

Florent Thouvenin

Ein Rechtsrahmen für KI in der Schweiz: Handlungsbedarf und Handlungsoptionen bei Transparenz und Datenschutz

Der Bundesrat hat am 12. Februar 2025 entschieden, wie das Schweizer Recht auf die Herausforderungen reagieren soll, die sich bei der Entwicklung und Verwendung von Künstlicher Intelligenz (KI) ergeben. Dieser strategische Entscheid sieht vor, dass die Konvention des Europarats zu Künstlicher Intelligenz (KI-Konvention) ratifiziert und die dafür notwendigen Anpassungen im Schweizer Recht vorgenommen werden. Dieser Beitrag untersucht den Handlungsbedarf und die Handlungsoptionen für Regelungen gegenüber Privatpersonen in den inhaltlich verschränkten Bereichen Transparenz und Datenschutz.

Beitragsart: Wissenschaftliche Beiträge
Rechtsgebiete: Datenschutzrecht

Zitiervorschlag: Florent Thouvenin, Ein Rechtsrahmen für KI in der Schweiz: Handlungsbedarf und Handlungsoptionen bei Transparenz und Datenschutz, in: Jusletter 30. März 2026

Inhaltsübersicht

1. Einleitung
2. Handlungsbedarf und Handlungsspielraum
3. Verschränkung von Transparenz und Datenschutz
4. Handlungsfelder
 - 4.1. Transparenz im Allgemeinen
 - 4.2. Identifikation von KI-generierten Inhalten
 - 4.3. Transparenz bei Interaktion mit KI-Systemen
 - 4.4. Informationspflicht und Auskunftsrecht bei automatisierten Entscheidungen
 - 4.5. Bearbeitung von Personendaten beim Training von KI-Modellen
5. Erkenntnisse

1. Einleitung

[1] Am 12. Februar 2025 hat der Bundesrat entschieden, wie das Schweizer Recht auf die Herausforderungen reagieren soll, die sich bei der Entwicklung und Verwendung von Künstlicher Intelligenz (KI) ergeben.¹ Dieser strategische Entscheid beruht auf einer Auslegeordnung zur Regulierung von KI,² die sich ihrerseits auf umfassende Analysen der Bundesverwaltung stützt, namentlich auf eine rechtliche Basisanalyse,³ eine Analyse der regulatorischen Entwicklungen im Bereich von KI in 20 ausgewählten Ländern⁴ und eine Analyse der bestehenden und geplanten Änderungen in sektorspezifischen Regelungen des Bundesrechts.⁵ Der Entscheid umfasst im Wesentlichen zwei Aspekte. Zum einen sollen die Konvention des Europarats zu Künstlicher Intelligenz (KI-Konvention)⁶ ratifiziert und die dafür notwendigen Anpassungen im Schweizer Recht vorgenommen werden.⁷ Bundesrat Albert Rösti hat die KI-Konvention bereits am 27. März 2025 im Namen der Schweiz unterzeichnet.⁸ Zum anderen sollen die erforderlichen Gesetzesanpassungen möglichst sektorbezogen und themenspezifisch erfolgen; eine allgemeine, sektorübergreifende Regelung soll sich auf zentrale, grundrechtsrelevante Bereiche beschränken, bspw. auf

¹ Medienmitteilung des Bundesrates, KI-Regulierung: Bundesrat will Konvention des Europarates ratifizieren, 12. Februar 2025, abrufbar unter: <https://www.news.admin.ch/de/nsb?id=104110>; Stand: 10. März 2026.

² Bundesamt für Kommunikation BAKOM, Auslegeordnung zur Regulierung von künstlicher Intelligenz, Bericht an den Bundesrat, 12. Februar 2025, abrufbar unter: <https://www.sem.admin.ch/bj/de/home/staat/gesetzgebung/kuenstliche-intelligenz.html>; Stand: 10. März 2026.

³ Bundesamt für Justiz BJ, Rechtliche Basisanalyse im Rahmen der Auslegeordnung zu den Regulierungsansätzen im Bereich der künstlichen Intelligenz, 31. August 2024, abrufbar unter: <https://www.bj.admin.ch/bj/de/home/staat/gesetzgebung/kuenstliche-intelligenz.html>; Stand: 10. März 2026.

⁴ Bundesamt für Kommunikation BAKOM, Analyse der Regulierung von künstlicher Intelligenz in verschiedenen Ländern und Weltregionen, Basisanalyse für die Auslegeordnung zur Regulierung von künstlicher Intelligenz, 16. Dezember 2024, abrufbar unter: <https://backend.bakom.admin.ch/fileservice/sdweb-docs-prod-bakom-adminch-files/files/2025/04/15/5a875c91-a2a9-4259-9cd1-c93d9ee61ead.pdf>; Stand: 10. März 2026.

⁵ Bundesamt für Kommunikation BAKOM, Überblick zu aktuellen sektoriellen Regulierungsaktivitäten im Zusammenhang mit Künstlicher Intelligenz, 16. Dezember 2024, abrufbar unter: <https://backend.bakom.admin.ch/fileservice/sdweb-docs-prod-bakomadminch-files/files/2025/04/15/33835e01-8869-4ff8-9f75-c1d0150c50dd.pdf>; Stand: 10. März 2026.

⁶ Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 5. September 2024, Council of Europe Treaty Series No. 225, abrufbar unter: <https://rm.coe.int/1680afae3c>; Stand: 10. März 2026.

⁷ Medienmitteilung des Bundesrates (Fn. 1), KI-Regulierung.

⁸ Medienmitteilung des Bundesamtes für Kommunikation BAKOM, Schweiz unterzeichnet Europaratskonvention zu KI, 26. März 2025, abrufbar unter: <https://www.news.admin.ch/de/nsb?id=104646>; Stand: 10. März 2026.

den Datenschutz.⁹ Damit hat der Bundesrat – zu Recht¹⁰ – die vieldiskutierte Option verworfen, KI als Technologie zu regulieren, wie das insb. die EU mit der KI-Verordnung¹¹ getan hat.

[2] In seinem strategischen Entscheid hat der Bundesrat auch das weitere Vorgehen festgelegt. Bis Ende 2026 wird das Eidgenössische Justiz- und Polizeidepartement (EJPD) zusammen mit dem Eidgenössischen Departement für Umwelt, Verkehr, Energie und Kommunikation (UVEK) und dem Eidgenössischen Departement für auswärtige Angelegenheiten (EDA) eine Vernehmlassungsvorlage erarbeiten. Darin werden die gesetzlichen Massnahmen festgelegt, die zur Umsetzung der KI-Konvention notwendig sind. Im Vordergrund stehen aus Sicht des Bundesrates die Bereiche Transparenz, Datenschutz, Nichtdiskriminierung und Aufsicht. Die gesetzlichen Massnahmen sollen sich dabei an drei Zielen orientieren: der Stärkung des Innovationsstandortes Schweiz, der Wahrung des Grundrechtsschutzes (inkl. Wirtschaftsfreiheit) und der Stärkung des Vertrauens der Bevölkerung in KI.¹²

[3] Dieser Beitrag analysiert den Handlungsbedarf, der sich für den schweizerischen Gesetzgeber aus den Vorgaben der KI-Konvention ergibt. Er beschränkt sich auf die Bereiche Transparenz und Datenschutz und die Vorgaben gegenüber Privatpersonen. Die Bereiche Nichtdiskriminierung und Aufsicht werden in weiteren Beiträgen untersucht. Näher untersucht wird auch die seit Längerem diskutierte Frage, ob die in Art. 21 DSGVO vorgesehene Informationspflicht bei vollständig automatisierten Entscheidungen auf teilweise automatisierte Entscheidungen ausgedehnt werden soll und wie die Bearbeitung von Personendaten beim Training von KI-Modellen datenschutzrechtlich zu beurteilen ist.

2. Handlungsbedarf und Handlungsspielraum

[4] Die Bestimmungen der KI-Konvention sind nicht unmittelbar anwendbar («non self-executing»); sie müssen deshalb ins innerstaatliche Recht überführt werden. Die Unterzeichnerstaaten haben dabei einen grossen Handlungsspielraum, weil die Konvention keine konkreten Massnahmen zur Erreichung der Ziele vorgibt. Das gilt ganz besonders für die Vorgaben für Privatpersonen, wie die Definition des Anwendungsbereichs (Art. 3 KI-Konvention) zeigt. Während die Vertragsparteien verpflichtet sind, die Vorgaben der KI-Konvention auf die Verwendung von KI-Systemen durch öffentliche Organe anzuwenden (Art. 3 Abs. 1 Bst. a KI-Konvention), müssen sie gegenüber Privatpersonen («*private actors*») lediglich die Risiken und Wirkungen adressieren, die sich aus der Verwendung von KI-Systemen ergeben (Art. 3 Abs. 1 Bst. b KI-Konvention). Bei der Umsetzung der Vorgaben gegenüber Privatpersonen besteht damit ein grosser Handlungs-

⁹ Medienmitteilung des Bundesrates (Fn. 1), KI-Regulierung.

¹⁰ Siehe dazu etwa: FLORENT THOUVENIN/STEPHANIE VOLZ, Ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz (KI) in der Schweiz, ZBJV 2024, 613 ff., 618 f. und 627 f.; Digital Society Initiative (DSI), Positionspapier: Ein Rechtsrahmen für Künstliche Intelligenz, November 2021, abrufbar unter: <https://www.itsl.uzh.ch/dam/jcr:4687cf40-85a8-47f6-88ea-7ba54ad31d34/dsi-strategy-lab-21-de.pdf>; Stand: 10.03.2026; LEONIE RITSCHER/SIMON RUESCH/LUKAS FEDERER, Der Schweizer Weg in die KI-Zukunft: drei Szenarien, in: PwC, Whitepaper: KI-Standort Schweiz: Chancen nutzen und die Zukunft gestalten, 2024, 25 ff., 26 f. und 29 f.; THOMAS HOCHSTRASSER/ALEXANDER GÖBEL, KI-Regulierung und rechtliche Auswirkungen im Fokus, Anwaltsrevue 2025, 23 ff., 27.

¹¹ Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierender Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz).

¹² Medienmitteilung des Bundesrates (Fn. 1), KI-Regulierung.

spielraum, der nur punktuell durch zwei vergleichsweise konkrete Bestimmungen eingeschränkt wird.¹³

[5] Hinzu kommt, dass die KI-Konvention keine umfassenden Vorgaben zur Verwendung von KI-Systemen macht. Vielmehr soll sie nach ihrem Zweckartikel (lediglich) sicherstellen, dass Aktivitäten innerhalb des Lebenszyklus' von KI-Systemen vollständig mit *Grundrechten, Demokratie und Rechtsstaatlichkeit* vereinbar sind (Art. 1 Abs. 1 KI-Konvention). Da Privatpersonen in der Schweiz grundsätzlich nicht an die Grundrechte gebunden sind und die Gewährleistung der Rechtsstaatlichkeit allein Aufgabe des Staates ist, müssen die Vorgaben der KI-Konvention für Privatpersonen in der Schweiz nur umgesetzt werden, wenn eine *direkte Drittwirkung von Grundrechten* besteht oder wenn sich die Tätigkeit von Privatpersonen auf das *Funktionieren der Demokratie* auswirkt. Ersteres betrifft (wohl)¹⁴ nur den Anspruch von Mann und Frau auf gleichen Lohn für gleichwertige Arbeit (Art. 8 Abs. 3 Satz 3 BV), der auch besteht, wenn KI-Systeme bei lohnrelevanten Entscheidungen verwendet werden. Zweiteres wird wohl nur bei der Tätigkeit von Medien und grossen Kommunikationsplattformen der Fall sein. Neben dem grossen Handlungsspielraum innerhalb der Vorgaben der KI-Konvention ist der *Handlungsbedarf* bei der Umsetzung der Konvention gegenüber Privatpersonen damit auch inhaltlich *auf einige wenige Konstellationen beschränkt*. Bei der Umsetzung dieser Vorgaben wird der Gesetzgeber den Grundrechten auch im Rahmen der indirekten Drittwirkung (Art. 35 Abs. 3 BV) Rechnung tragen, namentlich der Rechtsgleichheit und dem Diskriminierungsverbot (Art. 8 BV), dem Schutz der Privatsphäre (Art. 13 BV), der Meinungs- und Informationsfreiheit (Art. 16 BV), der Medienfreiheit (Art. 17 BV), der Wirtschaftsfreiheit (Art. 27 BV) und den politischen Rechten (Art. 34 BV), insb. dem Schutz der freien Willensbildung und der unverfälschten Stimmabgabe (Art. 34 Abs. 2 BV).

