2022

technische universität dortmund

sfs Sozialforschungsstelle Dortmund

BEITRÄGE AUS DER FORSCHUNG
Band 211

Thorben Krokowski & Hartmut Hirsch-Kreinsen

Vertrauenswürdige KI Made in Germany und Europe: Distinktionsmerkmal oder Verkaufstrick?

Impressum
Beiträge aus der Forschung, Band 211
ISSN: 0937-7379
Dortmund 2022

Sozialforschungsstelle Dortmund (sfs)
Fakultät Sozialwissenschaften | Technische Universität Dortmund
Evinger Platz 17
D-44339 Dortmund

Tel.: +49 (0)2 31 – 755-1
Fax: +49 (0)2 31 – 755-90205
Email: information.sfs@tu-dortmund.de
www.sfs.sowi.tu-dortmund.de

Thorben Krokowski & Hartmut Hirsch-Kreinsen

Vertrauenswürdige KI Made in Germany und Europe: Distinktionsmerkmal oder Verkaufstrick?

Thorben Krokowski & Hartmut Hirsch-Kreinsen

Vertrauenswürdige KI Made in Germany und Europe: Distinktionsmerkmal oder Verkaufstrick?

Trustworthy AI Made in Germany and Europe: Unique feature or sales gimmick?

Zusammenfassung

Um sich im weltweiten Wettrennen um die Technologieführung durchzusetzen oder zumindest nicht abgehängt zu werden, sind die Erforschung, Entwicklung und Anwendung von Künstlicher Intelligenz (KI) bzw. Artificial Intelligence (AI), als ausgemachter Schlüsseltechnologie des 21. Jahrhunderts, längst zu einer fundamentalen Säule moderner Industrienationen evolviert. Die deutsche Bundesregierung bzw. die Europäische Union zielen mit der Etablierung einer vertrauenswürdigen Künstlichen Intelligenz (Trustworthy AI) Made in Germany bzw. Made in Europe darauf ab, ein charakteristisch-europäisches Alleinstellungsmerkmal zu realisieren. Infolge erhofft man sich, gegenüber der Konkurrenz aus den USA und China einen entscheidenden Wettbewerbsvorteil zu generieren und kraft dessen die internationale Wettbewerbsfähigkeit der deutschen bzw. europäischen Wirtschaft zu unterstützen. Zudem ist hiermit die Absicht verbunden, europäische Wertmaßstäbe auf die internationale Ebene zu transferieren. Ziel des vorliegenden Beitrags ist es zu untersuchen, ob und inwieweit die Entwicklung einer Trustworthy AI Made in Germany/Europe tatsächlich als fundiertes Distinktionsmerkmal fungiert. Daneben wird den Fragen nachgegangen, welche Implikationen die Konstituierung einer vertrauenswürdigen KI mit sich bringt und welche Faktoren besonders berücksichtigt und ggf. darüber hinaus in Betracht gezogen werden müssen, um die kolportierten Wirkpotenziale einer vertrauenswürdigen KI bestmöglich ausschöpfen zu können. Hierbei gehen die Autoren des Beitrags davon aus, dass im deutsch-europäischen Kontext insbesondere ein bislang weitgehend vernachlässigtes Zusammenspiel aus industrieller und vertrauenswürdiger KI-Orientierung eine unentbehrliche und für die Förderung und Steigerung des KI-Nutzenpotenzials nachhaltige Verbindung stiften kann.

Die zugrundeliegenden Informationen basieren auf der Durchsicht einer großen Zahl von „grauen“ Dokumenten, Preprints, politischen Verlautbarungen, Websites und Fachpublikationen aus dem nationalen und internationalen Kontext. Empirische Basis ist weiterhin eine Reinterpretation vorliegender eigener Forschungsergebnisse über den gesellschaftlichen Digitalisierungsprozess der letzten Jahre sowie die Ergebnisse von 16 Interviews mit KI-Expertinnen und Experten aus Wirtschaft und Wissenschaft. Diese wurden zwischen Oktober 2021 und März 2022 im Rahmen einer Expertise zum Thema „Dynamik der Künstlichen Intelligenz“, die im Kontext des BMBF-geförderten Projektes KI.Me.Ge. am ISF München durchgeführt wird, unter der Federführung von Hartmut Hirsch-Kreinsen realisiert. Kürzel für Interviewpartner*innen: W = (KI-)Wissenschaftler*in / E = (KI-)Experte*in.

Abstract

In order to prevail in the global race for technological leadership or at least not to be left behind, the research, development and application of Artificial Intelligence (AI), as a key technology of the 21st century, has long since evolved into a fundamental pillar of modern industrial nations. With the establishment of Trustworthy AI Made in Germany and Made in Europe,

the German Government and the European Union are aiming to realize a characteristic European unique selling proposition. As a result, it is hoped to generate a decisive competitive advantage over the competition from the USA and China and, by virtue of this, to support the international competitiveness of the German and European economy. In addition, the intention is to transfer European value standards to the international level. The aim of this paper is to investigate whether and to what extent the development of a Trustworthy AI Made in Germany/Europe actually functions as a well-founded unique feature. In addition, questions are addressed as to which implications the constitution of a Trustworthy AI entails. Furthermore, it is examined, which factors must be particularly considered and, if necessary, additionally be integrated, in order to be able to exploit the colported impact potentials of a trustworthy AI in the best possible way. The authors of the article assume that especially in the German-European context an interplay between Industrial and Trustworthy AI orientation, which has been largely neglected so far, can create an indispensable and sustainable connection for the promotion and increase of the AI benefit potential.

The underlying information is based on a review of a large number of “gray” literature, preprints, political announcements, websites and specialist publications from the national and international context. Furthermore, the empirical basis is a reinterpretation of existing own research results on the societal digitization process of the last years as well as the results of 16 interviews with AI experts from business and science. These interviews were conducted between October 2021 and March 2022 as part of an expertise on the topic of “Dynamics of Artificial Intelligence”, which is being conducted in the context of the BMBF-funded project KI.Me.Ge. at ISF Munich, under the leadership of Hartmut Hirsch-Kreinsen. Abbreviations for interview partners: W = (AI) scientist / E = (AI) expert.

Schlüsselbegriffe

Künstliche Intelligenz · Vertrauenswürdige KI · Technologiesouveränität · Distinktionsmerkmal · Innovationspolitik · Schlüsseltechnologie

Keywords

Artificial Intelligence · Trustworthy AI · Technological Sovereignty · Unique Feature · Innovation Policy · Key Technology · Bottleneck Technology

Inhalt

1. Zum Label Made in Germany

Dass das in der Gegenwart noch immer für eine hohe Qualität und hohe Verarbeitungsstandards stehende Gütesiegel „Made in Germany“ (vgl. Bundesverband Deutscher Patentanwälte 2022) ursprünglich der Feder des britischen Parlaments entstammt, wissen heutzutage vermutlich nur die wenigsten Menschen. Dass die originäre Intention des Labels dabei (aus deutscher Perspektive) keineswegs wohlwollender Natur gewesen ist, sondern im Gegenteil, britische Verbraucher*innen mithilfe der Herkunftsangabe ein Warnsiegel erhielten, welches vor dem Erwerb mutmaßlich minderwertiger Waren aus dem Wilhelminischen Kaiserreich warnen sollte (vgl. Lutteroth 2012), ist womöglich noch umso weniger Personen bekannt. Dieser Intention zuwiderlaufend avancierte das Gütesiegel in der Folgezeit bekanntermaßen zu einem Qualitätsindikator, der das international wirkende Renommee des Industriestandorts Deutschland mit Qualitätsmerkmalen wie Zuverlässigkeit, Innovation und Nachhaltigkeit (vgl. Grillo 2015: 1) eindrucksvoll stärkte. Welch enorme Schlagkraft dem Label noch immer inhärent ist, verdeutlicht unter anderem der im Jahr 2017 von Statista in Zusammenarbeit mit dem Marktforschungsunternehmen Dalia Research entwickelte „Made-In-Country-Index“. Jener wird als Vergleichswert für die Markenstärke von Herkunftsländern herangezogen und listet Deutschland unter Berücksichtigung einer Konsumenten*innenbefragung von 43.034 Personen aus 52 unterschiedlichen Ländern auf Platz 1 der beliebtesten Label weltweit auf (vgl. Manager Magazin 2017; vgl. Statista 2017). Als bemerkenswert erweist sich hierbei, dass neben der Schweiz auf Platz 2 das Label „Made in Europe“ dicht gefolgt auf Rang 3 folgt. Bemerkenswert deshalb, da die Europäische Kommission bereits zwischen den Jahren 2003 und 2004 mit dem Versuch, ein einheitliches Label für die gemeinschaftliche Herkunftsbezeichnung von Waren aus der Europäischen Union einzuführen, unter dem Vetorecht eines Großteils der EU-Mitgliedstaaten gescheitert ist (vgl. Verbraucherportal Deutschland 2017). Besonders Stimmen aus der deutschen Industrie wie der Bundesverband der Deutschen Industrie e. V. (BDI) oder die deutsche Industrie- und Handelskammer (DIHK) traten als meinungsstarke Kritiker gegenüber der EU-Pläne auf. Grund hierfür war die Sorge, an Strahlkraft und damit einhergehendem wirtschaftlichen Nutzen infolge einer Verwässerung des Made in Germany-Labels einzubüßen (vgl. Ludwig 2015: 2).

Dessen ungeachtet existieren durchaus gute Gründe dafür, das Label Made in Germany zwar als Qualitätssiegel, indes eben nicht als Herkunftssiegel anzuerkennen, denn: So kann durchaus die Frage aufgeworfen werden, ob ein zum Großteil aus unterschiedlichen Bestandteilen aus differierenden Ländern stammendes und zumindest schon zum Teil in Fernost montiertes Endprodukt (vgl. ebd.: 1) tatsächlich noch als originär „deutsches“ Produkt betrachtet werden kann. Aus diesem Grund ist die EU trotz zuletzt mehrerer gescheiterter Versuche noch immer darin bestrebt, für industriell gefertigte Produkte aus EU-Mitgliedstaaten ein Gütesiegel als einheitlich-verpflichtendes Signum zur nachhaltigen Steigerung der Wertschöpfung im EU-Raum zu etablieren (vgl. Deighton 2017).

2. Versprechungen, Erwartungen und Ziele

2.1 KI Made in Germany/Europe

Eine intensivere Auseinandersetzung mit der vom EU-Parlament avisierten Einführung einer allgemein verpflichtenden „Made in EU"-Kennzeichnung innerhalb des europäischen Binnenmarktes sowie den daraus etwaig entstehenden Folgen für althegebrachte Qualitätslabel wie Made in Germany ist zwar nicht Gegenstand des vorliegenden Beitrags. Gleichwohl verdeutlicht das illustrierte Spannungsgefüge eine Grundproblematik, die sich gegenwärtig im Rahmen eines thematisch ähnlich gelagerten Diskurses widerspiegelt. So wird das Label Made in Germany seit einiger Zeit im Bereich der KI-basierten Technologieentwicklung seitens der deutschen Bundesregierung (sowie analog der europäischen Union) als Diktum und Vorzeigeprojekt veranschlagt. Dessen Nutzung obliegt die Einführung einer weltweit ersten KI-Strategie zur Etablierung internationaler Standards und Normen für die Implementierung und Anwendung von KI-Systemen (vgl. PLS 2020: 11) – der „*KI-Normungs-Roadmap für Künstliche Intelligenz*". Die Hauptintention offenbart sich darin, einen Handlungsrahmen für die Normung und Standardisierung von KI voranzutreiben, der die internationale Wettbewerbsfähigkeit der deutschen bzw. europäischen Wirtschaft unterstützt und europäische Wertmaßstäbe auf die internationale Ebene hebt (vgl. Wahlster & Winterhalter 2020). Unter dem Stichwort „KI (auch: *AI*[1]) Made in Germany/Europe" sollen national sowie international (und hier zuvorderst europäische Staaten) anschlussfähige Kriterien und Anforderungen an die Entwicklung und den Einsatz von Künstlicher Intelligenz in den verschiedensten Anwendungsbereichen entwickelt werden, um hierauf basierend das Vertrauen der Anwender*innen und Nutzer*innen in KI zu stärken.

2.2 Grundbestreben

Auf Grundlage dessen erhofft man sich, eine Sicherstellung des gesellschaftlichen Nutzenpotenzials sowie eine gemeinwohlorientierte Ausschöpfung der durch die Nutzung von KI bereitgestellten Potenziale zu ermöglichen (vgl. PLS 2020; vgl. Die Bundesregierung 2018b). Ziel der Entwicklung spezieller KI-Zertifizierungen von KI-Systemen ist es, besonders den wirtschaftlichen und wissenschaftlichen Bereich im internationalen KI-Wettbewerb zu stärken und darüber hinaus innovationsfreundliche Rahmenbedingungen für eine nachhaltige und zukunftssichere Wertschöpfung (vgl. BMWi 2020) im Bereich der KI-Technologie zu schaffen. Die primäre Intention der KI-Zertifizierungen lässt sich wie folgt zusammenfassen: „Eine gelungene Zertifizierung ermöglicht die Erfüllung wichtiger gesellschaftlicher und ökonomischer Prinzipien – wie etwa Rechtssicherheit (z. B. Haftung und Entschädigung), Interoperabilität, IT-Sicherheit oder Datenschutz. Zudem kann Zertifizierung Vertrauen schaffen, zu besseren Produkten führen und die nationale und internationale Marktdynamik beeinflussen" (PLS 2020: 10).

