HANSEATIC
BLOCKCHAIN
INSTITUTE

KONRAD
ADENAUER
STIFTUNG

Synergies of Blockchain and AI

#KAS4
INNOVATION

kas.de

Synergies of Blockchain and AI

At a Glance

- This study advocates for the development of a European AI ecosystem, supported by a sovereign data marketplace and federated learning. The proposed approach aims to bolster Europe's technological sovereignty in the AI era by outlining a blockchain-based vision for "AI Made in Europe."

- **Digitalization and Data:** Leveraging the technical synergies of AI and blockchain allows businesses to collaborate in a manner compliant with data protection regulations. This is increasingly crucial given the growing digitalization of the economy and heightened legal requirements for data management, particularly within the European Union (EU).

- **Enterprises:** Start-ups and small to medium-sized enterprises (SMEs) stand to gain significantly from the synergetic relations of AI and blockchain. These business entities have traditionally faced high market barriers and costs related to data protection and AI training, which has limited their ability to capitalize on the economic potentials of artificial intelligence when compared to larger corporations.

- **Digital Data Marketplace:** The combined use of blockchain and AI applications can facilitate the creation of decentralized digital infrastructures. These infrastructures can be utilized by individuals, start-ups, or companies to collaboratively train AI models and use their results or even trade on their data processing algorithms without necessarily revealing the contents of this data, therein ensuring privacy and data compliance. Such digital data exchanges can be managed by individual or corporate actors, promoting the establishment of publicly accessible data markets.

- **Use Cases:** Potential applications for digital data marketplaces based on blockchain and AI include data utilization in healthcare, the development of smart cities, industrial AI applications e.g. for improving machine production processes, and the use of mobile devices for AI training. Additional areas of application could be the metaverse and the traceability of data to verify compliance with AI regulations.

Table of Contents

Preface

Preface

The German economy in general, and particularly the SME sector – the backbone of society – is increasingly facing the challenge of keeping pace in AI-enhanced global competition. In addition to many factors, the need for effectively integrating technological innovations into company operations plays a significant role. However, the adoption of new technologies often involves significant personnel or resource expenditure and is not feasible for many SMEs. Also, the German startup scene is confronted with difficulty in facing the latest technological developments. Meaningful and necessary regulations on data protection and the development of digital products limit the enthusiasm for innovation, increase operational costs for young companies, and correlate with the exit of ideas and investments to other economic zones.

As this report shows, future technologies such as AI and blockchain offer promising foundations to support startups and SMEs within the EU in their technological goals and to enable data-protection-compliant collaborations. Thus, they have the potential to transform collaboration and efficiency between companies in the EU and the German economic area and allow startups and SMEs to effectively participate in the economic potential of artificial intelligence.

The authors discuss new and potentially transformative technologies that could enable the implementation of a comprehensive and collaboratively designed data economy, following the model of European regulation. Pioneering in this context is the idea of creating decentralized digital infrastructures based on the synthesis of blockchain and AI, which individuals, startups, and companies alike can use to collaboratively train AI models and utilize these models results, or even exchange algorithms for data processing, without necessarily disclosing the contents of their own data. Data protection-compliant AI development "Made in Europe and Germany" has the potential to become a hallmark of the German economy and to increase global competitiveness. With the solutions described, the technological maturity of the German industry and its "hidden champions" can be leveraged to enable developments based on high-quality data for startups and other companies. This can create an energetic ecosystem in Germany that benefits from the unique niche-leadership position of German SMEs and enables young startups to tap the competitive advantages of building on top of large, and distributed, and accessible AI infrastructures.

For Germany, it is now important to be creative and innovative in finding new pathways in the face of the current AI revolution. The solutions suggested in this report, are in the interest of the technological sovereignty of the European and German economy and thus a direct response to the objectives of the European Data Strategy and the EU AI Act, as well as the Federal Government's Data Strategy of 2021.

Prof. Dr. Isabell Welpe
Chair for Strategy and Organization
TUM School Of Management

0

Introduction

0 — Introduction

In recent years, blockchain and Artificial Intelligence (AI) have been among the top trending topics in the tech sector. Often mentioned with other 'techie' buzz-topics such as the Metaverse, the Internet of Things, Crypto Assets and Digital Tokens they are understood by many to constitute the next phase of our digital economies. While blockchain has become increasingly popular since the release of the Bitcoin Whitepaper by anonymous author (or authors collective) Satoshi Nakamoto in 2008, AI has become a buzzword the latest with the release of web application ChatGPT in late 2022 and, likely, will stay trending for the years to come. Estimated to reach a valuation of $420 billion USD by 2025 [1], total market size of the AI-based economy is expected to grow somewhere between $1,6 - $2 trillion USD by 2030 [1], [2]. And Statista predicts the global blockchain technology market to grow from $5,85 billion USD in 2021 to USD $1,235 trillion by 2030.

While these estimates should be taken with a word of caution, they do however speak to the enormous transformative potential that is expected to come from the implementations of blockchain and AI alike. At the core of this transformation is each technology's unique capacity to process or store data in new and innovative ways. As the following pages will outline, AI is essentially an excellent response to process the ever-growing amounts of data that constitute digitalized societies at the early 21st century, turning this data into human-readable content. While blockchain technology has been implemented in its current form to enhance consensus, communication, and trust between (mostly) pseudonymous third parties on the internet.

However, while the setup of blockchain technology is comparatively cheap, the successful development of AI models as well as their successful application is a cost-intensive endeavor. Recent estimates state that each training cycle of OpenAI's popular GPT-3 model has cost at least $5 Mio. USD [3], while completing model production required more than $100 Mio. USD of investments [3]. These costs are increasing with growing capabilities, and hence training complexity, of AI models [4]. Therefore, while large digital platform companies (such as Google, Apple, Facebook, Amazon, Microsoft (GAFAM)) are already preparing for the 'biggest revolution since the invention of the internet' [5], [6], [7], it remains an unsolved question how individuals, e.g. everyday internet users and citizens, as well as organizations with stronger budgets constraints, e.g., Small and Medium Enterprises (SME), can find avenues to participate in and be empowered by the next innovation wave of the digital economy [8].

As this report will outline, the successful implementation of blockchain as part of AI infrastructures could help to mitigate this effect of platform-centric AI and help to empower users, as well as smaller and medium size businesses and startups alike to collaboratively participate in the emerging AI-economy. For example, AI, a technology that is currently designed to operate in a rather closed-off, corporately-owned settings, could be produced in a more distributed and collectively shared and collaborative setting; while at the same time, blockchain can enhance AI systems by adding transparency, reliability, and equal access to digital infrastructures, additionally providing tools to trace ownership and assign authenticity to digital data, therein helping to establish advanced forms of trust within our digital environments.

At the same time, emphasizing the potential of blockchain and AI towards envisioning collaboratively shared digital infrastructures will need to be accompanied by a shift in how we think about data. Unfortunately, it appears that we have become used to thinking that user-centric data can be owned by corporations and governments, only. Subsequently, the question what would happen if each and every one of us could share and trade our data has gained increasingly less attention in the past years. Still, in the midst of the current AI revolution we could (and maybe should) take this question even further and possibly ask: what could happen if we used commonly owned and sourced data to feed AI models, which, also, would be commonly owned and shared? In embracing data within this new paradigm, we could initialize a debate that offers moving from "Big Data" to "Shared Data" [9]. Therein emphasizing the need for collaboration and participation in the digital value creation chain of our democratic societies.

As Alex Pentland, director of the MIT Connection Science initiative, argues, in working to think of data as commonly shared and accessible resource, we should also emphasize the fact that data is increasingly emerging as a new asset class; similar to how oil has been understood as driving the economies of the 20th century.[i] At the same time, and in a positive opposition to oil, data has the benefit that it does not deplete itself through usage. In fact, in being shared, compounded, and re-utilized it can keep on adding value over time; possibly working to train the AI models that could drive future visions of a digital commons. Quoting Mr. Pentland again, "data is now central to the economy, government, and health systems, so why are data and the AI systems that interpret the data in the hands of so few people? Communities without data about themselves and without the tools to use their data are at the mercy of those with data and tools" [9].

So far, the European Union has proven to be strong in understanding the potential economic and social imbalances that data can cause, if left unchecked in the hands of a few, powerful third parties, and has worked towards finding effective ways to regulate and frame these emerging technologies. In outlining a European Data Strategy in 2020 [10], the EU Data Act in 2024 [11], as well as the issuing of the AI Act in 2023 [12], the EU has therein established important foundations for a European vision of the Internet. The former document therein asserts the importance of European Data Spaces - genuine data markets, open to data from across the world [10, p. 4] - that provide an "open, fair, diverse, democratic, and confident" [10, p. 2] environment, while ensuring trust and personal privacy laws as defined by the GDPR [13]. On the other hand, the EU's AI Act provides a comprehensive regulatory framework for emerging AI technologies, therein emphasizing the importance of data regulation compliance, as well as data governance, record-keeping, transparency, and access control [14].

As we will outline in this report, the combination of blockchain and AI can help to realize this new vision of shared data, by fostering technology that is decentralized, commonly shared, and distributed. Usage of these technologies could therein support the EU's goal to reach a share of the global data economy by 2030, that "at least corresponds to its economic weight" [10, p. 2]. If well sourced, the combination of blockchain and AI might therefore offer a strong foundation to empower the collective of European societies.

Also, for Germany, a nation whose economic success is based on its "hidden champions", Small and Medium-Size Enterprises who have secured themselves a position as world leaders in their respective business and industry niches, the creation of such infrastructure could be promising. Normally, for many SMEs the creation of an internally sourced AI-model is usually overly cost-intensive, while simultaneously, the usage of centralized AI models, such as ChatGPT cannot be interpreted as fully compliant for securing corporate confidentiality. Additionally, many SME's do not have sufficient data, as well as expertise, in building internal solutions to maintain effective AI-model infrastructures. As stated earlier, if larger and more

[i] In this sense, data has become a new essential factor of production, which can be considered as valuable as other resources that drive our economies, such as maintaining a skilled workforce and being able to draw upon financial capital [9].

globalized corporations are increasingly able to source their own AI models for production, while smaller corporations and startups are being left behind, this might result in a productivity gap, causing economic deficits for those nations who are not catching up with the rapid pace of AI-innovation and -enhancement.

Highlighting the mutual benefits of blockchain and AI to empower individual users, as well as smaller and medium size businesses inside the EU, this report will then begin with a description of the core technologies in question. It will start with an initial and brief description of blockchain and AI technologies and highlight their historical context, in order to illustrate how these technologies have come to emerge as effective mediators for the many communicative crises our societies find themselves in today. Therein, we will discuss the central features and innovations that blockchain and AI do bring to the table. In this sense, this initial chapter will provide a brief summary and overview of the key technologies and innovations behind the latest bust in contemporary AI models, while at the same time showcase, how shortcomings in these models could possibly be approached by innovations in the blockchain sector. We will therein avoid discussing the technological specifications of blockchains in too much detail, as many of these have already been approached in the KAS innovation report #4 on tokenization [15].

The second chapter will focus more closely on the ways in which the essential affordances of both technologies in question can be utilized and merged in order to produce digital infrastructures that could truly help us in establishing a shared data economy, as well as to advance European ideals of equality, privacy, and civic participation. Here, we will specifically focus on discussing how the combination of blockchain and AI could open the unique opportunity to establish a European Data Space. This would imply the establishment of open and publicly owned data marketplaces, distributed machine learning infrastructures, and the clear assignment of authorship and subsequent rewarding of individual data contributions that could be made possible on top of distributed digital architectures that are powered by merging blockchains with AI.

The third chapter of this report will provide a deep dive into the technological foundations that would make such Data Exchanges possible, specifically discussing how a type of Machine Learning technique called Federated Learning could be implemented, to enable individual users and businesses to collaborate with their data, without compromising their digital privacy. Additionally, we will briefly discuss how the technologies in question could be applied in future industry application scenarios such as the Metaverse and the establishment of trustworthy forms of AI.

As we will see throughout this report, many of the technologies discussed are still under development – therefore finding concrete avenues for their implementation will need to be approached as a work in progress. We also cannot stress enough the importance of including diverse perspectives and a considerate approach into the establishment of the next phase of digital infrastructure early on, as many current implementations of our digital spaces have the unfortunate tendency to replicate bias and segregation [16], [17], [18], [19]. A tendency that, among other crucial elements to consider, operates counterintuitive to democratic ideals, including principles of equal access and fairness. The last section of this report will therefore provide a summary of our findings as well as a critical discussion of future avenues to consider, in order to re-emphasize the possible technical as well as social issues that come with the development of the technologies mentioned in this report. In closing our argument, we will offer concluding thoughts and avenues for future exploration that could help to productively implement blockchain and AI as part of our shared digital infrastructures.

1

Understanding Blockchain and AI

1 — Understanding Blockchain and AI

1.1 — Technological Foundations of Blockchain and AI

Visions of machine-simulated forms of intelligence can be traced to have captivated human imagination for many centuries. From its early beginnings manifesting in the idea of Golems [20], [21] – mechanical entities composed of clay and magic –more contemporary, popular examples are the 18th century's Mechanical Turk[ii], 19th century Mary Shelley's Frankenstein Monster, or 20th century's R2D2 and C3PO (Star Wars). Nevertheless, truly functional models, able to automate or simulate human thought, have been works of science-fiction for a long time. While, by early 2024 their possibilities and advancements are debated non-stop. Given these recent advancements in AI research and the immense success of AI-Chat Bots, such as OpenAI's ChatGPT, Anthropic's Claude, and Google's Gemini, they increasingly seem to become central features of our daily collective experiences. Considering the latest buzz around AI tech, it therein seems noteworthy to mention that still, today's AI models do not really "think" in the ways in which we are used to think about the meaning of this word. Rather, they refer to a subset of mathematical functions in order to compute the likeliness of outcomes, the preciseness of which is so close to our own conceptions of the world that we are likely to compare the output of their mechanics to patterns of human thought. Unsurprisingly, the technology behind what modern models are doing is so intricate and exciting in its possibilities that questions of when we will achieve AI models fully capable of modeling complex human thought, often referred to as Artificial General Intelligence (or AGI), have become increasingly more frequent in the last years. While discussions of the possibilities of AI and predictions of future impact are at an all-time high (s. Figure 1), we believe that any AI-maximization discourse should be taken with a word of caution.