[6] Der Schweizer Gesetzgeber sollte den *Handlungsspielraum*, den ihm die KI-Konvention belässt, in doppelter Hinsicht *nutzen*. Zum einen sollten die Vorgaben so umgesetzt werden, dass sie sinnvoll ins bestehende Schweizer Recht eingefügt werden können und inhaltlich den vom Bundesrat verfolgten Zielen – Stärkung des Innovationsstandortes Schweiz, Wahrung des Grundrechtsschutzes (inkl. Wirtschaftsfreiheit) und Stärkung des Vertrauens der Bevölkerung in KI¹⁵ – dienen. Zum anderen sollte der Gesetzgeber dort über die Vorgaben der Konvention hinausgehen, wo es mit Blick auf das bestehende Schweizer Recht und die mit den gesetzlichen Massnahmen verfolgten Ziele sinnvoll erscheint, bspw. bei der Informationspflicht bei automatisierten Entscheidungen (Art. 21 DSG)¹⁶ oder bei der Bearbeitung von Personendaten beim Training von KI-Modellen.¹⁷

¹³ Siehe dazu hinten, 4.2 und 4.3.

¹⁴ Zur umstrittenen Frage, ob eine unmittelbare Drittwirkung auch in weiteren Bereichen besteht, siehe: JÖRG PAUL MÜLLER, V. Teil Grund- und Menschenrechte, 1 Entstehung und Entwicklung der Grundrechte in der Schweiz, in: Diggelmann/Hertig Randall/Schindler (Hrsg.), Verfassungsrecht der Schweiz, Bd. II, Zürich/Basel/Genf 2020, Rn. 33 ff. und 39 ff.; BERNHARD WALDMANN, in: Waldmann/Belser/Epiney (Hrsg.), Basler Kommentar, Bundesverfassung, 2. Aufl., Basel 2025, BV 35 N 5, N 29; JÖRG PAUL MÜLLER, Verwirklichung der Grundrechte nach Art. 35 BV, Der Freiheit Chancen geben, Bern 2018, 126 ff.

¹⁵ Siehe dazu vorne, 1.

¹⁶ Siehe dazu hinten, 4.4.

¹⁷ Siehe dazu hinten, 4.5.

3. Verschränkung von Transparenz und Datenschutz

[7] Die Gewährleistung von Transparenz ist eine der zentralen Herausforderungen bei der Verwendung von KI. Menschen sind es seit jeher gewohnt, mit Menschen zu interagieren. Zudem bestand bisher die auf Erfahrung gestützte Erwartung, dass Entscheidungen grundsätzlich von Menschen getroffen und Texte, Bilder, Videos und andere Inhalte von Menschen geschaffen werden. Diese Erfahrungen und Erwartungen werden mit der Verwendung von KI zunehmend in Frage gestellt. Menschen interagieren nicht mehr nur mit Menschen, sondern auch mit Chatbots, Entscheidungen werden zunehmend von KI-Systemen getroffen, und Texte, Bilder und Videos werden von solchen Systemen generiert. Damit stellt sich die Frage, ob, in welchen Konstellationen und gegebenenfalls wie sicherzustellen ist, dass Menschen die Verwendung von KI-Systemen erkennen und deren Funktionsweise verstehen können.

[8] Angesichts der zentralen Bedeutung der Transparenz erstaunt es wenig, dass die KI-Konvention mehrere Vorgaben dazu enthält. Neben einer allgemein gefassten Bestimmung (Art. 8 KI-Konvention) finden sich auch zwei spezifische Vorgaben, eine zur Transparenz von KI-generierten Inhalten (Art. 8, zweiter Teilsatz KI-Konvention) und eine zur Transparenz bei der Interaktion mit KI-Systemen (Art. 15 Abs. 2 KI-Konvention). Wenn Menschen mit KI-Systemen interagieren oder diese nutzen, dann bearbeiten diese Systeme regelmässig Daten über diese Menschen, die als Personendaten im Sinne des DSG zu qualifizieren sind (Art. 5 lit. a DSG). Da die Bearbeitung von Personendaten umfassend durch das DSG geregelt wird und die Transparenz der Datenbearbeitung der wohl wichtigste Grundsatz des Datenschutzrechts (Art. 6 Abs. 2 f. DSG) ist, bestehen zwischen den Vorgaben der KI-Konvention zur Transparenz beim Einsatz von KI und den Vorgaben zur Transparenz im geltenden Datenschutzrecht enge Bezüge. Besonders deutlich wird dies bei der Informationspflicht bei automatisierten Entscheidungen (Art. 21 DSG), nach der Personen informiert werden müssen, wenn eine sie betreffende Entscheidung vollständig automatisiert gefällt wird, bspw. von einem KI-System. Aber auch die allgemeine Informationspflicht (Art. 19 f. DSG) und das Auskunftsrecht (Art. 25 ff. DSG) können zur Transparenz beim Einsatz von KI-Systemen beitragen. Soweit Personendaten bearbeitet werden, vermag das Datenschutzrecht deshalb die bei der Verwendung von KI-Systemen erforderliche Transparenz schon heute weitgehend zu gewährleisten. Wegen der engen inhaltlichen Bezüge und der Regelung zentraler Aspekte der Transparenz im Datenschutzrecht drängt es sich auf, die Bereiche Transparenz und Datenschutz in diesem Beitrag gemeinsam zu behandeln.

[9] Die KI-Konvention enthält neben den Vorgaben zur Transparenz auch eine allgemein gefasste Vorgabe zum Datenschutz. Art. 11 KI-Konvention verlangt, dass die Vertragsstaaten Massnahmen zum Schutz der Privatsphäre und für den Datenschutz (*«privacy and personal data protection»*) ergreifen oder aufrechterhalten. Diese Vorgaben werden durch das DSG bereits umfassend erfüllt, bei der Bearbeitung von Personendaten durch KI-Systeme der öffentlichen Hand ebenso wie bei der Bearbeitung durch Privatpersonen. Der EDÖB hat denn auch klargestellt, dass die Vorgaben des DSG auch bei der Verwendung von KI einzuhalten sind.¹⁸ Hier besteht grundsätzlich kein gesetzgeberischer Handlungsbedarf. Vielmehr werden der EDÖB und die Gerichte unter Rückgriff auf die Lehre eine sinnvolle Praxis zur Anwendung des DSG auf KI-Systeme entwickeln. Darüber hinaus könnte es sinnvoll sein, die Zulässigkeit der Bearbeitung von Personendaten für

¹⁸ Mitteilung des Eidgenössischen Datenschutz- und Öffentlichkeitsbeauftragten EDÖB, Update – Geltendes Datenschutzgesetz ist auf KI direkt anwendbar, 8. Mai 2025, abrufbar unter: <https://www.edoeb.admin.ch/de/update-geltendes-datenschutzgesetz-ist-auf-ki-direkt-anwendbar>; Stand: 10. März 2026.

das Training von KI-Modellen im DSG ausdrücklich zu regeln. Die KI-Konvention enthält dazu zwar keine Vorgaben, die Frage ist aber von zentraler Bedeutung und wird deshalb nachfolgend kurz untersucht (4.5). Näher zu untersuchen sind zudem die allgemeine Vorgabe zur Transparenz (4.1) sowie die spezifischen Vorgaben zur Identifikation von KI-generierten Inhalten (4.2) und zur Transparenz bei der Interaktion mit KI-Systemen (4.3). Zudem stellt sich die Frage, ob die Informationspflicht bei vollständig automatisierten Entscheidungen und das entsprechende Auskunftsrecht auf teilautomatisierte oder im Wesentlichen automatisierte Entscheidungen ausgedehnt werden sollen (4.4).

4. Handlungsfelder

4.1. Transparenz im Allgemeinen

[10] Nach Art. 8 KI-Konvention müssen die Vertragsstaaten Massnahmen vorsehen, die eine den spezifischen Kontexten und Risiken angemessene Transparenz während des gesamten Lebenszyklus' von KI-Systemen gewährleisten. Diese allgemeine Vorgabe zur Transparenz geht auf den ersten Blick äusserst weit, wird aber durch den Verweis auf eine dem Kontext und den Risiken angemessene Transparenz massgeblich eingeschränkt. Vorgaben ergeben sich für das Schweizer Recht hinsichtlich der Tätigkeit von Privatpersonen auch hier nur, wenn sich diese auf das Funktionieren der Demokratie auswirkt. Das ist vor allem bei der Verwendung von KI durch Medien und grosse Kommunikationsplattformen der Fall.¹⁹

[11] Bei der Verwendung von KI-Systemen lassen sich verschiedene Dimensionen von Transparenz unterscheiden. Zum einen stellt sich die Frage, ob für Menschen bei der Interaktion mit einem KI-System erkennbar sein muss, dass sie nicht mit einem Menschen, sondern mit KI interagieren; vergleichbare Fragen stellen sich bei Inhalten, die von einem generativen KI-System produziert worden sind («Erkennbarkeit»). Diese Fragen werden wohl von der allgemeinen Vorgabe zur Transparenz (Art. 8 KI-Konvention) mit umfasst, in der KI-Konvention aber spezifisch geregelt (Art. 15 Abs. 2 KI-Konvention; Art. 8 letzter Teilsatz KI-Konvention) und nachfolgend näher untersucht.²⁰ Zum anderen stellt sich die Frage, ob und inwiefern Menschen verstehen müssen, wie ein KI-System zu einer bestimmten Vorhersage, Empfehlung oder Entscheidung gelangt oder einen bestimmten Inhalt erstellt («Erklärbarkeit und Nachvollziehbarkeit»). Die Transparenz bezieht sich hier auf die Funktionsweise eines KI-Systems. Diese beiden Aspekte stehen bei der allgemeinen Vorgabe zur Transparenz im Vordergrund, wie sich aus dem *Explanatory Report* zur KI-Konvention ergibt.²¹

[12] Das Erfordernis der «Erklärbarkeit» («*explainability*») bezieht sich nach dem *Explanatory Report* auf die Möglichkeit, nach Massgabe des Standes der Technik ausreichend verständliche Erklärungen zu liefern, *warum* ein KI-System bestimmte Inhalte generiert, Vorhersagen oder Empfehlungen macht oder Entscheidungen trifft. Die Transparenz kann dabei bspw. durch eine Liste der Faktoren geschaffen werden, die das KI-System bei der Vorhersage, Empfehlung oder Ent-

¹⁹ Siehe dazu vorne, 2.

²⁰ Siehe dazu hinten, 4.3 und 4.2.

