HANSEATIC BLOCKCHAIN INSTITUTE

KONRAD ADENAUER STIFTUNG

Synergien von Blockchain und KI

#KAS4 INNOVATION

kas.de

Synergien von Blockchain und KI

Auf einen Blick

- Die Studie plädiert für den Aufbau eines europäischen KI-Ökosystems, das durch einen souveränen Datenmarktplatz und föderiertes Lernen gestützt wird. Der vorgestellte Ansatz zielt darauf ab, die technologische Souveränität Europas im Zeitalter der KI zu stärken, indem eine Blockchain-basierte Vision für „KI Made in Europe“ skizziert wird.

- **Digitalisierung und Daten:** Die Nutzung der technischen Synergien von KI und Blockchain ermöglicht Unternehmen eine datenschutzkonforme Zusammenarbeit, die angesichts der zunehmenden Digitalisierung der Wirtschaft und der gestiegenen gesetzlichen Anforderungen an das Datenmanagement insbesondere in der Europäischen Union (EU) zu einem geschäftsentscheidenden Faktor geworden ist.

- **Unternehmen:** Insbesondere Start-ups sowie kleine und mittlere Unternehmen aus dem Mittelstand können von den KI-Blockchain-Synergien profitieren, da diese aufgrund hoher Marktbarrieren sowie Kosten für Datenschutz und KI-Training bisher weniger an den wirtschaftlichen Potenzialen von Künstlicher Intelligenz partizipieren als große Konzerne.

- **Digitaler Datenmarktplatz:** Durch den kombinierten Einsatz von Blockchain und KI-Anwendungen können dezentrale digitale Infrastrukturen geschaffen werden, die Individuen, Start-ups oder Unternehmen gleichermaßen nutzen können, um gemeinsam KI-Modelle zu trainieren und deren Ergebnisse zu nutzen oder sogar Algorithmen zur Datenverarbeitung auszutauschen, ohne dabei notwendigerweise die Inhalte der eigenen Daten preiszugeben. Solche digitalen Datenmarktplätze können von individuellen und/oder unternehmerischen Akteuren betrieben werden und den Aufbau öffentlich zugänglicher Datenmarktplätze ermöglichen.

- **Anwendungsfälle:** Mögliche Anwendungsfelder für digitale Datenmarktplätze auf Basis von Blockchain und KI sind unter anderem die Datennutzung im Gesundheitswesen, die Entwicklung von Smart Cities, industrielle KI-Anwendungen wie die Verbesserung maschineller Produktionsprozesse und die Nutzung mobiler Geräte für das KI-Training. Weitere denkbare Bereiche sind das *Metaverse* und die Rückverfolgbarkeit von Daten zur Überprüfung der Einhaltung von KI-Vorschriften.

Inhalt

Vorwort

Vorwort

Die deutsche Wirtschaft generell und insbesondere der Mittelstand – das Rückgrat der Gesellschaft – stehen zunehmend vor der Herausforderung, im globalen Wettbewerb um den Einsatz technischer Innovationen, wie KI, Schritt zu halten. Neben vielen Faktoren spielt insbesondere die Notwendigkeit zur technologischen Weiterentwicklung eine große Rolle. Der Einsatz neuer Technologien erfordert dabei jedoch häufig großen Personal- oder Ressourcenaufwand und ist für viele Mittelständler nicht praktikabel. Auch die deutsche Start-up-Szene hat es schwer, mit den neuesten technologischen Entwicklungen Fuß zu fassen.

Sinnvolle und notwendige Regulierungen zum Datenschutz und zur Entwicklung digitaler Produkte schränken die Innovationsfreudigkeit ein, steigern die operationalen Kosten für junge Unternehmen und können mit einer Abwanderung von Ideen in andere Wirschaftsräume korrelieren.

Wie dieser Bericht zeigt, bieten die Zukunftstechnologien KI und Blockchain dabei vielversprechende Grundlagen, um Start-ups und den Mittelstand innerhalb der EU in ihren technologischen Zielen zu fördern und datenschutzkonforme Kollaborationen möglich zu machen. Damit hätten sie das Potenzial, die Zusammenarbeit und Effizienz zwischen Unternehmen im Wirtschaftsraum der EU und Deutschlands zu transformieren und auch Start-ups und den Mittelstand am wirtschaftlichen Potenzial von Künstlicher Intelligenz effektiv teilhaben zu lassen.

Die Autoren dieser Studie diskutieren neue und möglicherweise transformative Technologien, welche die Umsetzung einer umfassenden und kollaborativ gestalteten Datenökonomie nach Vorbild europäischer Regulierung möglich machen könnten. Die Idee, auf Basis der Synthese von Blockchain und KI dezentralisierte digitale Infrastrukturen zu entwickeln, ist wegweisend. Diese Infrastrukturen sollen von Individuen, Start-ups und Unternehmen genutzt werden können, um KI-Modelle gemeinsam zu trainieren, deren Ergebnisse zu teilen oder sogar Algorithmen für die Datenverarbeitung auszutauschen. Hervorzuheben ist dabei, dass die Inhalte der geteilten Datensätze nicht notwendigerweise offengelegt werden müssen.

Datenschutzkonforme KI-Entwicklung „Made in Europe and Germany“ hat das Potenzial, ein Markenzeichen der deutschen bzw. europäischen Wirtschaft zu werden und die globale Wettbewerbsfähigkeit lokaler Unternehmen zu erhöhen. Mit den beschriebenen Lösungen kann die technologische Reife der deutschen Industrie und ihrer *hidden champions* genutzt werden, um Entwicklungen auf Basis von qualitativ hochwertigen Daten für Start-ups und andere Unternehmen zu ermöglichen. Damit kann in Deutschland ein dynamisches Ökosystem entstehen, das von der historischen Stellung der Mittelständler in ihren Nischen profitiert und junge Start-ups mit einzigartigen Wettbewerbsvorteilen hervorbringt.

Für Deutschland ist es nun wichtig, kreativ und innovativ zu sein, um neue Wege im Angesicht der aktuellen KI-Revolution zu gehen. Die vermehrte Förderung von Lösungen, wie in diesem Bericht vorgeschlagen, ist im Sinne der technologischen Souveränität der europäischen und deutschen Wirtschaft und somit eine direkte Antwort auf die Erläuterungen der Europäischen Datenstrategie und des EU AI Acts, als auch der Datenstrategie der Bundesregierung aus dem Jahr 2021.

Prof. Dr. Isabell Welpe
Inhaberin des Lehrstuhls für Strategie und Organisation TUM School of Management

0

Einleitung

0 — Einleitung

In den letzten Jahren gehörten Blockchain und Künstliche Intelligenz (KI) zu den Top Trendthemen im Technologiesektor. Oft in einem Atemzug mit anderen Schlagworten wie dem Metaversum, dem Internet der Dinge, Krypto-Assets und Digitalen Tokens genannt, gelten sie vielen als die nächste Entwicklungsstufe unserer digitalen Wirtschaft. Während Blockchain seit der Veröffentlichung des Bitcoin-Whitepapers durch die anonyme Autorin oder den anonymen Autor (oder die Personengruppe) Satoshi Nakamoto im Jahr 2008 zunehmend an Popularität gewonnen hat, ist KI spätestens mit der Veröffentlichung der Webanwendung ChatGPT Ende 2022 zu einem Buzzword geworden und wird voraussichtlich in den kommenden Jahren weiterhin im Trend liegen. Es wird geschätzt, dass die auf KI basierende Wirtschaft bis 2025 einen Gesamtwert von 420 Milliarden USD erreichen wird [1], während sie bis 2030 auf zwischen 1,6 und 2 Billionen USD anwachsen soll [1], [2]. Statista prognostiziert, dass der globale Markt für Blockchain Technologie von 5,85 Milliarden USD im Jahr 2021 auf 1,235 Billionen USD bis 2030 wachsen wird.

Obwohl diese Schätzungen mit Vorsicht zu genießen sind, sprechen sie dennoch vom enormen transformativen Potenzial, das von der Implementierung von Blockchain und KI gleichermaßen erwartet wird. Im Kern dieser Transformation steht die einzigartige Fähigkeit jeder Technologie, Daten auf neue und innovative Weise zu verarbeiten oder zu speichern. Die folgenden Seiten werden erläutern, wie KI im Wesentlichen eine ausgezeichnete Möglichkeit darstellt, die ständig wachsenden Datenmengen, die digitalisierte Gesellschaften des frühen 21. Jahrhunderts ausmachen, zu verarbeiten, indem sie diese Daten in für Menschen lesbare Inhalte umwandelt. Währenddessen wurde die Blockchain-Technologie in ihrer aktuellen Form implementiert, um Konsens, Kommunikation und Vertrauen zwischen (meist) pseudonymen Drittparteien im Internet zu verbessern.

0 — Einleitung

Allerdings ist die Einrichtung von Blockchain-Technologie vergleichsweise kostengünstig, während die erfolgreiche Entwicklung von KI-Modellen sowie deren erfolgreiche Anwendung ein kostenintensives Unterfangen ist. Jüngste Schätzungen besagen, dass jeder Trainingszyklus von OpenAI's beliebtem GPT-3-Modell mindestens 5 Millionen USD [3] gekostet hat, während für die Fertigstellung des vollfunktionalen Modells Investitionen von mehr als 100 Millionen USD erforderlich waren [3]. Diese Kosten steigen mit den wachsenden Fähigkeiten und somit der Trainingskomplexität von KI-Modellen [4]. Daher bereiten sich große digitale Plattformunternehmen (wie Google, Apple, Facebook, Amazon, Microsoft (GAFAM)) bereits auf die „größte Revolution seit der Erfindung des Internets" [5], [6], [7] vor, doch bleibt die Frage unbeantwortet, wie Einzelpersonen, z. B. alltägliche Internetnutzerinnen und -nutzer, Bürgerinnen und Bürger sowie Organisationen mit stärkeren Budgetbeschränkungen, wie kleine und mittlere Unternehmen (KMU), Wege finden können, an der nächsten Innovationswelle der digitalen Wirtschaft teilzunehmen und davon zu profitieren [8].

Wie dieser Bericht darlegen wird, könnte die erfolgreiche Implementierung von Blockchain als Teil von KI-Infrastrukturen dazu beitragen, den Effekt plattformzentrierter KI abzumildern und Nutzerinnen und Nutzern sowie kleineren und mittleren Unternehmen und Start-ups die Möglichkeit zu geben, kollaborativ an der aktuell entstehenden KI-Wirtschaft teilzunehmen. Beispielsweise könnte KI, eine Technologie, die derzeit dazu konzipiert ist, in eher geschlossenen, unternehmenseigenen Umgebungen zu operieren, in einem distribuierten, kollektiven und kollaborativen Umfeld produziert werden; während Blockchain solche KI-Systeme durch das Hinzufügen von Transparenz, Verlässlichkeit und gleichem Zugang zu digitalen Infrastrukturen verbessern würde. Zusätzlich könnte Blockchain-Technologie Werkzeuge bereitstellen, um Eigentumsansprüche klar digital zu bezeichnen, als auch um die Authentizität von Daten und Information zu bestätigen. In der Folge können neue Formen des Vertrauens im digitalen Raum möglich gemacht werden.

Um eine kollektiv geteilte digitale Infrastruktur anhand von Blockchain und KI vorzustellen, ist es von Vorteil, über Eigenschaften von Daten nachzudenken. In der öffentlichen Debatte scheint man sich daran gewöhnt zu haben, dass nutzerzentrierte Daten in großen Mengen nur im Besitz von Unternehmen und Regierungen vorkommen. Folglich hat die Frage, was passieren würde, wenn jeder von uns seine eigenen Daten selbst teilen und handeln könnte, in den letzten Jahren zunehmend weniger Aufmerksamkeit erhalten. Dennoch könnten wir inmitten der aktuellen KI-Revolution diese Frage sogar erweitern und möglicherweise so formulieren: Was wäre, wenn wir KI kreieren könnten, die in kollaborativer Zusammenarbeit gestaltet und auf geteilten Datensätzen trainiert wurde und in der Folge von allen Kollaborierenden gleichermaßen zugänglich ist und genutzt werden kann?

0 — Einleitung

Indem Daten in diesem neuen Paradigma wahrgenommen werden, könnte eine Debatte angestoßen werden, die statt *Big Data* eine Vision von *Shared Data* [9] betont. Dabei wird die Notwendigkeit für Zusammenarbeit und Teilnahme an der digitalen Wertschöpfungskette unserer demokratischen Gesellschaften hervorgehoben.

Wie Alex Pentland, Direktor der MIT Connection Science Initiative, argumentiert, sollten wir bei dem Bestreben, Daten als gemeinsam genutzte und zugängliche Ressource zu betrachten, auch die Tatsache hervorheben, dass Daten zunehmend als eine neue Vermögensklasse auftreten; ähnlich wie Öl im 20. Jahrhundert als Treiber der Wirtschaften verstanden wurde.[i] Gleichzeitig und im positiven Gegensatz zu Öl haben digitale Daten den Vorteil, dass sie sich durch Gebrauch nicht erschöpfen. Tatsächlich können Daten, indem sie geteilt, zusammengeführt und wiederverwendet werden, über die Zeit hinweg kontinuierlich Mehrwert schaffen und möglicherweise dazu beitragen, KI-Modelle zu trainieren, die zukünftigen Versionen eines digitalen Gemeinwohls entsprechen. Um noch einmal Herrn Pentland zu zitieren: „Daten sind nun zentral für die Wirtschaft, die Regierung und das Gesundheitssystem, also warum sind Daten und die KI-Systeme, die die Daten interpretieren, in den Händen von so wenigen Personen? Gemeinschaften ohne Daten über sich selbst und ohne die Werkzeuge, ihre Daten zu nutzen, stehen in Abhängigkeit von denen mit Daten und Analysemöglichkeiten derer" [9].

Bisher hat die Europäische Union ihre Fähigkeit unter Beweis gestellt, zu verstehen, wie Daten in den Händen weniger, mächtiger Dritter potenziell ein wirtschaftliches und soziales Ungleichgewicht erzeugen können. Ihre Repräsentantinnen und Repräsentanten haben dabei effektive Wege und innovative Lösungen gefunden, um diese aufkommenden Technologien zu regulieren und einzugrenzen. Mit der Ausarbeitung einer Europäischen Datenstrategie im Jahr 2020 [10], dem EU-Datengesetz im Jahr 2024 [11] sowie dem Erlass des KI-Gesetzes im Jahr 2023 [12] hat die EU wichtige Grundlagen für eine europäische Vision des Internets geschaffen. Das erstgenannte Dokument betont darin die Bedeutung Europäischer Datenräume – in Form von Datenmärkten und offen für Daten aus aller Welt [10, S. 4]. Datenmärkte könnten eine „offene, faire, vielfältige, demokratische und selbstbewusste" [10, S. 2] Umgebung bieten, während sie Vertrauen und den Schutz persönlicher Daten gemäß der Datenschutzgrundverordnung (DSGVO) gewährleisten [13]. Andererseits bietet das KI-Gesetz der EU einen umfassenden regulatorischen Rahmen für aufkommende KI-Technologien, indem es die Bedeutung der Einhaltung von Datenregulierungen sowie Datenverwaltung, Dokumentation, Transparenz und Zugangskontrolle hervorhebt [14].

i In diesem Sinne sind Daten zu einem neuen, wesentlichen Produktionsfaktor geworden, der als ebenso wertvoll betrachtet werden kann wie andere Ressourcen. Etwa der Erhalt gut ausgebildeter Arbeitskräfte oder die Möglichkeit, auf ausreichendes finanzielles Kapital zurückgreifen zu können [9].

0 — Einleitung

Wie in diesem Bericht dargelegt wird, kann die Kombination aus Blockchain und KI dazu beitragen, diese neue Vision geteilter Daten zu verwirklichen, indem Technologie gefördert wird, die dezentralisiert, gemeinschaftlich geteilt und distribuiert ist. Die Nutzung dieser Technologien könnte somit das Ziel der EU unterstützen, bis 2030 einen Anteil an der globalen Datenwirtschaft zu erreichen, der „mindestens ihrem wirtschaftlichen Gewicht entspricht" [10, S. 2]. Die Kombination aus Blockchain und KI könnte somit eine starke Grundlage bieten, um das Kollektiv der europäischen Gesellschaften zu fördern.

Auch für Deutschland, einer Nation, deren wirtschaftlicher Erfolg auf seinen *Hidden Champions* basiert, kleinen und mittelständischen Unternehmen, die sich eine Position als Weltmarktführer in ihren jeweiligen Geschäfts- und Industrienische gesichert haben, könnte die Schaffung einer solchen Infrastruktur vielversprechend sein. In der Regel ist für viele KMUs die Erstellung eines intern entwickelten KI-Modells übermäßig kostenintensiv, während gleichzeitig die Nutzung zentralisierter KI-Modelle, wie ChatGPT, nicht als vollständig konform mit Standards der Unternehmensvertraulichkeit verstanden werden kann. Zusätzlich verfügen viele KMUs nicht über ausreichende Daten und Expertise, um interne Lösungen für den effektiven Betrieb von KI-Modellinfrastrukturen zu entwickeln. Wenn größere und globalisierte Unternehmen vermehrt in der Lage wären, ihre eigenen KI-Modelle für die Produktion zu erschließen, während kleinere Unternehmen und Start-ups zurückbleiben, könnte dies zu einer Produktivitätslücke führen, die wirtschaftliche Defizite für jene Nationen verursacht, die nicht mit dem schnellen Tempo der KI-Innovation und Verbesserung Schritt halten.

Mit einem Fokus auf gemeinsame Vorteile von Blockchain und KI und mit der Intention, einzelne Nutzerinnen und Nutzer sowie kleine und mittlere Unternehmen innerhalb der EU zu stärken, konzentriert sich Kapitel 1 dieses Berichts zunächst auf eine Beschreibung der betroffenen Technologien. Es beginnt mit einer einführenden und knappen Erläuterung der Blockchain- und KI-Technologien und beleuchtet deren historischen Kontext, um zu veranschaulichen, wie diese Technologien zu effektiven Vermittlern in den zahlreichen kommunikativen Krisen unserer Gesellschaften geworden sind. Dabei werden die zentralen Merkmale und Innovationen diskutiert, die Blockchain und KI mit sich bringen. In diesem Sinne bietet das einleitende Kapitel eine kurze Zusammenfassung und Übersicht über die Schlüsseltechnologien und Innovationen hinter dem neuesten Boom in zeitgenössischen KI-Modellen, während gleichzeitig aufgezeigt wird, wie Mängel in diesen Modellen möglicherweise durch Innovationen im Blockchain-Sektor angegangen werden könnten. Es wird dabei vermieden, die technischen Spezifikationen von Blockchain zu detailliert zu diskutieren, da davon viele bereits in der *Token Studie* der Konrad-Adenauer-Stiftung behandelt wurden [15].

0 — Einleitung

Das zweite Kapitel konzentriert sich darauf, wie die wesentlichen Möglichkeiten beider Technologien genutzt und kombiniert werden können, um digitale Infrastrukturen zu schaffen, die tatsächlich dabei helfen könnten, eine gemeinsame Datenwirtschaft zu etablieren sowie europäische Ideale von Gleichheit, Datenschutz und bürgerlicher Teilhabe voranzutreiben. Hierbei wird speziell darauf eingegangen, wie die Kombination von Blockchain und KI die einzigartige Gelegenheit bieten könnte, einen Europäischen Datenraum zu etablieren, der sich u. a. durch die Schaffung von offenen Datenmarktplätzen, distribuierten Machine-Learning-Infrastrukturen, der Zuweisung von Urheberschaftsrechten, der Belohnung individueller Datenbeiträge, sowie entsprechender digitaler Architekturen auszeichnet.

Das dritte Kapitel dieses Berichts wird tiefer in die technologischen Grundlagen eintauchen, die sogenannte *Data Exchanges* ermöglichen. Insbesondere wird diskutiert, wie Federated Learning (FL), eine Art der Machine-Learning-Technik, implementiert werden könnte, um einzelnen Nutzerinnen und Nutzern sowie Unternehmen die Zusammenarbeit mit ihren Daten zu ermöglichen, ohne ihre digitale Privatsphäre zu kompromittieren. Zusätzlich wird kurz erörtert, wie die hier diskutierten Technologien in zukünftigen Branchenanwendungsszenarien wie dem *Metaverse* und der Etablierung vertrauenswürdiger Formen von KI angewendet werden könnten.

Wie im Verlauf dieses Berichts gezeigt wird, befinden sich einige der diskutierten Technologien noch in der Entwicklung – daher muss das Erschließen ihrer konkreten Implementierung als ein fortlaufender Prozess angegangen werden. Die Bedeutung der Integration diverser Perspektiven und eine reflektierte Herangehensweise bei der Gestaltung der nächsten Phase digitaler Infrastrukturen kann nicht genug betont werden, da viele aktuelle Implementierungen digitaler Räume leider dazu neigen, Vorurteile und Segregation zwischen öffentlichen Gruppen zu replizieren [16], [17], [18], [19]. Eine Tendenz, die – neben anderen wesentlichen Elementen – demokratischen Idealen sowie Prinzipien der Gleichheit und Fairness nicht entspricht. Der letzte Abschnitt dieses Berichts wird eine Zusammenfassung aller Erkenntnisse sowie eine kritische Diskussion ihrer Anwendbarkeit bieten, um erneut die Bedeutung einer kritischen und weltoffenen Herangehensweise bei der Entwicklung der beschriebenen Technologien zu betonen. Schlussfolgernd wird diskutiert, welche Wege zu beschreiten wären, um Blockchain und KI als produktiven Bestandteil gemeinsamer, digitaler Infrastrukturen zu implementieren.

