Lernende Systeme

DIE PLATTFORM FÜR KÜNSTLICHE INTELLIGENZ

Zertifizierung von KI-Systemen

Impulspapier

Jessica Heesen et al.

AG IT-Sicherheit, Privacy, Recht und Ethik; AG Technologische Wegbereiter und Data Science

Zusammenfassung

Die Zertifizierung von Künstlicher Intelligenz (KI) gilt als eine mögliche Schlüsselvoraussetzung, um den Einsatz von KI-Systemen in verschiedenen Wirtschafts- und Lebensbereichen voranzutreiben. Für eine Vielzahl von KI-Systemen kann eine Zertifizierung dazu beitragen, das gesellschaftliche Nutzenpotential sicher und gemeinwohlorientiert auszuschöpfen. Eine gelungene Zertifizierung von KI-Systemen ermöglicht die Erfüllung wichtiger gesellschaftlicher und ökonomischer Prinzipien, wie etwa Rechtssicherheit (z. B. Haftung und Entschädigung), Interoperabilität, IT-Sicherheit oder Datenschutz. Zudem kann sie bei Bürgerinnen und Bürgern Vertrauen schaffen, zu besseren Produkten führen und die nationale und internationale Marktdynamik beeinflussen.

Damit sich Zertifizierungsverfahren aber nicht als Innovationshemmnis erweisen, gilt es, bestimmte Standards von KI-Systemen zu garantieren, Überregulierung zu vermeiden, Innovation zu ermöglichen und bestenfalls neue Entwicklungen für einen europäischen Weg in der KI-Anwendung auslösen zu können. Das Spannungsfeld aus Potentialen und Herausforderungen bei der Zertifizierung von KI-Systemen haben Expertinnen und Experten der Plattform Lernende Systeme in vorliegendem Impulspapier systematisiert. Das Papier, das unter der Leitung der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik sowie der Arbeitsgruppe Technologische Wegbereiter und Data Science entstand, beleuchtet verschiedene technische, juristische und ethische Herausforderungen und gibt zudem einen Überblick über bestehende Initiativen zur Zertifizierung von KI-Systemen in Deutschland.

Impulspapiere der Plattform Lernende Systeme sind knappe Ad-hoc-Analysen zu aktuellen Fragen rund um Künstliche Intelligenz. Ziel des Formats ist es, einen schnellen und kompakten Beitrag zur aktuellen Debatte zu leisten und eine inhaltliche Vertiefung – etwa in Folgepublikationen – anzuregen. Impulspapiere spiegeln nicht die Position einzelner Arbeitsgruppen oder der gesamten Plattform wider, sondern die Meinung einzelner Autorinnen und Autoren.

1 Einleitung

Die Zertifizierung von Lernenden Systemen und Künstlicher Intelligenz (KI) wird als eine der Schlüsselvoraussetzungen gehandelt, um die Anwendung und Implementierung von KI-Systemen voranzutreiben. So ist die Forderung nach der Etablierung einer Zertifizierung von KI-Systemen in zahlreichen Strategiekonzepten zu finden, wie etwa in der KI-Strategie der Bundesregierung oder dem Whitepaper der Europäischen Kommission zu Künstlicher Intelligenz. Im Anschluss an diese Forderung existieren bereits einige Konzepte und Initiativen zur Zertifizierung von KI-Systemen (siehe Kapitel 5). Zertifizierung ist allerdings kein Allheilmittel. Damit Zertifizierung nicht zum Hemmschuh für Innovationen wird, ist es entscheidend, ein geeignetes Maß festzulegen, das bestimmte Standards des KI-Systems garantiert und gleichzeitig eine Überregulierung vermeidet. Gelingt eine passende Zertifizierung von KI-Systemen „made in Europe", kann diese Vertrauen und Orientierung schaffen, zu „besseren" Produkten führen und auch die internationale Marktdynamik beeinflussen. Bis zu diesem Ziel müssen jedoch noch einige Herausforderungen und Problemstellungen bewältigt werden, die im Vergleich zur Prüfung von herkömmlichen IT-Systemen und Softwarelösungen neu auftreten, wie beispielsweise eine Lösung für die Problematik der Komplexität und Dynamik von KI-Systemen. Gleichzeitig existieren für bestimmte Einsatzgebiete bereits Normen und Standards, an welche angeknüpft werden kann, um die Lücke zwischen etablierten Prüfverfahren und neuen Anforderungen moderner KI zu schließen und die Zertifizierung auf diese Weise auch marktfähig zu gestalten. Zum gegenwärtigen Zeitpunkt sind diesbezüglich noch viele Fragen offen.

Die Mitglieder der Plattform Lernende Systeme möchten den aktuellen Prozess von der Standardisierung bis hin zu einer Zertifizierung (siehe Infobox, Seite 3) von KI-Systemen als Rahmengeber begleiten. Ziel ist es, am Ende eine Lösung zu finden, die es erlaubt, die Potentiale der Zertifizierung von KI-Systemen zu realisieren und gleichzeitig negative Effekte zu vermeiden. Hierfür werden der Nutzen einer Zertifizierung vorgestellt, die Herausforderungen ausgearbeitet und inhaltliche Kriterien definiert, etwa für Anforderungen an KI-Systeme.

Vorliegendes Impulspapier stellt einen ersten Schritt auf diesem Weg dar. Es umreißt die Problemstellung, auf deren Basis dann in Folgepapieren Lösungsvorschläge präsentiert werden können. Es zeigt, dass Zertifizierungen für eine Vielzahl von KI-Systemen dazu beitragen können, ihr volles gesellschaftliches Nutzenpotential sicher und gemeinwohlorientiert auszuschöpfen. Damit dies erreicht werden kann, muss eine Form von Zertifizierung gefunden werden, die Überregulierung vermeidet, Innovation ermöglicht und gegebenenfalls selbst zum Auslöser neuer Entwicklungen für einen europäischen Weg in der KI-Anwendung wird. Gelingt eine geeignete Zertifizierung von KI-Systemen, so besteht deren Nutzen vor allem in der Vermittlung von Vertrauen und Orientierung sowie einer veränderten Marktdynamik. Die Hauptherausforderungen liegen in der Komplexität und der Dynamik der KI-Systeme: Moderne Lernende System sind häufig sogenannte Black-Box-Systeme, deren Lernprozess und Resultate zumindest nur schwer beziehungsweise nur mit hohem Aufwand nachvollziehbar sind und die sich teilweise selbstständig weiterentwickeln (siehe Infobox Deep Learning, Seite 13; Continuous Learning Systems, Seite 14). Diese Herausforderungen führen zu zahlreichen offenen Fragen für eine geeignete Zertifizierung.

Zertifizierung

Eine Zertifizierung ist eine meist zeitlich begrenzte Bestätigung, dass vorgegebene Standards, Normen oder Richtlinien eingehalten werden. Diese Bewertung wird von unabhängigen Dritten (Zertifizierungsstellen) durchgeführt und ist die höchste von drei Stufen der Konformitätsbewertung. Grundlage sind unterschiedliche national oder auch international anerkannte und gültige branchenabhängige Standards und Richtlinien. Es können sowohl Produkte und Dienstleistungen als auch Systeme, Prozesse und Personen zertifiziert werden. Meistens erfolgt eine Zertifizierung auf freiwilliger Basis, um die Qualität des zertifizierten Gegenstands nachzuweisen.