²¹ Council of Europe, *Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, Council of Europe Treaty Series – No. 225, 2024, Rn. 60 f.

scheidung berücksichtigt.²² Das Erfordernis der «*Nachvollziehbarkeit*» («*interpretability*») verweist darauf, dass verständlich sein muss, *wie* ein KI-System Inhalte generiert, Vorhersagen oder Empfehlungen macht oder Entscheidungen trifft.²³ Neben der Formulierung dieser Anforderungen weist der *Explanatory Report* auch darauf hin, dass der Transparenz bei KI-Systemen aus technischen Gründen Grenzen gesetzt sind, weil diese Systeme nicht von Menschen vorgegebenen Regeln folgen, sondern aufgrund der Analyse von statistischen Zusammenhängen in Daten selbst Regeln entwickeln, um von einem bestimmten Input zu einem bestimmten Output zu gelangen. Wie dieser Vorgang genau abläuft, ist auch für die Entwickler und Anwender von KI-Systemen oft nicht ohne weiteres klar. Schon deshalb kann keine volle, sondern nur eine angemessene Transparenz («*adequate transparency*») verlangt werden. Die Anforderungen bestimmen sich dabei nach dem Stand der Technik, den Umständen und dem Kontext.²⁴

[13] Bei der Umsetzung der allgemeinen Vorgabe zur Transparenz gegenüber Privatpersonen ist Zurückhaltung angezeigt. Aus technischer, wissenschaftlicher und gesellschaftspolitischer Sicht mag es zwar interessant und wünschenswert sein, die Funktionsweise von KI-Systemen möglichst genau verstehen und nachvollziehen zu können; es erstaunt deshalb wenig, dass die Forderung nach «*Explainable AI*» vor allem von Entwicklern, Wissenschaftlern und Vertretern von NGOs erhoben wird.²⁵ Eine allgemeine Rechtspflicht zur Schaffung von Transparenz bei der Entwicklung und Verwendung von KI-Systemen durch Privatpersonen würde aber massgeblich in Grundrechte eingreifen, namentlich in die Wirtschaftsfreiheit (Art. 27 BV) und in die Eigentumsgarantie (Art. 26 BV), die auch den Schutz von Immaterialgüterrechten und Geschäftsgeheimnissen umfasst.²⁶ Die Einführung von Transparenzpflichten ist deshalb nur zulässig, wenn ein öffentliches Interesse an der Transparenz besteht und die Einschränkung der Grundrechte verhältnismässig ist (Art. 36 BV). Diese Anforderungen dürften sich bei einer *allgemeinen Transparenzpflicht* bei der Entwicklung und Verwendung von KI-Systemen durch Privatpersonen kaum erfüllen lassen; anderes gilt aber für *sektorspezifische Transparenzpflichten*, die konkreten öffentlichen Interessen dienen, bspw. dem Schutz der Gesundheit, so etwa bei der Verwendung von KI-Systemen in Medizinprodukten, autonomen Fahrzeugen und Drohnen oder kritischen Infrastrukturen. In diesen Bereichen sollte der Gesetzgeber prüfen, ob die bestehenden sektorspezifischen Regelungen um Vorgaben zur Transparenz bei der Verwendung von KI-Systemen zu ergänzen sind.

²² Council of Europe (Fn. 21), Explanatory Report, Rn. 60.

²³ Council of Europe (Fn. 21), Explanatory Report, Rn. 61.

²⁴ Ebenso Council of Europe (Fn. 21), Explanatory Report, Rn. 61.

²⁵ Siehe dazu bspw. DAVID GUNNING/MARK STEFIK/JAESIK CHOI/TIMOTHY MILLER/SIMONE STUMPF/GUANG-ZHONG YANG, XAI – Explainable Artificial Intelligence, *Science Robotics* 2019, 4(37), abrufbar unter: https://openaccess.city.ac.uk/id/eprint/23405/8/aay7120_article_nearly%20final.pdf; Stand: 10. März 2026; RITA MATULIONYTE/AMBREEN HANIF, A call for more explainable AI in law enforcement, 2021 IEEE 25th International Enterprise Distributed Object Computing Workshop (EDOCW), 75 ff., abrufbar unter: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9626253&tag=1>; Stand: 10.03.2026; IBM, Was ist Explainable AI, abrufbar unter: <https://www.ibm.com/de-de/think/topics/explainable-ai>; Stand: 10. März 2026. Für aktuelle Forschungsprojekte zu Explainable AI siehe: Mitteilung der European Health and Digital Executive Agency vom 15. Juli 2025, abrufbar unter: https://hadea.ec.europa.eu/news/discover-three-new-eu-funded-projects-explainable-and-robust-ai-ai-data-and-robotics-partnership-2025-07-15_en; Stand: 10. März 2026; für weitere Projekte siehe: <https://www.zhaw.ch/en/research/project/79113> und <https://xai-project.eu/>; Stand beide: 10. März 2026.

²⁶ BERNHARD WALDMANN, in: Waldmann/Belser/Epiney (Hrsg.), *Basler Kommentar, Bundesverfassung*, 2. Aufl., Basel 2025, BV 26 N 19; KLAUS VALLENDER/PETER HETTICH, in: Ehrenzeller et al. (Hrsg.), *St. Galler Kommentar, Die Schweizerische Bundesverfassung*, 4. Aufl., Zürich 2023, BV 26 N 21; GIOVANNI BIAGGINI, *BV Kommentar, Bundesverfassung der Schweizerischen Eidgenossenschaft*, 2. Aufl., Zürich 2017, BV 26 N 12.

[14] Aus der KI-Konvention ergeben sich, wie erwähnt, nur Vorgaben, wenn sich die Tätigkeit von Privatpersonen auf das Funktionieren der Demokratie auswirkt. Das dürfte nur der Fall sein, wenn KI-Systeme von *Medien oder grossen Kommunikationsplattformen* verwendet werden. Im Vordergrund stehen dabei das Erstellen von Inhalten durch KI-Systeme (dazu sogleich)²⁷ und der Einsatz von sog. *Empfehlungssystemen* («*recommender systems*»), die von Kommunikationsplattformen verwendet werden. Das geltende Schweizer Recht kennt keine Normen, welche die Erklärbarkeit und Nachvollziehbarkeit von Empfehlungssystemen sicherstellen, weder allgemein noch spezifisch für Kommunikationsplattformen. Das Datenschutzrecht vermag zwar eine gewisse Transparenz zu schaffen, namentlich über die allgemeine Informationspflicht (Art. 19 f. DSG) und das Auskunftsrecht (Art. 25 ff. DGS); diese beschränkt sich aber auf Informationen über die Personendaten, die von den Anbietern von Kommunikationsplattformen bearbeitet werden. Ein Anspruch auf Informationen über die Funktionsweise von Empfehlungssystemen lässt sich daraus nicht ableiten. Das gilt auch für das Auskunftsrecht bei automatisierten Einzelentscheidungen (Art. 25 Abs. 2 lit. f DSG), das an sich einen Anspruch auf das Erteilen von Informationen über die Logik automatisierter Entscheidung vermitteln würde, dessen Anwendung aber daran scheitert, dass die Vorschläge von Empfehlungssystemen nicht als automatisierte Einzelentscheidungen im Sinne von Art. 21 DSG qualifiziert werden können, weil sie für die Betroffenen nicht mit einer Rechtsfolge verbunden sind und diese auch nicht erheblich beeinträchtigen. Damit besteht Handlungsbedarf. Dieser wurde bereits erkannt, zwar (wohl) nicht mit Blick auf die Umsetzung der KI-Konvention, wohl aber hinsichtlich der Regulierung von grossen Kommunikationsplattformen und Suchmaschinen.

[15] Am 29. Oktober 2025 wurde der *Vorentwurf für ein Bundesgesetz über Kommunikationsplattformen und Suchmaschinen (KomPG)* veröffentlicht. Dieser sieht vor, dass grosse Kommunikationsplattformen und Suchmaschinen in ihren allgemeinen Geschäftsbedingungen (AGB) die wichtigsten Parameter, die ihren Empfehlungssystemen zugrunde liegen, und deren Gewichtung darlegen müssen (Art. 18 Abs. 1 lit. a VE-KomPG). Damit soll sichergestellt werden, dass die Nutzer:innen von grossen Kommunikationsplattformen und Suchmaschinen nachvollziehen können, aus welchen Gründen ihnen ein bestimmter Inhalt angezeigt wird.²⁸ Diese Bestimmung dient der Transparenz. Sie ist zu Recht technologieunabhängig formuliert und findet damit auf alle Empfehlungssysteme Anwendung, auch auf solche, die auf KI-Systemen beruhen. Die hier in Frage stehende Vorgabe der KI-Konvention kann durch die vorgesehene Bestimmung im KomPG umgesetzt werden.

4.2. Identifikation von KI-generierten Inhalten

[16] Neben den allgemeinen Anforderungen an die Transparenz sieht die KI-Konvention auch eine spezifische Anforderung vor, nach der die Vertragsstaaten angemessene Transparenz hinsichtlich der Identifikation von Inhalten sicherstellen müssen, die durch KI-Systeme geschaffen wurden (Art. 8 letzter Teilsatz KI-Konvention). Die Massnahmen zur Sicherstellung der Transparenz müssen dabei auf den spezifischen Kontext und die spezifischen Risiken zugeschnitten sein (Art. 8 erster Teilsatz KI-Konvention). Nach dem *Explanatory Report* zur KI-Konvention soll

²⁷ Siehe dazu hinten, 4.2.

²⁸ Vorentwurf des Bundesgesetzes über Kommunikationsplattformen und Suchmaschinen (VE-KomPG), Erläuternder Bericht zur Eröffnung des Vernehmlassungsverfahrens vom 29. Oktober 2025, 27.

diese Vorgabe das Risiko von Täuschungen verringern und die Unterscheidung zwischen authentischen, also von Menschen geschaffenen, und von KI-Systemen generierten Inhalten ermöglichen.²⁹ Vor diesem Hintergrund hat der Bundesrat ausgeführt, dass es sinnvoll sein könnte, eine allgemeine Kennzeichnungspflicht für KI-generierte Inhalte einzuführen, die auch für Privatpersonen gelten würde.³⁰ Vergleichbare Vorgaben finden sich auch in der KI-Verordnung der EU; diese gelten aber nicht allgemein, sondern nur für bestimmte Konstellationen: Zum einen müssen Anbieter von KI-Systemen, die synthetische Inhalte erzeugen, sicherstellen, dass die Ausgaben ihrer KI-Systeme als künstlich erzeugt oder manipuliert erkennbar sind (Art. 50 Abs. 2 Satz 1 KI-VO). Zum anderen müssen die Betreiber von KI-Systemen, die Bild-, Ton- oder Videoinhalte erzeugen oder manipulieren, die sog. *Deepfakes* sind, offenlegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden (Art. 50 Abs. 4 Satz 1 KI-VO). Zudem müssen Betreiber von KI-Systemen, die Texte erzeugen oder manipulieren, welche die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse informieren, offenlegen, dass der Text künstlich erzeugt oder manipuliert wurde; diese Pflicht entfällt allerdings, wenn die Texte einer redaktionellen Kontrolle unterzogen werden und eine natürliche oder juristische Person die redaktionelle Verantwortung für die Veröffentlichung trägt (Art. 50 Abs. 4 Unterabsatz 2 KI-VO).

