen bzw. die Bearbeitung der Daten einer Person zur Nutzung für eine andere Person umfassen.¹¹⁹

Ob eine solche Zweckänderung von den Rechtsgrundlagen für die ursprüngliche Datenbearbeitung gestützt wird, ist fraglich. Grundsätzlich ist die Bearbeitung von Personendaten zu einem anderen Zweck nicht durch die gesetzliche Grundlage gedeckt, welche die ursprüngliche Erhebung rechtfertigt.¹²⁰ Unzulässig ist auch eine Datenerhebung «auf Vorrat», um diese später in einem anderen Kontext zu verwenden.¹²¹ In gesetzlichen Bekanntgabebestimmungen kann hingegen eine Rechtfertigung der Zweckänderung mitenthalten sein.¹²² Wenn beispielsweise Steuerdaten von der Steuerverwaltung an andere Behördenstellen bekannt gegeben werden dürfen, dann regelmässig nicht zu dem Zweck, zu dem sie erhoben worden sind (nämlich zur Veranlagung der Steuern), sondern 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.

4. Profiling im Veranlagungsverfahren

Die Datenüberprüfung durch die Steuerbehörden hat zum Ziel, die Steuererklärungen der steuerpflichtigen Personen, mithin die Korrektheit und Vollständigkeit der Angaben, zu prüfen.¹²³ Dies ändert sich im vollautomatisierten Steuerverfahren grundsätzlich nicht. Allerdings kann das Beiziehen anderer

¹¹⁷ Ehrke-Rabel, Automatisierung (Fn. 90), S. 50.

¹¹⁸ Zum Zusammenspiel von Zweckbindung und automatisierter Datenverarbeitung (Profiling, siehe unten [IV.4](#)), Heuberger Olivier, Profiling im Persönlichkeits- und Datenschutzrecht der Schweiz, Diss. Luzern 2019, Zürich/Basel/Genf 2020, N 237 ff.

¹¹⁹ Für Österreich bereits Ehrke-Rabel, Automatisierung (Fn. 90), S. 50.

¹²⁰ Epiney (Fn. 116), § 9 N 29 f. Ferner Praxiskomm IDG-BS-Rudin (Fn. 42), § 12 N 10 ff.

¹²¹ Epiney (Fn. 116), § 9 N 33. Dies gilt auch auf kantonaler Ebene, siehe etwa Praxiskomm IDG-BS-Rudin (Fn. 42), § 12 N 4.

¹²² Praxiskomm IDG-BS-Rudin (Fn. 42), § 12 N 10 ff., insbesondere N 15.

¹²³ Siehe oben [III.1](#).

Datenquellen und der Abgleich dieser Daten mit jenen anderer steuerpflichtiger Personen, beispielsweise im Rahmen eines RMS-Einsatzes, dazu führen, dass Kontrollhandlungen gezielt und effizienzorientiert jeweils bei denjenigen steuerpflichtigen Personen vorgenommen werden, bei denen das Risiko der unwahrheitsgetreuen Steuererklärung am höchsten ist.¹²⁴ Mit anderen Worten soll vorhergesagt werden, welche Personen mit hoher Wahrscheinlichkeit die Steuererklärung unwahrheitsgemäss ausgefüllt haben.¹²⁵ Bei diesem automatisierten Vorgang handelt es sich um nichts anderes als um ein Profiling.¹²⁶

Gemäss Art. 5 lit. f revDSG ist Profiling «jede Art der automatisierten Bearbeitung von Personendaten, die darin besteht, dass diese Daten verwendet werden, um bestimmte persönliche Aspekte, die sich auf eine natürliche Person beziehen, zu bewerten, insbesondere um Aspekte bezüglich Arbeitsleistung, wirtschaftlicher Lage, Gesundheit, persönlicher Vorlieben, Interessen, Zuverlässigkeit, Verhalten, Aufenthaltsort oder Ortswechsel dieser natürlichen Person zu analysieren oder vorherzusagen». Die im Gesetz genannten Merkmale sind nur beispielhaft.¹²⁷

Mithin ist Profiling der Vorgang der Bewertung bestimmter persönlichkeitsrelevanter Attribute einer Person durch vollautomatisierte Auswertung der betroffenen Personendaten.¹²⁸ Kein Profiling liegt vor, wenn eine Datensammlung besteht, sie jedoch nicht automatisiert ausgewertet wird.¹²⁹ Die Bearbeitung bedarf ferner einer bestimmten Wertung (subjektive Komponente), z. B. einer Prognose; die objektive Sachverhaltsfeststellung durch Datenerhebung ist kein Profiling.¹³⁰

¹²⁴ Um dies genauer beurteilen zu können, müsste allerdings die genaue Funktionsweise des RMS bekannt sein, siehe Ehrke-Rabel, Profiling (Fn. 64), S. 54 f.

¹²⁵ Angelehnt an Heuberger Götsch Olivier, Der Wert von Daten aus juristischer Sicht am Beispiel des Profiling, in: Fasel Daniel/Meier Andreas (Hrsg.), Big Data, Wiesbaden 2016, S. 83 ff., S. 86. Zum Risikobegriff aus der Verwaltungstätigkeit Wanner Markus/Zuber-Hagen Tom, Risikomanagement in den öffentlichen Verwaltungen von Kantonen und Gemeinden, AJP 2011, S. 901 ff., S. 903 f.

¹²⁶ Die folgenden Ausführungen stützen sich auf die (künftige) Rechtslage auf Bundesebene und verweisen nur punktuell auf die kantonalen Rechtsgrundlagen. Zur Anwendbarkeit des Bundes- oder kantonalen Datenschutzrechts siehe vorne [III.2](#).

¹²⁷ Botschaft DSG 2017 (Fn. 19), S. 7022.

¹²⁸ Botschaft DSG 2017 (Fn. 19), S. 7022. Vertieft zum Begriff des Profiling Bühlmann Lukas/Schüepf Michael, Begriff und Rechtsfolgen des Profilings im nDSG und der DSGVO, Jusletter 12. September 2022, N 69 ff.

¹²⁹ Botschaft DSG 2017 (Fn. 19), S. 7022.

¹³⁰ Rosenthal David, Das neue Datenschutzgesetz, Jusletter 16. November 2020, N 24.

Der Fokus liegt beim Profiling nicht allein auf dem Ergebnis, sondern vielmehr auf dem gesamten Bewertungsprozess.¹³¹ Ziel und Zweck der Analyse ist die Vorhersage von Verhaltensweisen, Fähigkeiten oder Bedürfnissen einer bestimmten Person.¹³² Beim steuerrechtlichen Überprüfungsvorgang kann es – je nachdem, welche Art von RMS zugrunde gelegt wird – um die Bewertung, Analyse und Vorhersage persönlicher Aspekte der steuerpflichtigen Person – nämlich deren Arbeitsleistung, wirtschaftliche Lage, Verhalten¹³³ und auch weitere persönlichkeitsrelevante Aspekte¹³⁴ wie Ehrlichkeit oder Zuverlässigkeit gehen. Findet dieser Vorgang vollautomatisiert statt, handelt es sich um Profiling i. S. d. revDSG. Profiling kann ferner als Grundlage für den Erlass automatisierter Einzelentscheidungen i. S. v. Art. 21 revDSG dienen.¹³⁵ Werden durch die Erhebung der steuerrechtlich relevanten Daten automatisiert wesentliche Aspekte der Persönlichkeit bewertet und die betroffene steuerpflichtige Person schliesslich vollautomatisiert veranlagt, basiert diese Entscheidung auf Profiling. Die beiden Begriffe müssen aber nicht zwingend miteinander in Verbindung stehen.¹³⁶

Es bleibt festzuhalten, dass nicht jedes RMS ein Profiling darstellt. Basiert ein RMS etwa schlicht darauf, dass Angaben der steuerpflichtigen Person von Angaben anderer Steuerpflichtiger abweichen, wenn es mithin einfach darum geht, «Ausreisser» zu finden, dann handelt es sich nicht um Profiling.

Staatliches Handeln setzt das Beachten der Grundrechte voraus (Art. 35 Abs. 1 BV). Im Kontext von Profiling ist insbesondere das Grundrecht auf informationelle Selbstbestimmung (Art. 13 Abs. 2 BV) zu wahren.¹³⁷ Der verfas-

¹³¹ Vertieft Heuberger (Fn. 118), N 56.

¹³² Botschaft DSG 2017 (Fn. 19), S. 7022.

¹³³ Siehe die Definition in Art. 5 lit. f revDSG.

¹³⁴ Zum Begriff der Persönlichkeit siehe Meier Philippe, *Droit des personnes*, 2. Aufl., Genf/Zürich/Basel 2021, S. 303 ff.; Riemer Hans Michael, *Personenrecht des ZGB*, 2. Aufl., Bern 2002, S. 131 ff.

¹³⁵ Ein Beispiel führt Braun Binder Nadja, *Als Verfügungen gelten Anordnungen der Maschine im Einzelfall... – Dystopie oder künftiger Verwaltungsalltag?*, ZSR 2020, S. 253 ff. (zit. Braun Binder, *Dystopie*), S. 269, auf.

¹³⁶ Zur Abgrenzung siehe Bühlmann/Schüepp (Fn. 128), N 130 ff.

¹³⁷ Rypka Thomas, *Das Grundrecht auf Privatsphäre in Bezug auf Steuerdaten*, Diplomarbeit der Universität Graz, 2020, S. 26 ff., online aufgerufen am 17. Februar 2023 unter: <https://unipub.uni-graz.at/obvugrhs/download/pdf/4825100?originalFilename=true>.

Zum Grundrechtsgehalt siehe Schweizer Rainer, *Kommentierung von Art. 13 BV N 74*, in: Ehrenzeller Bernhard/Schweizer Rainer J./Schindler Benjamin/Vallender Klaus A., *Die Schweizerische Bundesverfassung*, St. Galler Kommentar, 3. Aufl., Zürich 2014; BSK BV-Dig-gelmann (Fn. 38), Art. 13 N 33.

sungsrechtliche Anknüpfungspunkt liegt dabei auf der Schwere der möglichen Grundrechtseinschränkung, die durch die Datenbearbeitung erfolgen kann. Schwerwiegende Einschränkungen müssen in einem Gesetz im formellen Sinn statuiert sein (Art. 36 Abs. 1 BV).¹³⁸ Zudem muss die Rechtsgrundlage genügend bestimmt sein; die betroffene Person muss erkennen und voraussehen können, inwiefern ein Grundrechtseingriff erfolgt.¹³⁹

Das Datenschutzrecht konkretisiert das Grundrecht auf informationelle Selbstbestimmung und stellt erhöhte Anforderungen auf, wenn Behörden Profiling betreiben. Dies betrifft insbesondere die gesetzliche Grundlage und die Informationspflichten. Das revDSG statuiert, dass Profiling durch ein Bundesorgan einer Grundlage in einem Gesetz im formellen Sinn (Art. 34 Abs. 2 lit. b revDSG) bedarf. Für Bundesorgane unbeachtlich ist, ob es sich um ein «Profiling mit hohem Risiko»¹⁴⁰ gemäss Art. 5 lit. g revDSG handelt, oder um ein «gewöhnliches» Profiling gemäss Art. 5 lit. f revDSG.¹⁴¹ 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 revDSG) oder der Bearbeitungszweck für die Grundrechte der betroffenen Person keine besonderen Risiken birgt (Art. 34 Abs. 3 lit. b revDSG). Die Ausnahme der Unentbehrlichkeit wird im Steuervollzug wohl schwierig zu beweisen sein,¹⁴² zumal dieser bisher ohne Profiling erfolgte. Die zweite Ausnahme wird sodann ebenfalls nicht greifen, da das Bundesgericht bereits das Erstellen eines Persönlichkeitsprofils als schwerwiegenden Eingriff in die Grundrechte bezeichnete, da wesentliche Aspekte der Persönlichkeit zusammengestellt und beurteilt werden.¹⁴³

Im Zuge der Revision des Datenschutzgesetzes auf Bundesebene werden mehrere Gesetze um den Begriff des Profilings ergänzt. Beispielsweise sieht das revidierte Zollgesetz (revZG¹⁴⁴) nun vor, dass die eidgenössische Zollverwal-

¹³⁸ Zum Begriff des Gesetzes im formellen Sinn siehe BSK BV-Wytenbach/Wyss (Fn. 38), Art. 164 N 3.

¹³⁹ BSK BV-Epiney (Fn. 38), Art. 36 N 38.

