www.jusletter-it.eu

Michal Cichocki

Guidelines für Künstliche Intelligenz (KI): Besteht aus rechtlicher Sicht Handlungsbedarf?

Der Aufsatz beleuchtet KI-Guidelines aus rechtlicher Sicht mit besonderer Berücksichtigung des Datenschutzrechts. Dabei wird analysiert, ob es aus der Perspektive des juristischen Praktikers Handlungsbedarf gibt und wie dieser in der Unternehmenspraxis operationalisiert werden kann.

Beitragsart: The DPO View
Region: Schweiz; EU
Rechtsgebiete: Artificial intelligence; Datenschutz; Datensicherheit

Zitiervorschlag: Michal Cichocki, Guidelines für Künstliche Intelligenz (KI): Besteht aus rechtlicher Sicht Handlungsbedarf?, in: Jusletter IT 25. Februar 2021

Inhaltsübersicht

1. Ausgangslage

[1] Künstliche Intelligenz (KI) ist ein *hot topic* – kaum vergeht eine Woche ohne aufsehenerregende Schlagzeilen zu KI, die grosse Hoffnungen wecken, aber auch ebenso grosse Ängste schüren. Dazu gehören beispielsweise die erfolgreiche Bekämpfung von Hunger[1], Krankheiten sowie Kriminalität aber auch die Abhängigkeit oder Versklavung der Menschen[2] durch KI. Auch wenn solche Prognosen aus heutiger Sicht utopisch klingen mögen, werden Stimmen nach zwingender Regulierung von KI zunehmend lauter[3]. «Freiwillige» Guidelines für den vermeintlich richtigen Umgang mit KI gibt es bereits heute wie Sand am Meer: Eine Google-Suche der Schlagwörter «AI» und «Guidelines» führt zu mehr als 0.5 Mio. Treffern. Zahlreiche staatliche Behörden, Think Tanks, NGO sowie Big Tech-Unternehmen haben Richtlinien, Strategien, Grundsatzpapiere etc. zu KI veröffentlicht.

[2] In diesem Zusammenhang stellen sich aus der Sicht des juristischen Praktikers folgende Fragen: Wie soll mit bestehenden KI-Guidelines umgegangen werden? Gibt es aus einer rechtlichen Perspektive Handlungsbedarf[4]? Wenn ja, welchen?

1 Vgl. Virginia Kirst, Diese KI-Software ist die Revolution im Kampf gegen den Welthunger, Welt vom 03. Januar 2021 (online).

2 Vgl. Aatif Sulleyman, AI is highly likely to destroy humans, Elon Musk warns, Independent vom 24. November 2017 (online).

3 Für die EU: vgl. Michal Cichocki, Europäische Kommission veröffentlicht White Paper zu Künstlicher Intelligenz (KI), 29. Februar 2020, http://www.lawblogswitzerland.ch/2020/02/europaische-kommission-veroffentlicht.html.

4 Sind KI-Guidelines verbindlich und damit gerichtlich durchsetzbar – sei es aufgrund einer vertraglichen Übernahme oder aufgrund einer anwendbaren Rechtsgrundlage? Insbesondere wenn sie aus der Feder staatlicher Behörden stammen? Oder handelt es sich dabei lediglich um eine unverbindliche Meinungsäusserung eines weiteren (KI-)Akteurs?

2. Meta-Guidelines

[3] Die erste ganz praktische Herausforderung besteht darin, sich eine Übersicht über die riesige Vielzahl und Vielfalt der veröffentlichten KI-Guidelines[5] zu verschaffen. Glücklicherweise existieren erste Meta-Guidelines, die sich mit der Analyse bestehender KI-Guidelines befassen. Beispielsweise haben Forscher der Harvard University sechsunddreissig KI-Guidlines namhafter Akteure aus verschiedenen Organisationen, Branchen sowie aus unterschiedlichen Kontinenten untersucht und verglichen[6]. Dabei sind die Forscher zu einem überraschenden Ergebnis gelangt: Trotz hoher Diversität der Akteure wurden acht Grundsätze identifiziert, denen beim Umgang mit KI-Technologien *übereinstimmend besonders grosse Bedeutung* zukommen soll: Privacy, Accountability, Safety and Security, Explainability and Transparency, Fairness and Non-Discrimination, Human Control of Technology, Professional Responsibility sowie Promotion of Human Values.

[4] Nachfolgend wird untersucht, wie diese acht KI-Grundsätze aus *rechtlicher Sicht* zu beurteilen sind, insbesondere mit Blick auf das geltende sowie totalrevidierte Datenschutzgesetz (DSG bzw. revDSG) und die Europäische Datenschutz-Grundverordnung (EU-DSGVO). In einem zweiten Schritt wird auf deren *Operationalisierung in der Unternehmenspraxis* eingegangen.

3. Wichtigste KI-Grundsätze aus der rechtlichen Perspektive

3.1. Privacy

[5] Privacy wurde als Grundsatz in 97% der untersuchten KI-Guidelines erwähnt. Vereinfacht gesagt, wird darunter der *Schutz der Privatsphäre* natürlicher Personen, die von KI-Technologien betroffen sind, verstanden. Dieser Schutz soll beispielsweise durch Ansprüche auf Löschung, Widerspruch, Berichtigung oder generell durch informationelle Selbstbestimmung gewährleistet werden.