[17] Auf den ersten Blick scheint eine allgemeine Kennzeichnungspflicht für KI-generierte Inhalte sinnvoll, zumal – ganz allgemein – wenig gegen Transparenz spricht und Studien zeigen, dass Menschen Schwierigkeiten haben, KI-generierte Inhalte von solchen zu unterscheiden, die von Menschen geschaffen wurden.³¹ Zudem erscheint der Aufwand für die zur Kennzeichnung verpflichteten Privatpersonen überschaubar. Bei näherer Betrachtung kommen allerdings Zweifel auf. Die Norm spiegelt die Perspektive der Vertragsstaaten bei den Arbeiten an der KI-Konvention, mithin die Annahmen und Auffassungen in den Jahren 2022 bis 2024. Zu Beginn dieser Arbeiten waren KI-generierte Inhalte noch selten und die Qualität genügte auch bescheidenen Ansprüchen kaum. Seit der Markteinführung von ChatGPT im November 2022 wird generative KI aber von immer mehr Menschen und in immer mehr Bereichen für das Erstellen von Texten, Bildern und Videos genutzt, und deren Qualität wurde rasch besser. Die meisten Inhalte werden dabei durch ein Zusammenwirken von Menschen und Maschinen erstellt. Aufgrund dieser Entwicklung werden die meisten Menschen schon heute nicht mehr ohne weiteres davon ausgehen, dass Texte, Bilder oder Videos, die sie sehen, (allein) von einem Menschen erstellt worden sind. Diese Erkenntnis dürfte sich in Zukunft noch verfestigen. Vor dem Hintergrund dieser Entwicklung ist fraglich, ob das im *Explanatory Report* zur KI-Konvention als Begründung für die Kennzeichnungspflicht angeführte Risiko von Täuschungen tatsächlich (noch) besteht. Angesichts der rasch wachsenden Menge KI-generierter Inhalte kann man sich vielmehr fragen, ob statt KI-generierter Inhalte künftig von Menschen geschaffene Inhalte als solche gekennzeichnet werden sollten. Eine Rechtspflicht ist hier aber nicht erforderlich. Vielmehr ist denkbar, dass vollständig von Menschen geschaffene Inhalte künftig freiwillig als solche gekennzeichnet und

²⁹ Council of Europe (Fn. 21), Explanatory Report, Rn. 59.

³⁰ Bundesamt für Justiz, Rechtliche Basisanalyse zu KI (Fn. 3), 44.

³¹ JOEL FRANK et al., A Representative Study on Human Detection of Artificially Generated Media Across Countries, 10. Dezember 2023, abrufbar unter: <https://arxiv.org/pdf/2312.05976>; Stand: 10. März 2026; DI COOKE/ABIGAIL EDWARDS/SOPHIA BARKOFF/KATHRYN KELLY, As Good as a Coin Toss: Human Detection of AI-Generated Content, Communications of the ACM 2025, 100 ff., abrufbar unter: <https://dl.acm.org/doi/pdf/10.1145/3729417>; Stand: 10. März 2026; KRISTINA RADIVOJEVIC/MATTHEW CHOU/KARLA BADILLO-URQUIOLA/PAUL BRENNER, Human Perception of LLM-generated Text Content in Social Media Environments, 10. September 2024, abrufbar unter: <https://arxiv.org/pdf/2409.06653>; Stand: 10. März 2026.

als besonders werthaltig vermarktet werden, wie das bei handwerklich hergestellten Produkten schon heute der Fall ist. Mit Blick auf die technische Entwicklung und die bereits absehbaren Gewöhnungseffekte bestehen erhebliche Zweifel, ob eine allgemeine Kennzeichnungspflicht für KI-generierte Inhalte künftig einen Nutzen stiften könnte – oder ob eine solche Kennzeichnung nicht schon bald als ein Relikt aus längst vergangenen Zeiten angesehen würde.

[18] Hinzu kommen praktische Schwierigkeiten bei der Ausgestaltung. So ist bereits unklar, in welchem Gesetz (oder in welchen Gesetzen) eine allgemeine Kennzeichnungspflicht für KI-generierte Inhalte eingeführt werden könnte. Gewichtiger erscheint, dass es schwierig wäre, eine sinnvolle und allgemeingültige Grenze zwischen KI-generierten und von Menschen geschaffenen Inhalten zu ziehen, zumal Inhalte regelmässig in einem Zusammenwirken von KI-Systemen und Menschen geschaffen werden. Will man eine Grenze ziehen, müsste man wohl auf einen hinreichend relevanten menschlichen Beitrag abstellen, ein simpler Prompt dürfte kaum genügen. Damit ist die Frage aber keineswegs geklärt. Wie schwierig es ist, eine Grenze zu ziehen, zeigt sich im Urheberrecht, wo sich diese Frage schon heute stellt. Dort ist allgemein anerkannt, dass Inhalte, die von einem KI-System automatisch aufgrund eines Prompts erstellt werden, nicht urheberrechtlich geschützt sind, weil das Urheberrecht nur geistige Schöpfungen schützt (Art. 2 Abs. 1 URG), also Inhalte, die von Menschen geschaffen wurden.³² Wie relevant und welcher Art der menschliche Beitrag sein muss, ist aber weitgehend unklar. Im Urheberrecht hat dies zur Folge, dass unklar ist, ob ein mithilfe eines KI-Systems erstellter Inhalt urheberrechtlich geschützt ist. Bei der vorliegenden Frage bezieht sich die Unsicherheit aber nicht (erst) auf die Rechtsfolge, sondern (bereits) auf die Tatbestandsvoraussetzung der Kennzeichnungspflicht. Dies führt zu erheblicher Rechtsunsicherheit, weil für die Anbieter von Inhalten unklar ist, ab wann ein Inhalt als KI-generiert gilt und sie diesen entsprechend kennzeichnen müssen.

[19] Insgesamt bestehen damit *erhebliche Zweifel am Sinn einer allgemeinen Kennzeichnungspflicht für KI-generierte Inhalte*. Zur Einführung einer solchen Kennzeichnungspflicht ist die Schweiz aufgrund der KI-Konvention aber auch gar nicht verpflichtet. Denn zum einen verlangt die Konvention lediglich, dass die Vertragsstaaten Massnahmen ergreifen, um eine angemessene, auf den spezifischen Kontext und die spezifischen Risiken zugeschnittene Transparenz hinsichtlich der Identifikation von KI-generierten Inhalten sicherzustellen (Art. 8 KI-Konvention). Die Konvention gibt damit nur das Ziel der Identifikation vor, verlangt aber nicht eine allgemeine Kennzeichnung von KI-generierten Inhalten als Mittel. Zum anderen muss eine Identifikation von KI-generierten Inhalten unter Privatpersonen nach dem Zweck der KI-Konvention nur sichergestellt werden, wenn Grundrechte verletzt werden oder das Funktionieren der Demokratie bedroht ist. Aus der Pflicht zum Schutz der Grundrechte ergeben sich keine Vorgaben für Privatpersonen, zumal nicht ersichtlich ist, wie das Erstellen oder Verwenden von KI-generierten Inhalten die Lohnungleichheit von Mann und Frau (Art. 8 Abs. 3 Satz 3 BV) betreffen könnte. Eine *Gefahr für das Funktionieren der Demokratie* können KI-generierte Inhalte wohl nur schaffen, wenn sie von vielen Menschen wahrgenommen werden, namentlich weil sie in den Medien oder auf grossen Kommunikationsplattformen verbreitet werden. Zu denken ist insb. an sog. *Deepfakes* (z.B. Videos, in denen Politiker:innen scheinbar Aussagen machen, die sie tatsächlich nie gemacht haben) und

³² Siehe dazu etwa: CONSTANZE SEMMELMANN, Künstliche Intelligenz und Urheberrecht: Stand zu Training und Output 2024, sic! 2024, 637 ff., 644; CYRIL R. DÖRFLER, Das Schöpferprinzip im Immaterialgüterrecht, Bern 2024, 221 f.; MANFRED REHBINDER/LORENZ HAAS/KAI-PETER UHLIG, Orell Füssli Kommentar, Urheberrechtsgesetz mit weiteren Erlassen und internationalen Abkommen, 4. Aufl., Zürich 2022, URG 2 N 2.

an KI-generierte Desinformation (z.B. automatisiert erstellte *Fake News* über angeblichen Wahlbetrug). Die Gefährdung ergibt sich dabei allerdings nicht aus dem Umstand, dass ein Inhalt mit KI generiert wurde, sondern aus seinem täuschenden Gehalt. Dieser entsteht, weil ein mit technischen Mitteln erzeugter oder veränderter Inhalt von Nutzenden fälschlicherweise als authentisch, also als echt und wahrheitsgemäss, wahrgenommen wird. Vor diesem Hintergrund drängt es sich auf, Medien und grosse Kommunikationsplattformen nicht allgemein zur Kennzeichnung von KI-generierten Inhalten, sondern nur zur Kennzeichnung von Inhalten zu verpflichten, die von den Nutzenden fälschlicherweise für authentisch gehalten werden und sich auf das Funktionieren der Demokratie auswirken können. Das kann durch gesetzliche Vorgaben oder durch administrative oder andere Massnahmen erfolgen (Art. 1 Abs. 2 KI-Konvention), bspw. in Form von Selbstregulierung.

[20] Die *Schweizer Medien* haben die Herausforderungen bei der Nutzung von KI bereits erkannt und adressiert. Der Schweizer Presserat hat im Januar 2024 einen Leitfaden zu KI im Journalismus erlassen und dort unter anderem eine Kennzeichnungspflicht für Inhalte vorgesehen, «die mithilfe eines KI-Programms erstellt wurden». Der Presserat empfiehlt zudem grösstmögliche Transparenz und fordert die Redaktionen auf, explizit anzugeben, wo und wie sie KI-Systeme einsetzen.³³ Da sich alle grossen, demokratierelevanten Schweizer Medien dem Presserat angeschlossen haben und die Vorgaben des Leitfadens – soweit ersichtlich – auch einhalten, besteht hier derzeit kein gesetzgeberischer Handlungsbedarf.

[21] Anderes gilt für grosse und damit ebenfalls demokratierelevante *Kommunikationsplattformen*. Für diese sollte der Gesetzgeber eine Regelung vorsehen, die sie zur Kennzeichnung von nicht-authentischen Inhalten verpflichtet, sofern die Inhalte politisch oder gesellschaftlich relevant und damit geeignet sind, sich auf das Funktionieren der Demokratie auszuwirken. Zur Umsetzung einer solchen Pflicht müssten die Kommunikationsplattformen ihre Nutzer:innen in den Nutzungsbedingungen verpflichten, nicht-authentische politisch und gesellschaftlich relevante Inhalte zu kennzeichnen; welche Inhalte erfasst sind, können die Plattformen in ihren Nutzungsbedingungen näher umschreiben und mit Beispielen illustrieren. In Übereinstimmung mit im EU-Recht etablierten und auch in der Schweiz anerkannten Grundsätzen der Plattformregulierung³⁴ sollten die Kommunikationsplattformen allerdings nicht zur Überwachung von Inhalten oder zur aktiven Nachforschung verpflichtet werden. Zum Schutz des Funktionierens der Demokratie dürfte es vielmehr genügen, wenn die Betreiber von Plattformen verpflichtet würden, auf Hinweis von Nutzer:innen oder Dritten zu prüfen, ob ein auf ihrer Plattform zugänglich gemachter politisch oder gesellschaftlich relevanter Inhalt authentisch ist, und diesen Inhalt gegebenenfalls als nicht-authentisch zu kennzeichnen. Dass der Anwendungsbereich einer solchen Pflicht mit dem Abstellen auf «politisch oder gesellschaftlich relevante Inhalte» äusserst weit

³³ Schweizer Presserat, «KI im Journalismus», Leitfaden vom 17. Januar 2024, abrufbar unter: https://presserat.ch/journalistenkodex/ki_leitfaden/; Stand: 10. März 2026.