Already today, there exists a growing shortage in energy [24] as well as graphic processor supply [25], [26], both central resources needed to power the current generation of highly-compute intensive AI models. Concerns around energy shortage, while keeping climate goals in check, are so strong that, for example, Microsoft is already discussing investing in their own nuclear strategy to power future AI growth [27]. At the same time, it seems clear to many observers of these new technologies, that the coming advancements in the field of AI will have a tremendous, and most likely, transformative impact on our everyday lives.

[ii] A mechanical chess machine, that claimed to operate similar to modern chess computers. While, in fact it was operated by a small human hiding inside its casing playing against the outer opponent.

Google Trends

Frequency of search requests for the terms "artificial intelligence", "AI" or "ChatGPT" on Google in Germany from 2004 to today

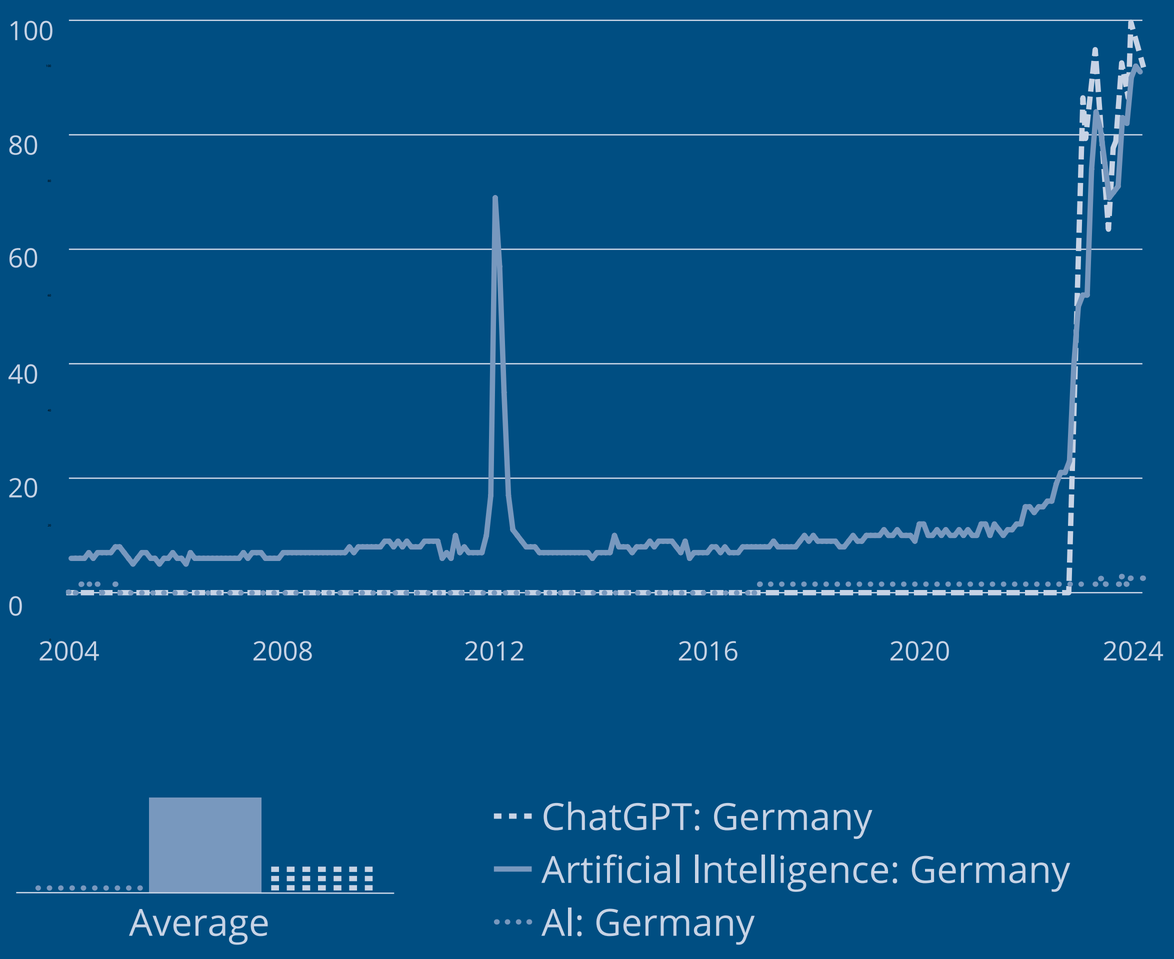


Figure 1

Source: Adapted from [22].

Deep neural network

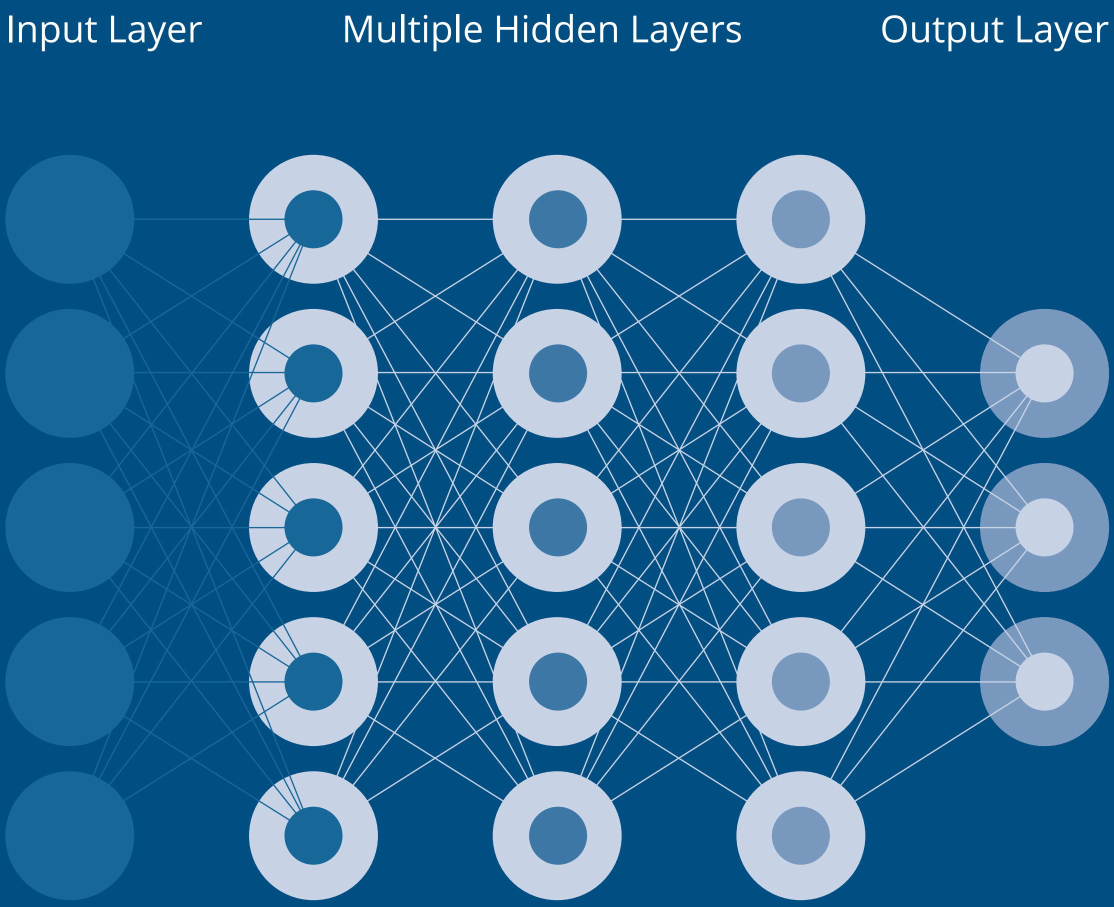


Figure 2

Source: Adapted from [23].

Understanding the art and science behind creating (both, simple and sophisticated) AI models is a process that would cover multiple books. Nevertheless, the following pages will aim to provide a general understanding of their core mechanics; possibly highlighting their potential shortcomings whenever appropriate. To begin with, we would need to emphasize that Artificial Intelligence and Machine Learning, while often used interchangeably, are not necessarily one and the same thing. Artificial Intelligence is generally concerned with the building of systems “that simulate intelligent behavior” [28, p. 1]. While Machine Learning, in fact, is a subfield of AI research that is particularly concerned with the question of how mathematical models, data, and algorithms can be utilized to mimic processes of human thought [29].

Both of these concepts are not new in the history of science and have been around in their current meaning at the least since the 1950s [30]. However, what had changed in this timeframe are the intricacies of the models. Early machine learning models were able to compute simple processes, as for example finding the shortest route between two points A and B in a given field [30], [31], or quickly skimming through a larger database, e.g., a phonebook, in order to find the single contact with the name “Jane Doe” as quickly as possible. The capacities of modern machine learning models have moved far beyond that and are now running on *neural networks*; digital data pipelines that are designed to mimic the operations of neurons in the human brain. These models are usually trained by enabling the neural network to balance its processes of meaning making by itself. Therefore, they include an interesting twist: As a part of model composition is targeted to allow a given models to train itself, no one today truly knows how the completed models work in detail [28, p. ix].

The functional logic of neural networks and successes of subsequent innovations, like ChatGPT, stand witness to the fact that contemporary advancements in AI research have managed to simulate human cognitive processes much more closely than many machine learning proponents in earlier phases of AI development - at least until the implementation of the World Wide Web in the 1990's and the subsequent vast amounts of data being generated - had thought possible [30]. Being the essential innovation behind this success, *neural networks* are a type of machine learning model that is designed to simulate the learning structure of the human brain. In simple terms, it is mimicking the way in which biological neurons send signals to each other [32]. To "learn", neural networks rely on large amounts of training data, which they process in order to improve the accuracy of their output over time [32], [33]. The process of training neural networks over many iterations, with the goal of improving their accuracy and to achieve effective outputs, is then called "Deep Learning". Deep Learning trained Neural Networks (s. Figure 2) are widely considered the most powerful and advanced machine learning models to date and are popular in everyday use-cases [28, p. 1].

In their essence the high quality of their output is powered through a multi-layered classification process, designed to identify, and describe intricate patterns and objects in the data. These objects can represent any type of "real world" input, which subsequently gets converted into human-readable output by the model. Deep Learning Model use-cases are for example, the translation of text from language A (say, German) to language B (say, French) (e.g., as implemented by the German StartUp Deepl.com), the transcription of recorded audio-content into written text (e.g. as exemplified by the Berlin-based Startup Speech-Text.AI), or, in a more advanced case, the conversion of textual input into images that are generated by the model (as exemplified by OpenAI's Dall-E, or the GenAI project "Stable Diffusion", whose underlying algorithm was conceived at LMU in Munich).[iii]

[iii] For a detailed overview of the various kinds of data generative AI models are able to process, we recommend Prince [28, p. 6].

A popular example of how pattern detection in Deep Learning Networks (or 'Deep Neural Networks') operates, is exemplified by the DeepDream software. DeepDream was released by Google engineer Alexander Mordvintsev in 2015 [34]. Figure 3 shows a sequence of images created by DeepDream [35]. In this particular example the model was trained to identify patterns of dogs and, subsequently, deliberatively overclocked in order to achieve a 'dreamlike' outcome for any given picture the model processes [36], [37]. While the top of the three pictures showcases the original depiction of three Moon Jellyfish in water, the middle picture highlights the output of dogs the model had "identified" after about 10 iterations. The last picture highlights the number of dogs identified after 50 iterations in total.

Despite its overclocked output algorithm, DeepDream is a good example to highlight the essential strengths and, at the same time, biggest weaknesses of Deep Learning models. As the middle section and last picture in Figure 3 show, Deep Learning models are excellent in detecting human-readable patterns in large amounts of seemingly unstructured data. While at the same time, they can be "overtrained" using highly homogenous or insufficient quality data samples in order to detect data patterns where there are none.[iv] The creation of effective Deep Learning models is therefore extremely dependent on ensuring that sufficient diversity in training data, as well as that quality of input is ensured for the model to function effectively, while minimizing bias.

The above-mentioned examples show that Deep Learning models represent a fundamental change in how the computation of data can be performed. Most of our online applications to date, have been written line by line in a given programming language; usually, by one or many (human) programmers. Deep Neural Networks, however, create their own structural patterns (or code) in order to compute data as input, and to predict the likelihood of outcomes as output. While previous software was designed to compute by executing lines of code of a human programmer, neural networks are now actually able to compute themselves [38]. They have become self-executing agents, able to perform advanced tasks in problem solving and predictions, and subsequently able to simulate processes which previously have been unique to the realm of human cognition.

[iv] The Deep Dream software therefore creatively utilizes the effect of Deep Learning models "hallucinating" (making up patterns to fit uncertainty) to create visually appealing, 'psychedelic' imagery.

Deep neural network

The original image (top) after applying ten (middle) and fifty (bottom) iterations of DeepDream, the network having been trained to perceive dogs and then run backwards.

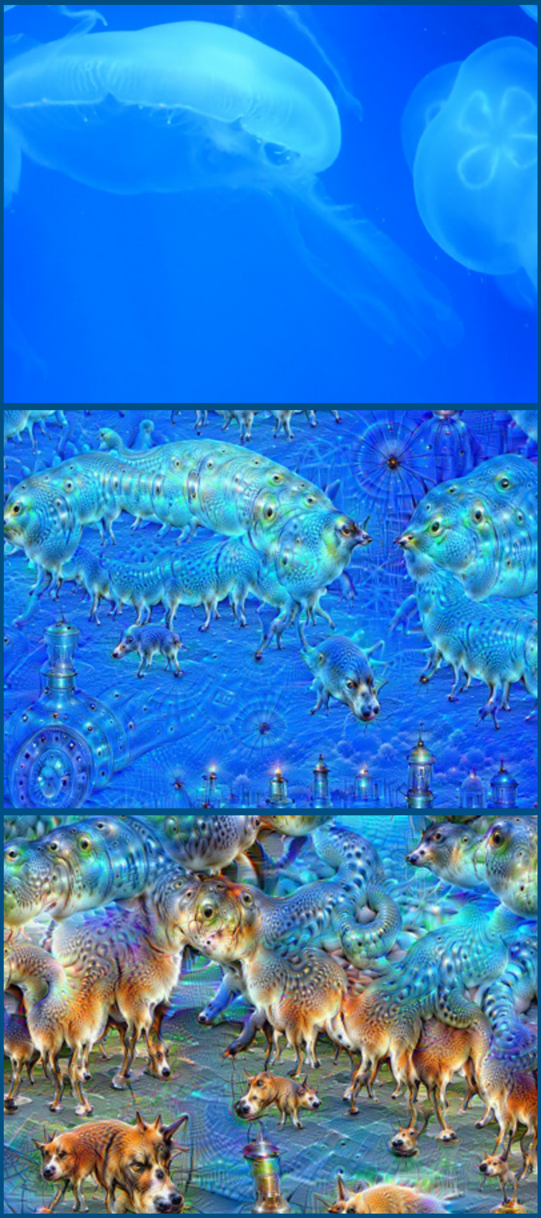


Figure 3

Source: [35].