[6] Mit Blick auf die Bestimmungen von Art. 28 ZGB, des derzeit geltenden DSG, dessen Totalrevision (revDSG), der EU-DSGVO (sofern anwendbar[7]) und einer Vielzahl spezialgesetzlicher Normen zum Schutz der Persönlichkeitsrechte, beispielsweise im Arbeits- oder Strafrecht[8], zeigt sich, dass der Privacy-Grundsatz heute bereits weitestgehend durch rechtliche Vorgaben abgedeckt ist. Aus rechtlicher Sicht besteht demnach kein zusätzlicher, KI-spezifischer Handlungsbedarf.

---

5 Vgl. Michal Cichocki, Aktuelle Guidelines zu Künstlicher Intelligenz (KI) – Eine Übersicht, 08. Januar 2019, http://www.lawblogswitzerland.ch/2019/01/guidelines-zu-kunstlicher-intelligenz.html.

6 Vgl. Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. «Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI.» Berkman Klein Center for Internet & Society, 2020.

7 Vgl. Art. 2 DSGVO für den sachlichen sowie Art. 3 DSGVO für den räumlichen Anwendungsbereich.

8 Z.B. Art. 328b OR (Schutz der Persönlichkeit der Arbeitnehmenden), Art. 26 ArGV 3 (Schutz vor unzulässiger Verhaltensüberwachung der Arbeitnehmenden), Art. 179 StGB (Verletzung des Schriftgeheimnisses), Art. 179bis StGB (Abhören und Aufnehmen fremder Gespräche), Art. 179ter StGB (unbefugtes Aufnehmen von Gesprächen), Art. 179quater StGB (Verletzung des Geheim- oder Privatbereichs durch Aufnahmegeräte), Art. 179novies StGB (unbefugtes Beschaffen von Personendaten) etc.

3.2. Security and Safety

[7] 81% der analysierten KI-Guidelines enthalten Security and Safety als KI-Grundsatz. Sie verstehen darunter im Wesentlichen die klassischen *Schutzziele der Informationssicherheit*, namentlich die Vertraulichkeit, Integrität sowie Verfügbarkeit.

[8] Security and Safety werden durch eine Vielzahl rechtlicher Vorgaben beispielsweise im Strafrecht[9], Datenschutzrecht[10] sowie durch spezialgesetzliche[11] und sektorspezifische Normen[12] weitgehend abgedeckt. Darüber hinaus gibt es eine grosse Anzahl etablierter (nicht rechtlicher) Regeln der Berufskunde[13], die ebenfalls Vorgaben für Security and Safety definieren. Auch hier ist kein KI-spezifischer Handlungsbedarf aus rechtlicher Perspektive erkennbar.

3.3. Fairness and Non-Discrimination

[9] Sämtliche (100%) untersuchten KI-Guidelines sehen Fairness and Non-Discrimination als Grundsatz für KI-Technologien vor und wollen damit die Angst vor beispielsweise «algorithmic bias» oder vor «technochauvinistischen»[14] Applikationen durch KI adressiert wissen.

[10] Einerseits dürfte der Gedanke von Fairness and Non-Discrimination mindestens teilweise durch rechtliche Vorgaben zu Treu und Glaube[15], Richtigkeit der Daten[16] aber auch zum Schutz vor unlauterem Wettbewerb[17], durch das Kartellgesetz[18] sowie Strafrecht[19] abgedeckt sein. Andererseits gilt es insbesondere im Zusammenhang mit dem Diskriminierungsverbot zu beachten, dass die *schweizerische* Rechtsordnung grundsätzlich *keine direkte Drittwirkung* von Grundrechten im Privatbereich vorsieht[20]; das Primat der Privatautonomie und der Vertragsfreiheit steht

9 Art. 143 StGB (unbefugte Datenbeschaffung), Art. 143bis StGB (unbefugtes Eindringen in ein Datenverarbeitungssystem), Art. 144bis StGB (Datenbeschädigung), Art. 147 StGB (Betrügerischer Missbrauch einer Datenverarbeitungsanlage), Art. 162 StGB (Geschäfts- und Fabrikationsgeheimnisse), Art 6 UWG (Verletzung von Fabrikations- und Geschäftsgeheimnissen) und Art. 321 StGB (Verletzung des Berufsgeheimnisses). Ferner ist denkbar, dass Art. 271 StGB (Handeln für einen fremden Staat) sowie Art. 273 StGB (wirtschaftlicher Nachrichtendienst) angerufen werden könnten.

10 Art. 7 DSG (Datensicherheit), Art. 35 DSG (Verletzung der beruflichen Schweigepflicht), Art. 8 revDSG (Datensicherheit), Art. 24 revDSG (Meldung von Verletzungen der Datensicherheit, «Data Breach»), Art. 62 revDSG (Verletzung der beruflichen Schweigepflicht), Art. 32 DSGVO (Sicherheit der Verarbeitung), Art. 33 f. DSGVO («Data Breach»).