¹⁴⁰ Dieses liegt vor, wenn es «ein hohes Risiko für die Persönlichkeit oder die Grundrechte der betroffenen Person mit sich bringt, indem es zu einer Verknüpfung von Daten führt, die eine Beurteilung wesentlicher Aspekte der Persönlichkeit einer natürlichen Person erlaubt» (Art. 5 lit. g revDSG).

¹⁴¹ Zum Begriff siehe den Beginn von [IV.4](#).

¹⁴² Roth Simon, Das Profiling im neuen Datenschutzrecht, SZW 2021, S. 34 ff., S. 38, generell kritisch zum Beweis der Unentbehrlichkeit von Profiling durch Bundesorgane.

¹⁴³ BGE 146 I 11 E. 3.2; 143 I 253 E. 3.2 und 4.9; 129 I 232 E. 4.3.2.

¹⁴⁴ Zollgesetz vom 18. März 2005 (SR 631.0), die revidierten Bestimmungen treten am 1.09.2023 in Kraft.

tung zur Erfüllung ihrer Aufgaben – der Erhebung von Personendaten zur Veranlagung nach Zollgesetz – Profiling (mit hohem Risiko) betreiben kann (Art. 110 Abs. 2 revZG). Dasselbe ist für das revidierte Mehrwertsteuergesetz (revMWSTG¹⁴⁵) vorgesehen: Die Eidgenössische Steuerverwaltung ist zum Profiling (mit hohem Risiko) befugt, wenn es der Erfüllung der gesetzlich vorgesehenen Aufgaben dient (Art. 76 Abs. 3 revMWSTG).

Die Kantone sind in ihrer Gesetzgebung im Datenschutzrecht unterschiedlich weit fortgeschritten. Es ist aber zu erwarten, dass sie mit Inkrafttreten des revDSG ihre Gesetze ebenfalls anpassen.¹⁴⁶ So wurde z. B. im Kanton Basel-Stadt im Zuge der Revision des IDG¹⁴⁷ eine ähnliche Definition des Profilings wie auf Bundesebene eingefügt.¹⁴⁸ Gemäss § 3 Abs. 7 rev IDG-BS ist Profiling jede automatisierte Auswertung von Informationen, um wesentliche persönliche Merkmale zu analysieren oder Entwicklungen, insbesondere bezüglich Arbeitsleistung, wirtschaftliche Lage, Gesundheit, persönliche Vorlieben, Interessen, Zuverlässigkeit, Verhalten, Aufenthaltsort oder Ortswechsel, vorherzusagen.¹⁴⁹

Unabhängig davon, ob die kantonalen Datenschutzgesetze das Profiling regeln, ist aus grundrechtlichen Erwägungen eine spezifische formell-gesetzliche Grundlage notwendig. Denn die Zusammenstellung und Beurteilung wesentlicher Aspekte der Persönlichkeit hat das Bundesgericht unter dem Begriff des «Persönlichkeitsprofils» als schwerwiegenden Grundrechtseingriff gewertet.¹⁵⁰ Zwar ging es in der bundesgerichtlichen Rechtsprechung um Persön-

¹⁴⁵ Bundesgesetz über die Mehrwertsteuer vom 12. Juni 2009 (SR 641.20), die revidierten Bestimmungen treten am 1.09.2023 in Kraft.

¹⁴⁶ Siehe Leitfaden der Konferenz der Kantonsregierungen «EU-Datenschutzreform/Moderernisierung der Europarats-Konvention 108: Anpassungsbedarf bei den kantonalen (Informations- und) Datenschutzgesetzen», N 3.8, online aufgerufen am 17. Februar 2023 unter: <https://www.dsb.bs.ch/datenschutz/privatim-und-kdk-leitfaden.html>.

¹⁴⁷ Siehe Ratschlag 21.1239.01 des Regierungsrates des Kantons Basel-Stadt zu einer Änderung des Gesetzes über die Information und den Datenschutz vom 9. Juni 2010 (Informations- und Datenschutzgesetz, IDG) und weiterer Gesetze (Anpassung an die europäischen Datenschutzreformen und weitere Anpassungen, zit. Ratschlag BS 2021).

¹⁴⁸ Die im Ratschlag BS 2021 (Fn. 147) enthaltene Fassung wurde im Zuge der Beratungen angepasst, siehe Bericht der Justiz-, Sicherheits- und Sportkommission 21.1239.02 vom 14. September 2022, S. 4 und 7.

¹⁴⁹ Kantonsblatt Basel-Stadt Nr. 85 vom 26.10.2022; die Referendumsfrist ist abgelaufen, das Inkrafttreten ist zum Zeitpunkt des Verfassens dieses Beitrages noch nicht beschlossen.

¹⁵⁰ BGE 146 I 11 E. 3.2; 143 I 253 E. 3.2 und 4.9; 129 I 232 E. 4.3.2.

lichkeitsprofile nach Art. 3 lit. d Datenschutzgesetz (DSG)¹⁵¹. Da das Profiling gemäss revDSG den bisherigen Begriff des Persönlichkeitsprofils ausdrücklich ablösen soll¹⁵², bietet sich eine analoge Anwendung der Rechtsprechung auf das Profiling an.

Die Anforderungen an die konkrete Ausgestaltung einer gesetzlichen Grundlage zum Einsatz von Profiling können nicht pauschal formuliert werden. Im Veranlagungsverfahren wird sie sich stark am Zweck der Datenbearbeitung orientieren müssen, um insbesondere auch die Anforderungen an die Normdichte einzuhalten. Die Durchführung eines Profilings müsste eindeutig ersichtlich sein,¹⁵³ ebenso wie die Tatsache, dass die Daten zum Zwecke des Abgleichs mit anderen Daten erhoben werden.

V. Fazit

Das internationale Umfeld zeigt, dass sich Steuerbehörden neuer Technologien bedienen, um die Steuererklärungen auf Plausibilität hin zu prüfen. Grundsätzlich ist dies auch in der Schweiz denkbar. Aus verfassungs- sowie datenschutzrechtlicher Sicht ist dies bedenklich, zumal insbesondere der Einsatz von Risikomanagementsystemen zu einem Profiling im Sinne des revidierten Datenschutzgesetzes führen kann und dieser entsprechender gesetzlicher Grundlagen bedarf.

¹⁵¹ Bundesgesetz über den Datenschutz vom 19. Juni 1992 (SR 235.1).

¹⁵² Botschaft DSG 2017 (Fn. 19), S. 6971.

¹⁵³ Bühlmann/Schüepf (Fn. 128), N 154.

Vertrauensschutz bei fehlerhaften Behördenauskünften durch Chatbots

Nadja Braun Binder / Liliane Obrecht / Grace Wittmer

Erstabdruck in: iusNet 30.05.2024

Inhalt

I.	Einleitung	188
II.	Grundlagen und Einsatzbereiche	189
1.	Eingrenzung	190
2.	Funktionsweise	190
2.1.	Regelbasierte Chatbots	191
2.2.	Datenbasierte Chatbots	192
a)	Grundsätze	192
b)	Large Language Models	193
III.	Vertrauensschutz bei Chatbot-Auskünften	195
1.	Eignung der Auskunft zur Begründung von Vertrauen	195
1.1.	Inhaltliche Bestimmtheit der Vertrauensgrundlage	195
1.2.	Vertrauenswürdigkeit der Auskunft	197
1.3.	Fehleranfälligkeit der datenbasierten Chatbot-Auskunft	199
a)	Bei datenbasierten Chatbots im Allgemeinen	199
b)	Zur Fehleranfälligkeit von LLMs	201
2.	Zuständigkeit der Behörde	203
2.1.	Grundsatz	203
2.2.	Herausforderungen bei datenbasierten Chatbots	204
3.	Vorbehaltlosigkeit der Auskunft	205
3.1.	Grundsatz	205
3.2.	Pauschale Vorbehalte	207
3.3.	Vorbehalte bei individuellen Auskünften	208
4.	Weitere Anforderungen	209
IV.	Fazit	210

I. Einleitung

Seit Joseph Weizenbaum in den 1960er Jahren «Eliza»¹ als virtuelle Psychotherapeutin programmierte,² wird an Chatbots als Interaktionskanal zwischen Menschen und Maschinen gearbeitet. Lange Zeit lieferten Chatbots aber ungenügende Resultate oder wirkten zu mechanisch.³ Erst Ende des 20. Jahrhunderts ermöglichten leistungsfähigere Computer mit höherer Speicher- und Rechenkapazität die weitere Entwicklung und Implementierung von Chatbots.⁴ Durch die Veröffentlichung von ChatGPT⁵ im November 2022 haben Chatbots einen weiteren Aufschwung erlebt.

Auch Schweizer Behörden sehen vermehrt Chatbots zur Informationsvermittlung und Auskunftserteilung vor. So setzt bspw. das Einwohneramt des Kantons Basel-Stadt einen Chatbot ein, der u. a. zur An- oder Abmeldung im Kanton oder zum Wohnungswechsel Auskunft erteilt.⁶ Chatbots der Sozialversicherungsämter der Kantone Aargau und St. Gallen beantworten Fragen rund um den Anspruch auf Prämienverbilligung.⁷ Die Zivilstandsämter der Kantone Basel-Stadt und Luzern setzen Chatbots u. a. zur Bestellung von Geburtsurkunden oder Heiratscheinen ein.⁸ Auf Bundesebene unterstützt ein Chatbot bei der Recherche

¹ Siehe <https://www.masswerk.at/elizabot/>.

² Vibhor Sharma/Monika Goyal/Malik Drishti, An Intelligent Behaviour Shown by Chatbot System, International Journal of New Technology and Research, 2017, S. 52 ff., S. 52.

³ Alexander Hoffmann, Chatbots: Einführung in die Zukunft von Marketing, PR und CRM, Haar bei München 2019, S. 19.

⁴ Beate Bruns/Cäcilie Kowald, Praxisleitfaden Chatbots: Conversation Design für eine bessere User Experience, Wiesbaden 2023, S. 6 f. Ausführlich zur Geschichte von Chatbots Eleni Adamopoulou/Moussiades Lefteris, Chatbots: History, technology, and applications, Machine Learning with Applications 2/2020, Nr. 100006; Robert Dale, The return of the chatbots, Natural Language Engineering 2016, S. 811 ff., S. 814; Toni Stucki/Sara D'Onofrio/Edy Portmann, Chatbots gestalten mit Praxisbeispielen der Schweizerischen Post, Wiesbaden 2020, S. 2 ff, S. 35.

⁵ «GPT» steht für «Generative Pre-trained Transformer», auf Deutsch: «Generativer Transformer mit Vortraining», siehe dazu Marcel Waldvogel, Wie funktioniert eigentlich ChatGPT?, dnp vom 30.01.2023, siehe <https://dnp.ch/2023/01/30/wie-funktioniert-eigentlich-chatgpt/>.

⁶ Bevölkerungsdienste und Migration Kanton Basel-Stadt, siehe <https://www.bdm.bs.ch/Wohnen/ausweis-im-kreditkartenformat.html>.

⁷ Aargau: Chatbot «Maxi», siehe <https://www.sva-ag.ch/pv>. St. Gallen: SVA St. Gallen, Der IPV-Chatbot, siehe <https://www.svasg.ch/news/meldungen/ipv-chatbot.php>.

⁸ Basel-Stadt: Bevölkerungsdienste und Migration, siehe <https://www.bdm.bs.ch/Zivilstand.html>. Luzern: Regionales Zivilstandsamt, siehe <https://www.stadtluern.ch/politik-verwaltung/stadtverwaltung/dienstabteilungenbereiche/8193>.

im Bundesarchiv,⁹ der Chatbot «StatBot.swiss» des Bundesamtes für Statistik interagiert nach dem Frage-Antwort-Prinzip mit den offenen Daten der öffentlichen Verwaltung (Open Government Data, OGD) in der Schweiz,¹⁰ und die Eidgenössische Stiftungsaufsicht hat unlängst den Chatbot «Esi» lanciert, der Stiftungen und Revisionsorganen helfen soll, unkompliziert Antworten auf Fragen rund um das Schweizer Stiftungswesen zu finden.¹¹ Es ist zu erwarten, dass im Zuge der allgemeinen Verwaltungsdigitalisierung¹² Chatbots weiter an Bedeutung gewinnen werden. Damit stellt sich die zentrale Frage, wie die Rechtslage aussieht, wenn Chatbots fehlerhafte Auskünfte erteilen.