Um jene Ambition bestmöglich zu verwirklichen, setzen die Verantwortlichen auf die Aufdeckung einer vermeintlichen Leerstelle innerhalb der mit der weltweit stattfindenden digitalen Transformation einhergehenden Umwälzungsprozesse. Diese manifestiert sich in der

[1] Die Termini *KI* (Künstliche Intelligenz) und *AI* (Artificial Intelligence) werden in Folgenden synonym verwandt. In der zugrundeliegenden Forschungsliteratur erweist sich die Verwendung deutsch-englischer Vermischungen für etwaige Schlagwörter als gängige Praxis.

Entwicklung einer nachvollziehbaren, nachprüfbaren, erklärbaren und vor allem *vertrauenswürdigen KI*. Seitens der Europäischen Union sowie nicht wenigen Vertretern*innen aus Industrie, Wirtschaft und Politik wird die ausgemachte Leerstelle als „Distinktionsmerkmal und Wettbewerbsvorteil europäischer Lösungen" (ebd.: 9) hochstilisiert. „The EU's focus on responsible and safe AI early on could give it an edge when it comes to setting ethical and regulatory standards. The EU's approach would not regulate the technology as such, nor its applications, but rather would help guide how the applications are developed and deployed. By reporting, monitoring, and analyzing the progress of AI, the EU could position itself to define quality, build alliances with like-minded partners, and lead multilateral initiatives" (Brattberg et al. 2020). Exemplarisch wird im EU-Förderprogramm *Digital Europe* eine Summe von 9,2 Mrd. Euro bereitgestellt, kraft derer die digitale Transformation der europäischen Wirtschaft und Gesellschaft angestoßen werden soll. Eines der Hauptangriffsziele wird im *Coordinated Plan on Artificial Intelligence* festgelegt und betrifft den Einsatz einer Gruppe von Expertinnen und Experten für KI (High-Level Expert Group, HLEG) mit dem Auftrag, Ethik-Leitlinien (im Orig.: *Ethics Guidelines for Trustworthy AI*) für den Einsatz von Künstlicher Intelligenz zu entwickeln. Die Konzipierung eines menschenzentrierten Ansatzes unter Rückgriff auf Ethikaspekte in der Normung und Standardisierung für KI – *ethics by Design* (EbD) –, „an approach to design that aims at the systematic inclusion of ethical values, principles, requirements and procedures into design and development processes" (SIENNA 2021), wird als tragende Säule beim Versuch, KI-„Schlüsseltechnologien und technologiebasierte Innovationen in Europa eigenständig zu entwickeln und hierfür eigene Produktionskapazitäten aufzubauen" (BMBF 2021b: 3), verstanden.

2.3 Vertrauenswürdige Künstliche Intelligenz: Distinktionsmerkmal

Aus deutscher Sicht wird mit dem Konzept einer ethischen KI Made in Germany die Erwartung verknüpft, das alteingesessene Leitmotiv Made in Germany in die Leittechnologie des 21. Jahrhunderts zu übertragen und hierdurch ein eigenständiges Merkmal deutscher Provenienz zu konsolidieren: „Ich glaube, dass wenn es gut läuft, das für mich so etwas wie eine ‚deutsche' KI im Kern definieren könnte" (W5, 05. November 2021). Der auch als *europäische Weg* (PLS 2020: 9) bezeichnete Vorgang der Etablierung einer vertrauenswürdigen Künstlichen Intelligenz setzt sich vor allem das Ziel, ein charakteristisch-europäisches Alleinstellungsmerkmal hauptsächlich gegenüber der Konkurrenz aus den USA und China im weltweiten Wettrennen um die Technologieführung herbeizuführen. Auf diese Weise soll ein Wettbewerbsvorteil generiert werden. Im bestmöglichen Fall soll dieser infolge einer erfolgreichen Bewährungsphase in den heimischen europäischen Gefilden zu einem „*Exportschlager*" (W11, 30. März 2022) formatiert werden.

In diesem Zusammenhang wird auf Ebene der EU in Anlehnung an das deutsche Narrativ das Label „KI Made in Europe/EU"[2] bemüht, um für eine „verantwortungsvolle und gemeinwohlorientierte Entwicklung und Anwendung von KI-Systemen" (BMBF 2021b: 13) zu werben.

[2] Eine eindeutige Trennung zwischen Mitgliedstaaten der EU sowie EU-unabhängigen Staaten wird im Kontext des Themas von vertrauenswürdiger/ethischer KI literaturübergreifend oftmals vernachlässigt, sodass der Eindruck entsteht, die seitens der EU-Kommission veranschlagte KI-Agenda entspreche beispielsweise der seitens der britischen Regierung avisierten Strategie zur Förderung des Einsatzes von Künstlicher Intelligenz. Dass eine Gleichsetzung von „europäischer KI" und „KI Made in

Integratives Element einer solchen *humanzentrierten* (im Orig.: „*human-centric approach on AI*"; European Commission 2019a: 9) „*Trusted KI* (auch: *Trustworthy AI*)" (Kersting & Tresp 2019: 13) ist die Orientierung an ethischen Richt- und Leitlinien[3], die im Rahmen der Entwicklung und Verwendung von KI Berücksichtigung erfahren. "The ethical dimension of AI is not a luxury feature or an add-on: it needs to be an integral part of AI development. By striving towards human-centric AI based on trust, we safeguard the respect for our core societal values and carve out a distinctive trademark for Europe and its industry as a leader in cutting-edge AI that can be trusted throughout the world" (European Commission 2019a: 9). Dabei impliziert das anberaumte Distinktionsmerkmal „automatisierte Erkennungs-, Vorschlags- und Entscheidungssysteme, die den Anforderungen von Transparenz, Verantwortlichkeit, Privatheit, Diskriminierungsfreiheit und Zuverlässigkeit gerecht werden" (Beckert 2021: 17). In Hinblick auf den erhofften wirtschaftlichen Nutzenvorteil des Distinktionsmerkmals stellt ein interviewter KI-Wissenschaftler relativ direkt heraus: „Vom Prinzip her ist es eine Policy und mit der kann man sich, wenn man sie clever anwendet, Zeit erkaufen… . Die Markteintrittskosten werden hoch für Player von außen" (W7I, 15. Februar 2022).

2.4 Vertrauenswürdige Künstliche Intelligenz: Technologiesouveränität

Neben der Steigerung der Wettbewerbsfähigkeit intendiert KI Made in Germany bzw. KI Made in Europe[4] die Steigerung der digitalen Souveränität. Man erhofft sich vor allem eine Reduktion struktureller Abhängigkeiten von Schlüsseltechnologien aus anderen, technologisch dominanten Staaten zu erreichen, um die eigene wirtschaftliche und technologische Handlungsfähigkeit sicherzustellen. Dieser Umstand wird unter dem Schlagwort *Technologi-*

---

the EU" mitunter unzutreffend ist, zeigt u. a. die *National Artificial Intelligence Strategy* des Vereinigten Königreichs, welche in einigen Aspekten erwähnenswerte Unterschiede und differierende Ansichten gegenüber dem *Coordinated Plan on Artificial Intelligence* der EU-Kommission bereithält (vgl. Pierides & Roxon 2021). Wird im Folgenden keine explizite Erläuterung betreffs etwaiger Differenzen zwischen europäischen Staaten angegeben, kann der jeweils verwandte Terminus pars pro toto für Europa bzw. die EU angesehen werden.

[3] Sowohl in der Fachliteratur als auch im öffentlichen medialen Diskurs wird das Schlagwort „ethische KI" häufig synonym für den Gegenstandsbereich um die Standardisierung und Zertifizierung von „vertrauenswürdiger KI" gebraucht. Hierbei gilt es zu beachten, dass ethische Leitlinien zwar ein essenzielles Element innerhalb des Lebenszyklus eines KI-Systems darstellen, um die Vertrauenswürdigkeit von KI-Systemen zu konsolidieren. Gleichwohl obliegen der Rahmenstruktur einer vertrauenswürdigen KI – vor dem Hintergrund der durch die HLEG der Europäischen Kommission entwickelten Leitlinien – neben der ethischen Komponente noch die beiden Komponenten der „Robustheit" und der „Rechtmäßigkeit". *Robustheit* meint die sichere und zuverlässige Funktionserfüllung eines KI-Systems sowie die Widerstandsfähigkeit gegen Angriffe und Sicherheitsverletzungen und die allgemeine Sicherheit, Präzision, Zuverlässigkeit und Reproduzierbarkeit der Technik. *Rechtmäßigkeit* meint die Einhaltung geltenden Rechts und gesetzlicher Bestimmungen. Jede dieser drei Komponenten ist notwendig, jedoch allein nicht ausreichend, um das Ziel einer vertrauenswürdigen KI zu erreichen (vgl. Europäische Kommission 2018).

[4] Im Folgenden wird der Terminus „KI Made in Europe" als handlungsleitendes Motiv verwandt. Dabei ist das Markenzeichen „KI Made in Germany" stets mit eingeschlossen, es sei denn, dass ein expliziter Unterschied zur EU-ausgerichteten perspektivischen Verortung vorliegt. U. a. Ringo Ossewaarde & Erdener Gülenç weisen auf die enge Verquickung zwischen deutscher und EU-ausgerichteter KI-Strategie hin, die oftmals kongruent ist: „The German government closely linking its AI strategy to the EU's" (Ossewaarde & Gülenç 2020: 57).

sche Souveränität zusammengefasst und meint den „Anspruch und die Fähigkeit zur kooperativen (Mit-)Gestaltung von Schlüsseltechnologien und technologiebasierten Innovationen. Dies umfasst die Fähigkeiten, Anforderungen an Technologien, Produkte und Dienstleistungen entsprechend der eigenen Werte zu formulieren, Schlüsseltechnologien entsprechend dieser Anforderungen (weiter) zu entwickeln und herzustellen sowie Standards auf den globalen Märkten mitzubestimmen" (BMBF 2021b: 3). Maßgebend erscheinen hier folglich die Fragen, wer letztlich die Kontrolle über Datenbestände und deren Verwertung ausübt; wie IT-Sicherheit generell gewährleistet werden kann; wer die Standards setzt und inwieweit und inwiefern eine Realisierung von Eigenständigkeit im Zuge der Entwicklung von Methodiken und Instrumenten möglich erscheint – aus Sicht vieler Expertinnen und Experten ein nicht immer eindeutig zu beantwortender Sachverhalt. Absicht ist mithin, eine selbstbestimmte Ausweitung der Gestaltungs- und Entscheidungsfreiheit im Rahmen der digitalen Transformation und bei der Wahl einer Technologie zu ermöglichen. Eine weitere Zielsetzung ist, die Akzeptanz und den Umgang europäischer Bürger*innen mit digitalen Technologien zu erhöhen und jene zu souveränem Handeln zu befähigen. Vor allem die Entwicklung von gesellschaftsübergreifendem Vertrauen gegenüber dem Einsatz von KI steht hier im Mittelpunkt (vgl. PLS 2020; vgl. Die Bundesregierung 2018b).

3. Akteure

In Anlehnung an das zumeist noch immer im industriellen Kontext verortete Narrativ deutscher Wertarbeit versucht die deutsche Bundesregierung, die Marke KI Made in Germany als *Zugpferd* und *Aushängeschild* aufzubauen. Handlungsleitend ist hierbei die Orientierung an einer qualitativ hochwertigen und vor allem den Grundwerten der freiheitlich-demokratischen Grundordnung der Bundesrepublik Deutschland entsprechenden Anwendung und Nutzung von KI-Technologien. So avisiert die Bundesregierung schon in ihrer *Strategie Künstliche Intelligenz*[5], dass Deutschland sowohl in der Forschung als auch in der Anwendung von KI unter dem Gütesiegel „Artificial Intelligence (AI) made in Germany" zum weltweit führenden KI-Standort werden soll (vgl. Die Bundesregierung 2018b: 8–38). Die Konzeptionierung spezifischer nationaler (bzw. europäischer) Kriterien und Anforderungen an die Entwicklung und den Einsatz von KI rührt zuvorderst aus unkalkulierbaren Abhängigkeiten deutscher/europäischer Hochschulen, Forschungsinstitute und Firmen von (zumeist) US-amerikanischen Internetfirmen wie Google oder Amazon und deren via Open Source-Struktur zur Verfügung gestellter Algorithmen und Software-Pakete (vgl. Kersting & Tresp 2019: 14). So nutzen deutsche und europäische Akteure aktuell in hohem Ausmaß Cloud-Infrastruktur-Komponenten

[5] Am 15. November 2018 hat die Bundesregierung ihre *Strategie Künstliche Intelligenz* verabschiedet, die als Rahmen einer ganzheitlichen politischen Gestaltung der weiteren Entwicklung und Anwendung von KI in Deutschland fungiert. Ziel ist es, Deutschland als Forschungsstandort für Künstliche Intelligenz zu stärken und die Förderung der Anwendung von KI in der Wirtschaft und hier insbesondere in KMU's voranzutreiben. Online: https://www.bundesregierung.de/breg-de/service/publikationen/strategie-kuenstliche-intelligenz-der-bundesregierung-2018-1551264 [08. Mai 2022].

amerikanischer IT-Konzerne (Infrastructure-as-a-Service), die einerseits starke Lock-In-Effekte[6] hervorrufen und die Autonomie der Anwender*innen und Nutzer*innen einzuschränken drohen.

