February 20th, 2024

AI x Crypto Primer

Mohamed Baioumy & Alex Cheema

Contents

Acknowledgements

Thanks to Arbion Halili for his substantial contributions to the writing, typesetting and editing of this document.

Thanks to the following people for feedback & discussions which helped shape this report: Sandeep Nailwal (Co-founder of Polygon), Harry Grieve (co-founder of Gensyn), Dima Romanov (co-founder of LayerN), Mohamed Fouda (AllianceDAO), Oliver Turnbull (PhD researcher at Oxford University), Altan Tutar (NEAR Protocol), DCBuilder (Contributor at WorldCoin), Pim & Bram van Roelen (Maven11), Will Papper (Syndicate.io), Bharat Krymo (6529 Capital), Justin McAfee (1kx), Rahilla Zafar (Stability AI), Raphael Doukhan (Giza), Jack Chong (Checker Finance), Yarco Haydek (Pragma), Piotr Saczuk (AlephZero), Will Robinson (AllianceDAO), Illia Polosukhin (co-founder of NEAR), David Ma (AllianceDAO), Nick Emmons (co-founder of Upshot), Siyuan Han (ABCDE) and Qiyun Lu (BeWater).

Special thanks to the BeWater team for translating this document to Chinese as well as offering us valuable feedback.

Additional thanks to our fellow HomeDAO members: Eito Miyamura, Josh Lavorini, Konrad Kopp, Joseph Rowell, Akshaiyram Reguram, James O'Leary, Lucas Schneider, Orlando Fraser, Alex PF and Simon Sällström.

Finally, thanks to everyone to helped review this report including the Foresight team, Tze Donn Ng, Owen Shen, Neel Daftary, Dylan Zhang, Bart Jaworski, Matt Garcia, Cathie So, Kartin Wang, Anand Iyer, Shafu, Lucian Zhu (DAO BASE), Morgan (Hopscotch), Young Ling, Will Nowak, Anna Bertha (DCG), Bobby (UMA Protocol), Philipp (GnosisDAO), Mikko Ikola (Veridise), Matt Garcia, Louis Guthmann (Starkware), Francis Zhan (Tribe Capital), Tim (flock.io), Gabriel Chua, Nitanshu Lokhande, 0xAA, Cesar Rodriguez (IQ), Ethan (MyShell), Ayush Patel (BuildersTribe), Chase Ando (Pickem), Leo (Supersight), Sid Rao, David Liu, Fred, Gabriel C and Tommy Shaughnessy.

Introduction

Artificial Intelligence (AI) will cause unprecedented levels of societal change.

The rapid growth in this space will inevitably lead to widespread economic disruption, with AI creating new possibilities in every industry. The Crypto industry is no exception. We observed three major DeFi attacks in the first week of 2024 and $76B stored in DeFi protocols at risk. Using AI, we can check smart-contracts for security holes, and integrate an AI-based security layer into blockchains.

The limitations of AI permit bad actors to abuse powerful models, as shown by the spread of malicious deepfakes. Thankfully, various advancements in cryptography will introduce new functionalities to AI models, greatly enriching the AI industry whilst addressing some of its most critical shortcomings.

The convergence of AI and Crypto will birth numerous high-traction projects. While some will provide a solution for the problems discussed above, others will combine AI and Crypto in superficial ways with no real benefit.

In this report, we introduce conceptual frameworks, supported by specific examples and actionable insights, to aid your understanding of this space's past, present and future. We will uncover the real opportunities and help you steer clear of inconsequential distractions.

A Note From The Authors:

We were caught off guard by the release of Sora, an AI model that can create realistic video from text instructions.

Sora was announced whilst this report was being written.

This is confirmation that the risks that we have highlighted in this report are real and that they are worsening in real-time.

Core Frameworks for AI x Crypto

This section will introduce some practical mental tools for understanding AI x Crypto, helping you to analyse individual projects in greater detail.

2.1 What Makes Something An AI x Crypto Project?

Let's review some examples of projects which use both Crypto and AI, then discuss whether they are actually AI x Crypto projects.

Case Study: Privasea

Privasea utilizes Fully Homomorphic Encryption (FHE) to train AI models for processing encrypted data, ensuring input privacy. This innovation facilitates a ChatGPT-like model that operates on encrypted prompts, meaning the service provider never sees unencrypted content. The system processes these encrypted inputs and generates encrypted outputs, which only the user can decrypt and understand. This method protects user data from exposure and is particularly crucial for sensitive sectors like finance, where safeguarding against potential information leaks is paramount. By enabling secure, encrypted interactions with AI, Privasea addresses significant privacy and security concerns in AI applications.

This case study shows Crypto helping to improve an AI product – a cryptographic method being used to change how the AI is trained. This results in a product that would not be possible using AI techniques alone: a model that accepts encrypted prompts. Adding encryption is an example of using a cryptographic technology to improve the capabilities of AI.

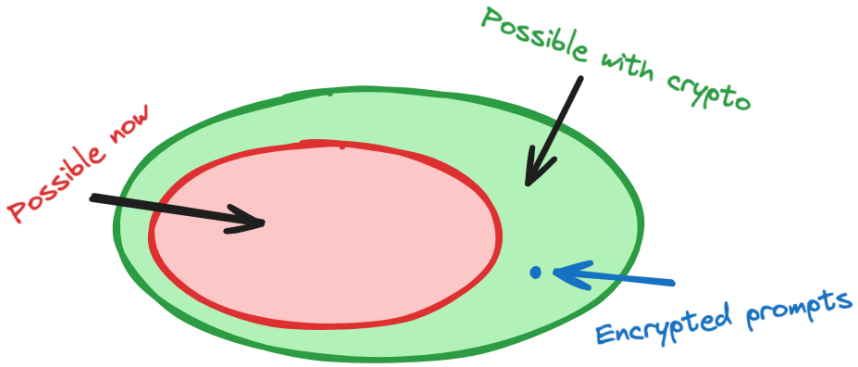

Figure 1: Using crypto to make internal changes to the AI stack results in novel capabilities. For example, FHE allows us to use encrypted prompts.

Case Study: Dorsa[1]

Dorsa is a blockchain security start-up, using AI models to audit and monitor smart contracts. They have developed a variety of tools, including reinforcement learning agents to extensively test the security of a smart contract. The tools developed by Dorsa can help both developers and auditors; these AI models can improve both the security and the reliability of smart contracts.

In this case, AI technology is being used to improve a Crypto product – the inverse of what we discussed earlier. Dorsa offers an AI model that makes the process of crafting a secure smart-contract faster and cheaper, and although it is off-chain, the use of the AI model still contributes to the functionality of Crypto projects: smart-contracts are often at the core of Crypto solutions.

Dorsa's AI can spot bugs that humans forget to check for, thus preventing future hacks. Yet, this particular example doesn't use AI to make Crypto products capable of things that they couldn't do before – it was always possible to write secure smart-contracts, Dorsa's AI just makes the process better and faster. Nonetheless, this is an example of AI technology (the model) improving a Crypto thing (the smart-contract).

Case Study: LoverGPT

LoverGPT is a Web2 AI girlfriend app. They don't censor explicit chats or images and thus might flag issues with Visa and Mastercard (similar to this company getting censored). To address this issue, they are accepting payment in crypto. This is the only way in which LoverGPT utilises crypto technologies.

LoverGPT is **not** an example of Crypto x AI. We've established that AI can help improve the crypto stack, and vice versa – this is illustrated by the examples of Privasea and Dorsa. However, in the example of LoverGPT, the Crypto components and AI components are not interacting – they simply co-exist within the product. For something to count as an AI x Crypto project, it is **not** enough for AI and Crypto to contribute to the same product or solution – the functionality provided by these technologies must interlock to produce the solution.

> AI and crypto are often combined in superficial ways that don't have tangible value.

[1]Dorsa is a company in stealth, there is no public infomation about this firm at the time of publication.

Crypto and AI are technologies which can be directly combined to produce better solutions. The use of one in combination with the other can make the other work better as part of the overall project. Only projects that involve this co-operation between technologies are classified as AI x Crypto projects.

2.2 How AI and Crypto Can Help Each-Other

Notice:

We use the word "Crypto" to refer to both blockchain technologies and cryptographic solutions across this report. Zero-knowledge proofs (ZKPs), cryptographic signatures and blockchain ledgers are all examples of "Crypto".[2] We'll clarify which aspect of Crypto we are referencing in each section where appropriate.

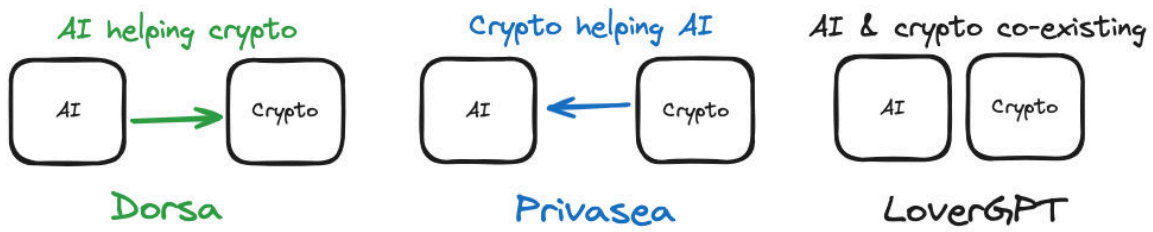

Figure 2: Illustration of the 3 ways AI and crypto interact within a product.

Let's recall the previous case-studies. In Privasea, FHE (which is Crypto) is used to produce AI models which are capable of accepting encrypted inputs. Therefore, we are using a Crypto (cryptographic) solution to improve the process of AI training; thus, Crypto is helping AI. Meanwhile, in Dorsa, an AI model is being used to make secure-contracts. Here, an AI solution is being used to improve an aspect of the Crypto product; thus, AI is helping Crypto. This brings us to an important distinction when evaluating a project at the AI x Crypto intersection: Is crypto being used to help AI, or is AI being used to help Crypto?

This simple question helps us uncover an important aspect of the use case at hand. What is the key problem being solved here? In the case of Dorsa, the desired outcome is to produce a secure smart-contract. This can be developed by a skilled human, but Dorsa happens to make this process more efficient with AI. The end product of a secure smart-contract remains constant. Once the key problem is clear, we can identify whether AI is helping crypto or crypto is helping AI. In some cases, there is no meaningful interaction between the two (e.g. LoverGPT).

[2]This is consistent with how the term is used in other contexts – Vitalik refers to "powerful crypto in the form of not just blockchain scaling solutions, but also ZKPs, FHE ..."

The table below provides several examples in each category.

Crypto Helping AI	AI Helping Crypto
1. FHE being used to train an AI model in a privacy-preserving manner 2. Federated learning for AI model training 3. Zero-knowledge proofs for verifiable inference 4. Token incentive for collecting and labelling data sets 5. Payment rails for AI agents 6. Decentralized GPU networks	1. AI models auditing and monitoring smart contracts 2. AI models being used for efficient search of on-chain data 3. LLMs for creating data analytics dashboards 4. Intent-based trading platforms 5. Customizable bots for on-chain games 6. AI agents managing a DAO

Table 1: How Crypto And AI Can Support Each Other.

You can find our directory of more than 150 AI x Crypto projects in the Appendix. If there is anything we've missed, or if you have any feedback, please reach out!

2.2.1 Mini-Summary

Both AI and Crypto are capable of supporting the other technology in serving its purpose. When you assess any product, it's key to understand whether it is an AI product at its core, or a Crypto product at its core.

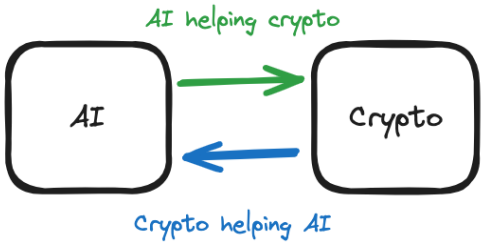

Figure 3: An illustration of the distinction.