Im Bereich KI-Systeme existieren aktuell (Stand März 2020) kaum gültige und anerkannte Standards und Normen, die konkret genug sind, um die Basis einer Zertifizierung bilden zu können.

2 Potentiale der Zertifizierung von KI-Systemen

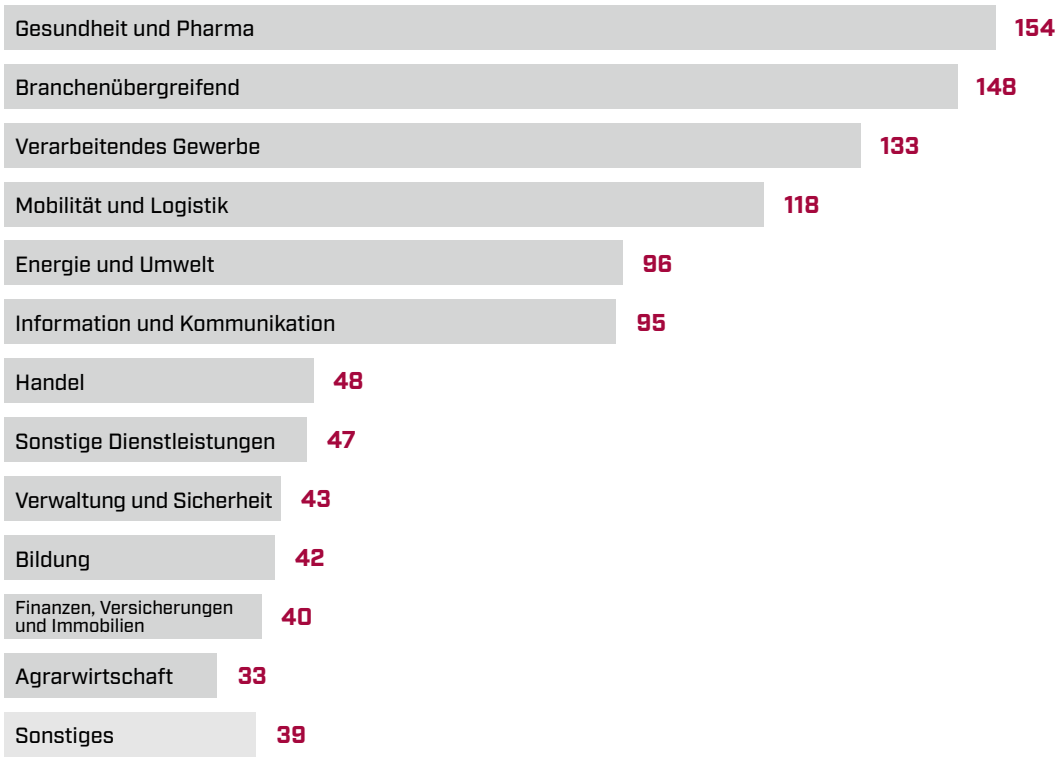
Der Einsatz von KI-Systemen verspricht einen hohen ökonomischen und gesellschaftlichen Nutzen. Die Plattform Lernende Systeme gibt unter anderem in Anwendungsszenarien einen ersten Ausblick auf mögliche Nutzengewinne in den Bereichen Mobilität, Medizin, Mensch-Maschine-Interaktion und Lebensfeindliche Umgebungen (siehe PLS 2019a, b, c, d). Einen Überblick über den zu erwartenden ökonomischen Nutzen gibt der AG-Bericht der Arbeitsgruppe Geschäftsmodellinnovationen (siehe PLS 2019e). Zertifizierung ist in vielen Fällen notwendig, damit KI-Systeme in die Anwendung gelangen und ihre Nutzenversprechen erfüllen.

Zertifizierung ist eine von mehreren Maßnahmen hin zum Einsatz von vertrauenswürdigen KI-Systemen. Sie ist eine Grundlage, um vor allem bei sehr komplexen Systemen Vertrauen in die Technik wie auch in die Hersteller und Anbieter beziehungsweise Bereitsteller aufzubauen. Mit Hilfe von Zertifizierung können auf der einen Seite negative Konsequenzen vermieden und auf der anderen Seite die Entwicklung von KI-Systemen im Sinne des Gemeinwohls begünstigt werden. Die Vermeidung potentieller unerwünschter Konsequenzen bei ordnungsgemäßem Einsatz ist eine Minimalanforderung an eine gute Zertifizierung – auch wenn durch Zertifizierung nie alle möglichen negativen Konsequenzen ausgeschlossen werden können. Zertifizierung kann jedoch auch Anreize schaffen, um über Minimalanforderungen hinaus KI-Systeme ethisch reflektiert und technisch zuverlässig zu entwickeln und einzusetzen. Da eine Zertifizierung die Vergleichbarkeit von KI-Systemen erhöht, besteht das Potential für einen fairen marktwirtschaftlichen Wettbewerb. Wenn ein Zertifizierungsprozess allerdings mangelhaft ist, kann das dazu führen, dass Potential blockiert wird. Die Frage nach der Effizienz der Zertifizierung von KI-Systemen lässt sich derzeit noch nicht für alle Einsatzgebiete abschließend beantworten. In manchen Branchen muss in einem ersten Schritt ein geeigneter Ansatz für die Zertifizierung von KI-Systemen entwickelt und erprobt werden, um in einem nächsten Schritt den Nutzen der Maßnahmen beurteilen zu können.

3 Notwendigkeit der KI-Zertifizierung

Heutzutage wird bereits in vielen Bereichen die Schlüsseltechnologie Künstliche Intelligenz eingesetzt (siehe Abbildung 1, Seite 4). Um das volle Nutzenpotential ausschöpfen zu können, sollen KI-Systeme aber in Zukunft noch häufiger eingesetzt werden. Eine wichtige Grundlage hierfür kann eine Qualitätsüberprüfung durch Dritte sein – eine Überprüfung, für die aktuell nur wenige übergeordnete Vorgaben existieren. Die Notwendigkeit einer Zertifizierung ergibt sich somit zum einen aus den Systemen selbst und zum anderen aus den ökonomischen und gesellschaftlichen Implementierungszielen für KI. Folgendes Kapitel gibt einen Überblick über die Notwendigkeit einer gut ausgestalteten Zertifizierung.