1

Verständnis von Blockchain und KI

1 — Verständnis von Blockchain und KI

1.1 — Technologische Grundlagen von Blockchain und KI

Visionen von maschinell simulierten Formen der Intelligenz faszinieren die menschliche Vorstellungskraft bereits seit vielen Jahrhunderten. In ihren frühesten Anfängen zeigen sie sich beispielsweise in der Idee der Golems [20], [21] – mechanischen Wesen, die aus Lehm und Magie geformt wurden. Zu den zeitgenössischeren und populäreren Beispielen zählen der sogenannte „Mechanische Türke“[ii] des 18. Jahrhunderts, das Frankenstein-Monster von Mary Shelley aus dem 19. Jahrhundert oder Star Wars‘ R2D2 und C3PO im 20. Jahrhundert. Dennoch waren funktionale Modelle, die menschliches Denken automatisieren oder simulieren könnten, lange Zeit Werke der Science-Fiction. Anfang 2024 werden genau diese Möglichkeiten und Fortschritte jedoch ununterbrochen diskutiert. Angesichts der jüngsten Entwicklungen in der KI-Forschung und dem immensen Erfolg von AI-Chatbots wie OpenAI‘s ChatGPT, Anthropic‘s Claude und Alphabet‘s Google Gemini scheinen sie zunehmend zu zentralen Elementen der täglichen kollektiven Erfahrungen zu werden. In Anbetracht des aktuellen Hypes um KI-Technologie ist es jedoch erwähnenswert, dass die heutigen KI-Modelle nicht wirklich „denken“, so wie wir diesen Prozess für gewöhnlich verstehen. Sie beziehen sich vielmehr auf eine Vielzahl mathematischer Funktionen, um die Wahrscheinlichkeit von Ergebnissen zu berechnen, deren Aussagegenauigkeit so nah an unseren menschlichen Weltvorstellungen liegt, dass zumeist die Resultate ihrer Mechanik mit Mustern menschlichen Denkens verglichen werden.

Es überrascht daher nicht, dass die Technologie hinter modernen KI-Modellen, aufgrund ihrer Komplexität und ihren Möglichkeiten, so faszinierend ist, dass in den letzten Jahren immer häufiger danach gefragt wurde, wann KI-Modelle entwickelt werden, die komplexe menschliche Gedankengänge vollständig abbilden können. Dieser Zustand wird oft bezeichnet als Künstliche Allgemeine Intelligenz (oder AGI). Während Diskussionen über die Möglichkeiten von KI und Vorhersagen über deren zukünftige Auswirkungen auf einem Allzeithoch sind (siehe Google Trends Abbildung 1), sollte jeder Diskurs zu KI-Maximierung mit Vorsicht angegangen werden. Bereits heute zeichnet sich ein zunehmender Mangel an Energie [22] sowie an Grafikprozessoren [23], [24] ab. Beides sind zentrale Ressourcen, die benötigt werden, um die derzeitige Generation von rechenintensiven KI-Modellen zu betreiben. Bedenken um den wachsenden Energiemangel, bei gleichzeitiger Einhaltung von Klimazielen, sind dabei groß, dass beispielsweise Microsoft bereits diskutiert, in eigene Kernenergie zu investieren, um zukünftiges KI-Wachstum zu unterstützen [25].

ii Der sogenannte „Mechanische Türke“ wurde als Schachmaschine präsentiert, die ähnlich wie moderne Schachcomputer funktionieren sollte. Tatsächlich wurde sie jedoch von einem kleinen Menschen bedient, der im Inneren versteckt saß und gegen die äußere Gegnerin oder den äußeren Gegner spielte.

Google Trends

Häufigkeit der Suchanfragen nach den Begriffen „Künstliche Intelligenz“, „AI“ oder „ChatGPT“ bei Google von 2004 bis heute

Abbildung 1

Quelle: Adaptiert von [105].

Deep neural network

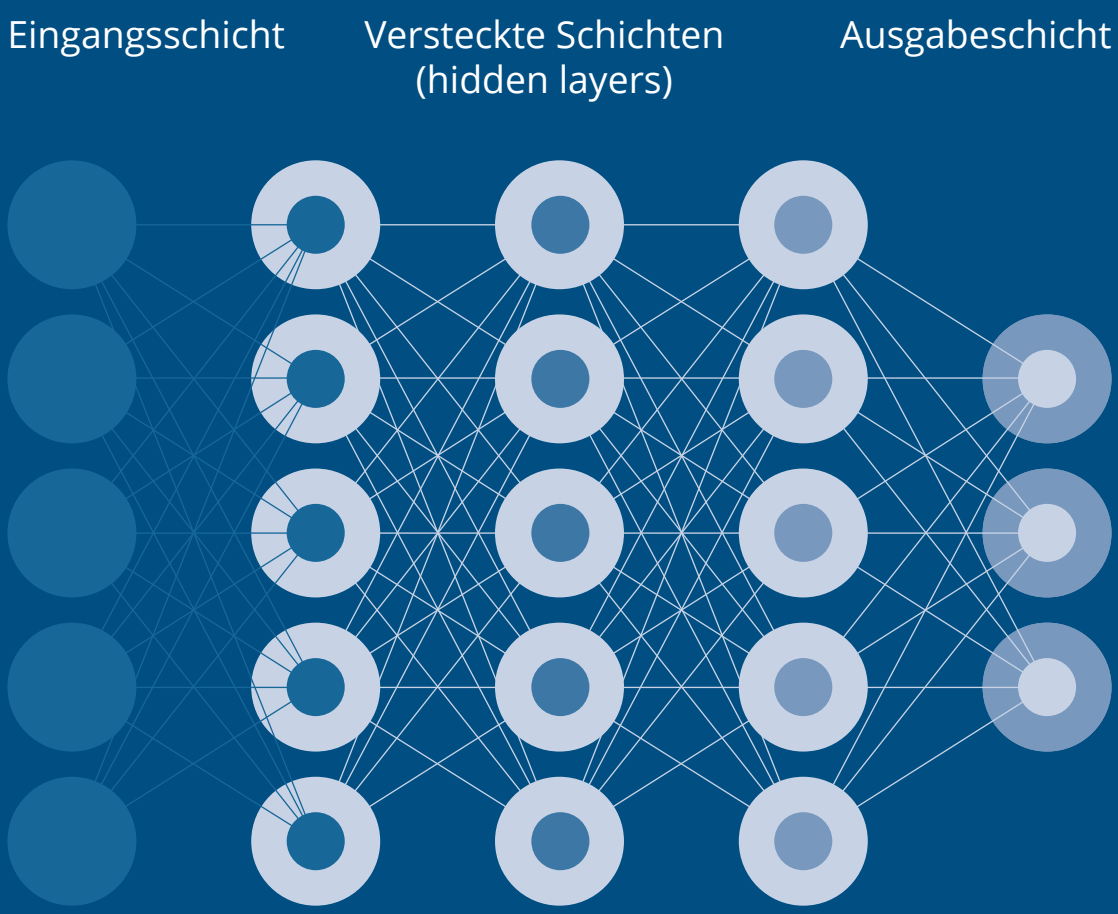

Abbildung 2

Quelle: Adaptiert von [106].

1 — Verständnis von Blockchain und KI

Die Erstellung (sowohl einfacher als auch ausgefeilter) KI-Modelle ist dabei eine Kunst und Wissenschaft, deren ausführliche Darstellung mehrere Bücher füllen könnte. Dennoch werden die folgenden Seiten darauf abzielen, ein allgemeines Verständnis ihrer Kernmechanismen zu vermitteln und dabei, wenn angebracht, potenzielle Schwachstellen hervorheben. Zunächst wird betont, dass künstliche Intelligenz und maschinelles Lernen, obwohl oft synonym verwendet, nicht notwendigerweise dasselbe sind. Künstliche Intelligenz befasst sich allgemein mit dem Bau von Systemen, „die intelligentes Verhalten simulieren" [26, S. 1]. Maschinelles Lernen hingegen ist tatsächlich ein Teilgebiet der KI-Forschung, das sich insbesondere mit der Frage beschäftigt, wie mathematische Modelle, Daten und Algorithmen genutzt werden können, um Prozesse menschlichen Denkens nachzuahmen [27].

Beide Konzepte sind in der Wissenschaftsgeschichte nicht neu und existieren in ihrer aktuellen Bedeutung mindestens seit den 1950er Jahren [28]. Was sich jedoch in diesem Zeitrahmen geändert hat, ist die qualitative Feinheit der Modelle. Frühe Modelle des maschinellen Lernens konnten einfache Prozesse berechnen, wie z. B. den kürzesten Weg zwischen zwei Punkten A und B in einem gegebenen Raum zu finden [28], [29] oder durch eine größere Datenbank, z. B. ein Telefonbuch, zu blättern, um den einzigen Kontakt mit Namen „Jane Doe" so schnell wie möglich zu finden. Die Kapazitäten aktueller Modelle des maschinellen Lernens gehen weit darüber hinaus und basieren nun auf neuronalen Netzwerken: digitale Datenpipelines, die darauf ausgelegt sind, die Funktionen von Neuronen im menschlichen Gehirn nachzuahmen. Diese Modelle werden in der Regel trainiert, indem dem neuronalen Netzwerk ermöglicht wird, seine bedeutungsbildenden Prozesse eigenständig auszubalancieren. Somit beinhalten sie eine interessante Eigenart: Da ein Teil der Modellkomposition darauf abzielt, dem Modell zu erlauben, sich selbst zu trainieren, weiß niemand genau, wie das fertige Modell arbeitet [26, S. ix].

1 — Verständnis von Blockchain und KI

Die funktionale Logik neuronaler Netzwerke und die Erfolge darauf aufbauender Innovationen, wie z. B. ChatGPT, zeugen davon, dass es die neuesten Fortschritte in der KI-Forschung geschafft haben, menschliche kognitive Prozesse viel genauer zu simulieren, als es einige Verfechterinnen und Verfechter des maschinellen Lernens in früheren Phasen der KI-Entwicklung – zumindest bis zur Implementierung des World Wide Web in den 1990er Jahren und der damit einhergehenden Menge an neuen, potenziell effektiven Trainingsdaten – für möglich gehalten hatten [28]. Die wesentliche Innovation hinter diesem Erfolg sind neuronale Netzwerke; ein Typ von maschinellem Lernmodell, das darauf ausgelegt ist, die Lernstruktur des menschlichen Gehirns zu simulieren. Einfach ausgedrückt, ahmt es die Art und Weise nach, wie biologische Neuronen einander Signale senden [30]. Um zu „lernen", verlassen sich neuronale Netzwerke auf große Mengen an Trainingsdaten, die sie verarbeiten, damit sich die Genauigkeit ihrer Ausgabe im Laufe der Zeit verbessert [30], [31]. Der Prozess des Trainierens neuronaler Netzwerke über viele Iterationen hinweg mit dem Ziel, ihre Genauigkeit zu verbessern und stetig effektivere Ergebnisse zu erzielen, wird dabei als „Deep Learning" bezeichnet. Durch Deep Learning trainierte neuronale Netzwerke (Deep Neuronal Networks) gelten bis heute als die leistungsfähigsten und fortschrittlichsten Modelle des maschinellen Lernens und sind in alltäglichen Anwendungsbereichen zu finden [26, S. 1].

Die hohe Qualität ihrer Ergebnisse wird durch einen Klassifizierungsprozess über multiple Ebenen ermöglicht, der darauf ausgelegt ist, komplexe Muster und Objekte in den Daten herauszuarbeiten und zu beschreiben. Diese Objekte können jede Art von „realweltlichem" Input repräsentieren, welcher anschließend vom Modell in für Menschen lesbaren Output umgewandelt wird. Deep-Learning-Modelle werden beispielsweise in den folgenden Bereichen umfassend genutzt: In der Übersetzung von Texten von Sprache A (z. B., Deutsch) nach Sprache B (z.B. Französisch), wie etwa durch das deutsche Start-up Deepl.com implementiert. In der Transkription von aufgezeichnetem Audioinhalt in geschriebenen Text, wie beispielsweise vom Berliner Start-up SpeechText. AI entwickelt. Oder eben auch in der Umwandlung von Texteingaben in vom Modell generierte Bilder, wie am Beispiel von OpenAI's Dall-E oder auch dem GenAI-Projekt „Stable Diffusion" verdeutlicht, dessen zugrundeliegender Algorithmus an der LMU in München entwickelt wurde.[iii]

iii Für einen detaillierten Überblick über die verschiedenen Arten von Daten, die von generativen KI-Modellen verarbeitet werden können, wird Prince [26, S. 6] empfohlen.

Ein populäres Beispiel wie Mustererkennung in Deep Learning Netzwerken (oder *Deep Neural Networks*) funktioniert, wird durch die Software DeepDream veranschaulicht. DeepDream wurde 2015 vom Google-Ingenieur Alexander Mordvintsev veröffentlicht [32]. Abbildung 3 zeigt eine Sequenz von Bildern, die mit DeepDream erstellt wurden. In diesem speziellen Beispiel wurde das Modell darauf trainiert, Muster von Hunden zu identifizieren und anschließend absichtlich übersteuert, um ein „traumähnliches" Ergebnis für jedes vom Modell verarbeitete Bild zu erzielen [33], [34]. Während das oberste der drei Bilder die ursprüngliche Darstellung von drei Mondquallen im Wasser zeigt, hebt das mittlere Bild die Gestalt von Hunden hervor, die das Modell nach etwa zehn Iterationen „identifiziert" hatte. Das letzte Bild veranschaulicht die Vielzahl an Hunden, die nach insgesamt 50 Iterationen identifiziert wurden.

Trotz des übersteuerten Ausgabealgorithmus ist DeepDream dabei ein gutes Beispiel, um die wesentlichen Stärken und gleichzeitig größten Schwächen von Deep Learning Modellen hervorzuheben. Wie das mittlere und letzte Bild zeigen, sind Deep Learning Modelle hervorragend darin, für Menschen erkennbare Muster in großen Mengen scheinbar unstrukturierter Daten zu erkennen. Gleichzeitig können sie jedoch durch die Verwendung von stark homogenen oder qualitativ unzureichenden Datensätzen „übertrainiert" werden und Datenmuster erkennen, wo keine sind.[iv] Die Erstellung effektiver Deep Learning Modelle ist daher sehr davon abhängig, dass eine ausreichende Vielfalt und Qualität der Trainingsdaten gewährleistet wird. Nur so kann sichergestellt werden, dass das Modell effektiv funktioniert und dabei Fehlausgaben minimiert werden.

Die genannten Beispiele zeigen, dass Deep Learning Modelle einen grundlegenden Wandel in der Datenverarbeitung darstellen. Die meisten der heutigen Online-Anwendungen wurden Zeile für Zeile in einer bestimmten Programmiersprache geschrieben, üblicherweise von einer/einem oder mehreren (menschlichen) Programmiererinnen oder Programmierern. *Deep Neural Networks* hingegen erstellen ihre eigenen strukturellen Muster (z. B. Code), um Daten als Eingabe zu verarbeiten und die Wahrscheinlichkeit von Ergebnissen als Ausgabe vorherzusagen. Während Software bisher darauf ausgelegt war, Programme Codezeile für Codezeile per Instruktion einer menschlichen Programmiererin oder eines menschlichen Programmierers auszuführen, sind neuronale Netzwerke nun in der Lage, sich eigenständig zu konfigurieren [28]. Sie sind zu selbstausführenden Agenten geworden, dazu fähig, fortgeschrittene Aufgaben der Problemlösung und Ereignisberechnungen zu übernehmen und somit Prozesse zu simulieren, die bisher speziell menschlicher Kognition vorbehalten waren.

iv Die Deep-Dream-Software nutzt somit auf kreative Weise den Effekt, dass Deep-Learning-Modelle „halluzinieren" (d. h. Unsicherheiten im Modell durch erfundene Muster ausfüllen), um visuell ansprechende, „psychedelische" Bilder zu erschaffen.

Deep neural network

Das Originalbild (oben) nach Anwendung von zehn (Mitte) und fünzig (unten) Iterationen von Deep-Dream, wobei das Netz auf die Wahrnehmung von Hunden trainiert wurde und sodann rückwärts läuft.

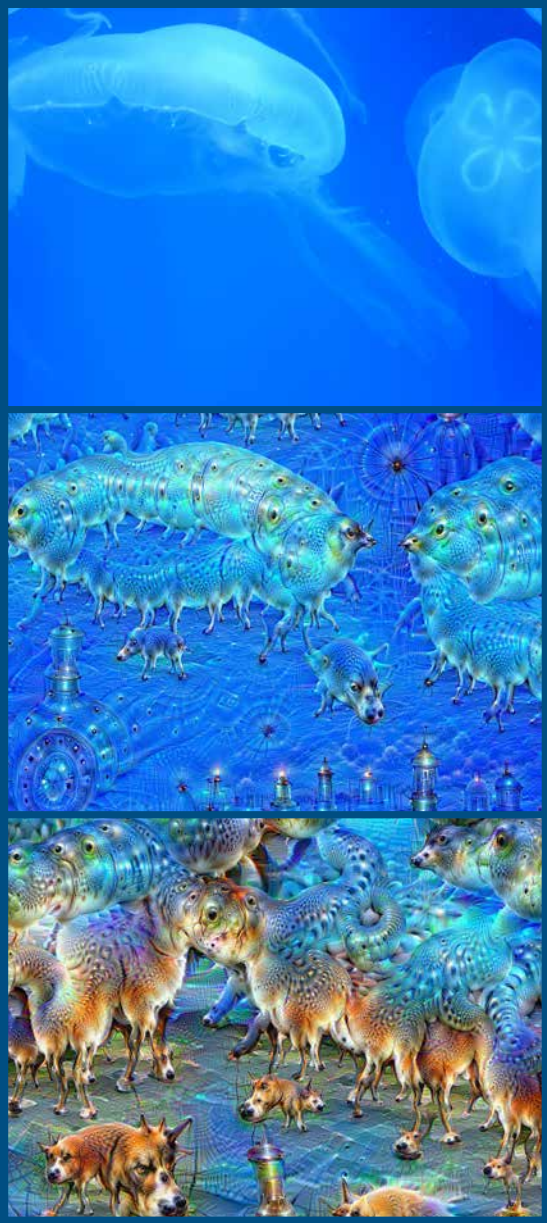

Abbildung 3

Quelle: [107].

In diesem Sinne stellen sie einen Wandel von dem, was zuvor als Software 1.0 bezeichnet wurde, zu Software 2.0 dar und auch einen Wandel in der Art und Weise, wie digitale Ökonomien in der Zukunft strukturiert und ausgeführt werden könnten. Software 2.0 repräsentiert dann das Paradigma, dass Algorithmen zunehmend in der Lage sein werden, sich selbst zu schreiben und zu entwickeln, wobei die Produktion und das Einspeisen von Daten in den Algorithmus zunehmend relevanter werden [36], [37]. Als Konsequenz könnte sich die Rolle von Software-Ingenieurinnen und -Ingenieuren von einer Produzentin oder einem Produzenten von Codezeilen, die ein Programm zusammensetzen (Software 1.0), hin zu einer Datenarchitektin oder einem Datenarchitekten verschieben, welche(r) Daten und Erkenntnisse in KI-Modelle einspeist, um hochwertige Vorhersageergebnisse zu erzeugen (Software 2.0). Natürlich wird das Programmieren per Codeeingabe in zukünftigen Softwareentwicklungsprozessen nicht vollständig ersetzt werden, könnte aber zunehmend von der KI selbst übernommen werden (aktuelle ChatGPT-Modelle sind bereits sehr gut darin, Codeausgaben für einfache Anwendungen zu erzeugen). In der Folge könnten Softwareanwendungen zunehmend auf KI-Vorhersagen und immer menschlicher anmutenden Interaktionen basieren (wie durch fortschrittliche KI-Chatbots wie Claude, Bard oder ChatGPT veranschaulicht). Ihre Effektivität und Leistung wird dabei weniger an klare Codeinstruktionen durch Programmiererinnen und Programmierer gebunden sein. Den Ausschlag geben dann eher die Menge und Qualität der Daten, die in das Modell integriert werden, um effektive Ausgaben zu liefern, Verzerrungen und das Risiko von Halluzinationen zu reduzieren, sowie Vorhersagen zu liefern, die zunehmend genaueren Darstellungen der tatsächlichen Welt entsprechen.

Wie zuvor erwähnt, wird eine wahrscheinliche Konsequenz dieser Dynamik die Bereitstellung von qualitativ hochwertigen Daten für das effektive und produktive Funktionieren von maschinellen Lernmodellen und deren umgebenden Softwareanwendungen wichtiger denn je sein. Folglich wird der reibungslose Ablauf digitaler Prozesse als auch die Leistung unserer Marktwirtschaft von Modellen abhängen, die hochwertige und repräsentative Daten verwenden und somit gut trainiert und präzise sind. Demnach fungieren Daten nicht nur als eine neue Ressource im 21. Jahrhundert („Daten sind das neue Öl“), sondern sie stellen auch ein neues Produktionsmittel dar, um das soziale Leben und die Wirtschaft zu verbessern.

Leider geht diese Vision mit einem Vorbehalt einher: Derzeitige Datenwirtschaftsmodelle sind darauf ausgelegt, Informationen auf den zentralisierten Serverinfrastrukturen digitaler Plattformbetreiber zu sammeln [5]. Dabei werden die Produzentinnen und Produzenten der Daten, Internetnutzerinnen und Internetnutzer oder Bürgerinnen und Bürger eines Landes in der Regel vom Prozess der Wertschöpfung ausgeschlossen [5]. Wenn jedoch Nutzerinnen und Nutzer und Bürgerinnen bzw. Bürger die Produzierenden ihrer Daten sind und diese im 21. Jahrhundert zu einem zentralen Treiber für effektives Marktwachstum werden, sollte die resultierende Wertschöpfung wesentlich diversifizierter gestaltet werden und den Nutzerinnen und Nutzern sowie ihren Gemeinschaften insgesamt zugutekommen.

