

Datenschutzrechtliche Herausforderungen bei der Entwicklung von KI-Modellen

FLORENT THOUVENIN*

Schlagwörter: Künstliche Intelligenz, Training von KI-Modellen, Datenbearbeitung, Datenschutz

A. Einleitung

Bei der Entwicklung und Verwendung von Modellen und Systemen der Künstlichen Intelligenz (KI) werden sehr grosse Datenmengen benötigt. In vielen Fällen handelt es sich dabei um Personendaten. Die Bearbeitung dieser Daten wird durch das Datenschutzgesetz (DSG) umfassend normiert. Damit stellt sich die Frage, wie die Bestimmungen des DSG ausgelegt und angewendet werden können, um die mit der Entwicklung und Verwendung von KI verbundenen Herausforderungen zu erfassen. Bei der rechtlichen Analyse ist zwischen der Bearbeitung von Personendaten bei der Entwicklung von KI-Modellen und der Bearbeitung beim Anbieten und bei der Verwendung von KI-Systemen zu unterscheiden.

Die Entwicklung von KI-Modellen lässt sich in drei Schritte unterteilen: das Trainieren, Validieren und Testen. Im ersten Schritt lernt das System aus den Trainingsdaten, bestimmte Aufgaben zu erfüllen. Resultat des Trainings ist das KI-Modell, das eine Sammlung von Korrelationen darstellt, die für Vorhersagen oder das Treffen von Entscheidungen nützlich sind.¹ Durch das Validieren wird die Qualität des Modells überprüft, insbesondere soll damit eine Überanpassung (*overfitting*) an die Trainingsdaten verhindert werden.² Nach dem Training

* Prof. Dr. iur., Rechtsanwalt, Ordinarius für Informations- und Kommunikationsrecht und Vorsitzender des Leitungsausschusses des Center for Information Technology, Society, and Law (ITSLS) an der Universität Zürich. Der Autor dankt ALEKSANDRA MARCHEWKA, BLaw, für das Erstellen eines Entwurfs für den Text dieses Beitrags auf der Grundlage von FLORENT THOUVENIN/STEPHANIE VOLZ, «Datenschutz», White Paper, Juni 2024, sowie für die Recherchen und das Erstellen des Fussnotenapparats.

1 DAVID LEHR/PAUL OHM, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, UC Davis Law Review, Dezember 2017, S. 653–717, S. 671 und 695 f.

2 FLORENT THOUVENIN/STEPHANIE VOLZ/SORAYA WEINER/CHRISTOPH HEITZ, *Diskriminierung beim Einsatz von Künstlicher Intelligenz (KI)*, Jusletter IT vom 4. Juli 2024, Rn. 53.

wird die Leistungsfähigkeit des Modells mit Testdaten geprüft. Diese haben dieselbe Struktur wie die Trainings- und Validierungsdaten, wurden aber nicht für das Trainieren und Validieren benutzt.³ In der Folge wird nur der Begriff des Trainings verwendet, der für die Zwecke dieses Beitrags das eigentliche Trainieren, das Validieren und das Testen umfasst.

Der Beitrag fokussiert auf die datenschutzrechtlichen Herausforderungen bei der Entwicklung von KI-Modellen, die seit einiger Zeit intensiv diskutiert werden. Er konzentriert sich dabei auf drei zentrale Aspekte: die Grundsätze der Datenbearbeitung (B.), die Rechtfertigung einer Verletzung eines Grundsatzes (C.) sowie die Rechte der betroffenen Personen und die Pflichten der Verantwortlichen (D.).

B. Grundsätze der Datenbearbeitung

Bei der Bearbeitung von Personendaten sind die Grundsätze der Datenbearbeitung (Art. 6 DSGVO) einzuhalten. Das gilt auch beim Entwickeln von KI-Modellen. Herausforderungen ergeben sich vor allem bei den Grundsätzen der Erkennbarkeit (Art. 6 Abs. 2 f. DSGVO), der Zweckbindung (Art. 6 Abs. 3 DSGVO) und der Verhältnismässigkeit (Art. 6 Abs. 2 DSGVO), dort namentlich bei der Datenminimierung und bei der Speicherbegrenzung (Art. 6 Abs. 4 DSGVO).

I. Grundsatz der Erkennbarkeit

Nach dem Grundsatz der Erkennbarkeit muss das Beschaffen und Bearbeiten von Personendaten für die betroffenen Personen erkennbar sein.⁴ Für das Training von KI-Modellen werden regelmässig Daten aus einer Vielzahl unterschiedlicher Quellen beschafft.⁵ Da die Daten kaum je direkt bei den betroffenen Personen beschafft werden, ist das Beschaffen und Bearbeiten der Personendaten für diese in der Regel nicht erkennbar.⁶

3 LEHR/OHM (Fn. 1), S. 685.

4 Botschaft des Bundesrates zum Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz, BBl 2017 6941, 7025; FLORENT THOUVENIN/NADIA BRAUN BINDER, Transparenz durch Datenschutzerklärungen von Behörden, ZSR 2022 I, S. 5–29, S. 16; RETO FANGER, in: Adrian Bieri/Julian Powell, Orell Füssli Kommentar, Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlässen, Zürich 2023, Art. 6 N 7.

5 LUKAS BÖHLMANN/MICHAEL REINLE, in: Gabor-Paul Blechta/David Vasella (Hrsg.), Basler Kommentar, Datenschutzgesetz, Öffentlichkeitsgesetz, 4. Aufl., Basel 2024, Art. 6 N 418.

6 DAVID ROSENTHAL/LIVIO VERALDI, Das Training von KI-Sprachmodellen mit fremden Inhalten und Daten aus rechtlicher Sicht, Jusletter vom 3. Februar 2025, Rn. 63.

