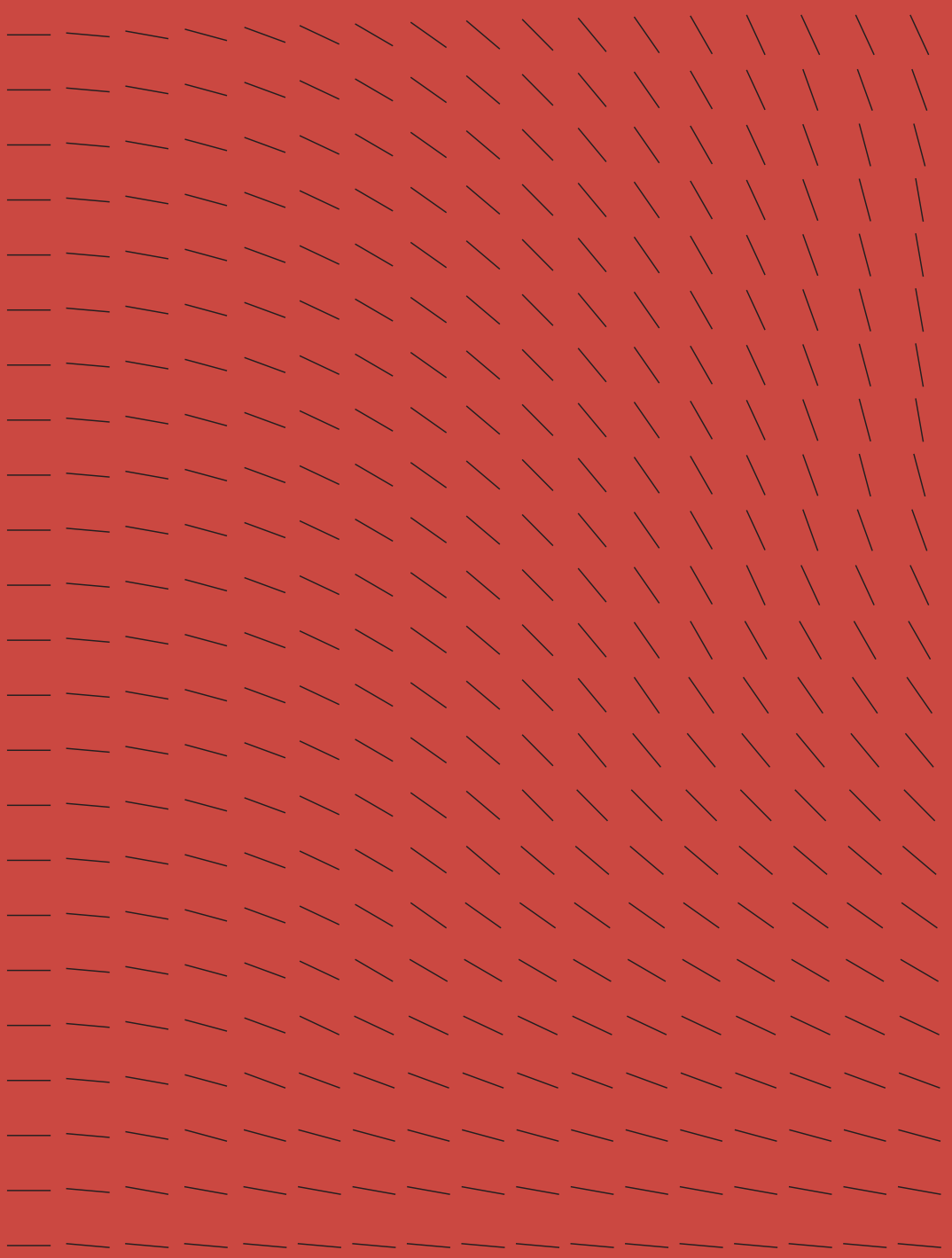
weizenbaum
institut



Weizenbaum Series #12

Thesenpapier

Die Regulierung Künstlicher Intelligenz - Neuer Rechtsrahmen für Algorithmische Entscheidungssysteme?

Ferdinand Müller, Martin Schüßler, Elsa Kirchner

Oktober 2020

HERAUSGEBER

Der Vorstand des Weizenbaum-Instituts e.V.