Vor diesem Hintergrund werden im Folgenden zunächst die massgebenden Unterschiede zwischen regelbasierten und datenbasierten Chatbots sowie deren potenzielle Einsatzbereiche erörtert (II). Sodann wird untersucht, ob das im Zusammenhang mit Behördenauskünften durch Menschen entwickelte Konzept des Vertrauensschutzes auch im Kontext von behördlichen Chatbot-Auskünften greift (III). Nicht Gegenstand dieses Aufsatzes sind Chatbots von privaten Anbietern, die mit öffentlich zugänglichen Daten von Behörden trainiert wurden.¹³ Der Beitrag schliesst mit einem Fazit (IV).

II. Grundlagen und Einsatzbereiche

Im Folgenden wird zuerst der Chatbot-Begriff für die Zwecke des vorliegenden Beitrages eingegrenzt (II.1). Sodann werden die verschiedenen Funktionsweisen erläutert (II.2).

⁹ Schweizerisches Bundesarchiv, Beratung, siehe <https://www.bar.admin.ch/bar/de/home/recherche/suchen/beratung.html>. Zum Chatbot siehe <https://chatbot.bar.smartive.cloud/>.

¹⁰ Bundesamt für Statistik, StatBot.swiss: mit offenen Daten diskutieren, siehe <https://www.bfs.admin.ch/bfs/de/home/dscc/blog/2023-02-statbot.html>.

¹¹ Eidgenössische Stiftungsaufsicht, Frag Esi, siehe <https://www.fragesi.ch/>.

¹² Exemplarisch sei auf das am 1. Januar 2024 in Kraft getretene Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben (EMBA) vom 17. März 2023 (SR 172.019) hingewiesen. Siehe zur allgemeinen Verwaltungsdigitalisierung auch Christian R. Ulbrich/Bruno S. Frey, Automated Democracy – Die Neuverteilung von Macht und Einfluss im digitalen Staat, Freiburg i. B. 2024.

¹³ Z. B. «ZüriCityGPT», der mit Daten der Domain stadt-zuerich.ch trainiert wurde, der aber nicht mit der Stadt Zürich assoziiert ist, siehe <https://zuericitygpt.ch/>. Siehe auch für Genf <https://geneva.gpt.liip.ch/> oder Lausanne <https://lausanne.gpt.liip.ch/>. Eine Übersicht liefert Christian Stocker, Swiss Government Chatbots, siehe <https://github.com/liip/SwissGovernmentChatbots>.

1. Eingrenzung

Bei einem Chatbot handelt es sich um ein Onlinedialogsystem, das in einer von Menschen gesprochenen Sprache (natürliche Sprache) und in Echtzeit rund um die Uhr mit gleichbleibender Qualität kommunizieren kann.¹⁴ Der Begriff setzt sich aus den englischen Worten *to chat* (sich unterhalten) und *bot* (Roboter) zusammen.¹⁵ Der Chatbot gibt Antworten auf spezifische Fragen oder leitet eine entsprechende Aktion ein.

Der vorliegende Beitrag beschränkt sich auf Chatbots, die zur (schriftlichen)¹⁶ Auskunftserteilung eingesetzt werden. Die interaktive Kommunikation erfolgt dabei in einer digitalen Umgebung – etwa einer Gemeindefachseite – und zielt darauf ab, den Austausch zwischen staatlichen Behörden und Privaten¹⁷ nachzubilden.¹⁸

2. Funktionsweise

Einem Chatbot können verschiedene Modelle zugrunde liegen. Für die vorliegenden Zwecke wird zwischen regelbasierten ([II.2.1](#)) und datenbasierten ([II.2.2](#)) Chatbots unterschieden.¹⁹

¹⁴ Bruns/Kowald (Fn. 4), S. 1 f.

¹⁵ Fabia Stöcklin, Robot Recruiting: Die Haftung des Arbeitgebers für Persönlichkeitsverletzungen beim Einsatz von Künstlicher Intelligenz im Bewerbungsprozess, *sui generis* 2023, S. 2, siehe <https://sui-generis.ch/article/view/4045>.

¹⁶ Sog. Voicebots können demgegenüber die Spracheingabe in Textbausteine umwandeln und die Textantwort des Bots als gesprochene Sprache wiedergeben, siehe dazu Nadja Braun Binder et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4 (zit. Braun Binder et al., Studie 2021), S. 57, siehe <https://www.zh.ch/de/news-uebersicht/medienmitteilungen/2021/04/kuenstliche-intelligenz-in-der-verwaltung-braucht-klare-leitlini.html>.

¹⁷ Terminologie angelehnt an Ulrich Häfelin/Georg Müller/Felix Uhlmann, Allgemeines Verwaltungsrecht, 8. A., Zürich 2020, Rz. 624 ff. Der Vertrauensschutz gilt für natürliche und juristische Personen gleichermaßen, weshalb im vorliegenden Beitrag der umfassende Begriff «Private» verwendet wird.

¹⁸ Zu den verschiedenen digitalen Umgebungen siehe [III.1.2](#).

¹⁹ Terminologie angelehnt an Stucki/D'Onofrio/Portmann (Fn. 4), S. 7; siehe auch Ulbrich/Frey (Fn. 12), S. 200 ff.

2.1. Regelbasierte Chatbots

Bei Chatbots, die auf regelbasierten Algorithmen beruhen, wird der Input (die Frage) anhand vorgegebener Regeln einer vordefinierten Antwort zugeordnet.²⁰ Der Chatbot kann z. B. auf einem Entscheidungsbaum basieren und zur Beantwortung der Anfrage einem Flussdiagramm folgen.²¹ Dabei kann die Benutzerin oder der Benutzer die Frage z. B. aus einer Liste auswählen.²² Ein solcher Ansatz liegt etwa dem Chatbot der Steuerverwaltung des Kantons Schwyz zugrunde. Dieser zeigt verschiedene Optionen an und leitet als Reaktion auf die Auswahl durch einen Fragenkatalog, um die entsprechende Antwort anzeigen zu können.²³ Die Benutzerin oder der Benutzer kann die Frage jedoch auch in Textform verfassen. Der Chatbot basiert in diesen Fällen i. d. R. auf der Schlüsselworterkennung.²⁴ Jeder hinterlegten Antwort sind entsprechende Schlüsselwörter zugeordnet.²⁵ Wird in der Frage keines der Schlüsselwörter genannt, stellt der Chatbot vorprogrammierte Rückfragen, um die fehlenden Worte zu ermitteln.²⁶ So setzt etwa das Migrationsamt des Kantons Zürich einen Chatbot ein, der die Worteingabe einem bestimmten Bereich zuordnet und daraufhin Rückfragen stellen kann.²⁷

Bei regelbasierten Chatbots sind die Antwortmöglichkeiten endlich. Es werden keine neuen Textantworten generiert.²⁸ Bei der Programmierung müssen folglich möglichst viele Dialoge antizipiert werden.²⁹ Dies führt gleichzeitig dazu, dass bei regelbasierten Chatbots nachvollziehbar ist, wie die Antwort auf eine bestimmte Frage zustande kommt.

²⁰ Wenn (if) A dann (then) B, siehe Christoph Beierle/Gabriele Kern-Isberner, Methoden wissensbasierter Systeme, 3. A., Wiesbaden 2006, S. 71 ff.

²¹ Julia Lehmann, Der Chatbot-Guide, in: Stefan Detscher (Hrsg.), Digitales Management und Marketing, Wiesbaden 2021, S. 306 ff., S. 309 ff.

²² Stucki/D'Onofrio/Portmann (Fn. 4), S. 7.

²³ Steuern Kanton Schwyz, siehe <https://www.sz.ch/behoerden/verwaltung/finanzdepartement/steuerverwaltung.html/8756-8758-8802-10332-10354>.

²⁴ Lehmann (Fn. 21), S. 310 f.

²⁵ Stucki/D'Onofrio/Portmann (Fn. 4), S. 7.

²⁶ Torsten Deutsch, Fake KI – welcher Algorithmus ist wirklich intelligent?, BigData Insider 18.08.2021, siehe <https://www.bigdata-insider.de/fake-ki-welcher-algorithmus-ist-wirklich-intelligent-a-f05efc1465bcb4f662e1341bf23394bc/>.

²⁷ Migrationsamt des Kantons Zürich, siehe <https://www.zh.ch/de/sicherheitsdirektion/migrationsamt.html>.

²⁸ Lehmann (Fn. 21), S. 311 ff.

²⁹ Stucki/D'Onofrio/Portmann (Fn. 4), S. 7.

2.2. Datenbasierte Chatbots

Im Unterschied zu den traditionellen Programmieransätzen können auch datenbasierte maschinell lernende Algorithmen – häufig mit dem Begriff «Künstliche Intelligenz» (KI) umschrieben³⁰ – eingesetzt werden.³¹ Im Folgenden wird zuerst auf einige Grundsätze betreffend datenbasierte Chatbots eingegangen ([II.2.2.a](#)), um anschliessend die Besonderheiten von Large Language Models ([II.2.2.b](#)) zu erläutern.

a) Grundsätze

Bei datenbasierten Chatbots ist das Ziel die Entwicklung eines Modells, das nach eigenen Regeln Wahrscheinlichkeiten berechnet, um von einem Input zu einem Output zu gelangen. Dieses Modell wird sodann auf unbekannte Datensätze angewendet.³² Solche Modelle beruhen somit auf wahrscheinlichkeitsbasierten Zusammenhängen und nicht auf einer Datenverarbeitung nach vorgegebenen Regeln.³³ Dabei können ganz unterschiedliche Methoden maschinellen Lernens, etwa künstliche neuronale Netze³⁴, zur Anwendung gelangen.³⁵ Das Fundament bilden i. d. R. Trainingsdaten, d. h. genügend grosse Datensätze (sog. *Big Data*³⁶), die auf verallgemeinerbare Muster und Korrelationen hin analysiert werden.³⁷ Der Vorteil ist, dass datenbasierte Chatbots dadurch weitreichendere Dialogelemente verarbeiten können als regelbasierte Chatbots.³⁸ Jedoch führen die Wahr-

³⁰ Siehe etwa Torsten Deutsch/Ralf Hülsmann, Fake-KI: Nicht jeder Algorithmus ist intelligent, *Digitale Welt* 2021, S. 16 ff.

³¹ Stucki/D'Onofrio/Portmann (Fn. 4), S. 7.

³² Hartmut Ernst/Jochen Schmidt/Gerd Beneken, *Grundkurs Informatik*, 7. A., Wiesbaden 2020, S. 801 ff.; Wolfgang Ertel, *Grundkurs Künstliche Intelligenz*, 4. A., Wiesbaden 2016, S. 265 ff.

³³ Ernst/Schmidt/Beneken (Fn. 32), S. 801 ff.

³⁴ Künstliche neuronale Netze sind durch biologische Prozesse inspiriert und werden insb. bei der Erkennung von Mustern und komplexen Zusammenhängen in Daten eingesetzt. Zum Begriff z. B. Michael Oettinger, *Data Science*, 2. A., Hamburg 2020, S. 119 ff.

³⁵ Siehe Oettinger (Fn. 34), S. 112 ff., S. 119 ff.

³⁶ Big Data enthält i. d. R. unstrukturierte oder semi-strukturierte Daten, die anhand der Dimensionen «volume» (enorme Datenmenge), «velocity» (hohe Generierungs- und Verarbeitungsgeschwindigkeit) und «variety» (Vielfältigkeit der Datentypen und -quellen) beschrieben werden, siehe nur Ronald Bachmann, *Big Data – Fluch oder Segen?*, Heidelberg/München 2014, S. 17 ff.

³⁷ Ernst/Schmidt/Beneken (Fn. 32), S. 801 ff.