³⁴ Siehe dazu statt vieler: ROLF H. WEBER, E-Commerce und Recht, Rechtliche Rahmenbedingungen elektronischer Geschäftsformen, 2. Aufl., Zürich/Basel/Genf 2010, 516; THOMAS HOEREN, Internetrecht, Ein Grundriss, 4. Aufl., Berlin/Boston 2018, Rn. 1001 ff.; PETER NOBEL/ROLF H. WEBER, Medienrecht, 4. Aufl., Bern 2021, 5. Kapitel Rn. 24 ff. Zum Problem der Desinformation im Besonderen: FLORENT THOUVENIN/STEPHANIE VOLZ/MARC EISENEGGER/DANIEL VOGLER/MARIELA JAFFÉ, Governance von Desinformation in digitalisierten Öffentlichkeiten, Jusletter vom 5. Februar 2024, Rn. 45; MICHEL BESSON/VÉRONIQUE BOILLET, The guarantee of political rights in view of misinformation: Is new regulation needed for Swiss referenda?, in: Baume/Boillet/Martenet (Hrsg.), Misinformation in referenda, London 2021, 251. Zur Umsetzung des Ansatzes des EU-Rechts in der Schweiz siehe: SWICO, Code of Conduct Hosting (CCH), abrufbar unter: <https://www.swico.ch/de/verband/oeffentlichkeitsarbeit/eigenverantwortung/code-conducts-hosting/>; Stand: 10. März 2026.

und offen ist, ist zwar nicht ideal, kann aber in Kauf genommen werden, weil die Rechtsfolge der Kennzeichnung wenig einschneidend ist. Eine spezifischere Definition des Anwendungsbereichs erscheint angesichts der Vielfalt der Inhalte und Konstellationen, die für das Funktionieren der Demokratie relevant sein können, kaum möglich. Im Zweifelsfall dürften sich die Plattformen für eine Kennzeichnung nicht-authentischer Inhalt entscheiden, was mit Blick auf das Ziel der Transparenz unproblematisch oder gar wünschenswert ist.

[22] Eine Pflicht der grossen Kommunikationsplattformen zur Kennzeichnung von politisch und gesellschaftlich relevanten nicht-authentischen Inhalten sollte im vorgesehenen *Bundesgesetz über Kommunikationsplattformen und Suchmaschinen (KomPG)* vorgesehen werden. Der Vorentwurf für das KomPG enthält weder eine solche Pflicht noch andere Massnahmen zur Identifizierung von nicht-authentischen Inhalten. Damit weicht er von der Regelung im *Digital Services Act (DSA)* der EU³⁵ ab. Dieser sieht zwar ebenfalls keine Pflicht vor, führt die Kennzeichnung von Inhalten aber als eine mögliche Massnahme von sehr grossen Online-Plattformen und sehr grossen Online-Suchmaschinen zur Minderung systemischer Risiken auf, wenn eine Einzelinformation bestehenden Personen, Gegenständen, Orten oder anderen Einrichtungen oder Ereignissen merklich ähnlich ist und einer Person fälschlicherweise als echt oder wahrheitsgemäss erscheint – und zwar unabhängig davon, ob es sich um einen erzeugten oder manipulierten Bild-, Ton- oder Videoinhalt handelt (Art. 35 Abs. 1 Bst. k DSA). Im Hinblick auf die Ratifizierung der KI-Konvention und die Umsetzung der Vorgabe zur Kennzeichnung KI-generierter Inhalte sollte das KomPG um eine vergleichbare Bestimmung ergänzt werden. Diese sollte im Sinne der vorstehenden Ausführungen aber enger gefasst werden und nur politisch und gesellschaftlich relevante nicht-authentische Inhalte erfassen. Im Gegensatz zum DSA sollte die Bestimmung aber als Pflicht ausgestaltet werden.

4.3. Transparenz bei Interaktion mit KI-Systemen

[23] Die KI-Konvention verpflichtet alle Vertragsstaaten, sicherzustellen, dass Personen, die mit einem KI-System interagieren, in einer dem Kontext angemessenen Weise darüber informiert werden, dass sie mit einem solchen System und nicht mit einem Menschen interagieren (Art. 15 Abs. 2 KI-Konvention). Erfasst werden direkte Interaktionen mit einem KI-System, bspw. mit einem Chatbot.³⁶ Eine vergleichbare Vorgabe findet sich auch in der KI-Verordnung der EU (Art. 50 Abs. 1 KI-VO). Die Bedeutung der Vorgabe der KI-Konvention für Privatpersonen wird zwar durch den Zweckartikel der KI-Konvention relativiert, zumal deren Interaktion mit KI-Systemen kaum je die Lohnleichheit von Mann und Frau (Art. 8 Abs. 3 Satz 3 BV) betreffen und sich nur in seltenen Fällen auf das Funktionieren der Demokratie auswirken wird. Dass Menschen bei einer Interaktion mit einem KI-System erkennen können, dass sie mit einem solchen System und nicht mit einem Menschen interagieren, erscheint aber allgemein als sinnvoll.³⁷ Diese Transparenz scheint einem allgemeinen Bedürfnis der meisten Menschen zu entsprechen. Abgesehen von den (überschaubaren) Kosten der Umsetzung sind keine Gründe ersichtlich, die gegen eine solche

³⁵ Verordnung (EU) 2022/2065 des europäischen Parlaments und des Rates vom 19. Oktober 2022 über den Binnenmarkt für digitale Dienste und zur Änderung der Richtlinie 2000/31/EG (Gesetz über digitale Dienste).

³⁶ Council of Europe (Fn. 21), Explanatory Report, Rn. 104.

³⁷ Ebenso: DSI (Fn. 10), Positionspapier, 3; ROLF WEBER, Künstliche Intelligenz: Regulatorische Überlegungen zum «Wie» und «Was», Zeitschrift für Europarecht 2022, B 1 ff., B 13.

Vorgabe sprechen würden. Die Vorgabe der KI-Konvention sollte deshalb im Schweizer Recht in allgemeingültiger Weise umgesetzt werden.

[24] Der Wortlaut von Art. 15 Abs. 2 KI-Konvention legt nahe, dass die Konvention eine aktive Information der betroffenen Personen verlangt (*«persons [...] are notified»*). Dieses Verständnis spiegelt sich auch in einigen Formulierungen des *Explanatory Report* (*«persons [...] should be duly notified»* und *«notification obligation»*). Art. 15 Abs. 2 KI-Konvention hält aber fest, dass es auf den jeweiligen Kontext (*«as appropriate for the context»*) ankommt, und der *Explanatory Report* präzisiert, dass die Vorgabe nicht Situationen erfassen soll, in welchen der Einsatz eines KI-Systems offensichtlich ist (*«where the use of the system is obvious from the context»*). Eine vergleichbare Einschränkung sieht auch die KI-Verordnung vor (Art. 50 Abs. 1 KI-VO). Damit ist klar, dass die KI-Konvention nicht immer eine aktive Information verlangt. Aus tatsächlicher Sicht wird eine aktive Information in vielen Fällen weder erforderlich noch sinnvoll sein, weil die betroffenen Personen regelmässig aus den Umständen erkennen können, dass sie mit einem KI-System interagieren, z.B. bei einem Kunden-Chatbot. Entscheidend ist nach dem Sinn und Zweck der Bestimmung einzig, dass Menschen erkennen können, ob sie mit einem KI-System oder mit einem anderen Menschen interagieren. Dazu wird in vielen Fällen keine aktive Information der betroffenen Personen (bspw. durch einen Hinweis in Textform, ähnlich wie bei «Cookie Banners») erforderlich sein. Erkennen die Betroffenen nicht schon aus den Umständen, dass sie mit einem KI-System interagieren, dann sollte die Interaktion bewusst so ausgestaltet werden, dass die Erkennbarkeit gewährleistet ist (*«transparency by design»*). Nur wenn auch dies nicht möglich oder nicht sinnvoll ist, sollte der Betreiber eines KI-Systems verpflichtet werden, die Erkennbarkeit aktiv durch Vermittlung einer entsprechenden Information sicherzustellen.

[25] Vor diesem Hintergrund und mit Blick auf die bestehende Informationspflicht für automatisierte Entscheidungen (Art. 21 DSGVO) und die weiteren Bestimmungen zur Transparenz im DSGVO drängt es sich auf, eine allgemeine Vorgabe zur Transparenz bei einer Interaktion mit einem KI-System im DSGVO zu regeln. Zwar lässt sich eine solche Informationspflicht mit etwas «juristischer Fantasie» bereits aus den Vorgaben des DSGVO ableiten.³⁸ Mit Blick auf die Bedeutung der Frage und die Rechtssicherheit erscheint es aber sinnvoll, eine Informationspflicht bei der Interaktion mit einem KI-System ausdrücklich im DSGVO zu regeln, bspw. in der Bestimmung über die Informationspflicht bei automatisierten Entscheidungen (Art. 21 DSGVO). Wie dort sollte auch hier nicht auf die verwendete Technologie, sondern auf den Umstand einer Interaktion mit einer Maschine abgestellt und die Informationspflicht technologieneutral formuliert werden. Für eine ausdrückliche Regelung spricht auch, dass es in gewissen Konstellationen möglich (und datenschutzrechtlich sinnvoll oder gar erforderlich) sein kann, die Interaktion mit einem KI-System so auszugestalten, dass keine Personendaten bearbeitet werden (bspw. bei einem Bedienungsroboter). In diesen Fällen liesse sich die Informationspflicht nicht mehr auf die bestehenden Normen des DSGVO stützen, wohl aber auf eine neue ausdrückliche Informationspflicht, die alle Interaktionen mit einer Maschine erfasst. Dass dann eine im DSGVO normierte Pflicht auch in Konstellationen greifen würde, in denen keine Personendaten bearbeitet werden, ist zwar gesetzessystematisch unbefriedigend. Da kein anderer Erlass für das Einfügen einer solchen allgemeinen Regelung ersichtlich ist, kann (und sollte) das aber hingenommen werden. Das gilt umso mehr, als die hier in

³⁸ Siehe dazu: Eidgenössischer Datenschutz- und Öffentlichkeitsbeauftragter (EDÖB), KI und Datenschutz, 24. September 2025, abrufbar unter: <https://www.edoeb.admin.ch/de/ki-und-datenschutz>; Stand: 10. März 2026.

Frage stehende Informationspflicht bei einer Interaktion mit einer Maschine der bereits im DSG geregelten Informationspflicht bei automatisierten Entscheidungen (Art. 21 DSG) nahesteht.

4.4. Informationspflicht und Auskunftsrecht bei automatisierten Entscheidungen

[26] Das DSG sieht eine besondere Informationspflicht und ein besonderes Auskunftsrecht bei vollständig automatisierten Entscheidungen vor. Die Bestimmungen wurden mit der Teilrevisi-
on vom 25. September 2020 eingeführt und gelten seit dem 1. September 2023. Die Informati-
onspflicht greift, wenn eine Entscheidung ausschliesslich auf einer automatisierten Bearbeitung
von Personendaten beruht und die Entscheidung für die betroffene Person mit einer Rechtsfolge
verbunden ist oder sie erheblich beeinträchtigt (Art. 21 Abs. 1 DSG). In diesem Fall haben die
betroffenen Personen die Möglichkeit, ihren Standpunkt darzulegen, und sie können verlangen,
dass die Entscheidung von einer natürlichen Person überprüft wird, was bisweilen als «Recht auf
menschliches Gehör»³⁹ bezeichnet wird (Art. 21 Abs. 2 DSG). Zudem können die Betroffenen ge-
stützt auf das Auskunftsrecht verlangen, dass ihnen neben der Information über das Vorliegen
einer automatisierten Entscheidung auch Informationen über die Logik mitgeteilt werden, auf
der die Entscheidung beruht (Art. 25 Abs. 2 lit. f DSG).

