Der Bundesbeauftragte
für den Datenschutz und
die Informationsfreiheit

Vortrag

des Bundesbeauftragten für den Datenschutz und die Informationsfreiheit

Prof. Ulrich Kelber

„Algorithmenregulierung nach Kritikalitätsstufen – der Ansatz der Datenethikkommission“

Tagung Regulierung für Algorithmen an der Universität Bonn

Bonn, den 7. September 2020

Es gilt das gesprochene Wort

Sehr geehrter Herr Professor Zimmer,

Sehr geehrte Damen und Herren,

I. Einleitung

Ich freue mich über die Gelegenheit, heute die Vorschläge der Datenethikkommission zur Regulierung von Algorithmen aus der Perspektive des Datenschutzes vorstellen zu dürfen.

Geplant war die heute und morgen stattfindende Tagung ja bereits für Mai dieses Jahres, doch leider hat uns das Corona-Virus einen Strich durch die Rechnung gemacht. Ich freue mich jedoch umso mehr, dass ich heute, wenn auch nur virtuell, hier sein kann, um Ihnen meine Perspektive näher zu bringen.

Die Thematik der Regulierung von Algorithmen begleitet mich seit Beginn meiner Amtszeit im Januar letzten Jahres. Mit Amtsantritt hatte ich auch die Mitgliedschaft in der Datenethikkommission (DEK) von meiner Vorgängerin, Frau Voßhoff, übernommen. Das Thema Algorithmen und ihre Regulierung war eines der zentralen Themen der DEK. Und auch im Abschlussgutachten der DEK, das im Oktober letzten Jahres veröffentlicht wurde, ist dem Thema Algorithmen ein eigenes Kapitel gewidmet.[1]

[1] Kapitel F des Abschlussgutachtens der DEK

Nach meiner Überzeugung sollten Algorithmen beziehungsweise algorithmische Systeme - dies ist der Begriff, den die DEK gewählt hat - auf europäischer Ebene reguliert werden.

Der Datenschutz als Leitmotiv dieser Regulierung soll technische Innovationen möglich machen, gleichzeitig aber auch Fehlentwicklungen vorbeugen. Wichtig ist es daher, einen risikobasierten Regulierungsansatz zu wählen. Das bedeutet, je größer die Risiken umso intensivere Regulierung.

Im Umkehrschluss bedeutet dies aber auch, dass bei all jenen algorithmischen Systemen, die ohne oder nur mit sehr geringen Risiken für den Einzelnen oder die Allgemeinheit verbunden sind, keine zusätzliche Regulierung erforderlich ist. Dies wird die weit überwiegende Zahl der AS umfassen. Die Innovation kann dort freien Lauf entfalten. Wenn jedoch ein unvertretbares Schädigungspotenzial festgestellt wird, müssen im Einzelfall auch bestimmte algorithmische Systeme verboten werden.

Im breiten Zwischenfeld kommen je nach Schädigungspotenzial verschiedene Ansätze, wie Codes of Conduct, Ex-post-Kontrollen oder aber auch Live-Schnittstellen mit Zugriffsmöglichkeiten für die zuständigen Aufsichtsbehörden in Betracht.

Diesen risikobasierten Ansatz habe ich auch als Mitglied der DEK vertreten und auf ihm gründen ebenfalls die entsprechenden Empfehlungen der DEK im Abschlussgutachten. Im Folgenden möchte ich daher näher auf die Empfehlungen der DEK zu algorithmischen Systemen eingehen.

II. Der Ansatz der Datenethikkommission zur Algorithmenregulierung

a. Der Begriff der algorithmischen Systeme

Die Mitglieder der DEK waren sich einig, in ihren Empfehlungen nicht zwischen Anwendungen sogenannter Künstlicher Intelligenz und „normalen“ Algorithmen zu unterscheiden. In der öffentlichen Debatte liegt der Fokus meist auf dem Einsatz Künstlicher Intelligenz und dem maschinellen Lernen. Doch die ethischen Fragestellungen für den Einsatz von Algorithmen stellen sich bei jeder Art, von der einfachen Anwendung im Getränkeautomat bis zur selbstständig tötenden Drohne. Daher macht die DEK in ihren Handlungsempfehlungen generell keine Unterscheidungen zwischen der Art des Algorithmus, sondern spricht allgemein von algorithmischen Systemen.