³⁸ Peter Ringeisen/Andrea Bertolosi-Lehr/Labinot Demaj, *Automatisierung und Digitalisierung in der öffentlichen Verwaltung: digitale Verwaltungsassistenten als neue Schnittstelle*

scheinlichkeitsberechnungen auch dazu, dass ein datenbasierter Chatbot ausgehend von identischen Inputs nicht immer dieselben Outputs generiert. Die Antworten sind somit nicht abschliessend vorhersehbar, d. h. der Weg, auf dem diese Antworten im konkreten Einzelfall zustande kommen, ist für Menschen kognitiv nicht nachvollziehbar.³⁹ Zudem sind die Antworten durch die Wahrscheinlichkeitsberechnungen auch entsprechend fehleranfällig. Immerhin sind sie in ihren Antwortmöglichkeiten dahingehend limitiert, als sie nicht über die Informationen in den Trainingsdaten, die ihnen zur Verfügung gestellt werden, hinausgehen können.⁴⁰

Bei dem von der Stadt St. Gallen eingesetzten Chatbot «Gallus» handelt es sich um einen solchen datenbasierten Chatbot.⁴¹ Gallus verfolgt das Ziel, einen einfacheren Zugang zu gewissen Basisdiensten der Verwaltung wie etwa der Abfrage von Öffnungszeiten oder der Parkplatzsituation zu ermöglichen.⁴² Dabei werden ständig neue Daten in das Chatbot-System eingepflegt, um dessen Anwendungsbereich zu vergrössern.⁴³ Auch die konsularische Direktion des Eidgenössischen Departements für auswärtige Angelegenheiten (EDA) setzte im Rahmen eines Pilotprojekts einen datenbasierten Chatbot ein. Dieser informierte (in französischer Sprache) über die konsularischen Dienstleistungen.⁴⁴

b) *Large Language Models*

Seit der Veröffentlichung von ChatGPT⁴⁵ werden sog. *Large Language Models* (LLMs)⁴⁶ im Kontext von Chatbots als besonders vielversprechend betrachtet.⁴⁷

zwischen Bevölkerung und Gemeinwesen, in: Swiss Yearbook of Administrative Sciences 2018, S. 51 ff., S. 55; Braun Binder et al., Studie 2021 (Fn. 16), S. 57; Stucki/D'Onofrio/Portmann (Fn. 4), S. 7.

³⁹ Ulbrich/Frey (Fn. 12), S. 201 f.

⁴⁰ Stucki/D'Onofrio/Portmann (Fn. 4), S. 7. Zur Manipulation der Datenbasis siehe [III.1.3.a](#).

⁴¹ Braun Binder et al., Studie KI 2021 (Fn. 16), S. 27.

⁴² Chatbot der Stadtverwaltung St. Gallen, siehe https://www.stadt.sg.ch/news/stsg_home/2020/12/chatbot.html.

⁴³ Künftig soll der Bot auch selbstständig Prozesse abwickeln können, siehe Braun Binder et al., Studie 2021 (Fn. 16), S. 27.

⁴⁴ Geschäftsstelle CNAI, Projektsteckbriefe Kompetenznetzwerk CNAI, Version 4.0 vom 11.04.2023, S. 8 f., siehe https://cna1.swiss/wp-content/uploads/2023/04/CNAI_Projekte_D_4_0.pdf.

⁴⁵ Siehe [I](#).

⁴⁶ Auf Deutsch: Grosse Computermodelle zur Sprachverarbeitung.

⁴⁷ Siehe etwa Philipp Hacker/Andreas Engel/Marco Mauer, Regulating ChatGPT and other Large Generative AI Models, Working Paper May 12 2023, <https://arxiv.org/abs/2302>.

Dabei handelt es sich um Systeme, die auf maschinell lernenden Algorithmen basieren, deren Ziel die Verarbeitung von natürlicher Sprache ist (*Natural Language Processing*, NLP).⁴⁸ Im Einsatz sind sog. *Transformer*, eine Weiterentwicklung von künstlichen neuronalen Netzen.⁴⁹ Dabei wird eine sehr grosse Menge an Trainingsdaten (hier in Textformat) in das Netz eingegeben.⁵⁰ Diese Daten werden zunächst numerisch abgebildet, wobei das System in einem Lernprozess eine Art Logik in der menschlichen Sprache herausarbeitet.⁵¹ Basierend auf dieser Logik wird selbständig ein (Sprach-)Modell entwickelt.⁵² So basiert der Chatbot *Esi* der Eidgenössischen Stiftungsaufsicht (ESA) auf dem LLM ChatGPT 3.5 Turbo.⁵³ *Esi* stützt sich bei der Auskunftserteilung auf Informationen und Dokumente, die von der ESA vorgegeben wurden.⁵⁴

Bei Chatbots, die LLMs nutzen, handelt es sich ebenfalls um datenbasierte Chatbots. Allerdings können LLMs selbständig Texte erzeugen (sie sind *generativ*), ohne auf zusätzliche Regeln oder Wissensquellen zurückgreifen zu müssen.⁵⁵ Dadurch sind sie noch flexibler einsetzbar und nicht auf einen bestimmten Kontext beschränkt.⁵⁶

[02337](#), S. 2 ff. Weitere Beispiele stellen etwa Google Bard oder GPT NEO dar. LLMs werden allerdings bereits seit Jahren entwickelt, es ist die Menge der zur Verfügung stehenden Trainingsdaten, die den Systemen zum Durchbruch verholfen hat, siehe Eva A. M. van Dis et al., ChatGPT: five priorities for research, *Nature* 2023, S. 224 ff., S. 224.

⁴⁸ Stucki/D'Onofrio/Portmann (Fn. 4), S. 3.

⁴⁹ Steffen Albrecht, ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen, Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, Berlin 2023, S. 20, siehe <https://www.bundes-tag.de/resource/blob/944148/30b0896f6e49908155fcd01d77f57922/20-18-109-Hintergrundpapier-data.pdf>.

⁵⁰ Albrecht (Fn. 49), S. 20.

⁵¹ Albrecht (Fn. 49), S. 20.

⁵² Siehe dazu Alexander Wehde, Regulierung von Large Language Models in DSA und AIA-E, *MMR* 2023, Nr. 455171.

⁵³ Siehe [I](#).

⁵⁴ Dazu zählen die Webseite der ESA (<https://www.edi.admin.ch/edi/de/home/fachstellen/eidgenoessische-stiftungsaufsicht/beratung/esi.html>), Kommentierungen aus dem Onlinekommentar zum Stiftungsrecht (<https://onlinekommentar.ch/de/kommentare>) und der Swiss Foundation Code (<https://www.swissfoundations.ch/publikationen/swiss-foundation-code-2021/>).

⁵⁵ Albrecht (Fn. 49), S. 22.

⁵⁶ Siehe zu generativen algorithmischen (KI-)Systemen etwa Natali Helberger/Nicholas Diakopoulos, ChatGPT and the AI Act, *Internet Policy Review* 16.02.2023, siehe <https://policyreview.info/es-say/chatgpt-and-ai-act>; Luciano Floridi, AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models, *Philosophy & Technology* 2023, S. 14 ff.

III. Vertrauensschutz bei Chatbot-Auskünften

Der Vertrauensschutz – als Teilaspekt von Treu und Glauben (Art. 9 BV)⁵⁷ – bedeutet, dass Private unter bestimmten Voraussetzungen – Eignung der Auskunft zur Begründung von Vertrauen (III.1); Zuständigkeit der Behörde (III.2); Vorbehaltlosigkeit der Auskunft (III.3); weitere Anforderungen (III.4) – in ihrem durch behördliches Verhalten geweckten Vertrauen geschützt werden.⁵⁸

1. Eignung der Auskunft zur Begründung von Vertrauen

Der Vertrauensschutz greift nur, wenn die erteilte Auskunft geeignet ist, Vertrauen zu begründen. Diesbezüglich sind im Zusammenhang mit behördlichen Chatbot-Auskünften insbesondere die inhaltliche Bestimmtheit der Vertrauensgrundlage (III.1.1), die Vertrauenswürdigkeit der Auskunft (III.1.2) und die Fehleranfälligkeit der Chatbot-Auskünfte (III.1.3) von Interesse.

1.1. Inhaltliche Bestimmtheit der Vertrauensgrundlage

Eine Vertrauensgrundlage liegt vor, wenn das Verhalten eines staatlichen Organs bei der oder dem Privaten eine bestimmte Erwartung auslöst.⁵⁹ Hierfür muss die Auskunft inhaltlich genügend bestimmt sein.⁶⁰ Der Vertrauensschutz bedingt grundsätzlich, dass eine individuell-konkrete Grundlage vorliegt; generell-abstrakte Auskünfte werden in der Regel nicht als genügende Vertrauensgrundlage erachtet.⁶¹ Das bedeutet, dass sich die Antwort auf einen konkreten Sachverhalt beziehen muss, zu dem die anfragende Person um Auskunft ersucht (individuell-konkrete Auskunft).

⁵⁷ 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. A., Basel 2016, Rz. 1996.

⁵⁸ Häfelin/Müller/Uhlmann (Fn. 17), Rz. 624, Rz. 627 ff.; Rhinow/Schefer/Uebersax (Fn. 57), Rz. 1998 ff.; Pierre Tschannen/Ulrich Zimmerli/Markus Müller, Allgemeines Verwaltungsrecht, 4. A., Bern 2014, § 22 N 478.

⁵⁹ BGE 132 II 218 E. 6.1, 6.2 S. 228 ff.; BGE 132 II 21 E. 2.2 S. 25 f.

⁶⁰ Häfelin/Müller/Uhlmann (Fn. 17), Rz. 668.

⁶¹ Siehe Tschannen/Zimmerli/Müller (Fn. 58), § 22 N 15; Pierre Tschannen/Markus Müller/Markus Kern, Allgemeines Verwaltungsrecht, 5. A., Bern 2022, Rz. 489; BSK BV-Tschentscher, Art. 9 N 16, in: Bernhard Waldmann/Eva Maria Belser/Astrid Epiney (Hrsg.), Bundesverfassung, Basler Kommentar, Basel 2015 (zit. BSK BV-Verfasser:in); BGE 143 V 341 E. 5.2.1 S. 346 f.; BGE 143 V 95 E. 3.6.2 S. 103; BGE 131 II 627 E. 6.1 S. 363 f.; BGE 125 I 267 E. 4b S. 274 f.; kritisch Beatrice Weber-Dürler, Neuere Entwicklung des Vertrauensschutzes, ZBl 2002 (zit. Neuere Entwicklung), S. 281 ff.; Giovanni Biaggini, BV Kommentar, 2. A., Zürich 2017, Art. 9 N 15; Chiariello (Fn. 57), S. 42.

Indes wird das Erfordernis einer individuell-konkreten Auskunft teilweise relativiert. So werden generell-abstrakte Aussagen oder Informationen auf einem amtlichen Formular, in einem Merkblatt oder auf einer amtlichen Internetseite unter gewissen Voraussetzungen als Vertrauensgrundlage gewertet. Dies gilt gemäss einem (restriktiveren) Teil der Lehre, wenn die Informationen so genau zum Einzelfall passen, dass sie als konkrete Auskunft eingestuft werden können.⁶² Ein anderer Teil der Lehre ist diesbezüglich grosszügiger und stellt darauf ab, ob eine Information allgemein darauf abzielt, rechtliche Klarheit zu schaffen.⁶³ Die Information muss sich dabei weder auf einen konkreten Sachverhalt beziehen noch an einen individuellen Adressaten oder eine individuelle Adressatin richten.⁶⁴ Dies wäre z. B. bei einer allgemeinen Rechtsauskunft in einem vorgedruckten Formular⁶⁵ oder einem allgemeinen Merkblatt⁶⁶ der Fall. Gemäss CHIARIELLO kommt es auf den Gesamtzusammenhang an, in den sich eine solche Information einfügt.⁶⁷

Zusammenfassend lässt sich festhalten, dass individuell-konkrete Auskünfte hinreichend bestimmt sind, um eine Vertrauensgrundlage zu bilden. Bei generell-abstrakten Informationen herrscht in der Lehre hingegen bislang keine Einigkeit, wann diese ausnahmsweise ebenfalls als Vertrauensgrundlage dienen können.⁶⁸ Auch in der Rechtsprechung zeigt sich gerade mit Blick auf Behördeninformationen auf amtlichen Webseiten (noch) keine klare Linie.⁶⁹

⁶² Tschannen/Zimmerli/Müller (Fn. 58), § 22 N 15; Tschannen/Müller/Kern (Fn. 61), Rz. 489; BSK BV-Tschentscher (Fn. 61) Art. 9 N 18; Pierre Moor/Alexandre Flückiger/Vincent Meritet, *Droit administratif*. Vol. I, Les fondements, 3. A., Bern 2012, S. 924.

⁶³ Häfelin/Müller/Uhlmann (Fn. 17) Rz. 669; siehe auch BGE 129 II 125 E. 5.6 S. 140 f.; Weber-Dürler, *Neuere Entwicklungen* (Fn. 61), S. 295.

⁶⁴ Häfelin/Müller/Uhlmann (Fn. 17) Rz. 669.

⁶⁵ Eine allgemeine Auskunft stellt hingegen keine hinreichende Vertrauensgrundlage dar, wenn sie etwa Dritten erteilt wird und durch diese weiter übermittelt wird, siehe Häfelin/Müller/Uhlmann (Fn. 17) Rz. 669.