In this sense, they represent a shift from what was previously referred to as *Software 1.0* to *Software 2.0*, and subsequently a shift in the ways our digital economies might be structured and executed in the future. *Software 2.0* then represents the paradigm that algorithms will increasingly be able to write and develop themselves, whereas the production and feeding of data into the algorithm will become increasingly more relevant [39], [40]. As a consequence, the role of the Software engineer might begin to shift from a producer of code lines that compose a program (*Software 1.0*), towards more of a Data Architect, who is feeding data and insights into AI models in order to produce accurate prediction outcomes (*Software 2.0*). Naturally, programming by lines of code will not become entirely replaced in future software development processes but might increasingly be handled by the AI itself (current ChatGPT models are already very good at producing code output for simple programs). Instead, software applications will come to be able to rely on predictions and increasingly more humane interfaces (as exemplified by Advanced AI chatbots, such as Claude, Bard, or ChatGPT). While their effectivity and performance will be less bound to the clear set of instructions a programmer has provided, but increasingly by the amount and quality of data that gets integrated into the model, in order to provide effective outputs, reduce bias and the risk for hallucinations, and to provide predictions that are ever more accurate representations of our actual world.

As stated previously, a likely consequence of this dynamic will be that high-quality input of data is becoming more important than ever for the effective and productive functioning of machine learning models and their surrounding software applications. Thereby, the effective functioning of our societies and performance of our market economy will come to depend on models that are accurate, use high-quality and representative data, and are well-trained. Subsequently, data does not only operate as a new resource in the 21st century (“data is the new oil”), but it also represents a new means of production to enhance social life and business.

Unfortunately, this vision comes with a caveat: the data economy models that are widely used nowadays are designed to collect and hoard data on the centralized server-infrastructures of digital platform providers [5]. While the producers of data, internet users and citizens of a nation, are usually excluded from the process of value extraction [5]. However, if users and citizens are the producers of their data, and as data is going to become a central driver for effective market growth in the 21st century, the value extraction of this data should become a lot more diversified and designed to benefit those users and their communities as a whole. Running future AI models on highly centralized infrastructures, will likely imply that the power of these AI models remains concentrated in the hands of a few, mostly, corporate actors. Among other factors, this could cause growing social inequality, induced by a gap in performance and opportunity of access between individuals, startups or smaller businesses, and large corporations [8]. Additionally, in the current technological landscape the creation of models is highly expensive and requires vast amounts of data, both factors, which only very large corporations can afford to date. However, if only a few organizations can afford development and ownership of this powerful technology, they can also control how it will be deployed in the future, effectively holding the potential power to overly influence the mindset and decisions of many of its users [41]. The downsides of these forms of highly centralized power have already been exemplified in the past. For example, when in 2012, then Facebook Inc. (Meta) experimented with over 680,000 users' News Feeds, to identify if negative or positive content would effectively affect their emotional state [42]. Or, e.g., when Amazon has been accused to allegedly leverage its market power by deterring sellers to offer lower-priced products on non-Amazon retail websites, while driving them to raise prices on Amazon by charging high fees [43]. One instance among several that have led to the U.S. government file an anti-trust lawsuit against Amazon at the end of 2023.

Lastly, governments will want to ensure that the data of their own citizens is kept safe and in respect of privacy laws. Past scandals like Cambridge Analytica have illustrated clearly how abuse of personal data can be weaponized to attempt and change political outcomes and campaigns. Establishing safe and locally owned AI models could subsequently ensure independence of these critical infrastructure for any societal and economic zone, as well as respect the autonomy and data privacy of citizens. One way to mitigate these risks, could be approached by rethinking how we can design machine learning models that are collaborative, fair, and structured to effectively gather data for advanced applications and AI modelling purposes. The following pages will provide a review of how the merging of AI technologies with blockchain could present an appropriate step in this direction.

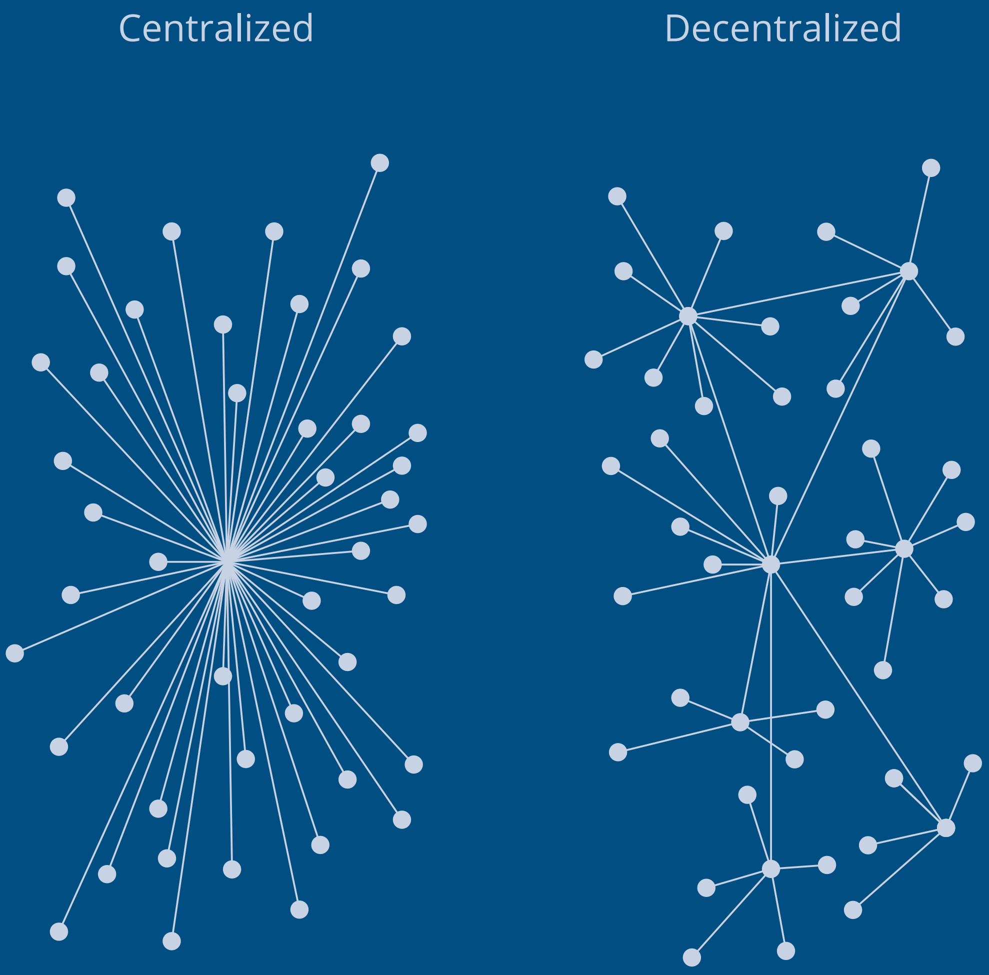
1.2 — Synergies of Blockchain and AI

With historical origins in the field of cryptography and the cypher punk scene, blockchain as a technology has gained widespread popularity with the introduction of Bitcoin, by its anonymous inventor (or inventors collective) known under the alias Satoshi Nakamoto [44]. In its essence, blockchain can be understood as a technology that allows the establishment of consensus between actors that do not necessarily know each other or are not in a position to trust each other. It does so, by introducing a distributed network architecture that records each interaction between these actors in an immutable database, therein effectively replacing the need for a middleman and offering a form of social exchange that has often been referred to as 'trustless' communication. As stated above, the previous Token Study [15] had already provided an essential deep dive into the technological specifications of blockchain technologies, which is why this report will aim to avoid discussing these specifications in depth. However, we would like to point out the essential features of blockchain technology that are relevant for the subject of this report. We will provide a brief introduction of each of these features in the following:

Decentralization

One of the core design features of (public) blockchains is their distributed approach to communication (s. Figure 4) [45]. In this distributed setting the blockchain is composed of a multitude of *nodes*, which ensure synchronicity and stable information exchange across the network. Users can therefore communicate directly with each other or send blockchain-based assets (more below), eliminating the need for a middleman [46]. In comparison to established online platform architectures, such as Google, Microsoft, Facebook, etc. blockchain-based architectures allow for a more egalitarian access to online infrastructures and provide the possibility to tokenize assets on the chain. Consequently, users can be compensated for their activities, or, as in the original Bitcoin use-case, use the network to transfer online currency for general payments and value storage.

Centralized vs. Decentralized Network



Centralized

Decentralized

Figure 4

Source: Adapted from [45].

Security & Immutability

Its origins in cryptography research have provided blockchains with the advantage of maintaining high degrees of security on-chain. At its core, each transaction on the chain is verified and secured by all nodes in the network. Subsequently, transactions and their timestamps are stored in the node as write-only, which provides the data entry with immutability once verified. As each node keeps a copy of each transaction and allows for changes only if a majority of nodes in the (decentralized) network record this change, blockchains are regarded as tamper-proof [47]. As they can only be modified if a given attacker gains control over more than 50% of nodes in the network[v]. Blockchains are also able to draw on strong cryptographic methods during their transactions, which adds another layer of data security [50], [51]. However, as many popular blockchains keep their databases with public reading access, and every node operator usually keeps a copy of the entire database of transactions, additional modifications to the system are required in order to ensure full data privacy for individual users or use-cases [51]. We will discuss these caveats in more detail in subsequent chapters.

Traceability

One of the core advantages of using blockchains to store information, is that, due to its sequential form of information processing and storage, every data point gets assigned a unique set of identifiers. On the one hand, these identifiers allow traceability and verifiability of information that is stored on-chain. For example, the originator (or 'author') of a given dataset can be identified and validity for information can be assured. At the same time, this feature of traceability allows for understanding how information flows on the network; subsequently, data abuse and leakage can be quickly identified, and overall risk of theft is decreased.

[v] This threshold accounts for Proof of Work blockchains. For comparison to an alternative consensus system, such as Proof of Stake, we recommend [48], [49].

Privacy

While Blockchain in its early days has been heralded as protecting the anonymity of its users, research, as well as recent developments and studies have shown that full privacy and anonymous use of blockchains is not fully granted in many cases [51]. One reason for this is that the above-described features of traceability and transparency in data transactions, also imply that with sufficient data and IT knowledge, the originator of a given transaction can be traced and possibly identified. To mitigate this effect, advanced encryption algorithms are available, which, due to their own inherent complexity, we cannot fully discuss in this report. For further reading, we recommend [52], [53].

One of the core tenets of many technologically founded discussions concerning blockchain, is then the idea that its affordances can be seen as an attempt to instill ethical accountability and possibly even fairness into digital infrastructures. Already with the publication of the Bitcoin White Paper, this ethical tenet has been a core momentum of Nakamoto's argument [44]. And despite the many cases of fraud and misconduct in the "crypto"-area - blockchain's widest-known and most capital-intensive application context to date - for many blockchain proponents it has become an integral part of their narrative to work on building digital infrastructures that are steadily becoming more accessible, egalitarian, and fair [54], [55]. In the previous chapters we discussed the technological foundations of AI, as well as possible shortcomings of overly centralized AI infrastructures. The goal of the following sections will be to provide an overview of how the promises of blockchain tech might help to mitigate these risks.

„We are embarking on a transition from **big data** to **shared data**, in which the knowledge that emerges from data is starting to securely move in our society."

Pentland et al, 2021

2

Data Exchanges

2 — Data Exchanges

2.1. — Introducing Data Exchanges

So far, this report has discussed the emergence of blockchain and Artificial Intelligence as two key technologies of early 21st century societies. We emphasized the relevance of AI to effectively process the vast amounts of data that digital infrastructures (the Internet being among these) produce every day and to generate insights from these data that are meaningful and human readable. At the same time, we have emphasized how blockchain was designed to address shortcomings of today's highly centralized digital platform architectures. We also highlighted the ethical mission that often is connoted with blockchains, once more technologically informed discussions are being considered. The ultimate relevancy of the fusion of blockchain and AI is therein the fact that data is becoming an increasingly central commodity of economic production cycles. While at the same time, it can allow for data to be provided for various kinds of market participants, big tech, large-size corporations, SMEs, and everyday Internet users alike.

The power of the blockchain is then that it allows to distribute ownership across its infrastructure and to manifest it, in the form of transparency, equal access, and traceability, for each of its users. Applying these features to AI infrastructures could prove a powerful combination and present an opportunity to build digital infrastructure that is more closely aligned with the European Union's set of values, fostering the protection of user privacy [14] while ensuring collaboration in emergent Data Spaces [9].