11 Z.B. im Produkthaftpflichtgesetz (PrHG).

12 Z.B. Art. 47 BankG (strafrechtlicher Schutz des Bankkundengeheimnisses) oder die einschlägigen Regulierungen der FINMA wie bspw. das FINMA-RS 2008/21 «Operationelle Risiken Banken», FINMA-RS 2018/3 «Outsourcing – Banken und Versicherer», die FINMA-Aufsichtsmitteilung 05/2020 zur Meldepflicht von Cyber-Attacken gemäss Art. 29 Abs. 2 FINMAG.

13 Z.B. die ISO/IEC-27000-Reihe, den IKT-Minimalstandard des Bundesamtes für wirtschaftliche Landesversorgung (BWL) oder Standards des deutschen Bundesamts für Sicherheit in der Informationstechnik (BSI-Standards).

14 Glaube, dass Technologie immer die bessere Lösung und generell dem Menschen überlegen sei.

15 Vgl. Art. 4 Abs. 2 ZGB, Art. 4 Abs. 2 DSG, Art. 6 Abs. 2 revDSG, Art. 5 Abs. 1 lit. a DSGVO (Treu und Glauben).

16 Vgl. Art. 5 DSG, Art. 6 Abs. 5 revDSG, Art. 5 Abs. 1 lit. d DSGVO (Richtigkeit der Daten).

17 Vgl. Art. 2 UWG (Grundsatz; Generalklausel) und allenfalls auch in Art. 3 Abs. 1 lit. h UWG (besonders aggressive Verkaufsmethoden), Art. 8 UWG (Verwendung missbräuchlicher Geschäftsbedingungen).

18 Vgl. Art. 5 KG (unzulässige Wettbewerbsabreden) sowie Art. 7 KG (unzulässige Verhaltensweisen marktbeherrschender Unternehmen).

19 Vgl. Art. 261bis StGB (Diskriminierung und Aufruf zu Hass).

20 Vgl. Benjamin Schindler, Zu Begriff und Verständnis der «Grundrechte» in der neuen Bundesverfassung, in: Thomas Gächter/Martin Bertschi (Hrsg.), Neue Akzente in der «nachgeführten Bundesverfassung», Zürich 2000, S. 55 f.

einem generellen Diskriminierungsverbot entgegen[21]. Beispielsweise verstösst die Festlegung unterschiedlicher (dynamischer) Preise für dieselben Güter oder Dienstleistungen unter Einbezug subjektiver Kriterien potenzieller Käufer im Privatbereich per se nicht gegen das Diskriminierungsverbot[22].

[11] Wo von Grundsätzen die Rede ist, gibt es auch Ausnahmen: So beispielsweise bei der Gleichstellung von Mann und Frau (Art. 8 Abs. 3 BV), die auch im Privatbereich direkt anwendbar ist. Darüber hinaus hat der schweizerische Verfassungsgeber in Art. 35 Abs. 3 BV vorgesehen, dass Grundrechte, soweit sie sich dafür eignen, insbesondere im Rahmen des Rechtsetzungsverfahrens für den Privatbereich berücksichtigt werden sollen[23]. Diesem Auftrag wurde in bestimmten Bereichen mit dem Mittel spezialgesetzlicher Diskriminierungsverbote[24] Rechnung getragen.

[12] Demnach werden Fairness and Non-Discrimination zwar durch rechtliche Vorgaben mindestens teilweise abgedeckt. Weil in der Schweiz ein generelles Diskriminierungsverbot fehlt, kann im Rahmen der Privatautonomie freiwillig auf das Ausschöpfen des rechtlich zulässigen Maximums verzichtet werden; beispielsweise indem die erwähnte dynamische Preisfestlegung nicht verwendet wird, obwohl diese «Ungleichbehandlung» rechtlich zulässig wäre. Solche Managemententscheide werden vermehrt als «*ethische*» *Entscheide* bezeichnet und können u.a. mit dem Geschäftsmodell, der Marktpositionierung oder der Reputation des jeweiligen Unternehmens begründet werden. Dieser Spielraum besteht unabhängig von KI-Technologien, dürfte aber bei deren Einsatz besonders sorgfältig geprüft werden, um allfällige KI-spezifische Diskriminierungsängste[25] im Voraus zu entschärfen. In diesem Zusammenhang könnte auch der Blick über das eigene Unternehmen hinaus interessieren: Die FINMA hat die *zulässige* Diskriminierung des zunehmend gläsernen Versicherungsnehmers als langfristiges Risiko für den Versicherungsmarkt *insgesamt* identifiziert[26]. Diese könne nämlich zu einer Entsolidarisierung, zu Missbrauch gegenüber Kunden sowie zu verschärfter Konkurrenz durch neue Dienstleister führen.