1 — Verständnis von Blockchain und KI

Die Implementierung zukünftiger KI-Modelle auf hochzentralisierten Infrastrukturen wird in aller Wahrscheinlichkeit dazu führen, dass die Macht dieser KI-Modelle in den Händen weniger, meist unternehmerischer Akteure konzentriert bleibt. Unter anderem könnte dies zu wachsender sozialer Ungleichheit, verursacht durch eine Kluft in Rechenleistung und Zugangsmöglichkeiten, zwischen Einzelpersonen, Start-ups oder kleineren Unternehmen und großen Konzernen führen [8]. Zudem ist die Erstellung von Modellen in der aktuellen technologischen Landschaft sehr kostspielig und erfordert riesige Datenmengen – beides Faktoren, die sich bislang nur sehr große Konzerne leisten können. Wenn jedoch nur wenige Organisationen die Entwicklung und den Besitz dieser mächtigen Technologie finanzieren können, können sie auch kontrollieren, wie sie in Zukunft eingesetzt wird und effektiv die potenzielle Macht haben, das Mindset und die Entscheidungen ihrer Nutzerinnen und Nutzer übermäßig zu beeinflussen [38].

Die Nachteile dieser Formen von hochzentralisierter Macht wurden in der Vergangenheit bereits veranschaulicht. Zum Beispiel, als Facebook Inc. (Meta) im Jahr 2012 mit den News Feeds von über 680.000 Nutzerinnen und Nutzern experimentierte, um herauszufinden, ob vermehrt negativer oder positiver Content den emotionalen Zustand signifikant beeinflussen kann [39]. Oder als Amazon beschuldigt wurde, mutmaßlich seine Marktmacht zu nutzen, um Verkäuferinnen und Verkäufer davon abzuhalten, auf Nicht-Amazon-Einzelhandelswebseiten Produkte zu niedrigeren Preisen anzubieten, während sie gleichzeitig dazu angehalten wurden, ihre Produktpreise auf Amazon mithilfe steigender Plattformgebühren zu erhöhen [40]. Ein Ereignis unter mehreren, das dazu führte, dass die US-Regierung Ende 2023 eine Antitrust-Klage gegen Amazon einreichte.

Zudem werden Regierungen sicherstellen wollen, dass die Daten ihrer Bürgerinnen und Bürger sicher aufbewahrt und Datenschutzgesetze respektiert werden. Vergangene Skandale wie Cambridge Analytica haben deutlich gezeigt, wie der Missbrauch persönlicher Daten dazu genutzt werden kann, politische Ergebnisse und Kampagnen zu beeinflussen. Die Einrichtung sicherer und lokal verwalteter KI-Modelle könnte folglich die Unabhängigkeit dieser kritischen Infrastruktur für spezifische gesellschaftliche und wirtschaftliche Räume gewährleisten sowie die Autonomie und den Datenschutz der Bürgerinnen und Bürger respektieren. Ein Weg, Risiken zu mindern, könnte darin bestehen, darüber nachzudenken, wie maschinelle Modelle aussehen, die kollaborativ, fair und so gestaltet sind, dass sie Daten effektiv für fortgeschrittene Anwendungen und KI-Modellierungszwecke sammeln. Die folgenden Seiten werden einen Überblick darüber geben, wie die Fusion von KI-Technologien und Blockchain einen angemessenen Schritt in diese Richtung darstellen könnte.

1.2 — Synergien von Blockchain und KI

Mit historischen Wurzeln im Bereich der Kryptografie und Cypherpunk-Szene und aufgrund der weitreichenden Verbreitung durch die Erfindung des Bitcoin-Netzwerks durch die anonyme Erfinderin oder den anonymen Erfinder (oder die Personengruppe) Satoshi Nakamoto [41], hat die Blockchain-Technologie weitreichende Popularität erlangt. Im Kern lässt sich Blockchain als eine Technologie verstehen, die es ermöglicht, Konsens zwischen Akteuren herzustellen, die sich entweder nicht kennen oder einander nicht vertrauen können. Dies wird erreicht, indem eine distribuierte Netzwerkarchitektur jede Interaktion zwischen diesen Akteuren in einer unveränderlichen Datenbank aufzeichnet und somit effektiv die Funktion einer Vermittlerin oder eines Vermittlers ersetzt. Auf diese Weise wird eine Form des sozialen Austauschs ermöglicht, die oft als „vertrauensunabhängige Kommunikation" bezeichnet wird. Wie bereits erwähnt, hat die vorangegangene *Token Studie* [15] bereits einen wesentlichen Einblick in die technischen Spezifikationen von Blockchain-Technologien gegeben, weshalb dieser Bericht nicht darauf abzielen wird, Blockchain im Detail zu diskutieren. Nichtsdestotrotz wird auf die wesentlichen Merkmale der Blockchain-Technologie hingewiesen, die für das Thema dieses Berichts relevant sind. Im Folgenden wird eine kurze Einführung in jede dieser Eigenschaften gegeben:

Dezentralisierung
Ein Kernmerkmal im Design von (public ledger) Blockchains ist ihr distribuierter Ansatz in der Kommunikation (siehe Abbildung 4: Zentralisiertes vs. Dezentralisiertes Netzwerk). In diesem distribuierten System besteht die Blockchain aus einer Vielzahl von *Nodes* (Knotenpunkten), die Synchronität und stabilen Informationsaustausch im Netzwerk gewährleisten. Dadurch können Nutzerinnen und Nutzer direkt miteinander kommunizieren oder Blockchain-basierte Vermögenswerte versenden, was den in herkömmlichen Transaktionen üblichen Mittelsmann ersetzt [42]. Im Vergleich zu etablierten Online-Plattformarchitekturen wie von Google, Microsoft, Facebook usw. ermöglichen Blockchain-basierte Architekturen dabei einen egalitäreren Zugang zu Online-Infrastrukturen und bieten die Möglichkeit, Vermögenswerte zu tokenisieren. In der Folge können Nutzerinnen und Nutzer für ihre Aktivitäten bezahlt werden oder, wie im ursprünglichen Bitcoin-Anwendungsfall, das Netzwerk nutzen, um allgemein Zahlungen und Wertespeicherung anhand von Online-Währungen zu tätigen.

Zentralisiertes vs. Dezentralisiertes Netzwerk

Zentralisiert (A)

Dezentralisiert (B)

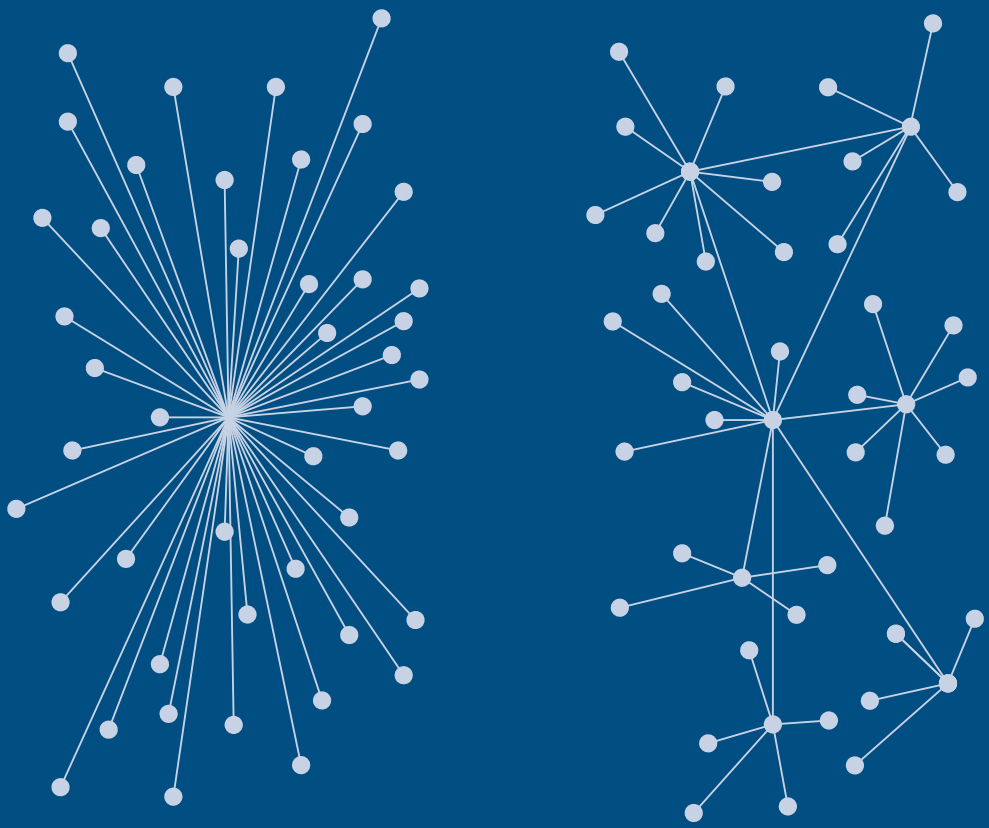

Abbildung 4

Quelle: Adaptiert von [108].

Sicherheit und Unveränderlichkeit
Ursprünge in der Kryptographieforschung haben Blockchains mit hohen Sicherheitsstandards versehen. Jede Transaktion in der Blockchain wird dabei von allen Servern (*Nodes*) im Netzwerk verifiziert und gesichert. Anschließend werden Transaktionen und ihre Zeitstempel schreibgeschützt und dezentralisiert auf mehreren *Nodes* gespeichert, was den Dateneinträgen Unveränderlichkeit verleiht, sobald sie verifiziert sind. Da jede *Node* eine Kopie jeder Transaktion aufbewahrt und Änderungen nur zulässt, wenn eine Mehrheit der *Nodes* im (dezentralisierten) Netzwerk diese Änderung verzeichnet, gelten Blockchains als manipulationssicher [43]. Sie können folglich nur modifiziert werden, wenn eine Angreiferin oder ein Angreifer die Kontrolle über mehr als 50 Prozent der *Nodes* im Netzwerk erlangt.[v] Blockchains können auch auf effektive kryptographische Methoden bei ihren Transaktionen zurückgreifen, was eine weitere Ebene der Datensicherheit hinzufügt [46], [47]. Da jedoch viele beliebte Blockchains ihre Datenbanken mit öffentlichem Lesezugriff pflegen und jede *Node*-Betreiberin bzw. jeder *Node*-Betreiber normalerweise eine Kopie der gesamten Transaktionsdatenbank aufbewahrt, sind zusätzliche Modifikationen am System erforderlich, um vollständige Datenprivatsphäre für einzelne Nutzerinnen und Nutzer oder Anwendungsfälle zu gewährleisten [37]. Diese Vorbehalte werden in den nachfolgenden Kapiteln ausführlicher diskutiert.

Rückverfolgbarkeit
Einer der Kernvorteile der Verwendung von Blockchains zur Speicherung von Informationen ist, dass aufgrund ihres sequenziellen Ansatzes in der Informationsverarbeitung und -speicherung jedem Datenpunkt eine einzigartige ID zugewiesen wird. Diese Kennzeichnung ermöglicht unter anderem die Rückverfolgbarkeit und Überprüfbarkeit von Informationen, die auf der Blockchain gespeichert sind. So kann beispielsweise die Urheberin oder der Urheber (aka. Autorin oder Autor) eines bestimmten Datensatzes identifiziert und die Gültigkeit von Informationen sichergestellt werden. Gleichzeitig erlaubt das Merkmal der Rückverfolgbarkeit nachzuvollziehen, wie sich Informationen im Netzwerk verbreiten. Dies trägt dazu bei, Datenmissbrauch und -lecks schnell zu erkennen, so dass das Gesamtrisiko weitreichender Datenkompromittierung verringert werden kann.

v Diese Schwelle betrifft Blockchains, die auf dem Proof-of-Work-Verfahren basieren. Zum Vergleich mit einem alternativen Konsenssystem, wie zum Beispiel Proof of Stake, empfehlen wir [44], [45].

Datenschutz
Während Blockchain in ihren Anfangstagen als Bewahrerin der Nutzeranonymität gepriesen wurde, haben Forschungen sowie jüngste Entwicklungen und Studien gezeigt, dass vollständige Privatsphäre und anonyme Nutzung von Blockchains in den meisten Fällen nicht umfassend gewährleistet sind [47]. Ein Grund hierfür ist, dass die oben beschriebenen Merkmale der Rückverfolgbarkeit und Transparenz in Datentransaktionen auch bedeuten, dass mit ausreichenden Daten und IT-Kenntnissen die Urheberin bzw. der Urheber einer bestimmten Transaktion zurückverfolgt und möglicherweise identifiziert werden kann. Um diesen Effekt abzumildern, stehen fortschrittliche Verschlüsselungsalgorithmen zur Verfügung, die jedoch, aufgrund ihrer eigenen inhärenten Komplexität, in diesem Bericht nicht vollständig diskutiert werden können. Für weiterführende Literatur, siehe [48], [49].

Eines der Kernprinzipien vieler technologisch fundierter Diskussionen über Blockchain ist ferner die Idee, dass ihre Möglichkeiten als Versuch gesehen werden können, ethische Verantwortlichkeit und möglicherweise sogar Fairness in digitale Infrastrukturen einzubringen. Bereits bei der Veröffentlichung des Bitcoin-Whitepapers war dieses ethische Prinzip ein zentraler Beweggrund Nakamotos [41]. Und trotz vieler Fälle von Betrug und Fehlverhalten im Bereich „Krypto“ – dem bisher bekanntesten und kapitalintensivsten Anwendungskontext der Blockchain – ist es für viele Befürworterinnen und Befürworter der Blockchain ein integraler Bestandteil ihres Enthusiasmus geworden, an dem Aufbau digitaler Infrastrukturen zu arbeiten, die zunehmend zugänglicher, egalitärer und fairer werden [50], [51]. In den vorangehenden Kapiteln wurden die technologischen Grundlagen der KI sowie mögliche Mängel übermäßig zentralisierter KI-Infrastrukturen diskutiert. Ziel der folgenden Abschnitte wird es sein, einen Überblick darüber zu geben, wie die vielversprechenden Möglichkeiten der Blockchain-Technologie helfen könnten, diese Risiken zu mildern.

„Wir befinden uns am Beginn des Übergangs von ***Big Data*** zu ***Shared Data***, infolgedessen das aus Daten gewonnene Wissen damit beginnt, **[als Bestandteil]** unserer Gesellschaft zu zirkulieren.“

Pentland et al, 2021; eigene Übersetzung

2

Data Exchanges

2 — Data Exchanges

2.1 — Einführung zu *Data Exchanges*

Bislang hat dieser Bericht das Aufkommen von Blockchain und künstlicher Intelligenz als zwei Schlüsseltechnologien der Gesellschaften des frühen 21. Jahrhunderts erörtert. Dabei wurde die Bedeutung von KI betont, um die riesigen Datenmengen, welche digitale Infrastrukturen (darunter das Internet) täglich produzieren, wirksam zu verarbeiten und daraus Einsichten zu gewinnen, die sinnvoll und für Menschen lesbar sind. Gleichzeitig wurde hervorgehoben, wie Blockchain-Technologie entworfen wurde, um u.a. Defizite in den aktuell dominanten, eher zentralisierten digitalen Plattformarchitekturen zu adressieren. Auch die ethische Mission, die oft mit Blockchain verbunden wird, sobald eine eher technisch fokussierte Perspektive in Betracht gezogen wird, wurde betont. Letztendlich liegt die Relevanz der Verschmelzung von Blockchain und KI darin, dass Daten zu einem zunehmend zentralen Gut wirtschaftlicher Produktionszyklen werden. Gleichzeitig kann eine solche Konvergenz es ermöglichen, Daten für verschiedene Arten von Marktteilnehmern bereitzustellen, seien es Big-Tech-Unternehmen, Großkonzerne, KMUs oder alltägliche Internetnutzerinnen und Internetnutzer.

Die Stärke der Blockchain-Technologie liegt darin, dass sie es ermöglicht, Eigentum und Zugang zu ihrer Infrastruktur zu distribuieren und dies in Form von Transparenz, gleichem Zugang und Nachverfolgbarkeit für ihre Nutzerinnen und Nutzer. Diese Merkmale auf KI-Infrastrukturen anzuwenden, könnte sich als eine wirksame Maßnahme erweisen, um digitale Infrastrukturen zu ermöglichen, die stärker mit den Wertvorstellungen der Europäischen Union übereinstimmen. Dabei würde der Schutz der Privatsphäre von Nutzerinnen und Nutzern gefördert [14] und gleichzeitig die Zusammenarbeit zwischen Akteuren innerhalb dieser neu entstehenden Datenräume sichergestellt [9].

Zeitgleich steht die aktuelle Datenwirtschaft vor zwei Herausforderungen beim Trainieren von Machine-Learning-Modellen: Einerseits existieren in vielen Branchen Daten in Form von isolierten Silos [52]. Diese Silos können zwischen Unternehmen, aber auch zwischen Abteilungen desselben Konzerns bestehen, da interner Wettbewerb Teams nicht immer dazu ermutigt, Daten zu teilen. Oder einfach, weil die Expertise und Ressourcen für die Einrichtung eines verbundenen Datensatzes nicht bereitgestellt werden können. Andererseits sehen sich Unternehmen mit wachsenden Anforderungen an die Einhaltung von Datenschutzvorschriften konfrontiert, was es zunehmend schwierig macht, kongruente Datenbanken zu erstellen und Machine-Learning-Modelle zu trainieren [52]. Datensilos sind ein Kernhindernis für die effektive Nutzung der Möglichkeiten von ML-Modellen. Wie zu Beginn des Berichts erwähnt wurde, sind Daten als Vermögensklasse im 21. Jahrhundert ebenso zentral wie es natürliche Ressourcen, etwa Öl, im 20. Jahrhundert waren. Es besteht jedoch ein Defizit darin, effektivere Lösungen zur Nutzung dieser neuen Vermögensklasse zu implementieren. Dabei bringen sie eine einzigartige Qualität mit sich: Während Öl jeweils nur in einer Einheit genutzt werden kann, können Daten unter Vielen geteilt werden [9], wodurch ihr positiver Gesamteffekt sogar verstärkt werden könnte.

2 — Data Exchanges

Ein Weg, das Potenzial von Daten zu erschließen, während die Einhaltung des Datenschutzes und die Kontrolle über die eigenen Datensätze sichergestellt wird, könnte die Einrichtung sogenannter *Data Exchanges* sein. *Data Exchanges* könnten als kollaborativ genutzte (deutsche, europäische und auch global zugängliche) digitale Kontenpunkte operieren, die es Unternehmen und Einzelpersonen gleichermaßen ermöglichen würden, ihre Daten als Vermögenswerte auf einem offenen Datenmarkt anzubieten. Anschließend könnten sie gehandelt werden, um fortschrittliche KI-Modelle zu trainieren, wobei den Datengeberinnen und -gebern eine zuvor vereinbarte Form der Kompensation im Austausch garantiert wird. Amateur-Wetterstationen tragen bereits heute zur meteorologischen Vorhersage und genaueren Wetterprognosen bei. Beispielsweise besteht in den USA das Citizen Weather Observer-Programm aus mehr als 7.000 Stationen, die nach Eigenangaben eine Anzahl von 50.000 bis 70.000 Beobachtungen pro Stunde senden [53]. Nach vollzogener Qualitätskontrolle werden diese Daten dann von großen US-Institutionen genutzt, einschließlich des National Weather Service, des National Ocean Service und der NASA [53].

Während Amateurwetterstationen derzeit größtenteils selbstfinanziert betrieben werden, könnte die Einrichtung von *Data Exchanges* einen Anreiz für verschiedene Nutzerinnen und Nutzer und Anwendungsfälle bieten, um ihre Daten mit Dritten zu teilen. Dadurch könnten Dienste und KI-basierte Vorhersagen verbessert werden. *Data Exchanges* fungieren als Vermittler, die effektiv die Lücke zwischen Datenproduzentinnen und -produzenten sowie Datennutzerinnen und -nutzern schließen [54] und letztere fair für ihr Engagement entschädigen. Darüber hinaus könnte die Vertraulichkeit der bereitgestellten Daten gewährleistet werden, indem in einem dezentralisierten Machine-Learning-Ansatz lokal gespeicherte Daten eines Marktteilnehmers verwendet werden, um Modelle lokal zu trainieren. Anschließend würden die Ergebnisse dieser jeweiligen Teilmodelle dem Netzwerk übermittelt und ein neuer iterativer Konsens über ein größeres, globales Modell etabliert [52].