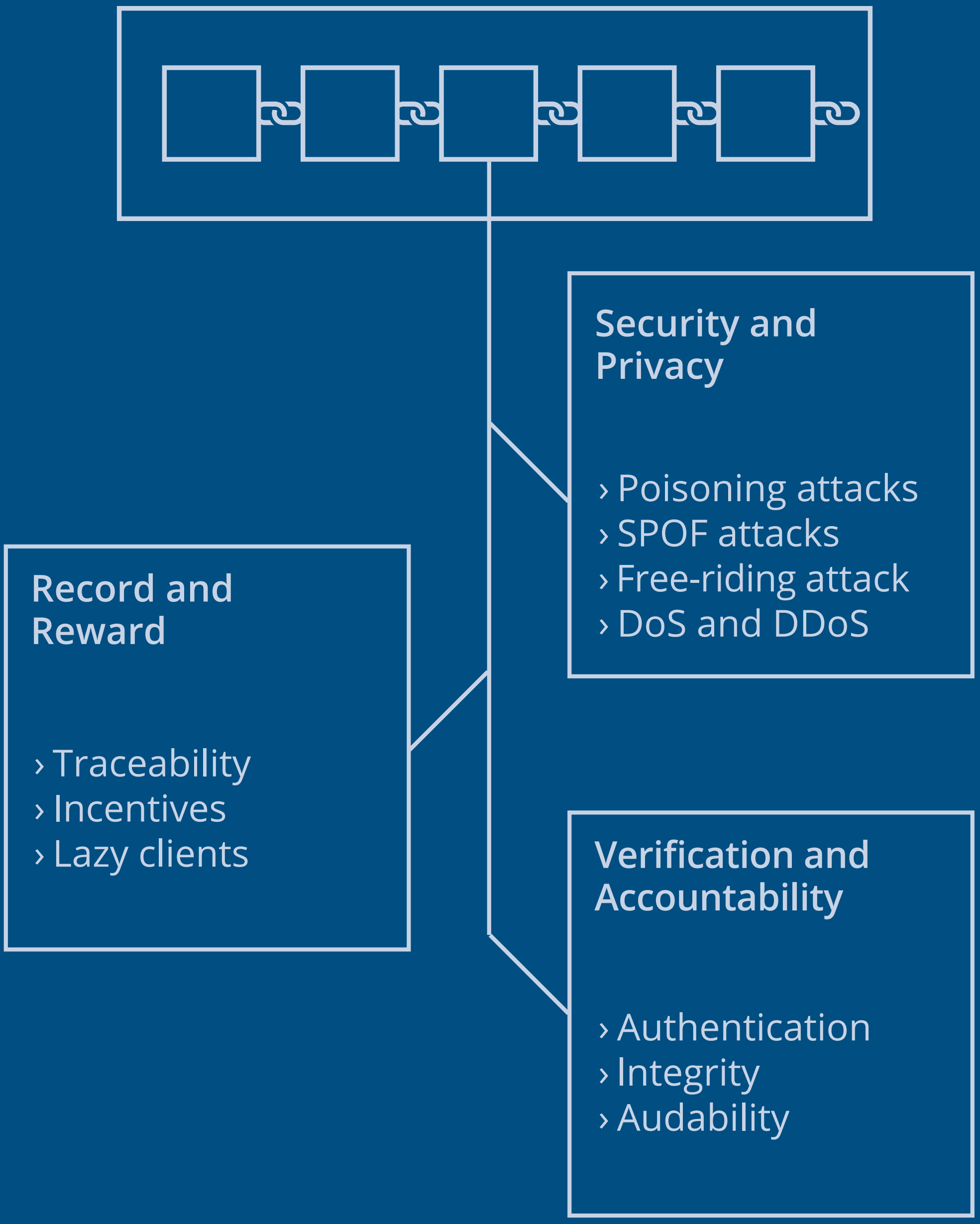
At the same time, in the current data-economy, the effective training of machine learning models is facing two challenges: On the one hand, in many industries data exists in the form of isolated islands [56]. These islands can exist between companies, but also between departments of the same corporation. As internal competition might not encourage teams to share data, or simply, the expertise and resources for establishing a connected dataset is not provided. On the other hand, companies are faced with growing demands for maintaining data-privacy regulations, which make it increasingly difficult to create connected databases and train Machine Learning models [56]. These 'Data Silos' are thereby one core-impediment for effectively unlocking the possibilities of ML models. As we mentioned at the beginning of the report, data as an asset class is as central to the 21st century as natural resources, such as oil, were to the 20th. However, there exists a need to implement more effective solutions towards leveraging this new asset class, which also comes with a unique quality: while oil can only be used one unit at a time, data could be shared among many [9], therein possibly enhancing its overall positive net-effect.

2 — Data Exchanges

One way to unlock the potential of data, while ensuring privacy compliance and control over one's own datasets, could be the establishment of Data Exchanges. Data Exchanges could be imagined as commonly shared (German, European, or even globally accessible) digital hubs that would allow anybody, businesses, and individuals alike, to provide their data as assets on an open data market. Subsequently, they could be traded to train advanced AI models, providing the data-issuers with some form of compensation in exchange. Amateur weather stations already contribute to meteorological forecast and more accurate weather predictions today. For example, in the US, the Citizen Weather Observer program consists of more than 7000 stations that are sending a self-reported number of 50.000 – 70.000 observations per hour [57]. Upon undergoing quality control, this data is then used by major US institutions, including the National Weather Service, the National Ocean Service, or NASA [57].

Whereas amateur weather stations are currently operating mostly self-funded, the establishment of Data Exchanges could provide an incentive for different users and use-cases to share their data with third parties that could improve their services and predictions as a consequence. Data Exchanges could hereby function as intermediaries, effectively bridging the gap between data producers and users [58] and compensating both sides fairly for their engagement. Furthermore, confidentiality of data provided could be ensured by using a distributed machine learning approach that is using locally stored data of a given market participant to locally train a given ML model. While subsequently, results of these respective sub-models are communicated on the network in order to establish a new iterative consensus on a larger, global model [56].

State-of-the-art: blockchain-based federated learning approaches



Security and Privacy

- › Poisoning attacks
- › SPOF attacks
- › Free-riding attack
- › DoS and DDoS

Record and Reward

- › Traceability
- › Incentives
- › Lazy clients

Verification and Accountability

- › Authentication
- › Integrity
- › Audability

Figure 5

Source: Adapted from [59]

An advantage of this so called ‘federated learning’ architecture is “the decoupling of [global] model training from the need for direct access to the raw training data” [59]. Privacy and security risks are hereby reduced due to minimized data transactions and the local processing of sensitive or confidential data [59], [60]. In this manner, large AI models can be trained step-by-step without revealing the content of local datasets, thereby minimizing the risks of data breaches or privacy infliction [59]. Merging this process of ‘federated learning’, with the unique features of blockchain technology (as described in the previous section), a shared European Data Exchange could benefit of the following features:

Data Authorship

The widespread use of Generative AI (GenAI) models has raised concerns about copyright and opened legal debate on whether training data can be used without reference to its origin, such as users, authors, or creative works. A notable example of this concern is The New York Times’ lawsuit against OpenAI and Microsoft for allegedly using its content without permission [61]. Similar cases have been opened by the US Author’s Guild (representing literary authors) and were among the main driving factors for the recent Writer’s Guild of America Strikes (representing screenplay authors) in the US film industry. All of the three parties mentioned are arguing for a breach of copyright that has occurred in AI training and are demanding means for compensation (e.g., a usage fee) if their content is used to advance AI models. These cases pinpoint at two current issues: Firstly, to date there exists no comprehensive record to trace the origin of the numerous data sources that go into the training of advanced AI models. And second, there is a lack of commonly agreed legal frameworks and infrastructure that allows legitimate access and equitable compensation for the use of personal or corporate data sets. To mitigate these issues, a blockchain based data marketplace could prove data authorship, enable fair and automated monetization, and allow the tracking of AI training data, establishing transparent infrastructure for AI advancement in the process. Similar to how stock-images are being purchased on the Internet today, Data Exchanges could, e.g., enable the acquisition of licensed-access to private and corporate data sets, fairly compensating the dataset originator in exchange.

At the same time, blockchain-based data marketplaces would empower data producers to register their authorship via unique identifiers, thus establishing irrefutable proof of *data ownership*. In December 2023 the US-based media service Fox News, already deployed its verification tool ‘Verify‘ [62], [63], which leverages cryptographic hashing and digital signatures to authenticate the originality of its content. Its main goal is to enable any end-user to ascertain the authenticity of content as being produced by Fox News. But models like the Fox News example could easily be adopted to ensure the authenticity of a diversity of data and content; Simultaneously, as of today no compensation mechanism for AI training is in place, this already hashed database could easily be integrated into a given data exchange for training purposes.

Accountability and Security

As shown before, blockchains connect data via blocks and secure these cryptographically [44], [64]. As a consequence, modifications by a malicious actor can easily be recognized [60]. In this sense, the system provides advanced security to external attacks and ensures that the models trained maintain their validity. Especially in times of growing numbers of hacking attempts worldwide [51], the enhancement of data pipelines through blockchain verification adds additional security and safety to running machine learning services. For example, the merging of blockchain and AI could ensure advanced security and provide safety from hacker attacks in networks that control and govern autonomous vehicles [65]. In this setting, the combination of both technologies can help to prevent “undesirable data modification in vehicular networks” [65] and possibly increase overall driver safety. On top, data that is protected by the cryptographic features of blockchains, can be tailored to comply with EU privacy law to be safely shared in a federated setting [66], e.g., to improve the AI algorithm for threat detection, and enhance overall vehicle safety performance.

Data Purity and Reputation

As previously discussed, machine learning models thrive on high quality and diversity of the data that is fed into their training. As blockchain allows to trace the originator of a given data set, this data can be used to rate authors and their data on a given data exchange [67]. Subsequently, enabling the establishment of a traceable and verifiable reputation system in a distributed setting. For example, going back to the previously mentioned hobbyist network to provide weather data, the origin and location of a measurement device (say, a webcam) could be authenticated by the original manufacturer and the current owner, adding to ensure the authenticity of data generated. Hereby, based on the hardware used, as well as the user's reputation score, the data providers' local contributions to a global model could be effectively evaluated. Through this configuration, entities without prior direct interactions could collaborate to effectively trade data, allowing advanced scenarios for data to emerge [68], [69].

Automated Payment Channels

As an additional benefit, data and payment transactions between market participants could happen fully automated. While training sources could include public infrastructure, IoT devices, industrial machine sensors, or even personal smartphones. Many of these devices engage in transactions that are characterized by their small size and high frequency. Such an environment allows for the deployment of automated payment channels capable of handling instant transactions, a requirement that traditional third-party verification and processing methods, such as inter-bank transfer and SWIFT, fail to meet due to their time-intensive and costly nature for microtransactions [55]. To address this, a blockchain-based data exchange would be equipped with instantaneous and automated payment mechanisms that facilitate direct financial exchanges between buyers and sellers. To add reliability to exchange dynamics, digital currency, for example a digital Euro, could be used to enable data proprietors to monetize their assets with reliable and verifiable compensation, ensuring a streamlined and secure transactional experience within the data exchange.

2.2 — Data Exchanges as Distributed Marketplaces

As the above sections have illustrated, the fusion of blockchain and AI can be utilized to establish an effective marketplace setting for datasets across a distributed set of participants. Data marketplaces could operate as intermediaries, effectively bridging the gap between data producers and users [58]. Also, market participants could utilize a wide variety of automated devices, ranging from public infrastructure devices and manufacturing machine sensors to personal smartphones, therein enabling a source of passive income for data-affine individuals as well as data-intensive sectors, such as manufacturing, production, or IoT operators. At the same time, Internet users could use this infrastructure to truly own their digital data and gain fair compensation for sharing their profile with third parties that want to process this data, e.g., companies that aim to gain more sophisticated costumer insights. On top, data exchanges could also have a supporting function to establish more safe forms of public information retrieval: content creators, such as newspaper outlets or media houses, could be enabled to declare ownership of their publications therein ensuring the factual correctness of content provided online via their reputation. Additionally, users could use a given reputation system to affirm the authenticity of the story described. Especially, in times of growing misinformation online, data verification outlets for the media could possibly provide an effective countermeasure and add another layer of distinction and trust to public media outlets [63], [70].

Lastly, the implementation of distributed data exchanges could provide a foundation towards a data-driven, competitive market environment where companies are enabled to innovate while upholding data sovereignty. Such an infrastructure would enable the development of new device or web applications, as well as data processing and Machine Learning services on basis of advanced output precision and tailored to specific market niches. Thus, the establishment of a blockchain-based data marketplace could provide significant economic benefits for the local economy, while enhancing the competitive environment and fostering a diverse data ecology, ultimately contributing to a vibrant European data ecosystem. Figure 6 displays a highly simplified version of how Data Exchanges could work. Each data transaction phase operationalized is illustrated in the subsequent summary. The subsequent section will then provide a Deep Dive into how distributed machine learning infrastructures could operate. Therein showing additional use-cases with a high potential for economic gains.

(A) **Offer Initialization:** In this foundational phase, participants, either as data providers or acquirers, generate offers that are recorded on the blockchain, ensuring transparency and immutability. These offers undergo rigorous validation processes that ascertain their authenticity and reliability. To facilitate automated settlement, data requests are coupled with corresponding payments in digital currency and held in escrow to guarantee transaction completion upon successful data delivery.

(B) **Matchmaking**: Offers are matched when participants reach a consensus, which may occur directly or through counteroffers that refine the terms of engagement. Upon agreement, all requisite information is securely acquired; for instance, in the context of federated learning, this would involve updating the global model via its decentralized model iterations.

(C) **Data Processing (incl. Training):** Depending on the type of data agreement, the data owner either conducts the local training of the AI algorithm or initiates the transaction of its offered sensitive data. The transaction culminates with the direct exchange of the agreed-upon data from data owner to data buyer.

(D) **Contract and Payment Settlement:** Post-delivery, the received data undergoes a verification process to confirm its integrity and adherence to the contract terms. Should discrepancies arise, a dispute resolution mechanism is activated. Following a satisfactory resolution, the payment is released from escrow and automatically processed. Additionally, transaction details and participant feedback are incorporated into a comprehensive reputation system, enhancing future trust and accountability.