⁶⁶ Weber-Dürler, *Neuere Entwicklungen* (Fn. 61), S. 295, forderte bereits 2002, Merkblätter vermehrt als taugliche Vertrauensgrundlage anzuerkennen, dafür betreffend die weiteren Voraussetzungen des Vertrauensschutzes einen strengeren Massstab zu setzen.

⁶⁷ Chiariello (Fn. 57), S. 43.

⁶⁸ Felix Uhlmann/Jasna Stojanovic, *Vertrauen im Finanzmarktrecht aus öffentlich-rechtlicher Sicht*, SZW 2017, S. 732 ff., S. 736; Pascal Hauenstein/Jonas Klemperer, *Vertrauensschutz bei behördlichem Webauftritt*, AJP 2023, S. 871 ff., S. 873 f.

⁶⁹ Die Tauglichkeit als Vertrauensgrundlage bejahend insb. die kantonale Rechtsprechung, etwa Verwaltungsgericht Graubünden, R 09 105, 16.3.2010, E. 1-3a; Versicherungsgericht Solothurn, VSBES.2013.159, 30.1.2014, E. 2.3.2., ferner auch BVGer, A-2575/2017, 14.1.2019, E. 3.4.1.; aufgrund der konkreten Umstände des Einzelfalles verneinend dagegen BGer,

Mit Blick auf Chatbots kann aufgrund des Gesagten zunächst einmal festgehalten werden, dass individuell-konkrete Chatbot-Auskünfte durchaus als hinreichend bestimmte Vertrauensgrundlage i. S. d. Vertrauensschutzes eingestuft werden können.⁷⁰ Gerade datenbasierte Chatbots können zur Beantwortung konkreter Fragen eingesetzt werden.⁷¹ Erhalten Private auf diese Art eine individualisierte Antwort, kann diese denselben Bestimmtheitsgrad aufweisen wie eine menschliche Behördenauskunft. Es spricht demnach nichts dagegen, diese als hinreichend bestimmte Vertrauensgrundlage zu qualifizieren.

Sodann ist hinsichtlich generell-abstrakter Chatbot-Auskünfte auf die in der Lehre unterschiedlich weit gehende ausnahmsweise Anerkennung allgemeiner Informationen als hinreichend bestimmte Vertrauensgrundlage hinzuweisen. Als Minimum ist diesbezüglich zu fordern, dass analog zur restriktiven Lehrmeinung mindestens jene allgemeinen Chatbot-Auskünfte als Vertrauensgrundlage zu qualifizieren sind, die so genau zum Einzelfall passen, dass sie als konkrete Auskunft eingestuft werden können. Somit könnte z. B. auch der Output (die Antwort) eines regelbasierten Chatbots, der anhand bestimmter Schlagwörter vorprogrammierten Abläufen folgt oder verschiedene Antworten in einem Flussdiagramm zur Auswahl anbietet,⁷² u. U. eine hinreichend bestimmte Vertrauensgrundlage bilden.

1.2. Vertrauenswürdigkeit der Auskunft

Eine weitere Voraussetzung für die Beurteilung, ob sich die Auskunft zur Begründung von Vertrauen eignet, bildet die Vertrauenswürdigkeit der erteilten Auskunft. Bei einem behördlichen Chatbot-Einsatz muss sich die Vertrauenswürdigkeit der Auskunft nach vorliegender Ansicht bereits durch die Tatsache ergeben, dass der Chatbot durch die Behörde eigens zur Information eingesetzt wird. Grundsätzlich scheint daher die Vertrauenswürdigkeit der Auskunft durch den Chatbot-Einsatz im Vergleich zur menschlichen Auskunft nicht beeinträchtigt zu werden. Hingegen ist im Folgenden zu diskutieren, ob das digitale Umfeld, in dem der Chatbot eingesetzt wird, die Vertrauenswürdigkeit der Auskunft tangiert.

1C_703/2020, 13.10.2022, E. 6.3; BGer, 2C_407/2012, 23.11.2012, E. 3.3; BVGer, B-5869/2014, 30.10.2017, E. 3.1.3; Verwaltungsgericht St. Gallen, B 2022/128, 16.1.2023, E. 4; Kantonsgericht Basel-Land, 715 16 233/49, 15.2.2017, E. 4.

⁷⁰ Braun Binder et al., Studie 2021 (Fn. 16), S. 61. Dem Grundsatz nach auch: Tschannen/Zimmerli/Müller (Fn. 58), § 22 N 15; BSK BV-Tschentscher (Fn. 61), Art. 9 N 16.

⁷¹ Siehe [II.2.2.](#)

⁷² Siehe [II.2.1.](#)

Öffentliche Organe können Chatbots – soweit technisch möglich – auf allen digitalen Umgebungen, mittels derer sie mit den Privaten in Kontakt treten, einsetzen. So ist nicht nur die Kommunikation über formelle Schnittstellen wie bspw. eine Gemeindefachseite möglich,⁷³ sondern auch der niederschwellige Kontakt über Messenger-Dienste sozialer Medien, die nicht von der Behörde bewirtschaftet werden, denkbar.⁷⁴ Es stellt sich die Frage, ob der formelle oder informelle Charakter der Plattform einen Einfluss auf die Vertrauenswürdigkeit der in diesem Umfeld erteilten Chatbot-Auskunft haben kann.

Der Vorteil, einen Chatbot über einen Messenger-Dienst zur Verfügung zu stellen, besteht darin, dass die weit verbreitete Nutzung der sozialen Medien auch mehr Private an das Chatbot-Angebot heranführt. So können andere, personalaufwändige Informationskanäle – bspw. Telefondienst und E-Mailverkehr – entlastet werden. Allerdings stellen die entsprechenden Plattformen keinen offiziellen staatlichen Berührungspunkt dar. Wie es der Name bereits sagt, sind die sozialen Medien eher für den unverbindlichen sozialen Austausch und nicht für die Erlangung behördlicher Informationen intendiert. Als kollaborative Plattformen werden sie primär von privaten Nutzerinnen und Nutzern gespiesen und sind kaum kontrolliert. Es könnte argumentiert werden, dass Informationen über die sozialen Medien von den Privaten nicht in guten Treuen als vertrauenswürdig eingestuft werden dürfen.⁷⁵

Daraus könnte gefolgert werden, dass sich die Auskünfte von Chatbots, die über informelle Plattformen abgegeben werden, nicht als Vertrauensgrundlage eig-

⁷³ Z. B. der Chatbot des Kantons Schwyz, der auf der Kantonswebseite Informationen zum Thema Steuern erteilt, siehe <https://www.sz.ch/behoerden/verwaltung/finanzdepartement/steuerverwaltung.html/8756-8758-8802-10332-10354>.

⁷⁴ Siehe z. B. die Schweizerische Post, die im Bereich PostNetz im Jahre 2017 eine Studie zum Thema *Smart Assistent* startete. Als Ergebnis der Studie wurde ein Chatbot-Prototyp über den Kanal Facebook-Messenger zur Verfügung gestellt, der den Kundinnen und Kunden Auskünfte zu den Filialen, Öffnungszeiten und zum Lieferstatus von Sendungen erteilen kann (Stucki/D'Onofrio/Portmann (Fn. 4), S. 13). Auch der Chatbot der SVA St. Gallen wurde ebenfalls über Facebook-Messenger eingesetzt, um Fragen rund um die individuelle Prämienverbilligung zu beantworten, siehe Braun Binder et al., Studie 2021 (Fn. 16), S. 61.

⁷⁵ Abgesehen von der Frage nach den Auswirkungen auf den Vertrauensschutz würde das Angebot der staatlichen Dienstleistungen auch Fragen des anwendbaren Rechts in Bezug auf den Datenschutz aufwerfen. Der Anspruch auf eine datenschutzrechtskonforme Ausgestaltung der Chatbots könnte allenfalls in Konflikt mit den individuellen Datenschutzrichtlinien und Nutzungsbedingungen der entsprechenden Plattformen treten, Ringeisen/Bertolosi-Lehr/Demaj (Fn. 38), S. 61; Roger Plattner, *Digitales Verwaltungshandeln*, Diss. Zürich 2021, Rz. 518.

nen.⁷⁶ Hingegen wäre es verfehlt, die Behörden durch die Verwendung dieser Plattformen von ihrer Pflicht zu qualitativ hochwertiger Information zu entbinden. Es sind die Behörden selbst, die sich zur Erschließung dieser neuen Kommunikationsplattform entscheiden und davon profitieren. Die Qualität und somit auch die Vertrauenswürdigkeit von Behördeninformationen können nicht von der gewählten Kommunikationsplattform abhängig gemacht werden.⁷⁷ Insgesamt erscheint daher die Vertrauenswürdigkeit der behördlichen Auskunft bei einem Chatbot-Einsatz durch das digitale Umfeld nicht beeinflusst.

1.3. Fehleranfälligkeit der datenbasierten Chatbot-Auskunft

Das Vorliegen einer Vertrauensgrundlage kann im Kontext eines Chatbot-Einsatzes auch durch die Fehleranfälligkeit von datenbasierten Chatbot-Auskünften in Frage gestellt werden. Dies betrifft alle datenbasierten Chatbots ([III.1.3.a](#)), ganz besonders aber solche, die LLMs nutzen ([III.1.3.b](#)).

a) Bei datenbasierten Chatbots im Allgemeinen

Dass datenbasierte Chatbots fehlerhafte Auskünfte geben können, wurde bereits weiter oben dargelegt.⁷⁸ Ein Ansatz, um dieser Problematik zu begegnen, ist u. a. die sorgfältige Vorauswahl überprüfter Trainingsdaten zur Entwicklung des Modells.⁷⁹ Es handelt sich mithin um eine *ex ante* Kontrolle der verwendeten Daten, wodurch die Antworten des Chatbots gewissermassen antizipierend eingehengt werden können. Die Entwicklung eines verlässlichen und vertrauenswürdigen Chatbots hängt somit zwangsläufig mit der Qualität der Trainingsdaten zusammen.⁸⁰ Die Korrektheit der Daten ist dabei von zentraler

⁷⁶ Im Grundsatz auch Ringeisen/Bertolosi-Lehr/Demaj (Fn. 38), S. 61.

⁷⁷ Anderer Ansicht Hauenstein/Klemperer (Fn. 68), S. 877, die für eine differenzierte Lösung plädieren.

⁷⁸ Siehe [II.2.2.a](#).

⁷⁹ Norwegian Consumer Council, Ghost in the machine – Addressing the consumer harms of generative AI, Juni 2023, S. 22, siehe <https://storage02.forbrukerradet.no/media/2023/06/generative-ai-rapport-2023.pdf>; Hacker/Engel/Mauer (Fn. 47), S. 2. Es kommt das verstärkte Problem hinzu, dass unweigerlich immer häufiger Datensätze zum Training verwendet werden, die durch andere «KI»-Systeme erstellt wurden, was sich ebenfalls negativ auf die Qualität der Chatbot-Antworten auswirkt, siehe Erich Moechel, Künftige KI-Modelle potenziell von Demenz bedroht, heise vom 07.07.2023, siehe https://www.heise.de/news/Kuenftige-KI-Modelle-potenziell-von-Demenz-bedroht-9209900.html?wt_mc=sm.red.ho.mastodon.mastodon.md.beitraege.md.beitraege.

⁸⁰ Hacker/Engel/Mauer (Fn. 47), S. 2. Es wird grundsätzlich von verschiedenen Seiten gefordert, Qualitätsstandards für algorithmische Systeme gesetzlich festzulegen und – etwa in

Bedeutung.⁸¹ Eine hohe Datenqualität setzt grundsätzlich voraus, dass der Ursprung der Daten bekannt ist und diese überprüfbar sind.⁸²

Überdies können in datenbasierten Chatbots auch Blockaden und Content-Filter programmiert werden, die diese in ihren Antworten limitieren.⁸³ Derartige programmierte Blockaden können jedoch überwunden werden.⁸⁴ So können Chatbots dahingehend manipuliert werden, dass sie Antworten geben, die bei der Programmierung eigentlich ausgeschlossen wurden.⁸⁵ Auf diese Weise können unerwünschte oder gar gesetzeswidrige Antworten produziert werden, obwohl dies ursprünglich ausgeschlossen werden sollte. *Jailbreaking* wird in diesem Zusammenhang als Begriff genutzt, um das Phänomen zu beschreiben, dass eine Anwendung so manipuliert wird, dass die Benutzenden Zugang zu allen Funktionen und Informationen erhalten, auch wenn diese eigentlich nicht zugänglich sein sollten.⁸⁶ Ein prominentes Beispiel unter vielen ist der Microsoft Chatbot Tay.⁸⁷ Dieser musste deaktiviert werden, da er durch gezielte Fragen – in nur wenigen Stunden – dazu gebracht wurde, rassistische Äusserungen zu verbreiten.⁸⁸ Bei datenbasierten Chatbots besteht demnach

einem Register – die Daten öffentlich zugänglich zu machen. Siehe dazu Jeremy Stevens, Datenqualität bei algorithmischen Entscheidungen, CR 2020, S. 73 ff., S. 74; Algorithm-Watch, Ein KI-Transparenzregister für die öffentliche Verwaltung, Konzeptpapier 2023, siehe https://algorithmwatch.org/de/wp-content/uploads/2023/03/Konzept_KI-Transparenzregister_AlgorithmWatch_2023.pdf.