[27] Die Informationspflicht und das Auskunftsrecht bei vollständig automatisierten Entschei-
dungen haben bisher kaum eine praktische Bedeutung erlangt, weil nur wenige Unternehmen
Entscheidungen, die mit einer Rechtsfolge oder einer erheblichen Beeinträchtigung der betroffe-
nen Personen verbunden sind, vollständig automatisiert fällen. KI-Systeme (und andere algorith-
mische Systeme) werden bis heute meist nur zur Unterstützung menschlicher Entscheidungen
verwendet (sog. «*decision support systems*»). In diesen Fällen besteht nach dem geltenden Recht
weder eine besondere Informationspflicht noch haben die betroffenen Personen ein besonderes
Auskunftsrecht. Das erscheint problematisch, weil sich gezeigt hat, dass Menschen dazu nei-
gen, den Empfehlungen automatisierter Systeme auch dann zu folgen, wenn menschliches Ur-
teilsvermögen auf Fehler hinweisen würde (sog. «*automation bias*»)⁴⁰. Dieser Effekt entsteht, weil
Menschen regelmässig den Weg des geringsten kognitiven Aufwands wählen und den Vorschlag
eines automatisierten Systems als verlässliche Grundlage für ihre Entscheidungen ansehen, be-
sonders unter Stress oder Zeitdruck, und erst recht, wenn sie Abweichungen vom Vorschlag des
Systems begründen müssten.⁴¹ Es ist deshalb fraglich, ob es sinnvoll ist, kategorisch zwischen
vollständig automatisierten und anderen Entscheidungen zu trennen. Vor diesem Hintergrund
wird in der Literatur schon seit Längerem gefordert, dass der Anwendungsbereich der Informa-

³⁹ DAVID ROSENTHAL, Das neue Datenschutzgesetz, Jusletter vom 16. November 2020, Rn. 111; FLORENT THOUVENIN/
ALFRED FRÜH, Automatisierte Entscheidungen: Grundfragen aus der Perspektive des Privatrechts, SZW 2020, 3 ff.,
15; KURT PÄRLI/NATHALIE FLÜCK, in: Baeriswyl/Pärli/Blonski (Hrsg.), Stämpflis Handkommentar, Datenschutz-
gesetz, 2. Aufl., Bern 2023, DSG 21 N 4.

⁴⁰ Siehe dazu: HANNAH RUSCHEMEIER/LUKAS J. HONDRICH, Automation bias in public administration – an inter-
disciplinary perspective from law and psychology, Government Information Quarterly 2024, Vol. 41, Issue 3,
1 ff., abrufbar unter: <https://www.sciencedirect.com/science/article/pii/S0740624X24000455>; Stand: 10. März
2026; HANNAH RUSCHEMEIER, The Problem of the Automation Bias in the Public Sector: A Legal Perspective, Pro-
ceedings of the Weizenbaum Conference 2023, 1 ff., abrufbar unter: [https://papers.ssrn.com/sol3/papers.cfm?](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4521474)
[abstract_id=4521474](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4521474); Stand: 10. März 2026; GIUSEPPE ROMEO/DANIELA CONTI, Exploring automation bias in
human-AI collaboration: a review and implications for explainable AI, AI & Society 2026, 259 ff., abrufbar unter:
<https://link.springer.com/article/10.1007/s00146-025-02422-7>; Stand: 10. März 2026.

⁴¹ ROMEO/CONTI (Fn. 40), 262 f.

tionspflicht und des Auskunftsrechts bei vollständig automatisierten Entscheidungen auf teilautomatisierte Entscheidungen ausgedehnt wird.⁴² Diese Forderung wurde schon vor Inkrafttreten des geltenden Art. 21 DSG in den politischen Diskurs eingebracht,⁴³ und auch das BJ kommt in seiner rechtlichen Basisanalyse zum Schluss, dass eine Ausweitung der Informationspflicht und des Auskunftsrechts auf teilautomatisierte Entscheidungen «von Vorteil sein [könnte]».⁴⁴

[28] Eine solche Ausdehnung ginge allerdings sehr weit und wäre nur schwer umsetzbar. Zum einen sind schon heute sehr viele Entscheidungen in Unternehmen in irgendeiner Weise automatisiert, und es ist absehbar, dass die Automatisierung rasch weiter zunehmen wird. Der Anwendungsbereich der heutigen Regelung würde jedenfalls dann äusserst weit ausgedehnt, wenn man den Begriff «teilautomatisiert» in Anlehnung an das heutige Verständnis der Wendung «ausschliesslich auf einer automatisierten Bearbeitung beruht» verstehen würde. Nach diesem schliesst jede menschliche Beteiligung bei der inhaltlichen Beurteilung eine ausschliesslich automatisierte Bearbeitung aus.⁴⁵ Im Umkehrschluss müsste damit jede Beteiligung einer Maschine bei der inhaltlichen Beurteilung zu einer Qualifikation als «teilautomatisierte» Entscheidung führen. Dieses Verständnis würde zwar erlauben, den Anwendungsbereich der besonderen Informationspflicht und des besonderen Auskunftsrechts klar zu definieren, die Bestimmungen würden aber sehr viele Fälle erfassen, in denen eine Information weder erforderlich noch sinnvoll erscheint, weil die Entscheidung im Wesentlichen von Menschen gefällt wird.

[29] Ausgangspunkt der Überlegungen zu einer allfälligen Ausdehnung der Informationspflicht und des Auskunftsrechts bei vollständig automatisierten Entscheidungen von Privatpersonen muss aus rechtlicher Sicht sein, dass es gegenüber Privatpersonen keinen Anspruch auf rechtliches Gehör gibt und diese nicht verpflichtet sind, ihre Entscheidungen zu begründen oder anderweitig offenzulegen, wie diese zustande kamen. Vor diesem Hintergrund stellt sich die grundsätzliche Frage, weshalb das geltende Recht Privatpersonen verpflichtet, andere Personen darüber zu informieren, dass eine sie betreffende Entscheidung vollständig automatisiert gefällt wird.⁴⁶ Erklären lässt sich das wohl nur mit einer historisch bedingten und weiterhin verbreiteten Skep-

⁴² RUSCHEMEIER/HONDRICH (Fn. 40), 5 f.; MARKUS WINKLER, Haftpflichtrecht für künstliche Intelligenz, Warum erst die starke KI unser Haftpflichtregime herausfordert, Zürich 2025, 36; NADJA BRAUN BINDER, Automatisierte Entscheidungen: Perspektive Datenschutzrecht und öffentliche Verwaltung, SZW 2020, 27 ff., 33 f.; THOUVENIN/FRÜH (Fn. 39), 16.

⁴³ Parlamentarische Initiative Wasserfallen (23.438) «Datenschutzgesetzgebung anpassen und um teilautomatisierte Entscheide aufgrund künstlicher Intelligenz ergänzen» vom 15. Juni 2023; Postulat Marti (21.4406) «Bericht zur Regulierung von automatisierten Entscheidungssystemen» vom 9. Dezember 2021.

⁴⁴ Bundesamt für Justiz, Rechtliche Basisanalyse zu KI (Fn. 3), 55 f., etwas zurückhaltender 44.

⁴⁵ 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, 7056 f.; SIMON HENSELER/DAVID VASELLA, in: Blechta/Vasella (Hrsg.), Basler Kommentar, Datenschutzgesetz/Öffentlichkeitsgesetz, 4. Aufl., Basel 2024, DSG 21 N 21 f.; PÄRLI/FLÜCK (Fn. 39), DSG 21 N 6; FABIENNE SUTER, Automatisierte Einzelentscheidungen im (Schweizer) Datenschutzrecht, Phänomen, Entstehungsgeschichte, Regelungsziele und Begriff, Zürich 2024, 297.

⁴⁶ Bezeichnenderweise findet sich weder in der Botschaft des Bundesrates noch in den Protokollen der parlamentarischen Beratungen eine inhaltliche Begründung für eine besondere Informationspflicht bei vollständig automatisierten Entscheidungen. Die Einführung einer solche Pflicht war allerdings unbestritten, wohl weil die sog. Konvention 108+ des Europarates vom 17./18. Mai 1981 («*Modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data*») ein Recht der Betroffenen vorsieht, keiner automatisierten Entscheidung unterworfen zu werden, ohne dass ihre Sicht berücksichtigt wurde («*Every individual shall have a right: not to be subject to a decision significantly affecting him or her based solely on an automated processing of data without having his or her views taken into consideration*»; Art. 9 Abs. 1 Bst. a Konvention 108+). Dass die im DSG eingeführte Informationspflicht nicht mit dem in der Konvention 108+ vorgesehenen Recht gleichgesetzt werden kann und die Einführung einer Informationspflicht deshalb einer Begründung bedurft hätte, scheint nicht diskutiert worden zu sein.

sis gegenüber der «Qualität» solcher Entscheidungen.⁴⁷ Vor einer Ausdehnung der besonderen Informationspflicht und des besonderen Auskunftsrechts auf teilautomatisierte Entscheidungen müsste deshalb erst einmal untersucht werden, was der Zweck der bestehenden Regelung ist und ob sie diesen Zweck auch erfüllt. Dazu wäre unter anderem zu prüfen, ob die Information von den betroffenen Personen überhaupt zur Kenntnis genommen bzw. nachgefragt und als relevant und nützlich erachtet wird. Das erscheint durchaus fraglich, lässt sich hier aber mangels empirischer Erhebungen nicht beantworten.

[30] Vor diesem Hintergrund und mit Blick auf den «*automation bias*», der meist als Grund für die Ausdehnung der Informationspflicht und des Auskunftsrechts bei vollständig automatisierten Entscheidungen angeführt wird, erscheint es sinnvoll, eine allfällige Ausdehnung zumindest sachlich einzuschränken. Statt auf den ebenso diffusen wie weiten Begriff «teilautomatisiert» abzustellen, könnte bei einer allfälligen Revision von Art. 21 DSG der Begriff «ausschliesslich» durch die Wendung «im Wesentlichen» ersetzt werden. Eine besondere Informationspflicht und ein besonderes Auskunftsrecht würden dann bei Entscheidungen bestehen, die «im Wesentlichen auf einer automatisierten Bearbeitung beruhen». Anders als heute würden damit auch Entscheidungen erfasst, bei denen zwar ein Mensch beteiligt ist, bei denen dieser aber keinen relevanten Beitrag leistet, weil sich die menschliche Beteiligung bspw. auf eine blosser Überprüfung und Absegnung der maschinell getroffenen Entscheidung beschränkt. Zugleich würde eine solche Formulierung sicherstellen, dass der Anwendungsbereich der Bestimmungen auf diejenigen Fälle beschränkt wäre, die der Gesetzgeber (wohl) als problematisch angesehen hatte.