Data Marketplace

Simplified depiction of a data marketplace

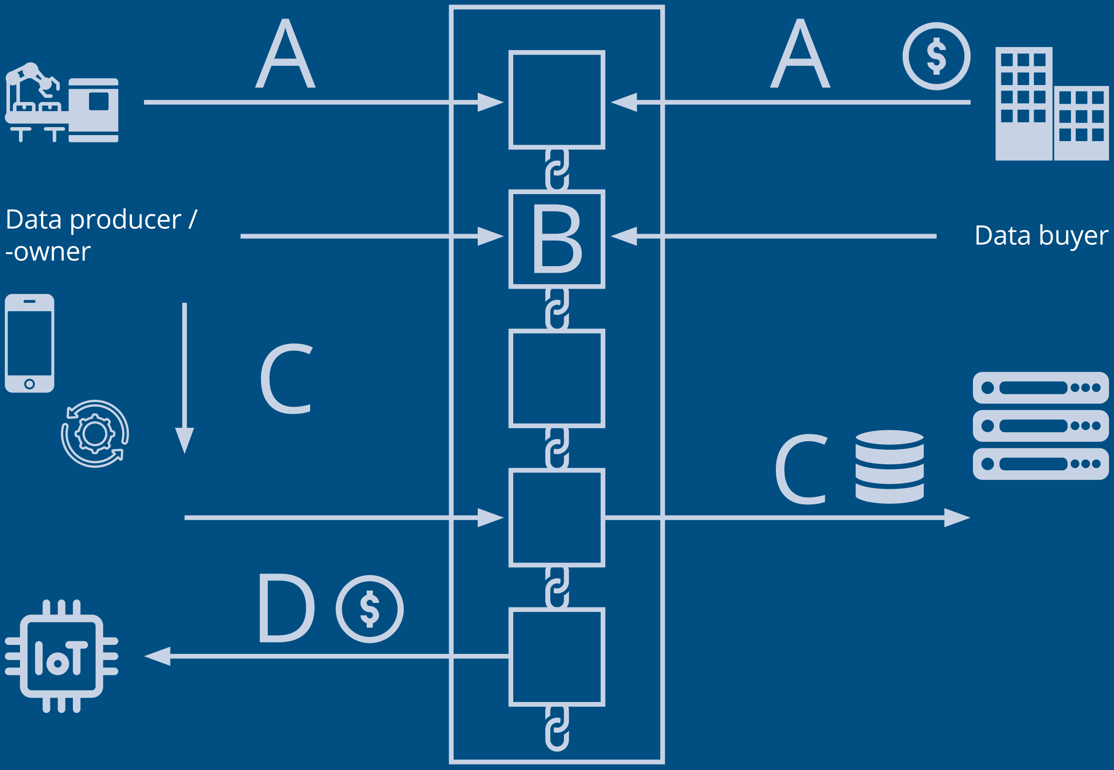


Figure 6

Source: Own illustration, adapted from [58]

3

Application Scenarios of Blockchain and AI

3 — Application Scenarios of Blockchain and AI

3.1 — Blockchain-enhanced federated learning – a technological deep dive

As previously discussed, traditional Machine Learning paradigms assume that data is stored in a centralized setting, such as a local server, and then fed into the model for training [59], [69], [71]. However, this approach has been limited by two, rather obvious, factors in the past. Firstly, as in the case of advanced AI models, with growing model size, the data needed to improve the model's next training cycle, can easily exceed the capacity of local storage facilities [59], [70]. And secondly, the idea of centralizing exceedingly large amounts of data, including the data of individuals and users, conflicts in many cases with general data privacy restrictions [59], [71]. For example, auto-complete models for text can be trained by using type-data from smartphone keyboard-interactions. However, as users' text and messaging data are usually containing highly personal and private information, it should be considered highly unethical to store this data in settings that allow centralized reading access to a data engineer. A better solution would be to train keyboard data locally and to update the results that compute the likeliness for a given word to occur in everyday language in the global model. For these reasons and to mitigate privacy concerns, researchers at Google developed the technique of *Federated Learning* [59].

In simple terms, Federated Learning (FL), facilitates model training on local (so called *edge*-) devices[vi] on which data is collected. Subsequently, it summarizes these local training results, to update an overarching, global model. The global model then contains the training outputs of all local models and, on that basis, finetunes its own predictions [59]. The larger sum of total training data as well as the combination of local model iterations that feeds into the global model, is therein enabling advanced output accuracy and, for larger samples, performs equal or faster in accuracy when compared to centralized training approaches [73], [74], [75].

However, while FL can leverage the benefits of distributed data collection, ownership and evaluation of the global model is still bound to a central entity in the training network [76]. Subsequently, the ownership of the central server also entails ownership of the global model and gives control over managing any training bias or model output weighting [76], [77] which may provide an unfair competitive advantage among other network participants. As a consequence, the incentive to contribute to model training by providing local data is limited in most application scenarios [78], [79]. Especially in a B2B context, the protection of corporate secrets is essential to maintaining a competitive advantage for a given market niche, making it undesirable for many companies to participate in a shared training process with only limited ability to owning their training results. Federated Learning with centralized data ownership is therefore an undesirable scenario for the establishment of Data Exchanges that point beyond the current GAFAM-type digital platform economy.

[vi] Devices such as smartphones, wearable health devices, and machine sensors that are located at the edge of a network between the data source and the cloud [72].

To advance coopetition and equal access to data sources, the paradigm of *Federated Learning on the basis of Distributed-Ownership* has been proposed [60], [68], [69], [76], [80], [81]. In distributed-ownership FL, while local training procedures stay the same, model aggregation and global model calculation is processed via a distributed ledger network and managed collaboratively via blockchains. Thereby, it offers advanced means for participation for all actors in the FL training network, while simultaneously ensuring protection from one single actor gaining control over the entire network and its model outputs [79]. Distributed-ownership FL then creates a (possibly) egalitarian and privacy preserving ML ecosystem that maintains training participants' collaborative sovereignty over model economy and outputs. In this system, participants can agree on rewards that are being issued for high-quality model inputs, therein increasing the chance to maintain accountability and produce positive network-effects among actors involved.

Under the conditions described above, and in consideration of ethical and regulatory standards [11], [12], a blockchain based FL framework could seamlessly integrate into visions of a European sovereign data marketplace. In this setting, while blockchain-based networks enhance the training of collaboratively managed global AI model, also training contributions by non-blockchain nodes could be added. In this way, the data marketplace is utilized to grant access to the blockchain based global model after the contract initialization between the distributed network and the data owner contributing to model training. Subsequently, a decentralized AI network that was created to develop a specific AI can be trained from data that is not owned by the network, while its owners can still receive compensation for their contribution. Especially, network participants that have limited computing resources can thereby participate and benefit from training contributions. Consequently, a data marketplace would increase access to data for decentralized AI projects therein enhancing the feasibility of more complex AI development projects.

Hypothetical architecture of a Blockchain based FL-network

To visualize the opportunities and limitations of a Blockchain based FL framework, Figure 7 illustrates a hypothetical decentralized training architecture based on [71]. A FL paradigm can utilize data from a variety of sources, but for simplification purposes we assume that every participating node possesses similar or equal computation power and connection bandwidth. Participating nodes can inherit either or both, the blockchain consensus algorithm and the AI training algorithm.

The committee fulfils two important functions for the network: On the one hand, as only committee members participate in the calculation of the global model, speed and scalability of the training network can be increased. At the same time, rotating committee selection ensures that all participants in the network can be held accountable for their contributions. Simultaneously, participants are encouraged to maintain high levels of data quality, as for example contributions that are flagged as 'valuable' by the committee could be paired with incentives, such as bonus payments, higher contribution rankings and increased trust for network participants. As a reference to each local model iteration is stored on the blockchain, all participants are enabled to audit the committee's actions and evaluation processes, which adds another layer of transparency to training dynamics.

However, it is noteworthy that also committee election processes can be biased and lead to increased concentration of evaluation power and computing resources, (possibly causing a re-centralization of network power), over time. For example, if only computing performance and quality of data contribution would account as determining factors for committee selection, less powerful network participants could be discriminated, while the committee would consist of an increasingly homogenous group of resourceful and compute-intensive, high-quality data providers. To avoid this effect of re-centralization and to ensure that increasingly homogenous data samples do not exacerbate output bias across model instances, mechanisms of diversification, monitoring and evaluation should be considered and effectively integrated into consensus computation processes. Other consensus evaluation mechanisms, such as *Delegated Proof of Stake*, could also be considered as feasible alternatives [80].

In the blockchain, the first block – the so-called genesis block – stores the fundamental AI model which marks the starting point of training.

(A) Each training participant downloads the corresponding model and conducts local training utilising their personal datasets. The improvement of the local AI model is measured via output functions, being mathematical representations of model performance for a given iteration.

(B) This update is then transacted to a so called ‘committee’ that consists of a sample of network participants, which pause training for a given time (e.g., one training iteration) in order to help compute model consensus in the global instance.

(C) As part of this process, the committee conducts a ranking on the quality of contributions, labels insufficient, or possibly, malicious data contributions, and calculates the global model. Subsequently, the next training round is launched, and new committee participants are chosen from the network of training participants.

Committee Consensus Architecture

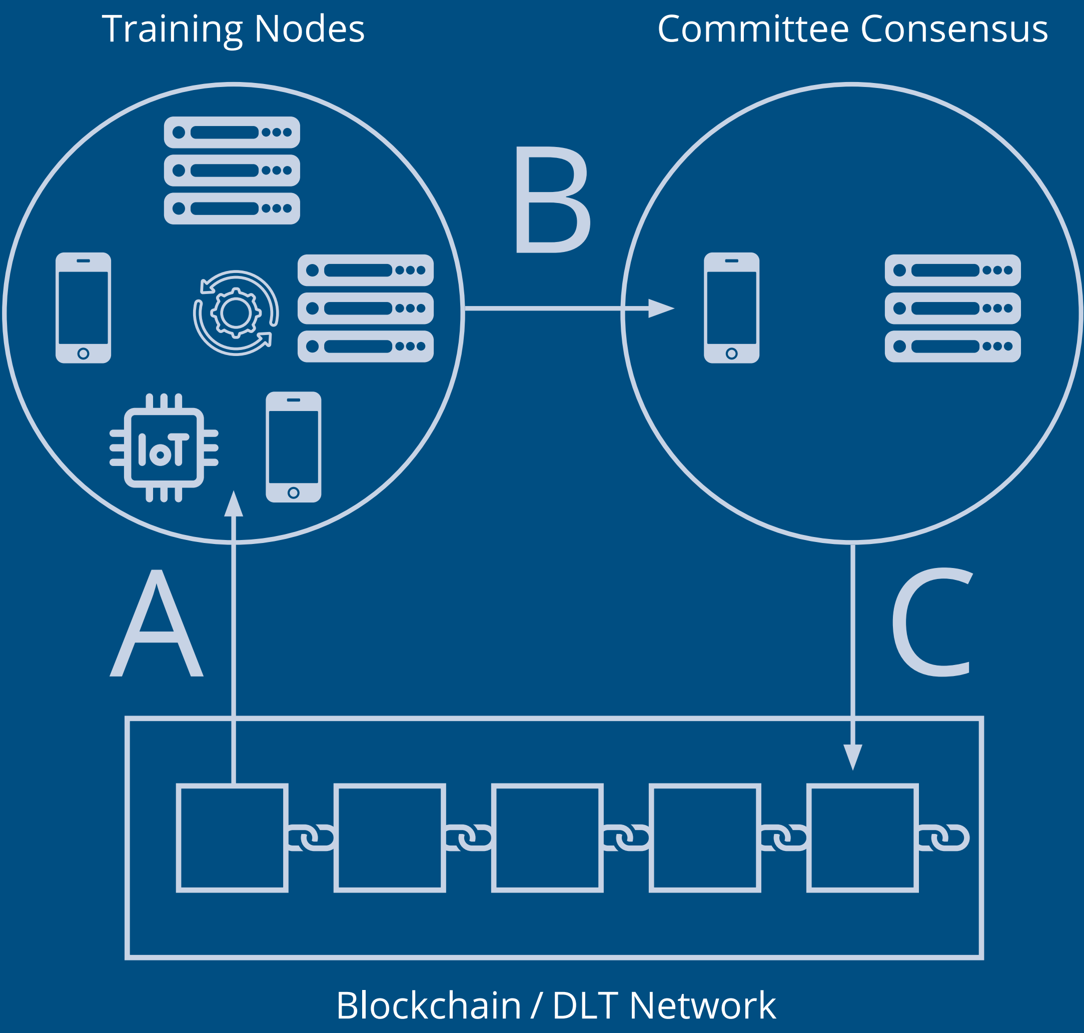


Figure 7

Source: Own illustration, adapted from [71]

Personalized Federated Learning (PFL)

Through the application of blockchain, the establishment of large-scale federated learning networks with diverse datasets becomes more feasible. However, the ensuing data heterogeneity in a privacy-preserving machine learning setting also depicts a challenge to the performance of the trained AI model [82], [83]. PFL enhances traditional federated learning by customizing models to individual users' unique data samples and use-cases. This approach ensures better model performance and user-specific results based on custom data characteristics.

Essentially, while a central model is developed across various devices (like in traditional federated learning), PFL adds a layer of customization by permitting each device to fine-tune this model based on its own unique data [79], [82], [83]. This means each user's device makes local updates to the model that, combined with the global model, add AI solutions that draw on diversified data input for better prediction accuracy, while ensuring the benefits of custom training output, similar to privately trained, local models. As a result, both, broad applicability, and personal relevance of model outputs are ensured. This effect encourages network participants to contribute their resources for decentralized AI development, while allowing for sufficient adaptivity to create AI models that are tailored to specific use cases. These specific models could then effectively be utilized in industrial or business settings that require high customizability. According to research [82], [84] this effect doubles, as through the specification of training data the performance of personalized decentralized AI models, exceeds those provided by generically trained federated solutions for many cases. For example, a manufacturing consortium can collaboratively train a decentralized AI from data produced by their specific machines. Hence, the resulting AI model is much more capable of managing and optimizing the process of these specific machines. Thereby, German (or European) SMEs could achieve a global competitive advantage by establishing highly specialized, collaboratively trained AI systems. To illustrate this case, the following will examine several industry scenarios that already are, or could be, benefitting from decentralized AI development.

Blockchain-based Federated Learning Sample Use Cases

Use-Case 1 – Health Care Research AI:
Health Care data is subject to strong privacy regulations as a result of which centralized data processing and AI training is subject of ethical and legal limitations. To advance the benefits of Big Data in health care, extensive research has been conducted in privacy-preserving federated learning approaches for health-care-AI development [85], [86], [87]. Despite being a promising technology, AI-based scan analysis is hard to develop in the naturally occurring data silos of the health care system [88]. Therefore, federated learning enables the local facilitation of private data to train an AI algorithm, for example in the detection of brain tumors [86]. A blockchain-based FL network could provide a promising avenue to enable collaboration between different health care institutions such as public hospitals, research institutes, and universities, to train an AI detection system that supports and accelerates in-field work or finds improvements to existing treatment procedures. However, especially in health care scenarios these implementations should be approached with caution. As we will discuss in the following chapter, despite the existence of privacy preserving methods such as differential privacy, a fraudulent attack on the global model might still reveal information about training inputs [89]. Especially, when it comes to highly sensitive personal data samples, these concerns must be addressed and solved, before FL architecture on health-care data can be implemented effectively and under consideration of ethical standards.

