

Decentralised artificial intelligence

Tom Marshall

June 2024



Introduction

Decentralised machine learning (ML) (sometimes just referred to as Decentralised Artificial Intelligence) differs greatly from centralised ML (and other types of machine learning). This article seeks to explain those differences and explore the benefits, use cases and regulatory challenges faced by Decentralised Artificial Intelligence.

Machine learning

To understand decentralised machine learning (and other types of machine learning) we must first understand what machine learning (ML) is and how it works. ML is a branch of artificial intelligence (AI) concerned with imitating human learning.³

The principal focus of ML is on using data-driven algorithms to empower AI models with abilities usually associated with human intelligence. It involves taking data and training AI models on that data, as well the queries that the AI models are given, so that patterns and/or predictions are made that improve in accuracy over time.⁴

AI models, empowered by ML, can focus on data of all kinds (e.g. language or images etc) with some AI models now being commonly used including Chat GPT (owned by OpenAI), Gemini (owned by Google) and Meta AI (owned by Meta). These are examples of AI models that are empowered by centralised ML (not decentralised ML). When a user uses an AI model empowered by centralised ML, they are engaging with a centralised entity (i.e. a company) which has collected data which it holds in its own central server for the training of an AI model that it owns and operates (typically for a commercial agenda).

Decentralised ML (sometimes just referred to as Decentralised Artificial Intelligence) differs greatly from centralised ML (and other types of machine learning) and this article seeks to explain those differences and explore the benefits, use cases and regulatory challenges faced by Decentralised Artificial Intelligence.

Types OF ML

There are many different classifications of ML – based on the sources, locations, and architectures of the data and the models involved. In order to further explain Decentralised ML, three of the most well-known classifications can be distinguished.

Centralised ML

Here, data is uploaded from (or collected from) various sources (and various parties) to a central source (such as a server), controlled by one entity, which trains an AI model that is owned and operated by that same entity.⁵ So, the computational burden of training is shifted away from the individual sources of the data to one centrally controlled server, but those individual sources give up (some) rights to their data and do not own and operate the AI model themselves.

Distributed ML

In distributed ML, AI models are trained on data that sits on multiple nodes (being individual devices or entities within that network which each hold a portion of the overall dataset). It uses the collective power of multiple AI models to expedite the training process (as data is processed in smaller subsets on different nodes simultaneously).⁶ Each node performs computations on its assigned subset and communicates the results with other nodes and in some cases a central source to collectively update and improve the AI model's parameters.⁷ So, distributed ML is designed for dealing with large-scale data in a short period of time.

Federated ML (which is a related type of distributed ML) takes the concept of distributed ML further by focussing on preserving data privacy, by allowing the AI model to train on data that is distributed across multiple nodes, in a way that ensures sensitive information always remains on (and never leaves) those local nodes. So, federated ML is particularly useful when training heterogenous datasets across various devices, such as mobile phones or tablets, where data privacy is crucial.⁸

Decentralised ML

Decentralised ML involves interactions between entities that sit within an AI network where control, decision-making, and content creation are distributed across a network of nodes (so, not only is the data set distributed, but control and governance are too).^{1,2} Each entity can make use of a pre-trained AI model on their local devices, and directly communicate AI model updates with their neighbours. So, there is no central server which holds all the data that is controlled by one entity (like there is for centralised ML) and there is no sending of outputs to a central source (like there can be for distributed ML). So, most importantly, whilst the other types of ML involve the role of a central entity, decentralised ML disperses communications, processing and decision making amongst all the entities in a network (without any central source whatsoever).⁹

These AI networks can enable total strangers to come together with self-securing databases (being databases capable of encryption, access control, automated patching, auditing, and the ability to secure data and users). This enables strangers to trust one another – or rather – it means they do not need to trust one another at all, because all interactions between entities in decentralised blockchain networks are done through smart contracts where no single entity has control over the contract, and all entities have equal access to the contract's details.

Benefits of decentralised Machine Learning approaches

Decentralised ML is significantly important in the realm of modern computing and AI for the reasons stated below.