2.3 Internal vs External Support

Let's consider an example of Crypto helping AI. When the collection of specific techniques which make up an AI technology changes, then what the AI solution as a whole is capable of changes. This collection of techniques is called a **Stack**. The AI stack includes the mathematical ideas and algorithms which make up every aspect of the AI

model. The specific techniques used for the processing of the training data, the training of the model, and model inference, are all part of the Stack.

> Stacks, as a whole, have properties and capabilities which are not present in any of its individual parts. It's the interaction between the parts, like cogs in the clock, that is responsible for the higher-level behaviour of the Stack.

Within a Stack, its parts are deeply interconnected – it's the way that the specific techniques are combined that is responsible for the features of the Stack. Therefore, by altering the Stack, you alter what the technology as a whole can achieve. Introducing new techniques into a stack can create new technological possibilities – Ethereum added new techniques within their Crypto Stack to make smart-contracts possible. Similarly, changes to the Stack can allow developers to bypass problems which were previously considered inherent to the technology – the changes made by Polygon to their version of the Ethereum Crypto Stack enabled them to reduce transaction fees to a level that was considered impossible.

> When the methods (techniques) used in one technology are used within the Stack of another technology, this represents a deeper and more technical integration of the technologies, which we will call "internal". If two technologies are used together, in support of one another, but the pieces which make up each stack do not change, then the integration will be "external".

- **Internal Support:** Crypto is used to make changes which are internal to an AI stack, such as changing the technological means by which models are trained. For instance, one can introduce FHE techniques into the AI stack, as Privasea does. Here, a cryptographic component is built directly into the AI stack, resulting in a modified AI stack. In the case of Privasea, the use of FHE techniques.
- **External Support:** Crypto is used to support the AI-based features without making changes to the AI Stack. Bittensor is an example of this, incentivizing users to contribute data - data which can then be used to train an AI model. Here, no change to how the model is trained or used under the hood; the AI stack doesn't undergo any changes. The use of financial incentives within the Bittensor network nonetheless helps the AI Stack serve its purpose better.

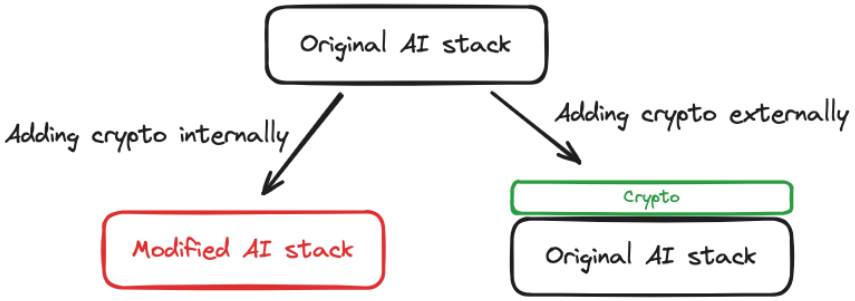

Figure 4: An illustration of the previous discussion.

Similarly, AI can help the Crypto either internally or externally:

- **Internal Support:** AI techniques are used within the Crypto Stack. The AI lives on-chain and interfaces directly with components inside the Crypto Stack. As an example, take an on-chain AI agent that governs a DAO. This AI doesn't just assist the crypto stack. It is an integral, deeply embedded part of the technological Stack which makes the DAO function.
- **External Support:** AI is external to the Crypto Stack. The AI is being used to support the crypto stack without making internal changes to it. Platforms such as Dorsa use an AI model to secure a smart contract. The AI is off-chain and is an external tool used to make the process of writing secure smart contracts faster and cheaper.

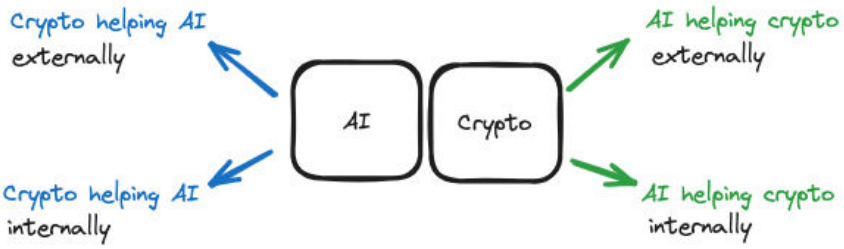

Figure 5: Here's the model, updated to include the distinction between internal and external support.

The first stage of analysis for any AI x Crypto project is deciding which category it belongs to.

2.4 Identifying Bottlenecks

Internal support, which features deep technological integration, tends to require more technical difficulties than examples of external support do. For instance, if we want to modify the AI stack by introducing FHE or zero-knowledge proofs (ZKPs), we will need staff who have considerable expertise in both cryptography and AI. Very few people fall into this intersection. Companies such as Modulus, EZKL, Zama and Privasea,

therefore, need substantial financial resources and rare talents to advance their solutions. Allowing users to integrate AI inside smart-contracts requires similarly in-depth knowledge; companies like Ritual and Ora must tackle complex engineering problems.

External support, conversely, has bottlenecks, but they often involve less technological complexity. For instance, adding Crypto payment features to an AI agent doesn't require us to overhaul the model. It's relatively easy to implement. Whilst building a ChatGPT plugin which allows ChatGPT to fetch statistics from the DeFi LLama webpage is not technically complex for an AI engineer, **very few AI engineers are members of the Crypto community.** Although the task is not technically complex, few AI engineers are willing to implement these tools, and many are unaware of these possibilities.

Description	The Main Bottleneck To Innovation
Crypto helping AI, internally	Fundamental deep-tech challenges need to be solved, including scaling ZKML and scaling homomorphic encryption.
Crypto helping AI, externally	Good economic models, aiming, for instance, to incentive data collection or reward users for contributing their GPUs to a network.
AI helping crypto, internally	Engineering challenges. AI tools are mature relative to the crypto stack, so adding AI requires solving complex engineering issues.
AI helping crypto, externally	The scarcity of AI talent in the crypto space.

2.5 Measuring Utility

There will be good projects across all four of these categories.

If AI is integrated inside the Crypto Stack, smart-contract developers will be able to access to AI models on-chain, thus increasing the number of possibilities and potentially resulting in a wide range of innovations. This is equally true of cases where Crypto is integrated into the AI Stack – deep technological fusion will produce new possibilities.

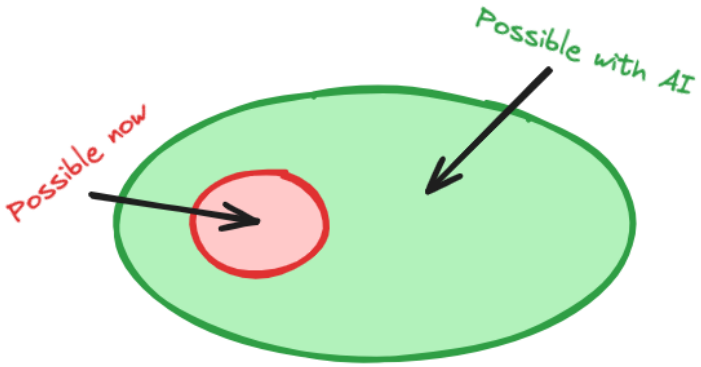

Figure 6: Adding AI to the crypto stack gives developers new capabilities.

In the case of AI helping Crypto externally, it is likely that an AI integration will improve an existing product whilst producing fewer breakthroughs and introducing fewer possibilities. For example, using an AI model for writing smart contracts might be faster and cheaper than before, and may also improve security, but it is unlikely to produce smart-contracts that weren't possible before. The same applies to Crypto helping AI externally – token incentives can be used around the AI stack, but that in itself is unlikely to redefine how we train AI models.

In summary, the integration of techniques from one technology into the tech Stack of another is likely to produce novel capabilities, whilst the use of a technology outside of the Stack is likely to improve usability and efficiency.

2.6 Assessing Projects

The amount of that a given project can be estimated partially based on what quadrant it fits into, since examples of internal support between the technologies can yield greater rewards, but estimating the total risk-adjusted upside of a project will require us to consider more factors and risks.

One factor to consider is whether the project under consideration can be useful within the context of Web2, Web3, or both. AI models with FHE features can be used in place of AI models that don't have FHE – introducing FHE is useful across both domains, in any context where privacy is valuable. However, an integration of AI models within smart-contracts is only relevant in the Web3 context.

As suggested previously, whether the technological integration between AI and Crypto within a project is internal or external will also determine upside potential – projects that involve internal support will tend to produce novel capabilities and more drastic efficiency gains, which are more valuable.

We must also consider the time horizon over which this technology can be expected to mature – this will determine how long one would have to wait for a return on their

investment into the project. This can be done by analysing current progress and considering the bottlenecks which are relevant to the project (see Section 2.4).

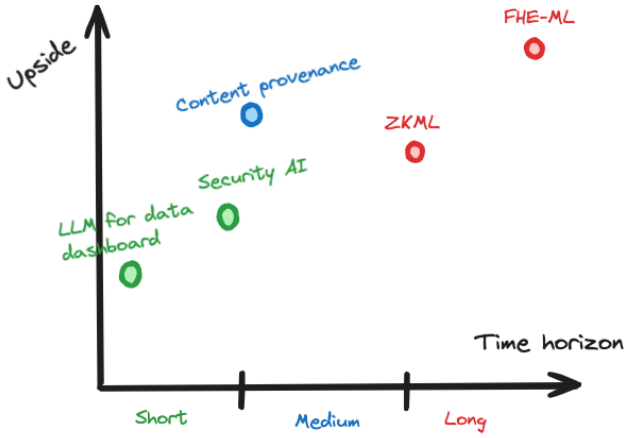

Figure 7: A hypothetical example illustrating potential upside compared to time horizon.

2.7 Understanding Complex Products

Some projects will involve a combination of the four categories that we have described, rather than just one. When this is the case, the risks and benefits associated with the project tend to multiply, and the time horizon over which the project can be executed will be longer.

Additionally, you will have to consider whether the whole of the project is better than the sum of its parts – it's often not enough to satisfy end users for a project to have a bit of everything all in one place. A focused approach tends to yield great products.

> Leveraging all four combinations of Crypto and AI can make a project more valuable, but only if the whole is greater than the sum of its parts.

2.7.1 Example One: Flock.io

Flock.io allows you to train a model such that it is "split" between multiple servers, and no party has access to all of the training data. Since you can participate in training the model directly, you can use your own data to contribute to the model without revealing any of it. This is good for user privacy. This involves crypto helping AI internally, as the AI stack (model training) is changed.

Additionally, they reward the people who participate in the training of models using Crypto tokens and also use smart-contracts to financially penalise people who sabotage the training process. This does not alter the processes involved in training the model, the underlying techniques stay the same, but there would be no way to keep the parties

involved in the training of models honest without the on-chain slashing mechanism. This is an example of Crypto helping AI externally.

Crucially, crypto helping AI internally introduces a new capability: a model can be trained across a decentralized network while keeping data private. However, crypto helping AI externally doesn't introduce a new capability, as the token is merely used to incentivise users to contribute to the network. The users can be compensated in fiat already. Incentivizing users with cryptocurrency is superior and increases the efficiency of the system, but it does not introduce new capabilities.

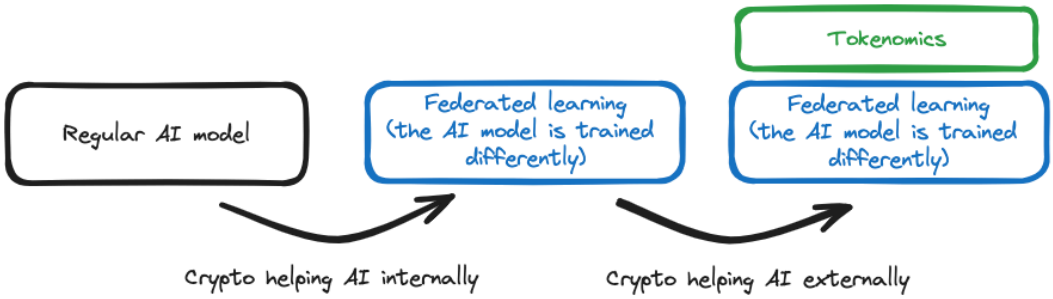

Figure 8: An illustration of Flock.io and how the stack changes, where color changes imply a change has been made internally.

