Oktober 2020 · Leonie Beining

Vertrauenswürdige KI durch Standards?

Herausforderungen bei der Standardisierung und Zertifizierung von Künstlicher Intelligenz

Stiftung Neue Verantwortung

Think Tank für die Gesellschaft im technologischen Wandel

Executive Summary

Künstliche Intelligenz ist mittlerweile zentral für viele Lebensbereiche. Dementsprechend werden auch zunehmend höhere Ansprüche an KI-Systeme formuliert. Die Europäische Kommission und die Bundesregierung arbeiten derzeit an konkreten Vorschlägen zur Regulierung von Künstlicher Intelligenz (KI). Technische Standardisierung und Zertifizierung sind dabei ein vieldiskutierter Ansatz, um in Europa sogenannte vertrauenswürdige KI zu entwickeln und einzusetzen.

Während die Standardisierung Anforderungen an ein Produkt spezifiziert, kann mittels einer Zertifizierung die Einhaltung dieser Standards sichergestellt werden. Grundsätzlich wird die technische Standardisierung von Unternehmen vorangetrieben. Im Vergleich zu Standards werden Normen eine größere Legitimität zugesprochen, da sie von staatlich anerkannten Organisationen und nach festgelegten Grundsätzen erarbeitet werden. Ihre Einhaltung ist zwar grundsätzlich freiwillig, es kann aber zum Beispiel mittels Gesetzgebung auf sie verwiesen werden.

Standardisierung und Zertifizierung als Instrumente stehen im Kontext von KI jedoch vor erheblichen Herausforderungen, die einer stärkeren Debatte bedürfen: Zum einen müssen die Stärken und Schwächen bestehender Zertifizierungsregime in Europa in den Blick genommen werden. In verschiedenen Bereichen haben Standardisierung und Zertifizierung in Europa bereits heute mit Problemen zu kämpfen, die sich im Kontext von KI potenziell weiter verschärfen. Dazu zählen beispielsweise der hohe zeitliche Aufwand der Standardisierungsprozesse, die einmalige Prüfung vor Inverkehrbringung eines Produktes. Darüber hinaus gibt es Probleme bei der Marktüberwachung, der angesichts der Möglichkeit für Hersteller eine Selbsterklärung zur Konformität abzugeben, eine große Bedeutung zukommt.

Zum anderen verfügt KI über grundlegende Merkmale, die die Anforderungen an Standardisierung und Zertifizierung verändern und bisherige Prozesse infrage stellen:

- KI ist eine Basistechnologie, die einer äußerst hohen Forschungs- und Entwicklungsdynamik unterliegt, wodurch technische Standards potenziell sehr schnell veralten.
- KI-Systeme agieren probabilistisch statt deterministisch, was die Definition und Überprüfung technischer Anforderungen erschwert.
- KI-Systeme sind soziotechnische Systeme, deren potenzielle Auswirkungen auf die Gesellschaft und Kontextabhängigkeit besondere Ansprüche an Standardisierung und Zertifizierung stellen.

Zukünftig muss im Hinblick auf die Rolle von technischer Standardisierung bei KI angesichts der potentiellen gesellschaftlichen Auswirkungen von KI-Systemen zunächst geklärt werden, welche Bereiche durch freiwillige Selbstregulierung abgedeckt werden können und wo es die in demokratische Prozesse eingebettete Normung braucht. Abgesehen davon ist es wichtig, mehr über gesellschaftliche Beteiligung an technischer Standardisierung zu reflektieren und die Partizipation diverser Stakeholdergruppen zu fördern. Insgesamt werden Standardisierung und Zertifizierung nur einen gesellschaftlichen Nutzen haben können, wenn Kriterien sowie Prozesse gefunden werden, die aussagekräftige Bewertungen von KI ermöglichen. Dazu gehört zum Beispiel, die Forschung zu verallgemeinerbaren Prüfkriterien voranzutreiben, von Prüf- und Zertifizierungsprozessen in anderen Bereichen zu lernen und soziotechnische Kompetenzen im Rahmen der Standardisierung und Zertifizierung aufzubauen.

Die Autorin bedankt sich beim Team der SNV, insbesondere bei Jan-Peter Kleinhans und Stefan Heumann für die Unterstützung bei der inhaltlichen Konzeption des Papiers und ihr wertvolles Feedback, sowie bei Maria Jacob, Andre Weisser und Johanna Famulok für das Lektorat.

Inhalt

1. Einleitung

Welche Anforderungen stellen wir in Europa an die Entwicklung und den Einsatz von Künstlicher Intelligenz (KI)? Während in den letzten zwei Jahren vor allem zahlreiche Prinzipienkataloge und Kriterien aufgestellt wurden, arbeitet die Politik gegenwärtig an konkreten Vorschlägen für einen Regulierungsrahmen für KI. Dieser soll die gesellschaftlichen Ansprüche an KI weiter konkretisieren. Im Zentrum der Diskussion stehen dabei auch technische Standardisierung und Zertifizierung als Erfolg versprechende Mittel, um die Anforderungen an sogenannte „vertrauenswürdige KI" effizient in die Praxis umsetzen zu können.[1]