Use-Case 2 – Industrial-AI-Application:
There are mainly two options for applying FL to industry use-cases:

(A) Large industrial manufacturers could integrate FL approaches into one of their tools to train an AI algorithm utilising the performance data from its customers without privacy infringement. A pilot of this implementation was conducted in Germany, by the Fraunhofer IPA and Lorch AG, a manufacturer of welding machines. As a promising effect of this collaboration, the federated training from multiple in-use welding machines resulted in an AI model that is able to proactively turn off a welding machine if an employee is in the process of making a potentially hazardous mistake [90]. Such applications of FL could be especially relevant for Germany's Small and Medium-Size Businesses landscape and help to advance global competitiveness for AI-enhanced implementation scenarios. Companies could facilitate local AI training to develop shared global models that draw on high-output accuracy, while being able to specialize on niche applications, tailored to individual business cases. Enhanced by a FL-scenario, Germany's current competitive edge of maintaining a larger cluster of world-leading industry and manufacturing businesses, could hold the opportunity to establish a federated network of AI-leaders in Industry and Manufacturing scenarios that will ensure its future competitive edge.

(B) In a second scenario, Industrial manufacturers could implement FL as part of their transnational machine infrastructure, creating a collaborative training setup for their machines in Germany and abroad, and under consideration of local data privacy laws. In a project with Siemens, the Start-Up Katulu created a centralized FL infrastructure for the improvement of Siemens' automated optical inspection systems in their factories in Erlangen [91]. According to Katulu, the successfully implemented FL system provides the foundation for a broader rollout to additional Siemens' factories, including those in China.

The following will highlight two instances in Germany and Europe that work on establishing blockchain-based FL application scenarios in an IoT setting. The German start-up deltaDAO is building an open infrastructure for a blockchain-based FL marketplace, on basis of which data, as well as algorithms, and entire ML models can be traded for industry application scenarios. Partnering with stakeholders such as Airbus and the Dutch Blockchain Coalition, deltaDAO is part of the European project Gaia-X. The start-up is active across several domains (among others, aviation, industry 4.0, mobility, manufacturing) and claims to successfully operate in accordance with EU law and regulations [92]. A second, related project is the fetch.ai foundation initiated by the German Bosch AG and the UK-based Startup fetch.ai and operating in partnership with Telekom MMS [93]. The resulting network aims “to foster innovation and cooperation between industry participants through collective research and development, collaborative applications, shared initiatives, and the discovery of valuable business models“ [94] in AI-based settings. Both projects represent early examples of how federated Data Marketplaces could be designed and should be monitored for future development.

Use-Case 3 –Smart City and IoT-AI:
Increasingly, machines used in city infrastructure produce massive amounts of data that in sum constitute the *Internet of Things*. These communication networks are inevitable for the deployment of smart devices. For example, the facilitation of autonomous driving vehicles in the city is based on access to large datasets of local environments and subsequently, their fast and intelligent analysis. Through the utilization of AI, important calculations can be conducted, for example to predict traffic flow [95].

In the pilot project "Heat" conducted in 2021 in the city of Hamburg, autonomous minibuses were successfully implemented facilitating an extensive AI and IoT network that derives real-time data from its surroundings [96]. Its follow-up project "ALIKE", funded by the German Federal Ministry for Digital and Transport, is set to provide publicly accessibly autonomous minibuses by 2024 in Hamburg [97]. These case-studies highlight that properly trained AI algorithms and data access are an important asset to advance autonomous driving. However, data islands and closed AI training systems could make development more difficult and complicate the establishment of transparency among stakeholders, e.g., in case of accidents or related events that require legal counsel. The implementation of a FL network based on distributed platform ownership could accelerate development of transparent and effective autonomous driving algorithms, by enabling privacy preserving access to IoT data that can be used to train AI algorithms whilst enabling the cooperation of a diverse set of entities. Thereby, blockchain supports collaborative smart city development in an egalitarian environment for every stakeholder and holds the potential to advance free market incentivization.

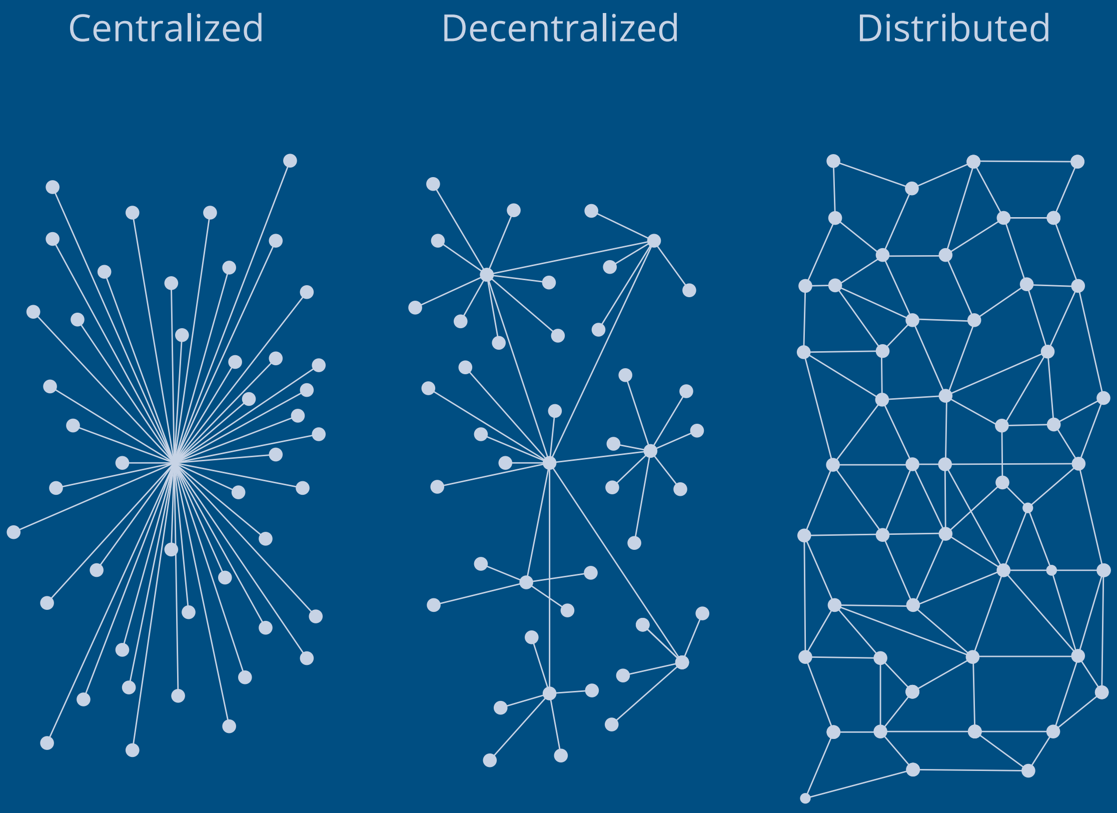
Use-Case 4 – Mobile-AI Applications:
FL can be applied on mobile devices to access training data that in itself would usually be too small in sample size to be considered for effective training [98]. To enable advanced training sample access, the authors of [98] designed a FL system that could enable a multitude of mobile devices to join the network and participate in training rounds to enable mobile-enhanced FL solutions. This can be useful, e.g., in e-commerce recommendation algorithms, that often utilize cloud-stored, but sensitive user-data which may infringe data privacy law. Considering this scenario, researchers successfully tested the application of a privacy-preserving FL system for Alibaba and Taobao, utilizing on-device data [99]. Such innovations could be very relevant for German e-commerce giants, like Zalando and AboutYou. Also, these approaches tend to become more feasible considering recent advancements such as PockEngine which significantly reduces required computational power for efficient AI fine-tuning [100].

Centralized Machine Learning: In centralized machine learning, a central server is the exclusive hub for aggregating data from various sources and processing it to train the machine learning algorithm.

Decentralized Federated Learning: Decentralized federated learning involves local model training at the data source, with individual updates then sent to a central server for global model aggregation.

Distributed Federated Learning: Distributed federated learning facilitates local model training at each data source, with the global model's aggregation and update being collaboratively managed by all nodes using distributed ledger technology.

Centralized vs. Decentralized vs. Distributed Network Architecture



Centralized

Decentralized

Distributed

Figure 8

Source: Adapted from [45]

Challenges and Open Research Questions for Federated Learning Applications

Challenge – 1: Privacy Leakage
While Federated Learning significantly improves privacy preservation in AI training, it is not a guarantee for privacy [89]. Although federated learning methods prevent the transfer of sensitive data, research has shown that some scenarios exist where adversarial attacks may reveal training data derived from model updates [89], [101]. To mitigate these effects and prevent data leakage, mechanisms such as differential privacy (DP) have been integrated. DP adds noise to the sensitive data which makes the retrieval of information unfeasible. However, DP is a trade-off between privacy and model accuracy because increased noise lowers the accuracy but improves the privacy [102], incurring a cost in usability and model precision [89]. Further research has to be conducted to combine DP with other techniques in order to maintain model accuracy while assuring privacy. Nevertheless, deliberate attempts to infringe user privacy can still impose a reasonable threat [89]. Advanced cryptography such as secure multi-party computation, fully homomorphic encryption and zero knowledge proofs are promising solutions but still require development to ensure practical application and reduce computational overhead [89], [103].

Challenge – 2: Poisoning Attacks
Inherent to Federated Learning is the challenge of identifying model updates that are malicious due to prior data poisoning or direct update poisoning [104]. Adversarial attacks can for example replace the labels of training data whereby the update parameters may lead to the convergence of sub-optimal global models or leave a backdoor for the attacker [89]. Research has proposed several solutions, such as byzantine-tolerant federated learning systems or median-based aggregators [89]. At the same time, researchers have highlighted that advanced impact of poisoning attacks on the global model can be mitigated through the facilitation of low cost defenses [104]. Subsequently, and in combination with a transparent and collaboratively managed global model, the risks of poisoning attacks could possibly be minimized [105].

Challenge – 3: Heterogeneous devices with varying network connections and computation power
The heterogeneity of large, distributed networks regarding network connection and computation power challenges the fair and open participation in federated learning training rounds. Powerful servers are underutilized because a linear training round can only be as fast as the slowest device [106], [107]. Varying degrees in network bandwidth and connection stability could also restrict the ability to participate in FL [107]. Therefore, research proposes asynchronous federated learning systems that enable the independent upload of training updates at random times, hence, improving network efficiency [89], [106], [107]. However, asynchronous FL also imposes the risk of overrepresenting model updates from computationally more powerful sources [107]. Thereby, asynchronous FL improves possibility of training access, and could advance the scalability of training networks, while simultaneously decreasing fair participation opportunities for less powerful devices; with the latter effect being an undesirable outcome. To mitigate these dynamics and reduce possibility of discriminating low-compute network participants, further research is needed to explore the capabilities of asynchronous federated learning frameworks.

3.2 — Metaverse and Trusted AI – Possible Future Scenarios for Blockchain & AI Implementations

To conclude this chapter, we briefly aim to highlight two additional application scenarios for which the fusion of blockchain and AI could prove beneficial. However, research in these areas is still very preliminary, making predictions difficult, as still several technical hurdles need to be overcome before we can speak of truly functional solutions. Therefore, we would need to emphasize that, to date, the following brief paragraphs can only be an object of (informed) speculation.

Metaverse

The first object of investigation to name in this context, would be the term Metaverse. At the latest, since platform giant Facebook renamed itself to Meta, this term has become a buzzword in itself. Representing a multifaceted layer of different meanings, their greatest commonality is probably that applications in the metaverse involve an *Augmented Reality* (AR) or *Virtual Reality* (VR) device that generates some form of projected (AR) or enclosed (VR) spatial computing environment. Throughout this report, we discussed the imbalances of power that might arise for users and smaller businesses if centralized forms of AI training were the only point of access. With the metaverse, similar issues could arise, bearing the difference that corporate control (and possibly ownership) over augmented reality or virtual reality projections, would directly affect a users' immediate visual field. In these terms, we could therefore conceptualize the Metaverse as a three-dimensional data environment that would possibly enhance an individual's daily experience and life, by producing projections that appear as productive to this user. But could also increasingly operate to create a flow of attention-grabbing nudges that abstracts value from users' immediate visual experiences. Subsequently, transmitting this value, in form of data, to the centralized servers of large, digital platform corporations. Framing the dynamics of the Metaverse in this context, it might become apparent that the difficulties that are growing more prominent concerning the current questions of AI training and ownership in a centralized data economy, might also apply to this yet-to-emerge spatial computing environment that is commonly referred to under the umbrella term ‘Metaverse’. As this report has shown, Blockchain can generally be understood as a technology that, if utilized properly, can offer more egalitarian and fair access to the digital infrastructures that constitute our lives today. While, at the same time, through its distributed and consensus-driven data processing infrastructure, holding the additional benefit of making our digital experiences more ‘safe’ - or at least more ‘tamperproof’. Providing this additional layer of data verification via the blockchain to Spatial Computing Environments could enhance the stability and safety of experiences in the Metaverse, making it more difficult for hackers to compromise the visual experiences of Metaverse users to their advantage [108].

Trustworthy AI

The recently issued EU AI Act, requires “data governance, record-keeping, transparency and access control” for AI [14]. To comply with this new regulation, blockchain could help to produce a trail of auditability, transparency, and traceability for AI models [14]. Throughout this report, we had discussed several possible scenarios for these features, but have not touched on the topic of auditability. As stated earlier, it is difficult, if not impossible, to fully comprehend the internal dynamics of Deep-Learning-enhanced AI-models. Hence to date, the exact decision processes of advanced AI models cannot be fully traced. With growing model size, and therefore growing model complexity, we expect this trend to continue, making it increasingly difficult for human observers to gain full insight into how a given AI model has come to make a specific decision or prediction. However, similar to the ways in which Github allows the tracking of code changes via so called ‘commits’, blockchain could allow the establishment of audit trails to track changes, enhancements, or data sources that have been fed into a given model – storing these records as immutable for as long as the blockchain exist [109]. While this approach might not provide comprehensive insight into the AI system, it might provide an initial orientation for auditors and regulators to gain insight into a given model’s black box dynamics of decision making. Future research will need to explore in more detail how such an attempt can be safely established, as well as which other means could be productively applied to create safer and more auditable AI in the future.