2.7.2 Example Two: Rockefeller Bot

Rockefeller Bot is a trading bot which operates on-chain. It uses AI to decide which trades to make, but since the AI model isn't run on the smart-contract itself, we rely on the service provider to run the model for us, then tell the smart-contract what the AI has decided, and prove to the smart-contract that they are not lying. If the smart-contract didn't check that the service provider wasn't lying, the service provider could make harmful trades on our behalf. Rockefeller Bot allows us to prove that the service provider is not lying to the smart-contract using ZK proofs. Here, ZK has been used to change the AI Stack. The AI Stack needs to adopt ZK techniques, otherwise, we could not use ZK to prove what the model decided for the smart-contract.

The resulting AI model, which has verifiable outputs because of ZK techniques, can now be queried from the blockchain on-chain, meaning that this AI model is used inside the Crypto stack. In this case, we have used AI models within the smart-contract to decide trades and prices in a fair manner. This was not possible without the AI.

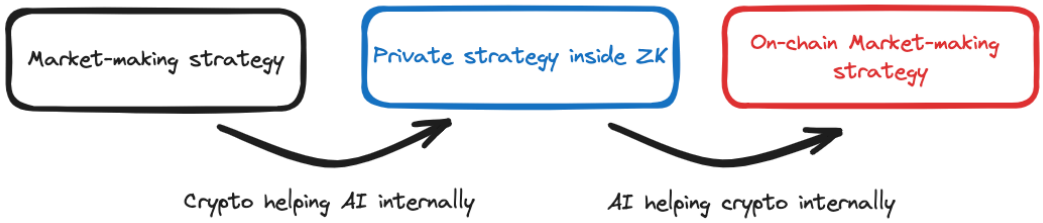

Figure 9: An illustration of Rockefeller Bot and how the Stack changes. Color changes imply a change in the Stack (internal support).

The Big Ideas

3.1 Crypto and The Deepfake Apocalypse

> *"One computer can create all this havoc … How are we going to operate without understanding some basic things about truth?"*
>
> – Eric Schmidt, Former CEO of Google to Sergey Nazarov, co-founder of Chainlink

On the 23rd of January, an AI-generated voice message falsely claiming to be President Biden discouraged Democrats from voting in the 2024 primary. Barely a week later, a finance worker lost $25 million to scammers through a deepfake video call mimicking his colleagues. On X (formerly known as Twitter), meanwhile, AI-created explicit images falsely attributed to Taylor Swift attracted 45 million views and sparked widespread outrage. These incidents, all taking place in the first two months of 2024, are only a snapshot of the diverse and damaging impact deepfakes can have across politics, finance, and social media.

3.1.1 How Did They Become a Problem?

Fake images are nothing new. In 1917, pictures of intricate cardboard cutouts designed to look like fairies were published in *The Strand* magazine; many believed that these photos were sound evidence for the existance of supernatural forces.

Figure 10: One of the photographs of The Cottingley Fairies. Sir Arthur Conan Doyle, the creator of Sherlock Holmes, used these forgeries as evidence for psychic phenomena.

Creating fakes has become easier and cheaper over time, drastically improving the speed at which misinformation spreads. For example, during the 2004 U.S. presidential election, a manipulated photo falsely showed Democratic nominee John Kerry attending a protest with Jane Fonda, a controversial American activist. While the Cottingley Fairies had required an elaborate set-up involving cardboard cut-outs of traced drawings from a children's book, this forgery was a simple task accomplished using Photoshop.

Figure 11: This photo purported to show John Kerry sharing a stage with Jane Fonda at an anti-Vietnam War rally. It later emerged that it was a forgery, created by combining two pre-existing images in Photoshop.

However, the risk posed by fake photos has been reduced by our learning of how to identify signs of editing. In the case of "tourist guy", amateurs were able to identify that an image was photoshopped based on the observation that the white balance was inconsistent between objects in the scene. This is a product of the increased public awareness regarding disinformation; people have learned to notice signs of image editing. The term "photoshopped" has entered common parlance: signs of image tampering are universally recognised, and photographic proof is no longer regarded as inalienable evidence.

3.1.1.1 Deepfakes Make Fakes Easier, Cheaper and More Realistic

Forgeries used to be easily detectable by eye, but deepfakes make it easy and cheap to create images almost indistinguishable from real photos. For example, the website "OnlyFake" uses deepfake technology to generate realistic photos of fake IDs in minutes for just $15. The photos have been used to bypass the anti-fraud safeguards, known as Know-Your-Customer (KYC), on OKX (a crypto exchange). In the case of OKX, the deepfake IDs fooled their staff, who are trained to spot doctored images and deepfakes. This highlights how it is no longer possible to detect deepfake-based fraud by eye, even for professionals.

We've increased our dependence on video evidence in response to the risk of image deepfakes, but deepfakes will severely undermine video evidence very soon. A researcher at UT Dallas managed to bypass the identity verification features implemented by KYC providers using a free deepfake face-swap tool. This is a drastic development - it used to be expensive and time-consuming to generate videos with the necessary level

of realism. In 2019, it would cost someone two weeks and $552 to create a 38 second deepfake video of Mark Zuckerberg, featuring noticeable visual defects. Today, we can create realistic deepfake videos for free, in minutes.

Figure 12: The OnlyFake Panel for creating fake IDs in minutes.

3.1.1.2 Why Video Matters So Much

Videos used to be reliable evidence before the advent of deepfakes. Unlike easily faked images, videos have historically been challenging to counterfeit, leading to their unquestioned acceptance as reliable evidence in our courtrooms. This makes video deepfakes particularly dangerous.

At the same time, scepticism of real videos can also lead to dismissing genuine content, as shown when a video of President Biden was wrongly called a deepfake. Critics cited his unblinking eyes and lighting differences as proof, though these claims were debunked. This leads to a catch-22 – deepfakes not only make the fake seem real, but the real seem fake, further blurring the line between truth and fiction and making accountability more difficult.

> We still trust videos, even though they can now be faked.

Deepfakes enable targeted advertising at scale. We may soon witness a version of YouTube where what is said, who says it, and where it is said are all personalized to the viewer. An early example of this is a hyper-localised Zomato ad that showed actor Hrithik Roshan ordering dishes from popular restaurants in the viewer's local city. A different deepfake ad, featuring restaurants local to the viewer, was generated based on the viewer's GPS location.

3.1.2 What's Wrong With The Current Solutions?

3.1.2.1 Awareness

Deepfakes are now advanced enough to the fool experts we trained to spot them. This has enabled hackers to bypass identity verification (KYC/AML) procedures, even those that feature human moderators. This suggests that we can't tell deepfakes apart from real images using our eyes. We cannot culturally prepare for deepfakes merely by being image-skeptical: we need additional tools to tackle the deepfake epidemic.

3.1.2.2 Platforms

Social media platforms are not willing to implement effective deepfake mitigations without intense social pressure. For example, Meta agreed to ban deepfake videos involving fake audio, but refused to ban deepfake videos falsely depicting people's actions. Against the advice of their own oversight board, they did not remove a deepfake video showing President Biden touching his granddaughter, a forgery falling into the latter category.

3.1.2.3 Policy

We need laws that effectively tackle the emerging risks of deepfakes without restricting the less problematic uses, such as in art or education, which do not try to deceive people. The spread of non-consensual deepfake images, like those of Taylor Swift, has motivated legislators to pass tighter laws against such deepfakes. It may be necessary to legally enforce enhanced moderation procedures online in response to such cases, but proposals to ban all AI-generated content have raised alarms among filmmakers and digital artists, who worry it could unjustly restrict their work. Finding the right balance is key, otherwise, those legitimate creative applications will be stifled.

By pushing legislators to increase the barrier to entry for the training of powerful models, big tech companies can secure their AI monopoly. This may lead to the irreversible concentration of power in the hands of a few corporations - for example, Executive Order 14110 covering AI contains a proposal for stringent requirements on companies with a large amount of compute.

Figure 13: US Vice President Kamala Harris applauds as US President Joe Biden signs the US' first AI executive order.

3.1.2.4 Tech

Building guardrails directly into AI models to prevent misuse is our first line of defence, but these guardrails have been continuously broken. AI models are hard to censor because we don't know how to reliably modify their high-level behaviours using our low-level tools. Despite this, the firms that train AI models can use the implementation of guardrails as an excuse to introduce undesirable censorship and biases to their models. This is problematic because big tech AI firms are not accountable to the public will – firms are free to influence their models to the detriment of their users.

Even if the creation of powerful AIs wasn't centralised in the hands of dishonest firms, it might still be impossible to build an AI which is both guardrailed and unbiased. Researchers struggle to pin down what constitutes misuse, which makes it difficult to prevent it whilst handling the user's requests in a neutral, balanced manner. If we can't define misuse, it might seem necessary to make the guardrails less stringent, thus reopening the possibility of abuse. Entirely prohibiting the abuse of AI models is, therefore, impossible.

One solution is to detect malicious deepfakes once they're in the wild instead of preventing their creation. But, deepfake-detecting AI models (such as those deployed by OpenAI) are becoming obsolete due to inaccuracies. Although deepfake detection methods have become more sophisticated, the techniques for creating deepfakes are becoming more sophisticated at a faster rate – the deepfake detectors are losing the technological arms race. This makes it difficult to identify deepfakes based on the media alone. AI is advanced enough to create fake footage so realistic AI itself cannot determine its inaccuracy.

Watermarking technologies allow us to invisibly mark deepfakes such that we can identify them wherever they appear. But home-made deepfakes won't always feature watermarks, since watermarks must be added deliberately. They are, however, still effective for identifying deepfakes that were generated by parties who are willing to mark their images as deepfakes (such as OpenAI). Regardless, watermarks can often be removed or forged with easy-to-use tools, thereby bypassing any watermark-based anti-deepfake solutions. They can also be removed accidentally: most social media platforms automatically remove them.

The most popular implementation of deepfake watermarking technology is C2PA (by the Coalition for Content Provenance and Authenticity). It is designed to prevent misinformation by tracking where media comes from and storing this information in the media metadata. It is supported by companies such as Microsoft, Google, and Adobe, so it is highly probable that C2PA will be rolled out across the entire content supply chain, unlike its less popular counterparts.

Unfortunately, C2PA has its own weaknesses. Since C2PA stores a full history of edits to the image, authenticating each edit using cryptographic keys that are controlled by C2PA-compliant editing software, we have to be able to trust this editing software. However, it is likely that people will accept the edited images due to the presence of valid C2PA metadata without considering whether they trust every party involved in the chain of edits. Therefore, if any editing software is compromised or enables malicious edits, it may be possible to convince others that fake or maliciously edited images are real.

> We need secure hardware, otherwise the keys which secure C2PA will be stolen.