Privacy preservation

Decentralised ML allows AI models to be trained directly on user devices without transferring sensitive data to centralised servers. Moreover, with the rise of edge computing (which moves computer storage and processing to the edge of the network), decentralised ML enables ML tasks to be performed directly on edge devices. This preserves individual privacy by keeping data local, whilst reducing latency and conserving bandwidth.¹⁰

Collaborative ML

Decentralised ML promotes collaborative ML among stakeholders without sharing sensitive data. This facilitates collaboration in sectors like healthcare, finance, and cybersecurity where data privacy regulations or anti-competitive laws may restrict data sharing. Moreover, decentralised ML fosters knowledge sharing and innovation across various organisations.¹¹

Regulatory compliance

Through minimising the need to transfer and store sensitive data centrally, decentralised ML reduces the risk of non-compliance with worldwide regulations. For example, decentralised ML can assist organisations in complying with data privacy laws, such as the General Data Protection Regulation 2018 (GDPR) and the California Consumer Privacy Act 2020 (CCPA).¹²

Resilience

On a more technical level, decentralised ML is not susceptible to disruptions of network caused by single points of failure (which are widely experienced by centralised ML participants – an example is when Chat GPT is at capacity, and unable to be used). Decentralised ML spreads tasks across multiple nodes, so, even if some nodes are faulty, the ML process can continue, ensuring constant operation.¹³

Scalability

Decentralised ML distributes computation and ML tasks across multiple devices, enhancing scalability (compared to centralised ML). This proves useful in large scale data-processing

scenarios, such as IoT networks or edge computing.¹⁴

Use cases

Many experts in the AI field agree that decentralisation is paramount to AI's success and ethical application.¹⁵ This is because decentralisation demonstrates a shift from siloed computational AI models to a more collaborative approach using open-source AI models and GPUs' (graphics processing unit) collective power across a network.¹⁶

Decentralised ML has many potential use cases, especially in domains where data is sensitive and it is already used and well known within the crypto and digital asset industry (e.g. some DeFi platforms, which are built on decentralised blockchain infrastructure and

enable direct control over financial assets without intermediaries are making use of decentralised ML to train AI models). There are domains where its use could be of significant benefit to participants such as the ones listed below.

Healthcare

In healthcare, data privacy laws typically prohibit personal data from being moved outside the country of origin. Decentralised ML will enable personal data to remain in its original country whilst being used to train AI models globally. A well-designed decentralised ML AI network could enable different medical entities (such as hospitals or medical centres) to collaborate to improve the quality of AI models, which enables democratic access to knowledge and sharing of the benefits, meaning the time taken to find solutions to illnesses may be reduced.¹⁷

Supply chains

Decentralised ML can improve product transparency and mitigate fraud by digitising paper-based processes, and by enabling the real-time tracking of goods from production to delivery. Therefore, AI-driven predictive analytics coupled with blockchains can enhance companies' demand patterns, inventory management, and decision-making, ultimately minimising costs.¹⁸

Financial services

The main advantage of regulatory compliance (discussed above) of decentralised ML could be seen in robo-advisory platforms (and centralised examples include Vanguard in the US and Revolut in Europe). If there was a shift towards using decentralised ML networks, a larger dataset would be available, and a decentralised robo-advisory platform could enhance the accuracy in predicting market trends, therefore limiting the financial risks retail investors are exposed to.

Aerospace and Defence Decentralised ML can play a transformative role in the aerospace and defence sector through its collaborative feature (discussed above). By integrating decentralised ML, military capabilities can benefit from a more robust operation system, particularly in multi-domain operations requiring co-ordination across space, sea, land, and air.¹⁹

Agriculture

Through its collaborative feature (discussed above), decentralised ML could provide farmers with real-time insights into crop health. It could help farmers identify areas that need irrigation, fertilisation, or pesticide treatment. Moreover, the privacy preservation advantage (discussed above) of decentralised ML has already been successfully tested in the agriculture sector. For example, blockchain-based platform AgroBLF is an innovative agricultural supply chain framework which operates through smart contracts, addressing challenges in agriculture such as price fluctuation, supply chain transparency, and storage facilities.²⁰