II. Grundsatz der Zweckbindung

Der Grundsatz der Zweckbindung verlangt, dass der Zweck der Beschaffung der Personendaten für die betroffene Person erkennbar ist (Art. 6 Abs. 3 Teilsatz 1 DSGVO). Zudem dürfen die Daten nur dann zu einem anderen Zweck bearbeitet werden, wenn dieser mit dem ursprünglichen Zweck vereinbar ist (Art. 6 Abs. 3 Teilsatz 2 DSGVO). Das ist nicht der Fall, wenn die Weiterbearbeitung der Daten unerwartet oder unangebracht ist.⁷

Beim Training von KI-Modellen werden Personendaten regelmässig für einen anderen Zweck bearbeitet als den, für den sie ursprünglich beschafft worden sind. Damit stellt sich die Frage, ob das Bearbeiten von Personendaten für das Training von KI-Modellen als neuer Zweck zu verstehen ist, der mit dem ursprünglichen Zweck der Bearbeitung nicht vereinbar ist. Das ist regelmässig der Fall, wenn das KI-Modell einem völlig anderen Zweck dient als der Zweck, zu dem die Daten beschafft worden sind, etwa wenn auf Social Media zugänglich gemachte Fotos verwendet werden, um eine Gesichtserkennungssoftware zu trainieren.⁸ Dasselbe gilt, wenn Daten für das Training eines KI-Modells verwendet werden, das für beliebige Zwecke eingesetzt werden kann (sog. «foundation model» oder «general-purpose AI model»). Kein Verstoß gegen den Grundsatz der Zweckbindung liegt dagegen vor, wenn Röntgenbilder von Knochenbrüchen verwendet werden, um ein KI-Modell zu trainieren, das Knochenbrüche auf Röntgenbildern erkennen kann.⁹

III. Grundsatz der Verhältnismässigkeit

Beim Training von KI-Modellen werden regelmässig sehr grosse Mengen an Daten bearbeitet.¹⁰ Denn grundsätzlich gilt: Je mehr Daten beim Training verwendet werden, desto besser sind die Ergebnisse des Modells. Der Bearbeitung möglichst vieler Daten steht – zumindest auf den ersten Blick – der Grundsatz der Datenminimierung entgegen. Danach dürfen nur diejenigen Daten verwendet werden, die für den verfolgten Zweck erforderlich sind;¹¹ teilweise wird gar gefordert, dass die Daten «absolut notwendig» sein müs-

7 Botschaft DSGVO (Fn. 4), S. 7025.

8 FLORENT THOUVENIN/STEPHANIE VOLZ, Ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz (KI) in der Schweiz, ZBJV 2024, S. 613–638, S. 631.

9 THOUVENIN/VOLZ (Fn. 8), S. 631; siehe dazu auch SAMUEL KLAUS, KI trifft Datenschutz – Risiken und Lösungsansätze, in: Astrid Epiney/Sophia Rovelli (Hrsg.), Künstliche Intelligenz und Datenschutz, Zürich 2021, S. 81–95, S. 82.

10 Siehe dazu auch CARLO PLITZ/ALEXANDER WEISS, Berechtigte Interessen als Rechtsgrundlage für das Training von KI-Modellen, DSB 2024, S. 291–296, S. 291.

11 BSK DSGVO-BÖHLMANN/REINLE (Fn. 5), Art. 6 N 53; STEPHANIE VOLZ, KI Sandboxen für die Schweiz?, SZW 2022, S. 51–68, S. 56; siehe dazu auch DAMIAN GEORGE, Prinzipien und Rechtmässigkeitsbedingungen im privaten Datenschutzrecht, Zürich 2021, S. 309 f.

sen.¹² Der Grundsatz der Datenminimierung kann hier aber auch so ausgelegt werden, dass kein Verstoß vorliegt, weil die Verwendung einer möglichst grossen Datenmenge für den Zweck des Trainings eines KI-Modells erforderlich oder gar notwendig ist.¹³

Dasselbe gilt für den Grundsatz der Speicherbegrenzung, welcher die Datenbearbeitung in zeitlicher Hinsicht beschränkt, indem er vorgibt, dass Personendaten nur so lange aufbewahrt werden dürfen, wie sie für den Zweck der Bearbeitung erforderlich sind (Art. 6 Abs. 4 DSGVO). Dabei sind zwei Fragen zu unterscheiden: Zum einen können Personendaten in den trainierten Modellen enthalten sein. In gewissen Fällen ist das gerade das Ziel des Trainings, bspw. bei einem Modell, das die Stimme einer natürlichen Person nachahmen soll. Meist sollen die beim Training verwendeten Personendaten aber nicht im trainierten KI-Modell gespeichert werden.¹⁴ Dennoch werden regelmässig Sequenzen aus den Trainingsdaten in KI-Systemen memorisiert, sodass es in vielen Fällen möglich ist, den trainierten Modellen mit einem gewissen Aufwand zumindest gewisse Personendaten zu entnehmen.¹⁵ Zum andern werden die beim Training verwendeten Personendaten regelmässig nach Abschluss der Entwicklung in separaten Datenbanken gespeichert,¹⁶ um ein späteres Nachtrainieren und Testen zu ermöglichen.¹⁷ Eine Speicherung kann auch aus anderen Gründen erfolgen, bspw. aufgrund von Vorgaben in Codes of Practice oder in tech-

12 PHILIPPE MEIER/NICOLAS TSCHUMY, in: Philippe Meier/Sylvain Météille (Hrsg.), *Commentaire Romand, Loi fédérale sur la protection des données*, Basel 2023, Art. 6 N 33; BSK DSGVO-BÜHLMANN/REINLE (Fn. 5), Art. 6 N 53.

13 FLORENT THOUVENIN/STEPHANIE VOLZ, «Datenschutz», White Paper, Juni 2024, S. 4; ROLF H. WEBER, *Künstliche Intelligenz und Datenschutz*, Jusletter vom 4. Juli 2024, Rn. 17.

14 ROSENTHAL/VERALDI (Fn. 6), Rn. 7; PHILIPP HACKER, *A legal framework for AI training data – from first principles to the Artificial Intelligence Act*, in: *Law, Innovation and Technology*, 13:2, S. 257–301, S. 277; siehe auch *Datatilsynet – The Norwegian Data Protection Authority, Artificial intelligence and privacy, Report*, January 2018, S. 10.

15 NICOLAS CARLINI/CHANG LIU/ULFAR ERLINGSSON/JERNEJ KOS/DAWN SONG, *The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks*, 28th USENIX Security Symposium 2019, S. 267 ff.; NICOLAS CARLINI/FLORIAN TRAMÈR/ERIC WALLACE/MATTHEW JAGIELSKI/ARIEL HERBERT-VOSS/KATHERINE LEE/ADAM ROBERTS/TOM BROWN/DAWN SONG/ULFAR ERLINGSSON/ALINA OPREA/COLIN RAFFEL, *Extracting Data from Large Language Models*, 30th USENIX Security Symposium 2021, S. 2633 ff.