State-of-the-art: Blockchain-basierte Federated Learning-Ansätze

Abbildung 5

Quelle: Eigene Darstellung, basierend auf [56]

Ein Vorteil dieser Architektur des sogenannten FL ist „die Entkopplung des [globalen] Modelltrainings vom direkten Zugriff auf die rohen Trainingsdaten“ [55]. Datenschutz- und Sicherheitsrisiken werden hierdurch minimiert, da Datenübertragungen reduziert und sensible oder vertrauliche Daten ausschließlich lokal verarbeitet werden [55], [56]. Auf diese Weise können große KI-Modelle schrittweise trainiert werden, ohne den Inhalt lokaler Datensätze preiszugeben, wodurch die Risiken von Datenlecks oder Datenschutzverletzungen minimiert werden [55]. Die Verbindung des FL- Prozesses mit den einzigartigen Funktionen der Blockchain-Technologie (wie im vorherigen Abschnitt beschrieben) könnte europäischen *Data Exchanges* folgende Vorteile bieten:

Datenurheberrechte

Die weit verbreitete Nutzung von generativen KI-Modellen (GenAI) hat Bedenken hinsichtlich des Urheberrechts geweckt und eine rechtliche Debatte darüber entfacht, ob Trainingsdaten ohne Bezugnahme auf deren Ursprung, wie Benutzerinnen und Benutzer, Autorinnen und Autoren oder anderen Formen der kreativen Urheberschaft, verwendet werden können. Ein bemerkenswertes Beispiel für diese Bedenken ist die Klage der *New York Times* gegen OpenAI und Microsoft wegen angeblicher Nutzung ihrer Inhalte ohne Befugnis [57]. Ähnliche Fälle wurden von der US-Autorengilde eingereicht und waren einer der Haupttreiber für die kürzlichen Streiks der Writers Guild of America (welche Drehbuchautorinnen und -autoren vertritt) in der US-Filmindustrie. Alle drei genannten Parteien argumentieren für eine Verletzung des Urheberrechts, die bei KI-Trainingsprogrammen stattgefunden hat, und fordern Mittel zur Entschädigung (z. B. eine Nutzungsgebühr), sobald ihre Inhalte bei der Entwicklung von KI-Modellen verwendet werden. Diese Fälle weisen auf zwei aktuelle Probleme hin: Erstens gibt es bisher kein umfassendes Verzeichnis, um den Ursprung der zahlreichen Datenquellen nachzuverfolgen, die in das Training fortgeschrittener KI-Modelle einfließen. Und zweitens fehlt es an allgemein anerkannten rechtlichen Rahmenbedingungen und Infrastrukturen, die einen legitimen Zugang und eine gerechte Entschädigung für die Nutzung persönlicher oder unternehmensbezogener Datensätze ermöglichen. Um diese Probleme zu mildern, könnte ein auf Blockchain basierender Datenmarktplatz die Datenurheberschaft nachweisen und eine faire und automatisierte Monetarisierung sowie das Tracking von KI-Trainingsdaten ermöglichen. So kann in der Folge eine transparente und offen zugängliche Infrastruktur für die KI-Entwicklung etabliert werden. Ähnlich wie heute Stockbilder im Internet gekauft werden, könnten *Data Exchanges* beispielsweise den Erwerb von lizenziertem Zugang zu privaten und unternehmensbezogenen Datensätzen ermöglichen und im Gegenzug Datengeberinnen und -geber fair entschädigen.

Gleichzeitig könnten Datenproduzentinnen und -produzenten mithilfe von Blockchain-basierten Datenmarktplätzen ihre Urheberschaft mittels eindeutiger Identifikatoren registrieren und so einen klar definierten, digitalen Beleg ihres Eigentums generieren. Im Dezember 2023 setzte der in den USA ansässige Mediendienst *Fox News* bereits sein Verifizierungstool „Verify" [58], [59] ein, das kryptographisches *Hashing* und digitale Signaturen nutzt, um die Originalität seiner Inhalte zu authentifizieren. Hauptziel ist es, jeder Endnutzerin und jedem Endnutzer zu ermöglichen, den Ursprung von Inhalten als von *Fox News* produziert zu erkennen. Doch Modelle wie das Beispiel von *Fox News* könnten leicht angepasst werden, um die Echtheit verschiedenster Daten und Inhalten zu gewährleisten. Während bis heute noch kein Entschädigungsmechanismus für das Training von KI existiert, könnte diese bereits gehashte Datenbank jedoch problemlos in einem gegebenen *Data Exchange* zu Trainingszwecken integriert werden.

Verantwortlichkeit und Sicherheit
Wie zuvor beschrieben, verbinden Blockchains Daten über Blöcke und sichern diese kryptographisch [41], [60]. Folglich können Modifikationen durch bösartige Akteure leicht erkannt werden [56]. In diesem Sinne bietet das System erweiterte Sicherheit gegen externe Angriffe und gewährleistet, dass die bereits trainierten Modelle ihre Gültigkeit behalten. Besonders in Zeiten weltweit zunehmender Hackerangriffe [47] fügt die Verbesserung von Datenpipelines durch Blockchain-Verifizierung zusätzliche Sicherheit und Schutz für laufende Machine-Learning-Dienste hinzu. Beispielsweise könnte die Verbindung von Blockchain und KI erweiterte Sicherheit und Schutz vor Hackerangriffen für Netzwerke bieten, die autonome Fahrzeuge steuern und verwalten [61]. In diesem Szenario kann die Kombination beider Technologien helfen, „unerwünschte Datenmodifikationen in Fahrzeugnetzwerken" [61] zu verhindern und damit die allgemeine Fahrsicherheit zu erhöhen. Darüber hinaus können durch kryptographische Eigenschaften der Blockchain geschützte Daten angepasst werden, um dem EU-Datenschutzrecht zu entsprechen und sicher in einem föderierten Umfeld geteilt zu werden [62]. Eine mögliche Anwendung wäre die Verbesserung des KI-Algorithmus zur Bedrohungserkennung und um die Gesamtsicherheitsleistung von Fahrzeugen zu erhöhen.

Datenreinheit und User-Reputation
Wie zuvor diskutiert, profitieren Machine-Learning-Modelle von hochwertigen und vielfältigen Daten, die in ihr Training einfließen. Da man bei Blockchains die Urheberin bzw. den Urheber eines bestimmten Datensatzes zurückverfolgen kann, können Autorinnen und Autoren als auch ihre Daten auf einem entsprechenden Datenmarktplatz bewertet werden [63]. Dies ermöglicht die Etablierung eines nachvollziehbaren und verifizierbaren Reputationssystems in einem distribuierten Umfeld. Beispielsweise könnte, im Kontext des zuvor erwähnten Hobbyisten-Netzwerk zur Bereitstellung von Wetterdaten, die Herkunft und der Standort eines Messgeräts (zum Beispiel eine Webcam) durch die Originalhersteller als auch die aktuelle Besitzerin bzw. den aktuellen Besitzer authentifiziert werden, um die Echtheit der generierten Daten zu gewährleisten. Basierend auf der verwendeten Hardware und der Reputation der Nutzerinnen und Nutzer könnten deren lokale Beiträge zum globalen Modell effektiv bewertet werden. Durch diese Konfiguration könnten Teilnehmende ohne die Notwendigkeit der Kenntnis voneinander (oder auch des Vertrauens zueinander) Daten effektiv handeln und zu fortgeschrittener Entwicklung einer Datenökonomie beitragen [64], [65].

Automatisierte Zahlungskanäle
Als zusätzlicher Vorteil könnten Daten- und Zahlungstransaktionen zwischen Marktteilnehmerinnen und -teilnehmern vollautomatisch erfolgen. Trainingsquellen könnten dabei öffentliche Infrastrukturen, IoT-Geräte, Sensoren industrieller Maschinen oder sogar persönliche Smartphones umfassen. Viele dieser Geräte tätigen Transaktionen, die durch ihre geringe Größe und hohe Frequenz gekennzeichnet sind. Ein solches Umfeld ermöglicht den Einsatz von automatisierten Zahlungskanälen, die in der Lage sind, Echtzeit-Transaktionen abzuwickeln. Dies bietet einen Vorteil gegenüber traditionellen Verifizierungs- und Abwicklungsmethoden durch Dritte, wie Banküberweisungen oder SWIFT, welche Mikrotransaktionen aufgrund ihres zeitaufwändigen und kostspieligen Charakters nicht ähnlich effizient abwickeln können [51]. Als Alternative würden Blockchain-basierte *Data Exchanges* mit sofortigen und automatisierten Zahlungsmechanismen ausgestattet, die direkte finanzielle Transaktionen zwischen Kaufenden und Verkaufenden erleichtern. Die Verwendung digitaler Währungen, wie beispielsweise eines digitalen Euros, wäre eine Möglichkeit, die Zuverlässigkeit von Austauschdynamiken zu optimieren. Denn sie ermöglichen Dateninhabern, ihre Vermögenswerte zu monetarisieren, wobei eine verlässliche und nachprüfbare Kompensation sowie ein reibungsloses und sicheres Transaktionserlebnis innerhalb von *Data Exchanges* gewährleistet wird.

2.2 — Data Exchanges als distribuierte Marktplätze

Wie die obigen Abschnitte gezeigt haben, kann die Verbindung von Blockchain und KI genutzt werden, um eine effektive Marktplatzumgebung für Datensätze einer distribuierten Gruppe von Teilnehmenden zu etablieren. Datenmarktplätze könnten als Vermittler fungieren und so die Lücke zwischen Datenproduzierenden und Nutzenden effektiv überbrücken [54]. Zudem könnten Marktteilnehmerinnen und Marktteilnehmer eine Vielzahl von automatisierten Mitteln nutzen – von öffentlichen Infrastrukturvorrichtungen und Sensoren für Produktionsmaschinen bis hin zu persönlichen Smartphones – um damit eine Quelle des passiven Einkommens zu generieren. Besonders profitieren könnten datenaffine Einzelpersonen sowie Unternehmen datenintensiver Sektoren wie dem Maschinenbau, der Fertigung und Produktion oder IoT-Betreiber.

Gleichzeitig könnten Internetnutzerinnen und -nutzer diese Infrastruktur verwenden, um wahrhaftige Eigentümer an ihren digitalen Daten zu werden oder auch um eine faire Entschädigung für die Weitergabe ihres Profils zur Verarbeitung von Dritten (z. B. Unternehmen) zu erhalten, die umfangreichere Kundenkenntnisse erlangen möchten. Darüber hinaus könnten Datenbörsen auch eine unterstützende Funktion haben, um sicherere Formen der öffentlichen Informationsbeschaffung zu etablieren: Erstellerinnen und Ersteller von Inhalten, wie Zeitungsverlage oder Medienhäuser, könnten in die Lage versetzt werden, ihre Veröffentlichungen als Eigentum zu markieren und so die faktische Richtigkeit der online bereitgestellten Inhalte über ihre Reputation zu gewährleisten. Zusätzlich könnten Nutzerinnen und Nutzer ein gegebenes Reputationssystem verwenden, um die Authentizität der Berichte zu bestätigen. Besonders in Zeiten zunehmender Falschinformationen online, könnten Datenverifizierungsstellen für Medien eine wirksame Gegenmaßnahme bieten und eine weitere Unterscheidungs- und Vertrauensebene zu öffentlichen Medien hinzufügen [59], [66].

Letztendlich könnte die Implementierung distribuierter *Data Exchanges* eine Grundlage für ein datengesteuertes, wettbewerbsfähiges Marktumfeld schaffen, in dem Unternehmen innovativ wirken können, während die eigene Datensouveränität gewahrt wird. Eine derartige Infrastruktur ermöglicht die Entwicklung neuer Geräte- oder Webanwendungen sowie Datenverarbeitungs- und Machine-Learning-Services. Diese können mit fortschrittlicher Ausgabegenauigkeit für spezifische Marktnischen maßgeschneidert werden. Die Einrichtung eines Blockchain-basierten Datenmarktplatzes könnte damit erhebliche Vorteile für die lokale Wirtschaft bieten, das Wettbewerbsumfeld verbessern und eine vielfältige Datenökologie fördern, was letztlich zu einem lebendigen europäischen Datenökosystem beitragen könnte. Abbildung 6 zeigt eine stark vereinfachte Version, wie Datenbörsen funktionieren könnten. Jede Phase der Datentransaktion wird in der nachfolgenden Zusammenfassung illustriert. Der anschließende Abschnitt wird dann einen tieferen Einblick geben, wie distribuierte Machine-Learning-Infrastrukturen funktionieren könnten und zusätzliche Nutzungsfälle mit hohem wirtschaftlichem Potenzial aufzeigen.

2 — Data Exchanges

(A) **Angebotseinleitung:** In dieser grundlegenden Phase generieren die Teilnehmerinnen und Teilnehmer, sei es als Datenanbietende oder -erwerbende, Angebote, die auf der Blockchain aufgezeichnet werden, was Transparenz und Unveränderlichkeit gewährleistet. Diese Angebote durchlaufen strenge Validierungsprozesse, die ihre Echtheit und Zuverlässigkeit sicherstellen. Um eine automatisierte Abwicklung zu erleichtern, werden Datenanfragen mit entsprechenden Zahlungen in digitaler Währung verbunden und treuhänderisch gehalten, um die Transaktionsvollendung bei erfolgreicher Datenlieferung zu garantieren.

(B) **Partnervermittlung:** Angebote werden abgeglichen, wenn die Teilnehmerinnen und Teilnehmer einen Konsens erreichen, was direkt oder durch Gegenangebote, die die Bedingungen der Zusammenarbeit verfeinern, geschehen kann. Bei Einigung werden alle erforderlichen Informationen sicher verarbeitet; beispielsweise würde dies im Kontext des föderierten Lernens ein Aktualisieren des globalen Modells über dezentrale Modelliterationen beinhalten.

(C) **Datenverarbeitung (inkl. Training):** Abhängig vom Typ des Datenabkommens führt die oder der Datenbesitzende entweder das lokale Training des KI-Algorithmus durch oder initiiert die Transaktion ihrer oder seiner angebotenen Daten. Die Transaktion gipfelt im direkten Austausch der Daten von Datenbesitzenden zu Datenkäufern.

(D) **Vertrags- und Zahlungsabwicklung:** Nach der Lieferung durchlaufen die erhaltenen Daten einen Verifizierungsprozess, um ihre Integrität und die Einhaltung der Vertragsbedingungen zu bestätigen. Sollten Unstimmigkeiten auftreten, wird ein Streitbeilegungsmechanismus aktiviert. Nach einer zufriedenstellenden Lösung wird die Zahlung aus dem Treuhandkonto freigegeben und automatisch verarbeitet. Zusätzlich werden Transaktionsdetails und Teilnehmendenrückmeldungen in ein umfassendes Reputationssystem eingebettet, was zu zukünftigem Vertrauen und der Einhaltung von Verantwortlichkeiten beiträgt.

Daten-Marktplatz

Vereinfachte Darstellung eines Datenmarktplatzes

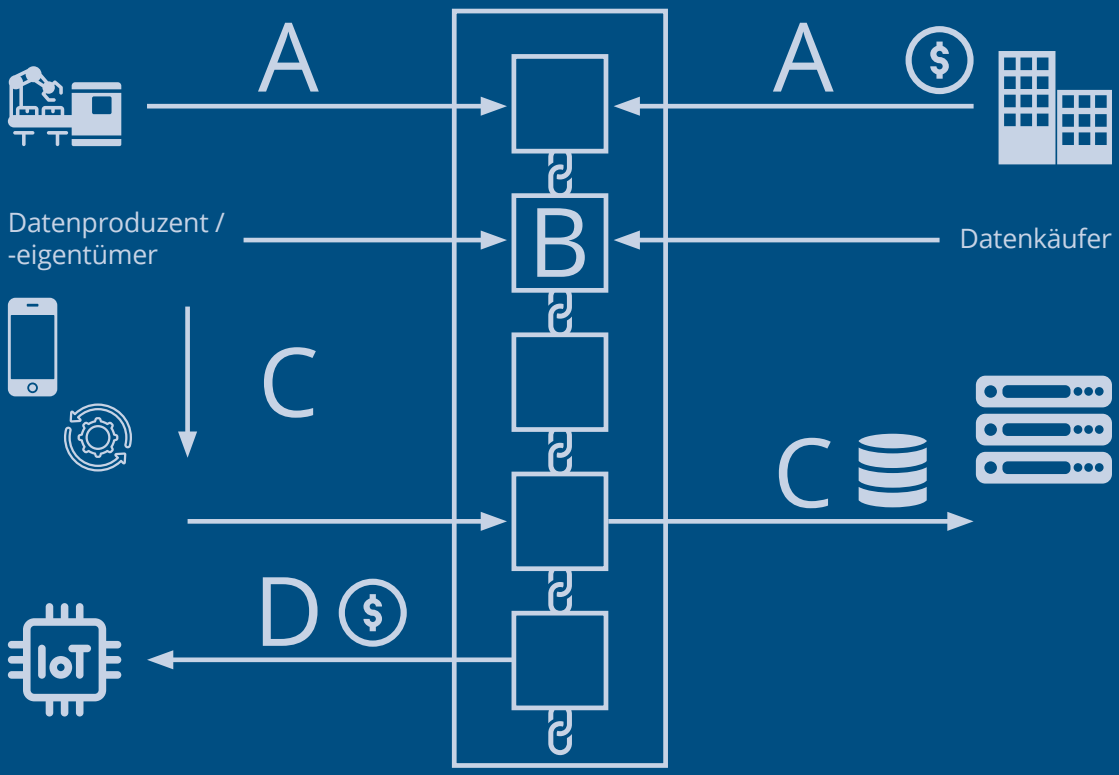

Abbildung 6

Quelle: Eigene Darstellung, basierend auf [47]

3

Anwendungs-szenarien von Blockchain und KI

3 — Anwendungsszenarien von Blockchain und KI

3.1 — Blockchain-gestütztes Federated Learning – ein technologischer Deep-Dive

Wie bereits erwähnt, gehen die traditionellen Verfahren des maschinellen Lernens davon aus, dass die Daten in einer zentralen Umgebung, z. B. auf einem lokalen Server, gespeichert und dann zum Training in das Modell eingespeist werden [55], [65], [67]. Dieser Ansatz wurde jedoch in der Vergangenheit durch zwei recht offensichtliche Faktoren eingeschränkt. Erstens kann bei fortgeschrittenen KI-Modellen die für den nächsten Trainingszyklus benötigte Datenmenge mit zunehmender Modellgröße die Kapazitäten lokaler Speichermedien überschreiten [55], [66]. Und zweitens steht die Idee der Zentralisierung großer Datenmengen, einschließlich der Daten von Einzelpersonen und Nutzerinnen und Nutzern, oft im Widerspruch mit geltenden Datenschutzrichtlinien [55], [67]. Beispielsweise können automatische Vervollständigungsmodelle für Text mithilfe von Tippdaten aus Smartphone-Tastaturinteraktionen trainiert werden. Da die Text- und Nachrichtendaten der Nutzerinnen und Nutzer jedoch in der Regel sehr persönliche und private Informationen enthalten, sollte es als höchst unethisch angesehen werden, diese Daten in Anwendungen zu speichern, die einer externen Dateningenieurin oder einem externen Dateningenieur zentralen Lesezugriff ermöglichen. Eine bessere Lösung wäre es, die Tastaturdaten lokal zu trainieren und die Ergebnisse zur Berechnung der Wahrscheinlichkeit, dass ein bestimmtes Wort in der Alltagssprache vorkommt, im globalen Modell zu aktualisieren. Aus diesen Gründen und zur Entschärfung von Datenschutzbedenken entwickelten Forscherinnen und Forscher bei Google 2016 die Technik des Federated Learning (FL).

Vereinfacht ausgedrückt, erleichtert FL das Training von Modellen auf lokalen (so genannten Edge-) Geräten, auf denen Daten gesammelt werden. Anschließend fasst es diese lokalen Trainingsergebnisse zusammen, um ein übergreifendes, globales Modell zu aktualisieren. Das globale Modell enthält dann die Trainingsergebnisse aller lokalen Modelle und verfeinert auf deren Grundlage seine eigenen Vorhersagen [55]. Somit unterliegt dem globalen Modell eine größere Anzahl an Trainingsdaten sowie die Kombination lokaler Modelliterationen, die in das globale Modell einfließen. Zusammengenommen ermöglicht dies eine höhere Ausgabegenauigkeit und bei größeren Stichproben die gleiche oder eine schnellere Präzision im Vergleich zu zentralisierten Trainingsansätzen [69], [70], [71].

FL kann zwar die Vorteile der distribuierten Datenerfassung nutzen, aber die Inhaberschaft und die Auswertung des globalen Modells sind immer noch an eine zentrale Einheit im Trainingsnetzwerk gebunden [72]. Die Eigentümerschaft des zentralen Servers bringt auch die Eigentümerschaft des globalen Modells mit sich und ermöglicht die Kontrolle über die Verwaltung von Fehlern beim Training oder die Gewichtung der Modellausgabe [72], [73], was zu einem unfairen Wettbewerbsvorteil gegenüber anderen Netzwerkteilnehmerinnen und Netzwerkteilnehmern führen kann. Infolgedessen ist der Anreiz, durch die Bereitstellung lokaler Daten zum Modelltraining beizutragen, in den meisten Anwendungsszenarien begrenzt [74], [75]. Insbesondere im B2B-Bereich ist der Schutz von Firmengeheimnissen für die Aufrechterhaltung von Wettbewerbsvorteilen in spezifischen Marktnischen entscheidend. Wenn die Vorteile nicht überwiegen, bevorzugen es daher viele Unternehmen, nicht an geteilten Trainingsprozessen teilzunehmen. FL (Federated Learning) mit zentralisiertem Dateneigentum und den damit einhergehenden beschränkten Zugriffsrechten für Datenproduzierende eignet sich daher nicht als Alternative zum Aufbau von *Data Exchanges*, um die aktuell dominante Version der digitalen Plattformökonomie des GAFAM-Typs zu optimieren.