Dies erscheint naheliegend, definieren technische Standards doch Anforderungen an Produkte oder Prozesse, deren Einhaltung anschließend im Rahmen einer Zertifizierung überprüft werden kann. Das sorgt für Klarheit über die Eigenschaften sowie die Zuverlässigkeit eines Produkts und schafft damit die Grundlage für Vertrauen – und so potenziell auch die Grundlage für vertrauenswürdige KI. Die Erwartungen an Standardisierung und Zertifizierung von KI sind entsprechend hoch. Sie werden nicht nur als Instrument zur Stärkung der Innovationskraft und Wettbewerbsfähigkeit, sondern auch als Mittel zur Gestaltung und Nutzung von KI im Sinne des Gemeinwohls gesehen. Allerdings stellt KI die technische Standardisierung und Zertifizierung vor Herausforderungen, die in der Debatte bislang unterbelichtet bleiben. Diese müssen jedoch dringend diskutiert werden, um zu bewerten, was Standardisierung und Zertifizierung bei der Regulierung von KI leisten können.

Das Ziel des vorliegenden Impulspapiers ist es, einen Blick auf das Thema Standardisierung und Zertifizierung zu werfen, verschiedene Herausforderungen zu benennen und damit technische Standardisierung und Zertifizierung im Kontext von KI einem Realitätscheck zu unterziehen. Dafür werden zunächst die grundlegenden Funktionslogiken von technischer Standardisierung und Zertifizierung sowie das in der EU etablierte System skizziert, um auf bereits bestehende Probleme aufmerksam zu machen. Daran schließt sich ein Überblick an, welche besonderen Herausforderungen KI-basierte Systeme potenziell an Standardisierung und Zertifizierung stellen werden. Das Papier dient als Einladung für Feedback und weitere Diskussionen, um darauf aufbauend zielführende Ideen und Impulse für einen angemessenen Umgang mit KI zu erarbeiten.

1 High-level Expert Group on AI (2019), Ethics guidelines for trustworthy AI https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. Abgerufen am 9.10.2020.

2. Technische Standardisierung und Zertifizierung in der EU

Um die Herausforderungen bei der Standardisierung und Zertifizierung von KI zu verstehen, ist es zunächst wichtig, die Zusammenhänge zwischen Standardisierung und Zertifizierung zu beleuchten. Darüber hinaus gilt es zu klären, auf welchen Funktionslogiken diese Zusammenhänge beruhen.

Technische Standardisierung verfolgt das Ziel, technische Anforderungen an Produkte zu vereinheitlichen, sie kompatibel zu gestalten und ihre Qualität sicherzustellen. Standards und Normen helfen auf diese Weise bei der Erschließung neuer Märkte, unterstützen den Handel und die Innovationskraft der Wirtschaft.[2] Die technische Standardisierung wird daher in erster Linie von Unternehmen vorangetrieben und mittlerweile gibt es eine Vielzahl von Organisationen, die als Foren für die technische Standardisierung dienen und die Entwicklung von Standards und Normen fördern.[3]

Dabei gilt, dass eine Norm immer ein Standard, ein Standard aber nicht immer eine Norm ist:[4] Normen werden nach festgelegten Grundsätzen durch anerkannte Organisationen erarbeitet, wie beispielsweise vom DIN in Deutschland, CEN/CENELEC auf EU-Ebene und der ISO auf internationaler Ebene.[5] Dazu gehört beispielsweise, dass Normen im Konsens und öffentlich, das heißt, auch unter Einbeziehung interessierter Stakeholder aus Wirtschaft, Wissenschaft, Politik und Zivilgesellschaft erstellt werden. Im Gegensatz zu Normen, kann die Erarbeitung von Standards durch einzelne oder meh-

2 Zum volkswirtschaftlichen Nutzen der Standardisierung insgesamt siehe: Blind, Knut, Jungmittag, Andre und Axel Mangelsdorf (2011): Der gesamtwirtschaftliche Nutzen der Normung. Eine Aktualisierung der DIN-Studie aus dem Jahr 2000, DIN Deutsches Institut für Normung e.V., Berlin. Abgerufen am 23.09.2020, von https://www.din.de/resource/blob/79542/946e70a818ebdaacce9705652a052b25/gesamtwirtschaftlicher-nutzen-der-normung-data.pdf.

3 Für einen Überblick über Standardisierungs- und Normungsorganisationen, die technische Standardisierung von KI vorantreiben: Lorenz, Philippe (2020): AI Governance through Political Fora and Standards Developing Organizations. Abgerufen am 07.10.2020, von https://www.stiftung-nv.de/sites/default/files/ai_governance_through_political_fora_and_standards_developing_organizations.pdf.

4 Rusche, Christian (2017): Potenziale von Standards für die deutsche Wirtschaft, IW policy paper 2/2017. Abgerufen am 07.10.2020, von https://www.iwkoeln.de/fileadmin/publikationen/2017/323185/IW-policy_paper_2017_2__Potenziale_von_Standards.pdf, S.6.