Andererseits wird demgemäß sowohl die Daten- als auch die Technologiesouveränität bei Soft- und Hardwarekomponenten entlang von KI-Wertschöpfungsketten (vgl. Die Bundesregierung 2018b: 15) tangiert, sodass die allgemeine digitale und technologische Souveränität im gesamteuropäischen technologisch-digitalen KI-Kontext durch externe Abhängigkeiten beschnitten und unterwandert wird. Nicht zuletzt deshalb ist es aus Sicht europäischer Staaten unabdingbar, Grundlagen für die Nutzung von autonomen „KI-Infrastrukturen, in denen Rechenleistung, Speicher und Werkzeuge bereitgestellt werden" (Kersting & Tresp 2019: 14), zu schaffen. Ein Beispiel hierfür stellt die jüngst bekannt gegebene Partnerschaft zwischen dem US-amerikanischen Halbleiterhersteller Intel und der Landeshauptstadt Magdeburg dar, welche die heimische Halbleiterindustrie mit einem Gesamtvolumen von 33 Milliarden Euro stärken und infolge zu einer unabhängigeren Position gegenüber ausländischen Halbleiterherstellern führen soll (vgl. Finsterbusch et al. 2022).

Wirtschaftliche Beratungsunternehmen wie die Wirtschaftsprüfungsgesellschaft KPMG Deutschland versuchen bereits jetzt als *First Mover* (Pionierunternehmen) auf dem Markt der vertrauenswürdigen KI aktiv zu werden, um sich mit Blick auf den (vermeintlich) bevorstehenden Run auf gesellschaftsverträgliche, vertrauenswürdige KI-Algorithmen frühzeitig einen pekuniär aussichtsreichen Platz zu sichern. Die kostenpflichtige Bereitstellung eines *Ethik-Kompass für Daten und Künstliche Intelligenz (KI)* soll „Unternehmen, die an der Integration von KI-Ethik in ihre Governance interessiert sind [...] [sowie] Unternehmen, die bereits eine Ethik-Governance aufgebaut haben und den Reifegrad ihrer KI-Ethik bemessen und erhöhen wollen" (KPMG 2022), unterstützen. Ein suggestiver Verweis auf die Möglichkeit einer bevorstehenden Emergenz humanoider Roboter wie im dystopisch konnotierten Hollywoodfilm Blade Runner tut hier im Werbetext sein Übriges, um potenzielle Akteure im KI-Umfeld für die Relevanz und Implementierung eines ethischen Rahmens mit regulatorischer Zielsetzung (vgl. ebd.) im eigenen Unternehmenskontext zu sensibilisieren.

Derweilen fungiert die vom Bundesministerium für Bildung und Forschung (BMBF) sowie von der Akademie für Technikwissenschaften (acatech) initiierte und koordinierte Plattform Lernende Systeme (PLS) ebenfalls als eines der zentralen politischen Instrumente im deutsch-europäischen KI-Diskurs. Mit Beteiligung von Expertinnen und Experten aus Wirtschaft, Politik und zivilgesellschaftlichen Organisationen aus den Bereichen Lernende Systeme und Künstliche Intelligenz ist PLS vor allem bestrebt, die Umsetzung der KI-Strategie der Bundesregierung (siehe *Hightech-Strategie 2025*, in: Die Bundesregierung 2018a) zu realisieren. Obschon zahlreiche heimische KI-Wissenschaftler*innen ebenfalls im Umfeld um PLS aktiv sind und eine solche Herangehensweise durchaus als lohnenswert ansehen, wird das Thema einer ethischen KI teilweise differenziert und kritisch gesehen. So attestiert unter anderem Thomas Metzinger, Philosoph und ehemaliges Mitglied der High-Level Expert Group on Artificial Intelligence, den von der EU-Kommission veröffentlichten „oberflächliche[n]

[6] *Lock-In-Effekt*: Unter dem Lock-in-Effekt versteht man in der Betriebswirtschaft die faktische Gebundenheit von Kunden*innen an einen Anbieter, die konsumenten*innenseitig als negativ empfundene Zwangsbindung definiert wird. Der Wechsel der Kunden*innen zu einem anderen Anbieter ist hierbei mit hohen Wechselkosten verbunden und somit unwirtschaftlich (vgl. Hages et al. 2017: 12).

Ethikkodizes“ (Jahn 2021) „Ethics Washing“, die ein seriöses und fundiertes Konzept vermissen lassen. Daher wird auch der Leitgedanke einer vertrauenswürdigen KI als bloße „Gute-Nacht-Geschichte für die Kunden von morgen“ (Metzinger 2019) eingestuft.

4. Evaluation

4.1 Vertrauenswürdige KI: Alleinstellungsmerkmal

Wie oben angesprochen, kann die Originalität deutscher Produkte, deren Komponenten eine internationale Herkunft aufweisen, in Frage gestellt werden. Das Label Made in Germany wird damit in seiner konstitutiven Qualität des Alleinstellungsmerkmals deutscher Provenienz relativiert. Eine vergleichbare Position lässt sich ebenso in Hinblick auf die Frage nach der Originalität einer ethischen KI Made in Europe aufwerfen. So lässt sich zunächst einmal der triviale, indes durchaus gewichtige Faktor herausstellen, dass von einer Eigentümlichkeit oder gar Einzigartigkeit einer auf Vertrauenswürdigkeit und Nachvollziehbarkeit gerichteten ethischen KI deutscher/europäischer Provenienz, wenn man diverse internationale KI-Strategiepapiere in den Blick nimmt, kaum die Rede sein kann.

Bereits in dem von Kanada im Jahr 2017 veröffentlichten und als weltweit erstem offiziellen KI-Strategiepapier geltenden Regierungspapier, „Pan Canadian AI Strategy“, wird die Entwicklung von ethischen Implikationen für Künstliche Intelligenz als eines der Hauptanliegen herausgestellt. 2018 veröffentlichte Indien seine „National Strategy on Artifial Intelligence: #AIforALL“ und hebt hierin unter anderem die Forschungsrelevanz bezüglich datenschutzrechtlicher, Privatsphäre schützender und ethischer, Bias-unterbindender KI-Bemühungen hervor. Brasiliens „Brazilian Artifical Intelligence Strategy“ aus dem Jahr 2020 sowie der im Jahr 2019 durch die australische Commonwealth Scientific and Industrial Research Organisation (CSIRO) publizierte „AI Ethics Framework“ befassen sich ebenfalls ausführlich mit der Bedeutung und Notwendigkeit von regulatorischen und ethischen Rahmenbedingungen und Orientierungsleitlinien im Zuge der Erforschung, Entwicklung und Verbreitung von KI-Produkten und -anwendungen (Zhang et al. 2021: 155–163). Aus westlicher Sicht zunächst erstaunlich anmutend und den (oftmals undifferenzierten) (vgl. Richter & Gebauer 2010) allgemeinpolitischen Einschätzungen zuwiderlaufend, wurde auch jüngst in China ein „White Paper on Trustworthy Artificial Intelligence“ (2021) durch die China Academy of Information and Communications Technology (CAICT) publiziert. Mit inhaltlich zentralen Elementen wie Vertrauenswürdigkeit, Nachvollziehbarkeit oder Fairness weist die chinesische KI-Strategie hinsichtlich vertrauenswürdiger KI deutliche Überschneidungen mit den adressierten Kerndimensionen der in den von der EU publizierten *Deliverables* (vgl. Kap. Perspektiven) auf. Eine Feststellung, die durch die Worte eines renommierten KI-Wissenschaftlers aus dem Sample unterstrichen wird. Der Interviewpartner konterkariert vor allem die in der westlichen Welt oftmals anzutreffende unterkomplexe, stereotypisierende und tendenziell negativfixierte China-Berichterstattung (vgl. Changbao 2021): „Im Moment ist China ganz vorne in der Bewahrung von Privatheit und Einschränkung der Macht der großen KI-Konzerne. Die sind gesetzlich ganz eng beschränkt worden in China und das wird auch streng durchgesetzt. Da ist China momentan auch gegenüber der Europäischen Union ein Vorreiter [...]. Trustworthy AI ist kein Alleinstellungsmerkmal der EU“ (W1, 12. Januar 2022). Eine Position, die durch die ebenfalls erst jüngst in China verabschiedeten Gesetze „The Personal Information Protection Law“

(2021) und „Data Security Law“ (2021) (vgl. Horwitz 2021) eine zusätzliche Untermauerung erfährt.

Vor diesem Hintergrund ist es ersichtlich, weshalb Jessica Newman, Programmleiterin für die KI-Sicherheitsinitiative (AISI) beim Zentrum für langfristige Cybersicherheit (CLTC) der UC Berkeley, deutsch-europäische Bemühungen, ethische, vertrauenswürdige und auf Standards basierende, reglementierte KI-Anwendungen als *hauseigene* Erfindung anzupreisen, als Verzerrung der Realität auffasst. Der Vorwurf, die EU inszeniere sich als *technologischer Wächter der Welt*, während die USA als *digitaler Westen* (vgl. Venkina 2021) dargestellt werde, ist unter Berücksichtigung der vorangegangenen Befunde tatsächlich nicht ganz von der Hand zu weisen. Zwar steht der zugrundeliegende Gegenstandsbereich, derart wie er in den EU-Deliverables behandelt wird, in Tiefe und Breite zuweilen weltweit tatsächlich keinem bekannten vergleichbaren Pendant gegenüber. Im Gegensatz dazu unterstützen Aussagen wie „Papiere sind geduldig“ (E5, 21. Februar 2022), in Bezug auf die Ernsthaftigkeit und Validität einer chinesischen Trustworthy AI, die These des technologischen Wächters jedoch eher, als ihr Wind aus den Segeln zu nehmen, denn: Selbiges Argument kann problemlos auch für die seitens der HLEG ausgearbeiteten Leitlinien angewandt werden. „The guidelines drafted by the AI high-level expert group are non-binding and as such do not create any new legal obligations” (European Commission 2019b: 3).

Darüber hinaus lässt sich anmerken, dass die europäischen Datenschutzrichtlinien und -gesetze gegenüber US-amerikanischen oder chinesischen zwar strenger sind und härtere Sanktionen androhen. Gleichwohl gilt dies nicht unbedingt für KI-basierte Technologien und im Zuge der Einführung des Kritikalitätsmodells (vgl. Kap. Ein problematisches Narrativ) bleibt ein „Großteil der KI-Nutzung [...] ungeregelt, und es werden lediglich freiwillige Leitlinien vorgeschlagen, um einen verantwortungsvollen Einsatz zu fördern” (vgl. Venkina 2021). Die Sinnhaftigkeit und Intention von KI-Standards wird durch diesen Umstand sicherlich in keiner Weise gemindert. Dennoch wirkt es im Hinblick auf die zugrundeliegenden Erkenntnisse disputabel, ob vertrauenswürdige und ethische KI-Ansätze als ein ausschließliches und originäres Produkt europäischer Provenienz angesehen werden können und den beiden hauptkonkurrierenden KI-Konzeptualisierungen und deren attestierter Hauptintention – China („Kontrolle“) und USA („Kommerz“) (Beckert 2021: 17) – ein Interesse an vertrauenswürdiger und auf ethischen Werten beruhender KI-Technologie abgesprochen werden kann.

4.2 Vertrauenswürdige KI: Innovationshemmnis?

Ungeachtet der Beanspruchung eines Alleinstellungsmerkmals ist gerade im Rahmen der Einführung der EU-Datenschutz-Grundverordnung (DSGVO)[7] immer wieder die Rede vom Innovationshemmnis eines zu stark regulierenden KI-Datenschutzrechts gewesen. Die Forderung des ehemaligen Präsidenten des Bundesverbandes der Deutschen Industrie, Dieter Kempf: „Keinesfalls darf Datenschutzrecht zum Innovationshemmnis und Standortnachteil werden“ (BDI 2018), ist in diesem Kontext sicherlich bedenkenswert. Dies gilt zumindest

[7] „Die Verordnung (EU) 2016/679 (EU-Datenschutz-Grundverordnung) löst die Europäische Datenschutzrichtlinie aus dem Jahr 1995 (RL 95/46/EG) mit dem Ziel der Harmonisierung und Modernisierung des europäischen Datenschutzrechts ab. Sie fördert den Schutz der Betroffenen bei der Verarbeitung personenbezogener Daten und den freien Verkehr solcher Daten (Artikel 1 Absatz 1 Datenschutz-Grundverordnung)“ (BMI 2021).

dann, wenn es um den Gegenstandsbereich vertrauenswürdiger, reglementierter KI-Technologien geht. Zweifelsohne vollführt Europa diesbezüglich einen Drahtseilakt, der maßgeblich durch zwei Entscheidungsoptionen gekennzeichnet ist, wie Pereira treffend herausstellt. Auf der einen Seite lässt sich die Frage aufwerfen, ob es wirklich der regulatorische Rahmen ist, der technologische Innovationen und somit auch KI-Innovationen hemmt und infolge die wirtschaftliche Wertschöpfung. Oder sind es nicht doch eben komplexe administrative Strukturmechanismen, die eine aussichtsreiche Skalierung von Wettbewerbsfähigkeit und Innovation limitieren. Auf der anderen Seite muss die Frage hinsichtlich der Positionierung zu den Aspekten Transparenz, Erklärbarkeit und Nachvollziehbarkeit bezüglich innovativer (KI-)Produkte in Hochrisikoszenarien in grundlegenden gesellschaftlichen Bereichen (Recht, Arbeit, Gesundheit und Versorgung etc.) gleichwohl einer intensiven Diskursanalyse unterzogen werden (vgl. Pereira 2021). Sicherlich ist es falsch zu behaupten, dass die Fokussierung auf ethische KI ausschließlich regulatorische Innovationshemmnisse, die eine Abschottung herbeiführen, fördert. Gleichermaßen falsch wäre es allerdings auch, durch zu generöse Deregulierungs- und Marktliberalisierungsmechanismen zu riskieren, das bereits bestehende Datenschutzniveau, die angestrebte technologische Souveränität sowie die Grundrechte und die Prinzipien der Rechts- und Sozialstaatlichkeit sowie das Demokratieprinzip für eine als „salvific power“ (Ossewaarde & Gülenç 2020: 53) hochstilisierte, mitunter politisch aufgeladene Heilsbringertechnologie zu opfern.