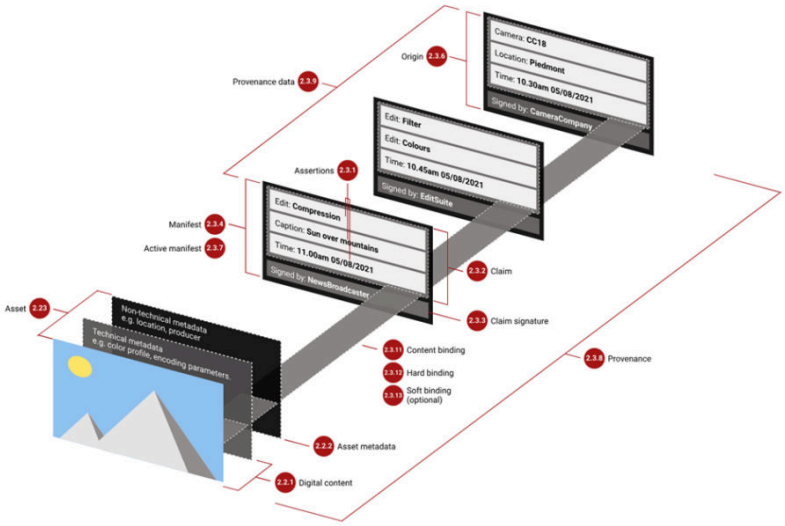

Figure 14: An example image with C2PA-compliant metadata containing a chain of edits. Each edit is signed by a different trusted party, but only the final edited image is public. Source: Real Photo vs AI-Generated Art: A New Standard (C2PA) Uses PKI to Show an Image's History

Further, the cryptographic signatures and metadata included within the C2PA watermark can be linked to specific users or devices. In some cases, the C2PA metadata could link all images that your camera has ever taken to each other: if we know that an image came from someone's camera, we can identify all other images which came from that camera. This could be used to deanonymize whistleblowers who have published images from their camera under their real names.

All potential solutions will face a unique set of challenges. However, these challenges vary drastically - including the limits of social awareness, the shortcomings of big tech firms, the difficulty of implementing regulatory policy, and our technological limitations.

> There is no silver bullet solution for tackling deepfake misinformation.

3.1.3 Can Crypto Solve This Problem?

Open-source deepfake models have already spread. Therefore, it is possible to argue that there will always be ways to abuse another person's likeness using deepfakes; even if the practice is criminalised, someone may choose to generate or platform unethical deepfake content anyway. However, we can almost entirely eliminate the problem by moving malicious deepfakes out of the mainstream. We can prevent people from thinking deepfakes are real images and enable people to create platforms where deepfakes are restricted. This section will describe a variety of crypto-based solutions to misinformation spread via malicious deepfakes, while highlighting the limitations of each approach.

3.1.3.1 Hardware Attestation

Hardware-attested cameras embed a unique proof with each photo they take, certifying that it was taken by that specific camera. This proof is created by a non-cloneable, tamper-proof chip unique to the camera, ensuring the image's authenticity. A similar process can be used for audio and video.

> It's cheaper to hack software, but expensive to compromise secure chips.

The attestation proof tells us that the image is taken by a real camera, meaning we can usually trust that it is a picture of a real object. We can flag images which don't have this proof. This fails if the camera takes an image of a fake scene designed to look like a real situation – you can just point the camera at a picture of a deepfake. We can currently tell if a picture was taken from a digital screen by checking for distortions in the captured images, but scammers will find ways of hiding these imperfections (for instance, by using better screens, or by limiting lens flare). Eventually, even AI tools will fail to identify this kind of fraud, since scammers can find ways to avoid all of these distortions.

Hardware attestation will minimise the number of deepfakes which are assumed to be real, but we still need extra tools to prevent the spread of deepfakes in those rare cases where a camera is compromised or otherwise misused. As we discussed earlier, it is still possible to create the impression that a deepfake is a real image using hardware-attested cameras, such as if the camera is hacked or is used to take pictures of deepfakes on computer screens. Additional tools will be needed to solve this problem, such as a camera blacklist.

Camera blacklists would enable social media platforms and apps to flag images which came from that particular camera which is known to have produced misleading images in the past. They can be implemented without requiring the camera to publicly reveal traceable data, such as the camera ID, with every published image. However, it's not

yet clear who will maintain the camera blacklist, or how we prevent them from taking bribes to blacklist cameras which belong to whistleblowers.[3]

3.1.3.2 Blockchain-based Image Chronology

Blockchains, being immutable, allow us to add images along with additional metadata to a timestamped chronology as they appear on the internet such that the timestamp and metadata cannot be tampered with. Since original, unedited images can be immutably stored on the blockchain by honest parties long before malicious edits proliferate, access to such a record would allow us to identify malicious edits and verify the original source. This technology was implemented on the Polygon blockchain network as part of Verify, a fact-checking tool built in collaboration with Fox News.

> Blockchain can prevent people from tampering with the chronology of images, but it can't prevent deceptive images from appearing in the first place.

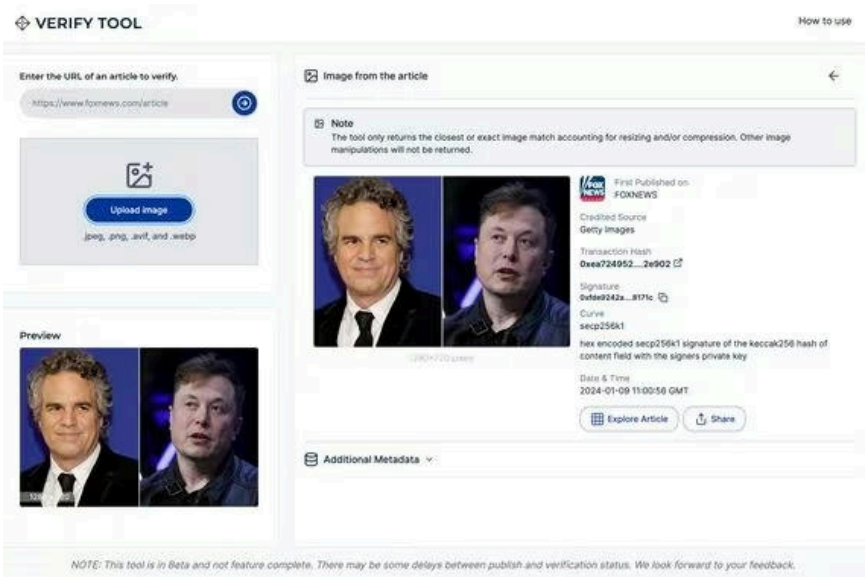

Figure 15: The user interface for Fox's blockchain-based tool, Verify. Articles can be looked up by their URL. The source, transaction hash, signature, timestamp and additional metadata are fetched from the Polygon blockchain and displayed.

3.1.3.3 Digital Identity

Trusted sources may become the only way of avoiding disinformation if deepfakes manage to destroy our trust in unverified images and videos. We already rely on trusted media sources to verify information because they employ journalistic standards, fact-checking processes, and editorial oversight to ensure the accuracy and credibility of the

[3]Reach out to Arbion if you are interested in the problems detailed in this paragraph.

content they publish. However, we will need a way of verifying whether what we see online comes from those sources that we trust. This is what cryptographic signature data is useful for: it can mathematically prove that one authored of a piece of content.

Signatures are generated using digital keys, which are only known by the person who owns the associated crypto wallet since wallets create and generate keys. This allows us to know who authored the data – you just check whether the signature corresponds to a key in one's crypto wallet, which belongs exclusively to them.

We can leverage crypto wallets to attach signatures to our posts, in a seamless and user-friendly way. If one uses a crypto wallet to log into a social media platform, they can take advantage of their wallet's ability to create and verify signatures on social media. Therefore, the platform will be able to warn us if a post is not from a trusted source – it will use automatic signature-checking to flag misinformation.

Additionally, zk-KYC infrastructure, which interfaces with wallets, can tie unknown wallets to identities verified by a KYC process without compromising user privacy and anonymity. However, as deepfakes become more sophisticated, the KYC process could be bypassed, allowing malicious actors to create fake, anonymous identities. This can be addressed using solutions such as Worldcoin's Proof of Personhood (PoP).

Proof of Personhood is the mechanism which WorldCoin uses to verify that their wallet belongs to a real human, and only permits one wallet per human. To achieve this, it uses the Orb, a biometric (iris) imaging device, to authenticate wallets. Since biometric data can not yet be faked, requiring social media accounts to be linked to a unique WorldCoin wallet is a viable way of preventing bad actors from producing multiple anonymous identities to conceal their unethical online conduct – it solves the deepfake KYC bypass problem, at least until hackers a way to fool biometric devices.

3.1.3.4 Economic Incentives

> *"(With Veracity Bonds), if you get caught with your hand in the cookie jar, you actually lose money, and, as a reader, I get to know how serious you are."*
>
> – Charles Hoskinson: Founder of Cardano; Co-Founder of Ethereum

Authors can be penalized for misinformation; users can be rewarded for identifying misinformation. For example, Veracity Bonds enables media outlets to stake money on the accuracy of their publications, facing financial penalties for misinformation. This offers these media companies a financial reason to be truthful.

Veracity Bonds will be an integral part of our "Truth Marketplace," where different systems compete to gain the trust of users by verifying the authenticity of content in the most efficient and robust manner. This is similar to a proof market, such as Succinct Network and =nil; Proof Market, but for the fuzzier problem of verifying truth

where cryptography alone isn't sufficient. Smart Contracts can be used as a means of enforcing the economic incentives required to make these truth markets work, thus blockchain technologies may play a central role in helping to fight disinformation.

3.1.3.5 Reputation Scores

> *"If we don't solve this reputation and authentication of content (problem), shit will get really weird."*
>
> – Illia Polosukhin: Co-Founder of NEAR; Co-Author of 'Attention Is All You Need'

We might use reputation as a proxy for credibility. For example, we might look at how many Twitter followers someone has to determine whether we should trust what they say. However, the reputation system could take into account each author's track record, not just their popularity. We don't want to confuse trustworthiness with their level of popularity.

We can't allow people to generate an unlimited number of anonymous identities, otherwise, they can ditch their identity whenever they tarnish their reputation to reset their level of social credibility. This will require us to use Digital Identities which can not be duplicated, as discussed in the last section.

We can also use evidence from Truth Markets and Hardware Attestation to determine ones reputation since these are robust ways of tracking their track record for factuality. Reputation Systems are the culmination of all the other solutions discussed thus far and are therefore the most robust and holistic family of approaches.

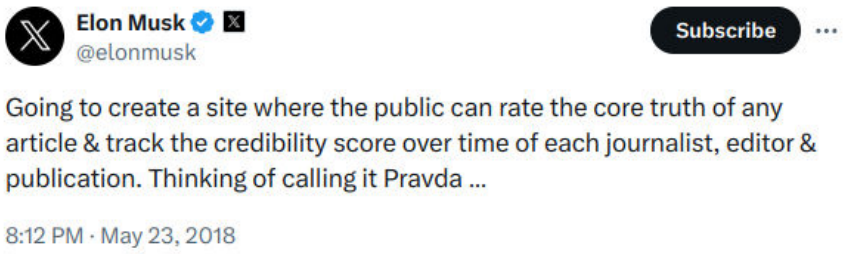

Figure 16: Elon in 2018 hinting at a website with credibility scores for journalists, editors and publications

3.1.4 Do Crypto Solutions Scale?

The aforementioned blockchain solution requires fast and high-storage blockchains – otherwise, we won't be able to fit all of our images into an on-chain verifiable chronological record. This will only get more important as the volume of online data published each day continues to grow exponentially. However, there are algorithms to compress data in a way where it can still be verified.

Additionally, the signatures produced through Hardware Attestation do not apply to edited versions of an image: zk-SNARKs must be used to produce a proof-of-edit. ZK Microphone is an implementation of proof-of-edit for audio.[4]

3.1.5 Deepfakes Aren't Inherently Bad

It's crucial to acknowledge that not all deepfakes are harmful. There are innocent uses of this technology, such as this AI-generated video of Taylor Swift teaching mathematics. More personalized experiences are also possible because of the low cost and accessibility of deepfakes. For example, HeyGen allows users to send personal messages with an AI-generated face that resembles their own. Deepfakes are also helping to bridge language gaps with dubbed translations.

> Deepfake technology is neutral. It can be used for good and bad.

3.1.5.1 Ways Of Controlling and Monetising Good Deepfakes

AI counterpart services based on deepfake technology are controlled by small teams that take high fees and lack accountability and oversight. Amouranth, a top earner on OnlyFans, recently released her own AI Influencer counterpart that fans can talk to in private. These services, often startups, can restrict or even shut down access, such as in the case of the AI companion service called Soulmate.