Autonomous vehicles

Decentralised ML could play a crucial role in shaping the future of autonomous vehicles. As data is processed on different nodes that act independently and integrate solutions through collaboration, decentralised ML could potentially improve tasks like perception, trajectory prediction, and route planning. Decentralised ML could combine data from various sensors (such as LiDAR, radar, and cameras) to create a comprehensive perception system.²¹

Challenges of decentralised ML

Whilst decentralised ML has many benefits such as the ones listed above, it contends with challenges related to data quality, security, and communication.

- Data challenges involve issues with data quality, data heterogeneity, or data biases of the healthcare participants.
- Privacy and security challenges include data poisoning attacks, adversarial attacks, membership inference attacks, or free-riding attacks in the collaborative network.
- Communication challenges might occur due to energy consumption issues, or computational and communication overheads prevalent within systems.²²

Open-source AI models, while promoting decentralisation, often rely on synthetic

data from commercial platforms (such as Chat GPT). Synthetic data, while useful for training AI models, is often generated by proprietary algorithms which are centrally controlled, posing a significant concern to true decentralisation which aims to move away from centralised systems.²³

An additional challenge to decentralised ML is the cost. Decentralised ML requires a robust infrastructure to handle the distributed nature of the system, which can be expensive for participants to set up and maintain. This is the same challenge faced by centralised ML, and whilst the goal of decentralised ML is to share the overheads, much of the venture capital and private equity funding goes to centralised entities which have the benefit of having harvested people's data, for free, for years. So, it is difficult for decentralised models to compete.

Decentralised ML and blockchain

AI combined with blockchain technologies could enhance data-driven decisions and resource management in various sectors such as,

healthcare, agriculture, urban planning, education, and energy.²⁴

Blockchains are immutable databases that use distributed networks of nodes instead of a centralised server.²⁵ So, they have a lot in common with the concept of decentralised AI. Blockchains use a data structure that interconnects each transaction, meaning any attempts to manipulate a transaction on a public blockchain typically require very significant computing power for any success. Anyone can view the history of transactions on a blockchain, so they create a transparent environment for data sharing. Decentralised AI integrates advanced technologies, including distributed computing and secure encryption methodologies, to facilitate a robust and efficient mechanism for safeguarding data privacy and security during collaborative AI model training endeavours.²⁶

SingularityNet: a decentralised project

The SingularityNET project aims to create a democratic, collaborative, and decentralised Artificial General Intelligence (AGI). This AGI will be designed to operate independently of any central authority, welcoming to everyone, and free from the narrow objectives of any individual corporation or nation.

As a precursor to this long-term vision, the project currently focusses on allowing anyone to create, share, and monetise AI services using a self-organising network of agents. This network of agents has a decentralised decision-making process, whereby AGIX token holders are empowered with governance rights (i.e. they can vote and decide on material changes to the network). This means changes to the network are user driven (rather than as a result of the commercial agenda of a single corporation).

Anyone in the network can create, share, and charge for AI services and these services can be used by anyone with an internet connection. The network is powered by blockchain technology, ensuring each transaction is secure, can execute automatically if conditions are met and can be completed without the need for an intermediary or central authority. Moreover, the marketplace exercises cross-chain interoperability, allowing the platform to function across different blockchain networks, further enhancing its reach and utility.

SingularityNET recently merged with Fetch.ai and Ocean Protocol to form the Artificial Superintelligence (ASI) Alliance, further demonstrating the growing interest in the decentralised AI industry and its growth.²⁷

Challenges of decentralised AI under the EU AI Act 2024

Of course, the SingularityNet project and other decentralised projects alike will face challenges as their participants navigate new and differing nation state approaches to both AI and blockchain regulation.²⁸

For example, the EU AI Act 2024 regulates AI through a risk-based approach which includes the following four levels of risks: unacceptable risk, high risk, limited risk, and minimal risk.