3.4. Explainability and Transparency

[13] 94% der Guidelines sprechen sich für Explainability and Transparency im Zusammenhang mit KI aus. Unter Explainability wird die Beantwortung der Fragen «wie funktionieren KI-Technologien » oder «auf welcher Berechnungsgrundlage bzw. Logik hat KI im Einzelfall entschieden» usw. verstanden. Mit Transparency ist dagegen gemeint, dass der Einsatz sowie die wesentlichen Eckwerte von KI-Technologien (z.B. Quellen, Verwendungszwecke, Empfänger) gegenüber den betroffenen Personen erkennbar, d.h. transparent, gemacht werden sollen.

---

21 Anders in der EU: bspw. haben die grundrechtsähnlichen *Grundfreiheiten* gemäss Vertrag über die Arbeitsweise der Europäischen Union (AEUV) unmittelbare Horizontalwirkung zwischen Privaten. Vgl. dazu Detlev W. Belling, Antje Herold, Marek Kneis, Die Wirkung der Grundrechte und Grundfreiheiten zwischen Privaten, in: AIUP (2014) 02, S. 98 ff.

22 Vgl. Florent Thouvenin, Dynamische Preise, in: Jusletter IT vom 22. September 2016.

23 Sog. mittelbare Drittwirkung von Grundrechten.

24 Z.B. im Gleichstellungsgesetz (GlG), Behindertengleichstellungsgesetz (BehiG), Art. 271 OR (Anfechtbarkeit der Kündigung im Mietrecht), Art. 328 und Art. 328b OR (Schutz der Persönlichkeit des Arbeitnehmers). Eingriffe in die Privatautonomie kann es auch aus anderen Gründen geben: Anleger-, Gläubiger-, Konsumentenschutz etc.

25 Vgl. Bitkom-Umfrage, Die Menschen wollen KI – und haben auch Angst vor ihr, 28. September 2020, https://www.bitkom.org/Presse/Presseinformation/Die-Menschen-wollen-KI-und-haben-auch-Angst-vor-ihr.

26 Vgl. FINMA-Risikomonitor 2020, S. 17, https://www.finma.ch/de/news/2020/11/20201111-mm-risikomonitor-2020.

[14] Erklärbarkeit (Explainability) und Transparenz (Transparency) nehmen insbesondere im Datenschutzrecht[27] eine herausragende Stellung ein und sind Ausfluss des Grundgedankens der informationellen Selbstbestimmung[28]. Diese wird vornehmlich durch Informationspflichten, Betroffenenrechte (inkl. Darlegung der Logik, auf der eine Entscheidung beruht, sofern die Voraussetzungen der automatisierten Einzelentscheidung erfüllt werden) sowie durch Datenschutz-Folgeabschätzungen (DSFA) oder Zertifizierungen sichergestellt. Je nach Anwendungsfall kommen spezialgesetzliche Vorgaben[29] hinzu.

[15] Demnach werden Explainability and Transparency ebenfalls durch rechtliche Vorgaben abgedeckt, vor allem wenn Personendaten bearbeitet werden. Ausserhalb der Bearbeitung von Personendaten[30], z.B. bei reinen Sachdaten (d.h. Angaben, bei denen von vornherein kein Personenbezug existierte), anonymisierten Daten (d.h. Angaben, bei denen sämtliche Identifikationsmerkmale einer Person entfernt wurden) oder synthetischen («erfundenen») Daten, könnte aus ethischen Gründen (vgl. 3.3.) *freiwillig* und damit gesetzesüberschiessend auf den Einsatz von KI-Technologien hingewiesen («transparent gemacht») und diese *soweit wie möglich*[31], erklärt werden.

3.5. Accountability

[16] Der Nachweis der Einhaltung von Vorgaben bzw. deren richtige Anwendung im Einzelfall wird von 97% der beleuchteten KI-Guidelines als guiding principle angesehen. Dieser Grundsatz soll beispielsweise mit den Mitteln eines Impact Assessments, Monitorings, Audits, menschlicher Intervention etc. sichergestellt werden.

[17] Das Konzept der Accountability (Rechenschaftspflicht) ist in der schweizerischen Rechtsordnung insbesondere im Datenschutzrecht[32] und in generell regulierten Branchen[33] fest verankert.

---

27 Vgl. Art. 4 Abs. 4 DSG (Erkennbarkeit), Art. 8 DSG und Art. 1 f. VDSG (Auskunftsrecht), Art. 19 Abs. 2 revDSG (Informationspflicht bei der Beschaffung von Personendaten), Art. 25 revDSG (Auskunftsrecht; vgl. Art. 25 Abs. 2 lit. f revDSG: zu automatisierten Einzelentscheidungen), Art. 5 Abs. 1 lit. a DSGVO (Transparenz), Art. 12 ff. DSGVO (Informationspflichten; vgl. Art. 13 Abs. 2 lit. f und Art. 14 Abs. 2 lit. g DSGVO zu automatisierten Entscheidungsfindungen einschliesslich Profiling), Art. 15 DSGVO (Auskunftsrecht; vgl. Art. 15 Abs. 1 lit. h DSGVO zu automatisierten Entscheidungsfindung einschliesslich Profiling, Art. 11 DSG (Zertifizierungsverfahren), Art. 13 revDSG (Zertifizierung), Art. 22 f. revDSG (Datenschutz-Folgenabschätzung), Art. 35 DSGVO (Datenschutz-Folgenabschätzung), Art. 42 DSGVO (Zertifizierung).