⁸¹ Rechtlich festgelegte Qualitätsmessstandards wird es für Daten wohl aber nicht allzu bald geben, siehe Liliane Obrecht/David Starchl, Tagungsbericht zur Veranstaltung Open Government Data und Verpflichtung zur Gewährung des Datenzugangs an den Staat, Jusletter 24.10.2022, Rz. 15.

⁸² Philipp Hacker/Lauri Wessel, KI-Trainingsdaten nach dem Verordnungsentwurf für Künstliche Intelligenz, in: Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz/Frauke Rostalski (Hrsg.), Künstliche Intelligenz, Tübingen 2022, S. 54 ff., S. 54 f.

⁸³ Albrecht (Fn. 49), S. 69.

⁸⁴ Kai Greshake et al., Not what you've signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection, eprint arXiv 2023, S. 2, siehe <https://ui.adsabs.harvard.edu/abs/2023arXiv230212173G/abstract>.

⁸⁵ Sog. «Jailbreaks», etwa die Aufforderung zu Gewalt (siehe dazu die Ausführungen sogleich). Für Beispiele siehe Fangzhao Wu et al., Defending ChatGPT against jailbreak attack via self-reminder, nature machine intelligence 2023, S. 1486 ff., siehe <https://www.nature.com/articles/s42256-023-00765-8>.

⁸⁶ Greshake et al. (Fn. 84), S. 2.

⁸⁷ Für aktuelle Beispiele von Jailbreaks siehe <https://www.jailbreakchat.com/>.

⁸⁸ Siehe etwa James Vincent, Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day, The Verge, 24.03.2016, siehe <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>; siehe auch Plattner (Fn. 75), Rz. 519.

das Risiko, dass diese – sei es durch einen allgemeinen Hacker-Angriff, sei es durch gezielte Manipulation, z. B. durch die anfragende Person – vermehrt falsche oder aus Sicht der Behörde ungewollte Auskünfte erteilen.

b) Zur Fehleranfälligkeit von LLMs

Die Fehleranfälligkeit von Chatbots wird insbesondere seit der Lancierung von ChatGPT konstatiert. Dabei wird von *Halluzination* gesprochen.⁸⁹ Als Halluzination wird in diesem Kontext eine linguistisch plausible Antwort eines LLMs bezeichnet, die allerdings inhaltlich nicht vollständig korrekt ist.⁹⁰ Dies kann insbesondere dann auftreten, wenn der Chatbot mit Inhalten (Sachstand) konfrontiert wird, bei der die Datenbasis unzureichend ist.⁹¹ Des Weiteren ist zu beachten, dass sich die Halluzinationen verstärken können, je stärker die Materie fachdomänenspezifisch ist.⁹² Es handelt sich mit anderen Worten um Falschaussagen durch den Chatbot, die durch Laien kaum erkannt werden können.⁹³

⁸⁹ Zihao Li, The Dark Side of ChatGPT: Legal and Ethical Challenges from Stochastic Parrots and Hallucination, arXiv 2023, siehe <https://arxiv.org/abs/2304.14347>; Hussam Alkaissi/Samy I. McFarlane, Artificial Hallucinations in ChatGPT: Implications in Scientific Writing, Cureus 2023, siehe <https://assets.cureus.com/uploads/editorial/pdf/138667/20230219-28928-6kcyip.pdf>; Gernot Beutel/Eline Geerits/Jan T. Kielstein, Artificial hallucination: GPT on LSD, Critical Care 2023, siehe <https://link.springer.com/article/10.1186/s13054-023-04425-6>.

⁹⁰ Zur Begriffsbestimmung und dem Stand der Forschung siehe Ziwei Ji et al., Survey of Hallucination in Natural Language Generation, ACM Computing Surveys 2022, S. 1 ff. Im Unterschied zum aus der Humanpsychologie verwendeten Term der Halluzination, der auf eine menschliche Sinneswahrnehmung, die nicht mit der Aussenwelt in Verbindung gebracht werden kann, abstellt. Zur Kritik an der Verwendung psychologischer Begriffe im Kontext von neuen Technologien Henry Shevlin/Marta Halina, Apply rich psychological terms in AI with care, nature machine intelligence 2019, S. 165 ff.

⁹¹ van Dis et al. (Fn. 47), S. 224. Es handelt sich dabei um eine Frage des Designs des Chatbots, siehe Carl Bergstrom, Your chatbot is not «hallucinating», post.news 2023, siehe <https://post.news/@/user/2Lr2DCy9lQz0pbzrVwrtgBD6l8l> [Der Beitrag ist unter dem angegebenen Link nicht mehr verfügbar. Ein leicht aktualisierter Beitrag findet sich hier: <https://undark.org/2023/04/06/chatgpt-isnt-hallucinating-its-bullshitting/>, abgerufen am 23.04.2025 – Anm. der Herausgeberschaft].

⁹² Albrecht (Fn. 49), S. 40 f. Beispiele aus dem juristischen Bereich zeigen Matthew Dahl et al., Hallucinating Law: Legal Mistakes with Large Language Models are Pervasive, Stanford University 2024, siehe <https://hai.stanford.edu/news/hallucinating-law-legal-mistakes-large-language-models-are-pervasive>.

⁹³ Pegah Maham/Sabrina Küspert, Governing General Purpose AI, Stiftung Neue Verantwortung, Juli 2023, S. 21, siehe https://www.stiftung-nv.de/sites/default/files/snv_go-

Überdies greifen LLMs häufig auf ungeprüfte Daten aus unsicheren Quellen zu. Dadurch können Daten (versteckt) manipuliert und unerwünschte Anweisungen für die LLMs hinterlegt werden (*Indirect Prompt Injections*).⁹⁴ Hinzu kommt, dass Studien zeigen, dass die Ergebnisse von LLMs über relativ kurze Zeiträume in ihrer Kohärenz abnehmen können.⁹⁵ Solche Chatbots müssen daher kontinuierlich evaluiert und bewertet werden.⁹⁶

Um den Schwächen und insbesondere der Halluzination von LLMs zu begegnen, setzen Chatbot-Anbieterinnen und Anbieter nun vermehrt auf *Retrieval Augmented Generation* (RAG). Dabei handelt es sich um eine Technik zur Verarbeitung natürlicher Sprache (NLP⁹⁷), die die Stärken von regelbasierten und datenbasierten generativen Modellen kombiniert.⁹⁸ Anders als bei der Generierung von Information (*generative*⁹⁹) geht es bei der Beschaffung von Informationen (*retrieval*) darum, relevante vorhandene Objekte aus einem grossen Datenpool, der kontinuierlich um Dokumente ergänzt werden kann, zu finden.¹⁰⁰ Die Kernidee ist dabei, dass eine Anfrage auf einem lokalen Textkorpus durch den *Retriever* bearbeitet wird und dieser mittels semantischer Suche die relevanten Textsegmente bestimmt.¹⁰¹ Das LLM (der *Generator*) wird daraufhin

[verning_general_purpose_ai_pdf.pdf](#). Beispiele (allerdings spezifisch im Kontext von ChatGPT) finden sich bei Bernard Marr, *ChatGPT: What Are Hallucinations And Why Are They A Problem For AI Systems?*, Bernard Marr & Co. 22. 03. 2023, siehe <https://bernardmarr.com/chatgpt-what-are-hallucinations-and-why-are-they-a-problem-for-ai-systems/>.

⁹⁴ Maham/Küspert (Fn. 93), S. 26. So sieht dies etwa das Bundesamt für Sicherheit in der Informationstechnik (BSI) in Deutschland als eine der grössten Herausforderungen beim Einsatz von LLMs, siehe dazu BSI, *Indirect Prompt Injections – Intrinsische Schwachstelle in anwendungsintegrierten KI-Sprachmodellen, Version 1.0 vom 18.07.2023*, siehe https://www.bsi.bund.de/SharedDocs/Cybersicherheitswarnungen/DE/2023/2023-249034-1032.pdf?__blob=publicationFile&v=3.

⁹⁵ Lingjiao Chen/Matei Zaharia/James Zou, *How is ChatGPT's behavior changing over time?*, arXiv Nr. 2307.09009, <https://arxiv.org/pdf/2307.09009.pdf>.

⁹⁶ Chen/Zaharia/Zou (Fn. 95), S. 14 f.

⁹⁷ Siehe bereits oben [II.2.2.b](#).

⁹⁸ Ivan Ilin, *Advanced RAG Techniques: an Illustrated Overview, Towards AI*, 17.12.2023, siehe <https://pub.towardsai.net/advanced-rag-techniques-an-illustrated-overview-04d193d8fec6>.

⁹⁹ Siehe [II.2.2.b](#).

¹⁰⁰ Penghao Zhao et al., *Retrieval-Augmented Generation for AI-Generated Content: A Survey*, arxiv 29.02.2024, S. 1, siehe <https://arxiv.org/abs/2402.19473>. Die Autoren sprechen von einem «adaptable data repository» (Deutsch für «anpassungsfähigen Datenspeicher»).

¹⁰¹ Zhao et al. (Fn. 100), S. 1.

aufgefordert, die Anfrage mittels der ihm zur Verfügung gestellten Informationen zu beantworten.¹⁰² Sowohl die ursprüngliche Anfrage der Nutzerin oder des Nutzers als auch der vom *Retriever* abgerufene Kontext werden in die an das LLM gesendete Eingabe eingefügt.¹⁰³ Das LLM wird u. a. dazu verwendet, die Anfrage umzuformulieren und die Antwort kohärent und strukturiert darzustellen.¹⁰⁴ Inwieweit die Fehlerhaftigkeit des LLM dadurch tatsächlich reduziert wird – insbesondere, ob sie gänzlich ausgeschlossen werden kann – ist bislang allerdings unklar.¹⁰⁵

2. Zuständigkeit der Behörde

Der Vertrauensschutz greift grundsätzlich für Auskünfte, die in den Zuständigkeitsbereich derjenigen Behörde fallen, die eine Auskunft erteilt.¹⁰⁶ Die Frage der Zuständigkeit wird im Folgenden zuerst allgemein im Hinblick auf den Chatbot-Einsatz betrachtet ([III.2.1](#)) und sodann hinsichtlich datenbasierter Chatbots diskutiert ([III.2.2](#)).

2.1. Grundsatz

Nach Lehre und Rechtsprechung genügt es, dass die Privaten in guten Treuen von der Zuständigkeit der Behörde ausgehen.¹⁰⁷ Der gute Glaube ist zerstört, wenn die Unzuständigkeit der Behörde offensichtlich, d. h. für die fragende Person eindeutig erkennbar war.¹⁰⁸ Die Frage nach der Offensichtlichkeit beurteilt sich jeweils im Einzelfall und richtet sich neben objektiven Gesichtspunkten auch nach subjektiven Kriterien, insbesondere dem persönlichen Hintergrund der fragenden Person.¹⁰⁹

¹⁰² Zhao et al. (Fn. 100), S. 3.

¹⁰³ Michael McTear/Marina Ashurkina, *Transforming Conversational AI: Exploring the Power of Large Language Models in Interactive Conversational Agents*, Berkeley 2024, S. 66; Zhao et al. (Fn. 100), S. 3.

¹⁰⁴ Zhao et al. (Fn. 100), S. 3 f.

¹⁰⁵ Zhao et al. (Fn. 100), S. 1, betonen, dass die Herausforderungen, die der Einsatz von LLMs mit sich bringt, durch RAG gemildert, wenn auch nicht vollständig gelöst werden können.

¹⁰⁶ Häfelin/Müller/Uhlmann (Fn. 17), Rz. 676 ff.; Chiariello (Fn. 57), S. 49.

