Deutscher Bundestag
Ausschuss für Digitales

Ausschussdrucksache
20(23)79

23.09.2022

„EU-Verordnung zu Künstlicher Intelligenz unter Einbeziehung von Wettbewerbsfähigkeit im Bereich Künstliche Intelligenz und Blockchain-Technologie“

Antworten von Jürgen Geuter <tante@tante.cc>
22.09.2022

1) Bei der fallweisen und sektoralen Erfassung von Risiken und Bedrohungen und dem Wunsch von z.B. Unternehmensverbänden bezüglich Sandboxen und einer Kombination von ex ante Risiko-Selbstbewertung und ex post Durchsetzung bei KI mit als hoch eingestuftem Risiko stellt sich insbesondere mit Blick auf die Empfehlung der Kommission zum Einsatz von KI in öffentlichen Diensten als sogenannte „Test Umgebung“ die Frage, inwiefern und wodurch sichergestellt werden kann, dass trotz Kombination mehrerer Anwendungen keine Schäden gerade bei Einsatz in kritischen Infrastrukturen entstehen, der Grad möglicher Gefährdungslagen auch unter Einbeziehung von sektorübergreifenden Schnittstellen erhoben und allgemeine Haftungsfragen geklärt, mögliche Sicherheitslücken gefunden, gemeldet und schnell behoben werden können?

Die “KI”-Domäne, genau wie schon die Software-Domäne im allgemeinen, strebt einer zunehmenden Spezialisierung und Kommodifikation entgegen, d.h. viele spezialisierte Produkte, die jeweils genau eine spezifische Frage per “KI” lösen, und die man dann zu Gesamtlösungen zusammensteckt. “KI” wird immer mehr COTS (“Commercial off the Shelf”). Diese Entwicklung bringt mit sich grundsätzlich die Herausforderung, dass die Laufzeitumgebung und der Einsatzkontext eines “KI”-Systems zum Entwicklungszeitpunkt kaum noch vorhersehbar (und damit bewertbar) ist. Eine Gesichtserkennung kann in einer Kamera helfen, bessere Fotos zu machen oder verwendet werden um grundrechtsinkompatible Massenüberwachung zu implementieren. Die “KI”-Komponente ist hier in beiden Fällen identisch, das Risiko entsteht aus dem Einsatzzweck (und den vorhandenen Machtgefälle zwischen denen, die die Software betreiben und denen, die von der Software ausgewertet werden).
Die Lieferketten auch im Softwarebereich werden immer komplexer und gerade bei “KI”-Systemen ist die Korrektur von “Bugs” grundsätzlich schon deutlich schwieriger als in klassischer Software. Auch auf dieser Basis ist es jetzt schon und wird in Zukunft noch schwerer werden, verlässliche Aussagen über die Eigenschaften einer Komponente, ihres Zusammenspiels mit irgendwelchen anderen Komponenten und die Wirkung des Gesamtsystems in einem Einsatzkontext zu machen.
Grundsätzlich wird die Haftung bei den Betreibern einer Lösung in einem realen Einsatzkontext liegen müssen; denn wie genau bestimmte Komponenten eingesetzt werden können, ist von den Anbietern nicht immer vorhersehbar. Sicherzustellen, dass keine Schäden passieren können, kann nur über aufwändige Tests stattfinden. Dabei können und dürfen aber reale, evtl.

sogar kritische öffentliche Dienste keinesfalls als “Test Umgebungen” verwendet werden: Test Umgebungen müssen getrennt von den sogenannten “Produktivsystemen” betrieben und in enger Überwachung und permanenter Evaluation betrieben werden. Es ist auch essentiell, dass Nutzer*innen und vor allem auch evtl. Andere Betroffene jederzeit merken, dass sie mit einem System “im Testbetrieb” interagieren bzw. von diesem verarbeitet werden mit einem direkten Weg, die Betreiber über Probleme zu informieren. Diese Information sollte sogar in eine Form von “Ticketsystem” fließen, so dass die Betroffenen auch über die Bearbeitung ihrer Anfrage informiert werden. Erst nach einer vorher definierten Dauer an fehler- bzw. Schadfreiem Betrieb kann genau die getestete Version in einen Produktivbetrieb überführt werden. Gerade im “KI”-Kontext wird gerne einfach so “in der realen Welt” getestet. Dieses Vorgehen ist unter der Bedingung, die Risiken und Schäden für Menschen möglichst klein zu halten, indiskutabel. Die Dokumentation des erfolgreichen und sicheren Betriebs mit repräsentativen Daten muss Voraussetzung für den Einsatz in öffentlichen Systemen sein.