5 Die Grundprinzipien sind Kohärenz, Transparenz, Offenheit, Freiwilligkeit der Anwendung, Unabhängigkeit von Einzelinteressen und Effizienz, siehe: Europäisches Parlament und Rat der EU (2012): Verordnung (EU) Nr. 1025/2012, 25.10.2012. Abgerufen am 23.09.2020, von https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32012R1025&from=DE, S. 12.

rere Stakeholder selbst initiiert werden, es bedarf dabei keiner Beteiligung der Öffentlichkeit. Organisationen, die solche Standardisierungsaktivitäten koordinieren, sind etwa IEEE oder IETF. Aufgrund der stärker institutionalisierten Verfahren wird Normen im Vergleich zu Standards eine größere Legitimität zugesprochen.[6] In beiden Fällen handelt es sich aber in erster Linie um das Ergebnis eines Aushandlungsprozesses von Unternehmen für Unternehmen.[7]

Neben den unmittelbaren wirtschaftlichen Vorteilen ist technische Standardisierung in erster Linie ein Instrument zur Selbstregulierung. Normen und Standards definieren zwar Anforderungen, ihre Einhaltung ist aber grundsätzlich freiwillig. Sie können „de-facto" oder „de-jure" jedoch auch verbindlich werden: Unternehmen können sich im Rahmen von Verträgen zur Einhaltung von Standards verpflichten oder sie sind durch ihre Marktmacht in der Lage, eigene Standards zu setzen, die für die übrigen Marktteilnehmer dann effektiv verbindlich werden. Darüber hinaus können auch Gesetze auf Normen verweisen, d.h. dass die Einhaltung einer Norm empfohlen oder sogar vorgeschrieben wird, um gesetzliche Vorschriften zu erfüllen.[8]

In der EU wurde eine solche Kopplung der technischen Standardisierung an die Regulierung durch den sogenannten "New Approach" als Public Private Partnership verankert.[9] Das bedeutet, dass der Gesetzgeber durch sogenannte technische Anforderungen den rechtlichen Rahmen schafft, aber die technischen Spezifikationen, die eine Erfüllung der technischen Anforderungen gewährleisten, nicht selbst vornimmt. Dies überlässt er den drei europäischen Normungsorganisationen CEN, CENELEC und ETSI, die gemeinsam mit der Industrie die Regulierungsziele durch die technische Standardisierung in Form sogenannter „harmonisierter Normen" konkretisieren. Die Einhaltung einer solchen harmonisierten Norm stellt in der Regel den einfachsten Weg dar, um die vom Gesetz vorgeschriebenen technischen Anforderungen zu erfüllen, wenn auch ihre Anwendung in den meisten Fällen nicht verpflichtend ist.[10]

6 Mangelsdorf, Axel (2019): Normen und Standards in der KI. In: Wittpahl V. (eds): Künstliche Intelligenz. Springer Vieweg, Berlin, Heidelberg. Abgerufen am 07.10.2020, von https://doi.org/10.1007/978-3-662-58042-4_3.

7 Rühlig, Tim Nicolas (2020): Technical Standardisation, China and the Future International Order. A European Perspective, Heinrich-Böll-Stiftung, Brüssel. Abgerufen am 09.09.2020, von https://eu.boell.org/sites/default/files/2020-03/HBS-Techn%20Stand-A4%20web-030320-1.pdf.

8 Europäische Union (2020): Europäische Normen. Abgerufen am 07.10.2020, von https://europa.eu/youreurope/business/product-requirements/standards/standards-in-europe/index_de.htm.

9 Aktualisiert im Jahr 2008 durch das New Legislative Framework.

10 ebd.

Ein Beispiel für europäische Normungsarbeit ist das CE-Kennzeichen zur Sicherstellung der physischen Produktsicherheit in der EU: Der Gesetzgeber hat festgelegt, dass bestimmte Produkte verpflichtend eine CE-Kennzeichnung benötigen, um im Binnenmarkt verkauft werden zu dürfen (Produktsicherheitsrichtlinie). Dazu erarbeiten die europäischen Normungsorganisationen gemeinsam mit der Industrie technische Standards und Unternehmen bescheinigen dann im Rahmen von Konformitätsbewertungen, dass ihre Produkte den Anforderungen genügen.

Eine solche Konformitätsbewertung kann in Form einer Selbsterklärung oder einer Zertifizierung durch eine dafür akkreditierte Prüfstelle erfolgen. Im Vergleich zur zeit- und kostenaufwändigen Zertifizierung durch Dritte stellt die Selbsterklärung für Unternehmen einen günstigen und unkomplizierten Weg dar, um ihre Produkte schnell an den Markt zu bringen. Die Selbsterklärung ist mit einem großen Vertrauensvorschuss verbunden, da das Produkt auf den Markt kommt, ohne dass ein unabhängiges Prüflabor sicherstellt, dass es tatsächlich alle relevanten Anforderungen erfüllt. Stattdessen wird der Aussage des Herstellers vertraut. Dies ermöglicht es den Herstellern grundsätzlich, die gestellten Anforderungen zu umgehen.[11]

Angesichts der Möglichkeit für Unternehmen, sich auf Selbsterklärungen zu beschränken, kommt daher der Marktüberwachung eine zentrale Rolle zu. In jedem EU-Mitgliedsstaat gibt es Marktüberwachung, um sicherzustellen, dass Produkte auf dem Markt tatsächlich den Anforderungen entsprechen. Die Institutionen der Marktüberwachung verfügen über das Mandat, Produkte stichprobenhaft zu kontrollieren und gegebenenfalls Sanktionen zu verhängen.[12]