3 — Anwendungsszenarien von Blockchain und KI

Um *Coopetition* und gleichberechtigten Zugang zu Datenquellen voranzutreiben, wurde das Konzept des FL auf Basis von distribuiertem Eigentum entwickelt [56], [64], [65], [72], [76], [77]. Bei FL mit distribuiertem Eigentum bleiben die lokalen Trainingsverfahren gleich, während die Modellaggregation und die globale Modellberechnung über ein distribuiertes Ledger-Netzwerk abgewickelt und kollaborativ über Blockchains verwaltet werden. Dadurch bietet es allen Akteuren des FL-Trainingsnetzwerks verbesserte Partizipationsmöglichkeiten und verhindert gleichzeitig, dass ein einzelner Akteur die Kontrolle über das gesamte Netzwerk und seine Modellergebnisse erlangt [75]. FL mit distribuierter Inhaberschaft kann somit ein potenziell egalitäres und die Privatsphäre schützendes maschinelles KI-Trainings-Ökosystem schaffen. Es bewahrt die kollaborative Souveränität der Trainingsbeteiligten über die Erstellung und Ergebnisse der Modelle. In diesem System können sich die Teilnehmerinnen und Teilnehmer auf Belohnungen für qualitativ hochwertige Beiträge einigen, was die Chancen für Zuverlässigkeit erhöht und positive Netzwerkeffekte unter den Akteuren fördert.

Unter den oben beschriebenen Bedingungen und unter Berücksichtigung ethischer und regulatorischer Standards [11], [12] könnte sich ein Blockchain-basiertes FL-System nahtlos in die Visionen eines europäischen Marktplatzes für souveräne Daten einfügen. In diesem Szenario könnten Blockchain-basierte Netzwerke das Training eines kollaborativ verwalteten globalen KI-Modells verbessern, wobei auch Trainingsbeiträge von Nicht-Blockchain-*Nodes* hinzugefügt werden können. Auf diese Weise wird der Datenmarktplatz genutzt, um nach der Vertragsinitialisierung zwischen dem distribuierten Netzwerk und den, zum Training des Modells beitragenden, Dateneigentümerinnen und -eigentümern Zugang zu dem Blockchain-basierten globalen Modell zu gewähren. Im Zuge dessen kann ein dezentrales KI-Netzwerk, welches zur Entwicklung einer bestimmten KI geschaffen wurde, anhand von Daten trainiert werden, die nicht dem Netzwerk gehören, während die Eigentümerinnen bzw. die Eigentümer der Daten weiterhin eine Entschädigung für den Beitrag erhalten. Insbesondere Netzwerkteilnehmende, die nur über begrenzte Rechenressourcen verfügen, können auf diese Weise partizipieren und von den Trainingsbeiträgen profitieren. Folglich würde ein solcher Datenmarktplatz den Zugang zu Daten für dezentralisierte KI-Projekte verbessern und die Durchführbarkeit komplexerer KI-Entwicklungsprojekte fördern.

3 — Anwendungsszenarien von Blockchain und KI

Hypothetische Architektur eines Blockchain basierenden FL-Netzwerks
Um die Möglichkeiten und Grenzen eines Blockchain-basierten FL-Frameworks zu veranschaulichen, wird im Folgenden eine hypothetische, dezentrale Trainingsarchitektur auf der Grundlage von Li et al. [72] dargestellt. Ein FL-Paradigma kann Daten aus einer Vielzahl von Quellen nutzen, aber zur Vereinfachung wird davon ausgegangen, dass jede teilnehmende *Node* über eine ähnliche oder gleiche Rechenleistung und Verbindungsbandbreite verfügt. Teilnehmende *Nodes* können entweder im Blockchain-Konsensalgorithmus oder im KI-Trainingsalgorithmus partizipieren oder beides.

Das Komitee erfüllt zwei wichtige Funktionen für das Netzwerk: Zum einen kann die Geschwindigkeit und Skalierbarkeit von FL-Netzwerken erhöht werden, da nur die Mitglieder des Komitees an der Berechnung des globalen Modells teilnehmen. Gleichzeitig wird durch die rotierende Auswahl der Komiteemitglieder sichergestellt, dass alle Teilnehmenden des Netzwerks für ihre Beiträge zur Verantwortung gezogen werden können. Darüber hinaus werden die Teilnehmenden ermutigt, ein hohes Maß an Datenqualität aufrechtzuerhalten, da beispielsweise Beiträge, die vom Komitee als „wertvoll“ eingestuft werden, mit Anreizen wie Bonuszahlungen, höherem Beitragsranking und größerem Vertrauen für die Netzwerkteilnehmenden verbunden werden könnten.

Es ist jedoch anzumerken, dass auch in Komitee-basierten Wahlverfahren es im Laufe der Zeit zu einer zunehmenden Konzentration von Evaluierungsmacht und Rechenressourcen kommen kann (was möglicherweise eine Rezentralisierung von Macht im Netzwerk zur Folge hat). Würden beispielsweise nur die Rechenleistung und die Qualität des Datenbeitrags als ausschlaggebende Faktoren für die Ausschusswahl herangezogen, könnten weniger leistungsfähige Netzwerkteilnehmende diskriminiert werden, während der Ausschuss aus einer zunehmend homogenen Gruppe von ressourcen- und leistungsstarken hochqualitativen Datenanbietenden bestehen würde. Um diesen Effekt der Rezentralisierung zu vermeiden und um sicherzustellen, dass die zunehmend homogenen Datenstichproben nicht zu einer Verzerrung der Ergebnisse über alle Modellinstanzen hinweg führen, sollten Mechanismen der Diversifizierung, Überwachung und Bewertung in Betracht gezogen und wirksam in die Konsensberechnungsprozesse integriert werden. Andere Konsensbewertungsmechanismen, wie z. B. *Delegated Proof of Stake*, könnten ebenfalls als praktikable Alternativen in Betracht gezogen werden [76].

Personalisiertes Federated Learning (PFL)
Durch die Anwendung von Blockchain wird die Einrichtung von großen FL-Netzwerken mit unterschiedlichen Datensätzen realistischer. Die daraus resultierende Datenheterogenität in einem datenschutzfreundlichen Lernumfeld stellt jedoch auch eine Herausforderung für die Leistung des trainierten KI-Modells dar [78], [79]. PFL verbessert das traditionelle FL, indem es die Modelle an die einzigartigen Datenproben und Anwendungsfälle der einzelnen Nutzerinnen und Nutzer anpasst. Dieser Ansatz gewährleistet eine bessere Modellleistung und benutzerspezifische Ergebnisse auf der Grundlage individueller Datenmerkmale.

3 — Anwendungsszenarien von Blockchain und KI

Über den ersten Block in der für das Training verwendeten Blockchain – dem sogenannten *Genesis-Block* – wird eine Referenz zum grundlegenden KI-Modell gespeichert, das den Ausgangspunkt für das Training bildet.

(A) Jeder Trainingsteilnehmende lädt das entsprechende Modell herunter und führt ein lokales Training unter Verwendung seiner persönlichen Datensätze durch. Die Verbesserung des lokalen KI-Modells wird über Output-Funktionen gemessen, die mathematische Darstellungen der Modellleistung für eine bestimmte Iteration bilden.
(B) Diese Aktualisierung wird dann an ein so genanntes Komitee weitergeleitet, das aus einer Stichprobe von Netzwerkteilnehmenden besteht, die das Training für eine bestimmte Zeit (z. B. eine Trainingsiteration) unterbrechen, um zur Berechnung des Modellkonsenses in der globalen Instanz beizutragen.
(C) Im Rahmen dieses Prozesses führt das Komitee ein Ranking über die Qualität der Beiträge durch, kennzeichnet unzureichende oder möglicherweise bösartige Datenbeiträge und berechnet das globale Modell. Anschließend wird die nächste Trainingsrunde eingeleitet, und neue Komiteeteilnehmende werden aus dem Netz der Trainingsteilnehmenden ausgewählt.

Komitee-Konsens-Architektur

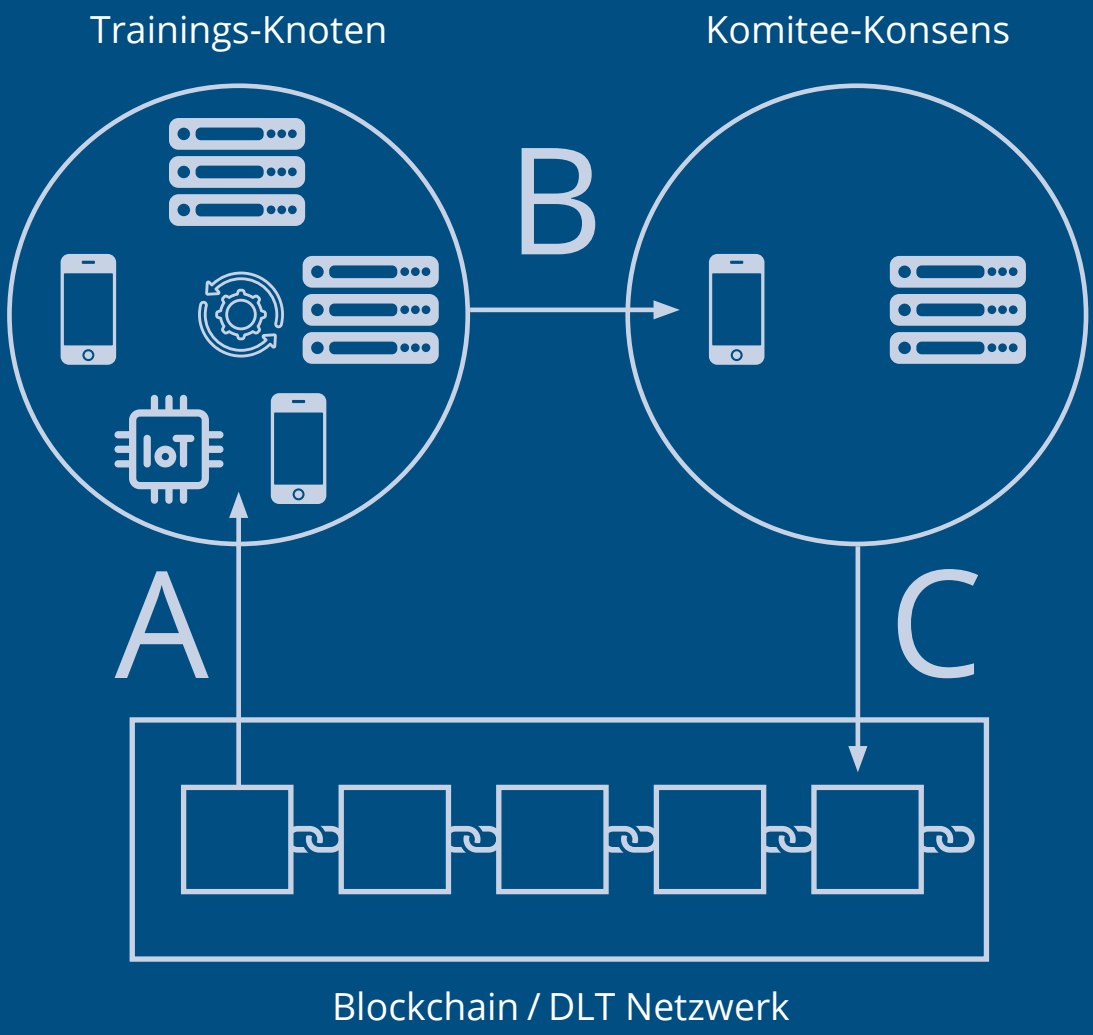

Abbildung 7

Quelle: Eigene Darstellung, basierend auf [67]

3 — Anwendungsszenarien von Blockchain und KI

Während ein globales Modell über verschiedene Geräte hinweg entwickelt wird (wie beim traditionellen FL), fügt PFL eine Ebene der Anpassung hinzu, indem es jedem Gerät erlaubt, dieses Modell auf der Grundlage seiner eigenen Daten fein abzustimmen [75], [78], [79]. Das bedeutet, dass lokale Aktualisierungen des Modells auf jedem Nutzergerät durchgeführt werden, welche im Anschluss mit dem globalen Modell kombiniert werden. Dadurch werden Modellergebnisse akkumuliert, die auf diversifizierten Modellberechnungen basieren, um eine verbesserte Vorhersagegenauigkeit zu erzielen. Gleichzeitig sichert dieses Verfahren die Vorteile eines maßgeschneiderten Trainings, ähnlich wie bei lokal trainierten, auf sehr spezifische Zwecke zugeschnittenen Modellen. In der Folge wird durch FL sowohl eine breite Anwendbarkeit als auch eine persönliche Relevanz der Modellergebnisse gewährleistet. Dieser Effekt ermutigt Netzwerkteilnehmende, ihre Ressourcen zur dezentralen KI-Entwicklung beizusteuern und ermöglicht gleichzeitig eine ausreichende Anpassungsfähigkeit, um KI-Modelle zu erstellen. Diese spezifischen Modelle könnten dann effektiv in industriellen oder geschäftlichen Umgebungen eingesetzt werden, die eine hohe Anpassungsfähigkeit erfordern.

Forschungsergebnissen zufolge [78], [80] wird der positive Trainingseffekt sogar noch verstärkt, da durch die Spezifikation von Trainingsdaten die Leistung personalisierter dezentraler KI-Modelle in vielen Fällen die von generisch trainierten föderierten Lösungen übertrifft. So kann beispielsweise ein Hersteller-Konsortium kollaborativ eine dezentrale KI anhand der Daten trainieren, die von ihren spezifischen Maschinen produziert werden. Das daraus resultierende KI-Modell ist daher besser in der Lage, den Prozess dieser spezifischen Maschinen zu verwalten und zu optimieren. Auf diese Weise könnten deutsche (oder europäische) KMUs einen globalen Wettbewerbsvorteil erzielen, indem sie hochspezialisierte, kollaborativ trainierte KI-Systeme entwickeln. Zur Veranschaulichung dieses Falles werden im Folgenden mehrere Industrieszenarien untersucht, die bereits von einer dezentralen KI-Entwicklung profitieren oder profitieren könnten.

3 — Anwendungsszenarien von Blockchain und KI

Anwendungsfälle für Blockchain-basiertes Federated Learning

Anwendungsfall 1 – KI in der Gesundheitsforschung
Die Daten des Gesundheitswesens unterliegen strengen Datenschutzbestimmungen, sodass die zentralisierte Datenverarbeitung und das KI-Training ethischen und rechtlichen Einschränkungen unterliegen. Um die Vorteile von Big Data im Gesundheitswesen zu fördern, wurden umfangreiche Forschungsarbeiten zu datenschutzfreundlichen FL-Ansätzen für die Entwicklung von KI im Gesundheitswesen durchgeführt [81], [82], [83]. Obwohl es sich um eine vielversprechende Technologie handelt, ist eine KI-basierte Scan-Software in den häufig vorkommenden Datensilos des Gesundheitswesens schwer zu entwickeln [84]. Daher ermöglicht FL die lokale Nutzung privater Daten, um einen KI-Algorithmus zu trainieren, z. B. bei der Erkennung von Hirntumoren [82]. Ein Blockchain-basiertes FL-Netzwerk könnte ein vielversprechender Weg sein, um die Zusammenarbeit zwischen verschiedenen Gesundheitseinrichtungen wie öffentlichen Krankenhäusern, Forschungsinstituten und Universitäten beim Training eines KI-Erkennungssystems zu ermöglichen. In Folge würde die Arbeit vor Ort unterstützt und beschleunigt oder bestehende Behandlungsverfahren verbessert werden. Gerade im Gesundheitswesen sind solche Implementierungen jedoch mit Vorsicht zu genießen. Wie das folgende Kapitel erörtert, kann ein betrügerischer Angriff auf das globale Modell trotz datenschutzfreundlicher Methoden wie *Differential Privacy* immer noch Informationen über die Trainingsdaten preisgeben [85]. Vor allem bei hochsensiblen personenbezogenen Daten müssen diese Bedenken berücksichtigt und ausgeräumt werden, bevor eine FL-Architektur für Gesundheitsdaten wirksam und unter Berücksichtigung ethischer Standards eingeführt werden kann.

Anwendungsfall 2 – Industrielle KI-Anwendung
Für die Anwendung von FL in der Industrie gibt es im Wesentlichen zwei Möglichkeiten:
(A) Große Industrieunternehmen könnten FL-Ansätze in Projekte integrieren, um einen KI-Algorithmus zu trainieren, der auf Performancedaten von Kundinnen und Kunden ohne Verletzung der Privatsphäre aufbaut. Ein Pilotprojekt für diese Implementierung wurde in Deutschland vom Fraunhofer IPA und der Lorch AG, einem Hersteller von Schweißmaschinen, durchgeführt. Ein vielversprechender Effekt dieser Zusammenarbeit war, dass das föderierte Training von mehreren sich in Betrieb befindenden Schweißmaschinen zu einem KI-Modell führte, das in der Lage ist, eine Schweißmaschine proaktiv abzuschalten, wenn eine Mitarbeiterin oder ein Mitarbeiter gerade dabei ist, einen potenziell gefährlichen Fehler zu begehen [86]. Solche Anwendungen von FL könnten besonders für Deutschlands kleine und mittelständische Unternehmen relevant sein und dazu beitragen, die globale Wettbewerbsfähigkeit für KI-gestützte Implementierungsszenarien zu verbessern. Unternehmen könnten lokales KI-Training nutzen, um gemeinsame globale Modelle zu entwickeln, die sich auf eine hohe maschinelle Leistungsgenauigkeit stützen, und sich gleichzeitig auf Nischenanwendungen spezialisieren, die auf individuelle Geschäftsfälle zugeschnitten sind. Durch ein FL-Szenario könnte Deutschlands derzeitiger Wettbewerbsvorteil, ein größeres Cluster von weltweit führenden Industrie- und Fertigungsunternehmen zu haben, die Möglichkeit bieten, ein Netzwerk von KI-Marktführern in Industrie- und Fertigungsszenarien aufzubauen und einen künftigen Wettbewerbsvorteil darstellen.

3 — Anwendungsszenarien von Blockchain und KI

(B) In einem zweiten Szenario könnten industrielle Hersteller FL als Teil ihrer länderübergreifenden Maschineninfrastruktur implementieren und so ein kollaboratives Training für ihre Maschinen in Deutschland und im Ausland unter Berücksichtigung der lokalen Datenschutzgesetze schaffen. In einem Projekt mit Siemens hat das Start-up Katulu eine FL-Infrastruktur für die Verbesserung der automatischen optischen Inspektionssysteme in den Siemens-Werken in Erlangen entwickelt [87]. Nach Angaben von Katulu bildet das erfolgreich implementierte FL-System die Grundlage für einen breiteren Rollout in weiteren Siemens-Werken, auch in China.

Im Folgenden werden zwei Beispiele in Deutschland und Europa vorgestellt, die an der Etablierung von Blockchain-basierten FL-Anwendungsszenarien im IoT-Umfeld arbeiten. Das deutsche Start-up deltaDAO baut eine offene Infrastruktur für einen Blockchain-basierten FL-Marktplatz auf, auf dessen Basis Daten, aber auch Algorithmen und ganze ML-Modelle für industrielle Anwendungsszenarien gehandelt werden können. In Zusammenarbeit mit Akteuren wie Airbus und der Dutch Blockchain Coalition ist deltaDAO Teil des europäischen Projekts Gaia-X. Das Start-up ist in mehreren Bereichen aktiv (Luftfahrt, Industrie 4.0, Mobilität und Produktion) und arbeitet nach eigener Aussage im Einklang mit EU-Recht und EU-Verordnungen [88].

Ein zweites, verwandtes Projekt ist die fetch.ai foundation, die von der deutschen Bosch AG und dem britischen Start-up fetch.ai initiiert wurde und in Partnerschaft mit Telekom MMS betrieben wird [89]. Das daraus resultierende Netzwerk zielt darauf ab, „Innovation und Zusammenarbeit zwischen Industrieteilnehmern durch kollektive Forschung und Entwicklung, kollaborative Anwendungen, gemeinsame Initiativen und die Entdeckung wertvoller Geschäftsmodelle“ [90] in KI-basierten Bereichen zu fördern. Beide Projekte sind frühe Beispiele dafür, wie föderierte Datenmarktplätze innerhalb des europäischen Raumes gestaltet werden könnten. Diese sollten im Hinblick auf ihre künftige Entwicklung beobachtet werden.

3 — Anwendungsszenarien von Blockchain und KI

Anwendungsfall 3 – Smart City und IoT-KI
Maschinen, die in der städtischen Infrastruktur eingesetzt werden, produzieren zunehmend große Datenmengen, die in ihrer Summe das Internet der Dinge bilden. Diese Kommunikationsnetze sind für den Einsatz intelligenter Geräte unumgänglich. Die Nutzung des autonomen Fahrens von Fahrzeugen in der Stadt basiert beispielsweise auf dem Zugang zu großen Datensätzen der lokalen Umgebung und der anschließenden schnellen und intelligenten Analyse. Durch den Einsatz von KI können wichtige Berechnungen durchgeführt werden, z. B. zur Vorhersage des Verkehrsflusses [91].

Im Pilotprojekt „Heat", das 2021 in der Stadt Hamburg durchgeführt wurde, wurden autonome Kleinbusse erfolgreich eingesetzt, die ein umfassendes KI- und IoT-Netzwerk nutzten, welches Echtzeitdaten aus der Umgebung ableitet [92]. Das vom Bundesministerium für Digitales und Verkehr geförderte Nachfolgeprojekt „ALIKE" soll ab 2024 öffentlich zugängliche, autonome Kleinbusse in Hamburg bereitstellen [93]. Diese Fallstudien zeigen, dass gut trainierte KI-Algorithmen und der Zugang zu Daten ein wichtiger Faktor sind, um das autonome Fahren voranzubringen. Dateninseln und geschlossene KI-Trainingssysteme könnten jedoch die Entwicklung erschweren und die Herstellung von Transparenz zwischen den Beteiligten verkomplizieren, z. B. im Falle von Unfällen oder ähnlichen Ereignissen, die eine rechtliche Klärung erfordern. Die Implementierung eines FL-Netzwerks, das auf distribuierten Eigentumsverhältnissen basiert, könnte die Entwicklung transparenter und effektiver Algorithmen für das autonome Fahren beschleunigen. Mit einem datenschutzfreundlichen Zugang zu IoT-Daten, die für das Training von KI-Algorithmen genutzt werden können und gleichzeitig die Zusammenarbeit einer Vielzahl von Akteuren fördern, unterstützt Blockchain die kollaborative Smart-City-Entwicklung in einem egalitären Umfeld für alle Beteiligten und fördert Anreize für den freien Markt.

