NETZWERK DATENSCHUTZEXPERTISE

Stellungnahme zum

Draft Ethics Guidelines for Trustworthy Artificial Intelligence

der European Commission´s High Level Expert Group on Artificial Intelligence (HLEG) vom 18.12.2018

Entwurf von Ethikrichtlinien über vertrauenswürdige Künstliche Intelligenz

der Hochrangigen Expertengruppe zu Künstlicher Intelligenz

der Europäischen Kommission

Stand: 21.02.2019

Thilo Weichert
Waisenhofstraße 41, 24103 Kiel
24103 Kiel
weichert@netzwerk-datenschutz-expertise.de
www.netzwerk-datenschutzexpertise.de

Stichworte

Die EU-Kommission hat eine hochrangige Expertengruppe (HLEG) zu Künstlicher Intelligenz (KI) einberufen. Diese veröffentlichte am 18.12.2018 ethische Grundsätze, die beim Einsatz von KI zu beachten sind.

https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=57112

https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai

Dieser Entwurf von Guidelines, also von Leitlinien, wurde der Öffentlichkeit zur Kommentierung bis zum 18.01.2019 bereitgestellt. Nach öffentlicher Kritik an der kurzen Kommentierungsphase über Weihnachten und über den Jahreswechsel wurde die Kommentierungsfrist bis zum 05.02.2019 verlängert. Danach soll durch die HLEG eine Überarbeitung der Guidelines erfolgen. Das Netzwerk Datenschutzexpertise nahm das Angebot zur Kommentierung an und übermittelte der EU-Kommission die folgende Stellungnahme.

(1) Das Ziel, auf EU-Ebene gemeinsame **ethische Grundlagen** zu erarbeiten, die bei der Erforschung, der Entwicklung und dem Einsatz von Künstlicher Intelligenz (KI) zu beachten sind, ist zu begrüßen.

(2) Zu kritisieren ist, dass sich die Leitlinien **auf KI beschränken**, also auf informationstechnische Systeme, die aus Sensoren oder anderen Quellen stammende Daten aufbereiten und hieraus „selbstlernend" ursprünglich von Menschen gestaltete Algorithmen verändern und auf dieser Grundlage automatisierte Schlussfolgerungen bzw. Ergebnisse gewinnen, die zur Grundlage von relevanten praktischen Entscheidungen genommen werden (können).

(3) KI ist eine Weiterentwicklung von **hochkomplexen Algorithmen.** Bei komplexen Algorithmen, die nicht auf selbstlernenden, sondern auf vorgegebenen ausdifferenzierten Datenauswertungsprozessen beruhen, bestehen ähnliche Herausforderungen, wie sie von der HLEG bzgl. KI in den Guidelines thematisiert werden. So werfen z. B. nicht auf KI-Basis funktionierende Scoring-Verfahren ethische Probleme fehlender Verantwortlichkeit, Zurechenbarkeit, Transparenz und Kontrollierbarkeit auf.

(4) Das Thema von ethischen Leitlinien sollte daher nicht auf „künstliche Intelligenz" beschränkt werden, sondern generell den Einsatz komplexer Algorithmen umfassen. Zielsetzung der EU sollte es demnach sein, generell einen normativen Rahmen für den Algorithmeneinsatz und die **Algorithmenkontrolle** zu definieren.

(5) Bei datengetriebener KI besteht wegen des dauernden Prozesses der **Selbstoptimierung durch Datenauswertung** und der Nachjustierung der Entscheidungsfindung eine noch geringere Bestimm- und Bachvollziehbarkeit als bei determinierten digitalen Prozessen. Dadurch verschärfen sich die generellen Probleme automatisierter Entscheidungen in Bezug auf Protokollierung, Transparenz, Verantwortlichkeit und Haftung. Diskriminierungseffekte können nicht nur durch die Programmierung bewirkt werden, sondern solche Effekte werden durch einfließende Daten aus realer Diskriminierung verstärkt. Eine Dokumentation und Nachvollziehbarkeit der Entscheidungsfindung ist nicht mehr gewährleistet. Eine individuelle Verantwortung für Einzelentscheidungen wird vorverlagert von der Systemgestaltung hin zur Entscheidung über das „Ob" eines Systemeinsatzes.

(6) Daher bedarf es hinsichtlich des KI-Einsatzes **weitergehender Restriktionen** oder Vorkehrungen. Für bestimmte Zwecke ist der Einsatz von KI-Technologie wegen der damit verbundenen Konsequenzen überhaupt nicht ethisch vertretbar und muss deshalb absolut ausgeschlossen werden.