Das europäische Regime aus technischer Standardisierung und Zertifizierung bringt allerdings einige Probleme mit sich: So handelt die Normung durch die Einbettung in die Regulierung zwar gewissermaßen im politischen und demokratisch legitimierten Auftrag. Gleichzeitig ist sie wegen der Abstimmungs- und Konsensbildungsprozesse aber auch langsam und langwierig. Obwohl Normung als flexibles Regulierungsinstrument gilt, dauert es

11 Kleinhans, Jan-Peter (2018): Strukturelle Schwächen des EU Cybersecurity Acts- Grundsätzliche Herausforderungen bei der Regulierung von IT-Sicherheit, BvD-News, 3/18. Abgerufen am 09.09.2020, von https://www.bvdnet.de/wp-content/uploads/2018/11/News18-3_web.pdf, S.69.

12 Aktueller Gesetzesentwurf des Europäischen Rates zum Ausbau einer einheitlichen europäischen Marktregulierung: Europäischer Rat (2020): Neue Vorschriften zu Produktsicherheit und Marktregulierung. Abgerufen am 07.10.2020, von https://www.consilium.europa.eu/de/policies/product-safety-market-surveillance/.

mehrere Jahre bis Normen fertiggestellt werden.[13] Zudem ist es für Akteure aus der Wissenschaft und Zivilgesellschaft in der Regel schwierig, an ihrer Entwicklung effektiv teilzunehmen, da es hierfür hoher technischer Kompetenz sowie finanzieller Ressourcen, aber auch Zugang zu rechtzeitiger Information und institutionellem Wissen bedarf.[14] Die Möglichkeit zur Partizipation, und letztlich Einflussnahme, ist damit an Voraussetzungen geknüpft, die in erster Linie die Industrie erfüllen kann. So bringen sich am Ende vor allem die ein, die starke wirtschaftliche Interessen an der Standardisierung haben und über die entsprechenden Ressourcen verfügen.

Zu diesen strukturellen Problemen der Normung kommt hinzu, dass die Regelungen zur Konformitätsbewertung bislang eine einmalige Prüfung *vor* dem Inverkehrbringen eines Produktes vorsehen. Das mag bei physischen Produkten, die sich einmal auf den Markt gebracht kaum verändern, unproblematisch sein. Bei Softwareprodukten etwa, die sich durch regelmäßige Updates auch nach ihrer Markteinführung ständig verändern können, ist ein solches Vorgehen jedoch wenig sinnvoll, da Zertifizierungen lediglich Momentaufnahmen sind, die schnell an Aussagekraft verlieren können.[15]

Erfahrungen im Rahmen der CE-Kennzeichnung verdeutlichen zudem die großen Probleme, vor denen die Marktüberwachung steht: Durch die unzureichende Ressourcenausstattung und die Heterogenität der Behörden in den Mitgliedsstaaten schafft sie es kaum, ihre Aufgabe angemessen zu erfüllen. Dies birgt das Risiko, dass auch mangelhafte Produkte auf den Markt kommen.[16]

Wenn es zukünftig um die Ausgestaltung eines europäischen KI-Zertifizierungsschemas für „vertrauenswürdige KI“ gehen soll, ist es von zentraler Bedeutung, die Funktionsweise und Defizite der bestehenden Systeme in den

13 Rühlig (2020)

14 Cihon, Peter (2019): Standards for AI Global Governance: International Standards to Enable Global Coordination in AI Research & Development. Abgerufen am 09.09.2020, von https://www.fhi.ox.ac.uk/wp-content/uploads/Standards_-FHI-Technical-Report.pdf.

15 Zu Problemen bei der Prüfung und Überwachung von physischer Produktsicherheit, siehe auch: ANEC und BEUC (2020): BEUC and ANEC Views for a Modern Regulatory Framework on Product Safety. Achieving a Higher Level of Consumer Safety through a Revision of the General Product Safety Directive, Brüssel. Abgerufen am 23.09.2020, von https://www.beuc.eu/publications/beuc-x-2020-068_beuc_and_anec_views_for_a_modern_regulatory_framework_on_product_safety.pdf.

16 Algemene Rekenkamer (2017): Products sold on the European market: unravelling the system of CE marking, Netherlands. Abgerufen am 23.09.2020, von https://bit.ly/303xVxw. Die EU Kommission hat unlängst einen Entwurf zur Reform der Marktüberwachung vorgelegt, siehe Fußnote 11.

Blick zu nehmen. Das bestehende Regime kommt an vielen Stellen an seine Grenzen. Diese Herausforderungen müssen grundsätzlich angegangen und auch bei der Entscheidung über einen Ordnungsrahmen für KI miteinbezogen werden.

Ein Aspekt, der häufig mit Standardisierung und Zertifizierung vermischt wird, sind Gütesiegel oder -label. Solche Label sollen dabei helfen, Nutzer:innen über Eigenschaften und Qualität eines Produkts aufzuklären, und sind damit in erster Linie ein Instrument zur Verbraucher:innenkommunikation. Das Vorhandensein eines solchen Siegels bedeutet zwar in manchen Fälllen, wie etwa beim europäischen Bio-Siegel, dass es auf einer Zertifizierung basiert oder gesetzliche Mindestanforderungen referenziert, dies ist jedoch nicht automatisch der Fall. Ebenso kann es das Ergebnis einer rein von der Privatwirtschaft getragenen Initiative sein.