*2) Im Bereich der Blockchain-Technologie fordern führende Kryptoexpert*innen in einem Brief an den US-Kongress die Regulierung von Kryptoassets sowie eine Überprüfung in Bezug auf den Mehrwert beim Einsatz von Blockchain-Technologien. Im Bereich der KI-Anwendungen wird von umfassenden Herausforderungen bei der Regulierung von Algorithmen zur Gestaltung sozio-ökonomischer und ökologischer Implikationen ausgegangen. Kann bei beiden Technologien gewährleistet werden, dass eine angemessene Anwendung nach Kosten-Nutzen-Abwägung erfolgt, z.B. im Bereich des ökologischen Ressourcenverbrauchs, oder dem Generieren eines echten Mehrwerts gegenüber klassischen IT-Anwendungen und wenn ja, anhand welcher messbarer Kriterien und werden gesellschaftliche/soziale Folgewirkungen ausreichend berücksichtigt?*

Sowohl “KI” als auch “Blockchain” sind nicht nur technologische Begriffe sondern Konzepte, die - insbesondere auch in der Politik - vor allem auch Hoffnungen und Fantasien beschreiben. Das macht es natürlich besonders schwierig sicherzustellen, dass sie seriös und sinnhaft strukturiert bewertet werden: Ein Problem mit den “Zukunfts”-Schlagwörtern “Blockchain” oder “KI” zu lösen bringt ungleich mehr Sichtbarkeit und Zuspruch als eine traditionellere Lösung.
Hier ist es - insbesondere sogar wenn man diese technologischen Imaginaries irgendwie fördern möchte - wichtig bei Ausschreibungen oder Anfragen eine immer anhand realer Anforderungen zu operieren: “Entwickeln Sie ein Konzept um X umzusetzen” anstelle von “Wie kann X mit einer Blockchain/KI umgesetzt werden”. Gerade in der Blockchain Domäne sehen wir den letzteren Fall sehr häufig und erzeugen so irgendwelche aufwändigen, fragilen Blockchain-Prototypen anstelle günstigerer, wartbarerer und robusterer Lösungen.
Auch die Kosten-Nutzen Frage ist leider oft nicht ganz so sinnhaft bewertet: Ein Vergleich einer halbfertigen Blockchain/KI-Lösung auf der grünen Wiese mit einer in diverse Bestandssysteme integrierten Umsetzung wird immer zu Lasten der Aufwertung der Bestandssysteme gehen, auch wenn genau diese Integrationsleistung ebenso für prototypische Insellösungen nachgezogen werden müsste.
Bei der Bewertung von “Potentialen” und “sozialen Auswirkungen” ist es extrem schwierig, belastbare Kriterien anzubringen, da die Auswirkungen oft erst nach Umsetzung der Lösung

erfahrbar werden können und sowohl “KI” als auch “Blockchain” als Narrative mit nahezu grenzenlosen Zukunftsversprechen verwoben sind.
Insbesondere für die Umsetzung von “KI”-Systemen muss es zum Standard-Procedere gehören, die ökologischen Impacts von Training und Einsatz im Vorfeld abzugrenzen und diese in die Bewertung einfließen zu lassen.
Gesellschaftliche Folgekosten sind schwer zu fassen. Was aber zu fassen wäre, sind Kriterien wie Transparenz und Fehlerbehandlung (wie konkret können im Falle von Problemen die Ursachen nachvollzogen, die Probleme behoben und eine Wiederholung dieser ausgeschlossen werden?), Wartbarkeit (wie einfach ist es, einen erkannten Fehler zu beheben/muss das ganze Netzwerk neu trainiert werden?), etc.
Grundsätzlich empfiehlt es sich - insbesondere als öffentliche Hand - nicht technologiespezifisch auszuschreiben sondern nachhaltige, wartbare Lösungen technologieoffen auszuschreiben und die Externalitäten (Umweltimpact, sozialer Impact, etc) in die Bewertung einfließen zu lassen.