16 THOUVENIN/VOLZ (Fn. 13), S. 4; AURÉLIEN GÉRON, *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems*, 2. Aufl., Sebastopol 2019, S. 46 ff.; ausführlich MARK ZHAO/NIKET AGARWAL/AARTI BASANT/BURA GEDIK/SATADRU PAN/MUSTAFA OZDAL/RAKESH KOMURAVELLI/JERRY PAN/TIANSHU BAO/HAOWEI LU/SUNDARAM NARAYANAN/JACK LANGMAN/KEVIN WILFONG/HARSHA RAS-TOGI/CAROLE-JEAN WU/CHRISTOS KOZYRAKIS/PARIK POL, *Understanding Data Storage and Ingestion for Large-Scale Deep Recommendation Model Training*, in: *Valentina Salapura, Proceedings of the 49th Annual International Symposium on Computer Architecture, ISCA '22*, New York 2022, S. 1042–1057, S. 1045.

17 ROSENTHAL/VERALDI (Fn. 6), Rn. 7 und 63.

nischen Standards¹⁸ oder zur Sicherung von Beweismitteln für allfällige Rechtsstreitigkeiten.¹⁹ Das Speichern der Daten ist damit für den Zweck der Bearbeitung regelmässig erforderlich und in der Regel nicht als Verstoß gegen den Grundsatz der Speicherbegrenzung zu qualifizieren.

C. Rechtfertigung

Die Ausführungen zu den Grundsätzen der Datenbearbeitung haben gezeigt, dass die Bearbeitung von Personendaten für das Training von KI-Modellen regelmässig gegen einen oder mehrere Grundsätze verstösst. Verstösse gegen die Bearbeitungsgrundsätze werden im DSGVO als Persönlichkeitsverletzungen qualifiziert (Art. 30 Abs. 2 lit. a DSGVO). Diese Persönlichkeitsverletzungen können allerdings durch Einwilligung der betroffenen Person, durch ein überwiegendes privates oder öffentliches Interesse oder durch das Gesetz gerechtfertigt werden (Art. 31 Abs. 1 DSGVO). Das Speichern der Daten kann zwar durch Gesetz gerechtfertigt sein, wenn gesetzliche Dokumentationspflichten bestehen, das eigentliche Training wird sich aber kaum je auf eine gesetzliche Grundlage stützen lassen. Die Rechtfertigung durch Einwilligung wird beim Training von KI-Modellen wegen der grossen Zahl der betroffenen Personen und der sehr grossen Menge an Daten sowie der verwendeten Quellen (bspw. Websites, Social Media Posts etc.) kaum eine Rolle spielen.²⁰ Im Vordergrund steht damit der Rechtfertigungsgrund der überwiegenden Interessen des Verantwortlichen (Art. 31 Abs. 1 f. DSGVO).²¹

Die Interessen des Verantwortlichen und der betroffenen Personen werden im DSGVO grundsätzlich frei und für die jeweilige Bearbeitung gegeneinander abgewogen.²² Für gewisse, besonders relevante Konstellationen hat der Gesetzgeber das Ergebnis der Interessenabwägung allerdings antizipiert (Art. 31 Abs. 2 DSGVO).²³ In diesen Fällen liegt regelmässig ein überwiegendes Interesse

18 THOUVENIN/VOLZ (Fn. 13), S. 4; ATENAS ANDERSON/BENEDETTA S. GALETTI, *La conservation des données personnelles: comment déterminer sa durée?*, sic! 2021, S. 103–119, S. 112 ff.; BSK DSGVO-BÜHLMANN/REINLE (Fn. 5), Art. 6 N 222.

19 Weiterführend und mit Bsp.: ANDERSON/GALETTI (Fn. 18), S. 112 ff.; ADRIAN BIERI/JULIAN POWELL, *Die Totalrevision des Bundesgesetzes über den Datenschutz, Übersicht der wichtigsten Neuerungen für Unternehmen*, Jusletter vom 16. November 2020, Rn. 5; BSK DSGVO-BÜHLMANN/REINLE (Fn. 5), Art. 6 N 222; ROSENTHAL/VERALDI (Fn. 6), Rn. 63.

20 WEBER (Fn. 13), Rn. 17; DAVID ROSENTHAL, *Datenschutz beim Einsatz generativer künstlicher Intelligenz*, Jusletter vom 6. November 2023, Rn. 34.

21 THOUVENIN/VOLZ (Fn. 8), S. 631.

22 MATTHIAS R. SCHÖNBÄCHLER, *Zum neuen Schweizer Datenschutzrecht*, ZBJV 2023, S. 171–195, S. 178; MONIKA PFÄFFINGER, in: Bruno Baeriswyl/Kurt Pärli/Dominika Blonski (Hrsg.), *Stämpfli's Handkommentar, Datenschutzgesetz*, 2. Aufl., Bern 2023, Art. 31 N 47 f.

23 JOSÉPHINE BOILLAT/STÉPHANE WERLY, in: Philippe Meier/Sylvain Météille (Hrsg.), *Commentaire Romand, Loi fédérale sur la protection des données*, Basel 2023, Art. 31 N 30.

des Verantwortlichen vor.²⁴ Diese Konstellationen umfassen unter anderem die Rechtfertigung von Bearbeitungen für nicht personenbezogene Zwecke, insbesondere für Forschung, Planung und Statistik (Art. 30 Abs. 2 lit. e DSGVO). Forschung, Planung und Statistik werden dabei exemplarisch genannt, der Rechtfertigungsgrund ist aber schon nach dem Wortlaut weiter und umfasst die Bearbeitung für nicht personenbezogene Zwecke allgemein.²⁵ Ein Personenbezug fehlt, wenn die Identität der betroffenen Personen für die Datenbearbeitung irrelevant ist und die Personen nicht (re-)identifizierbar sind.²⁶ Zwischen dem Zweck des Bearbeitens der Personendaten und der betroffenen Person darf also kein Zusammenhang bestehen.²⁷ Beim Training von KI-Modellen werden die Personendaten nicht bearbeitet, um Erkenntnisse über eine (oder mehrere) betroffene Person(en) zu gewinnen oder Folgen für diese Personen abzuleiten. Das Ziel ist vielmehr die Entwicklung eines Modells, das eine bestimmte Aufgabe (oder eine Vielzahl von unbestimmten Aufgaben) möglichst gut erledigen kann.²⁸ Die Bearbeitung ist damit nicht personenbezogen.²⁹

