

Die internationale Dimension des Einsatzes der Künstlichen Intelligenz

Dace L. Luters-Thümmel

Mitglied des djB-Arbeitsstabes "Digitales", Secretary General European Women Lawyers Association (EWLA)

Der Einsatz von Algorithmen, Künstlicher Intelligenz (KI) und digitalen Innovationen macht vor Landesgrenzen nicht Halt. Mit dem Siegeszug des Internets in sämtliche Bereiche des öffentlichen und privaten Lebens hat auch die Auswertung unserer Daten und Datenspuren im Netz ihren Einzug gehalten. Große Dienstleister – allen voran: Google, Amazon, Facebook und Apple – bieten ihren Service weltumspannend an und haben ihren weltweiten täglichen Nutzeranteil auf über eine Milliarde katapultiert.

Diese globale Dimension der Digitalisierung verlangt nach Antworten auf ethische und rechtliche Fragen, die sich von einem Landesgesetzgeber allein weder lösen noch durchsetzen lassen. Ein wichtiger Aspekt ist dabei auch die Frage, inwieweit durch maschinelles Lernen bestehende Diskriminierungen rezipiert und gar potenziert werden können.

Während der Deutsche Bundestag im Spätsommer 2018 eine Enquete-Kommission zur Thematik „Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale“ eingesetzt hat und die Bundesregierung im gleichen Jahr eine Datenethikkommission einberief, deren Handlungsempfehlungen wir im Herbst 2019 erwarten, haben sowohl der Council of Europe, die OECD als auch die Europäische Kommission in den vergangenen zwei Jahren ebenfalls Arbeitsgruppen ins Leben gerufen, die Studien, Grundsätze und Richtlinien für den vertrauenswürdigen Einsatz Künstlicher Intelligenz ausgearbeitet haben.

Wegen der Vielzahl an Dokumenten können im Folgenden nur die wesentlichen Schlaglichter aufgezeigt werden, um den Gesamtkontext der Bemühungen zu verdeutlichen ohne dabei auch nur im Ansatz einen Anspruch auf Vollständigkeit¹ zu erheben:

I. Council of Europe

Im März 2018 erschien die im Auftrag des Council of Europe erstellte „Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular Algorithms) and possible Regulatory Implications“². Die Studie warnt explizit davor, dass das maschinelle Lernen die Gefahr berge Vorurteile zu verstärken, wenn nicht bewusst dagegen gesteuert wird³. Dies vor allem im Kontext eines vorurteilsbehaftet programmierten Algorithmus, der dann im Anschluss (*programmierungsgemäß*)⁴ systematisch bestimmte Gesellschaftsgruppen (und damit unter anderem auch das Geschlecht) diskriminiere und diesen Diskriminierungseffekt repliziere und weiter multipliziere⁵.

Dem folgte im November 2018 „A Study of the Implications of Advanced Digital Technologies (including AI systems) for the Concept of Responsibility within a Human Rights Framework“⁶. Die Studie schlussfolgert, dass die Digitalwirtschaft eine moralische Verantwortung zur Einhaltung der Menschenrechte habe, die durch freiwillige Erklärungen allein nicht erfüllt werde. Vielmehr werden die Mitgliedstaaten dazu aufgefordert, Schutzmechanismen einzuführen, um den Menschenrechtsschutz zu gewährleisten.⁷

II. OECD

Die OECD verabschiedete ihrerseits im Mai 2019 ihre „Recommendation of the Council on Artificial Intelligence“⁸. Sie appelliert daran, dass die Stakeholder sich „proaktiv für eine verantwortungsvolle Steuerung vertrauenswürdiger KI einsetzen, die darauf abzielt, [...] eine Verringerung der wirtschaftlichen, sozialen, geschlechtsspezifischen und sonstigen Ungleichheiten [...] zu fördern“⁹ und sieht die KI-Akteure in der Pflicht, während des gesamten Lebenszyklus eines KI-Systems „das Prinzip der Rechstaatlichkeit, die Menschenrechte und demokratische Werte zu achten“¹⁰. Dabei wird auch das Prinzip der Nichtdiskriminierung und der Gleichbehandlung ausdrücklich miteinbezogen. Am 9. Juni 2019 hat der in Japan tagende G20-Rat anschließend eine entsprechende Erklärung¹¹ verabschiedet, die diese OECD Grundsätze eines „Menschen bezogenen Ansatzes der KI“ zur Richtschnur der G20-Staaten für die Anwendung Künstlicher Intelligenz erklärt, jedoch als nicht bindend deklariert.