Prof. Dr. Christoph Neuberger
Prof. Dr. Sascha Friesike
Prof. Dr. Herbert Zech
Dr. Karin-Irene Eiermann

Hardenbergstraße 32
10623 Berlin

Tel.: +49 30 700141-001
E-Mail: info@weizenbaum-institut.de
Web: www.weizenbaum-institut.de

AUTOREN

Ferdinand Müller
ferdinand.mueller@rewi.hu-berlin.de

Martin Schüßler
schuessler@tu-berlin.de

Elsa Kirchner
elsa.kirchner@dfki.de

REDAKTION / LAYOUT UND SATZ

Roland Toth
Filip Stiglmayer

DOI 10.34669/wi.ws/12

Diese Arbeit wurde durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert (Förderkennzeichen: 16DII121, 16DII122, 16DII123, 16DII124, 16DII125, 16DII126, 16DII127, 16DII128 – „Deutsches Internet-Institut“).

Weizenbaum Series #12
Working Paper

Die Regulierung Künstlicher Intelligenz - Neuer Rechtsrahmen für Algorithmische Entscheidungssysteme?

Ferdinand Müller[1], Martin Schüßler[2], Elsa Kirchner[3]

Oktober 2020

1 Ferdinand Müller arbeitet am interdisziplinären *Weizenbaum-Institut für die vernetzte Gesellschaft* und beschäftigt sich mit den Grundlagen künstlicher Intelligenz im Rechtsverkehr.

2 Martin Schüßler, ebenfalls vom *Weizenbaum-Institut*, ist Informatiker und Mensch-Maschine-Interaktions-Forscher im Bereich *explainable AI.*

3 Dr. Elsa Kirchner vom *Deutschen Forschungszentrum für Künstliche Intelligenz* (DFKI) ist Biologin und hat in der Informatik im Bereich der Mensch-Maschine-Interaktion promoviert. Sie arbeitet an Methoden des *Interactive Machine Learning.*

Die Diskussion um neue Gesetze für die Technologien der Künstlichen Intelligenz läuft auf Hochtouren. Allerorts werden Strategien, Leitlinien und Empfehlungen veröffentlicht. Doch was macht die Technik in ihrem Wesenskern aus? Was ist ihr spezielles Risiko? Die Autor*innen wollen fünf Thesen für eine Regulierung aufstellen, die bei der Suche und der Auswahl eines neuen Rechtsrahmens bedacht werden müssen.[1]

1. Technologie “des täglichen Lebens”

Täglich ist ein Großteil der Bevölkerung in entwickelten Ländern von der Anwendung solcher Technik betroffen - oftmals ohne, dass dies bewusst zur Kenntnis genommen wird. Soziale Netzwerke, Online-Shopping, Nachrichtenkonsum, Kommunikation - in all diesen Bereichen werden mittlerweile völlig routiniert *teil- und vollautomatisierte Algorithmische Entscheidungssysteme* (AES) eingesetzt. Auch in Europa dienen AES mittlerweile nicht nur Unternehmen zur Verfolgung kommerzieller Zwecke, sondern werden schon heute zur Durchführung und Unterstützung staatlicher Tätigkeiten verwendet.

Uns erwartet damit eine Zukunft, in der wir auch im Alltag immer mehr auf die Technik angewiesen sind. Möglicherweise könnten wir sogar in bestimmten Situationen zur Interaktion gezwungen sein. Umso drängender erscheint damit die Frage, ob die bestehenden Gesetze einen ausreichenden Schutz vor den Risiken dieser Technologie bieten. Gleichzeitig bietet die Technik ein enormes Potenzial.

2. Aktuelle Gesetzeslage ist unzureichend

In der Rechtswissenschaft wird bereits seit einiger Zeit diskutiert, welche neuen Risiken durch den zunehmenden Gebrauch der Technologie für immaterielle wie materielle Rechtsgüter entstehen und wie diese regulatorisch erfasst werden können.[2]

---

1 Das Thesenpapier beruht auf dem Beitrag „Ein “KI-TÜV” für Europa? Eckpunkte einer horizontalen Regulierung Algorithmischer Entscheidungssysteme”, in: Asmussen/Golla/Kuschel, Tagungsband GRUR Jr. 2020, Baden-Baden 2020, S. N.N.