¹⁰⁷ BSK BV-Tschentscher (Fn. 61), Art. 9 N 16; Häfelin/Müller/Uhlmann (Fn. 17), Rz. 677; Chiariello (Fn. 57), S. 49; siehe BGE 101 Ia 92 E. 3b S. 100.

¹⁰⁸ Chiariello (Fn. 57), S. 49; Häfelin/Müller/Uhlmann (Fn. 17), Rz. 864.

¹⁰⁹ Häfelin/Müller/Uhlmann (Fn. 17), Rz. 676 f.; siehe Chiariello (Fn. 57), S. 50. So sind regelmäßig höhere Anforderungen an fachkundige Personen – etwa mit juristischem Hintergrund – zu stellen, siehe auch Hauenstein/Klemperer (Fn. 68), S. 878.

Diese Grundsätze ändern sich bei einem Chatbot-Einsatz nicht. Chatbots, die in einem bestimmten Bereich zur Fragebeantwortung eingesetzt werden, müssen technisch so ausgestaltet sein, dass sie nur diejenigen Themen behandeln können, die in ihrem Zuständigkeitsbereich liegen.¹¹⁰ Somit sollen sich die Privaten zumindest in guten Treuen darauf verlassen können, dass der Chatbot nur Antworten aus diesem Zuständigkeitsbereich erteilt.¹¹¹ Werden Fragen ausserhalb des Zuständigkeitsbereichs gestellt, oder bestehen generell Unklarheiten, so muss der Chatbot auf menschliche Beratung verweisen.¹¹² Insbesondere bei regelbasierten Chatbots sollte dies unproblematisch sein, da sie vorgegebenen Abläufen folgen und nicht über ihre Programmierung hinausgehen.¹¹³ Der Zuständigkeitsbereich kann entsprechend mittels einer Kontrolle vor dem Einsatz des Systems eingeehtet werden. Bei datenbasierten Chatbots ist hingegen im Folgenden der erweiterte Antwortspielraum zu berücksichtigen.

2.2. Herausforderungen bei datenbasierten Chatbots

Bei datenbasierten Chatbots bestimmt grundsätzlich der Informationsgehalt der Trainingsdaten deren Antwortmöglichkeiten.¹¹⁴ Übersteigen die Trainingsdaten den Informationsgehalt des Zuständigkeitsbereichs einer Behörde, ist es demnach möglich, dass der (datenbasierte) Chatbot Auskünfte ausserhalb dieses Zuständigkeitsbereichs erteilt.¹¹⁵ Die Problematik verschärft sich beim Einsatz von LLMs. Ihre Stärke liegt gerade darin, linguistisch plausible Antworten zu liefern, auch wenn sie über keinen Zugang zu entsprechenden Wissensquellen verfügen.¹¹⁶ Wie die Antwort zustande gekommen ist, ist aufgrund

¹¹⁰ Bei regelbasierten Chatbots können Auskünfte ausserhalb des Zuständigkeitsbereiches bereits in der Programmierung verhindert werden, indem keine entsprechenden Informationsgrundlagen zur Verfügung gestellt werden. Bei Unklarheit soll auf menschliche Beratung verwiesen werden, Plattner (Fn. 75), Rz. 521. Bei datenbasierten Chatbots, insb. LLMs, kann versucht werden, durch sog. *fine tuning* entsprechende Auskünfte zu verhindern, siehe Daniel M. Ziegler et al., Fine-Tuning Language Models from Human Preferences, arxiv: 1909.08593, siehe <https://arxiv.org/abs/1909.08593>.

¹¹¹ Braun Binder et al., Studie 2021 (Fn. 16), S. 61; Plattner (Fn. 75), Rz. 521. Dies ist ein weiteres Argument, weshalb die aktuelle «Generation» generativer Chatbots nicht zur behördlichen Auskunftserteilung eingesetzt werden können soll.

¹¹² Plattner (Fn. 75), Rz. 521.

¹¹³ Siehe [II.2.1.](#)

¹¹⁴ Siehe [II.2.2.](#)

¹¹⁵ Siehe [II.2.2.a\).](#)

¹¹⁶ Siehe [II.2.2.b\).](#)

der technischen Natur von LLMs dabei nicht nachvollziehbar.¹¹⁷ Ob der Chatbot Antworten ausserhalb des Zuständigkeitsbereichs einer Behörde erteilen wird, kann deshalb nicht abschliessend vor dem Einsatz des Systems kontrolliert werden. Die Antworten können nur überprüft werden, nachdem sie generiert wurden. Diese Kontrolle kann von menschlichen Moderatorinnen und Moderatoren vorgenommen werden; auch auf maschinellem Lernen basierende Algorithmen können zu diesem Zweck eingesetzt werden.¹¹⁸

3. Vorbehaltlosigkeit der Auskunft

Schutzwürdiges Vertrauen können nur *vorbehaltlose* Auskünfte begründen.¹¹⁹ Sind die übrigen Voraussetzungen gegeben (siehe insbesondere [III.1](#) und [III.2](#)) und erteilt ein Chatbot die Auskunft ohne Vorbehalt, dann greift demnach der Vertrauensschutz. In der Praxis zeigt sich allerdings bisweilen das Bedürfnis, behördliche Chatbot-Auskünfte mit einem Vorbehalt zu versehen. Ob solche Vorbehalte angebracht werden dürfen, ist dabei fraglich. Eine Analogie zur Praxis bei menschlichen Behördenauskünften liefert keine zufriedenstellenden Anhaltspunkte ([III.3.1](#)). Es sind deshalb weitere Erwägungen anzustellen ([III.3.2–3.3](#)).

3.1. Grundsatz

Mit einem Vorbehalt bringt die Mitarbeiterin bzw. der Mitarbeiter einer Behörde zum Ausdruck, dass sie bzw. er sich nicht festlegen will.¹²⁰ Damit soll verhindert werden, dass aus spontanen Äusserungen unmittelbar Rechtswirkungen erwachsen. Dies mag etwa dann gerechtfertigt sein, wenn solche Äusserungen im Rahmen eines persönlichen Gesprächs, eines informellen Telefonats oder aus einem Interview hervorgehen. Die anfragende Person soll sich angesichts eines Vorbehalts nicht mehr auf den Vertrauensschutz berufen können.¹²¹ Voraussetzung ist, dass der Vorbehalt erkennbar ist.¹²² Im Zusam-

¹¹⁷ Siehe [II.2.2.b](#).

¹¹⁸ Siehe auch Hacker/Engel/Mauer (Fn. 47), S. 3.

¹¹⁹ Häfelin/Müller/Uhlmann (Fn. 17), Rz. 682; Tschannen/Müller/Kern (Fn. 61), Rz. 682.

¹²⁰ BGE 98 Ia 460 E. 2 S. 462 f.; Braun Binder et al., Studie 2021 (Fn. 16), S. 61.

¹²¹ Braun Binder et al., Studie 2021 (Fn. 16), S. 61; Häfelin/Müller/Uhlmann (Fn. 17), Rz. 670, Rz. 682; Uhlmann/Stojanovic (Fn. 68), S. 737.

¹²² Denkbar ist ein Hinweis auf der Plattform in der Nähe des Zugangs zum Chatbot. Der Chatbot könnte ebenfalls im Sinne einer Eröffnungsfloskel den Vorbehalt selbst in den Dialog einbringen. Weniger klar ist, ob die Erkennbarkeit gegeben ist, wenn der Vorbehalt nicht proaktiv bei der Benutzung des Chatbots eingeblendet wird, sondern auf der Plattform,

menhang mit Behördenauskünften sind Vorbehalte somit grundsätzlich anerkannt. Allerdings ist die Wirksamkeit dieses ursprünglich privatrechtlichen Motivs¹²³ bislang für Behördenauskünfte nicht abschliessend geklärt.¹²⁴ Auch gibt es in der Praxis keine Standards für die Ausgestaltung oder Platzierung eines Vorbehalts.¹²⁵

Die Situation zwischen einer spontanen Auskunftserteilung durch Behördenmitarbeitende ist jedoch nicht mit dem Einsatz von Chatbots vergleichbar. Stellt die Behörde einen Chatbot zur Verfügung, so ist dies eine bewusste Entscheidung. Die Konstellation einer spontanen Auskunft, einer Äusserung im Rahmen eines informellen Gesprächs oder Interviews kann so gar nicht erst entstehen. Vielmehr wird der Chatbot gezielt zur Beantwortung von Fragen eingesetzt. Die Frage-Antwort-Situation ist damit grundsätzlich vorhersehbar, auch wenn nicht alle Fragen an den Chatbot antizipiert werden können. Es ist zwar durchaus nachvollziehbar, dass eine Behörde einen Vorbehalt anbringen möchte, um sich im Falle falscher oder unzutreffender Antworten des Chatbots zu schützen. Ob dies gerechtfertigt werden kann, ist im Folgenden für das Anbringen von pauschalen ([III.3.2](#)) und individuellen ([III.3.3](#)) Vorbehalten zu diskutieren.

über die er zur Verfügung gestellt wird, gesucht werden muss (bspw. unter «Rechtliches», oder «Nutzungshinweise»), m. w. H. Hauenstein/Klemperer (Fn. 68), S. 878, insb. Fn. 72. Da aus öffentlich-rechtlicher Perspektive das konsensuelle Element des (privatrechtlichen) Zugangsprinzips nicht berücksichtigt werden muss (Robert George Briner, Haftung für Informationen auf Websites, sic! 2002, S. 231 ff, S. 239; Katja Favre, Sorgfaltspflichten bei der Datenübertragung, Diss. Zürich 2006, S. 239), genügt wohl die reine Möglichkeit der Kenntnisnahme nach den Umständen des Einzelfalles. Siehe dazu Häfelin/Müller/Uhlmann (Fn. 17), Rz. 682; CR Cst.-Dubey, Art. 9 N 82; in: Vincent Martenet/Jacques Dubey (Hrsg.), Commentaire Constitution fédérale, Basel 2021; Daniel Donauer/Fabian Loretan, Disclaimer – Haftungsausschluss bei Internetseiten, Jusletter 13.08.2018, Rz. 10.

¹²³ Briner (Fn. 122), S. 231 ff.

¹²⁴ Uhlmann/Stojanovic (Fn. 68), S. 736 f. Die Möglichkeit des Ausschlusses der Vertrauenshaftung mittels Vorbehalt bejahend, wenn auch als wenig zweckmässig erachtend Daniel Kettiger, Die Haftung des Staates für seine Geodaten, in: Verwaltungsorganisationsrecht – Staatshaftungsrecht – öffentliches Dienstrecht, Jahrbuch 2016/2017, Bern 2017, S. 103 ff., S. 118.

¹²⁵ Teilweise wird in gewissen Kantonen für bestimmte Bereiche ein Haftungsausschluss in generell-konkreten Regelungen im jeweiligen Verordnungsrecht statuiert, Kettiger (Fn. 124), S. 118 f.; Plattner (Fn. 75), Rz. 290. Gerade in Abgrenzung zum Privatrecht wird diskutiert, ob ein Vorbehalt *angenommen* werden muss, damit er im Verhältnis von Staat und Privaten zur Geltung kommt. Dankbar wäre ein Button, über den die Informationssuchenden bestätigen müssen, dass sie sich der Unverbindlichkeit der Auskunft bewusst sind und diese Prämisse akzeptieren. Siehe dazu Plattner (Fn. 75), Rz. 291 f.; Donauer/Loretan (Fn. 122), Rz. 10.

3.2. Pauschale Vorbehalte

Chatbots sollen gegenüber der traditionellen Behördenauskunft eine gleichwertige bzw. gleich verlässliche Alternative bieten. Dies gilt insbesondere, wenn ein Chatbot die traditionellen Kommunikationskanäle ganz oder teilweise ersetzen¹²⁶ und die Behörde damit entlasten¹²⁷ soll. Wird ein pauschaler Vorbehalt für alle Anfragen bzw. alle Antworten des Chatbots angebracht, so müssten Private, wenn sie eine verlässliche Auskunft erhalten wollen, einen Mehraufwand auf sich nehmen und weitere Quellen oder doch noch die Behörde konsultieren. Tatsächlich wäre es für jene Personen, die sich vorgängig mit dem Chatbot beschäftigen gegenüber jenen, die sich direkt an die Behörde wenden, schwieriger, an vertrauenswürdige Informationen zu kommen.¹²⁸ Dies würde im Widerspruch zum Grundgedanken der Entlastung der traditionellen Kommunikationskanäle stehen.¹²⁹ Hinzu kommt, dass mit einer Zunahme von Chatbots immer weniger Kapazitäten für den direkten Austausch mit den Behördenmitarbeitenden zur Verfügung stehen dürften.¹³⁰ Es ist daher fraglich, wie sinnvoll es überhaupt ist, einen «digitalen Verwaltungsassistenten»¹³¹ zur

¹²⁶ Für eine Entwicklung in diese Richtung spricht, dass bereits heute das Internet zunehmend zum primären Informationskanal der Behörden wird, siehe Uhlmann/Stojanovic (Fn. 68), S. 737; Braun Binder et al., Studie 2021 (Fn. 16), S. 61.