KI-Anwendungen nach Branchen (Übersicht aus KI-Landkarte)



Quelle: PLS: 2020 (Stand: März 2020, absolute Zahlen, Mehrfachzuordnung war möglich, N=720)

KI-spezifische Notwendigkeit

Die Notwendigkeit der Zertifizierung von KI-Systemen liegt in vielen der modernen KI-Systeme selbst begründet. Sie sind keine klassischen IT-Systeme, sondern sie sind oftmals dynamisch, das heißt, sie entwickeln sich (unter Umständen auch unvorhersehbar) weiter. Auch wenn das System am Ende des Lernvorgangs „eingefroren" wird, kann sich die Betriebsumgebung ändern (*concept drift*). Hinzu kommt, dass die von einem KI-System ausgegebenen „Entscheidungen" beziehungsweise Resultate oft probabilistischer Natur sind, sie nicht auf Kausalitäten, sondern auf statistischen Korrelationen beruhen, also unter Umständen fehlerbehaftet sein können (Fehlschlüsse), und zudem häufig schwer nachvollziehbar sind (Black-Box-Problem). Aufgrund dieser Spezifika ist aktuell offen, wie bisherige Zertifizierungsprozesse auf KI-Systeme angewendet werden können. Zum gegenwärtigen Zeitpunkt werden KI-Systeme vor allem als Werkzeug eingesetzt; in Zukunft aber werden autonome, auf KI basierende Systeme immer stärker an Relevanz gewinnen. Das führt dazu, dass der Einsatz von KI-Systemen die Interaktion zwischen Mensch und Technik noch stärker verändern wird. Der Einsatz von KI-Systemen führt also zu zahlreichen neuen Fragen und Herausforderungen, aber auch zu einem immensen Nutzenversprechen.

Zertifizierung als Voraussetzung für ökonomische und gesellschaftliche Ziele

Zertifizierung von KI-Systemen ist in vielen Fällen als eine der Schlüsselvoraussetzungen notwendig, um KI-Systeme in die Anwendung zu bringen, während in einigen Einsatzgebieten der KI möglicherweise auch keine oder lediglich eine weniger strikte Zertifizierung notwendig sein wird. Bei einer Nutzung von Zertifizierung für die Erreichung ökonomischer und gesellschaftlicher Ziele sind die folgenden drei Effekte von hervorgehobener Bedeutung:

- In manchen (besonders sensiblen und kritischen) Bereichen wird Zertifizierung eine **Zugangsvoraussetzung** werden (bzw. ist es heutzutage schon). Dies kann entweder in Form einer expliziten Zugangsbeschränkung erfolgen oder aber auch durch den öffentlichen Diskurs, der dazu führt, dass eine Zertifizierung zu einer Zugangsvoraussetzung werden kann: Wenn die Forderung nach einer Zertifizierung häufig und explizit geäußert wird, ist dies für Anwenderinnen und Anwender ein starker Anreiz, zertifizierte KI-Systeme einzusetzen. Einheitliche Vorgaben und Methodiken würden dabei unterstützen, KI-Systeme insgesamt in die Anwendung zu bringen.
- KI-Systeme können **gesellschaftliche und politische Konsequenzen** mit hoher Tragweite nach sich ziehen. Ein Beispiel ist das Problem der potentiellen Diskriminierung durch KI-Systeme, sei es bei Bewerbungsverfahren oder im Marketing (siehe Beck et al. 2019). KI-Systeme können also ethische oder gesellschaftlich relevante Fragestellungen berühren. Da diese Fragen oftmals nicht im Fokus der ökonomischen Perspektive stehen, müssen sie im Rahmen eines Zertifizierungsverfahrens und oftmals auch rechtlich adressiert werden. Eine Zertifizierung von KI-Systemen kann hier korrigierend und regulierend wirken.
- Grundlage für den Einsatz von KI-Systemen ist **Vertrauen**. Der Anreiz, neue KI-Anwendungen zu entwickeln, hängt davon ab, ob mit diesen Anwendungen die Erwartung verknüpft werden kann, dass potentielle Nutzerkreise existieren oder generiert werden können. Hierfür ist jedoch auf der Nutzerseite ein gewisses Vertrauen gegenüber solchen Anwendungen und den dahinter liegenden Akteuren und Prozessen notwendig. Zertifizierung kann hier vertrauensbildend wirken. Teil dieses Vertrauens kann das Wissen sein, gegebenenfalls Regressansprüche geltend machen zu können beziehungsweise Anspruch auf die Kompensation eines Schadens zu haben. Zertifizierung kann ein wichtiger Baustein für die **Regelung solcher Verantwortungs- und Haftungsfragen** sein.

4 Nutzen von KI-Zertifizierung

Der größte potentielle Nutzen einer Zertifizierung von KI-Systemen besteht darin, dass sie den Aufbau von Vertrauen in solche Systeme und in die mit ihnen verbundenen Akteure und Prozesse stärken kann. Das Vertrauen geht aus einer sachgerechten Prüfung durch Expertinnen und Experten im Rahmen eines Zertifizierungsprozesses hervor, aus der Reputation der üblicherweise akkreditierten zertifizierenden Organisation und aus den strukturierten und klaren Kriterien, die angelegt werden, aber auch aus dem Orientierungsangebot, das die Zertifizierung Akteuren anbietet. So können bei Nicht-Expertinnen und -Experten, Anwenderinnen und Anwendern, politischen Entscheidungsträgerinnen und Entscheidungsträgern sowie Behörden Unsicherheiten und Ängste beziehungsweise Bedenken gegenüber KI-Technologien abgebaut werden. Die Zertifizierung kann daher etwa die Aufgeschlossenheit von Beschäftigten in Betrieben gegenüber dem Einsatz von KI-Systemen fördern. Im geschäftlichen Kontext kann eine Zertifizierung von KI-Systemen vor allem bei Anbietern von Komponenten für KI-Systeme Vertrauen schaffen. Vertrauen ist ebenfalls eine wertvolle Ressource im (internationalen) Wettbewerb. Die Zertifizierung selbst kann wiederum zu einer Wettbewerbsdynamik führen, die Anreize setzt, bessere Produkte im Sinne der zugrundeliegenden Kriterien zu entwickeln, und somit in der Anwendungspraxis wiederum das Vertrauen in die Zertifizierung erhöhen. Im besten Fall wird ein wechselseitig vertrauensbildender Kreislauf angestoßen, der dabei hilft, die Potentiale von KI-Systemen für die Gesellschaft auszuschöpfen.

Prinzipien als Grundlage der Zertifizierung

Die Zertifizierung von KI-Systemen kann dazu führen, dass KI-Anwendungen und KI-Produkte bestimmte gesellschaftliche und ökonomische Kriterien erfüllen. Der Nutzen von Zertifizierung liegt also darin, dass die Umsetzung dieser Werte eingefordert werden kann. Es handelt sich hierbei um Prinzipien wie Rechtssicherheit (etwa bezüglich Haftung und Entschädigung), Verantwortlichkeit, Interoperabilität, Transparenz, Verständlichkeit der KI-Systeme, (Produkt-)Sicherheit und IT-Sicherheit (z. B. safety, security und robustness), Datenschutz oder Selbstbestimmung – sei es auf individueller oder gesellschaftlicher Ebene. Mit der Zertifizierung kann zudem einer „wilden" Standardisierung, die durch das Handeln von Unternehmen oder Staaten mit anderen Wertevorstellungen entstehen kann, entgegengewirkt werden. Die Gewissheit darüber, dass eine KI-Anwendung bestimmte wünschenswerte Kriterien erfüllt, kann wiederum die Akzeptanz und das Vertrauen in KI-Systeme fördern.