3) Inwieweit wird sich die KI-Verordnung auf die Wettbewerbsfähigkeit Europas im internationalen Vergleich auswirken?

Solche Vorhersagen sind immer mit einer gewissen Vorsicht zu genießen. Natürlich sind durch die KI-VO Unternehmen, die in den als Hochrisiko eingeschätzten Bereichen unterwegs sind, tendenziell etwas stärker eingeschränkt im Vergleich zum globalen Wettbewerb, der weniger Pflichten unterliegt. Andere Anbieter von “KI” Lösungen wären meiner Einschätzung nach nicht signifikant belastet, die Auswirkung wäre daher eher gering. Hoffnungen, dass solche Regulierung als globaler Wettbewerbsvorteil gesehen wird, halte ich für nicht besonders glaubwürdig, das hat sich auch schon bei der DSGVO nicht bewahrheitet.

4) Kann die KI-Verordnung in der Entwurfsfassung Diskriminierung zum Beispiel gegenüber Frauen oder PoC verhindern? Wo muss gegebenenfalls nachgesteuert werden?

Nein kann sie nicht. Natürlich sollen in Trainingsdaten befindliche Biases etc. identifiziert und mitigiert werden, aber diese Fällen decken nur die oberflächlichkeiten Diskriminierungsvektoren ab: Kombination von Werkzeugen oder unerkannte “Proxy-Variablen”, d.h. scheinbar unproblematische Daten, die aber zuverlässig eine besonders diskriminierungsrelevante Variable ersetzen können, haben weiterhin hohes Diskriminationspotenzial. Diskriminierung ist des weiteren kein rein-technisches Problem sondern haben immer mit politischem und sozialen Machtgefälle zu tun. Für die Nachsteuerung wäre es wichtig, allgemeine Antidiskriminierungsgesetze in ihrer Durchsetzung zu stärken und auch auf “KI”-Kontexte anzuwenden. Diskriminierung ist Diskriminierung, egal durch welchen Prozess oder Werkzeug sie erzeugt oder verstärkt wird.

5) Sind die in der DSGVO und im Verordnungs-Entwurf verankerten Regelungen zu Informations- und Beschwerderechten für Betroffene von KI-Entscheidungen ausreichend? Und wie könnten Betroffene für jene Rechte sensibilisiert werden?

Im VO-Entwurf sind die Informationspflichten beschränkt auf spezifische Arten von “KI”-Systemen. Diese sind alle als Hochrisikoanwendungen eingestuft. Für weniger hoch eingestufte Anwendungen sind die Informationspflichten daher sehr dünn. Aber erst im Einsatz und bezogen auf ihre eigene Lebensrealität können Betroffene ernsthaft einschätzen, ob sie mehr Informationen über die Systeme, mit denen sie interagieren, benötigen. Die Einschränkung ist daher etwas zu eng.
Zusätzlich muss man sich auch die Frage stellen, für wen die beschriebenen Informationen überhaupt verwendbar sind: Nicht nur aufgrund der hohen Abstraktion des Themas im Allgemeinen sondern auch aufgrund der für solche Informationen verwendeten juristischen Sprache sind solche Informationspakete oft nur für kleine Bevölkerungsteile überhaupt verständlich, geschweige denn in Aktionen überführbar. Hier findet eine Individualisierung von potenziell strukturellen Problemen statt, die im zweifel insbesondere Personen mit weniger (technischer) Expertise und Bildung im Regen stehen lässt.

6) Wie verlässlich ist eine Konformitätsbewertung von Hochrisiko-Anwendungen, die durch die Anbieter selbst durchgeführt wird? Brauchen wir gerade in sensiblen Bereichen eine externe Prüfung?

Die Verlässlichkeit ist auf jeden Fall - gerade in einem sehr kompetitiven globalen Markt - eingeschränkt. Hier sind auf jeden Fall Prüfungen nötig, gerade auch, wenn durch komplexe Zuliefererketten die Systemfunktionen zerlegt und potentiell auf der ganzen Welt verteilt sind.