Dies gilt z. B. für den militärischen KI-Einsatz bei tödlichen Waffen; dies gilt aber auch bei nicht-militärischen Nutzungen, wenn die per KI getroffenen Entscheidungen existenzielle Bedeutung für Menschen haben und keine Revidier- bzw. Kompensierbarkeit besteht.

(7) So wichtig ethische Standards sind, so bleiben diese unverbindlich, wenn sie nicht in bestimmte Gesetze oder sonstiges **zwingendes Regelungen** umgesetzt werden, die demokratisch zustande gekommen sind, deren Einhaltung unabhängig kontrolliert und deren Verletzung effektiv sanktioniert wird. Dieser Prozess der Operationalisierung der Leitlinien wird in den vorliegenden Guidelines nicht thematisiert. Ohne diese Operationalisierung besteht die Gefahr, dass den ethischen Leitlinien ein reiner Alibicharakter zukommt und dass diese zur Legitimation für ethisch problematische Techniknutzungen eingesetzt werden.

(8) Nicht nur der Prozess der Normsetzung wird in den Guidelines übergangen, sondern weitgehend auch die Relevanz von Normen generell: Um vertrauenswürdige KI zu erlangen, wird in erster Linie auf **technische Methoden** gesetzt. Dabei wird zutreffend differenziert zwischen Technikgestaltung, Architekturen, Testung, Bewertung, Dokumentation und Erklärbarkeit (S. 19 f.). Zu kurz kommen die Prozesse der regelmäßigen Kontrolle und Evaluation, die bei KI als Systemen, die auf lernenden, also sich ändernden Algorithmen basieren, besonders wichtig sind.

(9) Hinsichtlich der **nicht-technischen Methoden** wird auf Standarisierung, Governance, Verhaltensregeln, Erziehung und auf einen gesellschaftlichen pluralen Diskurs Bezug genommen (S. 21 f.). Diese Methoden sind zu ergänzen durch eine unabhängige Zertifizierung (s. u. Rn. 18) und eine unabhängige menschliche Kontrolle (s. u. Rn. 17). Die Notwendigkeit demokratisch getroffener Regeln bzw. Gesetze wird nicht ausdrücklich, sondern nur in sehr allgemeiner Form thematisiert (S. 21). Tatsächlich ist ein klarer, mit Verboten und Geboten, technisch-organisatorischen Vorgaben und prozeduralen Regeln festgelegter gesetzlicher Rahmen, dessen effektive Einhaltung gewährleistet wird, die zentrale Grundlage eines vertrauenswürdigen Einsatzes komplexer Algorithmen.

(10) Die Leitlinien benennen richtig als ethische Vorgaben die Grundrechte, insbesondere die Menschenwürde, die Freiheits- und Bürgerrechte, die Diskriminierungsverbote, sowie die Grundsätze von Demokratie, Rechtsstaatlichkeit und Solidarität (S. 7). Diese Grundsätze haben ihre verfassungsrechtliche Grundlage in der seit 2009 wirksamen europäischen Grundrechte-Charta (GRCh) gefunden. Nicht thematisiert wird die weitergehende Frage, inwieweit es durch die Digitalisierung einer Weiterentwicklung der verfassungsrechtlichen Normierung bedarf. Der insofern gestartete Prozess der **Formulierung digitaler Grundrechte** (https://digitalcharta.eu/) muss in den weiteren Diskussionen über die vorliegenden Guidelines ein zentraler Aspekt sein.

(11) Im Rahmen dieser verfassungsrechtlichen Diskussion bedarf es der Erörterung, inwieweit das Grundrecht auf Datenschutz, das vom deutschen Bundesverfassungsgericht (BVerfG) als „Recht auf informationelle Selbstbestimmung“ definiert wurde (BVerfG 15.12.1983 – 1 BvR 209/83 u. a.), um ein „**Recht auf digitale Souveränität**“ zu ergänzen ist, das auch juristischen Personen zusteht und nicht nur für von digitaler Verarbeitung Betroffene gilt, sondern auch für solche Techniken (verantwortlich) Anwendende (also Nutzende). Digitale Souveränität ist ein Ziel, das nicht nur für die Objekte von Datenverarbeitung (also Betroffene im datenschutzrechtlichen Sinn) realisiert werden muss, sondern auch für die Systemnutzenden als Subjekte. Ein zentrales Problem des Einsatzes künstlicher Intelligenz besteht darin, dass die diese (verantwortlich) Nutzenden auf die Technikbereitstellung durch Anbieter

angewiesen sind, deren Angebot sie weder bewerten und einschätzen, geschweige denn verantworten können. Die Idee wird in den Guidelines nur angedeutet (S. 9 f.).