2 *L. Käde/S. von Maltzan*, Die Erklärbarkeit von Künstlicher Intelligenz (KI), CR 2020, S. 66 ff.; *A. Allar*, Rechtliche Herausforderungen Künstlicher Intelligenz, ZUM2020, S. 325ff.; *H. Zech,* Risiken Digitaler Systeme: Robotik, Lernfähigkeit und Vernetzung als aktuelle Herausforderungen für das Recht, Weizenbaum Insights 2020, abrufbar unter: https://kurzelinks.de/i2sq;*B. Jakl,* Das Recht der Künstlichen Intelligenz - Möglichkeiten und Grenzen zivilrechtlicher Regulierung, MMR2019, S. 711ff.; *S. Meyer,* Künstliche Intelligenz und die Rolle des Rechts für Innovation, ZRP 2018, S. 233 ff.; *G. Borges*, Rechtliche Rahmenbedingungen für autonome Systeme, NJW 2018, S. 977 ff.; *T. Burri,* Künstliche Intelligenz und internationales Recht, DuD 2018, S. 603 ff.; *M. Herberger*, “Künstliche Intelligenz” und Recht, NJW 2018, S. 2825 ff.; *M. Martini*,

Herkömmliche Zertifizierungs- und Zulassungsverfahren stehen auf dem Prüfstand und müssten angepasst werden. Die Mehrheit der Befragten einer repräsentativen Umfrage in Deutschland im Januar 2020 spricht sich für die Einrichtung einer besonderen Prüfstelle und eine damit einhergehende staatliche Regulierung aus.[3] Die Europäische Kommission hat die Problemlage erkannt und arbeitet u.a. mit einer hochrangigen Expertengruppe an einer gesamteuropäischen Lösung. Zuletzt erfolgte eine Öffentliche Konsultation im Rahmen der Veröffentlichung eines Weißbuches zum aktuellen Stand der Debatte.[4]

Des Öfteren wird auch argumentiert, dass die 2018 in Kraft getretene europäische *Datenschutz-Grundverordnung* für die Handhabung der Technologie ausreichen würde. Die Verordnung konzentriert sich jedoch auf den Schutz personenbezogener Daten und die damit verbundene Wahrung des informationellen Selbstbestimmungsrechtes. Deswegen ist sie nur unzureichend geeignet für die Handhabung der durch AES entstehenden Risikobereiche.

3. Definition und Abgrenzung der Technologie entscheidend

Als eine Ausprägung moderner KI-Technologie stimmen die technischen Prämissen von *Algorithmischen Entscheidungssystemen* in den Grundlagen mit denen der elektronischen Datenverarbeitung überein: Maschinell lesbare Informationen in Form von Daten werden analog oder digital durch das System aufgenommen, verarbeitet und in einem sich analog oder digital manifestierenden Ergebnis ausgegeben. Die Verbesserung der Computertechnik und die massenhafte Nutzung von personenbezogenen Daten ermöglichten jedoch bedeutende Fortschritte.

Im Unterschied zur bisherigen Technik können solche Systeme im höheren Maße als bislang selbstständig bei der Informationsgewinnung, -verarbeitung und -ausgabe agieren, sodass auch die Bearbeitung großer Datenmengen effizient möglich ist ("Big Data"). Moderne AES verwenden zur Datenverarbeitung teilweise vielschichtige, nicht-lineare Modelle ("Artificial Neural Networks"), die auch die Abbildung äußerst komplexer Prozesse ermöglicht, sodass etwa das Verständnis menschlicher Sprache oder die Bilderkennung ermöglicht wird. Manche AES sind zudem in der

Algorithmen als Herausforderung für die Rechtsordnung, JZ2017, S. 1017ff.; zur Diskussion in den USA siehe *A. Tutt*, An FDA For Algorithms, Administrative Law Review 69 (2017), S. 83ff.

3 "Sicherheit und Künstliche Intelligenz", Studie des VdTÜV vom Januar 2020, abrufbar unter: https://kurzelinks.de/fm13

4 Dokument abrufbar unter: https://kurzelinks.de/e5uu.

Lage, ihre technischen Parameter im Laufe der Verwendung zu justieren und zu optimieren, so dass diese Systeme auch bei sich stetig verändernden Rahmenbedingungen effektiv eingesetzt werden können.
In den meisten Fällen ist ein Mensch Bestandteil des Systems, welcher die maschinell erstellte Entscheidungsempfehlung endgültig formuliert oder umsetzt, so dass ein *teilautomatisiertes* AES vorliegt. Einige AES agieren jedoch in ihrem Kernbetriebsablauf gänzlich ohne einen menschlichen Operator, so dass von *vollautomatisierten* AES gesprochen werden sollte.