¹²⁷ Bzgl. Entlastung durch KI in der Verwaltung generell, Braun Binder et al., Studie 2021 (Fn. 16), S. 15 f. Bzgl. Entlastung speziell durch Chatbots, Ringeisen/Bertolosi-Lehr/Demaj (Fn. 38), S. 54, S. 59; Mario Martini, Big Data als Herausforderung für das Datenschutzrecht und den Persönlichkeitsschutz, in: Hermann Hill/Mario Martini/Edgar Wagner (Hrsg.), Die digitale Lebenswelt gestalten, Baden-Baden 2015, S. 97 ff.; Plattner (Fn. 75), Rz. 513; Tobias Hänni, Digitalisierung: Der Beamten-Bot, St. Galler Tagblatt vom 28. Februar 2018, siehe <https://www.tagblatt.ch/ostschweiz/digitalisierung-der-beamten-bot-im-kanton-stgal-len-gibt-ein-chat-roboter-auf-facebook-auskunft-ueber-die-praemienverbilligung-ld.1006812>; Frederic Härri, Entlastung der Mitarbeiter: «Maxi» unterstützt die Sozialversicherung bei Anfragen zu Prämien, Aargauer Zeitung vom 21.01.2020, siehe <https://www.aargauerzeitung.ch/aargau/kanton-aargau/entlastung-der-mitarbeiter-maxi-unterstutzt-die-sozialversicherung-bei-anfragen-zu-pramien-ld.1406057>.

¹²⁸ Dem Grundsatz nach auch Weber-Dürler, Vertrauensschutz im öffentlichen Recht, Habil. Basel/Frankfurt a. M. 1983; Hauenstein/Klemperer (Fn. 68), S. 878.

¹²⁹ Dem Grundsatz nach auch Braun Binder et al., Studie 2021 (Fn. 16), S. 60; Hauenstein/Klemperer (Fn. 68), S. 878.

¹³⁰ Braun Binder et al., Studie 2021 (Fn. 16), S. 61.

¹³¹ Ringeisen/Bertolosi-Lehr/Demaj (Fn. 38), S. 54, verstehen unter einem digitalen Verwaltungsassistenten einen Chatbot, der die Rolle einer konversationsbasierten Schnittstelle zwischen Menschen, Maschine und der öffentlichen Verwaltung übernimmt. Er soll für die Privaten den gesamten Koordinationsaufwand übernehmen, damit diese ungehindert auf verfügbare Daten und Services zugreifen können.

Verfügung zu stellen, an den sich die Fragesuchenden richten sollen, auf dessen Auskünfte sie sich aber nicht verlassen können.

Eine Rechtfertigung für das Anbringen von pauschalen Vorbehalten bezüglich der Korrektheit der erteilten Auskünfte ist weder bei regelbasierten noch bei datenbasierten Chatbots ersichtlich. Die öffentliche Verwaltung schuldet der Bevölkerung bereits bei ihrem allgemeinen Internetauftritt ein erhöhtes Sorgfaltsmass.¹³² Private sollen sich möglichst einfach und rasch über die Rechtslage informieren und sich auf die zur Verfügung gestellten Informationen verlassen können.¹³³ Dies gilt auch beim Einsatz von Chatbots.

Denkbar bzw. zulässig erscheint hingegen, im Zusammenhang mit einem Chatbot zu präzisieren, was dieser leisten kann und was nicht. So erscheint beim baselstädtischen Migrationsamt-Chatbot auf Nachfrage hin die Information, dass der Chatbot nur generelle Antworten liefern kann.¹³⁴ Für einzelfallrelevante Informationen wird auf die Massgeblichkeit des Gesetzestextes verwiesen.

3.3. Vorbehalte bei individuellen Auskünften

Datenbasierte Chatbots können viel differenzierter antworten als regelbasierte Chatbots, da sie weitreichendere Dialogelemente verarbeiten können.¹³⁵ Datenbasierte Chatbots sind daher nicht starr an eine vordefinierte Antwortphrase gebunden und können spezifischer auf den Einzelfall bezogene Einschätzungen abliefern.¹³⁶ Wie bei behördlichen Auskünften durch einen Menschen¹³⁷ spricht diese konkrete Natur der Chatbot-Auskunft eher für die Möglichkeit, einen Vorbehalt im Einzelfall anbringen zu können.¹³⁸ Überdies kann auch der Datenverarbeitungsprozess dazu führen, dass ein Vorbehalt bei

¹³² Häfelin/Müller/Uhlmann (Fn. 17), Rz. 670; Uhlmann/Stojanovic (Fn. 68), S. 737.

¹³³ Der Grundsatz von Treu und Glauben verpflichtet die Behörde zu vertrauenswürdigem Handeln, siehe Braun Binder et al., Studie 2021 (Fn. 16), S. 61.

¹³⁴ Siehe <https://www.bdm.bs.ch/Wohnen/ausweis-im-kreditkartenformat.html?navopen=search>. Nach demselben Vorgehen agiert der Chatbot des Handelsregisteramtes des Kantons St. Gallen, siehe <https://www.sg.ch/recht/handelsregister-notariate.html>; siehe auch Plattner (Fn. 75), Rz. 290; bzgl. allgemein publizierten Informationen Hauenstein/Klempeper (Fn. 68), S. 878 ff.

¹³⁵ Ringeisen/Bertolosi-Lehr/Demaj (Fn. 38), S. 55; Braun Binder et al., Studie 2021 (Fn. 16), S. 57; Stucki/D'Onofrio/Portmann (Fn. 4), S. 7.

¹³⁶ Siehe [II.2.2.](#)

¹³⁷ Siehe [III.3.1.](#)

¹³⁸ Siehe auch Plattner (Fn. 75), Rz. 522.

datenbasierten Chatbot-Auskünften eher zulässig sein könnte, als bei regelbasierten. Bereits die Tatsache, dass datenbasierte Chatbots wahrscheinlichkeitsbasiert Auskünfte erteilen und jeweils diejenige Antwort auf eine Frage gegeben wird, die aufgrund der Berechnungen als am wahrscheinlichsten bezeichnet wurde,¹³⁹ führt dazu, dass die Chatbot-Antworten unter Umständen nicht korrekt sind.

Bei komplexen rechtlichen Anfragen kann es zudem durch kommunikative Missverständnisse, unvollständige oder falsche Trainingsdaten oder gar Missbrauch zu fehlerhaften Aussagen der Chatbots kommen.¹⁴⁰ Aufgrund dieser Fehleranfälligkeit und im Interesse des Schutzes der Privaten vor übereiltem Handeln scheint ein Vorbehalt im entsprechenden Bereich bzw. im Einzelfall eher gerechtfertigt.¹⁴¹ Allerdings würde diesfalls die Verantwortung für den Erhalt einer korrekten Auskunft auf die Privaten überwältigt. Es erscheint bedenklich, dass es den öffentlichen Organen ermöglicht werden soll, sich von der Verantwortung bezüglich der wohlbekannten Risiken zu entbinden.¹⁴² Es ist daher vielmehr dafür zu plädieren, dass unkontrollierbare bzw. risikobehaftete Chatbots gar nicht erst zur behördlichen Auskunftserteilung eingesetzt werden. Ein Vorbehalt vermag die Fehleranfälligkeit von Chatbots nicht zu kompensieren.

4. Weitere Anforderungen

Schliesslich werden weitere Anforderungen an den Vertrauensschutz gestellt. Diese können bei einem Chatbot-Einsatz allerdings nicht weniger gut eingehalten werden als bei menschlichen behördlichen Auskünften. Damit der Vertrauensschutz greift, muss die anfragende Person aufgrund der Auskunft Dispositionen getroffen oder unterlassen haben, die sich nicht ohne negative Konsequenzen rückgängig machen oder nachholen lassen (Vertrauensbetätigung).¹⁴³ Die Art der Auskunft, d. h. ob diese von einem Menschen oder einem Chatbot erteilt wird, wirkt sich auf diese Voraussetzung nicht aus.¹⁴⁴ Ausserdem dürfen sich der Sachverhalt sowie die Gesetzgebung nachträglich nicht verändert haben, ansonsten sind die Behörden nicht mehr an die erteilte Aus-

¹³⁹ Siehe I.

¹⁴⁰ Plattner (Fn. 75), Rz. 519. Zur Gefahr von Manipulation und Missbrauch siehe [III.1.3](#).

¹⁴¹ Braun Binder et al., Studie 2021 (Fn. 16), S. 62.

¹⁴² BSK BV-Epiney (Fn. 61), Art. 5 N 74 ff.

¹⁴³ Häfelin/Müller/Uhlmann (Fn. 17), Rz. 659; Tschannen/Zimmerli/Müller (Fn. 58), § 22 N 486.

¹⁴⁴ Braun Binder et al., Studie 2021 (Fn. 16), S. 62.

kunft gebunden.¹⁴⁵ Schliesslich bedarf es einer Interessenabwägung zwischen dem Schutz des Vertrauens in die unrichtige Auskunft und dem Interesse an der richtigen Rechtsanwendung.¹⁴⁶ Auch diese Voraussetzungen bleiben durch einen Chatbot-Einsatz unberührt.

IV. Fazit

Es hat sich gezeigt, dass Private auch beim Einsatz von Chatbots grundsätzlich in ihrem Vertrauen auf behördliche Auskünfte zu schützen sind. Zu den Risiken eines Chatbot-Einsatzes gehört die Fehleranfälligkeit von datenbasierten Chatbots, insbesondere beim Einsatz von LLMs. Es werden zwar Methoden entwickelt, um die Fehleranfälligkeit zu minimieren. Damit lassen sich die Risiken jedoch nicht vollständig ausschliessen. Diese Risiken sind nach der hier vertretenen Auffassung grundsätzlich von der Behörde zu tragen, die Chatbots einsetzt, und nicht von der Person, die sich mit einer Frage an den Chatbot wendet. Insbesondere ist es nicht vertretbar, durch pauschale Vorbehalte gegenüber der Richtigkeit von Chatbot-Auskünften einen Interaktionskanal zwischen Behörden und Privaten zu schaffen, der die Verlässlichkeit der Auskunft von vornherein ausschliesst. Dies gilt umso mehr, wenn die Auskunft auf die rechtliche Orientierung Privater abzielt. Denn Private sollen grundsätzlich auf die Richtigkeit behördlicher Auskünfte vertrauen können. Vorbehalte im Einzelfall gegenüber datenbasierten Auskünften von Chatbots erscheinen demgegenüber eher zulässig. Stossend ist aber wiederum deren Fehleranfälligkeit. Deshalb ist nicht jede Art von Chatbot für die Auskunftserteilung geeignet. Insbesondere bei LLM-basierten Chatbots ist grosse Vorsicht geboten.

¹⁴⁵ Nur Häfelin/Müller/Uhlmann (Fn. 17), Rz. 695.

¹⁴⁶ Häfelin/Müller/Uhlmann (Fn. 17), Rz. 699; Tschannen/Zimmerli/Müller (Fn. 58), § 22 N 487.

Der Bundesrat hat mit seiner Entscheidung, die KI-Konvention des Europarats zu ratifizieren, den Gesetzgebungsprozess zur KI-Regulierung ausgelöst. Dennoch bleiben noch einige Fragen offen: Wie wird die Schweiz Transparenz beim staatlichen Einsatz von algorithmischen Systemen gewährleisten? Welche Anforderungen wird sie an die Begründung automatisierter Entscheide stellen? Welche Massnahmen wird sie gegen Diskriminierung und Manipulation durch KI ergreifen?

Dieser Sammelband bündelt Beiträge aus dem Forschungsprojekt «Nachvollziehbare Algorithmen: Ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz». Die Aufsätze bieten einen Überblick über offene Fragen im öffentlichen Recht und beleuchten insbesondere die Verfügung im Kontext algorithmischer Entscheidungen.