By hosting AI models on-chain, we can use smart-contracts to fund and control models in a transparent way. This will assure users that they will never lose access to the model, and can help model creators to distribute profits amongst their contributors and investors. However, there are technical challenges. The most popular technology for implementing on-chain models, zkML (used by Giza, Modulus Labs and EZKL), makes the model run 1000x slower. Nonetheless, the research in this subfield is still ongoing, and the techniques are improving. For example, HyperOracle is experimenting with opML and Aizel is building a solution based on Multi-Party-Computation (MPC) and Trusted Execution Environments (TEE).

3.1.6 Chapter Summary

- The emergence of sophisticated deepfakes is eroding trust across politics, finance, and social media, highlighting the need for a "Verifiable Web" to maintain truth and democratic integrity.
- Deepfakes, once an expensive and skill-intensive endeavor, have become easily producible with advancements in AI, changing the landscape of misinformation.

[4]Reach out to Arbion if you are interested in this problem.

- Historical context shows us that manipulation of media is not a new challenge, but AI has made creating convincing fakes far easier and cheaper, necessitating new solutions.
- Video deepfakes present a unique danger as they compromise what has traditionally been considered reliable evidence, leading to a societal catch-22 where real actions can be dismissed as fake.
- Existing countermeasures are categorized into awareness, platforms, policy, and tech approaches, each facing challenges in effectively combating deepfakes.
- Hardware attestation and blockchain offer promising solutions by proving the origins of each image and creating a transparent, immutable record of edits.
- Crypto wallets and zk-KYC enhance the verification and authentication of online content, while on-chain reputation systems and economic incentives like Veracity Bonds propose a marketplace for truth.
- Acknowledging the positive uses of deepfakes, crypto also presents a method for whitelisting beneficial deepfakes, balancing innovation with integrity.

3.2 The Bitter Lesson

> "The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin."
>
> – *The Bitter Lesson, Professor Rich Sutton*

This is a statement which is counterintuitive, but true. The AI community refuses to accept that bespoke methods don't work as well, yet the Bitter Lesson is still true: using the most computing power always produces the best results.

We must scale: more GPU's, more datacenters, more training data.

An example of researchers getting this wrong is when computer-chess researchers tried building chess engines using insights from the best human chess players. The first chess programs often copied human opening strategies (using 'opening books'). Researchers expected that the chess engine would start with a strong position, without having to calculate the best moves from scratch. They also contained many 'tactical heuristics' – tactics used by human chess players, such as *forks*. To put it simply: the chess program was built according to the human insights regarding how to play chess successfully, not general computational methods.

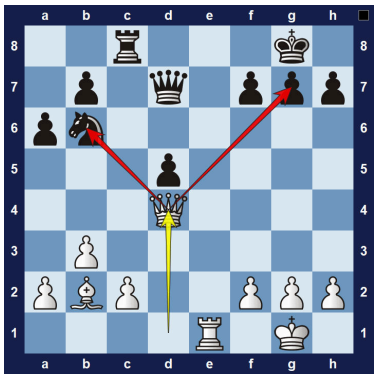

Figure 18: A Fork – the queen attacks two pieces.

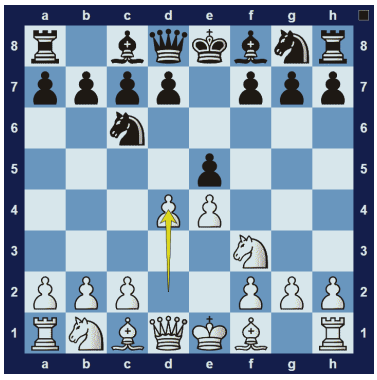

Figure 19: An example of a chess opening sequence.

In 1997, IBM DeepBlue, which combined enormous amounts of compute power with search-based technology, beat the reigning chess world champion. Even though it was superior to all 'human-engineered' chess engines, chess researchers shunned DeepBlue. They argued that DeepBlue's success was short-lived since it did not employ chess strategies – they saw it as a crude, brute-force solution. They were wrong: applying large swathes of compute to general problem-solving methods tends to produce better results than the bespoke approaches in the long-run. This high-compute ideology

spawned successful Go engines (AlphaGo), improved speech recognition, and more reliable computer vision.

The most recent triumph of the high-compute approach to AI is OpenAI's ChatGPT. Unlike previous attempts, OpenAI didn't try to encode human understanding of how languages work into the software. Instead, their model combined large swathes of data from the internet with an enormous amount of compute. They did not intervene nor embed any biases into the software, unlike other researchers. In the long-run, the best-performing methods are always based on *general methods* which take advantage of *large amounts of compute*. This is a historical observation; indeed, we probably have enough evidence to say it will *usually* be true.

The reason that combining enormous amounts of compute power with large swathes of data is the best approach in the long-run is Moore's law: the cost of compute will exponentially decrease over time. In the short-run, we may fail to ascertain a sizable increase in compute bandwidth, which can result in researchers trying to improve their technology by manually embedding human knowledge and algorithms into the software. This can work for some time, but it will not produce success in the long-run: embedding human knowledge into the underlying software makes the software more complicated and the models less capable of improving in response to extra compute power. This makes the manual approach short-sighted, and so Sutton recommends that we ignore manual techniques and focus on applying more compute power to general computational techniques.

The Bitter Lesson has serious implications for how we ought to build decentralized AI:

Building Large Networks: The above lesson highlights the urgency of developing larger AI models and assembling extensive computational resources to train them. These are key steps to advance into the new AI frontier. Companies like Akash, GPUNet and IoNet aim to provide that scalable infrastructure.

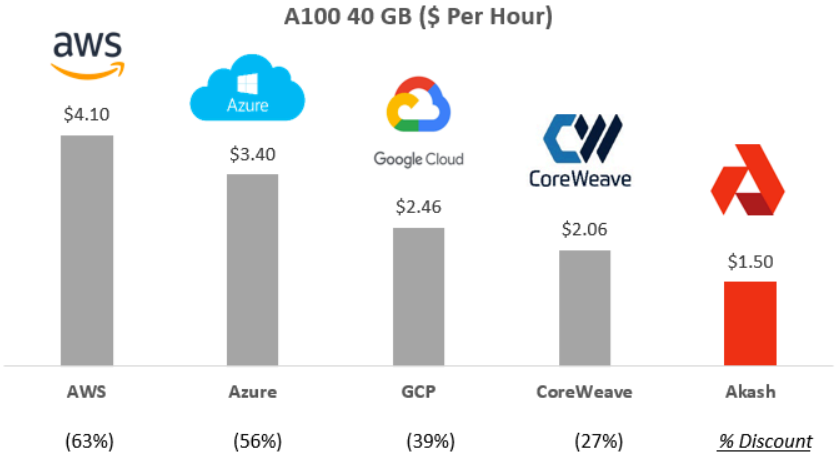

Figure 20: Akash price compared to other providers like Amazon AWS.

Innovate on Hardware: ZKML approaches have been criticized because they run 1000x slower than their non-ZK counterparts. This parallels the criticism faced by neural networks. In the 1990s, neural networks showed great promise. Yann LeCun's CNN model, a small neural network capable of classifying images of hand-written digits (see images below), found success. By 1998, banks were using this technology to read more than 10% of all checks in the US. However, these CNN models didn't scale, so interest in these neural networks decreased dramatically, and computer vision researchers started using human knowledge again to create better systems. In 2012, researchers developed a new CNN whilst leveraging the compute efficiency of GPUs, a popular piece of hardware usually used to generate computer graphics (gaming, CGI, etc). This allowed them to reach incredible performance, exceeding all the other methods available at the time. This network was called AlexNet, and it sparked the deep learning revolution.

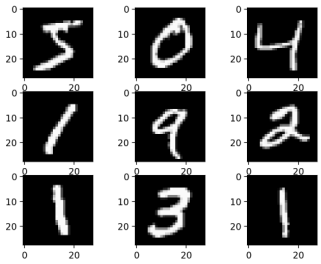

Figure 22: Neural networks in the 90s were limited to handle low-resolution images of digits.

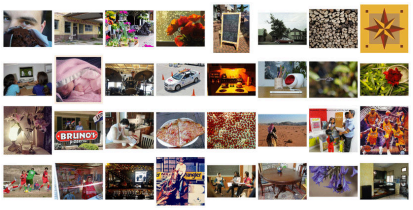

Figure 23: AlexNet (2012) was capable of handling complex images and outperformed all alternative methods.

The upscaling of AI technology is certain because compute is always getting cheaper. Custom hardware for technologies such as ZK and FHE will accelerate progress – companies such as Ingonyama, as well as academia, are paving the way. In the long-run, we will achieve large-scale ZKML by applying more compute power and introducing efficiency improvements. The only question is, what will we use these technologies for?

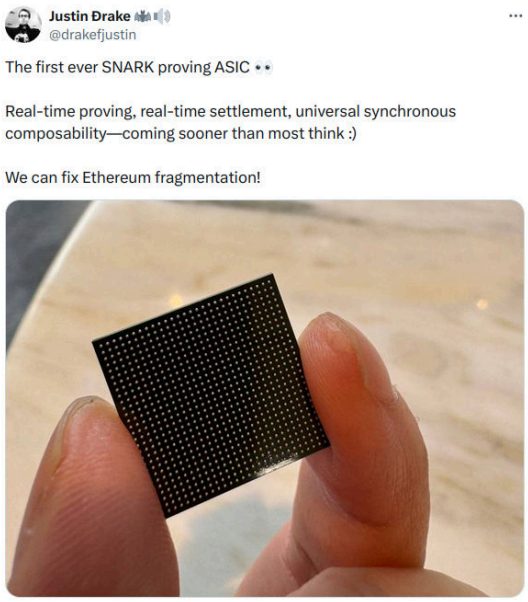

Figure 24: An example of hardware advancements for ZK provers. (Source)

Scale Data: As AI models grow in size and complexity, it is necessary to scale data sets accordingly. Generally, the size of the dataset should grow exponentially with the model size to prevent overfitting and ensure robust performance. For a model with billions of parameters, this often means curating datasets comprising billions of tokens or examples. Google's BERT model, for example, was trained on the entire English Wikipedia, containing over 2.5 billion words, and the BooksCorpus, with approximately 800 million words. Meta's LLama, meanwhile, was trained on 1.4 trillion tokens. These figures underscore the magnitude of the datasets we need – as models advance towards trillions of parameters, datasets must expand further. This scaling ensures that models can capture the nuance and diversity of human language, making the development of vast, high-quality datasets as critical as the architectural innovations in the models themselves. Companies like Giza, Bittensor, Bagel, and FractionAI are addressing specific needs in this space (see Chapter 5 for details on challenges in the data space such as model collapse, adversarial attacks and challenges in quality assurance).

Develop General Methods: In decentralized AI, the draw towards application-specific methods for technologies like ZKPs and FHE is driven by the quest for immediate efficiency. Tailoring solutions to particular architectures enhances performance but may sacrifice long-term flexibility and scalability, restricting broader system evolution. Conversely, focusing on general-purpose methods offers a foundation that, despite initial

inefficiencies, is scalable and adaptable to diverse applications and future developments. These methods are set to excel as computational power grows and costs decrease, thanks to trends like Moore's Law. The choice between short-term efficiency and long-term adaptability is crucial. Emphasizing general methods prepares decentralized AI for a future of robust, flexible systems that leverage computational advances, ensuring enduring success and relevance.

3.2.1 Conclusions

In the early stages of a product's development, it might be crucial to pick approaches that don't benefit from scale. This is relevant for both companies and researchers validating use cases and ideas. However, the bitter lesson suggests that we should always remember to prioritize general scalable methods in the long-run.

Here's an example of a manual approach being replaced with automatic, general solution: before automatic differentiation (autodiff) libraries, such as TensorFlow and PyTorch, gradients were often computed manually or using numerical differentiation — this was inefficient, prone to error, and caused problems which wasted the researchers time, unlike autodiff. Autodiff is now indispensable, because autodiff libraries have enabled faster experimentation and simplified the development of models. Thus, the general solution won – but the old, manual methods were necessary to enable ML research before autodiff became a mature and usable solution.