Die DEK hat den Begriff der algorithmischen *Systeme* gewählt, da es bei der Bewertung des Risikos nicht allein auf die *technischen* Eigenschaften des Algorithmus ankommt, sondern vielmehr das gesamte *sozio-technische* System berücksichtigt werden muss. Dies bedeutet, dass alle Komponenten der algorithmischen Anwendung einschließlich menschlicher Akteure und Einsatzszenario von der Entwicklungsphase über die Implementierung bis hin zur anschließenden Bewertung und ggf. Korrektur zu berücksichtigen sind.

b. Grundsätze zum Umgang mit Algorithmen

Bevor ich auf das Herzstück der DEK-Empfehlungen eingehe – die von der DEK entwickelten Kritikalitätsstufen, möchte ich zuvor noch mehrere allgemeine Grundsätze zum Einsatz von Algorithmen ansprechen, die bei der Entwicklung und dem Einsatz algorithmischer Systeme stets berücksichtigt werden sollten.

Die allgemeinen Anforderungen an algorithmische Systeme, welche die DEK aufstellt, decken sich überwiegend mit den Forderungen der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder (DSK) aus der Hambacher Erklärung. Die Hambacher Erklärung wurde von der DSK parallel zu den Arbeiten der DEK im April letzten Jahres verabschiedet.

Ich möchte drei Grundsätze herausgreifen, die sich sowohl im Gutachten der DEK als auch in der Hambacher Erklärung enthalten sind.

- Menschenzentriertes Design

Der wichtigste Aspekt, der beim Umgang mit Algorithmen immer mitgedacht werden muss, ist der *Mensch*. Algorithmen werden von Menschen geschaffen und sollen den Menschen dienen. Dies bedeutet, dass die Rechte und Freiheiten, aber auch das körperliche und emotionale Wohlbefinden aller Menschen, die von der Entscheidung des Algorithmus betroffen sind, auch mit einbezogen und geachtet werden müssen. Es ist an der Zeit, dass der Mensch bei der Entwicklung von Algorithmen in den Mittelpunkt gestellt wird.

- Vereinbarkeit mit gesellschaftlichen Grundwerten / Verfassung

Neben der Berücksichtigung von Rechten und Bedürfnissen der betroffenen Bürger, sind auch die gesellschaftlichen Grundwerte, die auch im Grundgesetz zum Ausdruck kommen, bei der Entwicklung und dem Einsatz von Algorithmen zu beachten. Algorithmische Systeme dürfen nicht im Widerspruch zu diesen Werten stehen oder durch diese unterminiert werden. Dieser Aspekt spielt beispielsweise beim Einsatz sogenannter Social Bots[2] eine Rolle. Sind diese nicht gekennzeichnet, so kann es zu einer versteckten Beeinflussung der Meinungsbildung zum Beispiel in sozialen Netzwerken kommen. Ein weit verbreiteter Einsatz von Social Bots könnte letztlich sogar die Demokratie gefährden. Aus diesem Grund spricht sich die DEK zu Recht für eine Kennzeichnung von Social Bots aus, auch bei hybriden Systemen aus KI und Mensch.

[2] Bots, die den Eindruck erwecken, menschliche Nutzer zu sein.

- Transparenz, Erklärbarkeit und Nachvollziehbarkeit

Der letzte Grundsatz, den ich ansprechen möchte betrifft die Transparenz, Erklärbarkeit und Nachvollziehbarkeit algorithmischer Systeme. Für die Legitimität und Akzeptanz algorithmischer Systeme sind diese Aspekte essentiell. Dies gilt nicht nur im Hinblick auf die Anwender algorithmischer Systeme, die diese nachvollziehen, erklären und kontrollieren können müssen. Auch im Hinblick auf alle von einem algorithmischen System Betroffenen müssen Transparenz, Erklärbarkeit und Nachvollziehbarkeit gewährleistet sein. Nur wenn die Betroffenen angemessen informiert werden, wissen sie überhaupt, wann sie Entscheidungen durch algorithmische Systeme ausgesetzt sind, können diese ggf. infrage stellen und ihre Rechte angemessen wahrnehmen. Ein Bereich, in dem mehr Transparenz herrschen sollte, ist z.B. individualisierte Online-Werbung. Es sollte für den Internetnutzer sofort klar erkennbar sein, wenn er mit individualisierter Werbung konfrontiert wird. Zudem sollte mit wenigen Klicks leicht auffindbar erfahren können, auf welchen Daten die Werbung beruht und aus welchen Quellen diese stammen.

Dies waren nur drei der allgemeinen Grundsätze, die beim Einsatz algorithmischer Systeme berücksichtigt werden sollten. Weitere Aspekte betreffen beispielsweise die Minimierung von Verzerrungen und Diskriminierung oder klare Verantwortlichkeiten.

c. Risikobasierter Ansatz – Systemkritikalität

Und nun zum Herzstück der DEK-Empfehlungen: die Einordnung algorithmischer Systeme in Kritikalitätsstufen.