[31] Die bestehende Informationspflicht bei vollständig automatisierten Entscheidungen wird von den Unternehmen, soweit ersichtlich, meist durch allgemeine Hinweise in Datenschutzerklärungen umgesetzt, nicht durch eine spezifische Information im Einzelfall.⁴⁸ Da das geltende Recht für Privatpersonen keine Vorgaben enthält, wie sie die Betroffenen über eine vollständig automatisierte Entscheidung informieren müssen,⁴⁹ erscheint dieses Vorgehen durchaus zulässig. Es wirft aber die Frage auf, ob sich der Zweck der Informationspflicht über vollständig oder im Wesentlichen automatisierte Entscheidungen auf diese Weise erfüllen lässt, zumal hinlänglich bekannt und empirisch belegt ist, dass die in Datenschutzerklärungen vermittelten Informationen von den meisten Betroffenen nicht zur Kenntnis genommen werden.⁵⁰ Das wirft grundlegende Fragen zu Sinn und Zweck von Informationspflichten auf, die hier allerdings nicht untersucht werden können. Geht man aber davon aus, dass sich nur vergleichsweise wenige Personen

⁴⁷ Zur Skepsis gegenüber automatisierten Entscheidungen siehe: SUTER (Fn. 45), 73; MARKUS CHRISTEN et al., Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz, Zürich 2020, 57. Zu Sinn und Zweck von Art. 21 DSG siehe: THOUVENIN/FRÜH (Fn. 39), 8 ff.; SUTER (Fn. 45), 207 ff.; HENSELER/VASELLA (Fn. 45), DSG 21 N 1 f.; OLIVIER HEUBERGER, Profiling im Persönlichkeits- und Datenschutzrecht der Schweiz, Zürich 2020, Rn. 189 ff.

⁴⁸ Siehe dazu bspw. die Datenschutzerklärung der SwissLife AG, abrufbar unter: <https://www.swisslife.ch/de/ueber-uns/footer/privacy/privatkunden/einzelversicherung.html>; Stand: 10. März 2026; und die Datenschutzerklärung der Swisscom AG, abrufbar unter: https://www.swisscom.ch/content/dam/swisscom/nl/rechtliches/res/datenschutzbestimmungen_de_clean_30-1-2026.pdf; Stand: 10. März 2026.

⁴⁹ Anderes gilt nach Art. 21 Abs. 4 DSG für Bundesorgane, die automatisierte Einzelentscheidungen als solche kennzeichnen müssen.

⁵⁰ BHANUKA SILVA et al., Entailment-Driven Privacy Policy Classification with LLMs, 25. September 2024, abrufbar unter: <https://arxiv.org/abs/2409.16621>; Stand: 10. März 2026; TOMÁ SIGMUND, Attention Paid to Privacy Policy Statements, Information 2021, 144 ff., abrufbar unter: <https://www.mdpi.com/2078-2489/12/4/144>; Stand: 10. März 2026; NILI STEINFELD, «I agree to the terms and conditions»: (How) do users read privacy policies online? An eye-tracking experiment, Computers in Human Behavior 2016, 992 ff., abrufbar unter: <https://www.sciencedirect.com/science/article/pii/S0747563215301692>; Stand: 10. März 2026.

dafür interessieren, ob eine sie betreffende Entscheidung vollständig oder im Wesentlichen automatisiert gefällt wurde, so ist eine allgemeine Information in Datenschutzerklärungen durchaus sinnvoll. Denn so können sich interessierte Personen leicht darüber informieren, ob ein Unternehmen vollständig oder im Wesentlichen automatisierte Entscheidungen trifft, und anschliessend entscheiden, ob sie gestützt auf ihr Auskunftsrecht Informationen darüber erhalten wollen, ob eine bestimmte sie betreffende Entscheidung vollständig oder im Wesentlichen automatisiert gefällt wurde, welche Personendaten dabei bearbeitet wurden und nach welcher Logik die Entscheidung gefällt wurde. Wie das allgemeine datenschutzrechtliche Auskunftsrecht könnte auch das besondere Auskunftsrecht von den Betroffenen jederzeit, kostenlos und ohne Angabe von Gründen geltend gemacht werden.⁵¹ Sinnvoll erschiene zudem, das in Art. 21 Abs. 2 DSG vorgesehene Recht auf Darlegung des eigenen Standpunkts und auf menschliche Überprüfung nur nach Ausübung des besonderen Auskunftsrechts zu gewähren und systematisch bei diesem Recht zu regeln.

4.5. Bearbeitung von Personendaten beim Training von KI-Modellen

[32] Die Bearbeitung von Personendaten für das Trainieren, Validieren und Testen von KI-Modellen verstösst regelmässig gegen einen oder mehrere Grundsätze der Datenbearbeitung, insb. gegen den Grundsatz der Erkennbarkeit (Art. 6 Abs. 2 f. DSG) und gegen den Grundsatz der Zweckbindung (Art. 6 Abs. 3 DSG). Damit stellt sich die Frage, ob diese Bearbeitungen durch Einwilligung der betroffenen Personen, durch ein überwiegendes privates oder öffentliches Interesse oder durch Gesetz gerechtfertigt werden können (Art. 31 Abs. 1 DSG). Da das Einholen einer Einwilligung in der Regel an der grossen Zahl der betroffenen Personen scheitern und eine gesetzliche Grundlage nur in wenigen Fällen (z.B. bei gesetzlichen Dokumentationspflichten) vorliegen wird, stellt sich regelmässig die Frage, ob die Bearbeitung durch ein überwiegendes Interesse des Verantwortlichen gerechtfertigt werden kann. Das wird meist der Fall sein.

[33] Das DSG sieht eine Reihe von Konstellationen vor, in denen ein überwiegendes Interesse des Verantwortlichen regelmässig in Betracht kommt (Art. 31 Abs. 2 DSG), unter anderem die Bearbeitungen von Personendaten für nicht personenbezogene Zwecke, insb. für Forschung, Planung oder Statistik (Art. 31 Abs. 2 lit. e DSG). Beim Trainieren von KI-Modellen werden Personendaten nicht bearbeitet, um Erkenntnisse über die betroffenen Personen zu gewinnen oder Folgen für diese Personen abzuleiten, sondern um ein Modell zu entwickeln, das eine bestimmte Aufgabe (oder eine Vielzahl unbestimmter Aufgaben) möglichst gut erfüllen kann. Die Bearbeitung ist damit nicht personenbezogen und lässt sich grundsätzlich durch das überwiegende Interesse des Verantwortlichen rechtfertigen.⁵² Die Rechtfertigung nach dem gesetzlichen Regelbeispiel (Art. 31 Abs. 2 lit. e DSG) setzt allerdings voraus, dass drei weitere Voraussetzungen erfüllt sind: Namentlich müssen die Daten so bald als möglich anonymisiert werden (Art. 31 Abs. 2 lit. e Ziff. 1 DSG), besonders schützenswerte Personendaten dürfen Dritten grundsätzlich nur so bekanntgegeben werden, dass die betroffenen Personen nicht identifiziert werden können (Art. 31 Abs. 2

⁵¹ Botschaft DSG, BBl 2017 6941, 7069; RALPH GRAMIGNA, in: Blechta/Vasella (Hrsg.), Basler Kommentar, Datenschutzgesetz/Öffentlichkeitsgesetz, 4. Aufl., Basel 2024, DSG 25 N 13; siehe dazu auch BGE 141 III 119 E. 7.1.1.

⁵² Ebenso THOMAS STEINER/CHRISTIAN LAUX, in: Bieri/Powell (Hrsg.), Orell Füssli Kommentar, Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlassen, Zürich 2023, DSG 31 N 37; ROLF H. WEBER, Künstliche Intelligenz und Datenschutz, Jusletter vom 4. Juli 2024, Rn. 19.

lit. e Ziff. 2 DSG), und die Ergebnisse dürfen nur so veröffentlicht werden, dass die betroffenen Personen nicht bestimmbar sind (Art. 31 Abs. 2 lit. e Ziff. 3 DSG). Diese Voraussetzungen werden sich bei der Bearbeitung von Personendaten für das Trainieren, Validieren und Testen von KI-Modellen in der Regel erfüllen lassen, obwohl sie auf diese Konstellation nicht wirklich passen.⁵³ Das gilt namentlich für die dritte Voraussetzung, die besonders deutlich auf die Bearbeitung von Personendaten für Forschung, Planung oder Statistik ausgerichtet ist. Für das Training von KI-Modellen passt diese Voraussetzung nicht wirklich, weil Personendaten in KI-Modellen nicht im Sinne dieser Vorschrift veröffentlicht werden. Da es allerdings durchaus möglich ist, einem trainierten KI-Modell Personendaten zu entnehmen, könnte das Veröffentlichen eines Modells als Veröffentlichung von Personendaten verstanden werden.⁵⁴ Ein Veröffentlichen sollte aber nur angenommen werden, wenn es der Verantwortliche unterlassen hat, geeignete Massnahmen zu ergreifen, um sicherzustellen, dass sich die Personendaten dem Modell möglichst gar nicht oder jedenfalls nur mit einem derart grossen Aufwand entnehmen lassen, dass nach der allgemeinen Lebenserfahrung nicht damit zu rechnen ist, dass ein Dritter diesen Aufwand auf sich nehmen wird.⁵⁵ In aller Regel wird sich der Verstoß gegen die Grundsätze der Datenbearbeitung damit durch das Regelbeispiel der Bearbeitung für nicht personenbezogene Zwecke (Art. 31 Abs. 2 lit. e DSG) rechtfertigen lassen.

[34] Eine Rechtfertigung ist allerdings auch möglich, wenn man zum Schluss kommen sollte, dass die Voraussetzungen des Regelbeispiels nicht erfüllt sind. Denn ein überwiegendes Interesse kann ohne weiteres auch dann bejaht werden, wenn die Bearbeitung nicht unter ein gesetzlich vorgesehenes Regelbeispiel fällt (Art. 31 Abs. 1 DSG). Da die Bearbeitung von Personendaten beim Trainieren, Validieren und Testen eines KI-Modells für die betroffenen Personen keinerlei Konsequenzen hat, dürfte das Interesse des Verantwortlichen an der Bearbeitung das Interesse der betroffenen Personen an der Nichtbearbeitung in aller Regel überwiegen.

[35] Auch wenn die Rechtslage nach dem geltenden DSG recht klar ist, wäre es wegen der grossen Bedeutung der Bearbeitung von Personendaten beim Training von KI-Modellen und mit Blick auf die Rechtssicherheit sinnvoll, die Frage ausdrücklich gesetzlich zu regeln. Dazu sollte der Gesetzgeber die Liste der Regelbeispiele für das Vorliegen eines überwiegenden Interesses des Verantwortlichen (Art. 31 Abs. 2 DSG) um ein weiteres Regelbeispiel ergänzen, das die Bearbeitung von Personendaten für das Trainieren, Validieren und Testen von KI-Modellen erfasst. Voraussetzung für die Annahme eines überwiegenden Interesses sollte dabei sein, dass der Verantwortliche hinreichende Massnahmen getroffen hat, um zu verhindern, dass dem trainierten Modell die beim Training verwendeten Personendaten entnommen werden können. Die EU sieht im Vorschlag für die sog. «Digital-Omnibus-Verordnung»⁵⁶ eine ähnliche Ergänzung der DSGVO vor. Namentlich soll in einem neuen Artikel (Art. 88c DSGVO) klargestellt werden, dass die Bearbeitung von

⁵³ Näheres dazu: FLORENT THOUVENIN, Datenschutzrechtliche Herausforderungen bei der Entwicklung von KI-Modellen, ZSR 2025, 343 ff.

⁵⁴ Siehe dazu: DAVID ROSENTHAL, Datenschutz und KI: Worauf in der Praxis zu achten ist, Jusletter IT vom 22. April 2022, Rn. 48 ff.; WEBER (Fn. 52), Rn. 19.

⁵⁵ THOUVENIN/VOLZ (Fn. 10), 632; DANIEL W. SEILER/MARCEL GRIESINGER, Spannungsfeld Künstliche Intelligenz (KI) und Datenschutzrecht, Jusletter vom 25. September 2023, Rn. 10.