Orientierung für Akteure durch Prinzipien

Wird durch die Zertifizierung bestätigt, dass bestimmte vorgegebene Kriterien und Prinzipien eingehalten werden, bietet dies Akteuren aus verschiedenen gesellschaftlichen Sektoren Orientierungspunkte im Umgang mit KI-Systemen. Das erleichtert die Beurteilung und Vergleichbarkeit unterschiedlicher Anwendungen, beziehungsweise macht dies unter Umständen sogar erst möglich. Je strukturierter und klarer ein Zertifizierungssystem ist, desto vorausschauender und sicherer können Hersteller, Anbieter, Bereitsteller und Nutzende von KI-Systemen handeln. So können die Verständlichkeit und Akzeptanz von KI-Systemen erhöht werden. Potentiellen Nutzerinnen und Nutzern fehlt häufig die notwendige Expertise, um die Funktionsweisen und Auswirkungen von KI-Systemen beurteilen zu können. Die Zertifizierung kann Entscheidungen im Einkaufsprozess von KI-Systemen erleichtern. Zudem können Konsumierende und andere Nutzende sich über die Zertifizierungskriterien eine gewisse KI-Kompetenz aneignen, indem sie die Qualitätsanforderungen an KI-Systeme kennenlernen und sie miteinander vergleichen und letztendlich selbst beurteilen können, was ein qualitativ wertiges KI-System auszeichnet. Entwicklerinnen und Entwicklern von KI-Systemen bietet die Zertifizierung Kriterien, die sie von Beginn an mitdenken und berücksichtigen können, ebenso wie KI-Forschern (vor allem in der anwendungsorientierten Forschung). Hierdurch werden zudem neue und anspruchsvolle Forschungsgegenstände entdeckt und es können von Beginn an Systemlösungen entwickelt werden, die aufgrund ähnlicher Standards zueinander passen. Für Unternehmen bietet die Zertifizierung von KI-Systemen ebenfalls Orientierung, um Unternehmensentscheidungen zu planen und umzusetzen.

Prinzipien und Standards im marktwirtschaftlichen Wettbewerb

Ein weiterer Mehrwert der Zertifizierung von KI-Systemen ist der wirtschaftliche Vorteil und die Wirkung am Markt. Eine vertrauenswürdige Marke „KI made in Europe", die einen bestimmten Katalog an zu definierenden vorgegebenen Kriterien erfüllt, kann einen Vorteil im internationalen Wettbewerb bieten. Die Zertifizierung von KI-Systemen kann ein Alleinstellungsmerkmal gegenüber chinesischen und US-amerikanischen Anbietern darstellen und so den Pfad für einen „europäischen Weg" ebnen. Aus Herstellersicht ist damit die Hoffnung auf eine Wertsteigerung der entwickelten KI-Produkte und einen größeren Absatz verbunden. Da der Kriterienkatalog eine gewisse Vergleichbarkeit zwischen KI-Systemen herstellt, kann dies den nationalen und internationalen Wettbewerb „befeuern". Im besten Falle entsteht weltweit eine entsprechende Nachfrage und ein Anpassungsdruck in Bezug auf die Kriterien. Somit könnten – aus europäischer Perspektive betrachtet – bessere und kriterienkonforme Produkte entstehen und sich weltweit durchsetzen.

5 Überblick über bestehende Projekte

Wie bereits dargestellt, existieren Stand heute (März 2020) keine gültigen Standards und Richtlinien für eine Konformitätsbewertung von KI-Systemen. Jedoch beschäftigen sich bereits eine Reihe von Projekten und Initiativen mit der Erarbeitung von Normen, Standards und Leitlinien für die Zertifizierung von KI-Systemen. In diesem Abschnitt wird ein Überblick über einige dieser Unternehmungen gegeben.

Die im folgenden Teil vorgestellten Projekte haben sich bereits im Rahmen eines Runden Tisches zur Zertifizierung der Plattform Lernende Systeme 2020 im März präsentiert. Die Diskussion auf der Veranstaltung zeigte, dass aktuell nicht parallel gearbeitet wird, ein regelmäßiger Austausch untereinander jedoch sehr wichtig ist. Im Laufe der Veranstaltung wurden zwei Arten von Projekten identifiziert:[1]

Projekte zur Meta-Ebene: Diese Projekte werfen einen umfassenden Blick auf den aktuellen Ist-Stand und formulieren Empfehlungen. Zielgruppe sind die politischen Entscheidungsträgerinnen und Entscheidungsträger sowie die fachlich interessierte Öffentlichkeit. Ein Beispiel ist die KI-Normungsroadmap (siehe Tabelle 1).

Tabelle 1 DIN e. V.: KI-Normungsroadmap

Laufzeit	Oktober 2019 – Dezember 2020. Die Normungsroadmap KI wird anschließend regelmäßig fortgeschrieben und weiterentwickelt.
Beteiligte	Das DIN leitet das Projekt, durchgeführt wird es gemeinsam mit dem Bundesministerium für Wirtschaft und Energie (BMWi) und DKE. Zahlreiche Expertinnen und Experten aus der Wirtschaft, Wissenschaft, Politik und Zivilgesellschaft sind beteiligt.
Ziel	Ziel ist die frühzeitige Entwicklung eines Handlungsrahmens für die Normung und Standardisierung. Die Roadmap wird eine Übersicht über bestehende Normen und Standards zu KI-Aspekten umfassen und Empfehlungen im Hinblick auf noch notwendige künftige Aktivitäten geben.
Arbeitsweise/ -vorgehen	Multi-Stakeholder-Prozess auf der Basis von Arbeitsgruppen: • AG Grundlagen (Begriffe, Klassifikationen, Entwicklung von KI, Vertrauenswürdigkeit) • AG Ethik/Responsible AI • AG Qualität & Zertifizierung • AG IT-Sicherheit bei KI-Systemen • AG industrielle Automation • AG Mobilität & Logistik • AG KI in der Medizin
Ergebnisse	Die Ergebnisse des Prozesses werden im Dezember 2020 zum Digitalgipfel veröffentlicht.
Weitere Informationen	https://www.din.de/de/forschung-und-innovation/themen/kuenstliche-intelligenz/normungsroadmap-ki

1 Es ist zu beachten, dass die Projekte schwerpunktmäßig den jeweiligen Projekttypen zugeordnet wurden, da sie nur schwer absolut trennscharf in die verschiedenen Kategorien einteilbar sind.

Projekte zur Definition von technischen und nicht-technischen Prüfkriterien: Hierbei handelt es sich um die Definition von Anforderungen: einerseits an die Prüf- und Bewertungstätigkeit bei der Prüfung von KI-Systemen sowie an die KI-Systeme selbst. Zudem beschäftigen sie sich (teilweise) mit dem Prozess der Zertifizierung von KI-Systemen. Sie stellen also die Frage: Wie kann die Zertifizierung von KI-Systemen in einem Bereich konkret gelingen? Die Zielgruppe sind Herstellerinnen und Hersteller sowie Anwenderinnen und Anwender, die fachlich interessierte Öffentlichkeit sowie politische Entscheidungsträgerinnen und Entscheidungsträger. Beispiele hierfür sind das Projekt Zertifizierung zur Sicherstellung einer vertrauenswürdigen KI (siehe Tabelle 2), das Projekt der AI Ethics Impact Group (siehe Tabelle 3) und das KI-Observatorium (siehe Tabelle 4).