4. Entstehen neuer Risiken

Diese von herkömmlicher Computertechnik zu unterscheidenden Eigenschaften führen auch zum Entstehen von neuartigen technischen Risiken. Bei der Verhütung und Bekämpfung vieler dieser Risiken steht die Forschung jedoch oft noch am Anfang.
Das Risiko einer unerkannten *Verzerrung des Systemergebnisses* (“Biased AI”), welches als Problem der Computertechnik als sich selbst bestätigendes Denksystem bereits bekannt war, verschärft sich mit dem zunehmenden Einsatz von AES. Gerade umfangreiche Modelle sind als Grundlage von AES durch ihre Komplexität anfällig für Fehler in der Auswahl und der Priorisierung bestimmter Parameter und Daten. Die subjektiven Vorurteile und Einstellungen von AnwenderInnen, EntwicklerInnen oder KonstrukteurInnen können sich in der erzeugten Computertechnik fortsetzen oder sogar verstärken. Dadurch kann es – gerade beim Einsatz der Technik im Rahmen hoheitlicher Tätigkeiten – zu ungewollten Diskriminierungen kommen, welche der Vorstellung einer “objektiven Maschine” widersprechen.
Von dem Risiko der *Intransparenz maschineller Prozesse* (“Blackbox AI”) spricht man, wenn die Ergebnisausgabe eines AES sich nicht durch einen Menschen interpretieren oder verstehen lässt. Wenn sich maschinelle Abläufe nicht mehr nachvollziehen lassen, ist die Aufklärung von Schäden im Nachhinein nicht möglich. Diese Intransparenz kann einerseits technische Ursachen haben, wenn etwa die Anzahl der Parameter in einem verwendeten Modell so groß ist, dass eine Vorhersagbarkeit von Ergebnissen nicht mehr gegeben ist. Auch können im Falle von Künstlichen Neuronalen Netzen Parameter in unüberschaubarer, nicht-linearer Weise zusammenwirken, was das Verständnis enorm erschweren kann. Diese Intransparenz kann jedoch auch wirtschaftlich-organisatorische Ursachen haben. Für Unternehmen, die AES entwickeln, kann es von Vorteil sein, sich auf das Argument der technisch geschuldeten Intransparenz zurückzuziehen oder absichtlich solche nicht- interpretierbaren Modelle zu verwen-

den, um den wirtschaftlichen Wert eines AES zu bewahren. Denn ein interpretierbares Modell kann prinzipiell auch rekonstruiert werden.

Die *Veränderungsfähigkeit* von AES kann starke Vorteile für viele Anwendungsbereiche mit sich bringen, aber auch Risiken entstehen lassen. Durch einen sich permanent wiederholenden Prozess kann erreicht werden, dass eine Abbildung, also ein Modell eines Prozesses einer Datenlage oder auch kognitive Phänomene auf Maschinen, sich Schritt für Schritt verbessert. Jedoch kann dadurch prinzipiell nicht mehr gewährleistet werden, dass die Funktion und Korrektheit, die zu Beginn des Einsatzes attestiert wurde, während der weiteren Verwendung gegeben bleibt.

5. Ausgleich zwischen Sicherheit und Innovation finden

Die Nutzung von Hochtechnologie birgt regelmäßig das Entstehen von Risiken. Der Gesetzgeber hat hier die Wahrung seiner grundgesetzlich verankerten Schutzpflicht mit dem Bedürfnis nach einer Steigerung des gesellschaftlichen Wohlstandes in Ausgleich zu bringen. Neue Regelungen dürfen die bestehenden und noch zu erwartenden Innovationen in diesem Bereich möglichst nicht verhindern oder ausbremsen.

Bei der Regulierung von AES tritt das Problem hinzu, dass diese Systeme in äußerst unterschiedlichen Anwendungsbereichen eingesetzt werden. Eine horizontale Regelung, die – ähnlich der DS-GVO – für alle AES gleichermaßen gelten soll, muss daher zwei Dimensionen bedenken:

- *Systembezogene Risiken:* Dies sind die bereits bei 3. dargestellten, technikspezifischen Risiken, die bei allen Anwendungsformen von AES auftreten können - Risiken der Verzerrung, der Transparenz und der Veränderungsfähigkeit.
- *Anwendungsbezogene Risiken*: Das sind zum anderen die anwendungsspezifischen Risiken, die sich aus dem jeweiligen technischen System ergeben. Ein autonomes Auto gefährdet andere Rechtsgüter stärker als bspw. eine Dating-App. Beides kann jedoch ein AES darstellen, wenn die technischen Gegebenheiten vorliegen.