(12) Das Prinzip der Erklärbarkeit bzw. der **Transparenz** von KI ist ein Grundanliegen der Guidelines (erstmals S. 10, dann z. B. S. 18). Dieses Prinzip ist eine Grundvoraussetzung nicht nur für KI, sondern für digitale Datenverarbeitung generell. Diese Prinzip ist auch grundlegend für die Realisierung des Grundrechts auf Datenschutz (Art. 8 GRCh) und ein zentrales Anliegen der dieses Grundrecht umsetzenden Europäischen Datenschutzgrundverordnung (DSGVO, dort z. B. Art. 5 Abs. 1 lit. a, 12 ff.).

(13) Entgegen einer weit verbreiteten Wahrnehmung dient die DSGVO nicht nur dem Schutz des Grundrechts auf Datenschutz, sondern dem Schutz „aller Grundrechte und Grundfreiheiten natürlicher Personen“ bei der personenbeziehbaren Datenverarbeitung (Art. 1 Abs. 1 DSGVO). Die Guidelines reduzieren die Anwendung der **DSGVO** auf den Respekt von Privatheit (Privacy, S. 17, 25). Dadurch wird auch ignoriert, dass die DSGVO sämtliche relevanten Bewertungskriterien vertrauenswürdiger KI einer Regulierung zuführt: Verantwortlichkeit, Design, Selbstbestimmung, die Verhinderung von Diskriminierung, Robustheit und Richtigkeit, Sicherheit und Transparenz (S. 24-27).

(14) Die DSGVO thematisiert umfassend den Grundrechtsschutz von Betroffenen bei personenbeziehbarer Datenverarbeitung sowie die damit verbundenen gesellschaftlichen Konsequenzen. Grundrechtsrelevante Wirkungen entfalten sich nicht nur bei der Verarbeitung personenbezogener Daten, sondern auch, wenn automatisierte Entscheidungen ganze Personenkollektive betreffen und hierbei **sächliche oder vollständig anonymisierte Daten verarbeitet** werden. In der weiteren Diskussion müssen anwendungs- und zweckbezogen die Bereiche identifiziert werden, in denen derartige Anwendungen eine derartige Relevanz entwickeln, dass regulative Ergänzungen zu den bestehenden Regelungen zur Verarbeitung personenbezogener Daten nötig sind (z. B. in den Bereichen Mobilität, Umweltschutz, Nahrungsmitteleinsatz, Biotechnologieeinsatz).

(15) Mit der DSGVO besteht bisher schon ein verbindlicher gesetzlicher Rahmen für den Einsatz von KI in Bezug auf **Profiling und automatisierte Entscheidungen** mit personenbeziehbaren Daten. In Art. 22 DSGVO werden Abwägungsanforderungen benannt, die bei der Gestaltung, dem Einsatz und der Nutzung von KI einfließen müssen: individuelle Selbstbestimmung (Einwilligung), Eingreif- und Revisionsmöglichkeit (z. B. Ausschaltknopf), Ersetzungsmöglichkeit durch einen menschlichen Entscheider, besonderer Schutz beim Einsatz sensitiver Daten, Rechtsschutz). In Art. 15 Abs. 1 lit. h DSGVO wird das Recht auf Auskunft über „die involvierte Logik sowie die Tragweite und die angestrebte Auswirkungen“ begründet.

(16) Ein weitergehender Rechtsrahmens für den Einsatz von Algorithmen im Allgemeinen und KI im Speziellen sollte daher auf **diesen bestehenden Normen aufbauen**. Die DSGVO gibt hierfür den notwendigen Spielraum (vgl. Art. 22 Abs. 2 lit. b DSGVO). Damit wird zugleich gewährleistet, dass weitere verfassungsrechtliche Anforderungen, die in den Guidelines nicht oder nur andeutungsweise erwähnt werden, beachtet werden. Dies gilt insbesondere für den Auskunftsanspruch der Betroffenen bzw. in einem erweiterten Verständnis der Anwendenden als „digitalen Souveräne“ (s. o. Rn. 11) sowie für die unabhängige staatliche Kontrolle (Art. 8 Abs. 2 S. 1 u. Abs. 3 GRCh).

(17) Die Notwendigkeit einer **unabhängigen staatlichen Kontrolle** wurde beim Datenschutz schon früh vom deutschen Bundesverfassungsgericht (BVerfG) verfassungsrechtlich begründet, insbesondere für den Einsatz digitaler Technik durch staatliche Einrichtungen (erstmals BVerfG 15.12.1983 – 1 BvR

209/83 u. a.). Sie wurde vom Europäischen Gerichtshof (EuGH) mehrfach eingefordert (EuGH 09.03.2010 – C-203/15 u. C-698/15, 16.10.2012, - C-614/10, 08.04.2014 – C-288/12). Die diese Rechtsprechung tragenden Erwägungen lassen sich auf die Kontrolle von KI generell im öffentlichen wie im privaten Bereich übertragen.