Fakt ist, dass die DSGVO auf multinationaler Ebene zweifelsfrei dazu geführt hat, dass einige Unternehmen, und hier auch solche außerhalb der EU wie Japan, Südkorea oder Israel, in ihrer datenschutzrechtlichen Grundausrichtung DSGVO-Bestimmungen aufgenommen und diese in ihre nationalen Rechtsvorschriften übersetzt und als verbindliche nationale Datenschutzbestimmungen zur Anwendbarkeit der DSGVO erlassen haben (vgl. Petrova 2019). Die seitens der EU geplante Einführung des „Artificial Intelligence Act“, dessen Ziel es ist, „den EU-Binnenmarkt durch verbindliche Regeln zum weltweiten Vorreiter für die Entwicklung sicherer, vertrauenswürdiger und innovativer KI-Systeme zu machen“ (Gröschel 2022), könnte in ähnliche Fußstapfen treten. Es ist ohne Frage vorstellbar, dass jener vergleichbare Auswirkungen auf Unternehmen außerhalb Europas und deren Nutzung oder Bereitstellung von KI-Systemen hat (vgl. Greenleaf 2021: 10). „Given the need to address the societal, ethical, and regulatory challenges posed by AI, the EU’s stated added value is in leveraging its robust regulatory and market power—the so-called “Brussels effect”[8]—into a competitive edge under the banner of “trustworthy AI” (Brattberg et al. 2020).

Zu berücksichtigen ist aber auch, dass etliche Staaten keine weiteren Schritte zur Harmonisierung der nationalen Rechtsvorschriften und der DSGVO unternommen haben (vgl. Petrova 2019) und Europa sich angesichts der starken Abhängigkeit von nicht-europäischen technologischen Infrastrukturherstellern (vgl. BMWi 2021) noch immer in einer eingeschränkten, asymmetrischen Verhandlungsposition befindet. Ein fundamentales und dem Hauptgegenstand der Debatte inhärentes Wesensmerkmal stellt die unauflösbare Einheit zwischen künstlicher Intelligenz und Datenfluss dar (vgl. Redman 2018): „Data is key to AI systems, since algorithms need to be trained how to operate by consuming and learning from data sets.

[8] „EU’s unilateral ability to regulate global markets by setting the standards in competition policy, environmental protection, food safety, the protection of privacy, or the regulation of hate speech in social media“ (Bradford 2021).

The size and quality of data sets on which an AI application has been trained therefore directly impact its real-world utility“ (Sahin & Barker 2021). Ein auch aus der Sicht europäischer Verantwortlicher handlungsleitendes Richtmaß, um die mit der Anwendung von KI-assoziierten Stärken und Vorteile zur Steigerung der Wettbewerbsfähigkeit zu nutzen. Eine zu weitgreifende Intensivierung der Datenschutzbestimmungen könnte der charakteristischen Funktionslogik von KI diametral entgegenlaufen, weshalb „[e]uropean lawmakers need to find ways to harness data for AI while preserving its strong track record and reputation on data protection“ (ebd.). In Bezug hierauf konstatiert eine KI-Expertin aus dem Interviewsample: „Wenn man versucht, Regulierungen abzuwehren, indem man sagt: ‚Das hilft ja gar nichts, wenn wir hier regulieren. Man muss immer global denken und da helfen uns Regulierungen nicht‘. Aus ethischer Perspektive ist so eine Argumentation erst einmal nicht zulässig, denn nur weil es faktische Probleme gibt, etwas nicht zu fordern, was normativ am besten wäre, wäre ethisch nicht zu begründen“ (E5, 21. Februar 2022). Eine stichhaltige, wenngleich auch in der Realität nur unter der Bedingung höchsten Feingefühls zu realisierende Position, die den europäischen Verantwortlichen zukünftig ein hohes Maß an kontextueller Ambidextrie bezüglich der Wahrung von Datenschutzrechten und der Nutzbarmachung von Daten für die KI-Entwicklung abverlangt.

4.3 Vertrauenswürdige KI: Ausrichtung und Umsetzung

Richtet man sein Augenmerk auf die verschiedenen KI-Strategien und -Programme der EU-Mitgliedstaaten, offenbart sich teilweise eine starke Divergenz hinsichtlich des Entwicklungsstatus, die vor allem auf die starke Fragmentierung des europäischen digitalen Binnenmarktes zurückzuführen ist. „Europe’s information technology assets are scattered across different countries. This is precisely why the establishment of the digital single market is considered so crucial for the EU’s global competitiveness“ (Brattberg et al. 2020). In der „Declaration of Cooperation on Artificial Intelligence“ haben sich die EU-Mitgliedstaaten im April 2018 darauf geeinigt, eine Zusammenarbeit im Hinblick auf die wichtigsten Fragen im Zusammenhang mit der Thematik der Künstlichen Intelligenz anzustreben. Hierzu gehören unter anderem die Sicherung der europäischen Wettbewerbsfähigkeit bei Forschung und Einsatz der KI sowie die Behandlung sozialer, wirtschaftlicher, ethischer und rechtlicher Fragen (vgl. European Commission 2018: 37). Ungeachtet dessen lässt sich vor allem zwischen den nordeuropäischen Staaten und denen im Süden und im Osten (bis auf ein paar wenigen Ausnahmen) eine signifikante Divergenz ausmachen (vgl. Brattberg et al. 2020).

Nicht selten wird das seitens der EU evozierte *europäische Mindset* vor diesem Hintergrund von Kritikern*innen versucht, als Illusion und Mythos (vgl. Bouchard 2016) zu entlarven. So wird den Themen „inequality, human rights, manipulation of information and disinformation, safeguards against massive surveillance, and the overall control and abuse of power“ (Brattberg et al. 2020) in den entsprechenden KI-Strategien tatsächlich nur eine untergeordnete Rolle gewährt. Zudem erfahren die einzelnen Themen zwischen den unterschiedlichen Staaten eine jeweils divergierende Relevanzzuschreibung. Gleichwohl scheint dies hinsichtlich der Einführung ethischer Leitlinien und einer auf Vertrauenswürdigkeit und Nachvollziehbarkeit beruhenden KI-Entwicklungsperspektive nicht der Fall zu sein. Diesbezüglich kann durchaus von einer gleichgerichteten verfolgten Zweck- und Zielsetzung europäischer Couleur gesprochen werden. Allerdings bezieht sich diese Erkenntnis auch wiederum nur auf das

bloße Vorhandensein einer Erwähnung des Gegenstandsbereichs in den jeweiligen KI-Strategien. Eine gezielte, akribische inhaltliche Auseinandersetzung im Kollektiv ist bislang nicht erfolgt. Dementsprechend weisen die europäischen Länder bislang keine einheitlich strukturierte und aneinander angepasste Rahmenvereinbarung auf, sodass zumeist eigene Vorstellungen und Anforderungen bezüglich der Formulierung, Ausgestaltung und Umsetzung von vertrauenswürdiger KI-Technologie in den landesspezifischen Ansätzen auftauchen. Dies hat zur Folge, dass europäische Verantwortliche zunächst einmal dazu aufgefordert sind, notwendige Infrastrukturen zu errichten, um die Interessen und Anstrengungen zu bündeln, nationale und europäische Strategien zu ergänzen und infolge global wirksame Impulse und globale Standards setzen zu können (vgl. BMBF 2021b; vgl. Brattberg et al. 2020).

Beim Blick auf die gegenwärtige KI-Landschaft Europas wird man vor allem gewahr, dass Deutschland sich das Kredo auferlegt hat, sowohl das Fundament als auch den Dreh- und Angelpunkt für europäische Bemühungen in der Auseinandersetzung mit einer vertrauenswürdigen, ethischen KI zu begründen. Umfangreiche Empfehlungen der Datenethikkommission, zahlreiche Veröffentlichungen der Plattform Lernende Systeme oder acatechs sowie vielfältige Initiativen und Zusammenschlüsse zur Steigerung der technologischen Souveränität und des Vertrauens der Bürger*innen in KI scheinen jenes Bild auch zu unterstreichen. Allen voran die deutsch-französische Initiative GAIA-X, mit dem Ziel, die Souveränität im Umgang mit Daten zu stärken, indem Menschen aus unterschiedlichen Wirtschaftszweigen und Wissenschaftsdisziplinen sowie Verwaltungen und Politik zusammengebracht werden und gemeinsam an der Zielerreichung arbeiten (vgl. BMBF 2021a), kann als ein zentrales Projekt der Initiativen hervorgehoben werden.

Interessanterweise wird im von der Europäischen Kommission im Jahr 2018 veröffentlichten Report „Artificial Intelligence – A European Perspective“ kein Bezug auf Deutschlands Bemühungen hergestellt, obschon die Initiierung eines ethischen Orientierungsrahmens für europäische KI-Technologien hier bereits als grundlegendes Ziel ausgegeben wird. Noch erstaunlicher erweist sich dann jedoch noch der Umstand, dass im aktuellen, rund 190 Seiten starken Gutachten der Experten*innenkommission Forschung und Innovation (EFI) aus dem Jahr 2022 die Begriffe „vertrauenswürdig“ und „ethisch“ im Zusammenhang mit KI an keiner Stelle Erwähnung finden (vgl. EFI 2022). Eine Feststellung, die zur berechtigten Kritik an der Glaubwürdigkeit und Nachhaltigkeit der kolportierten deutschen KI-Ausrichtung einlädt. So hat das Bundesministerium der Justiz und für Verbraucherschutz (BMJV) erst gegen Ende des vergangenen Jahres bekannt gegeben, mit dem „Zentrum für vertrauenswürdige Künstliche Intelligenz“ (ZVKI) ein Projekt zu avisieren und „Zivilgesellschaft, Wirtschaft und Wissenschaft in einem KI-Ökosystem zusammenzubringen – und so vertrauenswürdige Anforderungen und Eigenschaften von KI-Systemen zu definieren sowie darauf hinzuwirken, vertrauenswürdige KI-Systeme als Standard zu etablieren“ (BMJV 2021). Angesichts der offenkundig auseinanderklaffenden Lücke zwischen eigener Anspruchshaltung und (zumindest forschungs- und innovationsperspektivisch) nachgewiesener Realität erfahren kritische Stimmen, die dieser Art von KI-Aktivitäten nicht mehr als den Status eines „White Washing“ der KI und ihrer Anwendungen (vgl. z. B. Jansen & Cath 2021) attestieren, eine nicht von der Hand zu weisende Bekräftigung.

4.4 Vertrauenswürdige KI: Adoption

Ein weiterer, nicht zu vernachlässigender Aspekt im Zuge der Auseinandersetzung mit Trustworthy AI wird durch die gegenwärtig noch immer bestehende Diskrepanz aus Erwartungshaltung und Umsetzungsmöglichkeit hinsichtlich der Implementierung von erklärbarer, nachvollziehbarer und vertrauenswürdiger KI im wirtschaftsorientierten Anwendungsbereich hervorgerufen. Ein Interviewpartner erklärt, dass die derzeit im Bereich der erklärbaren KI zu realisierenden Methoden sich lediglich auf solcherlei beschränken würden, die den Data-Engineers und ML-Engineers hülfen nachzuvollziehen, warum eine etwaige Maschine krankt. Dies stelle allerdings keine Erklärbarkeit für den/die Endkunden*in dar, sondern sei lediglich eine Unterstützungshilfe für die Seite der Entwickler*innen. Der Seite der Endkunden*innen werde hiermit indes keine moralisch geartete Entscheidungshilfe mit an die Hand gegeben (W7II, 01. März 2022). Zudem weist auch er auf die Zwiespältigkeit hinsichtlich der Bemühungen hin, vertrauenswürdige KI in Form einer Datenschutz-Policy als Distinktionsmerkmal hervorzuheben. Die Ingenieurinnen und Ingenieure, die eigentlich neue Innovationen erbringen müssten, fingen an, für die Datenschutz-Policy zu arbeiten, wie im Zuge der DSGVO bereits wahrzunehmen sei. Infolge entstehe die Gefahr, der Volatilität der Märkte nicht gerecht zu werden. Zudem komme es zu einem Fokusverlust als Resultat einer Ausweitung der Geschäftsmodelle der Abmahnbriefe sowie bspw. Zertifizierungen als Geschäftsmodelle (vgl. ebd.).

Gleichwohl sind es vor allem die derzeit noch immer fehlenden Rahmenbedingungen im Bereich der KI, die häufig bereits auf vermeintlich banale Ursachen zurückzuführen sind. Ein Beispiel stellt das Fehlen eines einheitlichen Grundkonsenses bezüglich des Verständnisses über grundlegende inhaltliche Aspekte wie der Frage danach, was Künstliche Intelligenz in seiner grundlegenden Konstitution ist (und was nicht), dar. Als weiteres Beispiel kann die fehlende Trennschärfe im Unterschied zwischen KI-Problemen und Datenproblemen hervorgehoben werden, wie eine KI-Expertin anmerkt (vgl. E5, 21. Februar 2022). In Anbetracht dessen ist es kaum verwunderlich, dass die Frage danach, wie der gängige *Wald-und-Wiesenbetrieb* von nebenan die kolportierten Vorteile einer ethischen KI für sich nutzen kann – und dies gewiss auch immer vor dem Hintergrund eines ohnehin bestehenden Ressourcen- und Kompetenzmangels der KMU's –, wenn selbst unter den verantwortlichen Entscheidungsträgern*innen hinsichtlich der konstitutiven und fundamentalen Identitätsstrukturen der heilsbringenden Technologie bislang von Einigkeit und einem einheitlichen Grundverständnis keine Rede sein kann.