Entsprechend gibt es in vielen Bereichen zahlreiche verschiedene Siegel, was den Überblick für die Verbraucher:innen erschwert. Zudem geht es um die Frage, welche Produktaspekte durch ein Label erfasst und in welcher Form effektiv kommuniziert werden können. Positive Beispiele wie das Energieeffizienzlabel, können sich dabei auf klar quantifizier- und messbare Eigenschaften stützen, die nachvollziehbar und darüber hinaus mit klaren Vorteilen für Verbraucher:innen verbunden sind. Digitale Themen, wie IT-Sicherheit oder Datenschutz, sind demgegenüber wenig intuitiv und die Folgen für Nutzer:innen schwer zu erfassen. Bislang konnten sich in diesen Bereichen auch keine Siegel etablieren.[17] Ob dies im Kontext von KI anders sein wird, ist fraglich, zumal sehr viele Informationen in einem Siegel untergebracht werden müssten, um Qualität und potentielle Risiken einer Anwendung in Gänze abzubilden, was die Komplexität des Siegels erhöhen würde.

17 Balboni, Paolo und Dragan, Theodora (2018): Controversies and Challenges of Trustmarks: Lessons for Privacy and Data Protection Seals. In: Rodrigues, R., Papakonstantinou V. (eds) Privacy and Data Protection Seals. Information Technology and Law Series, Vol 28. T.M.C. Asser Press, The Hague. Abgerufen am 23.09.2020, von https://link.springer.com/chapter/10.1007/978-94-6265-228-6_6.

3. Warum es Standardisierung und Zertifizierung bei KI schwer haben

Während technische Standardisierung und Zertifizierung in Europa bereits heute mit grundsätzlichen Problemen zu kämpfen hat, stellt KI sie noch vor weitaus größere Herausforderungen. KI besitzt grundlegende Merkmale, die die Anforderungen an Standardisierung und Zertifizierung verändern und bisherige Prozesse infrage stellen:

- KI ist eine Basistechnologie, die einer äußerst hohen Forschungs- und Entwicklungsdynamik unterliegt, wodurch technische Standards potenziell schnell veralten.
- KI-Systeme agieren probabilistisch statt deterministisch, was die Definition und Überprüfung technischer Anforderungen erschwert.
- KI-Systeme sind soziotechnische Systeme, deren potenzielle Auswirkungen auf die Gesellschaft und Kontextabhängigkeit hohe Ansprüche an Standardisierung und Zertifizierung stellen.

Dynamische Basistechnologie

Bei KI handelt es sich um ein Forschungsfeld, das sich gegenwärtig durch eine hohe Dynamik auszeichnet und schnell weiterentwickelt. Das zeigen etwa der starke Anstieg von Veröffentlichungen in wissenschaftlichen Zeitschriften oder die wachsende Teilnehmendenzahl bei wissenschaftlichen KI-Konferenzen.[18] Insbesondere Anwendungen wie Bilderkennung oder Sprach- und Textverständnis bis hin zu Sprachgenerierung wurden erst im Laufe der letzten zehn Jahre zur Marktreife gebracht und verdeutlichen damit die schnell wachsende Leistungsfähigkeit von KI-Systemen. Im Umkehrschluss bedeutet dies, dass geltende technische Standards auch entsprechend schnell veralten können und die technische Standardisierung vor der Herausforderung steht, Standards schnell aktualisieren oder neu formulieren zu müssen. Darüber hinaus sind viele Fragen rund um KI, wie zum Beispiel bezüglich ihrer Erklärbarkeit und Nachvollziehbarkeit, nach wie vor Gegenstand intensiver Forschung und rücken gerade erst ins Zentrum der Aufmerksamkeit der Industrie. Das birgt das Potenzial, Standardisierungsbedarfe und Zertifizierungsverfahren zukünftig immer wieder zu verändern. Vor dem Hintergrund, dass Standardisierungs- und vor allem Normungspro-

18 HAI (2019): The 2019 AI Index Report. Abgerufen am 23.09.2020, von https://hai.stanford.edu/research/ai-index-2019.

zesse aufwändig sind und Zeit brauchen, stellt sich damit die Frage, inwieweit die Entwicklungszyklen von KI-Technologien einerseits und von technischen Standards andererseits in der Realität zusammenpassen.