Wenn man die system- mit den anwendungsbezogenen Risiken ins Verhältnis setzt, kann das spezifische Risiko *Algorithmischer Entscheidungssysteme* bestimmt werden. Diese Einstufung kann dann zur Ermittlung entsprechender Regulierungsmaßnahmendienen. Auch die Unterscheidung zwischen einem *teil*-und einem *vollautomatisierten AES* sollte Einfluss auf das Ausmaß der gesetzlichen Verpflichtung haben.

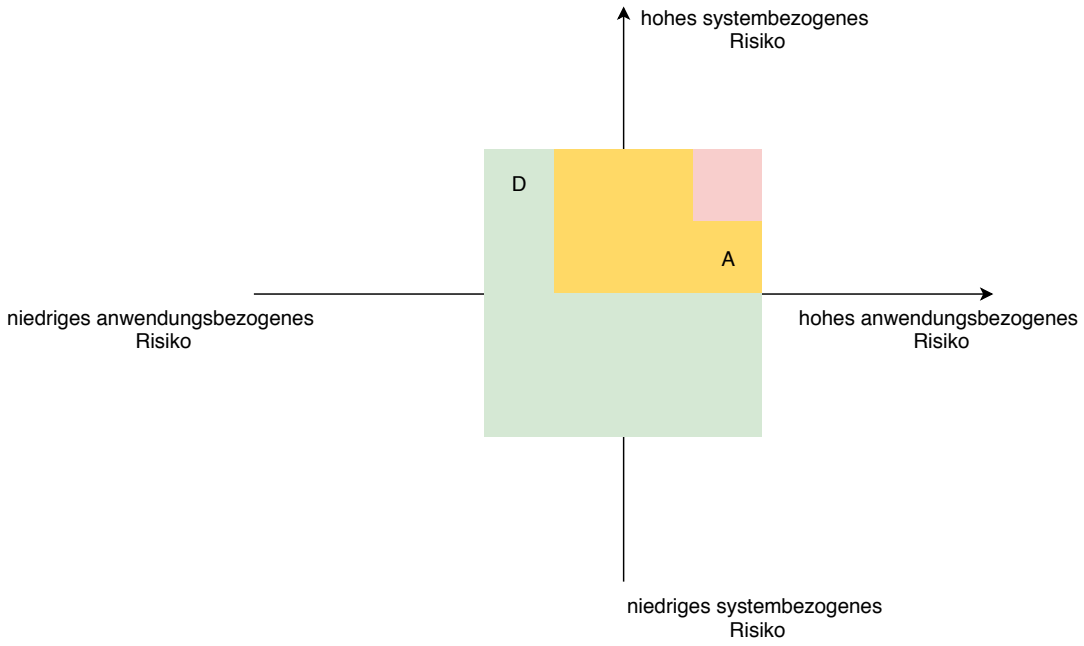


Abb. 1. Regulierungsmaßnahmen algorithmischer Entscheidungssysteme

Bei einem autonomen Auto (A) sind stärkere Regulierungsmaßnahmen zu ergreifen (Gelb). Bei einer Dating-App (D) sind weniger einschneidende Maßnahmen erforderlich (Grün). Das Bestehen eines hohen AES-Risikos (Rot) spricht für das Verbot oder starke Regulierung einer Anwendung.

Als Regulierungsmaßnahmen in Betracht kommen u.a.:

- Einführung einer Betreiberhaftung verbunden mit einer Pflichtversicherung,
- Einführung eines sektorspezifischen Registers,
- Überwachung durch eine Aufsichtsbehörde,
- Auflagen zur Evaluation der systembezogenen Risiken vor Einführung,
- Live-Schnittstellen, die bei Erreichen kritischer Limits einen Alarm senden,
- Technische Sicherheitsmaßnahmen, die beim Versagen des Systems einen Schutz vor physischen Schäden gewährleisten und die
- Einführung von speziellen Zertifizierungen und Normierungen.