Anwendungsfall 4 – Mobile KI-Anwendungen
FL kann auf mobilen Geräten angewandt werden, um auf Trainingsdaten zuzugreifen, die normalerweise zu klein wären, um für ein effektives Training in Frage zu kommen [94]. Um einen erweiterten Zugang zu Trainingsdaten zu ermöglichen, haben die Autoren Bonawitz et al. [94] ein FL-System entwickelt, das einer Vielzahl von mobilen Geräten ermöglicht, dem Netzwerk beizutreten und an Trainingsrunden teilzunehmen. Dies kann z. B. bei E-Commerce-Empfehlungsalgorithmen nützlich sein, die häufig in der Cloud gespeicherte, aber sensible Nutzerdaten verwenden, welche möglicherweise gegen das Datenschutzrecht verstoßen. In Anbetracht dieses Szenarios haben Forschende erfolgreich die Anwendung eines datenschutzfreundlichen FL-Systems für Alibaba und Taobao getestet, das On-Device-Daten nutzt [95]. Solche Innovationen könnten für deutsche E-Commerce-Plattformen wie Zalando und AboutYou alternative Infrastrukturen bieten. Außerdem werden diese Ansätze durch jüngste Fortschritte wie die FL-Anwendung PockEngine, die die erforderliche Rechenleistung für effizientes KI-fine-tuning erheblich reduzieren, zunehmend realisierbar [96].

3 — Anwendungsszenarien von Blockchain und KI

Zentralisiertes Machine Learning: Beim zentralisierten Machine Learning ist ein zentraler Server der ausschließliche Knotenpunkt für die Zusammenführung von Daten aus verschiedenen Quellen und deren Verarbeitung für das Training des maschinellen Lernalgorithmus.

Dezentrales Federated Learning: Beim dezentralen Federated Learning werden die Modelle lokal an der Datenquelle trainiert und die einzelnen Aktualisierungen an einen zentralen Server zur globalen Modellaggregation gesendet.

Distribuiertes Federated Learning: Distribuiertes Federated Learning ermöglicht ein lokales Modelltraining an jeder Datenquelle, wobei die Aggregation und Aktualisierung des globalen Modells von allen Knotenpunkten mithilfe der Distributed-Ledger-Technologie gemeinsam verwaltet wird.

Zentralisiertes vs. Dezentralisiertes vs. Distribuiertes Netzwerk

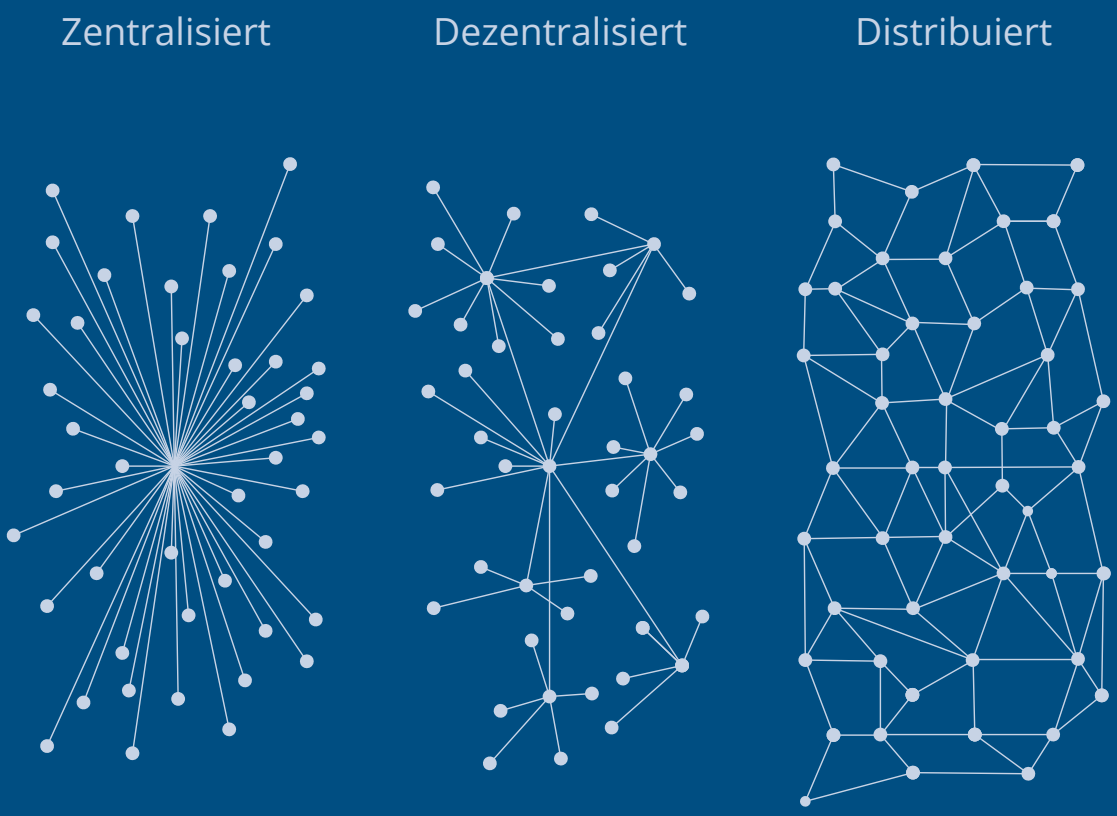

Abbildung 8

Quelle: Adaptiert von [108]

3 — Anwendungsszenarien von Blockchain und KI

Herausforderungen und offene Forschungsfragen für Anwendungen von Federated Learning

Herausforderung 1: Kompromittierung der Privatsphäre
FL verbessert zwar die Wahrung der Privatsphäre beim KI-Training erheblich, ist aber keine Garantie für den Datenschutz [85]. Obwohl FL-Methoden die Übertragung sensibler Daten verhindern, hat die Forschung gezeigt, dass es Szenarien gibt, in denen gegnerische Angriffe Trainingsdaten aus Modell-Updates offenlegen können [85], [97]. Um diese Auswirkungen abzuschwächen und Datenlecks zu verhindern, wurden Mechanismen wie die *Differential Privacy* (DP) integriert. *Differential Privacy* fügt den sensiblen Daten ein „Rauschen" hinzu, das die Abfrage von Informationen erschwert. Allerdings besteht ein Konflikt zwischen Datenschutz und Modellgenauigkeit, da erhöhtes Rauschen die Genauigkeit senkt, aber den Datenschutz verbessert [98], was zu Lasten der Benutzerfreundlichkeit und der Modellgenauigkeit führt [85]. Weitere Forschung sollte durchgeführt werden, die DP mit anderen Techniken kombiniert, um die Modellgenauigkeit beizubehalten und gleichzeitig Privatsphäre zu gewährleisten. Dennoch können bewusste Versuche, die Privatsphäre der Benutzerinnen und Benutzer zu verletzen, eine Bedrohung darstellen [85]. Fortgeschrittene Kryptographie wie sichere *Multi-Party-Computation*, vollständig homomorphe Verschlüsselung und *Zero Knowledge Proofs* sind vielversprechende Lösungen, müssen aber noch weiterentwickelt werden, um die praktische Anwendung zu gewährleisten und den Rechenaufwand zu reduzieren [85], [99].

Herausforderung 2: Poisoning-Angriffe
FL birgt die Herausforderung, Modell-Updates zu identifizieren, die aufgrund von vorherigem *Data Poisoning* oder direktem *Update Poisoning* manipuliert wurden [100]. Angriffe durch Dritte können beispielsweise die Labels der Trainingsdaten ersetzen, wodurch die Update-Parameter zur Entwicklung suboptimaler, globaler Modelle führen oder Schlupflöcher für die Angreiferin oder den Angreifer hinterlassen können [85]. Die Forschung hat mehrere Lösungen vorgeschlagen, wie z. B. byzantintolerante FL-Systeme oder medianbasierte Aggregatoren [85]. Gleichzeitig haben Forschende hervorgehoben, dass die weitreichenden Auswirkungen von Poisoning-Angriffen auf das globale Modell durch die Implementierung von kostengünstigen Verteidigungsmaßnahmen reduziert werden können [100]. In Kombination mit einem transparenten und gemeinschaftlich verwalteten globalen Modell könnten die Risiken von Poisoning-Angriffen weiter minimiert werden [101].

3 — Anwendungsszenarien von Blockchain und KI

Herausforderung 3: Heterogene Geräte mit unterschiedlichen Netzwerkverbindungen und Rechenleistungen
Die Heterogenität großer, distribuierter Netze in Bezug auf Netzanbindung und Rechenleistung stellt eine Herausforderung für die faire und offene Teilnahme an FL-Trainingsrunden dar. Leistungsstarke Server sind nicht ausreichend ausgelastet, da eine linear synchronisierte Trainingsrunde nur so schnell sein kann wie das langsamste Gerät [102], [103]. Unterschiedliche Netzwerkbandbreiten und Verbindungsstabilitäten können ebenfalls die Möglichkeit zur Teilnahme an FL einschränken [103]. Daher werden in der Forschung asynchrone FL-Systeme vorgeschlagen, die das unabhängige Hochladen von Trainings-Updates zu zufälligen Zeiten ermöglichen und damit das Netzwerk effizienter gestalten [85], [102], [103]. Asynchrones FL birgt jedoch auch das Risiko einer Überrepräsentation von Modell-Updates aus rechenstärkeren Quellen [103]. Somit verbessert asynchrones FL die Möglichkeit des Trainingszugangs und könnte die Skalierbarkeit von Trainingsnetzwerken steigern, während es gleichzeitig die fairen Partizipationsmöglichkeiten für weniger leistungsfähige Geräte verringert, was ein unerwünschtes Ereignis darstellt. Um diese Dynamik abzuschwächen und die Diskriminierung von Teilnehmerinnen und Teilnehmern mit geringer Rechenleistung zu minimieren, sind weiterführende Forschungen notwendig, die sich auf die Erkundung der Potenziale von asynchronen FL-Frameworks konzentrieren.

3.2 — *Metaverse* und vertrauenswürdige KI – Mögliche Zukunftsszenarien für Blockchain- und KI-Implementierungen

Zum Abschluss dieses Kapitels werden zwei weitere Anwendungsszenarien beleuchtet, für die sich die Kombination von Blockchain und KI als nützlich erweisen könnte. Allerdings befindet sich die Forschung in diesen Bereichen noch in einem sehr frühen Stadium, was Vorhersagen erschwert, da einige technische Hürden überwunden werden müssen, bevor von voll funktionalen Lösungen gesprochen werden kann. Daher muss betont werden, dass die folgenden kurzen Abschnitte bisher nur Gegenstand (fundierter) Spekulationen sein können.

Metaverse
Der erste in diesem Zusammenhang zu nennende Forschungsgegenstand ist der Begriff *Metaverse*. Spätestens seit sich der Plattformgigant Facebook in Meta umbenannt hat, ist dieser Begriff zu einem eigenen Buzzword geworden. Er repräsentiert eine vielschichtige Ebene unterschiedlicher Bedeutungen, deren größte Gemeinsamkeit wohl darin besteht, dass Anwendungen im *Metaverse* ein Augmented Reality (AR)- oder Virtual Reality (VR)-Gerät beinhalten, das eine Art projizierte (AR) oder geschlossene (VR) räumliche Computerumgebung erzeugt. In diesem Bericht wurden die Machtungleichgewichte erörtert, die für Nutzerinnen und Nutzer sowie kleinere Unternehmen entstehen könnten, wenn zentralisierte Formen des KI-Trainings die einzige Entwicklungsoption bleiben. Im *Metaverse* könnten ähnliche Probleme auftreten, mit dem Unterschied, dass die Kontrolle (und möglicherweise das Eigentum) der Unternehmen über Augmented-Reality- oder Virtual-Reality-Projektionen das unmittelbare Sichtfeld der Nutzerinnen und Nutzer direkt beeinflussen würde.

3 — Anwendungsszenarien von Blockchain und KI

In diesem Sinne kann man sich das *Metaverse* als eine dreidimensionale Datenumgebung vorstellen, die möglicherweise die täglichen Erfahrungen und das Leben eines Individuums verbessert, indem sie Projektionen erzeugt, die den Nutzenden als produktiv erscheinen. Es könnte aber auch in zunehmendem Maße dazu dienen, einen Strom aufmerksamkeitsfördernder Impulse zu erzeugen, die Werte aus den unmittelbaren visuellen Erfahrungen der Nutzenden abstrahieren. Anschließend werden diese Werte in Form von Daten an die zentralen Server großer, digitaler Plattformunternehmen übermittelt. Wenn man die Dynamik des *Metaverse* in diesem Kontext betrachtet, wird deutlich, dass die Schwierigkeiten, die bei den aktuellen Fragen der KI-Ausbildung und des Eigentums in einer zentralisierten Datenwirtschaft immer deutlicher hervortreten, auch in dieser noch im Entstehen begriffenen, räumlichen Datenverarbeitungsumgebung auftreten könnten. Wie dieser Bericht gezeigt hat, kann Blockchain allgemein als eine Technologie verstanden werden, die, wenn sie richtig eingesetzt wird, einen egalitäreren und gerechteren Zugang zu digitalen Infrastrukturen eröffnen kann. Gleichzeitig bietet sie durch ihre verteilte und konsensgesteuerte Datenverarbeitungsinfrastruktur den zusätzlichen Vorteil, dass sie die digitalen Erfahrungen sicherer – oder zumindest fälschungssicherer – macht. Die Bereitstellung dieser zusätzlichen Ebene der Datenüberprüfung über die Blockchain für räumliche Datenverarbeitungsumgebungen könnte die Stabilität und Sicherheit der Erfahrungen im *Metaverse* verbessern und es Hackern erschweren, die visuellen Erfahrungen der *Metaverse*-Nutzenden zu ihrem Vorteil zu beeinträchtigen.

Vertrauenswürdige KI

Der kürzlich erlassene EU AI-Act fordert klare Datenverwaltung, Aufzeichnungspflicht, Transparenz und Zugangskontrolle für KI [14]. Um diese neue Verordnung zu erfüllen, könnte Blockchain dabei helfen, eine Überprüfbarkeit, Transparenz und Rückverfolgbarkeit für KI-Modelle zu schaffen [14]. In diesem Bericht wurden mehrere mögliche Szenarien für diese Funktionen erörtert, sind aber nicht auf das Thema der Überprüfbarkeit eingegangen. Wie bereits erwähnt, ist es schwierig, wenn nicht gar unmöglich, die interne Dynamik von KI-Modellen, die auf neuronalen Netzen basieren, vollständig zu verstehen, weshalb die genauen Entscheidungsprozesse fortgeschrittener KI-Modelle bis heute nicht vollständig rückverfolgt werden können. Mit zunehmender Modellgröße und damit zunehmender Modellkomplexität wird sich dieser Trend voraussichtlich fortsetzen, sodass es für menschliche Beobachter immer schwieriger wird, einen vollständigen Einblick in die Art und Weise zu gewinnen, wie ein bestimmtes KI-Modell zu einer bestimmten Entscheidung oder Vorhersage gekommen ist. Ähnlich wie Github die Nachverfolgung von Code-Änderungen über sogenannte „Commits“ ermöglicht, könnte die Blockchain jedoch die Erstellung von Prüfpfaden ermöglichen, um Änderungen, Erweiterungen oder Datenquellen, die in ein bestimmtes Modell eingeflossen sind, nachzuverfolgen - und diese Aufzeichnungen als unveränderlich zu speichern, solange die Blockchain existiert [104]. Dieser Ansatz bietet zwar keinen umfassenden Einblick in das KI-System, könnte aber Prüfenden und Regulierungsbehörden eine erste Orientierung bieten, um die Blackbox-Dynamik der Entscheidungsfindung eines bestimmten Modells zu verstehen. Zukünftige Forschung sollte sich damit befassen, wie dieser Ansatz sicher implementiert und ob andere Methoden produktiv eingesetzt werden können, um zukünftig sicherere und besser überprüfbare KI-Systeme zu entwickeln.

Metaverse

Mögliche Angriffsvektoren auf das Nutzererlebnis im Metaverse im Vergleich mit Möglichkeiten, Blockchain und KI als Milderungstechnologien für erweiterte Sicherheit zu nutzen.

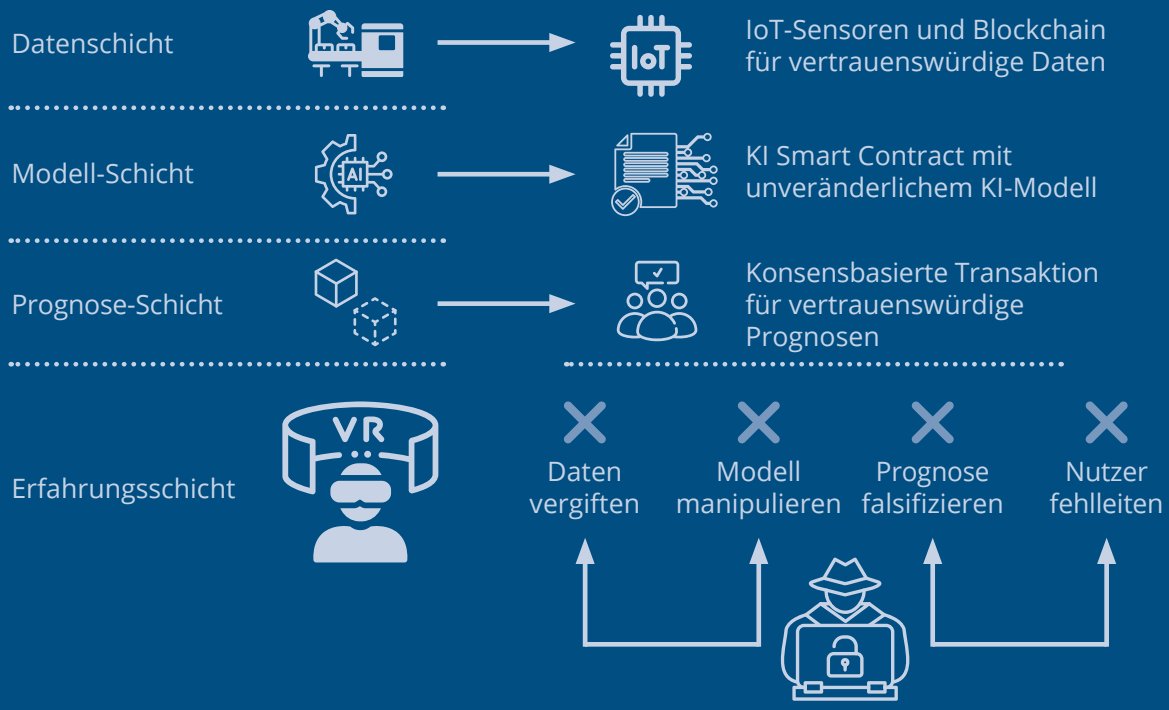

Abbildung 9

Quelle: Eigene Darstellung, basierend auf [96].

4

Fazit

4 — Fazit

In diesem Bericht wurde analysiert, wie die Konvergenz von Blockchain und KI produktiv umgesetzt werden kann, um den Datensektor in Deutschland und der EU zu stärken. Die Analyse ergab, dass die Stärke der jeweiligen Technologien u. a. darin liegt, zu einer besseren Verarbeitung von Daten beizutragen. Es wurde gezeigt, dass KI hervorragend geeignet ist, um die Übermenge an Daten, die unsere Gesellschaften zu Beginn des 21. Jahrhunderts ausmachen, produktiv zu meistern. Ein vielversprechendes Potenzial der Blockchain ist indessen, dass sie die Möglichkeit bietet, einen egalitäreren Zugang zu diesen Daten zu gewähren und den durch die Datennutzung generierten Reichtum gleichmäßiger und nach dem Prinzip individueller Beiträge zu verteilen.

Kapitel 1 bereitete eine vereinfachte, kurze Zusammenfassung und Übersicht darüber, wie Deep Neural Networks, die Infrastruktur hinter den neuesten Fortschritten in der KI-Entwicklung, funktionieren. Dabei wurde beschrieben, dass neuronale Netzwerke große Mengen an hochwertigen Daten benötigen, um sich zu entwickeln und immer genauere Vorhersagen bei gleichzeitiger Minimierung von Verzerrungen zu ermöglichen. Zusätzlich wurde durch deren selbstjustierende Dynamik aufgezeigt, dass die genaue Funktionsweise eines neuronalen Netzwerks in einer Black Box stattfindet, die von menschlichen Akteurinnen und Akteuren nur von außen untersucht werden kann. Hierbei ist die Kontrolle über Entscheidungsfindungs- und Vorhersageprozesse fortgeschrittener KI schwer zu erreichen und könnte mit zunehmender Modellkomplexität schwieriger werden. Anschließend wurden die möglicherweise negativen Auswirkungen der Zentralisierung von Macht und Kontrolle über KI in den Händen einer ausgewählten Gruppe international agierender, digitaler Plattformunternehmen diskutiert. Zuletzt wurde betont, wie die einzigartigen technologischen Merkmale der Blockchain dabei helfen könnten, eine digitale KI-Infrastruktur zu schaffen, die die Risiken einer übermäßig zentralisierten KI durch Mittel zur Dezentralisierung, Transparenz, Unveränderlichkeit und Nachverfolgbarkeit von Daten abschwächt.