Neben dem zentralen Aspekt des fehlenden Personenbezugs enthält der Rechtfertigungsgrund drei weitere Voraussetzungen:

- (i) Der Verantwortliche muss die Daten anonymisieren, sobald es der Bearbeitungszweck erlaubt. Ist die Anonymisierung unmöglich oder erfordert sie einen unverhältnismässigen Aufwand, muss der Verantwortliche angemessene Massnahmen ergreifen, um die Bestimmbarkeit der betroffenen Personen zu verhindern (Art. 31 Abs. 2 lit. e Ziff. 1 DSGVO). Beim Training von KI-Modellen wird die verlangte Anonymisierung in der Regel mit geeigneten technischen Massnahmen erreicht.³⁰
- (ii) Bei der Bearbeitung von besonders schützenswerten Personendaten dürfen diese Dritten nur so bekannt gegeben werden, dass die betroffene Person nicht bestimmbar ist; ist dies nicht möglich, ist zu gewährleisten, dass die Dritten die Daten nur zu nicht personenbezogenen Zwecken bearbeiten (Art. 31 Abs. 2 lit. e Ziff. 2 DSGVO). Auch diese Voraussetzung wird in den meisten Fällen keine relevante Hürde sein.³¹

24 CORRADO RAMPINI/REHANA C. HARASGAMA, in: Gabor-Paul Blechta/David Vasella (Hrsg.), Basler Kommentar, Datenschutzgesetz, Öffentlichkeitsgesetz, 4. Aufl., Basel 2024, Art. 31 N 27 f.

25 VOLZ (Fn. 11), S. 57.

26 BSK DSGVO-RAMPINI/HARASGAMA (Fn. 24), Art. 31 N 60; GABRIEL KASPAR, People Analytics in privatrechtlichen Arbeitsverhältnissen, Vorschläge zur wirksameren Durchsetzung des Datenschutzrechts, Zürich/St. Gallen 2021, S. 238.

27 SHK-PFAFFINGER (Fn. 22), Art. 31 N 82.

28 WEBER (Fn. 13), Rn. 19; THOMAS STEINER/CHRISTIAN LAUX, in: Adrian Bieri/Julian Powell (Hrsg.), Orell Füssli Kommentar, Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlassen, Zürich 2023, Art. 31 N 37.

29 Ebenso ROSENTHAL (Fn. 20), Rn. 35.

30 WEBER (Fn. 13), Rn. 19; ROSENTHAL/VERALDI (Fn. 6), Rn. 66.

31 THOUVENIN/VOLZ (Fn. 8), S. 632; WEBER (Fn. 13), Rn. 20.

- (iii) Die Ergebnisse sind zudem so zu veröffentlichen, dass die betroffenen Personen nicht bestimmbar sind (Art. 31 Abs. 2 lit. e Ziff. 3 DSGVO). Diese Voraussetzung ist auf die Regelbeispiele der Bearbeitung für Forschung, Planung oder Statistik gerichtet. In diesen Fällen ist ihre Erfüllung sinnvoll und in aller Regel auch erforderlich. Für das Training von KI-Modellen passt sie aber kaum. Ihre Einhaltung hängt davon ab, ob mit einem Teil der Lehre anzunehmen ist, dass die Trainingsdaten bei der Verwendung eines KI-Modells «veröffentlicht» werden.³² Diese Auffassung vermag jedoch nicht zu überzeugen. (Gewisse) Trainingsdaten dürften sich zwar mit hinreichenden technischen Kenntnissen und mit einem gewissen Aufwand aus dem trainierten KI-Modell ermitteln lassen (sog. «reverse-engineering»).³³ Die Personendaten werden aber nicht im Sinn von Art. 31 Abs. 2 lit. e Ziff. 3 DSGVO veröffentlicht. Das gilt auch, wenn das trainierte Modell veröffentlicht, also zur öffentlichen Benutzung zur Verfügung gestellt wird.³⁴ Denn auch hier werden die Personendaten nicht als solche veröffentlicht, sondern müssen dem Modell mit technischen Massnahmen entnommen werden. Voraussetzung für das Fehlen einer Veröffentlichung im Sinn von Art. 31 Abs. 2 lit. e Ziff. 3 DSGVO ist allerdings, dass die Anbieter und Betreiber eines KI-Systems hinreichende technische Massnahmen getroffen haben, die sicherstellen, dass sich dem KI-Modell die beim Training verwendeten Personendaten nur mit einem derart grossen Aufwand entnehmen lassen, dass nach der allgemeinen Lebenserfahrung nicht damit zu rechnen ist, dass Dritte diesen Aufwand auf sich nehmen werden.³⁵ Ist das der Fall, dann ist auch diese Voraussetzung erfüllt.

Selbst wenn Zweifel bestehen sollten, ob die Voraussetzungen für eine Rechtfertigung als nicht-personenbezogene Bearbeitung erfüllt sind, ist eine Rechtfertigung ohne weiteres möglich, weil die Rechtfertigungsgründe von Art. 31 Abs. 2 DSGVO nur Regelbeispiele sind, bei denen ein überwiegendes Interesse des Verantwortlichen meist in Betracht kommt.³⁶ Da es sich um Regelbeispiele und nicht um vorgegebene Rechtfertigungsgründe handelt, sind die in Art. 31 Abs. 2 DSGVO genannten Datenbearbeitungen umgekehrt auch bei Erfüllen aller Tatbestandsvoraussetzungen nicht *per se* gerechtfertigt. Die Regelbeispiele sind damit nur, aber immerhin, Anhaltspunkte für eine Interessenabwägung im Einzelfall.³⁷ Da die Bearbeitung von Personendaten beim Training

32 So namentlich: WEBER (Fn. 13), Rn. 20; DAVID ROSENTHAL, Datenschutz und KI: Worauf in der Praxis zu achten ist, Jusletter IT vom 22. April 2022, Rn. 48 ff.

33 THOUVENIN/VOLZ (Fn. 8), S. 632.

34 ROSENTHAL (Fn. 20), Rn. 36.

35 DANIEL W. SEILER/MARCEL GRIESINGER, Spannungsfeld Künstliche Intelligenz (KI) und Datenschutzrecht, Jusletter vom 25. September 2023, Rn. 22; THOUVENIN/VOLZ (Fn. 8), S. 632; FLORIAN WERKMEISTER/JONATHAN LAUX, Large-Language-Models und die DSGVO – «Wer den Datenschutz nicht ehrt, ist die KI nicht wert?», PinG 04/2023, S. 103–111, S. 104.