⁵⁶ Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Änderung der Verordnungen (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 und der Richtlinien 2002/58/EG, (EU) 2022/2555 und (EU) 2022/2557 hinsichtlich der Vereinfachung des digitalen Rechtsrahmens und zur Aufhebung der Verordnungen (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868 und der Richtlinie (EU) 2019/1024 (Digital-Omnibus-Verordnung), abrufbar unter: <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52025PC0837>.

Personendaten im Zusammenhang mit der Entwicklung und dem Betrieb von KI-Systemen und KI-Modellen auf der Grundlage eines überwiegenden Interesses im Sinn von Art. 6 Abs. 1 Bst. f DSGVO erfolgen kann. Zudem ist eine neue Ausnahme vom Verbot der Bearbeitung von besonderen Kategorien personenbezogener Daten (Art. 9 Abs. 1 DSGVO) vorgesehen, wenn diese Daten im Zusammenhang mit der Entwicklung und dem Betrieb von KI-Systemen und KI-Modellen bearbeitet werden (Art. 9 Abs. 2 Bst. k DSGVO); die Ausnahme greift allerdings nur, wenn gewisse Bedingungen erfüllt sind (Art. 9 Abs. 5 DSGVO).

5. Erkenntnisse

[36] Der Bundesrat hat am 12. Februar 2025 entschieden, wie das Schweizer Recht auf die Herausforderungen reagieren soll, die sich bei der Entwicklung und Verwendung von Künstlicher Intelligenz (KI) ergeben. Dieser strategische Entscheid sieht vor, dass die Konvention des Europarats zu Künstlicher Intelligenz (KI-Konvention) ratifiziert und die dafür notwendigen Anpassungen im Schweizer Recht vorgenommen werden. Die Gesetzesanpassungen sollen dabei grundsätzlich sektorbezogen und themenspezifisch erfolgen, eine sektorübergreifende Regelung ist nur in zentralen, grundrechtsrelevanten Bereichen vorgesehen, insb. im Datenschutz. Das Bundesamt für Justiz und das BAKOM werden nun bis Ende 2026 eine Vernehmlassungsvorlage erarbeiten. Im Vordergrund stehen dabei Regelungen in den Bereichen Transparenz, Datenschutz, Nichtdiskriminierung und Aufsicht.

[37] Dieser Beitrag untersucht den Handlungsbedarf und die Handlungsoptionen für Regelungen gegenüber Privatpersonen in den inhaltlich verschränkten Bereichen Transparenz und Datenschutz. Den Ausgangspunkt bilden die Vorgaben der KI-Konvention, die den Vertragsstaaten bei der Umsetzung gegenüber Privatpersonen viel Spielraum belässt. Hinzu kommt, dass die KI-Konvention keine umfassenden Vorgaben zur Verwendung von KI-Systemen macht, sondern lediglich sicherstellen will, dass Aktivitäten innerhalb des Lebenszyklus' von KI-Systemen vollständig mit Grundrechten, Demokratie und Rechtsstaatlichkeit vereinbar sind. Da Privatpersonen in der Schweiz grundsätzlich nicht an die Grundrechte gebunden sind und die Gewährleistung der Rechtsstaatlichkeit Aufgabe des Staates ist, müssen die Vorgaben der KI-Konvention gegenüber Privatpersonen in der Schweiz nur umgesetzt werden, wenn sich deren Tätigkeit auf das Funktionieren der Demokratie auswirken kann. Dem Schweizer Gesetzgeber steht es aber natürlich frei, weitere Fragen, die sich bei der Entwicklung und Verwendung von KI stellen, ebenfalls zu regeln.

[38] Vor diesem Hintergrund kommt der Beitrag zum Schluss, dass in fünf Bereichen Handlungsbedarf besteht:

[39] Erstens sollte durch eine geeignete Regelung sichergestellt werden, dass die Nutzer:innen von grossen und damit demokratierelevanten Kommunikationsplattformen und Suchmaschinen nachvollziehen können, weshalb ihnen ein bestimmter Inhalt angezeigt wird. Diese Vorgabe zur *Transparenz von Empfehlungssystemen* («*recommender systems*») ergibt sich aus der allgemeinen Bestimmung zur Gewährleistung von Transparenz bei der Verwendung von KI-Systemen, sollte aber nicht auf die Verwendung dieser Technologie beschränkt werden. Diese Vorgabe der KI-Konvention kann durch eine Regelung im vorgesehenen *Bundesgesetz über Kommunikationsplattformen und Suchmaschinen* (KomPG) umgesetzt werden, die grosse Kommunikationsplattformen und Suchmaschinen verpflichtet, in ihren allgemeinen Geschäftsbedingungen (AGB) die wich-

tigsten Parameter, die ihren Empfehlungssystemen zugrunde liegen, und deren Gewichtung darzulegen (Art. 18 Abs. 1 lit. a VE-KomPG).

[40] Zweitens sollten die *grossen Kommunikationsplattformen* verpflichtet werden, *politisch oder gesellschaftlich relevante nicht-authentische Inhalte zu kennzeichnen*. Gemeint sind damit namentlich (aber nicht nur) Inhalte, die mithilfe von generativer KI erstellt wurden. Auf eine umfassende Kennzeichnungspflicht für KI-generierte Inhalte sollte dagegen verzichtet werden, weil die meisten Menschen schon heute (und erst recht in Zukunft) nicht mehr ohne weiteres annehmen (werden), dass ein bestimmter Inhalt von einem Menschen geschaffen wurde. Ein relevantes Risiko für die Gefährdung des Funktionierens der Demokratie besteht jedenfalls nur bei politisch oder gesellschaftlich relevanten nicht-authentischen Inhalten. Für diese Inhalte erscheint eine Kennzeichnungspflicht wegen der potenziell gravierenden Konsequenzen einer erfolgreichen Täuschung durchaus sinnvoll – und von der KI-Konvention wohl auch vorgegeben. Zur Umsetzung einer solchen Pflicht müssten die Anbieter der grossen Kommunikationsplattformen ihre Nutzer:innen in den Nutzungsbedingungen verpflichten, nicht-authentische politisch oder gesellschaftlich relevante Inhalte zu kennzeichnen. Die Kommunikationsplattformen sollten allerdings nicht zur Überwachung von Inhalten oder zur aktiven Nachforschung verpflichtet werden. Zum Schutz des Funktionierens der Demokratie dürfte es vielmehr genügen, wenn die Betreiber von Plattformen verpflichtet würden, auf Hinweis von Nutzer:innen oder von Dritten zu prüfen, ob ein auf ihrer Plattform zugänglich gemachter politisch oder gesellschaftlich relevanter Inhalt authentisch ist, und diesen Inhalt gegebenenfalls als nicht-authentisch zu kennzeichnen. Zur Umsetzung der Vorgabe der KI-Konvention über die Kennzeichnung KI-generierter Inhalte sollte das vorgesehene KomPG deshalb um eine Bestimmung ergänzt werden, welche die grossen Kommunikationsplattformen verpflichtet, politisch oder gesellschaftlich relevante nicht-authentische Inhalten zu kennzeichnen, wenn sie von ihren Nutzer:innen oder von Dritten auf solche Inhalte hingewiesen werden.

[41] Drittens sollte durch eine geeignete Regelung sichergestellt werden, dass *Menschen erkennen können, wenn sie mit einer Maschine statt mit einem Menschen interagieren*. Die KI-Konvention enthält eine Vorgabe zur Transparenz bei Interaktionen von Menschen mit KI-Systemen. Diese Vorgabe sollte im Schweizer Recht allerdings nicht als solche, sondern durch eine technologie-neutrale Bestimmung umgesetzt werden, die jede Interaktion von Menschen mit Maschinen erfasst. In vielen Fällen wird für die betroffenen Personen ohne weiteres ersichtlich sein, dass sie mit einer Maschine und nicht mit einem Menschen interagieren. Eine aktive Information ist deshalb nur erforderlich, wenn die geforderte Transparenz nicht bereits durch die Ausgestaltung der Interaktion sichergestellt werden kann. Da das DSG bereits eine Pflicht zur Information über automatisierte Entscheidungen kennt (Art. 21 DSG), drängt es sich auf, eine allgemeine Vorgabe zur Transparenz bei einer Interaktion von Menschen mit einer Maschine ebenfalls im DSG zu regeln.

[42] Viertens sollte die bestehende *Informationspflicht bei vollständig automatisierten Entscheidungen* (Art. 21 DSG) und das entsprechende Auskunftsrecht (Art. 25 Abs. 2 lit. f DSG) auf «im Wesentlichen» automatisierte Entscheidungen ausgedehnt werden. Die KI-Konvention enthält dazu zwar keine Vorgabe, die Frage wird in der Schweiz aber seit einiger Zeit diskutiert, wobei meist eine Ausdehnung auf «teilautomatisierte» Entscheidungen gefordert wird. Eine solche Pflicht ginge allerdings zu weit, weil die meisten Entscheidungen in Unternehmen schon heute (und erst recht in Zukunft) teilweise automatisiert gefällt werden und eine derart weit gefasste Informationspflicht auch viele Fälle erfassen würde, die offensichtlich unproblematisch sind. Ohnehin ist unklar, ob die betroffenen Personen überhaupt ein Informationsbedürfnis haben und ob sie die

heute gesetzlich vorgeschriebene Information zur Kenntnis nehmen und als relevant und nützlich erachten. Diese Fragen müssten vor einer weitgehenden Ausdehnung der Informationspflicht auf «teilautomatisierte» Entscheide empirisch untersucht und geklärt werden. Trotz grundsätzlicher Zweifel am Sinn einer Informationspflicht über automatisierte Entscheidungen erscheint eine massvolle Ausdehnung der Pflicht durchaus sinnvoll, weil die heutige Regelung Entscheidungen nicht erfasst, bei denen sich die Beteiligung eines Menschen auf eine blosser Überprüfung und Absegnung einer maschinell getroffenen Entscheidung beschränkt. Dafür reicht es aber aus, die Informationspflicht auf Fälle auszudehnen, in denen eine Entscheidung «im Wesentlichen» automatisiert gefällt wurde.

[43] Fünftens sollte im DSG die Liste der Regelbeispiele für das Vorliegen eines überwiegenden Interesses an einer Datenbearbeitung um ein *weiteres Regelbeispiel* ergänzt werden, das klarstellt, dass *die Interessen des Verantwortlichen gegenüber den Interessen der betroffenen Personen in der Regel überwiegen, wenn Personendaten für das Trainieren, Validieren und Testen eines KI-Modells verwendet werden*. Zu diesem Ergebnis gelangt man zwar bereits heute aufgrund der Auslegung und Anwendung des Regelbeispiels zur Bearbeitung von Personendaten für nicht personenbezogene Zwecke, insb. für Forschung, Planung oder Statistik (Art. 31 Abs. 2 lit. e DSG). Angesichts der grossen Bedeutung der Bearbeitung von Personendaten beim Trainieren, Validieren und Testen von KI-Modellen und mit Blick auf die Rechtssicherheit wäre es aber sinnvoll, die Frage in einem neuen Regelbeispiel ausdrücklich gesetzlich zu regeln.

FLORENT THOUVENIN Prof. Dr. iur., Rechtsanwalt, Ordinarius für Informations- und Kommunikationsrecht und Vorsitzender des Leitungsausschusses des Center for Information Technology, Society, and Law (ITSL) an der Universität Zürich.

Der Autor dankt Dr. Stephanie Volz, wissenschaftliche Geschäftsführerin des ITSL, für wertvolle Inputs und Aleksandra Marchewka, MLaw, für die Unterstützung bei der Recherche und beim Erstellen der Fussnoten.