Es kommt hinzu, dass technische Standardisierung im Kontext von KI zu einer immer komplexeren Aufgabe wird. Als Basistechnologie kommt KI zum Beispiel in verschiedenen Branchen und Sektoren zur Anwendung, die unterschiedliche Ansprüche und Anforderungen mit sich bringen. Neben sektorspezifischer Standardisierungsarbeit macht dies vor allem viel Koordination und Abstimmung erforderlich, um Doppelungen und Widersprüche zu vermeiden.[19] Darüber hinaus fehlt es nach wie vor an einem gemeinsamen Vokabular. Letzteres gilt jedoch als Voraussetzung, um die Technologie und ihre Anforderungen präzise beschreiben zu können. Auch wenn international an definitorischen Grundlagen und Terminologiestandards gearbeitet wird, gibt es hier noch grundlegenden Handlungsbedarf. Trotz vielfältiger Standardisierungsaktivitäten weltweit steht die technische Standardisierung von KI damit noch ganz am Anfang.[20]

Hinzu kommt, dass es bereits heute oftmals schwierig ist, abzuschätzen, ob ein technischer Standard die erwartete Wirkung erzielt und die Einhaltung des Standards tatsächlich zum Erreichen des angestrebten Ziels führt. Aus diesem Grund wurde etwa in der EU damit begonnen, unabhängige Berater:innen einzusetzen, um sicherzustellen, dass die im Rahmen der europäischen Normungsorganisationen erarbeiteten harmonisierten Normen im Resultat auch die festgelegten Regulierungsziele erfüllen.[21] Es ist zu erwarten, dass auch diese Aufgabe im Kontext von KI noch einmal schwieriger wird.

Neue technische Qualitäten

Ein zentrales Wesensmerkmal heutiger KI-Systeme ist, dass sie lernende Systeme sind. Das bedeutet, dass sich die durch KI-Systeme produzierten Ergebnisse im Laufe der Zeit verändern. Darüber hinaus kann auch eine Veränderung der Umgebung, beziehungsweise der Eingabedaten, veränderte und unbeabsichtigte Ergebnisse zur Folge haben.

19 Kolliarakis, Georgios und Hermann, Isabella (2020): Towards European Anticipatory Governance for Artificial Intelligence, DGAP Report, 9/2020. Abgerufen am 09.09.2020, von https://dgap.org/sites/default/files/article_pdfs/dgap_report_no._9_april_29_2020_60_pp.pdf.

20 Joinup (2020): Rolling Plan for ICT Standardisation. Abgerufen 09.09.2020, von https://joinup.ec.europa.eu/collection/rolling-plan-ict-standardisation/artificial-intelligence.

21 StandICT.eu: https://2020.standict.eu/about-standicteu

Bereits klassische deterministische IT-Systeme, deren Verhalten sich durch kontinuierliche Updates ständig verändern kann, stellen bisherige Verfahren zur Zertifizierung vor Probleme, weil Zertifizierungen nur Momentaufnahmen sind und schnell veralten können. Bei KI-Systemen stellt sich aber nicht nur diese „Alterungsproblematik" und damit die zentrale Frage, ob überhaupt ein geeigneter Zeitpunkt für eine Zertifizierung gefunden werden kann. Während Veränderungen bei klassischen IT-Systemen immerhin nach festgelegten und damit exakt nachvollziehbaren Anforderungen – also deterministisch – erfolgen, etwa wenn IT-Produkten eine neue Funktionalität hinzugefügt wird, ist dies bei KI in der Regel nicht der Fall.

KI-Systeme sind zwangsläufig probabilistischer Natur, das heißt die durch sie berechneten Aussagen basieren auf Wahrscheinlichkeiten und nicht ausschließlich auf deterministischen Regeln. Nur so sind KI-Systeme überhaupt in der Lage, unerkannte Zusammenhänge und Muster zu entdecken. In der Praxis hat dies zur Folge, dass sich die Güte von KI-Systemen nie allgemein durch das Verfahren selbst, sondern immer nur durch Anwendung auf einen konkreten Satz an Eingabedaten bewerten lässt.

Während sich klassische Standardisierungsverfahren bislang auf universell und unabhängig überprüfbare Kriterien stützen konnten – ein DIN-A4-Blatt muss eben exakt 210 × 297 mm groß sein – fehlt es bei KI-Systemen an solchen generalisierbaren Qualitätsmaßen, die über einen konkreten Eingabe-Datensatz ihre Geltung behalten. Beispielsweise kann ein KI-System, das Straßenbegrenzungen bei Sonnenschein („Datensatz A") möglicherweise mit extrem hoher Sicherheit erkennt, in einer verregneten Umgebung („Datensatz B") gänzlich versagen. Aus letzterem Beispiel ergibt sich also unmittelbar die Frage, wie realistisch es ist, für Zertifizierungen geeignete Beispieldatensätze zu finden, oder ob sich eine derartige Zertifizierung immer auf sehr eng definierte Anwendungsfälle beschränken müsste. Bislang sind allgemeingültige, generalisierbare und unabhängig überprüfbare Qualitätsmaße, die über einen konkreten Eingabe-Datensatz hinweg ihre Geltung behalten, für KI-Systeme noch nicht bekannt.

KI als teilhaberelevantes soziotechnisches System

Als Gesellschaft sollten wir besonders hohe Ansprüche an die Qualität von KI-Systemen stellen, die in zentralen gesellschaftlichen Bereichen eingesetzt werden und so das Leben von Menschen maßgeblich beeinflussen können. Dazu zählen etwa die Bereiche Gesundheit, Bildung, Arbeitsmarkt oder öffentliche Sicherheit. In diesen Bereichen wird eine besonders eingehende

Prüfung erforderlich sein. Als soziotechnische Systeme hängt die Funktions- und Wirkungsweise und damit die Qualität von KI-basierten Systeme aber stark von ihrer Einbettung – das heißt von der Frage, wer nutzt die Technologie zu welchem Zweck – und auch von der Passgenauigkeit für den jeweiligen Kontext ab. Bei einer Zertifizierung reicht es daher nicht aus, isoliert die Technologie in den Blick zu nehmen und Anforderungen ausschließlich an die technischen Komponenten zu stellen. Vielmehr muss im Rahmen der Standardisierung und Zertifizierung idealerweise der gesamte Kontext der Technologieentwicklung und -nutzung berücksichtigt werden, um zu aussagekräftigen Bewertungen kommen zu können.