36 BSK DSGVO-RAMPINI/HARASGAMA (Fn. 24), Art. 31 N 27.

37 Botschaft DSGVO (Fn. 4), S. 7073; BSK DSGVO-RAMPINI/HARASGAMA (Fn. 24), Art. 31 N 27.

des Verantwortlichen vor.²⁴ Diese Konstellationen umfassen unter anderem die Rechtfertigung von Bearbeitungen für nicht personenbezogene Zwecke, insbesondere für Forschung, Planung und Statistik (Art. 30 Abs. 2 lit. e DSGVO). Forschung, Planung und Statistik werden dabei exemplarisch genannt, der Rechtfertigungsgrund ist aber schon nach dem Wortlaut weiter und umfasst die Bearbeitung für nicht personenbezogene Zwecke allgemein.²⁵ Ein Personenbezug fehlt, wenn die Identität der betroffenen Personen für die Datenbearbeitung irrelevant ist und die Personen nicht (re-)identifizierbar sind.²⁶ Zwischen dem Zweck des Bearbeitens der Personendaten und der betroffenen Person darf also kein Zusammenhang bestehen.²⁷ Beim Training von KI-Modellen werden die Personendaten nicht bearbeitet, um Erkenntnisse über eine (oder mehrere) betroffene Person(en) zu gewinnen oder Folgen für diese Personen abzuleiten. Das Ziel ist vielmehr die Entwicklung eines Modells, das eine bestimmte Aufgabe (oder eine Vielzahl von unbestimmten Aufgaben) möglichst gut erledigen kann.²⁸ Die Bearbeitung ist damit nicht personenbezogen.²⁹

Neben dem zentralen Aspekt des fehlenden Personenbezugs enthält der Rechtfertigungsgrund drei weitere Voraussetzungen:

- (i) Der Verantwortliche muss die Daten anonymisieren, sobald es der Bearbeitungszweck erlaubt. Ist die Anonymisierung unmöglich oder erfordert sie einen unverhältnismässigen Aufwand, muss der Verantwortliche angemessene Massnahmen ergreifen, um die Bestimmbarkeit der betroffenen Personen zu verhindern (Art. 31 Abs. 2 lit. e Ziff. 1 DSGVO). Beim Training von KI-Modellen wird die verlangte Anonymisierung in der Regel mit geeigneten technischen Massnahmen erreicht.³⁰
- (ii) Bei der Bearbeitung von besonders schützenswerten Personendaten dürfen diese Dritten nur so bekannt gegeben werden, dass die betroffene Person nicht bestimmbar ist; ist dies nicht möglich, ist zu gewährleisten, dass die Dritten die Daten nur zu nicht personenbezogenen Zwecken bearbeiten (Art. 31 Abs. 2 lit. e Ziff. 2 DSGVO). Auch diese Voraussetzung wird in den meisten Fällen keine relevante Hürde sein.³¹

- (iii) Die Ergebnisse sind zudem so zu veröffentlichen, dass die betroffenen Personen nicht bestimmbar sind (Art. 31 Abs. 2 lit. e Ziff. 3 DSGVO). Diese Voraussetzung ist auf die Regelbeispiele der Bearbeitung für Forschung, Planung oder Statistik gerichtet. In diesen Fällen ist ihre Erfüllung sinnvoll und in aller Regel auch erforderlich. Für das Training von KI-Modellen passt sie aber kaum. Ihre Einhaltung hängt davon ab, ob mit einem Teil der Lehre anzunehmen ist, dass die Trainingsdaten bei der Verwendung eines KI-Modells «veröffentlicht» werden.³² Diese Auffassung vermag jedoch nicht zu überzeugen. (Gewisse) Trainingsdaten dürften sich zwar mit hinreichenden technischen Kenntnissen und mit einem gewissen Aufwand aus dem trainierten KI-Modell ermitteln lassen (sog. «reverse-engineering»).³³ Die Personendaten werden aber nicht im Sinn von Art. 31 Abs. 2 lit. e Ziff. 3 DSGVO veröffentlicht. Das gilt auch, wenn das trainierte Modell veröffentlicht, also zur öffentlichen Benutzung zur Verfügung gestellt wird.³⁴ Denn auch hier werden die Personendaten nicht als solche veröffentlicht, sondern müssen dem Modell mit technischen Massnahmen entnommen werden. Voraussetzung für das Fehlen einer Veröffentlichung im Sinn von Art. 31 Abs. 2 lit. e Ziff. 3 DSGVO ist allerdings, dass die Anbieter und Betreiber eines KI-Systems hinreichende technische Massnahmen getroffen haben, die sicherstellen, dass sich dem KI-Modell die beim Training verwendeten Personendaten nur mit einem derart grossen Aufwand entnehmen lassen, dass nach der allgemeinen Lebenserfahrung nicht damit zu rechnen ist, dass Dritte diesen Aufwand auf sich nehmen werden.³⁵ Ist das der Fall, dann ist auch diese Voraussetzung erfüllt.

Selbst wenn Zweifel bestehen sollten, ob die Voraussetzungen für eine Rechtfertigung als nicht-personenbezogene Bearbeitung erfüllt sind, ist eine Rechtfertigung ohne weiteres möglich, weil die Rechtfertigungsgründe von Art. 31 Abs. 2 DSGVO nur Regelbeispiele sind, bei denen ein überwiegendes Interesse des Verantwortlichen meist in Betracht kommt.³⁶ Da es sich um Regelbeispiele und nicht um vorgegebene Rechtfertigungsgründe handelt, sind die in Art. 31 Abs. 2 DSGVO genannten Datenbearbeitungen umgekehrt auch bei Erfüllen aller Tatbestandsvoraussetzungen nicht *per se* gerechtfertigt. Die Regelbeispiele sind damit nur, aber immerhin, Anhaltspunkte für eine Interessenabwägung im Einzelfall.³⁷ Da die Bearbeitung von Personendaten beim Training

24 CORRADO RAMPINI/REHANA C. HARASGAMA, in: Gabor-Paul Blechta/David Vasella (Hrsg.), Basler Kommentar, Datenschutzgesetz, Öffentlichkeitsgesetz, 4. Aufl., Basel 2024, Art. 31 N 27 f.