In summary, Rich Sutton's Bitter Lesson teaches us that our AI's will improve more rapidly if we maximise the computing power we put into them, rather than trying to make them mimic well-understood human approaches. We must scale existing compute, scale data, innovate on hardware and develop general methods – the adoption of this approach will have numerous implications for the decentralized AI space. Although the Bitter Lesson does not apply in the earliest stages of research, it will probably usually be true in the long-run. Thi

3.3 AI Agents Will Disrupt Google and Amazon

3.3.1 The Google Monopoly

Online content creators generally depend on Google to distribute their content. In return for allowing Google to index and display their work, they are provided with a consistent flow of attention and revenue from advertisements. However, this relationship is imbalanced; Google holds a monopoly (over 80% of search engine traffic), a market share that content creators on their own could never hope to match. As a result, content creators are heavily reliant on Google, as well as other tech giants, for their revenue. A single decision by Google has the potential to end an individual's business.

Google's introduction of featured snippets - which display answers to user queries without requiring clicking through to the original site - highlights this problem, as information can now be obtained without ever leaving the search engine. This disrupts the arrangement content creators had come to rely on. In exchange for Google indexing their content, creators expected referral traffic and eyeballs on their sites. Instead, featured snippets allow Google to summarise content while cutting creators out of the traffic flow. The fragmented nature of content producers leaves them largely powerless to take collective action opposing Google's decisions; with no unified voice, individual sites lack any bargaining power.

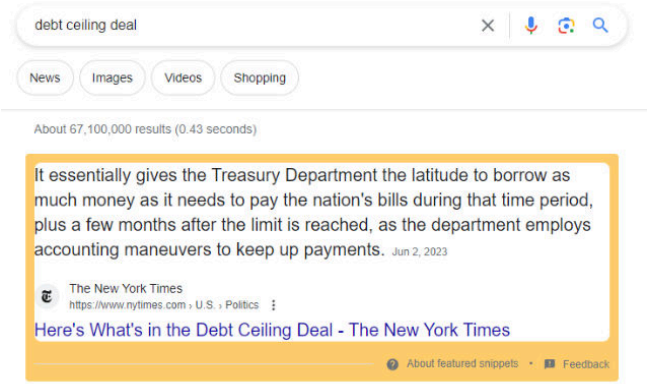

Figure 25: Example of a simple featured snippet.

Google experimented further by providing a list of sources for the answers to the user's query. The example below contains sources from the NYT, Wikipedia, MLB.com and more. These sites would not be getting as much traffic, since Google provides the answer directly.

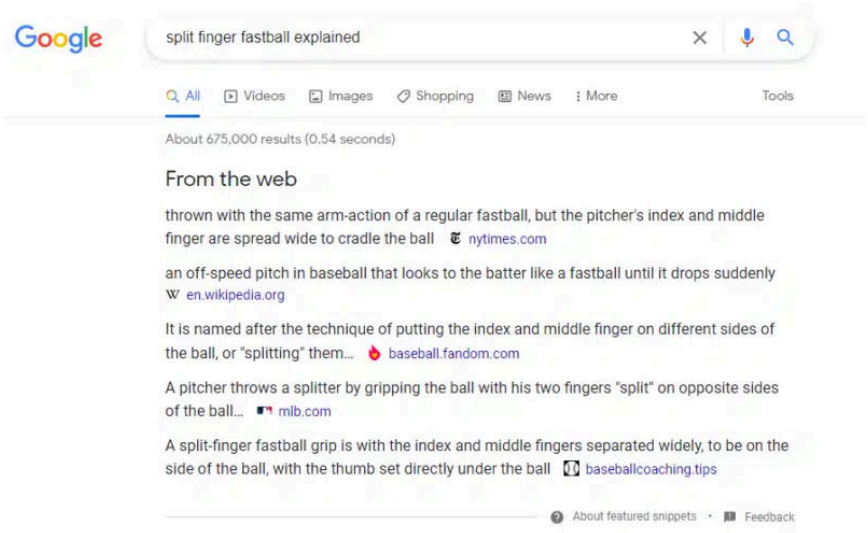

Figure 26: Example of a featured snippet with 'From the Web' functionality.

3.3.2 The OpenAI Monopoly

Google's introduction of featured snippets represented the beginning of a concerning trend - reducing the attribution of original content creators. ChatGPT has taken this concept to its logical extension, functioning as an omniscient information agent without any links or references back to the source material.

Language models like ChatGPT can answer virtually any question by summarising content scraped from the internet, but they fail to direct users to visit the original publishers. Instead, the model accumulates knowledge derived from copyrighted works into a singular interface entirely controlled by OpenAI.

The success of these models relies on the vast amounts of data that make up the internet, and content creators are not rewarded at all for their essential input to model training. Some larger publishers have managed to reach agreements with the likes of OpenAI, but such a solution is impossible for smaller content creators. Others have decided to simply block AI models from trawling their content, but with close sourced models there is no guarantee of this being respected.

AI companies try to justify their lack of compensation by arguing that the AI system is simply learning from the content - like a human might learn about the world from reading - but does not infringe upon content creators when producing an output. However, this assertion is questionable, as ChatGPT can regurgitate entire *New York Times* articles verbatim. Both Midjourney and DALL-E are similarly capable of generating copyrighted content.

The implications are clear - big tech continues to consolidate power, while unaffiliated content creators see their leverage diminish. The asymmetric relationship that launched lawsuits against Google has only grown more extreme. Established publishers like the

New York Times have already taken legal action, as have a range of content creators from digital artists to coders, in the form of class action lawsuits.

One proposed solution is Retrieval Augmented Generation (RAG), which allows the language model to provide sources for its answers. However, this shares issues with featured snippets - it doesn't give the user any reason to visit the original site at all. Worse, tech giants could leverage RAG disingenuously as legal cover while still denying creators traffic and ad revenues.

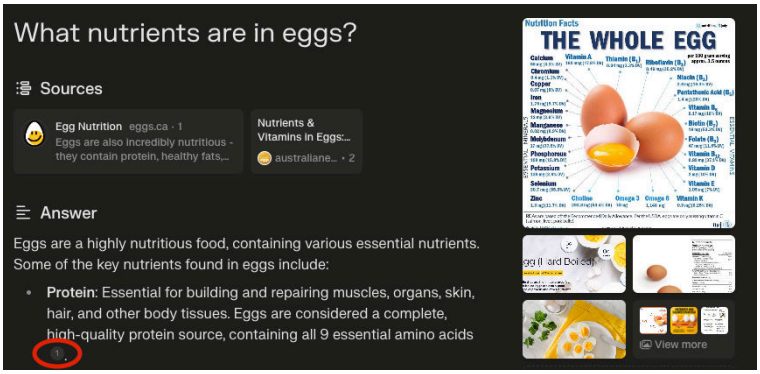

Figure 27: RAG includes links as sources. This result was generated by Perplexity AI.

3.3.3 Crypto's Potential Solution

As "information agents" like ChatGPT emerge, big tech seems poised to repeat the past and monopolise generative AI, despite their reliance on unaccredited creator content. But with AI now disrupting the market, there is an opportunity to redistribute power and build a fairer system for compensating creators. In the early 2000s, policymakers missed the chance to establish an equitable model, resulting in today's distribution system - a system monopolised by Google. The current rise of AI presents a crossroads — will we address past mistakes, or will history repeat itself with OpenAI and other tech giants gaining unilateral control?

To incentivise quality content production over the long term, we must explore ways to continue providing fair payment to creators. As Chris Dixon argues, crypto offers a solution with the blockchain by serving as a collective bargaining machine to solve large-scale economic coordination problems like this, especially in the current case of asymmetric power. Their governance could be jointly managed by creators and AI providers.

In the context of AI, creators can leverage smart contracts to encode software-enforced usage terms and restrictions - for instance, by setting conditions around commercial applications like model training. Smart contracts would then automatically enforce attribution systems, allocating portions of revenue generated by AI systems back to contributing creators. Even if AI companies wanted to compensate creators at present,

the sheer number of individual creators would make this virtually impossible - but smart contracts make compensation feasible. The composability of smart contracts would also reduce lock-in to any single model, enabling a freer AI marketplace. This competition would lead to more favourable profit-sharing for creators. Facing unified terms enforced by an impartial protocol, AI companies would need to either accept the collective bargain set by creators or go without their content; no longer could tech giants exert unilateral leverage over individuals.

Centralised control over information agents like ChatGPT also raises concerns over issues such as embedded advertising. While Google clearly demarcates ads separately at the top of search results, an AI agent could seamlessly integrate paid recommendations into responses. In contrast, crypto-based solutions allow for auditable AI agents.

3.3.4 AI Agents & Amazon

The natural extension from language models like ChatGPT answering queries is AI programmes that can take actions on the user's behalf, a move from information agents to action agents (simply referred to as 'agents'). These systems wouldn't just find you the best Bluetooth speaker, but order it directly to your house. Relying on closed-source agents from companies like OpenAI for these tasks will give them immense power beyond just the content creation market, potentially allowing them to dominate industries like the $6.3 trillion e-commerce market. OpenAI would become not just the next Google, but the next Amazon as well.

If a handful of big tech companies own the most powerful and widely used AI agents, it gives them huge leverage over consumers and industries. These agents would mediate a growing portion of our digital lives - shopping, travel, and finances. With no alternatives, there would be little choice but to rely on companies like OpenAI or Google. Their closed-source agents would become virtual gatekeepers, controlling our access to crucial services and information, and they could modify agent behaviours with little accountability.

This concentration of power in proprietary AI has parallels to the rise of big tech monopolies like Google and Facebook. But, the implications are exponentially greater when AI agents can take action seamlessly across domains. That is why decentralised blockchain alternatives are so vital - they introduce competition, user empowerment, and transparency, acting against the risk of AI agent monopolisation in the hands of big tech.

In summary, AI models like ChatGPT provide us with information agents, which can read content and answer questions on our behalf. This will disrupt not only the way we consume information but also Google's business model and the livelihood of creators. With this disruption in front of us, we have an opportunity to build a new internet that is fairer to creators and rewards them for their work and contributions. AI agents

capable of taking action (for instance, the purchasing of goods online) will disrupt e-commerce. Again, blockchain gives a chance to establish an equitable model for the internet. Will we learn from the mistakes of the Google and Amazon era?

3.4 Crypto Will Accelerate Open Source AI

Note: This is a draft.

This section will be updated in Version 2 of this report.

3.4.1 What AGI Means For Us

AI has become a staple of the 21st century. It has uses ranging from art to education, to finance, to politics, to programming and many more applications. It is able to generate photorealistic videos and images in seconds with rather limited prompts. Programmers have resorted to using AI to develop efficient and useful code, outsourcing their own labour to AI. The line between the real world and science fiction is blurring – and this leaves a reduced role for humans in the labour market.

Therefore, it is likely that we will face a crisis in labour. If we can outsource a significant level of our intellectual labour to AI, it may soon become economically preferable to use AI rather than human labour. If human labour is being replaced by AI, we will need to adapt to this new state of the labour market. Such adaptations will result in a disruption in the way our current economic systems work. This section will deal with the direction AI is heading and ways to prevent such crises from occurring through Crypto solutions.

3.4.2 Why We Should Care About Open Source AI?

Generally, an open source software is software that is open for anyone to use, usually attached with a license that states what can and cannot be done with the software. Open source AI is for AI software to be open for anyone to use, with similar limitations provided by licenses. Open source projects are typically organised with free participation in the development of the product in mind. It is centered around a community which welcomes contributions of codes and error fixing, they tend to be collaborative in nature. This section will provide reasons for why open source AI is significant.