III. EU

Die genannten Initiativen zur Erarbeitung von Grundsätzen einer vertrauenswürdigen KI blieben bislang auf der Ebene nicht bindender Empfehlungen und Feststellungen im Kreise von Expert*innen stehen. Der Europäischen Kommission hingegen steht grundsätzlich ein weiter gehendes Instrumentarium zur Verfügung.

- 1 Ergänzend sei hier nur das gemeinsame ALI-ELI Projekt "Principles for a Data Economy" erwähnt. ALI: American Law Institute; ELI: European Law Institute.
- 2 DGI (2017) 12, Council of Europe.
- 3 a.a.O. S.27.
- 4 Anmerkung der Verfasserin.
- 5 a.a.O. S. 28.
- 6 MSI-AUT (2018) 05, Council of Europe.
- 7 a.a.O. S.74.
- 8 C/MIN (2019)3/FINAL, online: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (Zugriff: 26.6.2019).
- 9 s. IV. 1.1.
- 10 s. IV. 1.2.
- 11 Online: https://g20trade-digital.go.jp/dl/Ministerial_Statement_on_Trade_and_Digital_Economy.pdf (Zugriff: 26.6.2019).

Auf europäischer Ebene wurde im Sommer 2018 zunächst eine High Level Expert Group on Artificial Intelligence ins Leben gerufen, die auf der Grundlage eines Grundsatzpapiers „Statement on Artificial Intelligence, Robotics and ‚Autonomous‘ Systems“¹² der European Group on Ethics in Science and New Technologies vom März 2018 an eigenen „Ethics Guidelines for Trustworthy AI“ gearbeitet hat. Im Vorfeld der Besetzung dieser High Level Expert Group on Artificial Intelligence hatte sich auch die European Women Lawyers Association (EWLA) um eine Mitarbeit in diesem Gremium bemüht. Die High Level Expert Group wurde dann zwar paritätisch männlich – weiblich besetzt, jedoch bestand das Gremium vornehmlich aus Vertreter*innen der (digitalen) Wirtschaft, der Gewerkschaften, einiger weniger Verbraucherorganisationen und von Universitäten. Eine Interessenvertretung zur Wahrung der Grundrechtecharta und einer geschlechtergerechten KI-Anwendung war nicht zugegen. Nach mehreren (nicht-öffentlichen) Arbeitsgruppensitzungen wurde am 18. Dezember 2018 ein erster Entwurf der Ethics Guidelines¹³ veröffentlicht und zur Diskussion gestellt. Bis zum 1. Februar 2019 bestand Gelegenheit diesen Entwurf im Rahmen der European Artificial Intelligence Alliance zu kommentieren. EWLA, die zu vertreten ich im Rahmen der Konsultation die Ehre hatte, kritisierte den fehlenden Gender-Bezug des Papiers.

Das Kapitel „3. Fundamental Rights of Human Beings“ und dort „3.4 Equality, non-discrimination and solidarity“ erwähnte im Zusammenhang mit dem Grundsatz der Gleichbehandlung aller Menschen mit keiner Silbe einen Geschlechteraspekt. So enthielt der Entwurf nur die Aussage, dass Gleichbehandlung auch die Inklusion von traditionell ausgeschlossenen Minderheiten beinhalte, insbesondere von Arbeitnehmer*innen und Verbraucher*innen, womit sich zugleich auch die Zusammensetzung der Expert*innengruppe spiegelte.