28 Vgl. Botschaft vom 15. September 2017 zum Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz, BBl 2017 7010 f. Ziff. 9.1.2.

29 Z.B. Art. 400 (Rechenschaftslegung im Auftragsrecht), Art. 72 f. FIDLEG (Herausgabe von Dokumenten), Art. 97 FIDLEV (Herausgabe von Dokumenten).

30 Vgl. Michal Cichocki, Big Data und Datenschutz: Ausgewählte Aspekte, in: Jusletter IT vom 21. August 2015, S. 5.

31 Zur Problematik der Erklärbarkeit von KI-Technologien: vgl. Ron Schmelzer, Understanding Explainable AI, Forbes vom 23. Juli 2019, https://www.forbes.com/sites/cognitiveworld/2019/07/23/understanding-explainable-ai/.

32 Vgl. Art. 11a DSG (Register der Datensammlungen), Art. 10d DSG (Datenbearbeitung durch Dritte), Art. 11 DSG (Zertifizierungsverfahren), Art. 11 revDSG (Verhaltenskodizes), Art. 12 revDSG (Verzeichnis der Bearbeitungstätigkeiten), Art. 13 revDSG (Zertifizierung), Art. 22 f. revDSG (Datenschutz-Folgenabschätzung), Art. 24 revDSG (Meldung von Verletzungen der Datensicherheit), Art. 5 Abs. 2 DSGVO (Rechenschaftspflicht), Art. 30 DSGVO (Verzeichnis von Verarbeitungstätigkeiten), Art. 26 DSGVO (Gemeinsam Verantwortliche), Art. 28 DSGVO (Auftragsverarbeiter), Art. 33 f. DSGVO (Meldung von Verletzungen des Schutzes personenbezogener Daten an die Aufsichtsbehörde), Art. 35 DSGVO (Datenschutz-Folgenabschätzung), Art. 42 DSGVO (Zertifizierung) – vgl. auch Art. 400 OR (Rechenschaftslegung im Auftragsrecht).

33 Vgl. z.B. FINMA-RS 17/1 «Corporate Governance – Banken», FINMA-RS 2017/02 Corporate Governance – Versicherer», FINMA-RS 2018/03 «Outsourcing – Banken und Versicherer», Art. 15. f FIDLEG (Dokumentation und Rechenschaft), Art. 18.f FIDLEV (Dokumentation und Rechenschaft).

Dabei spielen Dokumentationspflichten wie Inventare, Assessments, Datenschutz-Folgeabschätzungen (DSFA), Monitoring sowie Kontrollen eine zentrale Rolle. Voraussetzung für eine robuste Accountability sind jedoch die Grundsätze der Transparency und vor allem der Explainability. Nur wenn feststeht, wer, was, wie macht, kann auch darüber Rechenschaft abgelegt werden.

[18] Folglich wird Accountability als rechtliche Vorgabe im datengetriebenen Umfeld von KI-Technologien weitgehend durch datenschutzrechtliche und/oder sektorspezifische Vorgaben hinreichend abgedeckt. Auch hier ist ein KI-spezifischer Handlungsbedarf aus rechtlicher Perspektive kaum erkennbar.

3.6. Human Control of Technology

[19] Die Befürchtung, KI könne den Menschen überflügeln und müsse deswegen rechtzeitig mit einer Art «Kill Switch» abgestellt werden können, ist nach wie vor verbreitet und fusst offenbar auf einem generellen Misstrauen[34] gegenüber Vorgängen ohne (vollständige) menschliche Kontrolle. Aus diesem Grund wird Human Control of Technology in 69% der KI-Guidelines als Grundsatz vorgegeben und soll insbesondere mit den Mitteln des Widerspruchs bzw. der menschlichen Überprüfung automatisierter Entscheide sichergestellt werden.

[20] Überall wo der Mensch und seine (Personen)Daten betroffen sind, gelangt das Datenschutzrecht zur Anwendung. Dieses teilt offenbar das erwähnte Misstrauen und sieht eine Reihe rechtlicher Vorgaben[35], insbesondere im Bereich der Betroffenenrechte vor, welche sich gegen die automatisierte aber auch manuelle Bearbeitung seiner Personendaten richten. Demnach ist Human Control of Technology bereits heute bzw. spätestens mit dem Inkrafttreten des revDSG weitgehend durch datenschutzrechtliche Vorgaben abgedeckt.

3.7. Professional Responsibility

[21] Wenn es nach 78% der untersuchten KI-Guidelines geht, steht der Mensch am Anfang einer Idee und übt einen wesentlichen Einfluss auf Design, Entwicklung, Funktionsweise sowie Zielsetzung von KI-Technologien aus. Damit dieser Einfluss «richtig» ausgeübt wird, sollen Grundsätze beispielsweise für (Daten)Richtigkeit, für die Berücksichtigung langfristiger Konsequenzen, für den Beizug unterschiedlicher Stakeholder oder für wissenschaftliche Integrität («scientific integrity») konkretisiert und anschliessend verwendet werden.