5. Impact, Reichweite und Grenzen

5.1 Ein politisch motivierter Nischendiskurs

Der Forschungsüberblick zeigt, dass eine Auseinandersetzung mit ethischen Richtlinien für die KI-Technologie bislang tatsächlich höchstens nur einen ephemeren Stellenwert erfährt. „Though a number of groups are producing a range of qualitative or normative outputs in the AI ethics domain, the field generally lacks benchmarks" (Zhang et al. 2021: 127). Zudem erhärten die vorangegangenen Befunde den Eindruck, die Erschaffung einer ethisch motivierten, vertrauenswürdigen KI sei ein politisch vorangetriebener Diskurs. Durchaus etwas mokant stellt einer der interviewten KI-Wissenschaftler diesbezüglich fest: „Tatsächlich wird

dieser Bereich insofern momentan mit Leben gefüllt, dass Leute in Brüssel sitzen und versuchen, diese Dinge in Gesetze zu gießen. Das halte ich jedoch nicht für besonders zielführend... . Da formulieren Leute Gesetze über KI-Systeme, die nicht wissen, was ein KI-System überhaupt ist (denn das ist sehr schwierig zu definieren). Wenn Juristen zusammensitzen, kommen die auf andere Ideen als es KI-ler tun würden. Mein Eindruck ist da ein wenig, dass das von den Juristen allein verhandelt wird, die Gesetzgebung, ohne dass der Fachbezug durch die KI-ler abgesichert wird" (W4, 25. November 2021).

Tatsächlich werden immer wieder Stimmen laut, die dem Diskurs um eine vertrauenswürdige KI Made in Europe nicht mehr als den Status einer „künstlich getriebene[n] Blase einer Wissenschaftscommunity" (W7I, 15. Februar 2022) attestieren. Obendrein wird die vermeintlich eigentümliche KI-Ausrichtung hin zu nachvollziehbaren, nachprüfbaren, erklärbaren und vertrauenswürdigen KI-Algorithmen als ein mehr oder minder einer politischen Programmatik obliegendes, „künstlich gemachtes Distinktionsmerkmal" (ebd.) abgetan. Ein Argument, welches mit Blick auf die vorgestellten Inhalte international verorteter KI-Regierungspapiere untermauert wird und die kolportierte Eigentümlichkeit, Spezifität sowie die tatsächliche gesellschaftliche und wirtschaftspolitische Schlagkraft einer ethischen KI zur Disposition stellt.

5.2 Umsetzungsdefizite

Dass eine intensivierte Beschäftigung mit ethisch orientierten Grundregeln für KI sowie die Steigerung von technologischer Souveränität demungeachtet mehr als einleuchtend und wünschenswert erscheint, verdeutlicht nicht zuletzt die medial aufkeimende Flut an Berichten über Bewerbungs-Bots, die Biases im Gender- oder Racialkontext aufweisen. Ebenso, wenngleich nicht noch stärker, medial präsent präsentieren sich durch Whistleblower*innen wie Frances Haugen hervorgebrachte Enthüllungen, die belastende Beweise über den Zusammenhang zwischen KI-programmierter Algorithmen von Social Media-Plattformen und deren negativen Einfluss auf die geistige Gesundheit jugendlicher Nutzer*innen (vgl. Raidl 2021) liefern. Unter Rückgriff der von Beckert erhobenen Befunde lassen sich verschiedene Gründe für ein bestehendes Umsetzungsdefizit bzw. einen Mangel an Umsetzungsprojekten, die sich mit der Umsetzung von vertrauenswürdiger KI beschäftigen, identifizieren. Zunächst wäre dort das vorherrschende Kredo des First Mover. Obschon der gesellschaftliche Nutzen einer Berücksichtigung von Aspekten wie Vielfalt, Nichtdiskriminierung und Fairness beim Einsatz von KI quasi offenkundig ist, überwiegt der Status als Pionierprodukt auf dem Markt oftmals der Missachtung ethischer Kriterien (vgl. Beckert 2021: 20). Ein prominentes Beispiel markiert die US-Software HireVue mittels derer KI-gestützten Technik Konzerne wie Unilever oder Banken wie Goldman Sachs und JP Morgan Bewerber*innen hinsichtlich algorithmisch festgelegter Merkmale analysieren, um so den bzw. die bestmögliche Kandidaten*in zu finden. Zahlreiche Studienergebnisse dokumentieren Probleme im Kontext der Berufung auf das KI-gesteuerte Bewerbungsverfahren. Sie beruhen zumeist auf Verzerrungen im Algorithmus wie racial bias und gender bias und führen nicht selten zur Nichtbeachtung, Isolation und Ausgrenzung vulnerabler und kultureller Minoritäten im Auswahlprozess. Ungeachtet der nachweislich diskriminierenden Wirkung und manipulationsanfälligen Struktur haben sich die Wegbereiterunternehmen dieser Art von KI-Software und -Lösungen, aller Kritik zum Trotz, einen signifikanten Marktvorteil sichern können (vgl. Beckert 2021; vgl. Gupta et al.

2021; vgl. Noll 2019; vgl. Butcher 2016). Daneben sind die Relevanz sowie der logische Zusammenhang zwischen vertrauenswürdiger KI und konkreter Anwendung, sei es im Fertigungskontext (vorausschauende Wartung, Maschinen- und Logistikoptimierung, neue Materialien) oder im Forschungskontext (Modelloptimierung, Visualisierung), für die Entwickler*innen, Manager*innen und letztlichen Anwender*innen und Nutzer*innen des KI-Produkts oftmals unklar. Eine Berücksichtigung ethischer Aspekte im etwaigen Kontext wird bisweilen nicht zuletzt deshalb als obsolet betrachtet. Darüber hinaus fehlt es nicht selten an entsprechendem Know-how, spezifischer Planung und zusätzlichen Ressourcen (vgl. Beckert 2021), um der KI-genuinen Konstitution, „der Fähigkeit, jede kontinuierliche Funktion beliebig genau zu approximieren" (Kersting & Tresp 2019: 4), in adäquatem Maß nachzukommen, bzw. den Anforderungen einer vertrauenswürdigen KI innerhalb der Arbeitsabläufe und Prüfprozesse in identischer Intensität gerecht zu werden (vgl. Beckert 2021: 20).

Besonders der letzte Punkt offenbart ein grundlegendes Manko, da Unternehmen in Anbetracht der Implementierung von KI ohnehin schon mit einer Vielfalt neuer Spannungsfelder im Organisations- und Arbeitskontext konfrontiert werden, die häufig nicht richtig eingeschätzt werden können. Eine prominente KI-Wissenschaftlerin hierzu: „Für die meisten Unternehmen ist es sowohl von der Technologie her eine schwierige Umsetzung, aber was noch viel schwieriger ist und wo sie sich noch viel schwerer mit tun, ist es, tatsächlich die Beschäftigten ordentlich mitzunehmen und in die Prozesse mit einzubinden [...]. Da muss man sicherstellen, dass diese Daten im besten Fall gar nicht erhoben werden und gar nicht abgerufen werden können. Da man aber in diese Systeme nicht richtig reingucken kann, braucht man andere Werkzeuge und Mechanismen, die in diese Lösungen reinintegriert werden, sodass Leute, die nicht in die Entwicklungsprozesse mit eingebunden sind, wissen, was dort gespeichert wird und wie mit den Daten umgegangen wird" (W11, 30. März 2022).

5.3 Hindernisse und Grenzen

Zweifelsohne sind der Realisation einer KI-Ethik und ihrer Werte auch Grenzen gesetzt. Die AI Ethics Impact Group (AIEI Group) sieht jene vor allem in nicht zu vernachlässigenden Einflussfaktoren wie kulturellen Kontextabhängigkeiten und differierenden Anwendungsfeldern, die unterschiedliche Handlungs- und Gestaltungslogiken aufweisen und infolge die Berücksichtigung von Heterogenität hinsichtlich der Realisierung bestimmter Werte sowie ihrer praktischen Umsetzung erfordern. Ähnlich verhält es sich mit Gruppen unterschiedlicher Nutzer*innen und Anwender*innen, die oftmals divergierende Bedürfnisse und Anforderungen an ethische KI-Leitlinien erheben und somit einen ihren Anforderungen entsprechenden *simplify but not oversimplify* Orientierungsrahmen beanspruchen. Daneben müssten handlungsfähige Rahmenwerke für vertrauenswürdige KI-Algorithmen dem komplexen Entwicklungs- und Implementierungsprozess, der soziotechnischen Natur von KI-Systemen sowie der damit zusammenhängenden Verantwortung von Systementwicklern*innen sowie -anwendern*innen und -nutzern*innen Rechnung tragen (vgl. AIEI Group 2019: 10). Diese Befunde legen die Deutung nahe, dass sich inzwischen eine teilweise nur schwer überschaubare Vielfalt entsprechender Vorschläge zur Realisation jenes ethischen KI-Versprechens sowie der damit einhergehenden Regelungen auf nationaler und EU-Ebene herauskristallisiert hat, von einer allgemein anerkannten Position in absehbarer Zeit indes noch keine Rede sein kann. Gleichwohl ist neben den zersplitterten nationalen Regelungsansätzen auf der EU-Ebene, wie in etwa dem Data Governance Act oder dem Artificial Intelligence Act, vor allem die DSGVO

als konstitutiver Rahmen hervorzuheben, welcher einen ersten Schritt in Richtung Harmonisierung und Modernisierung des europäischen Datenschutzrechts und infolge Etablierung eines auf Vertrauenswürdigkeit und gemeinsamen ethischen Werten beruhendes KI-Programm anstrebt.

Eine interviewte KI-Expertin gelangt hinsichtlich dieser Umstände zu der Schlussfolgerung: „Es kommt darauf an, was man unter Ethik und Normen versteht. US-Amerikaner geben sich beispielsweise starke Hausregeln. Google und Facebook haben relativ strikte Vorgaben darin, was sie auf ihren Plattformen erlauben. Gerade spezifisch die USA mit der schwarzen Bürgerrechtsbewegung, Diversität etc., das sind politische Themen, die dann in die unternehmensinternen Regeln mit einfließen. Das sind aber keine regulatorisch-gesetzlichen Vorgaben. Da sind weder straffe Gesetze noch Standardisierungsvorgaben von betroffen. In Deutschland will man aber gerade genau das in den Vordergrund stellen. Man will einen bestimmten Regulierungskanon haben – ‚humanzentrierte KI' – oder der Begriff ‚Technologische Souveränität'. Das ist ein politisches Mittel, um zum Ausdruck zu bringen, dass die EU gegenüber den USA eine Art Sonderstellung haben wollen. Die Datenschutz-Grundverordnung wäre solch ein Tool. Hier haben sich die großen amerikanischen Konzerne gegen gewehrt. Trotzdem hat es funktioniert, dass ein neuer Standard gesetzt wurde und nichts passiert ist – keine Verhinderung von Innovationen" (E5, 21. Februar 2022).

6. Perspektiven

In Anbetracht der vom Center for Data Innovation erhobenen Befunde offenbart sich, dass Europa im Vergleich zu den USA und China insbesondere aufgrund des identifizierten „disconnect between the amount of AI talent in the EU and its commercial AI adoption and funding" (Castro et al. 2019) die Rolle des Nachzüglers im internationalen KI-Wettrennen einnimmt. Dies gilt jedenfalls dann, wenn man die der Untersuchung zugrunde gelegten sechs Hauptkategorien („talent", „research", „development", „adoption", „data" und „hardware") als Richtindex für die globale Positionierung innerhalb der internationalen *AI economy* geltend macht. Obschon Europa gemäß der Auswertung in den Bereichen Talent, Forschung und Entwicklung bislang noch vor China rangiert, wird es in keiner der analysierten Kategorien an der Spitze aufgeführt und zudem in vielen Bereichen auf lange Sicht als durch chinesische KI-Bemühungen in seiner globalen Position überflügelt angesehen. Bemerkenswert erweist sich vor allem die Feststellung, dass Europa im Bereich der Akzeptanz und damit zusammenhängenden Rezeption und Bereitschaft (*adoption*), KI-Lösungen als „Potenziale für Wertschöpfung, Wettbewerbsfähigkeit und bessere Lebensumstände von Bürgerinnen und Bürger[n]" (BMBF 2021b: 5) anzuerkennen, hinter seinen Mitbewerbern aus Fernost und den Vereinigten Staaten zurückbleibt. Dies gilt einerseits für die gesellschaftlich verbreitete Skepsis hinsichtlich der Auffassung und Beurteilung von KI – „Individuals in the European Union typically have more negative feelings toward AI in the workplace than workers in the United States, and significantly more negative feelings than employees in China" (Castro et al. 2019). Andererseits gilt dies für wirtschaftlich und industriell verortete Relevanzzuschreibungen bezüglich des Einsatzes von KI im Produktions- und Herstellungsprozess – „Other surveys have found that Europe has shown less urgency to adopt AI. For example, a survey of executives found that 40 percent of European respondents thought AI ‚is still nascent and unproven', compared with just 27 percent and 30 percent of North American and Asian-Pacific

firms, respectively“ (ebd.). Die EU erhofft sich vor allem, durch Evokation des intensivierten Umgangs mit KI-bezogenen Dimensionskriterien wie dem Vorrang menschlichen Handelns und menschlicher Kontrolle, technischer Robustheit und Sicherheit, Privatsphäre und Datenqualitätsmanagement, Transparenz und Erklärbarkeit, Vielfalt, Nichtdiskriminierung und Fairness, gesellschaftliches und ökologisches Wohlergehen sowie Rechenschaftspflicht (vgl. European Commission 2019b) sowie den dazugehörigen Vorschriften und Maßnahmen, Europas Position als *globales Zentrum für vertrauenswürdige KI* (Europäische Kommission 2021) zu sichern, um im Umkehrschluss ein *Ökosystem für Exzellenz und Vertrauen* (ebd.) zu errichten.