4 — Fazit

Unter der Prämisse, dass Daten im 21. Jahrhundert zu einer zunehmend zentralen Vermögensklasse werden, führte Kapitel 2 dann die Idee von *Data Exchanges* ein. Data *Exchanges* wurden als mögliche Lösung zur Optimierung zentralisierter KI-Infrastrukturen beschrieben. Dabei könnten sie einer Vielzahl von deutschen, europäischen und internationalen Unternehmen ermächtigen, an einem kollaborativen Datenmarkt teilzunehmen, der Beitragende fair entlohnt, gleichzeitig die Datenschutzrechte der Nutzerinnen und Nutzer wahrt und das effektive Training von KI-Modellen im Einklang mit dem europäischen Datenschutzgesetz sicherstellt. Es wurde betont, dass die Kerntechnologie, die diesen Paradigmenwechsel von Big Data zu Shared Data begleiten könnte, eine Form des FL sein würde, welche die Bedeutung des distribuierten Eigentums bei der Ausbildung fairer, ethischer und vorurteilsminimierender KI-Modelle hervorhebt. Anschließend diskutierten wir, wie ein von FL angetriebener Datenmarkt Start-ups, KMUs und größere Unternehmen in Deutschland sowie Europa stärken könnte, indem fortschrittliche und kollektive ML-Modelle produziert werden, die durch den gemeinsamen Zugang zu einem globalen Trainingspool Geschäftsinnovationen fördern und einen Wettbewerbsvorteil gewährleisten, indem sie die Freiheit bieten, auf den Anwendungsfall zugeschnittene und spezialisierte lokale Modellinstanzen zu erstellen.

Kapitel 3 fokussierte sich auf einen technologischen Deep-Dive in die Funktionsweise des FL und diskutierte die möglichen Vorteile der Anwendung dieser Technologie in vier spezifischen Industrieszenarien. Abschließend wurden die potenziellen Vorteile von Blockchains und KI-Technologie für die neuesten Trends in der technischen Innovation, insbesondere das *Metaverse* und die mögliche Weiterentwicklung von vertrauenswürdiger KI, diskutiert, die im Einklang mit den EU-Standards für Daten-Governance, Prüfbarkeit, Transparenz, Nachverfolgbarkeit und Zugangskontrolle für KI-Modelle operiert.

Wie zu Beginn dieses Dokuments erwähnt, ist diese Analyse nicht ohne ihre Limitationen. Aufgrund der Komplexität des Themas sowie des anhaltenden Tempos von Innovationen und neuen Entdeckungen in den beiden technischen Feldern Blockchain & KI, hat dieser Bericht zunächst vor allem das Ziel, einen konzeptionellen Überblick über die produktive Konvergenz von Blockchain und KI-Technologie zu geben. Obwohl es das Ziel war, diese Konvergenz möglichst spezifisch zu beschreiben, wurde teilweise die tatsächliche Mechanik dieser Technologie im Sinne der Zugänglichkeit vereinfacht. Außerdem ist die hier skizzierte Vision keineswegs allumfassend. Während Blockchain und KI eine vielversprechende Konvergenz zugunsten einer Verarbeitung von Daten bieten, die in stärkerem Einklang mit europäischen Werten steht, müssen noch viele technologische Hürden überwunden werden, bis die zugrundeliegende Vision Wirklichkeit werden kann. So ist beispielsweise die Blockchain-basierte Datenverarbeitung tendenziell langsamer und erfordert größere Datenmengen, da in jedem Validierungsknoten im Netzwerk eine Kopie aller Datenpunkte gespeichert wird. In der Folge müssten die Ingenieurinnen und Ingenieure von *Data Exchanges* überlegen, welche Teile der Daten genau in der Blockchain gespeichert werden müssen und welche Teile über ein alternatives Datenübertragungsnetzwerk verarbeitet werden können. Gleichzeitig löst FL nicht alle Datenschutzprobleme der heutigen KI-Modelle. Wie bereits erörtert, gibt es eine datenschutzfreundliche Dynamik des FL-Lernens, die jedoch für groß angelegte Anwendungen weiter verfeinert werden müsste. Außerdem können KI-Systeme gehackt oder manipuliert werden, um sensible Informationen sowohl für lokale als auch für globale Instanzen preiszugeben. Es gibt zwar Techniken zur Abschwächung dieses Effekts, aber weitergehende Forschung ist nötig, um einen fortgeschrittenen Schutz der Privatsphäre und der Sicherheit für Nutzerinnen und Nutzer föderierter Netze für maschinelles Lernen sowie für die Empfängerinnen und Empfänger der Ergebnisse, die auf der Grundlage dieser Netze erstellt werden, zu gewährleisten.

4 — Fazit

Zusätzlich zu den technischen Einschränkungen wurde die Bedeutung einer ethischen KI-Entwicklung betont, einschließlich der Notwendigkeit, nicht-diskriminierende KI-Ökosysteme zu entwickeln. Da KI ihre Ergebnisse auf der Grundlage dessen produziert, worauf sie trainiert wurde, ist das Datenmanagement ein wichtiger Hebel, um soziale Voreingenommenheit und Diskriminierung zu überwinden. Vielfalt bei Eingabemustern und der KI-Trainingsinfrastruktur wird daher Teil der wichtigsten Schritte sein, um die von KI-Algorithmen erzeugten Vorurteile zu verringern. Entwicklerinnen und Entwickler sind zu ermutigen, die KI so weiterzuentwickeln, dass die Vielfalt und die Einbeziehung verschiedenster Stimmen und Akteure von Anfang an gefördert werden. Außerdem sollten Mechanismen für den gleichberechtigten Zugang und die gleichberechtigte Beteiligung von Nutzerinnen und Nutzern sowie Empfängerinnen und Empfänger von KI-Ergebnissen in die Infrastrukturen für Governance und Datenmanagement integriert werden. In der Hoffnung, dass diese Maßnahmen dazu beitragen werden, diese wichtige Debatte auch auf der Ebene der digitalen Plattforminfrastrukturen voranzutreiben.

Trotz der oben genannten Einschränkungen konnten die Leserinnen und Leser den vorliegenden Exkurs in den Bereich der Blockchain und KI hoffentlich genießen und haben ein Gefühl dafür bekommen, wie diese Technologien produktiv eingesetzt werden könnten, um die deutsche und europäische Wirtschaft zu stärken und die Ressource Daten stärker in den Mittelpunkt zu rücken. In diesem Sinne könnte die Konvergenz von Blockchain und KI eine Alternative zu den hochgradig zentralisierten Modellen der Datenverarbeitung bieten, nach denen wir im derzeitigen Status Quo von digitalen (Plattform-)Ökonomien arbeiten und leben. Wie genau diese Alternative aussehen könnte, muss in Gesprächen zwischen Forschenden, Ingenieurinnen und Ingenieuren, Entscheidungstragenden, Politikerinnen und Politikern, europäischen Regierungen und der Öffentlichkeit erkundet werden.

5

Literatur-verzeichnis und Glossar

5 — Literaturverzeichnis

[1] B. Thormundsson, ‘Artificial Intelligence market size 2030’, Statista. Letzter Zugriff 14. Dezember 2023 [Online]. https://www.statista.com/statistics/1365145/artificial-intelligence-market-size/.

[2] World Economic Forum, ‘This is the AI-environmment balancing act — it’s delicate | World Economic Forum’. Letzter Zugriff 14. Dezember 2023 [Online]. https://www.weforum.org/agenda/2023/04/balancing-ais-carbon-footprint-and-its-potential-for-transformative-positive-climate-impact/.

[3] C. S. Smith, ‘What Large Models Cost You – There Is No Free AI Lunch’, Forbes. Letzter Zugriff 14. Dezember 2023 [Online]. https://www.forbes.com/sites/craigsmith/2023/09/08/what-large-models-cost-you--there-is-no-free-ai-lunch/.

[4] P. Wang et al., ‘Learning to Grow Pretrained Models for Efficient Transformer Training’, 2023, doi: 10.48550/ARXIV.2303.00980.

[5] P. Ahluwalia und T. Miller, ‘The next big thing – artificial intelligence’, Soc. Identities, vol. 29, no. 1, S. 1–4, Jan. 2023, doi: 10.1080/13504630.2023.2236372.

[6] B. Gates, ‘The Age of AI has begun’, gatesnotes.com. Letzter Zugriff 14. Dezember 2023 [Online]. https://www.gatesnotes.com/The-Age-of-AI-Has-Begun.

[7] M. Manela, ‘I think AI is perhaps the biggest revolution since the invention of the internet’, ctech. Letzter Zugriff 14. Dezember 2023 [Online]. https://www.calcalistech.com/ctechnews/article/qc1x9jcho.

[8] P. Olson, ‘Google, Microsoft Will Dominate AI as Computing Costs Surge’, Bloomberg.com, Feb. 19, 2024. Letzter Zugriff 26. Februar 2024[Online]. https://www.bloomberg.com/opinion/articles/2024-02-19/artificial-intelligence-microsoft-google-nvidia-win-as-computing-costs-surge.

[9] A. Pentland, A. Pentland, A. Lipton, und T. Hardjono, Building the new economy: data as capital. Cambridge, Massachusetts; London, England: The MIT Press, 2021.

[10] European Commision, ‘Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions - A European strategy for data’, European Commision, Brussels, Belgium, Feb. 2020.

[11] European Commision, ‘Data Act | Shaping Europe’s digital future’. Letzter Zugriff 27. Februar 2024 [Online]. https://digital-strategy.ec.europa.eu/en/policies/data-act.

[12] European Commision, ‘AI Act | Shaping Europe’s digital future’. Accessed: Feb. 27, 2024. [Online]. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

5 — Literaturverzeichnis

[13] European Parliament and Council of the European Union, 'REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)'. Official Journal of the European Union, Apr. 27, 2016. [Online]. https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[14] S. Ramos und J. Ellul, 'Blockchain for Artificial Intelligence (AI): enhancing compliance with the EU AI Act through distributed ledger technology. A cybersecurity perspective', Int. Cybersecurity Law Rev., vol. 5, no. 1, S. 1–20, Mar. 2024, doi: 10.1365/s43439-023-00107-9.

[15] P. Sandner und B. Schaub, Token Studie: Grundlagen und Anwendungsszenarien der Blockchain-Technologie. Berlin: Konrad-Adenauer-Stiftung e. V, 2023.

[16] C. Apprich, F. Cramer, W. Hui Kyon Chun, und H. Steyerl, Pattern Discrimination. DE: meson press, 2018. Letzter Zugriff 27. Februar 2024 [Online]. https://doi.org/10.14619/1457.

[17] N. Shah, 'Refusing Platform Promises: A Gendered Rewriting of Digital Imaginaries', 2023, doi: 10.25595/2298.

[18] A. Juhasz, G. Langlois, und N. Shah, Really Fake. in In search of media. Minneapolis: University of Minnesota Press, 2021.

[19] S. U. Noble, Algorithms of oppression: how search engines reinforce racism. New York: New York University Press, 2018.

[20] A. Vudka, 'The Golem in the age of artificial intelligence', NECSUSEuropean J. Media Stud., vol. 9, no. 1, pp. 101–123, Jul. 2020, doi: 10.25969/MEDIAREP/14326.

[21] E. D. Bilski, 'Meet The Golem: The First "Artificial Intelligence"', Google Arts & Culture. Letzter Zugriff 29. Februar 2024 [Online]. https://artsandculture.google.com/story/meet-the-golem-the-first-artificial-intelligence/BAXhTNxULrWYKg.

[22] D. Erdenesanaa, 'A.I. Could Soon Need as Much Electricity as an Entire Country', The New York Times, New York, Oct. 10, 2023. Letzter Zugriff 29. Februar 2024 [Online]. https://www.nytimes.com/2023/10/10/climate/ai-could-soon-need-as-much-electricity-as-an-entire-country.html.

[23] E. Griffith, 'The A.I. Industry's Desperate Hunt for GPUs Amid a Chip Shortage', The New York Times, New York, Aug. 16, 2023. Letzter Zugriff 27. Februar 2024 [Online]. https://www.nytimes.com/2023/08/16/technology/ai-gpu-chips-shortage.html

[24] D. Paresh, 'Nvidia Chip Shortages Leave AI Startups Scrambling for Computing Power', Wired Magazine, Aug. 24, 2023. Letzter Zugriff 29. Februar 2024 [Online]. https://www.wired.com/story/nvidia-chip-shortages-leave-ai-startups-scrambling-for-computing-power/.

5 — Literaturverzeichnis

[25] J. Calma, ‘Microsoft is going nuclear to power its AI ambitions’, The Verge. Letzter Zugriff 26. Februar 2024 [Online]. https://www.theverge.com/2023/9/26/23889956/microsoft-next-generation-nuclear-energy-smr-job-hiring.

[26] S. J. D. Prince, Understanding deep learning. Cambridge, Massachusetts: The MIT Press, 2023.

[27] M. Chui et al., ‘The economic potential of generative AI - The next productivity frontier’. McKinsey & Company, Jun. 2023.

[28] #OxfordAI, ‘A history of AI’. University of Oxford, 2023. Letzter Zugriff 26. Februar 2024 [Online]. https://oxford.shorthandstories.com/ai-a-history/.

[29] J. Schmidhuber, ‘Annotated History of Modern AI and Deep Learning’, 2022, doi: 10.48550/ARXIV.2212.11279.

[30] IBM, ‘What is Machine Learning? | IBM’. Letzter Zugriff 27. Februar 2024 [Online]. https://www.ibm.com/topics/machine-learning.

[31] C. M. Bishop, Pattern recognition and machine learning. in Information science and statistics. New York: Springer, 2006.

[32] A. Mordvintsev, C. Olah, und M. Tyka, ‘Inceptionism: Going Deeper into Neural Networks’, Google Research. Letzter Zugriff 27. Februar 2024 [Online]. https://blog.research.google/2015/06/inceptionism-going-deeper-into-neural.html.

[33] A. Mordvintsev, M. Tyka, und C. Olah, ‘deepdream/dream.ipynb at master · google/deepdream’, GitHub. Letzter Zugriff 27. Februar 2024 [Online]. https://github.com/google/deepdream/blob/master/dream.ipynb.

[34] A. Gordić, ‘gordicaleksa/pytorch-deepdream’, GitHub. Letzter Zugriff 27. Februar 2024 [Online]. https://github.com/gordicaleksa/pytorch-deepdream.

[35] N. Shah, ‘The unbearable oldness of generative artificial intelligence: Or the re-making of digital narratives in times of ChatGPT’, Eur. J. Cult. Stud., p. 13675494231223572, Jan. 2024, doi: 10.1177/13675494231223572.

[36] A. Karpathy, ‘Software 2.0’, Medium. Letzter Zugriff 27. Februar 2024 [Online]. https://karpathy.medium.com/software-2-0-a64152b37c35.

[37] A. Pentland, J. Werner, und C. Bishop, ‘Blockchain+AI+Human: Whitepaper and Invitation’. The MIT Trust::Data Consortium for blockchain+AI research, 2021. Letzter Zugriff 2. Februar 2023 [Online]. https://connection.mit.edu/sites/default/files/publication-pdfs/blockchain%2BAI%2BHumans.pdf.

5 — Literaturverzeichnis

[38] V. Kennedy, 'Exploring the future of AI: The power of decentralization', Cointelegraph. Letzter Zugriff 27. Februar 2024 [Online]. https://cointelegraph.com/news/the-crucial-role-of-decentralization-in-shaping-ai-s-future.

[39] R. Meyer, 'Everything We Know About Facebook's Secret Mood-Manipulation Experiment', The Atlantic. Letzter Zugriff 27. Februar 2024 [Online]. https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/

[40] Le Monde, 'Amazon sued in the US for anti-competitive practices', Le Monde, Franckreich, Sep. 27, 2023. Letzter Zugriff 27. Februar 2024 [Online]. https://www.lemonde.fr/en/international/article/2023/09/27/amazon-sued-in-the-us-for-abusing-its-position_6140177_4.html.

[41] S. Nakamoto, 'Bitcoin: A Peer-to-Peer Electronic Cash System'. 2008. Letzter Zugriff 2. Februar 2023 [Online]. https://bitcoin.org/bitcoin.pdf.

[42] S. Ding und C. Hu, 'Survey on the Convergence of Machine Learning and Blockchain', 2022, doi: 10.48550/ARXIV.2201.00976.

[43] H. Taherdoost, 'Blockchain Technology and Artificial Intelligence Together: A Critical Review on Applications', Appl. Sci., vol. 12, no. 24, S. 12948, Dez. 2022, doi: 10.3390/app122412948.

[44] C. Schwarz-Schilling, J. Neu, B. Monnot, A. Asgaonkar, E. N. Tas, und D. Tse, 'Three Attacks on Proof-of-Stake Ethereum', 2021, doi: 10.48550/ARXIV.2110.10086.

[45] V. Buterin, 'A Proof of Stake Design Philosophy', Medium. Letzter Zugriff 8. Januar 2024 [Online]. https://medium.com/@VitalikButerin/a-proof-of-stake-design-philosophy-506585978d51.

[46] X. Zhu, H. Li, und Y. Yu, 'Blockchain-Based Privacy Preserving Deep Learning', in Information Security and Cryptology, vol. 11449, F. Guo, X. Huang, and M. Yung, Eds., in Lecture Notes in Computer Science, vol. 11449., Cham: Springer International Publishing, 2019, S. 370–383. doi: 10.1007/978-3-030-14234-6_20.

[47] S. Heister und K. Yuthas, 'How Blockchain and AI Enable Personal Data Privacy and Support Cybersecurity', in Blockchain Potential in AI, T. M. Fernández-Caramés and P. Fraga-Lamas, Eds., IntechOpen, 2022. doi: 10.5772/intechopen.96999.

[48] X. Sun, F. R. Yu, P. Zhang, Z. Sun, W. Xie, und X. Peng, 'A Survey on Zero-Knowledge Proof in Blockchain', IEEE Netw., vol. 35, no. 4, S. 198–205, Jul. 2021, doi: 10.1109/MNET.011.2000473.

[49] A. Acar, H. Aksu, A. S. Uluagac, und M. Conti, 'A Survey on Homomorphic Encryption Schemes: Theory and Implementation', ACM Comput. Surv., vol. 51, no. 4, pp. 1–35, Jul. 2019, doi: 10.1145/3214303.

5 — Literaturverzeichnis

[50] V. Buterin, ‘The Meaning of Decentralization’, Medium. Accessed: Feb. 27, 2024. [Online]. https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274.

[51] A. M. Antonopoulos, The internet of money: a collection of talks. Volume 1. Vereinigte Staaten von America: Merkle Bloom LLC, 2016.

[52] Y. Cheng, Y. Liu, T. Chen, und Q. Yang, ‘Federated learning for privacy-preserving AI’, Commun. ACM, vol. 63, no. 12, S. 33–36, Nov. 2020, doi: 10.1145/3387107.

[53] CWOP, ‘Citizen Weather Observer Program’, Citizen Weather Observer Program. Letzter Zugriff 27. Februar 2024. [Online]. http://wxqa.com/.

[54] L. D. Nguyen, S. R. Pandey, S. Beatriz, A. Broering, und P. Popovski, ‘A Marketplace for Trading AI Models based on Blockchain and Incentives for IoT Data’, no. arXiv:2112.02870. arXiv, Dec. 06, 2021. Accessed: Jan. 29, 2024. [Online]. http://arxiv.org/abs/2112.02870.

[55] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, und B. A. y Arcas, ‘Communication-Efficient Learning of Deep Networks from Decentralized Data’, 2016, doi: 10.48550/ARXIV.1602.05629.

[56] A. Qammar, A. Karim, H. Ning, und J. Ding, ‘Securing federated learning with blockchain: a systematic literature review’, Artif. Intell. Rev., vol. 56, no. 5, S. 3951–3985, 2023, doi: 10.1007/s10462-022-10271-9.

[57] M. M. Grynbaum und R. Mac, ‘New York Times Sues OpenAI and Microsoft Over Use of Copyrighted Work - The New York Times’, The New York Times, New York, Dez. 27, 2023. Letzter Zugriff 27. Februar 2024 [Online]. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

[58] F. Corporation, ‘Verify tool’, verify.fox. Letzter Zugriff 27. Februar 2024 [Online]. https://www.verify.fox.

[59] Polygon Labs, ‘Fox Corporation Taps Polygon PoS to Power Verify, an Open Protocol for Content and Image Verification’. Letzter Zugriff 27. Februar 2024 [Online]. https://polygon.technology/blog/fox-corporation-taps-polygon-pos-to-power-verify-an-open-protocol-for-content-and-image-verification.

[60] T. Laurence, Blockchain for Dummies, 3rd ed. Indianapolis: John Wiley & Sons Inc, 2023.

[61] G. Bendiab, A. Hameurlaine, G. Germanos, N. Kolokotronis, und S. Shiaeles, ‘Autonomous Vehicles Security: Challenges and Solutions Using Blockchain and Artificial Intelligence’, IEEE Trans. Intell. Transp. Syst., vol. 24, no. 4, S.. 3614–3637, Apr. 2023, doi: 10.1109/TITS.2023.3236274.

[62] A. Giannaros et al., ‘Autonomous Vehicles: Sophisticated Attacks, Safety Issues, Challenges, Open Topics, Blockchain, and Future Directions’, J. Cybersecurity Priv., vol. 3, no. 3, S.. 493–543, Aug. 2023, doi: 10.3390/jcp3030025.