In der nun am 8. April 2019 veröffentlichten Endfassung der „Ethics Guidelines for a Trustworthy AI“¹⁴ wurde diese Formulierung nachgebessert. In der neu nummerierten Fassung ist nunmehr unter „2. From fundamental rights to ethical principles“ unter „2.1. Fundamental rights as a basis for Trustworthy AI“¹⁵ zu lesen:

“In an AI context, equality entails that the system’s operations cannot generate unfairly biased outputs (e.g. the data used to train AI systems should be as inclusive as possible, representing different population groups). This also requires adequate respect for potentially vulnerable persons and groups, such as workers, women, persons with disabilities, ethnic minorities, children, consumers or others at risk of exclusion.”

– Übersetzung¹⁶: “Im Kontext der KI, beinhaltet Gleichbehandlung, dass die Systemvorgänge keine unfairen vorurteilsbehafteten Ergebnisse hervorbringen dürfen (z.B. Daten, die zum Anlernen von KI-Systemen verwendet werden, sollten so inklusiv wie möglich sein und verschiedene Bevölkerungsgruppen widerspiegeln). Das erfordert eine entsprechende Achtung von potentiell besonders verletzbaren Personen und Gruppen, wie beispielsweise Arbeitnehmer*innen, Frauen, Behinderten, ethnischen Minderheiten, Kindern, Verbraucher*innen oder anderen Personen, die der Gefahr eines Ausschlusses unterliegen.“

Dies dürfte insoweit allerdings nur einen Teilerfolg darstellen, da die Guidelines bislang noch nicht in einen verpflichtenden Rechtsakt eingeflossen sind. Dennoch ist es ein Beleg dafür, dass wir uns auch auf europäischer Ebene rechtspolitisch engagieren müssen, um rechtzeitig weitere Weichen für die europäische Rechtsetzung zu stellen. Ich möchte mich ausdrücklich für die Unterstützung, die ich vom Deutschen Juristinnenbund e.V. (djb) und von EWLA hierbei erfahren habe, bedanken.

Zusammenfassend formulieren die „Ethics Guidelines for a Trustworthy AI“ nunmehr folgende Grundsätze¹⁷ einer vertrauenswürdigen KI:

„Vertrauenswürdige KI hat **drei Komponenten**, die während des gesamten Lebenszyklus eines Systems zu beachten sind:

1. Sie sollte **rechtmäßig** sein, und alle anwendbaren Gesetze und Vorschriften beachten,
2. sie sollte **ethisch** sein und eine Orientierung an ethischen Grundsätzen und Werten sicherstellen und
3. sie sollte **robust** sein, sowohl aus technischer als auch aus gesellschaftlicher Sicht, da, obwohl mit guter Absicht, KI-Systeme auch unbeabsichtigten Schaden anrichten können.“¹⁸

Im weiteren Fortgang ist vorgesehen, dass Kapitel 3 der Guidelines – die sogenannte Trustworthy AI Assessment List¹⁹, die von der Expert*innengruppe zusammengestellt wurde – von der Praxis im zweiten Halbjahr 2019 getestet und beurteilt werden soll. Die High Level Expert Group wird diese Assessment List dann bis Anfang 2020 anhand der Rückmeldungen aus der Praxis überarbeiten. Auch hier wird Wachsamkeit gefordert sein, um die Weiterentwicklung der Beurteilungsgrundsätze für KI aus Sicht des djb und von EWLA zu kommentieren.

Nachdem die „Ethics Guidelines for a Trustworthy AI“ den Praxistest durchlaufen haben, wird die Europäische Kommission Vorschläge zum weiteren Vorgehen auf EU-Ebene unterbreiten. Wir bleiben dran!

12 Online: https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf (Zugriff: 26.6.2019).

13 Online: <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai> (Zugriff: 26.6.2019).

14 Online: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (Zugriff: 26.6.2019).

15 a.a.O. S.13.

16 der Verfasserin.

17 a.a.O. S.7.

18 Übersetzung der Verfasserin.

19 ein Fragenkatalog, anhand dessen die Vertrauenswürdigkeit von KI aus verschiedenen Perspektiven überprüft werden soll.