34 Vgl. Kathleen Walch, Will There Be A «Kill Switch» For AI?, Forbes vom 05. März 2020 (Online).

35 Vgl. Art. 12 Abs. 2 lit. b DSG (Widerspruchsrecht), Art. 15 DSG (Rechtsansprüche), Art. 21 Abs. 2 revDSG (Informationspflicht bei einer automatisierten Einzelentscheidung; «menschliches Gehör), Art. 30 Abs. 2 lit. b revDSG (Widerspruchsrecht), Art. 32 revDSG (Rechtsansprüche), Art. 17 DSGVO (Recht auf Vergessenwerden), Art. 18 DSGVO (Recht auf Einschränkung der Verarbeitung), Art. 21 DSGVO (Widerspruchsrecht), Art. 22 Abs. 3 DSGVO (Automatisierte Entscheidungen im Einzelfall einschliesslich Profiling).

[22] Rechtliche Vorgaben beispielsweise im Arbeitsrecht[36], Auftragsrecht[37], Datenschutzrecht (vgl. auch Privacy by Design und Default sowie Ziff. 4 unten)[38], Haftungsrecht[39] und Strafrecht[40] geben vor, dass sorgfältig, richtig und vorausschauend gearbeitet werden soll. Diese werden i.d.R. durch (nicht rechtliche) Regeln der Berufskunde, Branchenstandards, Usanzen usw. konkretisiert[41]. Damit dürfte der Grundsatz der Professional Responsibility aus rechtlicher Sicht hinreichend abgedeckt sein.

3.8. Promotion of Human Values

[23] Der letzte gemeinsame Nenner lautet Promotion of Human Values und wird in 69% der ausgewerteten Guidelines als KI-Grundsatz erwähnt.

[24] Dieser Grundsatz besagt im Wesentlichen, dass KI-Technologien einen positiven Einfluss auf den Menschen haben und seinem Fortkommen dienen sollten. Was genau darunter verstanden werden soll, ist jedoch jedem Unternehmen selbst überlassen.

[25] Innerhalb des anwendbaren Rechtsrahmen ist jedes Unternehmen frei, selbst zu definieren, wozu es KI einsetzen will und wie damit eine positive Wirkung erzielt werden kann. KI ist kein Selbstzweck, sondern lediglich ein (machtvolles) Mittel zur Erreichung eines oder mehrerer Ziele. Insofern ist der Grundsatz von Promotion of Human Values keine rechtliche Vorgabe, sondern ein *Auftrag* zur «freiwilligen» Definition sinnvoller geschäftspolitischer bzw. «ethischer» Ziele, welche mit dem Mittel von KI-Technologien erreicht werden sollen. Beispielsweise hat eine Gruppe von Wissenschaftlern, Ingenieuren sowie Unternehmern im Rahmen einer privaten Initiative[42] öffentlich versprochen, KI nicht für die Entwicklung, Herstellung und Handel von tödlichen, autonomen Waffen zu verwenden.

3.9. Fazit

[26] Sechs[43] der acht beleuchteten KI-Grundsätze sind weitgehende Wiederholungen anerkannter sowie fest verankerter, rechtlicher Vorgaben, die gerichtlich überprüft und gegebenenfalls

36 Vgl. Art. 321a OR (Sorgfalts- und Treuepflicht im Arbeitsrecht).

37 Vgl. Art. 321a OR (Sorgfalts- und Treuepflicht im Arbeitsrecht), Art. 398 OR (Sorgfaltspflicht beim Auftrag).

38 Vgl. Art. 5 DSG (Richtigkeit der Daten), Art. 6 Abs. 5 revDSG (Richtigkeit der Daten), Art. 5 Abs. 1 lit. d DSGVO (Richtigkeit), Art. 7 revDSG (Datenschutz durch Technik und datenschutzfreundliche Voreinstellungen), Art. 25 DSGVO (Datenschutz durch Technikgestaltung und durch datenschutzfreundliche Voreinstellungen), Art. 15 DSG (Rechtsansprüche), Art. 34 f. DSG (Strafbestimmungen), Art. 32 revDSG (Rechtsansprüche), Art. 60 ff. revDSG (Strafbestimmungen), Art. 77 ff. DSGVO (Rechtsbehelfe, Haftung und Sanktionen), Art. 11 DSG (Zertifizierungsverfahren), Art. 11 revDSG (Verhaltenskodizes), Art. 13 revDSG (Zertifizierung), Art. 22 f. revDSG (Datenschutz-Folgenabschätzung), Art. 35 DSGVO (Datenschutz-Folgenabschätzung), Art. 42 DSGVO (Zertifizierung).

39 Vgl. Art. 28 ff. ZGB (Persönlichkeitsverletzungen), Art. 41 ff. OR (Haftung für unerlaubte Handlungen), Art. 97 ff. OT (vertragliche Haftung).