Metaverse

Possible Layers of Attack to disrupt the metaverse user experience vs. ways to use Blockchain & AI as mitigating technologies for advanced security.

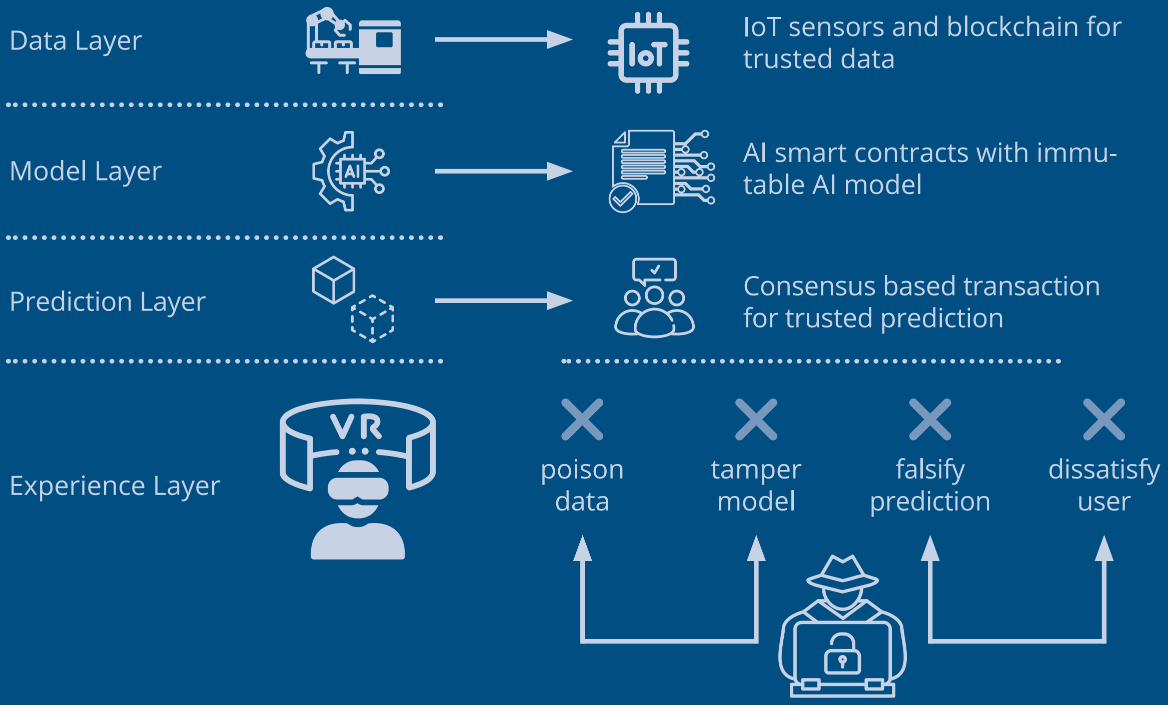

Data Layer
IoT sensors and blockchain for trusted data
Model Layer
AI
AI smart contracts with immutable AI model
Prediction Layer
Consensus based transaction for trusted prediction
Experience Layer
VR
poison data
tamper model
falsify prediction
dissatisfy user

Figure 9

Source: Adapted from [108].

4

Conclusion

4 — Conclusion

In this report we have analyzed how the convergence of Blockchain and AI can be productively implemented to strengthen the data sector of Germany and the EU. During our analysis we have found that each technology's strength is bound to the fact that it advances how data can be processed. We found that AI proves an excellent match to productively master the yottabytes of data that compose our societies at the beginning of the 21st century, whereas a promising potential of blockchain is that it can offer the opportunity to provide more egalitarian access to this data, as well as distribute the wealth generated by use of this data more evenly and according to the principle of each individual's contributions.

Chapter 1 has therein provided a simplified and brief summary and overview of how Deep Neural Networks, the infrastructure behind the latest advancements in AI development, operate. Thereby we have shown that Neural Networks require large amounts of high-quality data to develop and to provide ever increasing accuracy in their predictions – while avoiding and/or minimizing bias. Additionally, due to their self-adjusting dynamics, we have also shown that the exact workings of a neural network operate inside a black box that human operators can only investigate from its outside. Therein, control over decision making and prediction processes of advanced AI is difficult to achieve and could increase in difficulty with growing model complexity. We then discussed the possibly negative effects of centralizing the power and control over AI in the hands of a selected group of internationally operating digital platform corporations. Lastly, we highlighted how the unique technological features of blockchain could help in creating digital AI infrastructure that mitigates the risks of overly centralized AI by offering means for the decentralization, transparency, immutability, and traceability of data.

4 — Conclusion

Drawing upon the idea that data is becoming an increasingly central asset class of the 21st century, Chapter 2 introduced the idea of Data Exchanges. We thereby described data Exchanges as a possible solution to overcoming overly centralized AI infrastructure, while empowering a multitude of German, European, or even international, businesses to partake in a collaborative data market that fairly rewards contributors while protecting users' data privacy and ensures the effective training of AI models in accordance with European data law. We have highlighted that the core technology that could accompany this paradigm shift from "big data" to "shared data" would be a form of Federated Learning that emphasizes the importance of distributed ownership in training fair, ethical, and bias-minimizing AI models. Subsequently, we discussed how a FL-driven Data Exchange could empower startups, SMEs, and larger corporations in Germany and Europe by producing advanced and collaboratively owned ML-models that can drive business insights by shared access to a global training pool and ensure a competitive edge by allowing the freedom to create use-case-tailored and specialized local model instances.

Chapter 3 then provided a technological deep dive into the workings of Federated Learning and discussed the possible benefits in the application of this technology in four specific industry scenarios. Lastly, we discussed the potential benefits of blockchains and AI technology for the latest trends in tech innovation, specifically the Metaverse and the possible advancement of Trustworthy AI that operates in accordance with EU standards of achieving data governance, auditability, transparency, traceability, and access control for AI models.

As stated at the beginning of this document, our analysis does not come without its limitations. Firstly, due to the complexity of this topic as well as the ongoing pace of innovation and rate of new discoveries, this report has only aimed to provide a conceptual overview of the productive convergence of blockchain and AI technology. While it has been our goal to provide a highly accurate depiction of this convergence, at times we simplified the actual mechanics of this technology in order to ensure greater accessibility of this report. Additionally, the vision outlined is by no means comprehensive. While blockchain and AI offer a promising convergence to the establishment of a vision of data that operates in stronger accordance with European values, still many technological hurdles need to be overcome until this vision can become a reality. For example, blockchain-based data processing tends to be slower and require larger amounts of data, as a copy of every data point tracked is stored in every validator node in the network. Subsequently, engineers of Data Exchanges would need to consider which parts of the data exactly needs to be stored on-chain, and which parts can be processed via the regular data-transmission network. At the same time, FL does not solve all privacy issues of contemporary AI models. As we had discussed, privacy preserving dynamics of FL-learning are existent, but would need further refinement for large-scale applications. Also, AI-systems can still be hacked or manipulated to reveal sensitive information for both, local and global instances. While techniques exist to mitigate this effect, additional research needs to be done to ensure advanced privacy and security for the users of federated machine learning networks, as well as for the recipients of outputs that are built on basis of these networks.

In addition to the technical limitations, we emphasize the importance of ethical AI development including the imperative to develop non-discriminating AI ecosystems. As AI produces its output based on what it is trained on, data management is an important lever to address in overcoming social bias and discrimination. Diversity in input patterns and AI training infrastructure could therefore be one of the essential steps in reducing the output of bias by AI algorithms. We sincerely encourage developers to pursue AI advancement that fosters diversity and the inclusion of diverse voices from the very beginning, and secondly, to integrate mechanisms for equal access and participation among users and the recipients of AI output into governance and data management infrastructures. We hope these measures will contribute towards advancing this important debate also on the layer of digital platform infrastructures.

Despite the constraints mentioned above, we do hope that the reader has enjoyed this brief exploration in the field of blockchain and AI. As well as has gotten a feel of how these technologies could possibly be applied productively to empower German and European economies, by putting data as a resource closer to its center. In this sense, the convergence of blockchain and AI might offer an alternative to the highly centralized models of data processing that we operate by in the current mode of our digital (platform) economies. How exactly this alternative might look like is something that would need to be explored in conversation between researchers, engineers, decision makers, policy makers, European governments, and the public.

5

References and Glossary

5 — References and Glossary

[1] B. Thormundsson, ‘Artificial Intelligence market size 2030’, Statista. Accessed: Dec. 14, 2023. [Online]. Available: 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’. Accessed: Dec. 14, 2023. [Online]. Available: 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. Accessed: Dec. 14, 2023. [Online]. Available: 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 and T. Miller, ‘The next big thing – artificial intelligence’, Soc. Identities, vol. 29, no. 1, pp. 1–4, Jan. 2023, doi: 10.1080/13504630.2023.2236372.

[6] B. Gates, ‘The Age of AI has begun’, gatesnotes.com. Accessed: Dec. 14, 2023. [Online]. Available: 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. Accessed: Dec. 14, 2023. [Online]. Available: https://www.calcalistech.com/ctechnews/article/qc1x9jcho.

[8] P. Olson, ‘Google, Microsoft Will Dominate AI as Computing Costs Surge’, Bloomberg.com, Feb. 19, 2024. Accessed: Feb. 26, 2024. [Online]. Available: 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, and 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’. Accessed: Feb. 27, 2024. [Online]. Available: 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]. Available: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

[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]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[14] S. Ramos and 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, pp. 1–20, Mar. 2024, doi: 10.1365/s43439-023-00107-9.

[15] P. Sandner and 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, and H. Steyerl, Pattern Discrimination. DE: meson press, 2018. Accessed: Feb. 27, 2024. [Online]. Available: 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, and 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. Accessed: Feb. 29, 2024. [Online]. Available: https://artsandculture.google.com/story/meet-the-golem-the-first-artificial-intelligence/BAXhTNxULrWYKg.

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

[23] IBM, ‘What is a Neural Network? | IBM’. Accessed: Apr. 22, 2024. [Online]. Available: https://www.ibm.com/topics/neural-networks.

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

[25] E. Griffith, ‘The A.I. Industry’s Desperate Hunt for GPUs Amid a Chip Shortage’, The New York Times, New York, Aug. 16, 2023. Accessed: Feb. 27, 2024. [Online]. Available: https://www.nytimes.com/2023/08/16/technology/ai-gpu-chips-shortage.html.

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

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

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

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

[30] #OxfordAI, ‘A history of AI’. University of Oxford, 2023. Accessed: Feb. 26, 2024. [Online]. Available: https://oxford.shorthandstories.com/ai-a-history/.

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

[32] IBM, ‘What is Machine Learning? | IBM’. Accessed: Feb. 27, 2024. [Online]. Available: https://www.ibm.com/topics/machine-learning.

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

[34] A. Mordvintsev, C. Olah, and M. Tyka, ‘Inceptionism: Going Deeper into Neural Networks’, Google Research. Accessed: Feb. 27, 2024. [Online]. Available: https://blog.research.google/2015/06/inceptionism-going-deeper-into-neural.html.

[35] Wikipedia, ‘DeepDream’, Wikipedia. Dec. 22, 2023. Accessed: Apr. 22, 2024. [Online]. Available: https://en.wikipedia.org/w/index.php?title=DeepDream&oldid=1191335336.

[36] A. Mordvintsev, M. Tyka, and C. Olah, ‘deepdream/dream.ipynb at master · google/deepdream’, GitHub. Accessed: Feb. 27, 2024. [Online]. Available: https://github.com/google/deepdream/blob/master/dream.ipynb.

[37] A. Gordić, ‘gordicaleksa/pytorch-deepdream’, GitHub. Accessed: Feb. 27, 2024. [Online]. Available: https://github.com/gordicaleksa/pytorch-deepdream.

[38] 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.

[39] A. Karpathy, ‘Software 2.0’, Medium. Accessed: Feb. 27, 2024. [Online]. Available: https://karpathy.medium.com/software-2-0-a64152b37c35.

[40] A. Pentland, J. Werner, and C. Bishop, ‘Blockchain+AI+Human: Whitepaper and Invitation’. The MIT Trust::Data Consortium for blockchain+AI research, 2021. Accessed: Feb. 02, 2023. [Online]. Available: https://connection.mit.edu/sites/default/files/publication-pdfs/blockchain%2BAI%2BHumans.pdf.

[41] V. Kennedy, ‘Exploring the future of AI: The power of decentralization’, Cointelegraph. Accessed: Feb. 27, 2024. [Online]. Available: https://cointelegraph.com/news/the-crucial-role-of-decentralization-in-shaping-ai-s-future.

[42] R. Meyer, ‘Everything We Know About Facebook’s Secret Mood-Manipulation Experiment’, The Atlantic. Accessed: Feb. 27, 2024. [Online]. Available: https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/.

[43] Le Monde, ‘Amazon sued in the US for anti-competitive practices’, Le Monde, France, Sep. 27, 2023. Accessed: Feb. 27, 2024. [Online]. Available: https://www.lemonde.fr/en/international/article/2023/09/27/amazon-sued-in-the-us-for-abusing-its-position_6140177_4.html.

[44] S. Nakamoto, ‘Bitcoin: A Peer-to-Peer Electronic Cash System’. 2008. Accessed: Feb. 02, 2023. [Online]. Available: https://bitcoin.org/bitcoin.pdf.

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

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

[47] H. Taherdoost, ‘Blockchain Technology and Artificial Intelligence Together: A Critical Review on Applications’, Appl. Sci., vol. 12, no. 24, p. 12948, Dec. 2022, doi: 10.3390/app122412948.

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

[49] V. Buterin, ‘A Proof of Stake Design Philosophy’, Medium. Accessed: Jan. 08, 2024. [Online]. Available: https://medium.com/@VitalikButerin/a-proof-of-stake-design-philosophy-506585978d51.

[50] X. Zhu, H. Li, and 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, pp. 370–383. doi: 10.1007/978-3-030-14234-6_20.