25 VOLZ (Fn. 11), S. 57.

26 BSK DSGVO-RAMPINI/HARASGAMA (Fn. 24), Art. 31 N 60; GABRIEL KASPAR, People Analytics in privatrechtlichen Arbeitsverhältnissen, Vorschläge zur wirksameren Durchsetzung des Datenschutzes, Zürich/St. Gallen 2021, S. 238.

27 SHK-PFAFFINGER (Fn. 22), Art. 31 N 82.

28 WEBER (Fn. 13), Rn. 19; THOMAS STEINER/CHRISTIAN LAUX, in: Adrian Bieri/Julian Powell (Hrsg.), Orell Füssli Kommentar, Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlassen, Zürich 2023, Art. 31 N 37.

29 Ebenso ROSENTHAL (Fn. 20), Rn. 35.

30 WEBER (Fn. 13), Rn. 19; ROSENTHAL/VERALDI (Fn. 6), Rn. 66.

31 THOUVENIN/VOLZ (Fn. 8), S. 632; WEBER (Fn. 13), Rn. 20.

32 So namentlich: WEBER (Fn. 13), Rn. 20; DAVID ROSENTHAL, Datenschutz und KI: Worauf in der Praxis zu achten ist, Jusletter IT vom 22. April 2022, Rn. 48 ff.

33 THOUVENIN/VOLZ (Fn. 8), S. 632.

34 ROSENTHAL (Fn. 20), Rn. 36.

35 DANIEL W. SEILER/MARCEL GRIESINGER, Spannungsfeld Künstliche Intelligenz (KI) und Datenschutzrecht, Jusletter vom 25. September 2023, Rn. 22; THOUVENIN/VOLZ (Fn. 8), S. 632; FLORIAN WERKMEISTER/JONATHAN LAUX, Large-Language-Models und die DSGVO – «Wer den Datenschutz nicht ehrt, ist die KI nicht wert?», PinG 04/2023, S. 103–111, S. 104.

36 BSK DSGVO-RAMPINI/HARASGAMA (Fn. 24), Art. 31 N 27.

37 Botschaft DSGVO (Fn. 4), S. 7073; BSK DSGVO-RAMPINI/HARASGAMA (Fn. 24), Art. 31 N 27.

von KI-Modellen für die betroffenen Personen im Normalfall keinerlei Konsequenzen hat, dürfte das Interesse des Verantwortlichen an der Bearbeitung das Interesse der Betroffenen an der Nicht-Bearbeitung aber in aller Regel überwiegen.³⁸

Aufgrund der schon heute grossen und weiter wachsenden Bedeutung von KI-Modellen wäre es sinnvoll, diese Rechtslage im DSG ausdrücklich klarzustellen, indem der Gesetzgeber die Liste der Regelbeispiele für das Vorliegen eines überwiegenden Interesses des Verantwortlichen um ein weiteres Beispiel ergänzen würde, das die Bearbeitung von Personendaten für das Trainieren, Validieren und Testen von KI-Modellen erfasst. Alternativ könnte es genügen, wenn der Eidgenössische Datenschutz- und Öffentlichkeitsbeauftragte (EDÖB) in einem Leitfaden oder Merkblatt präzisieren würde, dass die Inbetriebnahme eines KI-Modells nicht als Veröffentlichung von Ergebnissen im Sinne der dritten Voraussetzung des Rechtfertigungsgrundes der Bearbeitung für nicht personenbezogene Zwecke (Art. 31 Abs. 2 lit. e Ziff. 3 DSG) gilt³⁹ oder bei der Bearbeitung von Personendaten für das Training von KI-Modellen allgemein ein überwiegendes Interesse des Verantwortlichen vorliegt.

D. Rechte und Pflichten

Das DSG sieht Pflichten des Verantwortlichen und Rechte der betroffenen Personen vor. Bei der Entwicklung von KI-Modellen stehen die Informationspflicht des Verantwortlichen und das Auskunftsrecht der Betroffenen im Vordergrund. Zudem stellt sich die Frage, ob Betroffene das Recht auf Berichtigung, den Anspruch auf Löschung und das Recht auf Widerspruch gegen das Bearbeiten von Personendaten beim Training von KI-Modellen geltend machen können.

Diese Fragen stellen sich nicht nur vor und beim Training, sondern auch danach. Datenschutzrechtliche Rechte und Pflichten bestehen nach dem Training aber nur, wenn das trainierte KI-Modell Personendaten enthält. Auch wenn in KI-Modellen regelmässig Sequenzen aus den Trainingsdaten memorisiert werden und es in vielen Fällen möglich ist, den trainierten Modellen zumindest gewisse Personendaten zu entnehmen,⁴⁰ lässt sich daraus nicht allgemein schliessen, dass trainierte KI-Modelle Personendaten im Sinn des Datenschutzrechts enthalten. Denn ob Personendaten vorliegen, bestimmt sich stets relativ aus der Perspektive des jeweiligen Bearbeiters.⁴¹ Entscheidend ist dabei, ob der Bearbeiter ein Interesse an der Identifizierung betroffener Personen hat und in der

Lage ist, diese mit vernünftigem Aufwand zu identifizieren.⁴² Ob ein trainiertes KI-Modell Personendaten im Sinn des Datenschutzrechts enthält, kann deshalb nicht allgemein, sondern nur aus Sicht des jeweiligen Anbieters oder Betreibers beantwortet werden.