40 Vgl. Art. 111 ff. StGB (Strafbare Handlungen gegen Leib und Leben), Art. 137 ff. StGB (Strafbare Handlungen gegen das Vermögen), Art. 173 ff. StGB (Strafbare Handlungen gegen die Ehre und den Geheim- oder Privatbereich), Art. 180 ff. StGB (Verbrechen und Vergehen gegen die Freiheit).

41 Vgl. z.B. Richtlinien für Integrität in der Forschung der ETH Zürich.

42 Vgl. Future of Life Institute, Lethal Autonomous Weapons Pledge, https://futureoflife.org/lethal-autonomous-weapons-pledge.

43 Privacy, Accountability, Safety and Security, Explainability and Transparency, Human Control of Technology, Professional Responsibility.

mit staatlichen (Zwangs)Mitteln durchgesetzt werden können. Die übrigen zwei[44] KI-Grundsätze entsprechen einer generellen Aufforderung, ethische Entscheide nach Massgabe individueller (Unternehmens)Ziele innerhalb des Rechtsrahmens als freiwillige Einschränkung des rechtlich zulässigen Maximums zu treffen. Folglich scheint es aus rechtlicher Sicht keinen Handlungsbedarf hinschlich der erwähnten KI-Grundsätze zu geben; der Mehrwert der untersuchten KI-Guidelines mag für den juristischen Praktiker eher gering ausfallen. Wieso ist das so?

3.10. Ein Erklärungsversuch

[27] Ob und inwieweit rechtliche Vorgaben anwendbar sind, bestimmt sich nach der Erfüllung ihrer Tatbestandsvoraussetzungen. Diese sind generell-abstrakt sowie technologieneutral formuliert, dadurch viel beständiger und müssen nicht laufend angepasst werden. Rechtliche Vorgaben knüpfen folglich in der Regel nicht an die Verwendung bestimmter Technologien an, sondern an das Vorliegen konkreter Lebenssachverhalte und Anwendungsfälle. Ob dabei (KI)-Technologien verwendet werden oder nicht, spielt mangels spezifischer Rechtsgrundlagen für KI mindestens derzeit[45] keine entscheidende Rolle[46]. Ferner ist die Anwendung von (KI-)Technologien auf mannigfache Lebenssachverhalte denkbar. Im Gegenzug sind auch die potenziell anwendbaren rechtlichen Vorgaben vielseitig. Dieser Umstand macht die Formulierung fassbarer (KI-)Guidelines zur Regelung allgemeiner Lebenssachverhalte sehr schwierig. Es gilt nämlich der Lieblingssatz des Juristen: Es kommt auf den Einzelfall an. Schliesslich ist in der Praxis eine Abgrenzung zwischen rechtlichen sowie («freiwilligen») ethischen Vorgaben wichtig: Während die Einhaltung rechtlicher Vorgaben mittels staatlicher Zwangsmittel durchgesetzt und deren Nichteinhaltung sanktioniert werden kann, ziehen Verstösse gegen ethische Ziele (vgl. Ziff. 3.3) vorwiegend Reputationsschäden nach sich.

4. Aspekte der Operationalisierung in der Unternehmenspraxis

4.1. Three Lines (of Defense) Model

[28] Im nicht-rechtlichen Bereich gibt es jedoch sehr wohl Handlungsbedarf im Zusammenhang mit KI-Guidelines. Dabei spielt das Three Lines[47] (of Defense) Model[48] für ein effektives Risikomanagement und eine wirksame Corporate Governance eine wichtige Rolle: Wurde ein Use

44 Non-Discrimination sowie Promotion of Human Values.

45 Für die EU: vgl. Michal Cichocki, White Paper zu Künstlicher Intelligenz (KI) Europäische Kommission veröffentlicht Konsultationsergebnisse, 19. Juli 2020, http://www.lawblogswitzerland.ch/2020/07/white-paper-zu-kunstlicher-intelligenz.html.

46 Vgl. Ziff. 4.1/(v) Bericht zu «Herausforderungen der künstlichen Intelligenz» der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat, 13. Dezember 2019.

47 Das breit anerkannte «Three Lines Model» (bis 2020: «Three Lines *of Defense Model*») des Institute of Internal Auditors (IAA) hilft Organisationen aus allen Branchen, Strukturen und Prozesse zu identifizieren, welche die individuelle Zielerreichung am besten unterstützen und dabei eine starke Governance und ein starkes Risikomanagement ermöglichen. Dieses Modell ist für alle Organisation anwendbar. Vgl. Das Drei-Linien-Model des IAA, Eine Aktualisierung der Three Lines of Defense, 2020, https://global.theiia.org/about/about-internal-auditing/Pages/Three-Lines-Model.aspx.

48 Vgl. Ruud, Flemming & Kyburz, Adrian (2014) Gedanken zum Three Lines of Defense Model – Was ist mit Verteidigung gemeint? – Analyse des Governance-Models aus der Sicht des internen Audits. Der Schweizer Treuhänder, Band 88, S. 762.