[51] S. Heister and 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.

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

[53] A. Acar, H. Aksu, A. S. Uluagac, and 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.

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

[55] A. M. Antonopoulos, The internet of money: a collection of talks. Volume 1. United States of America: Merkle Bloom LLC, 2016.

[56] Y. Cheng, Y. Liu, T. Chen, and Q. Yang, 'Federated learning for privacy-preserving AI', Commun. ACM, vol. 63, no. 12, pp. 33–36, Nov. 2020, doi: 10.1145/3387107.

[57] CWOP, 'Citizen Weather Observer Program', Citizen Weather Observer Program. Accessed: Feb. 27, 2024. [Online]. Available: http://wxqa.com/.

[58] L. D. Nguyen, S. R. Pandey, S. Beatriz, A. Broering, and 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]. Available: http://arxiv.org/abs/2112.02870.

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

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

[61] M. M. Grynbaum and R. Mac, 'New York Times Sues OpenAI and Microsoft Over Use of Copyrighted Work – The New York Times', The New York Times, New York, Dec. 27, 2023. Accessed: Feb. 27, 2024. [Online]. Available: https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

[62] F. Corporation, 'Verify tool', verify.fox. Accessed: Feb. 27, 2024. [Online]. Available: https://www.verify.fox.

[63] Polygon Labs, 'Fox Corporation Taps Polygon PoS to Power Verify, an Open Protocol for Content and Image Verification'. Accessed: Feb. 27, 2024. [Online]. Available: https://polygon.technology/blog/fox-corporation-taps-polygon-pos-to-power-verify-an-open-protocol-for-content-and-image-verification.

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

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

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

[67] A. Dixit, A. Singh, Y. Rahulamathavan, and 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.

[68] Z. Zhou, C. Guo, X. Zhang, R. Wang, L. Zhang, and 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, pp. 1041–1044. doi: 10.1109/ICBC56567.2023.10174981.

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

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

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

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

[73] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson, and M. Jirstrand, ‘A Performance Evaluation of Federated Learning Algorithms’, in Proceedings of the Second Workshop on Distributed Infrastructures for Deep Learning, Rennes France: ACM, Dec. 2018, pp. 1–8. doi: 10.1145/3286490.3286559.

[74] M. Asad, A. Moustafa, and T. Ito, ‘Federated Learning Versus Classical Machine Learning: A Convergence Comparison’, 2021, doi: 10.48550/ARXIV.2107.10976.

[75] I. Khitrov and C. Harth-Kitzerow, ‘Accuracy Tradeoffs of Federated Learning approaches’. Technical University Munich, 2022. Accessed: Nov. 30, 2023. [Online]. Available: https://www.net.in.tum.de/fileadmin/TUM/NET/NET-2022-11-1/NET-2022-11-1_04.pdf.

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

[77] Z. Yang, Y. Shi, Y. Zhou, Z. Wang, and K. Yang, ‘Trustworthy Federated Learning via Blockchain’. arXiv, Aug. 12, 2022. Accessed: Jan. 28, 2024. [Online]. Available: http://arxiv.org/abs/2209.04418.

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

[79] J. Yun, Y. Lu, and X. Liu, ‘BCAFL: A Blockchain-Based Framework for Asynchronous Federated Learning Protection’, Electronics, vol. 12, no. 20, p. 4214, Oct. 2023, doi: 10.3390/electronics12204214.

[80] H. Chen, S. A. Asif, J. Park, C.-C. Shen, and M. Bennis, ‘Robust Blockchained Federated Learning with Model Validation and Proof-of-Stake Inspired Consensus’. arXiv, Jan. 09, 2021. Accessed: Feb. 21, 2024. [Online]. Available: http://arxiv.org/abs/2101.03300.

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

[82] J. Mills, J. Hu, and G. Min, ‘Multi-Task Federated Learning for Personalised Deep Neural Networks in Edge Computing’. arXiv, Jul. 22, 2021. Accessed: Jan. 25, 2024. [Online]. Available: http://arxiv.org/abs/2007.09236.

[83] A. Z. Tan, H. Yu, L. Cui, and Q. Yang, ‘Towards Personalized Federated Learning’, IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 12, pp. 9587–9603, Dec. 2023, doi: 10.1109/TNNLS.2022.3160699.

[84] A. Rakotomamonjy, M. Vono, H. J. M. Ruiz, and L. Ralaivola, ‘Personalised Federated Learning On Heterogeneous Feature Spaces’. arXiv, Jan. 26, 2023. Accessed: Feb. 09, 2024. [Online]. Available: http://arxiv.org/abs/2301.11447.

[85] R. S. Antunes, C. André Da Costa, A. Küderle, I. A. Yari, and 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.

[86] W. Li et al., ‘Privacy-preserving Federated Brain Tumour Segmentation’, no. arXiv:1910.00962. arXiv, Oct. 02, 2019. Accessed: Jan. 28, 2024. [Online]. Available: http://arxiv.org/abs/1910.00962.

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

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

[89] P. Kairouz et al., ‘Advances and Open Problems in Federated Learning’. arXiv, Mar. 08, 2021. Accessed: Feb. 09, 2024. [Online]. Available: http://arxiv.org/abs/1912.04977.

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

[91] M. Kuehne-Schlinkert, K. Schmidt, E. Schwulera, B. Scharinger, T. Blumauer-Hiessl, and 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.

[92] deltaDAO, ‘deltaDAO Homepage’. Accessed: Feb. 29, 2024. [Online]. Available: https://www.delta-dao.com/.

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

[94] Fetch.ai Foundation, ‘Home | Foundation Web’. Accessed: Feb. 27, 2024. [Online]. Available: https://fetchai.foundation/.

[95] Y. Kim, P. Wang, and L. Mihaylova, ‘Structural Recurrent Neural Network for Traffic Speed Prediction’, no. arXiv:1902.06506. arXiv, Feb. 18, 2019. Accessed: Jan. 28, 2024. [Online]. Available: http://arxiv.org/abs/1902.06506.

[96] C. Latz, V. Vasileva, and 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, pp. 479–496. doi: 10.1007/978-3-031-15086-9_31.

[97] Hochbahn, 'Autonome On-Demand-Shuttles - Wie das Projekt ALIKE mit autonomen Kleinbussen den ÖPNV ergänzen soll'. Accessed: Feb. 27, 2024. [Online]. Available: https://www.hochbahn.de/de/projekte/autonome-on-demand-shuttles.

[98] K. Bonawitz et al., 'Towards Federated Learning at Scale: System Design', no. arXiv:1902.01046. arXiv, Mar. 22, 2019. Accessed: Jan. 28, 2024. [Online]. Available: http://arxiv.org/abs/1902.01046.

[99] 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, pp. 1–14. doi: 10.1145/3372224.3419188.

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

[101] L. Lyu, H. Yu, and Q. Yang, 'Threats to Federated Learning: A Survey'. arXiv, Mar. 04, 2020. Accessed: Feb. 09, 2024. [Online]. Available: http://arxiv.org/abs/2003.02133.

[102] A. E. Ouadrhiri and A. Abdelhadi, 'Differential Privacy for Deep and Federated Learning: A Survey', IEEE Access, vol. 10, pp. 22359–22380, 2022, doi: 10.1109/ACCESS.2022.3151670.

[103] 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, pp. 619–640, Feb. 2021, doi: 10.1016/j.future.2020.10.007.

[104] V. Shejwalkar, A. Houmansadr, P. Kairouz, and D. Ramage, 'Back to the Drawing Board: A Critical Evaluation of Poisoning Attacks on Production Federated Learning'. arXiv, Dec. 13, 2021. Accessed: Feb. 24, 2024. [Online]. Available: http://arxiv.org/abs/2108.10241.

[105] C. Fung, C. J. M. Yoon, and I. Beschastnikh, 'Mitigating Sybils in Federated Learning Poisoning'. arXiv, Jul. 2020. Accessed: Feb. 29, 2024. [Online]. Available: http://arxiv.org/abs/1808.04866.

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

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

[108] S. Badruddoja, R. Dantu, Y. He, M. Thompson, A. Salau, and K. Upadhyay, ‘Trusted AI with Blockchain to Empower Metaverse’, in 2022 Fourth International Conference on Blockchain Computing and Applications (BCCA), San Antonio, TX, USA: IEEE, Sep. 2022, pp. 237–244. doi: 10.1109/BCCA55292.2022.9922027.

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

5 — Glossary

Artificial Intelligence (AI) is the capability of software to perform tasks that traditionally require human intelligence.

Deep Learning is a subfield of machine learning that employs deep neural networks, i.e., layers of interconnected (digital) "neurons" whose connections have parameters or weights that can be trained. It is particularly effective at learning from unstructured data such as images, text, and audio.

Generative AI refers to AI that is typically built on so-called Foundation Models and possesses capabilities that previous AI did not have, e.g., the ability to generate content.

Foundation Models are a kind of general models trained on vast amounts of data and can also be used for non-generative purposes (e.g., classifying user sentiment as negative or positive based on conversation logs). Their pre-trained capabilities offer significant improvements over previous models.

Large language models (LLMs) are a class of Foundation Models that can process vast amounts of unstructured text and learn the relationships between words or parts of words, called tokens. In this way, LLMs can generate natural language texts and perform tasks such as summaries or knowledge extraction. GPT-4 (which underlies ChatGPT) and LaMDA (the model behind Bard) are examples of LLMs.

Machine Learning (ML) is a subfield of artificial intelligence where a model gains abilities after being trained or shown many example data points. Machine learning algorithms recognize patterns and learn to make predictions and recommendations by processing data and experiences rather than receiving explicit programming instructions. The algorithms also adapt and can become more effective in response to new data and experiences.

Prompt Engineering refers to the process of designing, refining, and optimizing input prompts to steer a generative AI model to produce the desired (i.e., accurate) results. Centralized AI systems denote models where decision-making and computing processes are concentrated on a single, central entity. In this system, all data processing and algorithmic processes occur in one place, often using a central server or database. This concentration can lead to efficient data handling but also poses risks such as single points of failure and can have scaling problems with large volumes of data.

Blockchain is a distributed database technology consisting of a chain of blocks that record transaction data. Each block contains a cryptographically secured hash of the previous block, transaction data, and a timestamp. This structure makes blockchain inherently secure and resistant to manipulation, as any change in a block would invalidate the entire chain.

Centralized AI systems refer to models in which decision-making and computing processes are concentrated on a single, central entity. In this system, all data processing and algorithmic processes occur in one place, often using a central server or database. This concentration can lead to efficient data handling but also carries risks such as single points of failure and may have scaling issues with large volumes of data.

Distributed AI systems denote models that distribute data processing and decision-making across multiple interconnected nodes. In these systems, there is no central authority, leading to increased fault tolerance and scalability. While some of these systems may use technologies like blockchain, they generally rely on a network of nodes that collaborate to tackle complex tasks and make decisions.

Federated Learning (FL) is a machine learning concept where models are trained on distributed devices without sensitive data leaving these devices. After the AI models have been trained locally, all updates are aggregated and processed into an improved global model. A central or distributed AI system can be used for the aggregation and calculation of the global model.

Personalized Federated Learning (PFL) is an extension of Federated Learning, where the trained model is based not only on common data but also on individual data of the participants to deliver personalized results. This allows for finer adjustment to the specific needs or preferences of individual users.

Edge devices are devices at the edge of a network that often perform data processing tasks locally, rather than sending them to central servers. These devices, such as smartphones, IoT devices, or local servers, perform data analysis and processing close to the data source, which reduces latency and increases efficiency. Edge computing plays a significant role in distributed AI systems and the Internet of Things (IoT).

About the Authors

Tom Haverland

Tom Haverland is a research fellow and project lead at the Hanseatic Blockchain Institute, leading the W3NOW project which stands out as the first extensive research project funded by the German Federal Ministry of Economy and Climate Protection to examine the adoption of blockchain technology in the German economy. Haverland's further research delves into the synergistic potential between blockchain technology and artificial intelligence. As a graduate of Leuphana University with a focus on digital politics, Haverland has researched the transformation of traditional paper-based voting systems into the digital realm, leveraging blockchain technology and cryptographic proof systems to enhance security and integrity in digital voting processes.

Lukas Beckenbauer

Trained in Media Studies, Machine Learning, and Philosophy, Lukas Beckenbauer is a Doctoral Candidate at the Technical University in Munich, where he is looking at the role blockchain can play for decentralized modes of ownership and online governance. In the past, he has worked as a researcher with the University of California Irvine, USA, the Arizona State University, USA, and the Leuphana University, Lüneburg, Germany. He has also been Junior Research Fellow at ArtEZ University in the Netherlands, working to understand what exactly drives social change narratives. Lukas holds a Research Master's in Media Studies from the University of Amsterdam.

Imprint

Published by: Konrad-Adenauer-Stiftung e. V., 2024, Berlin

Contact:
Dr. Christian Hübner
Künstliche Intelligenz
Analysis and Consulting
christian.huebner@kas.de
Tel. +49 30 26996-3264

Design: Kaluza+Schmid Studio GmbH, Berlin
Typesetting: Janine Höhle, Konrad-Adenauer-Stiftung e. V.

This publication was published with financial support of the Federal Republic of Germany.

This publication of the der Konrad-Adenauer-Stiftung e. V. is solely intended for information purposes. It may not be used by political parties or by election campaigners or supporters for the purpose of election advertising. This applies to federal, state and local elections as well as elections to the European Parliament.

Equal authorship contribution. Please cite this report as follows:
Beckenbauer, L., & Haverland, T. (2024). Blockchain & AI Report. KAS4 Innovation Series. Konrad-Adenauer-Stiftung e.V. Berlin, Deutschland.

ISBN 978-3-98574-244-8

This report discusses a vision of the European AI ecosystem that offers an effective alternative to existing centralized data paradigms. At its core stands the integration of blockchain-based data marketplaces and federated learning to enable privacy-compliant approaches to AI training. Special attention is given to supporting startups and SMEs in fully harnessing the potential of AI and blockchain, and adapting training methods to the specific data structures of these companies.

The report emphasizes the importance of strengthening Europe's technological sovereignty in the age of Artificial Intelligence and presents practical solutions for realizing "AI Made in Europe."

Konrad-Adenauer-Stiftung e. V.