Aufgrund der technischen Komplexität, der gesellschaftlichen Relevanz sowie der Kontextabhängigkeit von KI-Systemen ist bislang unklar, welche Rahmenbedingungen es bräuchte, um die Güte von KI-Systemen im Kontext einer Zertifizierung überhaupt realistisch überprüfen zu können. Den kompletten KI-Lebenszyklus im Auge zu behalten macht die ohnehin schon langen und teuren Prüfungen noch komplexer und aufwändiger. Die Anforderungen, die sich zukünftig an die Prüfenden (und die Institutionen der Marktüberwachung) stellen, werden noch weiter wachsen. Das heißt, Zertifizierungen werden nicht nur ressourcenintensiver, sie erfordern auch noch zusätzliche Kompetenzen. So braucht es nicht nur hohe technische, statistische und juristische Kompetenz, sondern auch soziotechnische Expertise und Erfahrungen aus den jeweiligen Anwendungskontexten. Aussagekräftige Prüfungen setzen zudem den Zugang zu umfangreichen Informationen über einzelne Prozessschritte, wie zum Beispiel über Entscheidungen im Entwicklungsprozess, voraus. Hier stellt sich die Frage, inwieweit die bestehende Prüfinfrastruktur “fit for purpose” ist und inwieweit sie in die Lage versetzt werden kann, ihre anspruchsvollen Aufgaben gut zu bewältigen.

Angesichts der gesellschaftlichen Relevanz und der potentiellen Risiken von KI-Systemen, arbeiten verschiedene Standardisierungs- und Normungsinitiativen daran, auch ethische Anforderungen, wie Fairness und Verantwortung an KI, im Rahmen der Standardisierung zu konkretisieren und sie damit zertifizierbar zu machen. Allerdings ist die Frage, welche Werte wann wichtig sind und was Werte definiert, höchst normativ, abhängig vom gesellschaftlichen Kontext und nie abschließend zu beantworten. Es ist daher höchst fraglich, inwieweit unklare, normative Konzepte, für die es mitunter kein gesellschaftlich geteiltes Verständnis gibt, im Rahmen der technischen Standardisierung überhaupt ausgewählt und in eindeutige, allgemeingültige An-

forderungen überführt werden können.[22]

Vor diesem Hintergrund erscheinen auch die strukturellen Probleme der technischen Standardisierung in einem neuen Licht. Obwohl die transparente Beteiligung interessierter Kreise zu den Grundsätzen der Normung gehört, fehlt es im Rahmen von Standardisierungsprozessen in der Regel an diversen Stimmen. Aus gesellschaftlicher Sicht ist das schon jetzt problematisch. Als Technologie, die sich potenziell stark auf unser gesellschaftliches Zusammenleben auswirkt und Risiken für soziale Teilhabe birgt, braucht es bei KI die Einbindung unterschiedlicher gesellschaftlicher Interessen noch dringender. Umfassende Partizipation ist insofern auch eine notwendige Voraussetzung, wenn mit Hilfe technischer Standardisierung „ethische KI" vorangetrieben werden soll.

4. Fazit

Technische Anforderungen zu definieren und deren Einhaltung durch eine Zertifizierung zu garantieren, erscheint in der Theorie als ein sinnvoller Weg in Richtung vertrauenswürdiger KI. Standards und Normen bieten die Möglichkeit, eine gemeinsame Sprache zu finden, die dringend benötigt wird, wenn es darum geht, über unseren Umgang mit KI zu bestimmen. Indem sie gesellschaftliche Anforderungen spezifizieren und überprüfbar machen, haben Standardisierung und Zertifizierung das Potenzial Transparenz, Verlässlichkeit und damit Vertrauen in Technologie zu schaffen. Die Realität ist aber oftmals komplexer und das gilt insbesondere für das sich nach wie vor im Entstehen befindende Feld der KI.

Zunächst darf nicht vergessen werden, dass technische Standardisierung vornehmlich Verantwortlichkeit der Industrie ist und von Unternehmen vorangetrieben wird. Das ist grundsätzlich gut und richtig, schließlich sind sie es, die die Produkte, um die es geht, entwickeln. Hinsichtlich der gesellschaftlichen Auswirkungen von KI-Systemen ist aber zu klären, welche Bereiche durch freiwillige Selbstregulierung abgedeckt werden können und wo es die in demokratische Prozesse eingebettete Normung braucht. Gesellschaftliche Fragen rund um den Einsatz von KI lassen sich nicht allein durch die Formulierung technischer Standards lösen, sondern brauchen de-