Perspektivisch gesehen besteht kein Zweifel darin, dass die EU „ethischen Prinzipien der Achtung der menschlichen Autonomie, der Schadensverhütung, der Fairness und der Erklärbarkeit“ (PLS 2020: 10) die Rolle desjenigen Handlungsmoments zuspricht, welches zur Resilienz – der Fähigkeit, unerwartet widrigen Umständen adäquat entgegentreten zu können und diese schnellstmöglich zu verkraften, indem Prognostizierbarkeit und Prävention kultiviert werden (vgl. acatech 2020; vgl. Thoma 2014) – der Wirtschafts- und Gesellschaftssysteme Europas führen soll. Dabei soll zeitgleich der Aufbau und Erhalt eigener Fähigkeiten sowie die Vermeidung oder Verringerung von einseitigen, strukturellen Abhängigkeiten von anderen Wirtschaftsräumen ermöglicht werden (vgl. Fraunhofer ISI 2020: 4; vgl. PLS 2020b: 5–6; vgl. Schieferdecker & March 2020). So werden die deutsche Bundesregierung und Europäische Kommission nicht müde, die Unabdingbarkeit jener eingeschlagenen Perspektive mit emphatischen Nachdruck zu bekräftigen: „Europe can distinguish itself from others by developing, deploying, using and scaling Trustworthy AI, which we believe should become the only kind of AI in Europe, in a manner that can enhance both individual and societal well-being” (European Commission 2019b: 6). Zahlreiche KI-Strategiepapiere der letzten Zeit versuchen diesen postulierten Status quo sowie die Legitimierung des gewählten politischen Wegs zu konsolidieren. Dabei wird die vorherrschende Version einer einheitlich europäischen Identität proklamiert, deren mutmaßlich gleichgerichtetes Zielvorhaben, wie ökonomisches Wachstum, Sicherheit, Nachhaltigkeit, Armut- und Krankheitsbekämpfung usw., durch eine nachhaltige Fokussierung auf KI und spezieller: auf ethische/vertrauenswürdige KI, als magischem Schlüssel (im Orig.: „magical key“) zur Lösung nationaler, europäischer und globaler Probleme (vgl. Ossewaarde & Gülenç 2020: 55), erreicht werden könnten. Durch die HLEG realisierte Deliverables der vergangenen vier Jahre wie die bereits erwähnten Ethics Guidelines for Trustworthy AI, in denen die sieben Kerndimensionen vertrauenswürdiger KI festgelegt wurden sowie drei weitere, hierauf aufbauenden Deliverables seit 2018[9], scheinen diesen Eindruck nochmals zu verstärken. Fasst man das Urteil Pereiras, dass „[b]y trying to innovate in AI just in the same way as the US or China, Europe has already lost the competitiveness race“ (Pereira 2021), als faktisch gegeben auf, erfährt der Impuls, einen eigenständigen Weg in der KI-Ausrichtung einzuschlagen, zusätzliche Nachvollziehbarkeit.

[9] *Deliverable II*: Policy and Investment Recommendations for Trustworthy AI
Deliverable III: The final Assessment List for Trustworthy AI (ALTAI)
Deliverable IV: Sectoral Considerations on the Policy and Investment Recommendations

7. Resümee und Ausblick

7.1 Ein legitimes Bestreben

Mit dem Essay wurde die Frage verfolgt , ob eine Trustworthy AI Made in Germany/Europe als ein relevantes, konstruktiv einzustufendes und vor allem seinen Ansprüchen gerecht werdendes Distinktionsmerkmal identifiziert werden kann. Zudem wurde versucht zu eruieren, ob ein solches Distinktionsmerkmal den europäischen Weg für KI anleiten und im bestmöglichen Fall als handlungs- und orientierungsleitendes Richtmaß im internationalen KI-Wettbewerb herangezogen werden kann. Diesbezüglich gilt es zunächst einmal festzuhalten, dass sich der Entwicklungs- und Ausbauwunsch autonomer, autarker und gar europäisch-eigenständiger KI-Strukturen auf europäischem Boden als nachvollziehbar und plausibel erweist. Besonders die noch immer in beträchtlichem Ausmaß bestehende technologisch-digitale Abhängigkeit gegenüber der Konkurrenz aus USA und Ostasien sowie damit verbundene Unsicherheiten unterstreichen die Plausibilität des Bestrebens. Eine Rücksichtnahme ethischer Richtlinien im Zusammenhang des Umgangs mit Algorithmen und Künstlicher Intelligenz im Kontext der Initiierung und Etablierung von Trustworthy AI (vgl. BMI 2018) stellt sich (zumindest aus politischer Sicht) als ein gleichermaßen sinnvolles Unterfangen dar, denn: Aus deutsch-europäischer Sicht offenbart sich der gegenwärtige Ist-Zustand in der bestehenden Divergenz zwischen der Regulierungshoheit seitens privater, nichteuropäischer *De-facto-Regulierer*, deren für Europa bis dato nicht substituierbare digital-technologische Infrastrukturen wie Plattformen fernab von demokratischer Kontrolle und ohne Rekurs auf europäische Werte sowie der hiermit einhergehenden Wirksamkeit des geltenden (Datenschutz-)Rechts dargeboten werden. Andererseits wird der Aspekt der Datenhoheit durch die Abhängigkeit gegenüber US-amerikanischen und chinesischen Daten- und Analyseinfrastrukturen, deren Datenschutzregularien dem (Daten-)Souveränitätsverständnis europäischer Bürger*innen, Unternehmen sowie der Zivilgesellschaft zuwiderlaufen, konterkariert (vgl. Kagermann & Wilhelm 2020: 7).

Hierbei ist die angestrebte Vision einer möglichst eigenständigen und besonderen KI-Entwicklung nicht zuletzt als Ausdruck eines Technologieversprechens aufzufassen. Gemäß jenem werden weitreichende Entwicklungspotenziale der KI prognostiziert, sofern die Etablierung einer vertrauenswürdigen KI, als international geltender Leitlinie, Konkurrenzvorteile für Europa und die Bundesrepublik begründen und infolge gesellschaftspolitisch wünschenswerte Trends ermöglichen und fördern kann (vgl. Hirsch-Kreinsen 2023). Wie aufgezeigt wurde, erweist sich dieser Weg aus verschiedenen Perspektiven heraus betrachtet als legitim. So verdeutlichen empirische Ergebnisse, dass europäische Staaten in den entscheidenden Kategorien, die eine Überprüfung des gegenwärtigen Leistungs- und Entwicklungsstands im Bereich Künstlicher Intelligenz gewähren, hinter der Hauptkonkurrenz aus den USA und China zurückhängt (vgl. Castro et al. 2019). Besonders die mangelnde Rezeption und Bereitschaft, KI-Lösungen als „Potenziale für Wertschöpfung, Wettbewerbsfähigkeit und bessere Lebensumstände von Bürgerinnen und Bürger[n]“ (BMBF 2021b: 5) anzuerkennen, hat sich als einer derjenigen Hauptfaktoren hervorgetan, welcher die weitere Erforschung, Entwicklung und Verbreitung von KI-Produkten und -anwendungen bremst.

Hiermit einher geht auch die unter dem Stichwort der *Black Box-Problematik* (Müller 2022) bekannt gewordene Diskussion, die gleichermaßen als Hemmnisfaktor für die Verbreitung

von Akzeptanz auftritt. Unter der Black Box-Problematik wird im Diskurs um KI-Technologie ein Zustand opaker Funktionsweisen verstanden. Zwecks Einführung klarer Richt- und Leitlinien sowie Normen und Zertifizierungen wird versucht, der besagten Problematik mit Transparenzbemühungen zu begegnen, um infolge den Grad an Kontrollierbarkeit hinsichtlich der im Rahmen des Blackbox-Problems entstandenen Unsicherheiten zu erhöhen. In Kombination mit dem ferneren Bestreben, europäische Werte wie demokratische Kontrolle, Datenschutz und ein (Daten-)Souveränitätsverständnis zu konsolidieren, erscheint dieses Vorhaben auch seitens Vertreter*innen der Industrie auf positive Resonanz zu stoßen, wie eine KI-Expertin hervorhebt: „Tatsächlich gibt es aber auch hier Stimmen aus den Unternehmen selbst, die sagen: ‚Wir brauchen bestimmte Rahmenbedingungen.' Dann ist es kein Innovationshindernis, sondern eine klare Marktstruktur mit bestimmten Normen und rechtlichen Vorgaben, auf denen sich alle bewegen. Da kann man dann auch Impulse für neue Innovationen setzen" (E5, 21. Februar 2022). Davon ausgehend ist der nachgezeichnete Weg, eben jenes mangelnde Vertrauen der Bürger*innen, Anwender*innen und Nutzer*innen zu gewinnen, vielversprechend und durchaus überzeugend.

7.2 Ein problematisches Narrativ

Richtet man den Blick auf das avisierte Ziel, ein spezifisches Distinktionsmerkmal konstituieren zu wollen, lässt sich zunächst konstatieren, dass KI-Zertifizierungen im internationalen Kontext bisweilen kaum anzutreffen sind und auch ein Diskurs um solche zumeist nur dann zutage gefördert wird, wenn jener durch Dritte, wie die Europäische Union bzw. Kommission, angestoßen wird. Dies gilt vor allem für multinationale Unternehmen wie Amazon, Apple, Facebook, Google oder Intel, die zwar im Rahmen der „Partnership on AI"[10] Richtlinien erstellt haben, welche indes keiner Zertifizierung durch Dritte obliegen bzw. keine Zertifizierung durch Dritte vorsehen (vgl. Mangelsdorf et al. 2021: 3). Die Vorstellung einer KI Made in Europe erscheint vor diesem Hintergrund – zumindest dann, wenn sie ein bislang weitgehend unbeachtetes Terrain betritt und ein mutmaßlich bestehendes Desiderat im KI-Diskurs identifiziert und jenes mit innovativen Ideen versucht, aufzuladen – als gerechtfertigtes Bestreben, ein spezifisches Alleinstellungsmerkmal zu begründen.

Konträr hierzu und mit Blick auf die ausgeprägten internationalen Verflechtungen der KI-Entwicklung und ihrer Community beharren einige KI-Expertinnen und Experten allerdings darauf, dass es wenig Sinn ergeben würde, von einer *europäischen/deutschen KI*, als Sonderweg im globalen KI-Wettbewerb, zu sprechen (W11, 29. März 2022). In Anbetracht dessen sei die Grundlegung einer deutschen KI als solche als ein nahezu an Unsinnigkeit zu bezeichnendes Konstrukt einzustufen. Die Entwicklung von KI sei von Anfang an keine deutsche bzw. europäische, sondern eine ausgeprägt internationale, d. h. besonders von den USA geprägte, Entwicklung gewesen (W1, 17. November 2021). Obendrein muss auf die bereits skizzierte Fülle an internationalen KI-orientierten Verlautbarungen verschiedenster Regierungen und Länder verwiesen werden. Wie aufgezeigt, halten diese (mal mehr und mal weniger) kongru-

[10] „Partnership on AI (PAI) is a non-profit partnership of academic, civil society, industry, and media organizations creating solutions so that AI advances positive outcomes for people and society. By convening diverse, international stakeholders, we seek to pool collective wisdom to make change. We are not a trade group or advocacy organization." Online: https://partnershiponai.org/about/#mission [10. Mai 2022].

ente Bekanntmachungen und Erklärungen über die Konzeptionierung, Gestaltung und Einführung von vertrauenswürdigen KI-Kriterien bereit. Dieser Sachverhalt wirft Zweifel am avisierten Alleinstellungsmerkmal europäischer KI-Bemühungen auf, relativiert den beanspruchten Status als globales Zentrum für vertrauenswürdige KI und widerstrebt dem Narrativ eines exklusiven KI-Sonderwegs. Dies gilt auch dann, wenn das bis hierhin erarbeitete Konzept in seiner Tiefe und Ausprägung weltweit keinem vergleichbaren Pendant gegenübersteht, denn: Quantität hinsichtlich politisch formulierter Richt- und Leitlinien als Indikator für die Legitimation einer exklusiven oder gar eigenständigen Stellung heranzuziehen, erweist sich als ebenso abwegig wie widersinnig, wie die Berufung auf qualitative Handlungs- und Gestaltungsparameter im Kontext moralisch getriebener Geltungsansprüche. Die alte Frage, inwiefern (und vor allem: welche) moralische(n) Werte kulturübergreifend als einheitlich und allgemein anerkannt zu betrachten bzw. aufzufinden sind (vgl. Haridy 2019), drängt sich an dieser Stelle in den Vordergrund, ohne mit Blick auf den limitierten Rahmen des vorliegenden Beitrags jedoch einer eingehenderen Untersuchung unterzogen werden zu können.