- The unacceptable risk category bans AI models that pose clear threats to human safety or rights. Examples include services such as biometric identification and social scoring.
- The high-risk category covers AI models that threaten human safety or fundamental rights. Examples include the use of AI for hiring and recruitment.
- The limited risk category targets AI systems with specific transparency needs. For instance, chatbots for customer service require users to know that they are conversing with a machine.
- The minimal risk category covers AI-powered video games and email spam filters. However, these systems have the lowest level of risk, and there are no specific obligations under the EU AI Act 2024.

One of the biggest challenges is determining the difference between high risk and unacceptable risk AI systems (especially as the list of activities set out in the EU AI Act 2024 as having an unacceptable risk are just examples – and so there is room for arguing edge cases that may or may not be unacceptable). In practice, this means businesses will need to ensure that high risk activities (which are not prohibited) do not

suddenly become unacceptable risk activities (which are prohibited).

The EU AI Act 2024 demands extensive control over development processes, including risk mitigation and data governance. Decentralised systems (and their participants), which may not have centralised data storage or clear lines of responsibility, could find it challenging to meet these requirements. Moreover, the

EU AI Act 2024 requires significant record keeping, which can be burdensome for decentralised AI, where development and decision-making processes are spread across multiple nodes.

Proof of Honesty (PoH)

One of the biggest challenges for projects that seek to combine blockchains technology and decentralised AI networks will relate to the appropriateness of the consensus mechanism that facilitates economic activity on the platforms they develop. Whilst in centralised AI, there is a central server which controls data and ensures accurate activity in line with participants' wishes, decentralised AI models could be susceptible to the dishonest practices of some participants.

The Proof of Honesty (PoH) protocol is a consensus mechanism which seeks to solve issues of dishonesty, achieving high security, with low computational resources. Transfers are made when a node proves its honesty with the PoH network requiring only one honest node to function.²⁹

The PoH protocol incentivises service providers with payments proportional to network contributions, allowing decentralised AI models to store datasets in geographically distributed locations. Then, users collect dataset pieces from service providers who receive a reward directly proportional to their contribution to the dataset pieces. The PoH protocol encourages service providers to cache frequently accessed datasets in close proximity and deploy them in high-bandwidth zones, thereby optimising network topology and resource allocation in a decentralised way.

The PoH protocol incorporates test tasks (i.e. "phishing tasks") within real tasks. Therefore, if a service provider does not accurately and honestly complete tasks, the "PoH officers" catch them

using test tasks, deterring dishonesty in AI computing. So, by incorporating test tasks into regular tasks, the PoH protocol promotes integrity and motivates providers to deliver their best endeavour, as any task could be a test – and in this way, it could play a significant role in keeping participants within decentralised AI networks honest.

Carbon footprint/energy consumption

It is estimated that AI models use four to five times more energy than conventional web searches, and within years, they could consume as much energy as entire nations. Moreover, AI models will likely need large amounts of fresh water to generate electricity and cool their processors.³⁰

Decentralised ML powered AI models could be beneficial in combating this energy consumption issue. Through their decentralised nature, these AI models can distribute the workload across multiple nodes. Each node performs local computations, resulting in lower energy consumption per device (as instead of relying on a few large servers, decentralised ML leverages many smaller devices). Moreover, traditional centralised data centres consume significant amounts of water for cooling, whereas decentralised ML with its distributed architecture minimises water usage, thereby achieving sustainable computing.

As decentralised ML systems are inherently decentralised, they can withstand disruptions (such as natural disasters and power outages) better than centralised models. This resilience coupled with the environmental benefits discussed above contribute to long-term environmental sustainability.