(18) In der DSGVO ist in den Art. 42 f. der rechtliche Rahmen für die **Zertifizierung** komplexer informationstechnischer Systeme durch eine freiwillige unabhängige Überprüfung festgelegt. Die hierfür nötigen Instrumente müssen umgehend in der Realität umgesetzt und angewendet werden.

(19) Für den grundwertekonformen Einsatz von KI genügt in vielen Bereichen eine freiwillige Zertifizierung nicht. Es bedarf, wie beim Technikeinsatz in anderen gesellschaftlichen Bereichen üblich (z. B. bei der Mobilität, bei Emissionen, beim Gentechnikeinsatz, bei Arzneimitteln) einer darüber hinausgehenden bereichsspezifischen Regulierung mit **Melde-, Genehmigungs- und Evaluationspflichten** und einer einsprechenden hoheitlichen Kontrolle.

(20) Aus der Diskussion in den USA kommend, wird auch in Europa teilweise die Position vertreten, von Algorithmen errechnete Ergebnisse könnten den Schutz der **Meinungsfreiheit** (Art. 11 GRCh) für sich in Anspruch nehmen. Diese Argumentation wird eingesetzt, um eine stärkere Regulierung von KI bzw. eine verstärkte Algorithmenkontrolle zurückzuweisen. Es ist notwendig, in den Guidelines klarzustellen, dass die Nutzung von Ergebnissen digitaler Datenverarbeitung, insbesondere von KI, für sich nicht das Grundrecht auf Meinungsfreiheit in Anspruch nehmen kann.

(21) Die Guidelines vermeiden bei dem Ziel der Herstellung vertrauenswürdiger KI bzw. generell von vertrauenswürdigen digitalen Entscheidungsprozessen eine Aussage zu einer grundlegenden Fragestellung: Von Verantwortlichen wird dem Transparenzerfordernis der Schutz von **Betriebs- und Geschäftsgeheimnissen** entgegengesetzt. Tatsächlich hat z. B. das oberste deutsche Zivilgericht, der Bundesgerichtshof, entschieden, dass Betriebs- und Geschäftsgeheimnisse selbst Transparenzforderungen an digitale Prozesse entgegen gehalten werden können, die von datenschutzrechtlich Betroffenen geltend gemacht werden (BGH 28.01.2014 – VI ZR 156/13, BGH 22.02.2011 – VI ZR 120/10). Berechtigte Verfassungsklagen hierzu wurden bisher vom deutschen BVerfG nicht angenommen (dazu Weichert DatenschutzNachrichten 2/2018, 134). Der EuGH hat sich mit dieser Problematik bisher nicht befasst.

(22) Ein zentrales Problem beim Einsatzes von KI ist, dass die in den Guidelines aufgeführten ethischen Werte bei vielen konkreten KI-Einsätzen in der Praxis nicht beachtet werden, weil die diese Verfahren einsetzenden Unternehmen, bei denen es sich sehr oft um mächtige Wirtschaftsunternehmen aus den USA wie Google, Facebook, Amazon oder Microsoft handelt, sich bisher erfolgreich weigern, die zur Umsetzung der Grundrechte und der demokratischen Kontrolle nötige Transparenz herzustellen. Durch diese **Transparenzverweigerung**, für die angeblich bestehende ökonomische Rechte ins Feld geführt werden, wird eine wirksame Rechtskontrolle unmöglich gemacht. Um dieses zentrale Problem in den Griff zu bekommen, bedarf es klarer gesetzlicher Offenlegungspflichten der einsetzenden Unternehmen gegenüber der demokratischen Öffentlichkeit bzw. gegenüber staatlichen Stellen sowie einer hinreichenden Ausstattung der unabhängigen Aufsicht, damit diese Pflichten auch praktisch durchgesetzt werden können. Eine europäische Regulierung ist wegen der europa-, ja weltweiten Bedeutung des Problems wünschenswert. Demokratie- und Grundrechtskonformität muss Vorrang haben vor Marktverfügbarkeit, Wettbewerb und ökonomischem Nutzen.

(23) Durch eine Vorverlagerung des Risikos beim KI-Einsatz von der Gestaltung des Einsatzes digitaler Technik hin zur Entscheidung, ob diese eingesetzt wird, muss im Sinne eine Gefährdungshaftung zumindest eine **zivil- und verwaltungsrechtliche Verantwortlichkeit** gesetzlich begründet werden. Die KI einsetzenden Stellen müssen per Gesetz spezifischen Gestaltungs- und Unterlassungspflichten sowie einer umfassenden Haftung unterworfen werden.

Eingereicht bei der Europäischen Kommission am 23.01.2019