Gleichermaßen kritisch betrachtet werden muss die von der Datenethikkommission auf den Weg gebrachte Kritikalitätspyramide (vgl. Datenethikkommission 2019: 177). Das übergreifende Modell zur Bestimmung des Kritikalitätsgrades algorithmischer Systeme bemisst die Kritikalität von KI-Systemen nach dem „Ausmaß möglicher Verletzungen von Rechtsgütern und Menschenleben durch den Einsatz eines KI-Systems sowie de[n] Umfang der Handlungsfreiheiten des Individuums bei seiner Auswahl und Nutzung“ (PLS 2020: 10). Wie mit der Plastizität und zugrundeliegenden Gestaltbarkeit und Bewertung der Anwendungsrisiken in unterschiedlichen Feldern umgegangen werden soll, inwiefern der Terminus der Kritikalität hierbei als Maß für potenzielle Gefahren definiert werden soll und kann und welche normierenden Konsequenzen daraus gezogen werden sollten, wird bisweilen nur spärlich bis gar nicht begründet. Erforderlich ist hier vor allem auch eine Abstimmung mit der Risikodefinition der EU Kommission im sog. AI Act (vgl. z. B. Müller 2022). In diesem Zusammenhang ist es notwendig darauf hinzuweisen, dass eine differenzierte Sicht auf KI als konstitutiv-handlungsleitendes Element beibehalten werden muss, da Risiken bzw. Kritikalitätsstufen in Abhängigkeit vom Anwendungskontext unterschiedliche Verbindungen stiften, ungeachtet des Vorliegens desselben KI-Systems (vgl. Heesen et al. 2020).

7.3 To-Do‘s

Die Einstufung einer vertrauenswürdigen KI Made in Europe als bloßem Gimmick eines politisch-programmatisch aufgeladenen Diskurses mutet zugegebenermaßen als etwas überspitzt formulierte Charakterisierung an. Ob jenes Gimmick als Verkaufstrick anzusehen ist oder möglicherweise doch ein fundiertes Distinktionsmerkmal darstellt, ist unter Rücksichtnahme der vorliegenden Datenlage indes nur schwierig mit einem eindeutigen „Ja“ oder „Nein“ zu beantworten, ohne sich Ambivalenzen auszusetzen. Der Slogan „Trustworthy AI – Made in Europe“ kann – unter Rücksichtnahme der erörterten Anforderungen und Herausforderungen der europäischen KI-Landschaft – zweifelsohne als attraktiver Entwicklungspfad identifiziert werden, wenn es darum geht, bestehende Konflikte, Disharmonien und Widerstände infolge mangelnder Adoption und Akzeptanz potenziell zu überwinden. So deuten zahlreiche europäisch aufgesetzte KI-Förder- und Entwicklungsinitiativen, von denen die DSGVO als prominentes und richtungsweisendes Beispiel hervorgehoben werden kann, be-

reits das Wirkpotenzial des Motivs einer vertrauenswürdigen KI Made in Europe an. Der Vorwurf, dass es sich lediglich um politische Programmatik oder um ein bloßes Gimmick handelt, wird dem Einfluss- und Inkubationspotenzial dieser Entwicklungen nicht gerecht.

Konträr hierzu erweist sich die oftmals propagierte Exklusivstellung angesichts der Vielfalt an ähnlichen und vergleichbaren Erklärungen und Verlautbarungen im außereuropäischen Kontext als kaum haltbar. Stand bzw. steht Made in Germany insbesondere für höchste qualitative Standards und Präzision, stellt dies auf Basis ethischer Kriterien innerhalb einer KI Made in Europe, die mit anderen ethischen Kriterien in Wettstreit treten, ein diffuses und teils nur schwer nachzuvollziehendes Konzept dar. Des Weiteren ist der Forschungs- und Entwicklungsprozess in Europa noch immer von einem defizitären disziplin-, akteurs- und länderübergreifenden Zusammenarbeitsstatus geprägt, der vielfach als eine der zentralen Herausforderungen angesehen werden kann (vgl. Castro et al. 2019), wenn es um den Entwicklungs- und Leistungsstatus im Vergleich zu anderen Nationen geht. Die Fokussierung auf eine "[c]ross-border cooperation in AI" (Brattberg et al. 2020) zwischen und innerhalb der europäischen Staaten wird ein wesentlicher Parameter im weltweiten KI-Wettlauf sein, um seinen Platz an der Spitze beizubehalten und nicht von der Konkurrenz verdrängt zu werden. Die gemeinsame Entwicklung ethischer KI-Kriterien kann diesbezüglich zwar als orientierungsleitendes Handlungsmoment fungieren. Nichtsdestotrotz kann eine ausschließliche Konzentration auf diesen Aspekt im KI-Diskurs nur bedingt als erfolgsversprechender Faktor ausgemacht werden. Vielmehr entscheidend wird es sein, rechtliche und ethisch orientierte Regulationskriterien derart zu gestalten, dass sie kein Innovationshemmnis darstellen. Diese müssen von einem gemeinsamen europäischen Grundbestreben getragen werden und eine symbiotische Verbindung aus industrieller und vertrauenswürdiger KI-Orientierung zulassen, um infolge das volle KI-Entwicklungspotenzial ausschöpfen zu können.

Zudem kann der Slogan „Trustworthy AI Made in Europe" in gewisser Weise als defizitär angesehen werden, da ein für das beabsichtige Zielvorhaben grundlegender Erfolgsindikator nicht mit einbezogen wird – der explizite Bezug auf industrielle Anwendungen. Für die Stabilisierung und Konsolidierung einer europäischen KI-Perspektive ist es vermutlich fundamental, einer potenziellen Modifizierung bzw. Neujustierung des Slogans vor allem eine Ausweitung des thematischen Gegenstandsbereichs um industrielle KI-Anwendungen voranzustellen. Die Hauptintention muss hier in der Sicherstellung einer zielgerichteten Verbindung von industriell orientierter und auf Vertrauenswürdigkeit gründender KI liegen. Bislang ungeklärt erweist sich die Frage, wie mit Hilfe von regulativen Normen die hohe Entwicklungsdynamik der KI sowie daraus resultierende neue Methoden und Anwendungspotenziale erfasst werden können. Dabei ist nicht von der Hand zu weisen, dass gerade der Industriesektor als Epizentrum deutsch-europäischer KI-Entwicklungsbemühungen fungieren sollte. Der Grund hierfür liegt in dem ohnehin schon immer bestehenden KI-Anwendungsbezug auf die Industrie (zumindest in der Bundesrepublik) sowie der gegenüber USA und China im Allgemeinen stärker ausgeprägten industriellen Orientierung (vgl. hierzu v. a. Hirsch-Kreinsen 2023). Vertreter*innen von acatech unterstreichen diese Feststellung, wenn sie erklären, dass „deutsche Forschungseinrichtungen und Unternehmen durch deutliche Vorteile in der industriellen Anwendung" (acatech 2020: 8) glänzen würden. Weiter betont ein interviewter KI-Wissenschaftler das Potenzial Deutschlands vice versa Europas, welches sich aus der im Vergleich dominie-

renden Rolle der Industrie ergeben würde: „In Deutschland ist man mit industriellen Anwendungen auch schon sehr weit, jedenfalls wesentlich weiter als in den USA“ (W11, 30. März 2022).

Obschon die Existenz einer genuin europäischen KI, wie angedeutet, vielfach auf Unverständnis trifft und auch der Etablierungsversuch eines distinktiven, ethisch orientierten KI-Diskursraumes auf geteilte Resonanz stößt, verspricht eine verstärkte Miteinbeziehung des industriellen KI-Bezugs in die teils noch opake und inkohärente Profilbildung eines eigenständigen KI-Profils Europas ein vielversprechender Ansatz zu sein. Dies gilt umso mehr mit Blick auf die besonderen industriellen Strukturbedingungen des deutschen Innovationssystems – hohes Ausbildungsniveau technischer Fachkräfte und Ingenieure*innen, die die Nutzung von KI-Werkzeugen erleichtern (vgl. acatech 2020). Für den europäischen KI-Diskurs impliziert dies, bestimmte technologische Bereiche „über eine gezielte Positionierung in bestimmten Bereichen strukturell arbeitsteilig mitzugestalten und Abhängigkeiten vorzubeugen“, wobei „Alleinstellungsmerkmale […] am ehesten in Schnittstellenbereichen jener Technologien zu entwickeln [sein] [sollten], die aktuell durch die softwaregetriebene Digitalisierung revolutioniert werden. Im besonderen Fokus stehen hierbei die Produktionstechnologien“ (Kroll et al. 2022: 8). Ziel sollte es also sein, diejenigen „Nischen“ ausfindig zu machen, wie ein Interviewpartner erläutert, die auf Grund von fehlender Größe für die internationalen Hyperscaler unattraktiv, gleichermaßen jedoch groß genug für europäische Akteure seien, um KI-Entwicklungen in diesem Bereich voranzutreiben (vgl. W7I, 15. Februar 2022). Eine intensivierte industrielle Orientierung in Verbindung mit ethisch orientierten, vertrauenswürdigen KI-Anwendungen kann eine solche Nische repräsentieren und den Vorwurf eines programmatisch aufgeladenen, künstlichen Gimmicks zusätzlich und nachhaltig entkräften.

Literatur

acatech (Hrsg.) (2020): *Corona-Krise: Volkswirtschaft am Laufen halten, Grundversorgung sichern, Innovationsfähigkeit erhalten – intervenieren – stabilisieren – stimulieren*. acatech IMPULS, München

AIEI Group (2019): *From Principles to Practice – An Interdisciplinary Framework to operationalise AI ethics*. Artificial Intelligence Ethics Impact Group. Bertelsmann Stiftung, Gütersloh

Beckert, Bernd (2021): *Vertrauenswürdige Künstliche Intelligenz – Ausgewählte Praxisprojekte und Gründe für das Umsetzungsdefizit*. In: Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis, Vol. 30, No. 3, 17–22

Bouchard, Gérard (2016): *Europe in Search of Europeans - The Road of Identity and Myth*. Jacques Delors Institute, Paris

Bradford, Anu (2021): *The European Union in a globalised world: the "Brussels effect"*. In: Revue Européenne du Droit, Vol. 2, 75–79

Brattberg, Erik/ Csernatoni, Raluca & Rugova, Venesa (2020): *Europe and AI: Leading, Lagging Behind, or Carving its own Way?* Stand: 09. Juli 2020. Online: https://carnegieendowment.org/2020/07/09/europe-and-ai-leading-lagging-behind-or-carving-its-own-way-pub-82236 [19. Dezember 2021]

Bundesministerium der Justiz und für Verbraucherschutz (BMJV) (2021): *Zentrum für vertrauenswürdige Künstliche Intelligenz (KI) soll Verbraucherinteressen in der digitalen Welt stärken*. [Pressemitteilung] Stand: 20. Oktober 2021. Online: https://www.bmj.de/SharedDocs/Archiv/DE/Pressemitteilungen/2021/1020_Zentrum_KI.html [14. April 2022]

Bundesministerium des Innern und für Heimat (BMI) (2018): *Leitfragen der Bundesregierung an die Datenethikkommission*. Berlin

Bundesministerium des Innern und für Heimat (BMI) (2021): *Datenschutz-Grundverordnung.* Berlin

Bundesministerium für Bildung und Forschung (BMBF) (2021a): *Daten, Infrastrukturen und Technologien im EFR*. Stand: 2021. Online: https://www.forschungsraum.eu/forschungsraum/de/forschungsthemen/themenfeatures/daten-infrastrukturen-und-technologien-im-efr/daten-infrastrukturen-und-technologien.html?utm_campaign=sea&utm_term=data [06. Januar 2022]

Bundesministerium für Bildung und Forschung (BMBF) (2021b): *Technologisch souverän die Zukunft gestalten – BMBF-Impulspapier zur technologischen Souveränität*. Herausgegeben durch Bundesministerium für Bildung und Forschung (BMBF). Berlin

Bundesministerium für Wirtschaft und Energie (BMWi) (2020): *„KI – Made in Germany" etablieren.* Stand: 30. November 2020. Online: https://www.bmwi.de/Redaktion/DE/Pressemitteilungen/2020/11/20201130-ki-made-in-germany-etablieren.html [14. März 2022]

Bundesministerium für Wirtschaft und Energie (BMWi) (2021): *Schwerpunktstudie digitale Souveränität – Bestandsaufnahme und Handlungsfelder*. Herausgegeben durch Bundesministerium für Wirtschaft und Energie (BMWi), Berlin

Bundesverband der Deutschen Industrie e.V. (BDI) (2018): *Datenschutz darf keinesfalls zum Innovationshemmnis und Standortnachteil werden*. Stand: 20. Mai 2018. Online: https://bdi.eu/artikel/news/datenschutz-darf-keinesfalls-zum-innovationshemmnis-und-standortnachteil-werden/ [11. April 2022]

Bundesverband Deutscher Patentanwälte (2022): *130 Jahre „Made in Germany" – Vom Makel zum must-have*. Stand: 2022. Online: https://www.bundesverband-patentanwaelte.de/patente/130-jahre-made-germany-vom-makel-zum-must/#:~:text=Ungeachtet%20der%20nicht%20eindeutig%20geregelten,hohen%20Verarbeitungsstandards%20und%20hochwertigen%20Materialien. [23. Februar 2022]

Butcher, Sarah (2016): *HireVue: Die schöne neue Welt der Vorstellungsgespräche*. eFinancialCareers, Stand: 23. August 2016. Online: https://www.efinancialcareers.de/nachrichten/2016/08/hirevue-die-schone-neue-welt-der-vorstellungsgesprache [17. März 2022]

Castro, Daniel/ McLaughlin, Michael & Chivot, Eline (2019): *Who is winning the AI Race: China, the EU or the United Sates?* Stand: 19. August 2019. Online: https://datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/ [24. März 2022]

Changbao, Jia/ Leutner, Mechthild & Minxing, Xiao (2021): *Die China-Berichterstattung in deutschen Medien im Kontext der Corona-Krise*. Herausgegeben von der Rosa-Luxemburg-Stiftung, Berlin