Intellectual Property rights

What data can be used to build and develop decentralised ML powered AI models, who can use the data to train the AI models, and who owns the AI models' output are all affected by intellectual property rights, such as copyright, trademarks and patents. The law in this area will

likely be shaped by the results of legal cases that are in progress.³¹

— **Copyright:** A work must be “original” to attract copyright protection. In particular, the raw data that each party holds may be subject to copyright protection, depending on the nature, originality, and jurisdiction of the data and in turn if the data used to train the AI model is still discernible in the AI model's output, this may still be protected by copyright (s178 Copyright Designs and Patents Act 1988 (CDPA)). For example, a database of literary works may be copyrighted, while a database of factual and confidential information may not be. By using decentralised ML, the parties can ensure that their copyrighted or sensitive data remains at its original source, preserving its ownership and control over their data. The case of *Getty Images v Stability AI* [2023] EWHC 3090 (Ch) highlights the risks posed by the use of copyrighted material for a dataset in the context of generative AI. Getty accused Stability (which has a centralised AI model) of using its dataset of photos without a licence to train their AI model (called “Stable Diffusion”) to produce more accurate outputs.³² If a decentralised AI model had been used in the above-mentioned case, the AI model may be able to analyse content to determine if it is original or if it potentially infringes on existing copyrighted works (and it may be able to be trained in a way whereby the data used to train it never leaves the source). This can be particularly useful in cases where content is remixed or transformed, as decentralised AI processes data across multiple nodes. However, the very nature of decentralised AI can raise problems when assessing the responsibility for copyright infringement across parties.³³

— **Trademarks:** The parties may have their own trademarks that they use to distinguish themselves from their competitors or to establish their brand identity. By using decentralised ML, the parties can avoid disclosing their trademarks to each other or to the central server, thus preventing potential infringement, dilution, or confusion of their trademarks. Moreover, decentralised ML could potentially improve

transparency in cases like the MetaBirkins dispute,³⁴ through its distributed approach to managing and verifying trademarks. The MetaBirkins case involved artist Rothschild being sued by Hermès for creating and commercialising around 100 NFTs (non-fungible tokens) extremely similar to the “Birkin” bag.³⁵ By using decentralised ML, the authenticity of the MetaBirkins NFTs could have been verified by analysing the metadata and the history of the transactions associated with the MetaBirkins NFTs. Additionally, decentralised ML could be integrated with smart contracts to automatically execute terms agreed by the parties, such as ceasing the sale of the MetaBirkin NFT if it was found to infringe on trademark rights.

- **Patents:** The parties may have their own patents that cover their inventions related to their data, AI models, or applications. By using decentralised learning, the parties can avoid infringing or disclosing their patents to each other or to the central server, thus maintaining their competitive advantage or innovation potential. However, complications can arise as to the infringement of patents, as UK patent laws impose liability for “making, disposing of, offering to dispose of, using, importing or keeping a patented product without consent” (s 60(1) Patents Act 1977). Issues regarding infringement are also raised by the use of decentralised ML in processes which are patented. A patented process might consist of multiple steps. The completion of one of these steps by a party who doesn’t have, say, a licence from the patent holder to carry out that step of the process (or the process in its entirety) might constitute an infringement of that patent. If

such a step was to be completed by a decentralised ML algorithm, a question arises over who has infringed the patent: the user of that decentralised ML system, the original creator of the system, or even the system itself?

While regulation catches up to innovation, only for innovation to inevitably race ahead again, those developing and using AI should be vigilant in managing their activities to safeguard against intellectual property-related risks.³⁶ This is especially important in the context of decentralised AI, where the use of a vast number of data sources increases the possibility that one of those sources might infringe an intellectual property right (such as copyright, trademarks or patents). If the AI model used by all sources has been trained in part by a source that has committed an infringement whilst doing so, then all sources using the AI model are potentially benefiting from the infringement. This issue may be easier to safeguard against in centralised ML approaches, as all the AI model training happens in one centralised location, which intrinsically can be overseen more easily.

Conclusion

- Experts believe that decentralisation is key to the success of AI. AI models are always most effective when they can draw from a larger dataset. Decentralised ML aims to democratise the AI creation and application by shifting from monolithic computational systems to distributed, collaborative approaches, making this transformative technology more accessible to people across the world.