Primärer Adressat der Rechte und Träger der Pflichten ist der Verantwortliche. Als *Verantwortlicher* gilt, wer allein oder zusammen mit anderen über den Zweck und die Mittel der Bearbeitung entscheidet (Art. 5 lit. j DSG). Verantwortlicher ist damit regelmässig das Unternehmen, welches das KI-Modell trainiert und über die dafür notwendigen Daten verfügt, in der Terminologie der KI-Verordnung der EU⁴³ also der Anbieter («provider»)⁴⁴. Dasselbe gilt für die Weiterentwicklung von KI-Modellen. Entwickelt oder trainiert der Anbieter ein KI-Modell aber spezifisch für einen Betreiber («deployer»), werden der Anbieter und der Betreiber regelmässig als gemeinsam Verantwortliche zu qualifizieren sein.⁴⁵

Der Verantwortliche ist verpflichtet, die betroffenen Personen über das Beschaffen von Personendaten angemessen zu informieren (Art. 19 Abs. 1 Satz 1 DSG). Das gilt auch, wenn die Daten nicht bei der betroffenen Person beschafft werden (Art. 19 Abs. 1 Satz 2 DSG); in diesem Fall entfällt die *Informationspflicht*, wenn die Information nicht möglich ist oder einen unverhältnismässigen Aufwand erfordert (Art. 20 Abs. 2 DSG). Werden Daten direkt bei der betroffenen Person beschafft, wird der Verantwortliche die Person in aller Regel informieren können. Beim Training von KI-Modellen ist das allerdings kaum je der Fall, weil der Verantwortliche die Daten in der Regel bei Dritten beschafft, bspw. durch das Einkaufen grosser Datensätze oder das Crawlen von Websites.⁴⁶ Da der Verantwortliche die Betroffenen regelmässig nicht oder nur mit unverhältnismässigem Aufwand direkt informieren kann, muss es genügen, wenn er die Öffentlichkeit in allgemeiner Weise informiert, bspw. durch eine Mitteilung auf seiner Website.⁴⁷

42 GABOR-PAUL BLECHTA/LUCA DAL MOLIN/KIRSTEN WESIAK-SCHMIDT, in: Gabor-Paul Blechta/David Vasella (Hrsg.), Basler Kommentar, Datenschutzgesetz, Öffentlichkeitsgesetz, 4. Aufl., Basel 2024, Art. 5 N 29; PHILIP GLASS, in: Adrian Bieri/Julian Powell (Hrsg.), Orell Füssli Kommentar, Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlassen, Zürich 2023, Art. 5 N 14; SHK-RUDIN (Fn. 41), Art. 5 N 10 f.

43 Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 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) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz).

44 Art. 3 Ziff. 3 KI-Verordnung; siehe dazu: ROSENTHAL (Fn. 20), Rn. 12 ff.; WEBER (Fn. 13), Rn. 38 f.

45 THOUVENIN/VOLZ (Fn. 13), S. 5.

46 MARCEL GRIESINGER, Datenschutz-Compliance beim Einsatz künstlicher Intelligenz, RR-COMP 6/2024, S. 5–7, S. 6.

47 Siehe dazu auch BIERI/POWELL (Fn. 19), Rn. 20.

38 ROSENTHAL/VERALDI (Fn. 6), Rn. 63, 66.

39 Siehe dazu auch: WEBER (Fn. 13), Rn. 22; THOUVENIN/VOLZ (Fn. 8), S. 632.

40 Siehe dazu oben, Kapitel B.III.

41 Statt vieler: BEAT RUDIN, in: Bruno Baeriswyl/Kurt Pärli/Dominika Blonski (Hrsg.), Stämpfli Handkommentar, Datenschutzgesetz, 2. Aufl., Bern 2023, Art. 5 N 12.

Jede Person kann vom Verantwortlichen *Auskunft* darüber verlangen, ob Personendaten über sie bearbeitet werden (Art. 25 Abs. 1 DSG). Das gilt auch beim Bearbeiten von Personendaten für das Training von KI-Modellen. Da die hierzu verwendeten Daten in der Regel in einer Datenbank gespeichert sind, ist der Anbieter («provider») eines KI-Modells durchaus in der Lage, einem Auskunftsbeghären zu genügen. Fraglich ist, ob die betroffene Person auch beim Betreiber («deployer») eines KI-Systems Auskunft darüber verlangen kann, ob ihre Personendaten für das Training des in diesem System verwendeten Modells genutzt wurden. Ein solcher Anspruch besteht jedenfalls nur, wenn das trainierte Modell Personendaten enthält. Aus Sicht der Betreiber von KI-Systemen ist das regelmässig nicht der Fall, weil sie meist weder ein Interesse haben noch in der Lage sind, einem trainierten Modell mit vernünftiger Aufwand Personendaten zu entnehmen. Die Betroffenen haben deshalb in aller Regel keinen Anspruch auf Auskunft gegenüber dem Betreiber eines KI-Systems.

Das *Recht auf Berichtigung* wird im DSG zwar in der Bestimmung über die Rechtsansprüche normiert, die bei einem Verstoß gegen die Vorgaben des DSG zur Verfügung stehen (Art. 32 Abs. 1 DSG). Es ist aber anerkannt, dass die betroffenen Personen dieses Recht auch geltend machen können, wenn kein solcher Verstoß vorliegt.⁴⁸ Werden beim Training von KI-Modellen unrichtige Personendaten bearbeitet, können die Betroffenen damit an sich deren Berichtigung verlangen. Voraussetzung ist allerdings, dass die Daten im datenschutzrechtlichen Sinn unrichtig sind, dass sie also eine Tatsache oder einen Umstand im Hinblick auf den Bearbeitungszweck nicht richtig wiedergeben.⁴⁹ Da es beim Training von KI-Modellen in aller Regel nicht auf den Gehalt eines einzelnen Datums ankommt, ist das kaum je der Fall. Kein Anspruch auf Berichtigung besteht gegenüber einem trainierten KI-Modell. Selbst wenn ein trainiertes Modell im datenschutzrechtlichen Sinn unrichtige Personendaten enthalten sollte, muss ein solcher Anspruch daran scheitern, dass eine Berichtigung einzelner Daten in einem Modell nicht oder jedenfalls nur mit unverhältnismässigem Aufwand möglich wäre. Die Bestimmung über das Recht auf Berichtigung sieht zwar keine solche Ausnahme vor, sie lässt sich aber durch einen Analogieschluss zu den entsprechenden Ausnahmen bei der Informationspflicht (Art. 20 Abs. 2 DSG), bei der Information der Betroffenen über eine Verletzung der Datensicherheit (Art. 24 Abs. 5 lit. b DSG) und beim Recht auf

Datenherausgabe (Art. 28 Abs. 2 DSG) oder mit dem Rechtsmissbrauchsverbot begründen.⁵⁰

Anders als das Recht auf Berichtigung besteht ein *Recht auf Löschung* von Personendaten nur, wenn ein Verstoß gegen die Vorgaben des DSG vorliegt (Art. 32 Abs. 2 lit. c DSG). Das ist beim Training von KI-Modellen, wie vorstehend ausgeführt, kaum je der Fall, weil sich allfällige Verstöße gegen die Grundsätze der Datenbearbeitung in aller Regel rechtfertigen lassen.⁵¹