5 — Literaturverzeichnis

[63] A. Dixit, A. Singh, Y. Rahulamathavan, und M. Rajarajan, 'FAST DATA: A Fair, Secure, and Trusted Decentralized IIoT Data Marketplace Enabled by Blockchain', IEEE Internet Things J., vol. 10, no. 4, Art. no. 4, Feb. 2023, doi: 10.1109/JIOT.2021.3120640.

[64] Z. Zhou, C. Guo, X. Zhang, R. Wang, L. Zhang, und M. Imran, 'A Blockchain-based Data Sharing Marketplace with a Federated Learning Use Case', in 2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Dubai, United Arab Emirates: IEEE, May 2023, S.. 1041–1044. doi: 10.1109/ICBC56567.2023.10174981.

[65] B. Eom, S. Lim, Y.-H. Suh, S. Woo, und C. Park, 'Federated Learning Using Blockchain-based Marketplace', in 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN), Paris, Frankreich: IEEE, Jul. 2023, S. 795–797. doi: 10.1109/ICUFN57995.2023.10199626.

[66] X. Wang, H. Xie, S. Ji, L. Liu, und D. Huang, 'Blockchain-based fake news traceability and verification mechanism', Heliyon, vol. 9, no. 7, S. e17084, Jul. 2023, doi: 10.1016/j.heliyon.2023.e17084.

[67] Y. Li, C. Chen, N. Liu, H. Huang, Z. Zheng, und Q. Yan, 'A Blockchain-Based Decentralized Federated Learning Framework with Committee Consensus', IEEE Netw., vol. 35, no. 1, S.. 234–241, Jan. 2021, doi: 10.1109/MNET.011.2000263.

[68] W. Shi, J. Cao, Q. Zhang, Y. Li, und L. Xu, 'Edge Computing: Vision and Challenges', IEEE Internet Things J., vol. 3, no. 5, Art. no. 5, Okt. 2016, doi: 10.1109/JIOT.2016.2579198.

[69] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson, und M. Jirstrand, 'A Performance Evaluation of Federated Learning Algorithms', in Proceedings of the Second Workshop on Distributed Infrastructures for Deep Learning, Rennes Frankreich: ACM, Dez. 2018, S. 1–8. doi: 10.1145/3286490.3286559.

[70] M. Asad, A. Moustafa, und T. Ito, 'Federated Learning Versus Classical Machine Learning: A Convergence Comparison', 2021, doi: 10.48550/ARXIV.2107.10976.

[71] I. Khitrov und C. Harth-Kitzerow, 'Accuracy Tradeoffs of Federated Learning approaches'. Technical University Munich, 2022. Letzter Zugriff 30. November 2023 [Online]. https://www.net.in.tum.de/fileadmin/TUM/NET/NET-2022-11-1/NET-2022-11-1_04.pdf.

[72] Y. Li, C. Chen, N. Liu, H. Huang, Z. Zheng, und Q. Yan, 'A Blockchain-based Decentralized Federated Learning Framework with Committee Consensus', IEEE Netw., vol. 35, no. 1, S. 234–241, Jan. 2021, doi: 10.1109/MNET.011.2000263.

[73] Z. Yang, Y. Shi, Y. Zhou, Z. Wang, und K. Yang, 'Trustworthy Federated Learning via Blockchain'. arXiv, Aug. 12, 2022. Letzter Zugriff 28. Januar 2024 [Online]. http://arxiv.org/abs/2209.04418.

5 — Literaturverzeichnis

[74] Y. Liu, F. R. Yu, X. Li, H. Ji, und V. C. M. Leung, 'Blockchain and Machine Learning for Communications and Networking Systems', IEEE Commun. Surv. Tutor., vol. 22, no. 2, S. 1392–1431, 2020, doi: 10.1109/COMST.2020.2975911.

[75] J. Yun, Y. Lu, und X. Liu, 'BCAFL: A Blockchain-Based Framework for Asynchronous Federated Learning Protection', Electronics, vol. 12, no. 20, S. 4214, Okt. 2023, doi: 10.3390/electronics12204214.

[76] H. Chen, S. A. Asif, J. Park, C.-C. Shen, und M. Bennis, 'Robust Blockchained Federated Learning with Model Validation and Proof-of-Stake Inspired Consensus'. arXiv, Jan. 09, 2021. Letzter Zugriff 21. Februar 2024 [Online]. http://arxiv.org/abs/2101.03300

[77] D. C. Nguyen et al., 'Federated Learning Meets Blockchain in Edge Computing: Opportunities and Challenges', 2021, doi: 10.48550/ARXIV.2104.01776.

[78] J. Mills, J. Hu, und G. Min, 'Multi-Task Federated Learning for Personalised Deep Neural Networks in Edge Computing'. arXiv, Jul. 22, 2021. Letzter Zugriff 25. Januar 2024 [Online]. http://arxiv.org/abs/2007.09236.

[79] A. Z. Tan, H. Yu, L. Cui, und Q. Yang, 'Towards Personalized Federated Learning', IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 12, S. 9587–9603, Dez. 2023, doi: 10.1109/TNNLS.2022.3160699.

[80] A. Rakotomamonjy, M. Vono, H. J. M. Ruiz, und L. Ralaivola, 'Personalised Federated Learning On Heterogeneous Feature Spaces'. arXiv, Jan. 26, 2023. Letzter Zugriff 9. Februar 2024 [Online]. http://arxiv.org/abs/2301.11447.

[81] R. S. Antunes, C. André Da Costa, A. Küderle, I. A. Yari, und B. Eskofier, 'Federated Learning for Healthcare: Systematic Review and Architecture Proposal', ACM Trans. Intell. Syst. Technol., vol. 13, no. 4, Art. no. 4, Aug. 2022, doi: 10.1145/3501813.

[82] W. Li et al., 'Privacy-preserving Federated Brain Tumour Segmentation', no. arXiv:1910.00962. arXiv, Oct. 02, 2019. Letzter Zugriff 28. Januar 2024 [Online]. http://arxiv.org/abs/1910.00962.

[83] D. C. Nguyen et al., 'Federated Learning for Smart Healthcare: A Survey', ACM Comput. Surv., vol. 55, no. 3, Art. no. 3, März. 2023, doi: 10.1145/3501296.

[84] R. O. Ogundokun, S. Misra, R. Maskeliunas, und R. Damasevicius, 'A Review on Federated Learning and Machine Learning Approaches: Categorization, Application Areas, and Blockchain Technology', Information, vol. 13, no. 5, S. 263, May 2022, doi: 10.3390/info13050263.

[85] P. Kairouz et al., 'Advances and Open Problems in Federated Learning'. arXiv, Mar. 08, 2021. Letzter Zugriff 9. Februar 2024 [Online]. http://arxiv.org/abs/1912.04977

5 — Literaturverzeichnis

[86] Franhofer IPA, ‘Fehler beim Schweißen: Schnell und automatisch erkannt’, Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA. Letzter Zugriff 27. Februar 2024 [Online]. https://www.ipa.fraunhofer.de/de/presse/presseinformationen/fehler-beim-schweissen-schnell-und-automatisch-erkannt.html.

[87] M. Kuehne-Schlinkert, K. Schmidt, E. Schwulera, B. Scharinger, T. Blumauer-Hiessl, und T. Kaufmann, ‘Overcoming the data deadlock - Federated Learning in Industry: Challenges, experiences, and take-aways from a real-world implementation of federated learning in electronics manufacturing.’, 2023.

[88] deltaDAO, ‘deltaDAO Homepage’. Letzter Zugriff 29. Februar 2024 [Online]. https://www.delta-dao.com/.

[89] Deutsche Telekom, ‘KI im Fokus: Telekom kooperiert mit Bosch und der Fetch.ai Foundation’. Letzter Zugriff 27. Februar 2024 [Online]. https://www.telekom.com/de/medien/medieninformationen/detail/telekom-kooperiert-mit-bosch-und-der-fetch-ai-foundation-1058942.

[90] Fetch.ai Foundation, ‘Home | Foundation Web’. Letzter Zugriff 27. Februar 2024 [Online]. https://fetchai.foundation/.

[91] Y. Kim, P. Wang, und L. Mihaylova, ‘Structural Recurrent Neural Network for Traffic Speed Prediction’, no. arXiv:1902.06506. arXiv, Feb. 18, 2019. Letzter Zugriff 28. Januar 2024 [Online]. http://arxiv.org/abs/1902.06506.

[92] C. Latz, V. Vasileva, und M. A. Wimmer, ‘Supporting Smart Mobility in Smart Cities Through Autonomous Driving Buses: A Comparative Analysis’, in Electronic Government, vol. 13391, M. Janssen, C. Csáki, I. Lindgren, E. Loukis, U. Melin, G. Viale Pereira, M. P. Rodríguez Bolívar, and E. Tambouris, Eds., in Lecture Notes in Computer Science, vol. 13391. , Cham: Springer International Publishing, 2022, S. 479–496. doi: 10.1007/978-3-031-15086-9_31.

[93] Hochbahn, ‘Autonome On-Demand-Shuttles – Wie das Projekt ALIKE mit autonomen Kleinbussen den ÖPNV ergänzen soll’. Letzter Zugriff 27. Februar 2024 [Online]. https://www.hochbahn.de/de/projekte/autonome-on-demand-shuttles.

[94] K. Bonawitz et al., ‘Towards Federated Learning at Scale: System Design’, no. arXiv:1902.01046. arXiv, März 22, 2019. Letzter Zugriff 28. Januar 2024 [Online]. http://arxiv.org/abs/1902.01046.

[95] C. Niu et al., ‘Billion-scale federated learning on mobile clients: a submodel design with tunable privacy’, in Proceedings of the 26th Annual International Conference on Mobile Computing and Networking, London United Kingdom: ACM, Sep. 2020, S. 1–14. doi: 10.1145/3372224.3419188.

[96] L. Zhu et al., ‘PockEngine: Sparse and Efficient Fine-tuning in a Pocket’, in 56th Annual IEEE/ACM International Symposium on Microarchitecture, Okt. 2023, S. 1381–1394. doi: 10.1145/3613424.3614307.

5 — Literaturverzeichnis

[97] L. Lyu, H. Yu, und Q. Yang, ‘Threats to Federated Learning: A Survey’. arXiv, März 04, 2020. Letzter Zugriff 9. Februar 2024 [Online]. http://arxiv.org/abs/2003.02133.

[98] A. E. Ouadrhiri und A. Abdelhadi, ‘Differential Privacy for Deep and Federated Learning: A Survey’, IEEE Access, vol. 10, S. 22359–22380, 2022, doi: 10.1109/ACCESS.2022.3151670.

[99] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, ‘A survey on security and privacy of federated learning’, Future Gener. Comput. Syst., vol. 115, S. 619–640, Feb. 2021, doi: 10.1016/j.future.2020.10.007.

[100] V. Shejwalkar, A. Houmansadr, P. Kairouz, und D. Ramage, ‘Back to the Drawing Board: A Critical Evaluation of Poisoning Attacks on Production Federated Learning’. arXiv, Dec. 13, 2021. Letzter Zugriff 24. Februar 2024 [Online]. http://arxiv.org/abs/2108.10241.

[101] C. Fung, C. J. M. Yoon, und I. Beschastnikh, ‘Mitigating Sybils in Federated Learning Poisoning’. arXiv, Jul. 2020. Letzter Zugriff 29. Februar 2024 [Online]. http://arxiv.org/abs/1808.04866.

[102] L. Feng, Y. Zhao, S. Guo, X. Qiu, W. Li, und P. Yu, ‘BAFL: A Blockchain-Based Asynchronous Federated Learning Framework’, IEEE Trans. Comput., vol. 71, no. 5, S. 1092–1103, Mai 2022, doi: 10.1109/TC.2021.3072033.

[103] C. Xu, Y. Qu, Y. Xiang, und L. Gao, ‘Asynchronous federated learning on heterogeneous devices: A survey’, Comput. Sci. Rev., vol. 50, S. 100595, Nov. 2023, doi: 10.1016/j.cosrev.2023.100595.

[104] J. Weng, J. Weng, J. Zhang, M. Li, Y. Zhang, und W. Luo, ‘DeepChain: Auditable and Privacy-Preserving Deep Learning with Blockchain-based Incentive’, IEEE Trans. Dependable Secure Comput., S. 1–1, 2019, doi: 10.1109/TDSC.2019.2952332.

[105] Google Trends, 2024. [Online]. https://trends.google.de/trends/.

[106] IBM, ‘What is a Neural Network? | IBM’. Letzter Zugriff: 22. April 2024 [Online]. https://www.ibm.com/topics/neural-networks.

[107] Wikipedia, ‘DeepDream’, Wikipedia. 22. Dezember 2023. Letzter Zugriff: 22. April 2024 [Online]. https://en.wikipedia.org/w/index.php?title=DeepDream&oldid=1191335336.

[108] A.-L. Barabási, Linked: the new science of networks. Cambridge, Mass: Perseus Pub, 2002.

5 — Glossar

Künstliche Intelligenz (KI) ist die Fähigkeit von Software, Aufgaben auszuführen, die traditionell menschliche Intelligenz erfordern.

Deep Learning ist ein Teilbereich des maschinellen Lernens, bei dem tiefe neuronale Netze zum Einsatz kommen, d. h. Schichten von miteinander verbundenen (digitalen) „Neuronen", deren Verbindungen über Parameter oder Gewichte trainiert werden können. Deep Learning ist besonders effektiv im Umgang mit unstrukturierten Daten wie Bildern, Text und Audio.

Generative KI basiert in der Regel auf sog. Foundation Models und besitzt Fähigkeiten, die frühere KI-Modelle nicht hatten, wie etwa die Generierung von Inhalten. Foundation Models können auch für nicht-generative Zwecke verwendet werden (z. B. zur Klassifizierung der Nutzerstimmung als negativ oder positiv auf der Grundlage von Gesprächsprotokollen) und bieten gleichzeitig erhebliche Verbesserungen gegenüber früheren Modellen.

Large Language Models (LLMs) bilden eine Klasse von Foundation Models, die riesige Mengen unstrukturierten Textes verarbeiten und die Beziehungen zwischen Wörtern oder Wortteilen, sogenannten Tokens, lernen können. Auf diese Weise können LLMs natürlich wirkende Texte generieren und Aufgaben wie Zusammenfassungen oder Wissensextraktion durchführen. GPT-4 (das ChatGPT zugrunde liegt) und LaMDA (das Modell hinter Bard) sind Beispiele für LLMs.

Maschinelles Lernen (ML) ist ein Teilbereich der künstlichen Intelligenz, in dem ein Modell mit Hilfe vieler ihm trainierter oder demonstrierter Beispieldaten Fähigkeiten entwickelt. Algorithmen des maschinellen Lernens erkennen Muster und lernen, Vorhersagen und Empfehlungen zu machen, indem sie Daten und Erfahrungen verarbeiten, anstatt ausdrückliche Programmieranweisungen zu erhalten. Die Algorithmen können sich anpassen und in Reaktion auf neue Daten und Erfahrungen effektiver werden.

Prompt-Engineering bezieht sich auf den Prozess des Entwurfs, der Verfeinerung und der Optimierung von Eingabeaufforderungen, um ein generatives KI-Modell so zu steuern, dass es die gewünschten (d. h. genauen) Ergebnisse produziert.

Blockchain ist eine verteilte Datenbanktechnologie, die aus einer Aneinanderreihung von Blöcken besteht, in denen Transaktionsdaten aufgezeichnet werden. Jeder Block enthält einen kryptografisch gesicherten Hash des vorherigen Blocks, Transaktionsdaten und einen Zeitstempel. Diese Struktur macht die Blockchain inhärent sicher und widerstandsfähig gegen Manipulationen, da jede Änderung in einem Block die gesamte Kette ungültig machen würde.

5 — Glossar

Zentralisierte KI-Systeme konzentrieren Entscheidungsfindung und Rechenprozesse auf eine einzige zentrale Einheit, meist über einen zentralen Server oder eine Datenbank. Dies ermöglicht effiziente Datenverarbeitung, birgt jedoch Risiken wie Single-Points-of-Failure und kann bei hohem Datenvolumen zu Skalierungsproblemen führen.

Distribuierte KI-Systeme sind Modelle, die die Datenverarbeitung und Entscheidungsfindung auf mehrere miteinander verbundenen *Nodes* (Knotenpunkten) verteilen. In diesen Systemen gibt es keine zentrale Kontrollinstanz, was zu erhöhter Ausfallsicherheit und Skalierbarkeit führt. Während einige dieser Systeme Technologien wie Blockchain nutzen können, basieren sie allgemein auf einem Netzwerk von kollaborierenden *Nodes*, welche komplexe Aufgaben bewältigen und Entscheidungen treffen.

Federated Learning (FL) ist ein maschinelles Lernkonzept, bei dem Modelle auf distribuierten Geräten trainiert werden, ohne dass sensible Daten diese Geräte verlassen. Nach dem lokalen Training der KI-Modelle, werden alle Updates zusammengenommen und zu einem verbesserten globalen Modell verarbeitet. Für die Aggregation und die Kalkulation des globalen Modells kann ein zentrales oder distribuiertes KI-System genutzt werden.

Personalisiertes Federated Learning (PFL) ist eine Erweiterung des Federated Learning, bei der das trainierte Modell nicht nur auf gemeinsamen Daten, sondern auch auf individuellen Daten der Teilnehmerinnen und Teilnehmer basiert, um personalisierte Ergebnisse zu liefern. Dies ermöglicht eine feinere Anpassung an die spezifischen Bedürfnisse oder Präferenzen der einzelnen Benutzer.

Edge-Devices sind Geräte am Rand eines Netzwerks, die oftmals Datenverarbeitungsaufgaben lokal durchführen, anstatt sie an zentrale Server zu senden. Diese Geräte, wie Smartphones, IoT-Geräte oder lokale Server, führen Datenanalyse und -verarbeitung nahe der Datenquelle durch, was Latenzzeiten verringert und die Effizienz erhöht. Edge-Computing spielt eine wichtige Rolle in distribuierten KI-Systemen und im Internet der Dinge (IoT).

Über die Autoren

Tom Haverland

ist Referent für Blockchain und Projektleiter am Hanseatic Blockchain Institute und leitet das Projekt W3NOW, das als erstes umfangreiches, vom Bundesministerium für Wirtschaft und Klimaschutz gefördertes Forschungsprojekt den Einsatz der Blockchain-Technologie in der deutschen Wirtschaft untersucht. Haverlands weitere Forschung beschäftigt sich mit dem Synergiepotenzial zwischen Blockchain-Technologie und künstlicher Intelligenz. Als Absolvent der Leuphana Universität mit dem Schwerpunkt digitale Politik hat Haverland die Transformation traditioneller papierbasierter Wahlsysteme in die digitale Welt unter Nutzung der Blockchain-Technologie und kryptografischer Beweis-systeme erforscht, um die Sicherheit und Integrität digitaler Wahlprozesse zu verbessern.

Lukas Beckenbauer

ausgebildet in Media Studies, Machine Learning und Philosophie, ist Doktorand an der Technischen Universität München, wo er untersucht wie Blockchain dezentrale Formen von Eigentum und Online Governance ermöglichen kann. Lukas hat als Forscher mit der University of California Irvine, USA, der Arizona State University, USA, und der Leuphana Universität Lüneburg, Deutschland, gearbeitet. Zudem war er Junior Research Fellow an der ArtEZ Universität in den Niederlanden, im Zuge dessen er Social Change Narrative untersucht hat. Er hält einen Research Master in Media Studies von der Universität Amsterdam.

Impressum

Herausgeberin: Konrad-Adenauer-Stiftung e. V. 2024, Berlin

Kontakt in der Konrad-Adenauer-Stiftung:
Dr. Christian Hübner
Referent Künstliche Intelligenz
Analyse und Beratung
christian.huebner@kas.de
Tel. +49 30 26996-3264

Gestaltung und Satz: KALUZA+SCHMID Studio GmbH, Berlin

Hergestellt mit finanzieller Unterstützung der Bundesrepublik Deutschland.

Diese Veröffentlichung der Konrad-Adenauer-Stiftung e. V. dient ausschließlich der Information. Sie darf weder von Parteien noch von Wahlwerbern oder -helfenden zum Zwecke der Wahlwerbung verwendet werden. Dies gilt für Bundestags-, Landtags- und Kommunalwahlen sowie für Wahlen zum Europäischen Parlament.

Der Text dieses Werkes ist lizenziert unter den Bedingungen von „Creative Commons Namensnennung-Weitergabe unter gleichen Bedingungen 4.0 international”, CC BY-SA 4.0 (abrufbar unter: https://creativecommons.org/licenses/by-sa/4.0/legalcode.de).

Geteilte Autorenschaft. Zitationsinformation:
Beckenbauer, L., & Haverland, T. (2024). Synergien von Blockchain und KI.
KAS4 Innovation. Konrad-Adenauer-Stiftung e.V. Berlin, Deutschland.

ISBN 978-3-98574-223-3

Dieser Bericht diskutiert die Vision eines europäischen KI-Ökosystems, das eine wirkungsvolle Alternative zu bestehenden zentralisierten Paradigmen bietet. Kern des Konzepts ist die Integration von Blockchain-basierten Datenmarktplätzen und föderiertem Lernen, um datenschutzkonforme Ansätze des KI-Trainings zu ermöglichen. Ein besonderes Augenmerk liegt auf der Unterstützung von Start-ups und KMUs, um die Potenziale von KI und Blockchain voll auszuschöpfen und die Trainingsmethoden an die spezifischen Datenstrukturen der Unternehmen anzupassen.

Der Bericht betont die Bedeutung der Stärkung der technologischen Souveränität Europas im Zeitalter der Künstlichen Intelligenz und stellt praxisnahe Lösungen für die Realisierung einer „KI Made in Europe“ vor.

Konrad-Adenauer-Stiftung e. V.