Wie bereits erwähnt sind bei der Bewertung eines algorithmischen Systems alle Komponenten zu berücksichtigen: Die eingesetzte Technik, das Anwendungsgebiet und die beteiligten menschlichen Akteure. All diese Aspekte sind zu prüfen, um die Kritikalität und damit das Schädigungspotenzial zu erfassen, das von einem algorithmischen System ausgeht.

Die DEK empfiehlt hierfür, ein übergreifendes Modell zu entwickeln, nach dem algorithmische Systeme Kritikalitätsstufen zugeordnet werden. Die DEK hat dazu fünf Kritikalitätsstufen entwickelt. Diese Stufen können Sie sich als spitz zulaufende Pyramide vorstellen, wobei die erste Stufe die breiteste unten ist und die Stufen nach oben hin immer kleiner werden. Dies bedeutet, dass die DEK davon ausgeht, dass die meisten algorithmischen Systeme der ersten Stufe zuzuordnen sind und nur wenige der fünften Stufe.

Auf der *ersten Stufe* stehen Anwendungen ohne oder mit geringem Schädigungspotenzial. Hier bedarf es weder spezieller Qualitätsanforderungen noch besonderer Kontrollmechanismen. Beispiele hierfür sind etwa ein herkömmlicher Getränkeautomat oder aber auch hochkomplexe Anwendungen wie Alpha-Go[3]. Gedankenexperiment Getränkeautomat.

Auf *Stufe 2* stehen Anwendungen mit einem gewissen Schädigungspotenzial. Hier sollten erste Regulierungen greifen, wie etwa Ex-post-Kontrollen bei dem begründeten Verdacht eines Fehlverhaltens des Systems, die Veröffentlichung angemessener Risikofolgeabschätzungen oder branchenspezifische, von den Aufsichtsbehörden genehmigte Codes of Conduct. Dieser Stufe sind beispielsweise dynamische Preissetzungen ohne Personalisierung zuzuordnen.

Bei regelmäßigem oder deutlichem Schädigungspotenzial ist ein algorithmisches System der *dritten Stufe* zuzuordnen. Für Anwendungen dieser Stufe sollten spezielle Zulassungsverfahren eingeführt werden, die z.B. bestimmte Qualitätsanforderungen in der Entwicklung voraussetzen. Auf dieser Stufe sind beispielsweise personalisierte Preise anzusiedeln.

[3] AlphaGo ist ein Computerprogramm, das das Brettspiel Go spielt. AlphaGo kombiniert Techniken des maschinellen Lernens und der Traversierung. Quelle: https://de.wikipedia.org/wiki/AlphaGo

Die vorletzte, *vierte Stufe* würde bei Anwendungen mit erheblichem Schädigungspotential greifen. Hier wären neben der formalen Zulassung weitere Kontrollmöglichkeiten zu schaffen. Denkbar ist etwa die Möglichkeit der kontinuierlichen behördlichen Kontrolle über eine Live-Schnittstelle. Dieser Stufe greift zum Beispiel für die Beurteilung der Kreditwürdigkeit durch Akteure mit massiver Marktmacht.

Die *fünfte* und damit höchste Stufe betrifft Anwendungen mit einem unvertretbaren Schädigungspotenzial. Diese müssten verboten werden. Beispiel hierfür sind algorithmen-determinierte Tötungen durch den Einsatz von autonomen Waffensystemen. Diese sind von der bloßen Unterstützung bei der Objekterkennung zu unterscheiden.

An diesem extremen Beispiel wird bereits klar, dass die letzte Stufe nur in äußert wenigen Fällen zum Einsatz kommen würde. Die meisten algorithmischen Systeme wären auch bei Anwendung dieser Kritikalitätsstufen weiterhin zulässig, würden jedoch abhängig vom Schädigungspotenzial einer weitergehenden Regulierung und Kontrolle unterworfen.

d. Umsetzung

Um das skizzierte Regulierungsmodell umzusetzen müsste eine horizontale Algorithmen-Verordnung auf EU-Ebene geschaffen werden. Diese Verordnung müsste unter anderem die zentralen Grundprinzipien für algorithmische Systeme enthalten. Geregelt würden in einer solchen Verordnung unter anderem die Zulässigkeit und Gestaltung algorithmischer Systeme, Transparenzaspekte und Betroffenenrechten.

Die aktuelle deutsche Ratspräsidentschaft könnte die Bundesregierung zum Anlass nehmen, eine europäische Algorithmen-Verordnung zu initiieren.

Ich danke Ihnen für Ihre Aufmerksamkeit.