Tabelle 2 CERTIFIED AI: Zertifizierung zur Sicherstellung einer vertrauenswürdigen KI

Laufzeit	Das Programm ist auf fünf Jahre angelegt.
Beteiligte	Von BSI und Fraunhofer IAIS geleitet, weitere Forschungsinstitutionen und Industriepartner sowie die Kompetenzplattform KI.NRW sind beteiligt. Generell wird ein nationaler Beteiligungsprozess angestrebt.
Ziel	Ziel des Programms ist die standardisierungsreife Entwicklung von technischen Prüfkriterien, Prüfmethoden und Prüfwerkzeugen für KI-Anwendungen und die Etablierung der erforderlichen Prüfinfrastruktur im Bundesgebiet. Parallel dazu wird die internationale Harmonisierung eingeleitet.
Arbeitsweise/ -vorgehen	Bidirektionaler, iterativer und breiter Beteiligungsprozess: Top-down: Die Entwicklung eines Prüf-Frameworks (Kriterien, Methoden, Werkzeuge, Strukturen), das die Vergleichbarkeit von Prüfungen mit hoher Qualität für unterschiedliche KI-Anwendungen leistet. Bottom-up: Die Anwendung von spezifischen Prüfgrundlagen in Pilotprojekten (Use Cases nach der KI-Normungsroadmap). Hierzu werden Anwenderkreise mit Industriebeteiligung betreut.
Ergebnisse	1) Konzept- und Positionspapiere für die Handlungsfelder vertrauenswürdiger KI und für Aufbau, Funktionsweise und Kooperation des Programms im Kontext des Standardisierungsprozesses. 2) Prüfansatz in Anlehnung an internationale Standardprüfverfahren wie zum Beispiel die „Common Criteria for Information Technology Security Evaluation" als Diskussionsgrundlage. 3) Abstimmung mit Akteuren des Beteiligungsprozesses und Initialisierung der Programmstrukturen.
Weitere Informationen	https://www.iais.fraunhofer.de/ki-zertifizierung

Tabelle 3 AI Ethics Impact Group: Integrierter Ansatz zur wirksamen und normungsfähigen Verankerung von Ethik beim Einsatz von Künstlicher Intelligenz

Laufzeit	September 2019 – April 2020
Beteiligte	VDE, Bertelsmann Stiftung, Hochleistungsrechenzentrum Stuttgart, Institut für Technikfolgenabschätzung und Systemanalyse (KIT), Internationales Zentrum für Ethik in den Wissenschaften, Universität Tübingen, iRights.Lab, Universität Kaiserslautern
Ziel	Entwicklung eines Ethics Ratings für KI-Systeme, um KI-Ethik messbar zu machen.
Arbeitsweise/ -vorgehen	Konsortium von Wissenschaftlerinnen und Wissenschaftlern arbeitet zusammen, um das Rating anhand von Methoden der Technikfolgenabschätzung und der Ingenieurethik zu entwickeln.
Ergebnisse	Ethik-Kennzeichnungen, die sich an der Energieeffizienzkennzeichnung von Haushaltsprodukten orientieren. Für diese Kennzeichnungen sind „Ethics Ratings“ entlang der Werte Transparenz, Accountability, Privatsphäre, Gerechtigkeit, Verlässlichkeit und Nachhaltigkeit vorgesehen. Die Werte werden über eine Indikatoren-basierte Systematik von Kriterien gemessen und über quantitative und qualitative Observablen fundiert. Die Ethics Ratings basieren auf einer Risiko-Matrix für KI-Anwendungen, die den Bedarf für eine Zertifizierung identifiziert. Die Ergebnisse werden im Mai 2020 in Brüssel vor Vertreterinnen und Vertretern der EU-Kommission vorgestellt.
Weitere Informationen	https://www.vde.com/de/presse/pressemitteilungen/vde-ethik-kennzeichnung-ki

Tabelle 4 Denkfabrik Digitale Arbeitsgesellschaft: KI-Observatorium

Laufzeit	Eröffnet März 2020
Beteiligte	Das KI-Observatorium wird von der „Denkfabrik Digitale Arbeitsgesellschaft" des Bundesministeriums für Arbeit und Soziales betreut (BMAS).
Ziel	Das Observatorium will die Auswirkungen von KI auf Leben und Arbeit in Deutschland monitoren und versteht sich als „Kartograph" eines neuen technologischen Ökosystems. Leitfragen: • Wie prägt Künstliche Intelligenz (KI) die Arbeitswelt? • Wie verändert sie unsere Gesellschaft? • Welche Leitlinien brauchen wir für eine sichere, transparente und nachvollziehbare KI? • Wie gelingt ein menschenzentrierter und partizipativer Einsatz von KI in der Arbeitswelt und darüber hinaus?
Arbeitsweise/ -vorgehen	Je nach Handlungsfeld nationale und internationale Zusammenarbeit mit externen Expertinnen und Experten unter anderem aus Gewerkschaften, Unternehmen, Wissenschaft, Zivilgesellschaft. Interne themenbezogene Kooperation mit anderen Projektteams der Denkfabrik und den Fachabteilungen des BMAS. Die Zusammenarbeit findet zum Beispiel in Form von zeitlich begrenzten Labs, Workshops oder Projektgruppen statt.
Ergebnisse	Regelmäßige Publikation von Ergebnissen
Weitere Informationen	https://www.denkfabrik-bmas.de/die-denkfabrik/ki-observatorium

6 Herausforderungen der KI-Zertifizierung

KI-Systeme stellen die Zertifizierung vor große Herausforderungen. Sie bestehen etwa darin, das richtige Maß von Zertifizierung zu finden, um mögliche Kosten für Unternehmen gering zu halten und damit Markteintrittsbarrieren zu verhindern. Weiterhin resultieren sie aus den Spannungsverhältnissen, die sich aus den Eigenschaften von KI-Systemen ergeben, aber auch aus den Ansprüchen, die an eine Zertifizierung von KI-Systemen angelegt werden können. So ergeben sich Herausforderungen für die KI-Zertifizierung beispielsweise aufgrund der Dynamik moderner Lernender Systeme, der anzulegenden Maßstäbe und Metriken oder hinsichtlich der Markt- und Innovationsfähigkeit. Folgendes Kapitel stellt diese Herausforderungen anhand der Spannungsverhältnisse dar.

Statik der Zertifikate und Dynamik der KI und ihres Umfeldes

Während eine Zertifizierung generell nur Momentaufnahmen darstellen kann, sind viele moderne KI-Systeme von Dynamik geprägt, ebenso wie ihr Umfeld. Da diese Form der Veränderbarkeit in bisherigen Zertifizierungssystemen nicht in diesem Ausmaß mitberücksichtigt werden musste, fehlt es an wichtigen Erfahrungswerten.

„Verhaltensänderungen" von KI-Systemen sind grundsätzlich möglich und nicht immer vorhersehbar. KI-Systeme dahingehend zu überprüfen und zu verifizieren, ob sie genau das machen, was sie machen sollen, ist daher eine grundlegende Herausforderung. Dies gilt vor allem dann, wenn der Input für die Systeme komplex ist oder durch systematische und permanente Fehler geprägt ist.