7) Sehen Sie wesentliche begriffliche Unklarheiten in der KI-VO und, falls ja, welche regulatorischen Komplikationen ergeben sich möglicherweise daraus, und wie ließen sich derartige Komplikationen vermeiden oder beheben?

Es wurde zwar versucht, die Definition von “KI”-Systemen recht klar an Technologien zu binden, Annex I sorgt aber dafür, dass faktisch jede Software betroffen ist, denn jedes IT-System ist zu einem gewissen Grade “logic and knowledge-based”. Der Begriff “KI” selbst ist extrem weich und undefiniert. Hier wäre es evtl. Sinnvoller, eher über die Aufgaben der Systeme zu kategorisieren (z.B. Bilderkennung, Vorhersage, etc) zu argumentieren.

8) Laut einer aktuellen Umfrage von Bitkom betrachten 49 Prozent der befragten Unternehmen rechtliche Unsicherheiten als Hemmnis für die Einführung von KI-Anwendungen. Wird die KI-VO Ihrer Meinung nach zu einer Verbesserung der Situation führen, oder könnte sie sie ggfs. Sogar verschärfen – insbesondere für KMUs?

Ich sehe in der Hinsicht keinen grundlegenden Impact der VO auf die Situation. Welche konkreten “rechtlichen Unsicherheiten” bestehen denn für die Unternehmen? Für die spezifischen Fälle, wo genau eine der regulierten Fragen betroffen war, wird sich die Situation für die Unternehmen natürlich nun klären, aber wie hoch der Anteil dieser Fälle ist, kann man kaum bewerten. In dieser Allgemeinheit kann man diese Frage nicht beantworten.

9) Die Bundesregierung hat deutlich gemacht, dass sie KI-Anwendungen für Sicherheitsbehörden vom Hauptvertragstext getrennt regeln will. Denken Sie, dass dies sinnvoll ist und was sind die Vor- und Nachteile einer getrennten Regulierung?

Gerade in Anbetracht des hohen Missbrauchspotenzials von “KI” Anwendungen und der besonderen Bedeutung von Vertrauen in diesem Kontext sendet eine getrennte Regulierung von “KI”-Anwendungen für Sicherheitsbehörden das falsche Signal. Wenn die Bundesregierung Vertrauen in “KI” aufbauen will, dann sollte sie nicht den Eindruck erwecken, für sich andere - potentiell für Bürger*innen bedrohlich wirkende - Regeln einführen zu wollen.

10) In welchem Maße ermöglicht die KI-Verordnung Bürgerinnen und Bürgern, den Einsatz von KI-Systemen zu erkennen, zu verstehen und ihre Rechte wahrzunehmen, wenn sie von Entscheidungen oder Entscheidungsvorbereitungen durch KI betroffen sind und sind die Transparenzanforderungen aus Artikel 52 Satz 1 ausreichend, um darüber zu informieren, dass KI-Systeme automatisiert oder halb-automatisiert Entscheidungen treffen oder vorbereiten oder beeinflussen?

Der Rückgriff auf individuelle Informiertheit ist grundsätzlich ein politisches Problem: Um mit solchen Informationen sinnvoll umgehen zu können, sind nicht nur eine signifikante Bildung, eine hohe technische Kompetenz und viel Zeit nötig. Diese Vorbedingungen sind in der Realität für viele Menschen nicht gegeben. Das führt - analog zu den bekannten Cookie-Bannern - dazu, dass Menschen formal in alle möglichen Dinge einwilligen bzw. Ihre Informiertheit bestätigen, ohne diese wirklich mit Leben füllen zu können. Die Wahrnehmung der eigenen Rechte wird damit zum Privileg.

11) Mit Blick auf große Datensets für Gemeinwohl/Forschungsdaten: Gehen Sie davon aus, dass Forschungssandboxes so gestaltet werden können, dass keine strukturellen Einschränkungen von Datenschutz durch Nutzung von Forschungsdaten (Beispiel europäischer Gesundheitsdatenraum) erfolgen kann? Wenn ja wie und wenn nicht, könnten Sie bitte die Gründe ausführen?