Case innerhalb der ethischen Unternehmensziele definiert, erfolgt die Identifizierung sowie Auslegung der relevanten rechtlichen Vorgaben durch die Legal-Funktion als 2nd Line (of Defense). Damit steht fest «was» gemacht werden soll. Doch «wie» soll dafür vorgegangen werden?

[29] Die eigentliche Arbeit steht also erst noch bevor: Die rechtlichen Vorgaben sowie ethischen Ziele («was») müssen in operationalisierbare Vorgaben für technische und/oder organisatorische Massnahmen (TOM-Vorgaben) abgeleitet werden. Diese TOM-Vorgaben beantworten die Frage, «wie» vorgegangen werden soll. Zu diesem Zweck kommen neue, dedizierte Funktionen[49] der 1st Line (of Defense), d.h. das «Business» (Risikobewirtschafter), und damit ausserhalb der Legal-Funktion zum Zug. Sie verfügen über das erforderliche prozessuale sowie technische Know-How, setzen sich mit den erwähnten rechtlichen Vorgaben und ethischen Zielen auseinander, ziehen eigene, nicht-rechtliche Regeln ihrer jeweiligen Berufskunde (vgl. Ziff. 3.2 und 3.7), Branchenstandards etc. bei und leiten daraus eigenverantwortlich angemessene TOM-Vorgaben als Steuerungs- und Kontrollelemente ab. Schliesslich stellt das interne Audit die 3rd Line (of Defense) und damit eine unabhängige, objektive Prüfung sowie Beratung sicher.

4.2. Beispiel

[30] Am Beispiel der rechtlichen Vorgabe[50], wonach Personendaten u.a. dann gelöscht werden sollen, sobald ihr Bearbeitungszweck erreicht ist, kann dieses Zusammenspiel wie folgt skizziert werden: die Legal-Funktion (2nd Line) legt beispielsweise aus, was ein Personendatum ist und wann es aus rechtlicher Sicht als gelöscht[51] gilt. Die 1st Line leitet daraus z.B. ein Löschkonzept zwecks Berechnung konkreter Aufbewahrungs- bzw. Löschfristen (d.h. eine Vorgabe für organisatorische Massnahmen) ab und definiert eine Funktionalität, z.B. mittels Architecture Building Blocks, zur Durchführung einer hinreichenden Löschung (d.h. eine Vorgabe für technische Massnahmen). Diese eigenständigen TOM-Vorgaben der 1st Line werden anschliessend im Einzelfall entlang der bestehenden Aufbau- und Ablauforganisation der jeweiligen Unternehmung umgesetzt.

[31] Dieser Mechanismus aus der Unternehmensrealität ist auch dem Gesetzgeber nicht fremd. Er entspricht einem Teilaspekt der rechtlichen Vorgabe von Privacy by Design und Default[52] und verpflichtet den datenschutzrechtlichen Verantwortlichen (d.h. das Unternehmen – nicht den Datenschutzbeauftragten)[53] frühzeitig, die Bearbeitung von Personendaten mittels TOM-Vorgaben so «*auszugestalten, dass die Datenschutzvorschriften eingehalten werden*».

49 Business Analysten, Business Engineers usw. Die Bezeichnung variiert je nach Organisation.

50 Vgl. Art. 6 Abs. 4 revDSG, Art. 4 Abs. 2 DSG oder Art. 5 Abs. 1 lit. e DSGVO.

51 Vereinfacht: wenn der Personenbezug fehlt bzw. nur noch mit unverhältnismässig grossem Aufwand wiederhergestellt werden kann (s. auch den «Logistep-Entscheid» zur sog. *relativen Methode*: BGE 136 II 508).

52 Vgl. Art. 7 revDSG (Datenschutz durch Technik und datenschutzfreundliche Voreinstellungen) oder Art. 25 DSGVO (Datenschutz durch Technikgestaltung und durch datenschutzfreundliche Voreinstellungen).

53 Diejenige natürliche oder juristische Person, welche «über den Zweck und den Inhalt der Datensammlung» (Art. 3 lit. I DSG) bzw. «über den Zweck und die Mittel der Bearbeitung» (Art. 5 lit. j revDSG) bzw. « über die Zwecke und Mittel der Verarbeitung von personenbezogenen Daten» (Art. 4 ziff. 7 DSGVO) entscheidet.

4.3. Ausblick

[32] In Zeiten komplexer Lebenssachverhalte, zunehmender Regulierung sowie vermehrter ethischer Fragestellungen stellt der Einsatz neuartiger Technologien (wie z.B. KI) immer höhere Anforderungen an die Ableitung und anschliessende Operationalisierung praktikabler Vorgaben für technische/organisatorische Massnahmen (TOM-Vorgaben). Gerade in diesem Bereich könnten (KI)-Guidelines einen wertvollen Beitrag leisten, um zur Entwicklung von Standards oder Branchenusanzen beizutragen.

Michal Cichocki, MLaw, Rechtsanwalt, ist Leiter Legal Datenschutz & Outsourcing und DPO bei einer schweizerischen Bank. Der vorliegende Aufsatz gibt ausschliesslich die persönliche Auffassung des Autors wieder.