Die betroffenen Personen können gegen die Bearbeitung ihrer Daten beim Training eines KI-Modells *Widerspruch* erheben (Art. 30 Abs. 2 lit. b DSG). Die Bearbeitung der Personendaten ist damit aber nicht *per se* unzulässig, weil es auch nach einem Widerspruch möglich bleibt, die Bearbeitung durch ein überwiegendes Interesse zu rechtfertigen (Art. 30 Abs. 2 i.V.m. Art. 31 Abs. 1 DSG).⁵² Die Voraussetzungen für die Annahme eines überwiegenden Interesses des Verantwortlichen sind dann zwar erhöht, weil die betroffene Person mit dem Widerspruch unzweideutig zum Ausdruck gebracht hat, dass sie die Bearbeitung ihrer Personendaten verhindern möchte.⁵³ Dies schliesst eine Rechtfertigung im Einzelfall aber keineswegs aus. Nicht nur im Einzelfall, sondern allgemein wird man ein überwiegendes Interesse des Verantwortlichen bejahen, wenn eine betroffene Person erst nach dem Training eines KI-Modells Widerspruch gegen die (weitere) Bearbeitung ihrer Personendaten erheben sollte, weil Personendaten aus einem trainierten Modell gar nicht oder nur mit unverhältnismässigem Aufwand gelöscht werden können.

E. Fazit

Die Analyse zeigt, dass das Bearbeiten von Personendaten bei der Entwicklung von KI-Modellen regelmässig gegen einen oder mehrere Grundsätze der Datenbearbeitung verstösst, namentlich gegen den Grundsatz der Erkennbarkeit und der Zweckbindung, bisweilen auch gegen den Grundsatz der Verhältnismässigkeit. Diese Verstöße lassen sich aber mit den überwiegenden Interessen des Verantwortlichen rechtfertigen. Im Vordergrund steht dabei das Regelbeispiel des überwiegenden Interesses bei einer Bearbeitung für nicht personenbezogene Zwecke. Angesichts der grossen Bedeutung der Frage wäre es allerdings

48 Botschaft DSG (Fn. 4), S. 7076; APOLLO DAUAG/LILIANE OBRECHT, in: Adrian Bieri/Julian Powell (Hrsg.), Orell Füssli Kommentar, Kommentar zum Schweizerischen Datenschutzgesetz mit weiteren Erlassen, Zürich 2023, Art. 32 N 20; BSK DSG-RAMPINI/HARASGAMA (Fn. 24), Art. 32 N 10; a.M. SHK-PFAFFINGER (Fn. 22), Art. 32 N 12.

49 BSK DSG-RAMPINI/HARASGAMA (Fn. 24), Art. 32 N 9; OFK-DAUAG/OBRECHT (Fn. 48), Art. 32 N 20; OFK-FANGER (Fn. 4), Art. 6 N 11; JOSEPHINE BOILLAT/STÉPHANE WERLY, in: Philippe Meier/Sylvain Métile (Hrsg.), Commentaire Romand, Loi fédérale sur la protection des données, Basel 2023, Art. 32 N 10.

50 Siehe zum zweiten Begründungsansatz: DAVID ROSENTHAL, Das neue Datenschutzgesetz, Jusletter vom 16. November 2020, Rn. 140; SHK-PFAFFINGER (Fn. 22), Art. 32 N 14; BSK DSG-RAMPINI/HARASGAMA (Fn. 24), Art. 32 N 11.

51 Siehe dazu oben, Kapitel C.

52 BGE 127 III 481 E. 3b; BSK DSG-RAMPINI/HARASGAMA (Fn. 24), Art. 30 N 17; OFK-STBIENER/LAUX (Fn. 28), Art. 30 N 20; SYLVAIN MÉTILE, Le traitement de données personnelles sous l'angle de la (nouvelle) loi fédérale sur la protection des données du 25 septembre 2020, SJ 2021 II, S. 1–46, S. 33; ROSENTHAL (Fn. 50), Rn. 138.

53 SHK-PFAFFINGER (Fn. 22), Art. 30 DSG N 51; ROSENTHAL (Fn. 50), Rn. 38, 41.

wünschenswert, die Zulässigkeit der Bearbeitung von Personendaten für das Training von KI-Modellen in einem neuen gesetzlichen Regelbeispiel ausdrücklich zu erwähnen.

Zusammenfassung

Bei der Entwicklung und Verwendung von Modellen und Systemen der Künstlichen Intelligenz (KI) werden regelmässig sehr grosse Mengen von Personendaten verwendet. Von zentraler Bedeutung ist die Frage, ob und unter welchen Voraussetzungen Personendaten beim Training von KI-Modellen verwendet werden dürfen. Die dabei erfolgenden Bearbeitungen von Personendaten verstossen regelmässig gegen einen oder mehrere Grundsätze der Datenbearbeitung, namentlich gegen die Grundsätze der Erkennbarkeit und der Zweckbindung, bisweilen auch gegen den Grundsatz der Verhältnismässigkeit. Diese Verstösse lassen sich aber mit den überwiegenden Interessen des Verantwortlichen rechtfertigen. Im Vordergrund steht dabei das Regelbeispiel des überwiegenden Interesses bei einer Bearbeitung für nicht personenbezogene Zwecke. Besondere Herausforderungen bestehen auch bei den Rechten der betroffenen Personen und den Pflichten der Verantwortlichen, namentlich bei der Informationspflicht, beim Auskunftsrecht und beim Recht auf Berichtigung.

Résumé

Le développement et l'utilisation de modèles et de systèmes d'intelligence artificielle (IA) impliquent régulièrement l'utilisation de très grandes quantités de données personnelles. La question de savoir si et dans quelles conditions des données personnelles peuvent être utilisées lors de l'entraînement de modèles d'IA est pertinente. Les traitements de données personnelles qui en découlent violent régulièrement un ou plusieurs principes du traitement des données, notamment le principe de la reconnaissabilité et de la finalité, parfois aussi le principe de la proportionnalité. Ces violations peuvent toutefois être justifiées par les intérêts prépondérants du responsable du traitement, notamment par l'intérêt prépondérant lors d'un traitement à des fins qui ne se rapportent pas à des personnes. Des défis particuliers se posent également en ce qui concerne les droits des personnes concernées et les obligations des responsables, notamment l'obligation d'information, le droit d'accès et le droit de rectification.