Das hängt massiv von den betroffenen Datenkategorien ab. Die vergangenen Jahre haben gezeigt, dass die De-anonymisierung bzw. De-pseudonymisierung von personenbezogenen Daten in großen Datensätzen mit großer Wahrscheinlichkeit möglich ist, selbst wenn Best Practices eingehalten wurden. Insbesondere, wenn man unterschiedliche Datensätze mit Querverweisen verbindet und externe Datenbanken hinzuzieht. Sobald personenbezogene Daten in den Datensets stecken, muss man davon ausgehen, dass irgendwelche unvorhergesehenen Leaks entstehen können und werden.

12) Sind die sozialen Auswirkungen von KI derzeit ausreichend erforscht, oder benötigt es eine spezifische Forschungsethik und strukturelle wissenschaftliche Forschung/Evaluation, um Anwendungsbeispiele z.B. aus dem Bereich der Sicherheitstechnik kritisch zu hinterfragen und sicherzustellen, dass KI nicht diskriminiert und Ungleichheit verfestigt?

“KI” als singuläre Technologiefamilie herauszustellen halte ich an dieser Stelle für einen Holzweg: Wie die letzten Jahre zeigten sind “KI” Systeme am Markt oft gar keine im engeren Sinne sondern basieren zum Beispiel auf unsichtbar gemachter menschlicher Arbeit. Die dort teilweise stattfindende Diskriminierung ist davon aber nicht betroffen: Diskriminierung ist Diskriminierung. Es geht also viel mehr darum, Machtverhältnisse im allgemeinen zu verstehen, beschreiben und zu regulieren, nicht sich mit einer mehr oder weniger spezifisch beschriebenen Form von Technologie zu beschäftigen, auch wenn sie gerade ein Hype ist.

13) An welchen Stellen sehen Sie bei den Kompromissvorschlägen der tschechischen Ratspräsidentschaft vom 15. Juli 2022 noch Verbesserungsbedarf, was die Definition von KI, die Bestimmung von Hoch-Risiko-Systemen und die Einstufung von KI-Anwendungen in Annex III Betrifft?

Der Kompomissvorschlag ist in der Hinsicht gut, dass er die “KI” Definition enger an autonomes Handeln knüpft. Für die Risikobewertung geht der Vorschlag aber ein wenig zu weit und schränkt “Hochrisiko” zu sehr ein: Indem alle Systeme mit einem Menschen “im loop” aus der Hochrisikokategorie entfernt werden, hat man einen einfachen Weg geschaffen, Regulierung zu unterlaufen. Aus diversen Kontexten wissen wir, dass selbst wenn ein Mensch eine finale Entscheidung eines Softwaresystems formal akzeptieren muss, der Widerspruch gegen das automatisierte System oft faktisch nicht passiert (z.B. weil die Person dadurch im Job Risiken auf sich nimmt oder sich extreme Mehraufwände einhandelt).

14) Gibt die KI-Verordnung aus Ihrer Sicht genügend Freiraum für deutsche und europäische KI-Forschung, um mit den Forschungsbedingungen in den USA und China konkurrenzfähig zu sein und wo sehen Sie gegebenenfalls Bestimmungen der KI-Verordnung, die, gerade auch mit Blick auf die vorgesehenen Regelungen für Sandboxes, einschränkend auf zukünftige KI-Forschungsvorhaben, aber auch für den Transfer aus der Forschung in für den Markt zugelassene Produkte wirken könnten?

Die Dominanz von US und chinesischen Forschungszentren in der “KI” wird durch die VO nicht signifikant berührt. Am Ende geht es in dem Bereich um Datensätze und Ressourcen (“KI” ist reine Materialschlacht). Ob durch die VO nun einige wenige Datensätze in der EU schwieriger zu verwenden sind oder nicht, macht im Großen und Ganzen glaube ich wenig Unterschied.

15) Wie kann aus Ihrer Sicht eine einheitliche Auslegung des AI-Acts in allen EU-Mitgliedstaaten erreicht werden?

Vor allem natürlich durch eine enge Zusammenarbeit der Union mit den Mitgliedsstaaten. Hierbei muss besonderer Fokus darauf liegen, kein regulatives “Race to the Bottom” zu erzeugen, in dem sich einige Staaten durch deutliches Aufweichen der Vorgaben einen Wettbewerbsvorteil gegenüber den anderen verschaffen wollen.