Da sich die Kontexte, in denen KI-Systeme lernen, unvorhersehbar verändern können und es zugleich auch schwierig ist, umfassend adaptive Lernende Systeme zu entwickeln, sind die möglichen Risiken (z. B. Diskriminierung, Filterblasen etc.) oft schwer im Vorhinein zu bestimmen. Dies liegt auch daran, dass die Modelle, die aus den Algorithmen hervorgehen, Verallgemeinerungen beziehungsweise Vereinfachungen darstellen und so beispielsweise komplexe soziale Handlungen nicht vollständig und auch nicht immer adäquat abbilden. Es ist zudem schwierig vorherzusagen oder zu steuern, welche Daten das Lernende System in der Einsatzzeit erhält und welche Veränderungen und Fehlfunktionen somit nach der Testphase auftreten können. Daten können sich beispielsweise verändern, weil sich die Umwelt oder der Nutzungszusammenhang des KI-Systems verändert, aber auch, weil unter Umständen Sensoren aufgrund von Verschleißerscheinungen schlechte Daten liefern. Da die Forschung zu KI und zu Technologien, auf denen KI basiert, ebenfalls ein dynamisches Feld ist, kann auch dies für die Zertifizierung zur Herausforderung werden.

Dynamik von Produkten und ihrem Umfeld ist zwar kein Phänomen, das ausschließlich auf KI-Systeme begrenzt ist. So können beispielsweise regelmäßige Software-Updates ebenfalls zu einer Art neuem Produkt führen und auch für nicht KI-basierte Systeme können Störungen der Sensorik problematisch sein. Die Frage nach dem Verhältnis von Dynamik und Zertifizierung stellt sich allerdings bei modernen, selbst- und weiterlernenden KI-Systemen in einem neuen Licht dar. Grund ist einerseits die Unsicherheit über mögliche Risiken, die durch die probabilistische Natur und mangelnde Nachvollziehbarkeit von KI-Systemen bedingt wird. Andererseits kann es notwendig werden zu bestimmen, wie und auf welche Weise sich die Systeme verändert haben, um einen praktikablen Ansatz für die Zertifizierung zu entwickeln.

Allgemeine Zertifizierung vs. umfassende, kleinteilige Zertifizierung

Sowohl die Wahl der Ansatzpunkte für die KI-Zertifizierung als auch die Wahl der Maßstäbe und Metriken, die angelegt werden, stellen eine komplexe Aufgabe dar. Soll eine KI-Zertifizierung jedem Verfahren, jedem System und jeder Branche Rechnung tragen, so führt dies unter Umständen zu einer Vielzahl unterschiedlich starker Zertifikate mit verschiedenen Kriterien und zu einer Vielzahl von Zertifizierungsstellen. Die Herausforderung besteht demnach darin, ein allgemeines Prüfsystem zu entwickeln, um die Zertifizierung hochgradig unterschiedlicher KI-Systeme mit unterschiedlichen Einsatzkontexten vergleichbar zu gestalten, ohne zu allgemeine oder zu heterogene Anforderungen heranzuziehen. Dabei sollten neben den KI-spezifischen Anforderungen auch bestehende Standards, Prüfverfahren und -infrastrukturen miteinbezogen werden, sodass die Lücke zwischen etablierten Normen und Standards und tatsächlich neuen Anforderungen geschlossen werden kann. Allerdings kann ein allgemeines Prüfwerkzeug nicht jeden spezifischen Fall berücksichtigen, sodass sich viele Anforderungen lediglich in spezifischen Anwendungsfällen (Use Cases) als gestaltbar darstellen. Im Folgenden werden die Hintergründe für das skizzierte Spannungsverhältnis mit Blick auf Ansatzpunkte, Maßstäbe und Metriken genauer betrachtet.

Wahl der Ansatzpunkte

Die Zertifizierung kann an unterschiedlichen Punkten ansetzen, seien es Spezifika der KI-Systeme selbst oder Anforderungen, die aus gesellschaftlicher oder sektoraler Sicht an die Systeme angelegt werden.

1. Die Schwierigkeit liegt darin, zu definieren, was eine Künstliche Intelligenz darstellt. Bestehende Definitionen sind oftmals unterschiedlich und zu abstrakt gehalten, um sich für die Praxis zu eignen. Ebenso ist es schwierig zu bestimmen, wo bei KI-Systemen die Systemgrenzen liegen, das heißt, was das System umfasst und deshalb Gegenstand der Zertifizierung sein soll. Dies ist auch deshalb eine schwierige Frage, weil zum einen bei den Datensätzen, Algorithmen und beim Systemverhalten angesetzt werden könnte, zum anderen aber auch umfassender bei ganzen Anwendungen oder Diensten, der Systemarchitektur oder dem gesamten sozio-technischen

Ensemble eines KI-Systems. Dieses Problem wird verschärft, weil KI sehr variantenreich in physische Systeme und in Softwaresysteme integrierbar ist und in der Praxis häufig als hybrides System umgesetzt wird (vgl. hybride KI-Systeme, siehe Infobox).

2. Die Aufgabenstellung, einen adäquaten Ansatzpunkt für die Zertifizierung zu finden, wird noch schwieriger, wenn unter Umständen unterschiedliche Kategorien von Diensten definiert werden müssen. So kann es nötig sein, dass die Zertifizierung beispielsweise je nach Risiko oder Reichweite der Konsequenzen, die ein KI-System zeigen könnte, bei unterschiedlichen Klassen beziehungsweise Kategorien von KI-Systeme ansetzen muss. Wie tiefgreifend eine Zertifizierung umgesetzt wird, ist dann beispielsweise von einer Risikoeinschätzung abhängig, sodass in sehr kritischen Fällen eine umfassende Zertifizierung notwendig würde, während in Fällen mit niedrigem Risiko eine weniger tiefgreifende beziehungsweise unter Umständen auch keine Zertifizierung notwendig würde. Die Definition von Risikoklassen ist hierfür eine potentiell folgenreiche Voraussetzung und entsprechend anspruchsvoll. Für die jeweiligen Klassen müssten entsprechend bestimmte Assurance Level (Prüftiefen) formuliert werden. Beispiele für Prüftiefen finden sich schon in IT-Sicherheit (vgl. Common Criteria).

3. Diese Aufgabe wird noch umfassender, wenn die Zertifizierung nach unterschiedlichen Branchen und Bereichen und deren Spezifika differenziert werden muss, da es nicht möglich ist, ein Zertifikat für alle gesellschaftlichen Sektoren übergreifend zu definieren. So bestehen schon heute zum Beispiel für Medizinprodukte, die Luft- und Raumfahrt, aber auch in der Produktion und der Automotive Branche zahlreiche Normen und Standards, die auch bei der Zertifizierung von KI in solchen Einsatzgebieten berücksichtigt werden müssen.