Biog box

Tom Marshall is a senior associate at CMS and is currently on secondment with the institutional legal team at Coinbase. He plays a key role in CMS’s crypto and digital assets team advising exchanges, wallet providers, financial institutions, protocols, DeFi projects, NFT issuers and platforms, DAOs and metaverse projects as well as most of the top VCs and hedge funds in crypto. Tom also works with traditional clients on their digital assets projects, including accepting crypto payments, NFTs, Web3 development, sponsorship, partnering, metaverse and the tokenisation of real world assets. Email: tom.marshall@cms-cmno.com.

References

1. Decentralized and collaborative machine learning framework for IoT (sciencedirect.com)
2. Decentralized federated learning based on blockchain: concepts, framework, and challenges” (<https://www.sciencedirect.com/science/article/abs/pii/S0140366423004851>)
3. What is machine learning (ML)? (ibm.com)
4. Machine learning, explained (mitsloan.mit.edu)
5. Federated Learning Versus Classical Machine Learning: A Convergence Comparison (Nagoya Institute of Technology)
6. Elevating ML to new heights with distributed learning
7. From distributed machine to distributed deep learning: a comprehensive survey (Journal of Big Data)
8. Exploring the distributed learning on federated learning and cluster computing via convolutional neural networks
9. Decentralized vs. Distributed Organization: Blockchain, Machine Learning and the Future of the Digital Platform (journals.sagepub.com)
10. What is decentralized machine learning and top 3 amazing ways blockchain can benefit from it
11. What is decentralized machine learning and top 3 amazing ways blockchain can benefit from it
12. What is decentralized machine learning and top 3 amazing ways blockchain can benefit from it
13. What is decentralized machine learning and top 3 amazing ways blockchain can benefit from it
14. What is decentralized machine learning and top 3 amazing ways blockchain can benefit from it
15. What is Decentralized AI? (moralis.io)
16. The Case for Decentralizing AI | Built In
17. AI Should Be Decentralized, but How? (coindesk.com) <https://www.coindesk.com/consensus-magazine/2023/09/22/ai-should-be-decentralized-but-how/>
18. Blockchain and AI - Use Cases | Chainlink
19. AI Revolutioni.
20. AgroBLF: blockchain-based framework for smart agriculture
21. Fayyad, Jaradat, Gruyer and Najjaran, 'Deep Learning SensorFusion for Autonomous Vehicle Perception and Localisation: A review' (2020)
22. Decentralized_Learning_in_Healthcare_A_Review.pdf
23. Decentralized AI will play a pivotal role in shaping the future of AI | TechRadar
24. Blockchain and AI - Use Cases | Chainlink
25. What is Decentralized AI? (moralis.io)
26. Decentralized federated learning based on blockchain: concepts, framework, and challenges - ScienceDirect
27. SingularityNET to merge crypto network with Fetch.AI and Ocean Protocol

28. AMB Crypto, 'What is AGIX? SingularityNET Ecosystem Explained (Crypto Meets AI).
<https://medium.com/singularitynet/what-is-agix-singularitynets-native-token-explained-d82e37e0d76a>
29. Geeq Review: Proof of Honesty, Infinitely Scalable, High TPS
30. Generative AI's environmental costs are soaring – and mostly secret
31. Artificial intelligence toolkit (UK) | Practical Law (thomsonreuters.com) Artificial intelligence and copyright: checklist | Practical Law (thomsonreuters.com) Section I. - Introduction | Practical Law (thomsonreuters.com)
32. Getty Images lawsuit says Stability AI misused photos to train AI (reuters.com)
33. The copyright status of AI-generated works.
34. Hermès International and Hermès of Paris v Mason Rothschild [2023].
35. Hermes wins U.S. trademark trial over 'MetaBirkin' NFTs.
36. Using generative AI and protecting against copyright issues | World Economic Forum (weforum.org)

Further reading

- Blockchain interoperability: connecting decentralized infrastructure for traditional finance (2022) 7 JIBFL 470.
- Artificial intelligence in syndicated lending (2024) 5 JIBFL 327.
- Lexis+® UK: Company Secretary's Review: Artificial Intelligence: sustainability risks and opportunities.