16) Wie sollte Ihrer Meinung nach die Governance bei der Aufsicht und Kontrolle für KI-

Anwendungen aussehen, konkret, was die Ausgestaltung des europäischen AI-Boards angeht, dessen Zusammenarbeit mit den nationalen Behörden und die Kompetenzverteilung zwischen dem AI-Board und den nationalen Behörden, und welche Kriterien sollten für die Auswahl der nationalen Behörden Ihrer Meinung nach angesetzt werden?

Bei der Zusammensetzung des Boards steht für mich im Zentrum, dass insbesondere Vertreter*innen von besonders marginalisierten Gruppen dort eine einflussreiche Stimme erhalten. “KI”-Systeme der Art wie sie heute gebaut werden, haben eine grundsätzlich strukturerhaltende Qualität (da sie diese Strukturen aus den Trainingsdaten internalisieren), damit tragen sie in sich auch immer zu einem gewissen Maße die bestehenden Macht- und Diskriminierungsstrukutren in sich. Hier die potentiell am ehesten negativ Betroffenen auf oberster Ebene zu integrieren ist die große Herausforderung. Bei der Auswahl nationaler Behörden ist es wichtig, dass der Fokus nicht auf “Technologie” gelegt wird sondern auf die rechte und Würde der Menschen, d.h. “KI”-Regulierung ist kein Thema im Kontext “Digitalisierung” sondern im Kontext “Menschenrechte/Menschenwürde”.

17) Wäre es Ihrer Auffassung nach sinnvoll, die geplante Verordnung um einen eigenen Titel zu „Normen und Standards“ zu erweitern? Schließlich hat die Normung im Entstehen respektive Wachstum begriffener Technologien entscheidenden Anteil an der Marktfähigkeit konkreter Lösungen und damit Marktchancen einzelner gegenwärtiger und künftiger Anbieter. Sollte die Kommission es als ihre Aufgabe begreifen, die (Normungs-)Interessen deutscher und europäischer Akteure im Bereich der Künstlichen Intelligenz in den einschlägigen internationalen Gremien mit Nachdruck zu vertreten?

Grundsätzlich ist das keine schlechte Idee, explizit ist immer besser als implizit. Im Gegenzug findet im Bereich “KI” natürlich viel Normierung/Standardisierung statt auf Basis der technischen Frameworks und dominanten Technologien, weniger durch klassische Gremienarbeit. Hier wäre es wichtig sich zu überlegen, auf welchen Ebenen man hier Normierung gesetzlich festzurren kann und will.

18) Wäre es Ihrer Auffassung nach sinnvoll, die möglichen KI-Lösungen in der geplanten Verordnung nicht nur defensiv in Risikoklassen einzuordnen, sondern komplementär in Chancen- oder Wertigkeitsklassen? Ließe sich auf diese Weise nicht das enorme Innovations- und Schöpfungspotential von KI auf einem hoch dynamischen Markt betonen, was in Deutschland und in der EU im Gegensatz zu den USA und zu China zu selten und zu zaghaft geschieht?

Die Bewertung von Technologien muss immer ganzheitlich passieren, sowohl positive wie negative Externalitäten müssen gemeinsam erfasst, bewertet und gegeneinander abgewogen werden.
Unter dem Eindruck eines immer noch deutlich spürbaren “KI”-Hypes und einer aggressiven Präsenz von Anbietern am Markt ist es denke ich nicht das Problem, dass Potenziale oder Chancen von “KI”-Systemen unterrepräsentiert werden: Gerade aufgrund des doch immer signifikanten finanziellen Aufwandes zur Implementierung solcher Lösungen werden die

Chancen und Hoffnungen zuvor im Detail betrachtet und bewertet worden sein. Die Aufgabe der “KI”-Regulierung gerade muss es primär sein, den extremen Wildwuchs den wir insbesondere in Bereichen wie Biometrie, Affective Computing, Behavior Prediction und ähnlichen sehen, einzudämmen. Daher ist der Fokus der Regulierung auf Missbrauchspotenziale und Risiken der heutigen Situation absolut angemessen. Die EU und Deutschland argumentieren in Abgrenzung gegen USA/China oft mit “europäischen Werte”, die in EU/deutscher “KI” stecken würden. Die Regulierung über die möglichen Schäden trägt genau diesen Werten Rechnung.