Hybride KI-Systeme

Systeme, die auf einer Kombination verschiedener KI-Verfahren basieren oder auch nicht KI-basierte Verfahren mit KI verbinden, wie z. B. unterschiedliche Verfahren des Maschinellen Lernens, unterschiedliche KI-Ansätze oder etwa konventionelle Software und Lernverfahren. So können etwa wissensbasierte Ansätze mit datengetriebenen Ansätzen kombiniert werden. Hierbei nutzt das hybride System zum einen menschliches Wissen über die Welt, um daraus Schlussfolgerungen zu ziehen, und zum anderen Lernverfahren, die auf statistischen Zusammenhängen beruhen. Ein weiteres Beispiel sind regelbasierte Systeme, die ihre Regeln selbst lernen. Schließlich sind in der Praxis Maschinelle Lernverfahren häufig in konventioneller Software, wie etwa grafischen Benutzeroberflächen, integriert, sodass eine Interaktion zwischen KI und nicht KI-basierten Lösungen stattfindet.

Wahl der Maßstäbe und Metriken

Herausforderungen können auch die Wahl der anzulegenden Maßstäbe und der tolerierten Systemantworten sowie die Definition von Schwellenwerten und Metriken für bestimmte Prüfkategorien sein. Je nach Lernverfahren (z. B. überwachtes, unüberwachtes Lernen oder verstärkendes Lernen, siehe Infobox Seite 13) oder auch Branche kann es durchaus schwierig sein, sinnvolle Metriken anzulegen, weil unter Umständen die Systemantworten nicht mehr intuitiv nachvollziehbar sind, die Komplexität des Algorithmus zu hoch ist oder die spezifische Branche besonders strenge und hohe Anforderungen an das KI-System stellt (z. B. Gesundheit). Die Kompatibilität mit ethischen und sozialen Werten und Normen ist nur schwer über Technologie und somit auch durch die Zertifizierung von KI-Systemen herstellbar. Dies kann in bestimmten Branchen ein Problem darstellen. Zwar können manche gesellschaftlichen Problemstellungen möglicherweise durch eine technische Implementierung von Policies gelöst werden, etwa um einem missbräuchlichen Gebrauch von KI-Systemen vorzubeugen, jedoch muss auch hier der technische Durchsetzungsmechanismus in der Einsatzzeit stets funktionsfähig bleiben. Schließlich kann in einigen Fällen die Umsetzung bestimmter Kriterien der Zertifizierung schwierig werden, da einige Problemfelder in der Forschung selbst noch nicht geklärt sind, wie etwa die Erklärbarkeit und Nachvollziehbarkeit

bestimmter Lernverfahren, theoretische Grundlagen des Tiefen Lernens (vgl. Kersting/Tresp 2019: Seite 4, 11) oder die Verifikation und Validierung solcher Verfahren (vgl. Deep Learning, siehe Infobox).

Markt- und Innovationsfähigkeit vs. umfassende, kleinteilige Zertifizierung

Der Nutzen der Zertifizierung von KI-Systemen hängt entscheidend davon ab, ob ein richtiges Maß für eine EU-weite Zertifizierung gefunden werden kann: Eine umfassende, kleinteilige Zertifizierung könnte die bestmöglichen Standards garantieren, allerdings könnte die Zertifizierung dadurch zu einem Hemmnis im Markt und für Innovation werden. So können die (fixen) Kosten einer Zertifizierung vor allem für Start-ups und kleine und mittelständische Unternehmen eine Hürde für den Zugang zum Markt darstellen, da sie im Gegensatz zu großen, etablierten Unternehmen nicht über dieselben Ressourcen verfügen. Mit Blick auf die Wirtschaft ist zum gegenwärtigen Zeitpunkt ebenfalls schwer zu beurteilen, ob zertifizierte Produkte tatsächlich am Markt attraktiv sind. Dies hängt auch von der Qualität der Zertifizierung ab. Nicht-zertifizierte Produkte könnten billiger und schneller in der Markteinführung sein und trotzdem eine gute Qualität bieten. Der Aufwand einer KI-Zertifizierung könnte sehr hoch sein, etwa wenn diese auf der Code-Ebene ansetzt. Dies könnte die Marktfähigkeit und die Entwicklung innovativer Produkte beeinträchtigen, aber unter Umständen auch die Entstehung und den Erfolg neuer Geschäftsmodelle. Schließlich könnte sich eine zu ausufernde Zertifizierung auch in der Forschung als Hemmnis erweisen, wenn dadurch bestimmte Forschungsprojekte zu sehr erschwert würden. Somit ist aktuell noch offen, ob beziehungsweise inwieweit sich zertifizierte Produkte am Markt durchsetzen würden. Die Herausforderung ist es daher, eine marktfähige Zertifizierung zu entwickeln, welche Qualitätsstandards garantiert.

Tiefes Lernen (Deep Learning)

Methode des Maschinellen Lernens in künstlichen neuronalen Netzen. Diese umfassen mehrere Schichten – typischerweise eine Eingabe- und Ausgabeschicht sowie mehr als eine „versteckte" dazwischenliegende Schicht. Die einzelnen Schichten bestehen aus einer Vielzahl künstlicher Neuronen, die miteinander verbunden sind und auf Eingaben von Neuronen aus der jeweils vorherigen Schicht reagieren. In der ersten Schicht wird etwa ein Muster erkannt, in der zweiten Schicht ein Muster von Mustern und so weiter. Je komplexer das Netz (gemessen an der Anzahl der Schichten von Neuronen, der Verbindungen zwischen Neuronen sowie der Neuronen pro Schicht), desto höher ist der mögliche Abstraktionsgrad – und desto komplexere Sachverhalte können verarbeitet werden. Angewendet wird Deep Learning z. B. bei der Bild-, Sprach- und Objekterkennung sowie dem verstärkenden Lernen.

Überwachtes Lernen

Lernalgorithmen, die als Trainingsmaterial neben Rohdaten auch die erwarteten Ergebnisse erhalten. Weicht die Ausgabe des trainierten Modells vom gewünschten Ergebnis ab – wenn beispielsweise eine Tulpe als Rose identifiziert wird –, passt der Lernalgorithmus das Modell an. Ziel ist es, dem Netz durch unterschiedliche Ein- und Ausgaben die Fähigkeit anzutrainieren, selbst Verbindungen herzustellen.

Unüberwachtes Lernen

Wenn im Training eines Modells die Zielgrößen (z. B. Klassifikationslabels) nicht zur Verfügung stehen, spricht man von unüberwachtem Lernen. Ziel kann es sein, in einem großen, unstrukturierten Datensatz interessante und relevante Muster zu erkennen oder die Daten kompakter zu repräsentieren. So können beispielsweise Kundendaten nach Zielgruppen segmentiert werden, die man auf ähnliche Weise adressieren möchte.

Verstärkendes Lernen

Prozess, bei dem ein Lernendes System Entscheidungen trifft, auf deren Basis es anschließend handelt. Dazu verwendet es einen Algorithmus, der lernt, die Erfolgsaussichten der einzelnen Aktionen in den verschiedenen Situationen besser einzuschätzen. Für die gewählten Aktionen erhält es positives oder negatives Feedback. Ziel des Systems ist, möglichst viel positives Feedback zu erhalten. Beim Deep Reinforcement Learning werden dazu künstliche neuronale Netze als Modelle verwendet, die man erfolgreich in Spielen eingesetzt hat (z. B. Go, Poker, Atari).