China Academy of Information and Communications Technology (2021): *White Paper on Trustworthy Artificial Intelligence*. China Academy of Information and Communications Technology (CAICT) JD Explore Academy, Beijing

Datenethikkommission (2019): *Gutachten der Datenethikkommission*. [Gutachten] Herausgegeben durch Datenethikkommission der Bundesregierung und Bundesministerium des Innern, für Bau und Heimat. Berlin

Deighton, Ben (2017): *'Made in Europe' label could help EU competitiveness*. European Commission, HORIZON – The EU Research & Innovation Magazine. Stand: 01. März 2017. Online: https://ec.europa.eu/research-and-innovation/en/horizon-magazine/made-europe-label-could-help-eu-competitiveness [24. Februar 2022]

Die Bundesregierung (2018a): *Forschung und Innovation für die Menschen – Die Hightech-Strategie 2025*. [Positionspapier] Herausgegeben durch das Bundesministerium für Bildung und Forschung (BMBF). Stand: September 2018, Berlin

Die Bundesregierung (2018b): *Strategie Künstliche Intelligenz der Bundesregierung*. [Strategiepapier] Stand: November 2018, Berlin

Europäische Kommission (2018): *Ethik-Leitlinien für eine vertrauenswürdige KI*. Unabhängige hochrangige Expertengruppe für Künstliche Intelligenz, Brüssel

Europäische Kommission (2021): *Ein Europa für das digitale Zeitalter: Kommission schlägt neue Vorschriften und Maßnahmen für Exzellenz und Vertrauen im Bereich der künstlichen Intelligenz vor*. Europäische Kommission, Pressemitteilung. Brüssel

European Commission (2018): *Artificial Intelligence – A European Perspective*. Published by Joint Research Centre (JRC), the European Commission's Science and Knowledge Service. Ispra

European Commission (2019a): *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the committee of the Regions: Building Trust in Human-Centric Artificial Intelligence*. Brussels

European Commission (2019b): *Policy and Investment Recommendations for Trustworthy AI: HLEG on AI*. Brussels

Expertenkommission Forschung und Innovation (EFI) (2022): *Gutachten 2022*. Gutachten zu Forschung, Innovation und technologischer Leistungsfähigkeit. Berlin

Finsterbusch, Stephan von/ Lindner, Roland & Kafsack, Hendrik (2022): *Intel investiert 17 Milliarden Euro in Deutschland*. Frankfurter Allgemeine Zeitung, Stand: 15. März 2022. Online: https://www.faz.net/aktuell/wirtschaft/unternehmen/intel-us-chiphersteller-investiert-17-milliarden-euro-in-deutschland-17879159.html [17. März 2022]

Fraunhofer-Institut für System- und Innovationsforschung (ISI) (2020): *Technologiesouveränität – Von der Forderung zum Konzept*. Online verfügbar unter: https://www.isi.fraunhofer.de/content/dam/isi/dokumente/ publikationen/technologiesouveraenitaet.pdf, 1–31

Greenleaf, Graham (2021): *The 'Brussels effect' of the EU's 'AI Act' on data privacy outside Europe.* In: Privacy Laws & Business International Report 1, 1–10

Grillo, Ulrich (2015): *Made in Germany ist und bleibt eine starke Marke*. Executive Letter, Bundesverband der Deutschen Industrie e.V., Berlin

Gröschel, Philippe (2022): *Künstliche Intelligenz – Wie steht es um den AI-Act der EU?* Basecamp, Stand: 18. Januar 2022. Online: https://www.basecamp.digital/kuenstliche-intelligenz-wie-steht-es-um-den-ai-act-der-eu/ [12. April 2022]

Gupta, Manjul/ Parra, Carlos M. & Dennehy, Denis (2021): *Questioning Racial and Gender Bias in AI-based Recommendations: Do Espoused National Cultural Values Matter?* In: Information System Frontiers, 1–17

Hages, Larissa/ Oslislo, Christoph/ Recker, Clemens & Roth, Steffen J. (2017): *Digitalisierung, Lock-in-Effekte und Preisdifferenzierung*. Otto-Wolff Institut für Wirtschaftsordnung, Otto-Wolff-Discussion Paper, 5/2017, Köln

Haridy, Rich (2019): *Oxford anthropologists identify seven universal rules of morality*. New Atlas – Science. Stand: 13. Februar 2019. Online: https://newatlas.com/seven-universal-moral-rules-oxford-study/58474/#:~:text=%22People%20everywhere%20face%20a%20similar,the%20right%20thing%20to%20do.%22 [10. Mai 2022]

Heesen, Jessica et al. (2020): *Ethik-Briefing – Leitfaden für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen*. [Whitepaper] Herausgegeben durch Plattform Lernende Systeme – Die Plattform für Künstliche Intelligenz, Berlin

Hirsch-Kreinsen, Hartmut (2023): *Das Versprechen der Künstlichen Intelligenz – Zur Genese eines Technologiefeldes*. [In Veröffentlichung] Frankfurt am Main: Campus Verlag

Horwitz, Josh (2021): *China passes new personal data privacy law, to take effect Nov. 1*. Reuters, Stand: 20. August 2021. Online: https://www.reuters.com/world/china/china-passes-new-personal-data-privacy-law-take-effect-nov-1-2021-08-20/ [12. April 2022]

Jahn, Thekla (2021): *Künstliche Intelligenz und „Ethics Washing" – „Wir haben keine seriösen Experten für angewandte Ethik der KI"*. Deutschlandfunk, 11. Juni 2021. Stand: 2021. Online: https://www.deutschlandfunk.de/kuenstliche-intelligenz-und-ethics-washing-wir-haben-keine-100.html [10. August 2022]

Jansen, Fieke & Cath, Corinne (2021): *Algorithmic registers and their limitations as a governance practice*. In: Kaltheuner, Frederike (Hrsg.): Fake AI. Manchester: Meatspace Press, 183–192

Kagermann, Henning & Wilhelm, Ulrich (Hrsg.) (2020): *European Public Sphere – Gestaltung der digitalen Souveränität Europas*. acatech IMPULS, München

Kersting, Kristian & Tresp Volker (2019): *Maschinelles und Tiefes Lernen – Der Motor für „KI made in Germany"*. [Whitepaper] Herausgegeben durch Plattform Lernende Systeme – Die Plattform für Künstliche Intelligenz, Berlin

KPMG Deutschland (2022): *Künstliche Intelligenz ohne Gewissen? KPMG unterstützt Unternehmen bei der Entwicklung gesellschaftsverträglicher Algorithmen*. Stand: 2022. Online: https://klardenker.kpmg.de/digital-hub/kuenstliche-intelligenz-ohne-gewissen/ [04./05. April 2022]

Kroll, Henning et al. (2022) : *Schlüsseltechnologien*. In: Studien zum deutschen Innovationssystem, No. 7-2022, Expertenkommission Forschung und Innovation (EFI), Berlin

Ludwig, Thomas (2015): *EU drückt bei Label-Pflicht auf die Bremse*. Handelsblatt, International. Stand: 29. Mai 2015. Online: https://www.handelsblatt.com/politik/international/made-in-eu-drueckt-bei-label-pflicht-auf-die-bremse/11843702.html [24. Februar 2022]

Lutteroth, Johanna (2012): *Qualitätssiegel „Made in Germany"*. Stand: 24. August 2012. Online: https://www.spiegel.de/geschichte/made-in-germany-vom-stigma-zum-qualitaetssiegel-a-947688.html [23. Februar 2022]

Manager Magazin (2017): *„Made in Germany" ist das beliebteste Label der Welt*. Stand: 27. März 2017. Online: https://www.manager-magazin.de/politik/deutschland/made-in-germany-ist-das-beliebteste-label-der-welt-a-1140546.html [23. Februar 2022]

Metzinger, Thomas (2019): *EU-Ethikrichtlinien für Künstliche Intelligenz – Nehmt der Industrie die Ethik weg!* Der Tagesspiegel. Stand: 08. April 2019. Online: https://www.tagesspiegel.de/politik/eu-ethikrichtlinien-fuer-kuenstliche-intelligenz-nehmt-der-industrie-die-ethik-weg/24195388.html [10. August 2022]

Müller, Angela (2022): *„Der Artificial Intelligence Act der EU: Ein risikobasierter Ansatz zur Regulierung von Künstlicher Intelligenz"*. In: EuZ Zeitschrift für Europarecht, Ausgabe 1|2022, 1–25

Noll, Andreas (2019): *KI ersetzt Personaler – Software analysiert Bewerber*. Deutschlandfunk Nova, Stand: 01. Oktober 2019. Online: https://www.deutschlandfunknova.de/beitrag/job-bewerbung-kuenstliche-intelligenz-entscheidet [17. März 2022]

Ossewaarde, Ringo & Erdener, Gülenç (2020): *National Varieties of Artificial Intelligence Discourses: Myth, Utopianism, and Solutionism in West European Policy Expectations*. In: Computer, Vol. 53, No. 11, 53–61

Pereira, David (2021): *AI Ethics Sells…But Who's Buying?* Towards Data Science, Stand: 19. April 2021. Online: https://towardsdatascience.com/ai-ethics-sells-but-whos-buying-c050054ec44 [06. April 2022]

Petrova, Anastasia (2019): *The impact of the GDPR outside the EU*. Ius Laboris, Stand: 17. September 2019. Online: https://www.lexology.com/library/detail.aspx?g=872b3db5-45d3-4ba3-bda4-3166a075d02f [12. April 2022]

Pierides, Mike & Roxon, Charlotte (2021): *Legislative Approaches to AI: European Union v. United Kingdom*. Morgan Lewis, Tech & Sourcing. Stand: 25. Oktober 2021. Online: https://www.morganlewis.com/blogs/sourcingatmorganlewis/2021/10/legislative-approaches-to-ai-european-union-v-united-kingdom [24. März 2022]

Plattform Lernende Systeme (PLS) (2020): *Zukunftsfähigkeit mit KI sichern – Ansätze für mehr Resilienz und digitale Souveränität*. [Positionspapier] Lernende Systeme – Die Plattform für Künstliche Intelligenz, Berlin

Raidl, Melanie (2021): *Digitale Ethik –„Beachten, welche Folgen KI für den Menschen hat": Diese Frau möchte den Umgang mit KI regeln*. Handelsblatt, Stand: 16. November 2021. Online: https://www.handelsblatt.com/technik/it-internet/digitale-ethik-beachten-welche-folgen-ki-fuer-den-menschen-hat-diese-frau-moechte-den-umgang-mit-ki-regeln-/27803136.html [05. April 2022]

Redman, Thomas C. (2018): *If your Data is bad, your Machine Learning Tools are useless*. Harvard Business Review, Stand: 02. April 2018. Online: https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless [13. April 2022]

Richter, Carola & Gebauer, Sebastian (2010): *Die China-Berichterstattung in den deutschen Medien*. Heinrich-Böll-Stiftung – Bildung + Kultur, Bd. 5, Berlin

Sahin, Kaan & Barker, Tyson (2021): *Europe's Capacity to Act in the Global Tech Race – Charting a Path for Europe in Times of Major Technological Disruption.* DGAP, Stand: 22. April 2021. Online: https://dgap.org/en/research/publications/europes-capacity-act-global-tech-race#Artifical%20Intelligence [12. April 2022]

Schieferdecker, Ina & March, Christoph (2020): *Digitale Innovationen und Technologiesouveränität*. In: Wirtschaftsdienst 100, 30–35

SIENNA (2021): *Ethics by Design and Ethics of Use approaches for Artificial Intelligence, Robotics and Big Data*. The SIENNA Project, Stakeholder-informed ethics for new technologies with high socio-economic and human rights impact. Stand: 21. März 2021. Online: https://sienna-project.eu/public-consultation/ai/ethics-by-design/#:~:text=Ethics%20by%20Design%20(EbD)%20is,into%20design%20and%20development%20processes. [16. März 2022]

Statista (2017): *Made-In-Country-Index*. Stand: 2022. Online: https://de.statista.com/page/Made-In-Country-Index [23. Februar 2022]

Thoma, Klaus (Hrsg.) (2020): *Resilien-Tech – „Resilience-by-Design": Strategie für die technologischen Zukunftsthemen*. Acatech STUDIE, München

Venkina, Ekaterina (2021): *„Die Kluft zwischen den USA und der EU im Bereich KI ist übertrieben"*. BigData-Insider. Stand: 29. November 2021. Online: https://www.bigdata-insider.de/die-kluft-zwischen-den-usa-und-der-eu-im-bereich-ki-ist-uebertrieben-a-1077581/ [15. März 2022]

Verbraucherportal Deutschland (2004): *Made in Germany – Der Ursprung von Made in Germany*. Stand: 21. September 2017. Online: https://web.archive.org/web/20100318064607/http://www.verbraucherportal-deutschland.de/aboutus_madein.php [24. Februar 2022]

Wahlster, Wolfgang & Winterhalter, Christioph (2020): *Deutsche Normungsroadmap – Künstliche Intelligenz*. DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE, Berlin

Zhang, Daniel/ Mishra, Saurabh/ Brynjolfsson, Erik/ Etchemendy, John/ Ganguli, Deep/ Grosz, Barbara/ Lyons, Terah/ Manyika, James/ Niebles, Juan Carlos/ Sellitto, Michael/ Shoham, Yoav/ Clark, Jack & Perrault, Raymond (2021): *The AI Index 2021 Annual Report*. AI Index Steering Committee, Human-Centered AI Institute. California: Stan-ford University

Band 211

Beiträge aus der Forschung

sfs