22 Dazu siehe auch: Hottenrott, Moritz, Thorwarth, Susanne und Christian Wey (2016): Gegenstandsbereich der Normung, DICE Ordnungspolitische Perspektiven, 83/März 2016. Abgerufen am 07.10.2020, von https://www.dice.hhu.de/fileadmin/redaktion/Fakultaeten/Wirtschaftswissenschaftliche_Fakultaet/DICE/Ordnungspolitische_Perspektiven/083_OP_Hottenrott_Thorwarth_Wey.pdf, S. 16 ff.

mokratisch legitimierte Antworten. Dazu zählt auch, dass die Ziele, die durch den Einsatz von KI erreicht werden sollen, zunächst gesellschaftlich reflektiert und entsprechend eingegrenzt werden müssen. Es muss klar sein, dass KI-Systeme, die dem gesellschaftlichen Interesse zuwiderlaufen, auch durch Standardisierung und Zertifizierung nicht zu akzeptablen Systemen werden. Gleichzeitig muss zukünftig aber auch mehr über zivilgesellschaftliche Beteiligung an technischer Standardisierung nachgedacht werden. Es müssen Formate gefunden werden, die verschiedene Stakeholder und Expert:innen aus betroffenen Anwendungsfeldern effektiv einbinden, damit verschiedene gesellschaftliche Stimmen im Normungsprozess besser Gehör finden.

Standardisierung und Zertifizierung werden darüber hinaus nur einen gesellschaftlichen Nutzen haben, wenn Kriterien sowie Prozesse identifiziert werden, die aussagekräftige Bewertungen von KI ermöglichen. Es braucht Forschung zu verallgemeinerbaren Prüfkriterien und es ist zu klären, ob und in welcher Form sich gesellschaftliche Anforderungen wie zum Beispiel Fairness oder Sicherheit durch Standards formulieren lassen. Klar ist auch, dass etwa eine einmalige Prüfung von KI-Systemen nur bedingt aussagekräftig ist und Prüfprozesse zukünftig anders gedacht werden müssen, als dies bei klassischer Zertifizierung zumeist der Fall ist. Zum Beispiel als dauerhaftes Monitoring oder angelehnt an Verfahren, wie es sie etwa im Bereich der Steuer- oder Wirtschaftsprüfung gibt. Sicher ist, dass es spezieller Kompetenzen, Kontextwissen und entsprechender Ressourcen bedarf, um soziotechnische KI-Systeme angemessen prüfen zu können.

Da die Entwicklung von technischen Standards, und damit die Grundlage einer jeden Zertifizierung von KI-Systemen, aufwändig ist und Zeit benötigt, ist es wichtig, schon frühzeitig über diese Fragen nachzudenken. Hier gilt es auch von Erfahrungen mit Standardisierung und Zertifizierung aus anderen Bereichen, wie etwa der physischen Produktsicherheit und vor allem der IT-Sicherheit, zu lernen, um schließlich bewerten zu können, wie tragfähig bereits etablierte Ansätze bei KI tatsächlich sind. [23]

23 Europäische Kommission (2020): White Paper On Artificial Intelligence - A European approach to excellence and trust. Abgerufen am 23.09.2020, von https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf, S. 23.

Über die Stiftung Neue Verantwortung

Die Stiftung Neue Verantwortung (SNV) ist ein gemeinnütziger Think Tank, der an der Schnittstelle von Technologie und Gesellschaft arbeitet. Die Kernmethode der SNV ist die kollaborative Entwicklung von Politikvorschlägen und -analysen. Die Expert:innen der SNV arbeiten nicht allein, sondern entwickeln und testen Ideen gemeinsam mit Vertreter:innen aus Politik und Verwaltung, Technologieunternehmen, Zivilgesellschaft und Wissenschaft. Unsere Expert:innen arbeiten unabhängig von Interessengruppen und Parteien. Unsere Unabhängigkeit gewährleisten wir durch eine Mischfinanzierung, zu der viele verschiedene Stiftungen, öffentliche Mittel und Unternehmensspenden beitragen.

Über die Autorin

Leonie Beining ist Projektleiterin bei der Stiftung Neue Verantwortung (SNV) und setzte in der Funktion verschiedene Projekte an der Schnittstelle zwischen Technologie und Gesellschaft um. Zuletzt arbeitete sie im Rahmen des Projektes „Algorithmen fürs Gemeinwohl“ zur Nachvollziehbarkeit von algorithmischen Entscheidungssystemen. Zuvor war sie bereits als Projektmanagerin im Projekt „Gemeinwohl im digitalen Zeitalter“ tätig und entwickelte Strategien, wie sich zivilgesellschaftliche Akteure stärker in die Gestaltung der Digitalisierung einbringen können. Ab Herbst 2020 arbeitet Leonie im Konsortialprojekt ExamAI an der Entwicklung von politischen Handlungsempfehlungen zur Umsetzung von Prüfverfahren für KI-Systeme.

So erreichen Sie die Autorin

Leonie Beining
+49 (0) 30 81 45 03 78 81
lbeining@stiftung-nv.de
Twitter @LeonieBeining

Impressum

Stiftung Neue Verantwortung e. V.
Beisheim Center
Berliner Freiheit 2
10785 Berlin

T: +49 (0) 30 81 45 03 78 80
F: +49 (0) 30 81 45 03 78 97

www.stiftung-nv.de
info@stiftung-nv.de

Design:
Make Studio
www.make-studio.net

Layout:
Jan Klöthe