Open source AI can be used to create a more competitive AI market. Competition is good because there are a variety of people with different talents and varying levels of skill at developing the product. Suppose that there is a firm which generally has a lot of talent, but has made a massive mistake, a competitive market would encourage people to fix that error and so limit the damage of that error. Open source AI lowers the barrier for people to enter into the AI market, anyone can use the AI software and contribute. Since anyone can enter through open source AI, it means that there are more competitors in the market making the industry a much more competitive one than it is at present.

Open source AI can be used to safeguard the industry against bad actors seeking to do harm. Technology is a powerful tool but itself is neutral. It can be used for the purposes of benefitting or harming humanity, depending on whose hands the technology is gripped by. It is preferable that we have good actors wielding this technology, especially to counter those that are using technology to harm. Open source AI makes it so that more people and talent enter the AI industry. This gives us a better chance of good actors entering the AI industry who work to counter those that wish to use AI for harm. A reduction in the bad actors using AI will lead us to a safer world.

Furthermore, there is a lowering of the barrier of entry to the AI industry, unlocking a bank of talent and skill which only serves to develop the industry further. AI can provide progress for us as a species. We have already outlined its wide usage today, but the potential for development is immense, this is great for our progress as a species. Open source AI will accelerate the development of AI through lowering the barrier of entry for talent to enter into the industry. More talent means that we can make better AI for even wider applications in society. Open source AI provides us the tool to access a wellspring of talent and ability to develop this technology.

Open source AI provides developers with the freedom to specify and customise AI as they see fit. One feature of AI is that it can be customised for specific purposes. A customised AI can be directed to deal with specific needs and demands which improves the quality of the product significantly. Developers' ability to customise AI software is often bound by the closed nature of the AI industry at present. What open source AI can provide is that it allows the developer to freely customise and make their product the best it can be for their purpose. Such freedom will make for a better market and for better products to be brought out from it.

To conclude, we find that the Techno-Capital machine is one of perpetual growth. A virtuous cycle is formed where development of technology is fuelled by capital gained from the market and the market creates more capital through the development of technology. Open source AI provides a space for competition, low barriers to entry, freedom and collaboration which only serves to increase innovation and diffuse AI technologies stimulating economic growth and more markets. It is essential for both the industry's ability to grow and also for its benefit that it can provide to humanity.

3.4.3 OpenAI versus Open Source AI

In the AI industry, the leading company is OpenAI. Since releasing ChatGPT in 2022, OpenAI has led the AI industry in profit and in knowledge. They have the backing of Microsoft and big tech, and thus are well entrenched in the AI market. From this, it seems that it is quite an uphill battle to have open source AI compete on the level of OpenAI from our present perspective. Yet, we have good reasons to suggest that open source AI has the promise to challenge and beat OpenAI in the market.

Open source AI is able to evade the inevitable regulations that will be placed upon the AI industry by fearful governments. Presently, governments and regulatory bodies around the world are scrambling to regulate and limit the emerging AI industry. The target for this regulation will be the traditional AI that is used by OpenAI which is centralised and kept by a single organisation. This regulation will limit the development of the industry. However, open source AI has the advantage of being without a keeper and being decentralised. This means that governments will find it difficult to regulate open source AI. This provides open source AI the advantage of not being as limited by regulations as OpenAI will be in the future.

Furthermore, we see that open source AI can benefit from OpenAI, whereas OpenAI cannot benefit from open source AI. OpenAI's main objective is to maximise profit for themselves, this means they keep a lot of their models and data private to prevent the competition from using it to their advantage. Open source AI can use licenses to prevent OpenAI from benefitting from their less regulated developments. OpenAI will find themselves isolated in the market, since they will be just one company against a host of many different organisations which use open source AI. This will mean that open source can benefit from OpenAI's wealth of data and knowledge through limiting OpenAI's access initially. A new license may be needed, however, to eventually provide OpenAI this access.

Finally, open source AI has a better chance of attracting ideologically minded people with a passion for the betterment of AI than OpenAI. This is because open source AI has such a low barrier of entry. However, one might suggest that this would result in an difficult atmosphere to coordinate and develop products. Yet, it is not necessary to have a lot of people on such projects. Linus Torvalds at Linux is such an example where a single person contributing continuously to a project can have such a massive impact. OpenAI's high barrier of entry makes it difficult for passionate people to join.

Despite it seeming right now that OpenAI has a tight grip over the AI industry, there are several routes that open source AI can rely upon to challenge OpenAI's grip. Open sources AI's flexibility, accessibility and community-centred over profit-centred approach to AI means that it has great tools to beat OpenAI through isolating them in a consistently evolving and dynamic market. Certainly, open source AI has the potential to beat OpenAI in the AI industry.

3.4.4 The Problem with Open Source AI.

However, there are still obstacles open source AI must overcome if it is able to take the fight to big tech and OpenAI. These obstacles can be divided into three themes. First, there is the lack of talent present in open source AI. Second, there is the lack of compute present for open source AI to physically work. Third, there is a lack of data for AI to train themselves upon and develop. This section will be dedicated to elaborating upon these three problems.

For any organisation to function, there needs to be talent to provide the innovative ideas and work necessary for the development of a product. Open source AI faces a major problem in that there is no profit or monetary incentive to work in these communities. The majority of AI engineers, when deciding on how they want to work in the industry are forced to choose between either a high-paying job at Big tech or run the risk of entrepreneurship. The safe option, which is what most people take, is to work at big tech and earn a living there. The best talent go to OpenAI rather than work in an open source AI community with no monetary incentive. Therefore, open source AI is unable to attract that best talent necessary for it to develop innovative new products which can challenge big tech.

Another problem is that open source AI lacks access to the computing power necessary for the scale at which OpenAI has access to. Larger scale AI requires more GPUs to scale up the operation. GPUs are expensive, effectively only produced by a single company in Nvidia. The lack of money in open source AI mean that it is very difficult to physically generate enough compute for AI models to compete on the same level as OpenAI's ChatGPT. Even Linux, which has high quality standards in software engineering is limited by the compute scale for their programs. Open source AI lacks the ability to access a supercomputer at will, and so it becomes difficult to compete with OpenAI which does have that privilege.

AI models require data for them to train. Despite labelling themselves as 'open' or 'open source', big tech such as Meta and OpenAI keep the data that they use to train their AI as private and kept only to them. They only publish the finished AI models afterwards, such as ChatGPT. This data tends to be of high quality and high quantity too, sourced from the mass userbase of both Meta and OpenAI. Open source AI has the disadvantage of not having access to this high quality data in high quantity and therefore they cannot train their AI models with the best and most amount of data necessary for them to compete with Meta or OpenAI. Therefore, open source AI is not in a state to develop products which can stand against OpenAI or Meta's products.

What open source AI requires is a way to overcome these three major problems which hinder their potential at challenging big tech's grip over the AI inudstry. We suggest that Crypto has a solution which can tackle these issues for open source AI.

3.4.5 Crypto's Solution for Open Source AI.

We believe that Crypto has solutions for all three of the problems outlined in the previous section. This section will be dedicated to proposing these solutions to each specific problem.

First, Crypto can solve the talent problem through creating a revenue/reward system in open source AI. Crypto can help open source AI through providing intrinsic rewards for contributions to the project. An example of this is the AGI Guild, which is a group

of ideologically open source developers who have formed a license which can reward developers for contributing to open source projects.

This works through having the license be governed by an AGI token. Companies which are over 1 million USD must acquire a stake in the AGI token to use the license. This stake is then distributed to contributors to the open source AI project. This provides the monetary reward for the developer and gives the token value itself, encouraging more contributions. There are voting systems as well for members to collectively decide on how the project should go, encouraging meritocracy and democracy in open source AI whilst garnering the ability to make an earning from open source AI.

Since there is now a monetary incentive, prospective and present talent will not need to view their path as either safely choosing Meta or risking all on a venture but seeing a third way where you can earn as you develop AI and not have to work with the proprietary restrictions of a big tech company. Talent will be attracted to open source AI projects as a viable competitor against big tech.

Second, Crypto can solve the compute problem by lowering the barrier for access to servers and powers. Cloud providers could reject permission for developers to use their servers for their products in the past. With crypto, it now becomes permissionless which means that anyone can get the compute they need to use for their project. This means that developers, using open source AI and Crypto, can now freely use as much compute as they need.

Crypto also provides developers and communities with the bargaining power aginst the cloud providers. In the short past, cloud providers could increase their prices for usage since the open source AI developers needed their servers to run their programs. Now, with Crypto, we can decentralise this systen and begin to challenge the cloud providers' prices and open the doors to many more who want to develop open source AI.

Now, communities have the means and capabilities of challenging the cloud providing companies like AWS to decrease their prices and therefore increase the quality of the product and the work cone in the end.

Finally, Crypto can help solve the data problem by providing rewards for users who generate data for an open source AI project. An example of this is Grass, which has a browser extension which generates tokens as someone is browsing the internet. This reward system means that open source AI is able to gain more and more data as time goes on, as Crypto reveals more data. This is done in a decentralised way which means that we don't fall into the problem of regulators or big tech figuring this out and taking measures to prevent it like banning your IP address. This can't be effective against this way because there is a host of people with this extension, banning one person has little effect on the rest being able to scrape that data. This allows for open source AI developers to get as good data as the big companies are able to attain. This will certainly improve the training of the AI model. Even for data which is meant to be

kept secret, Crypto can both keep the data secret and scrape it too given that the user to whom the data belong is alright with it, it is able to be used for the betterment of the AI model.

Overall, what we find is that Crypto is able to scrape and attend to the sensitivities of data collection whilst benefitting the open source AI programs significantly through giving them the data they need to train their AI models.

The ideal that we are seeking, where an open source AI system has beaten the private propriety software based system is this. On the supplier side, open source AI will lead to the best AI being present in the marketplace since there will be the most diverse and best engineers working on developing AI with the best data. This continues to push forward the AI industry and progresses it even further.

On the consumer side, we will see that consumers can choose from a wide variety of specific and customised AI models which then allows for consumers to form an AI for their needs rather than for the company's needs. This customisation effectively brings to the consumer the ability to train an AI themselves for their own purposes.

Opportunities for Builders

Note: This is a draft.

This section will be updated in Version 2 of this report.

While writing this report, we received many suggestions from the community of potential ideas for builders in the AI x Crypto space. We've compiled a curated list of suggested ideas we're excited about to get you started. We present each **problem** and a **potential solution**. If you're interested in any of these ideas, get in touch.

4.1 A Curated List of Suggested Ideas

4.1.1 AI DAO

Category: AI helping Crypto internally -> AI helping

Problem: DAOs rely on an engaged community to vote on proposals thoughtfully. Every proposal needs to be voted on manually which slows things down and leads to DAO fatigue.

Potential Solution: Automate proposals by giving members of a DAO an AI Agent that votes on their behalf based on their values. The model could use a Bayesian approach which takes into account the uncertainty in the decision being made. If the uncertainty is above a threshold, then bring the proposal to the attention of the members of the DAO and require a manual vote.

The onboarding process for creating these AI agents needs to be streamlined for users to adopt this. A combination of on-chain and off-chain data from each user can be used to define the preferences of a user without too much manual work. For example, DAO Base is looking into solutions around the idea that what you do in the past defines you.

One potential approach is to query an LLM on-chain, passing in the proposal along with the DAO member's prompt so that they can adjust according to their requirements.

```
pragma solidity ^0.8.24;

contract Example {
  function decide(string memory proposal) public pure returns (Choice, string memory) {
      string memory result = llmInference(string.concat(prompt, proposal));

      string memory v = substring(result, 0, 1);
      string memory reason = substring(result, 2, bytes(result).length);

      if (keccak256(abi.encodePacked((v))) == keccak256(abi.encodePacked(("Y")))) {
          return (Choice.For, reason);
      } else if (keccak256(abi.encodePacked((v))) == keccak256(abi.encodePacked(("N")))) {
          return (Choice.Against, reason);
      } else if (keccak256(abi.encodePacked((v))) == keccak256(abi.encodePacked(("A")))) {
          return (Choice.Abstain, reason);
      }

      revert(string.concat("Unknown vote: ", v, " with reason: ", reason));
  }
}
```

,

Listing 1: This is a simple example but tooling would be needed to make it easy for users to specify their requirements and automate the voting itself.