7 Offene Fragen

Aus den dargestellten Herausforderungen bei der Zertifizierung von KI-Systemen ergeben sich einige offene Fragen, die für eine gelungene Zertifizierung von KI-Systemen zu klären sind:

- Welche Vorgaben existieren bereits, die auf die Zertifizierung von KI-Systemen angewendet werden könnten?
- Welche Kriterien sollten zertifiziert werden?
- Wann sollte die Zertifizierung erfolgen (Zeitpunkt, Risikoklasse)?
- Was sollte zertifiziert werden (Produkte, Prozesse, Personen)?
- Wie detailliert sollte zertifiziert werden?
- Wie sollte mit weiterlernenden KI-Systemen (Continuous Learning Systems, siehe Infobox) umgegangen werden?
- Wer sollte die Zertifizierung vornehmen?
- Wie könnte ein Standard entworfen werden, der europa- und im nächsten Schritt auch weltweit akzeptiert werden würde?
- Wo beziehungsweise in welchen KI-Anwendungsfeldern liegt der nationale wie internationale zukünftige Bedarf an Zertifizierung?

Diese Fragen bilden die Grundlage weiterer Aktivitäten und einer Anschlusspublikation der Plattform Lernende Systeme.

Continuous Learning Systems

Lernende Systeme, die in der Lage sind, ihre Vorhersagemodelle reibungslos zu aktualisieren, um verschiedene Aufgaben zu erfüllen und Datenverteilungen zu berücksichtigen, aber zugleich nützliches Wissen und Fähigkeiten im Lauf der Zeit wiederzuverwenden.

Literaturverzeichnis

Bundesregierung (2018): Strategie Künstliche Intelligenz der Bundesregierung. www.bmbf.de/files/Nationale_KI-Strategie.pdf (abgerufen am 31.03.2020).

Europäische Kommission (2020): White Paper on Artificial Intelligence – A European approach towards excellence and trust. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (abgerufen am 31.03.2020).

Kersting/Tresp (2019): Maschinelles und tiefes Lernen – Whitepaper aus der Plattform Lernende Systeme. https://www.plattform-lernende-systeme.de/publikationen.html (abgerufen am 31.03.2020).

Plattform Lernende Systeme (2019a): Anwendungsszenario – Intelligent vernetzt unterwegs. https://www.plattform-lernende-systeme.de/umfeldszenario-intelligent-vernetzt-unterwegs.html (abgerufen am 31.03.2020).

Plattform Lernende Systeme (2019b): Anwendungsszenario – Mit Künstlicher Intelligenz gegen Krebs. https://www.plattform-lernende-systeme.de/anwendungsszenario-onkologie.html (abgerufen am 31.03.2020).

Plattform Lernende Systeme (2019c): Anwendungsszenario – Information Butler fürs Büro. https://www.plattform-lernende-systeme.de/anwendungsszenarien.html (abgerufen am 31.03.2020).

Plattform Lernende Systeme (2019d): Anwendungsszenario – Schnelle Hilfe beim Rettungseinsatz. https://www.plattform-lernende-systeme.de/anwendungsszenario-rettungseinsatz.html (abgerufen am 31.03.2020).

Plattform Lernende Systeme (2019e): Neue Geschäftsmodelle mit Künstlicher Intelligenz – Bericht der Arbeitsgruppe Geschäftsmodellinnovationen. https://www.plattform-lernende-systeme.de/publikationen.html (abgerufen am 31.03.2020).

Plattform Lernende Systeme (2020): KI-Landkarte. https://www.plattform-lernende-systeme.de/ki-in-deutschland.html (abgerufen am 31.03.2020).

Über dieses Impulspapier

Beteiligte

Vorliegendes Impulspapier wurde auf der Basis von Experteninterviews mit Mitgliedern und Vertreterinnen und Vertretern der in der Plattform Lernende Systeme beteiligten Forschungseinrichtungen und Unternehmen erstellt. Federführend waren die Arbeitsgruppe Technologische Wegbereiter und Data Science und die Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik. Beteiligt waren darüber hinaus Mitglieder der Arbeitsgruppe Arbeit/Qualifikation, Mensch-Maschine-Interaktion, Mobilität und intelligente Verkehrssysteme und der Arbeitsgruppe Lebensfeindliche Umgebungen.

Autorinnen und Autoren

PD Dr. Jessica Heesen, Universität Tübingen
Prof. Dr. Jörn Müller-Quade, Karlsruher Institut für Technologie
Prof. Dr. Stefan Wrobel, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (IAIS)

Befragte Expertinnen und Experten

Prof. Dr. Jürgen Beyerer, Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung (IOSB)
Dr. Gunnar Brink, ROSEN Technology and Research Center GmbH
Jan-Henning Fabian, ABB AG
Dr. Tim Gutheit, Infineon Technologies AG
Dr. Martin Hoffmann, ABB AG
Dr. Detlef Houdeau, Infineon Technologies AG
Dr. Norbert Huchler, Institut für sozialwissenschaftliche Information und Forschung e.V.
Dr. Elsa Kirchner, Deutsches Forschungszentrum für Künstliche Intelligenz
Prof. Dr. Tobias Matzner, Universität Paderborn
Roland Norden, Robert Bosch GmbH
Dr. Matthias Peissner, Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO)
Dr. Christoph Peylo, Robert Bosch GmbH
Thomas Schauf, Deutsche Telekom AG
Dr. Sirko Straube, Deutsches Forschungszentrum für Künstliche Intelligenz
Oliver Suchy, Deutscher Gewerkschaftsbund

Die Erstellung des Impulspapiers wurde unterstützt durch
Dr. Maximilian Poretschkin, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (IAIS)
Stephanie Dachsberger, Geschäftsstelle der Plattform Lernende Systeme
Maximilian Hösl, Geschäftsstelle der Plattform Lernende Systeme

Redaktion
Dr. Ursula Ohliger, Geschäftsstelle der Plattform Lernende Systeme
Stephanie Dachsberger, Geschäftsstelle der Plattform Lernende Systeme
Maximilian Hösl, Geschäftsstelle der Plattform Lernende Systeme

Impressum
Herausgeber: Lernende Systeme – Die Plattform für Künstliche Intelligenz | Geschäftsstelle | c/o acatech | Karolinenplatz 4 | D-80333 München | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de | Folgen Sie uns auf Twitter: @LernendeSysteme |
Stand: April 2020 | Bildnachweis: Vertigo3d / iStock

Die Autorinnen und Autoren sind Mitglieder der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik und der Arbeitsgruppe Technologische Wegbereiter und Data Science der Plattform Lernende Systeme. Alle Publikationen der Plattform Lernende Systeme sind online verfügbar unter: https://www.plattform-lernende-systeme.de/publikationen.html

Empfohlene Zitierweise: Jessica Heesen et al. (Hrsg.): Zertifizierung von KI-Systemen – Impulspapier aus der Plattform Lernende Systeme, München 2020.

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

acatech
DEUTSCHE AKADEMIE DER TECHNIKWISSENSCHAFTEN