See here for demo code.

4.1.2 Verifiable Training

Problem: It's not always possible to tell what data a model was trained on, even if you have access to the model weights, because training models is a compression of training data. This introduces several challenges that do not exist in traditional software:

- Complicates copyright
- Harder to fairly compensate the owners of the data
- Harder to know who trained each part of the model in settings where training is done by more than one party, e.g. Model Zoo
- Easier to add biases in models

Potential Solution: Make the training process itself verifiable. Build tools to break down how a model was trained and check if it contains a given piece of data. Several approaches can be explored:

Integrate cryptographic primitives into the training process itself. For example, Pytorch NFT Callback hashes the current network weights, some metadata (data, accuracy, etc…) and your eth address every N epoch, which proves who did the model training. Note: This approach introduces a performance overhead to training models.

An alternative solution would be to train the model on a purpose-built decentralised network. There are solutions based on traditional consensus mechanisms such as BFT. However, BFT requires more than $\frac{2}{3}$ of nodes to be reliable and honest, leading to the

result that the minimum number of nodes for a BFT consensus is $N = 3f + 1$ where f gives the number of nodes that either (i) fails or (ii) maliciously acts. This introduces a lot of redundant work (linear in N e.g. 60x for $N = 60$). An example of this approach is Proof of Training. We suggest an alternative decentralised network that is built from the ground up for this purpose.

4.1.3 Alternative Paths to Verifiable Inference

Problem: Most of the research in verifiable ML is on Zero-Knowledge Machine Learning (zkml). However, zkml currently has a performance overhead of 1000x and cannot run large models yet.

Potential Solution: Several approaches are being explored. This is a relatively new field and there is an opportunity for different approaches that make different tradeoffs.

Ora is experimenting with opML. This approach involves a single party 'optimistically' inferencing a model, putting the result on-chain, and incentivising verifiers to challenge incorrect results by paying them tokens.

Aizel is building a solution based on Multi-Party-Computation (MPC) and Trusted Execution Environments (TEE). Their aim is to do verifiable inference at the same cost as normal inference.

EZKL is working on splitting and parallelising zk proofs, which would make proving large models feasible.

4.1.4 DePin 2.0

Problem: As DePin (Decentralized Personal Internet) technologies and robotics intersect, the unique challenges of integrating dynamic, autonomous systems like robot swarms become evident. Unlike static, sensing devices such as solar panels, robot swarms possess the ability to act independently, which introduces vulnerabilities. Specifically, a single malfunctioning or malicious (Byzantine) robot within a swarm can compromise the integrity of the entire system. Given the irreversible nature of actions taken by robots, ensuring the reliability and security of these swarms is paramount.

Potential Solution: A potential solution involves leveraging zero-knowledge proofs (ZKPs) to certify the execution of specific swarm strategies without revealing the strategy's details. This cryptographic technique can validate that a swarm operated according to a pre-defined algorithm, based on hardware-attested data collected from the swarm and cryptographically signed information from external sensors. By incorporating ZKPs, we can establish a trustless environment where the behavior of robot swarms is verified in real-time, mitigating the risk posed by Byzantine robots. This approach not only enhances the security and reliability of robot swarms but also aligns with the decentralized ethos of DePin, ensuring transparency and trust without compromising sensitive operational details.

4.1.5 Transparent LLM Drift

Problem: Large Language Models (LLMs) are central to many applications, yet they suffer from unpredictability and performance drift over time. Traditional benchmarks, like the Open LLM Leaderboard by HuggingFace, are criticized for being gameable and lacking historical data, making it difficult to track and understand LLM performance changes.

Potential Solution: Integrate automated benchmarks with community voting, using tokens as incentives for participation. Results are stored on a blockchain, ensuring transparency and immutability. This approach aims to provide a reliable, transparent history of LLM performance, addressing issues of unpredictability and performance drift.

4.1.6 Truth Marketplace

Problem: As AI-generated content becomes more prevalent, distinguishing between true and false information online grows more complex. This ambiguity challenges the traditional verification methods, making it hard to maintain trust in digital content.

Potential Solution: An idea to consider is extending the concept of a proof marketplace to address the verification of truth in a more nuanced, fuzzy environment. By introducing economic incentives, such as those seen in existing proof marketplaces, but tailored to handle ambiguity, this system could encourage the discovery and validation of truth. Bittensor's Yuma Consensus would be the ideal venue for a Truth Marketplace as deals well with ambiguity. This approach wouldn't prescribe a specific method but would create a framework where various strategies for identifying truth are economically incentivized, fostering an ecosystem where truth has a tangible value.

4.1.7 Trustless Prediction Markets Resolved by AI

Problem: The challenge with prediction markets is that participants often disagree on the meaning of the events being bet on, including what counts as evidence for winning or losing, leading to confusion and disputes.

Potential Solution: Resolve prediction markets by feeding information on a pre-agreed date into a pre-agreed AI model. The information comes from a pre-agreed data source (which could leverage the prior idea of a Truth marketplace).

How Does Machine Learning Work?

Before diving into the intersection of artificial intelligence (AI) and crypto, it's first crucial to cover some concepts in the field of AI separately. Since this report is written for readers in the crypto space, a deep understanding of concepts in AI and machine learning isn't always present. This is crucial so that the reader can assess which ideas at the intersection of AI and crypto have substance, and to accurately asses the tech risk of projects. This section focuses on AI concepts; however, hints for how this relates to crypto are highlighted throughout.

> Hints throughout this chapter give an idea as to how AI concepts relate to crypto.

Summary of topics covered in this section:

- Machine learning (ML) is a branch of artificial intelligence where a machine is taught to make decisions from data without being explicitly programmed.
- The ML pipeline is divided into three steps: data, training and inference.
- Training a model is very expensive computationally, while inference is relatively cheap.
- There are three main types of learning: supervised learning, unsupervised learning and reinforcement learning.
- Supervised learning refers to learning from examples (provided by a teacher). A model can be shown pictures of dogs and told by the teacher that these are dogs. Then the model learns to distinguish dogs from other animals.
- Many popular models, such as LLMs (e.g. GPT-4 and LLaMa) are trained using unsupervised learning. In this mode of learning, no guidance or examples are provided by a teacher. Rather, the model learns to find patterns in the data.
- Reinforcement learning (learning from trial and error), is mainly used in sequential decision-making tasks such as robotic control and playing games (e.g. chess or Go).

1.1 Artificial Intelligence and Machine learning

In 1956 some of the brightest minds of the time came together for a workshop. They aimed to come up with general principles of intelligence. They noted that:

"Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it."

In these early days of AI, researchers were full of optimism. In some sense, they were aiming for Artificial General Intelligence (AGI), quite ambitious. As we now know, these researchers didn't manage to create AI agents with general intelligence. Neither did AI researchers in the 70′s and 80′s. In that period AI researchers attempted to develop 'Knowledge-based systems'.

The key idea in knowledge-based systems is that we can write very precise rules for a machine to follow. Essentially, we extract very specific and precise domain knowledge from an expert and write it down for a machine in the form of rules. The machine can then use these rules to reason and make sound decisions. For instance, we can attempt to distil all principles of playing chess from Magnus Carlson and build an AI to play chess.

However, this is extremely hard to do and even if it's possible, a tremendous amount of manual effort is required to create those rules. Imagine trying to write down rules for how to recognise a dog to a machine. How can the machine go from having pixels to knowing what a dog is?

Recent progress in Artificial Intelligence came from a branch referred to as Machine Learning. In this paradigm, rather than writing precise rules for a machine, we use data and allow the machine to learn from it. Modern AI tools using machine learning can be found everywhere, for example, GPT-4, FaceID on your iPhone, Gaming bots, Gmail spam filter, models for medical diagnosis, self-driving cars … etc.

1.2 The Machine Learning Pipeline

The machine learning pipeline can be divided into three main steps. You have data, we train the model and then when we have the model, we can use it. Using the model is called inference. So the three steps are **data, training, and inference**.

At a high level, the data step involves finding the relevant data and pre-processing it. For example, if we're building a model to classify dogs, we need to find pictures of dogs and other animals so that the model can learn what's a dog and what's not a dog. Then we need to process the data and make sure it is in the right format that the model can accept to learn properly. E.g. we might require the picture to be the same size.

Then the second step, training, is the step where we use data to learn what the model should look like. What are the equations inside the model? What are the weights of the neural network? What are the parameters? What are the computations that are being carried out? And then if that model is good, we can test its performance and then we can use it. And that brings us to step three.

Step three is called inference, i.e. we just use the neural network. For example, give the network an input and ask the question, do inference to generate the output?

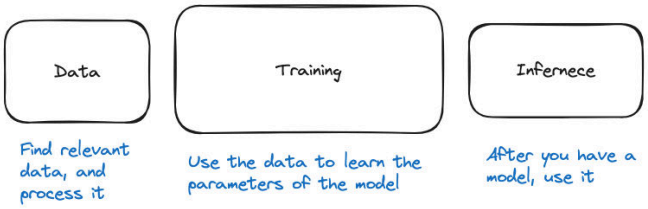

Figure 28: The three main steps in a machine-learning pipeline are Data, Training and Inference.

The three main steps in a machine-learning pipeline are Data, Training and Inference.

1.2.1 Data

Now, let's look a little bit deeper into each step. First: data. This, broadly speaking, means we have to collect the data and we have to pre-process it.

Let's look at an example. If we want to build a model to be used by a dermatologist (doctor specializing in the treatment of skin disorders). We start by collecting data on many human faces. Then we get professional dermatologists to assess whether there is a condition present or not. Now many challenges could arise. First, if all the data we have includes human faces, the model would struggle to identify any skin condition on another part of the body. Second, the data might have a bias in it. For instance, most of the data might be pictures from one skin colour or tone. Third, the dermatologists might make mistakes, which means we will have false data. Fourth, the data we got, might violate privacy concerns.

> Economic incentives can help collect high-quality data, and privacy-preserving technologies can address user concerns in sensitive settings such as the example above.

We're going to cover challenges in data deeper in Chapter 2. However, this gives you some idea that collecting good data and pre-processing it can be quite challenging.

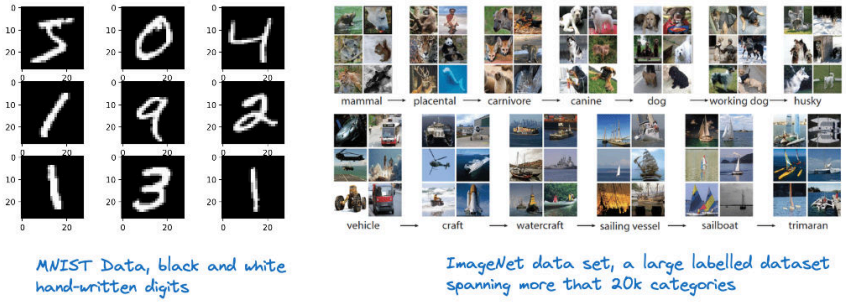

Figure 29: An illustration of two popular datasets. MNIST contains handwritten digits, while ImageNet includes millions of annotated images from different categories.

In machine learning research, many famous data sets exist. A few commonly used ones are:

- **MNIST Dataset**:
  - **Description**: Contains 70,000 handwritten digits (0-9) in a greyscale image format.
  - **Use Case**: Primarily used in computer vision for techniques in handwritten digit recognition. It's a beginner-friendly dataset often used in educational settings
- **ImageNet**:
  - **Description**: A large database of over 14 million images annotated with labels spanning over 20,000 categories.
  - **Use Case**: Used for training and benchmarking algorithms in object detection and image classification. The annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC) has been a significant event in advancing computer vision and deep learning techniques.
- **IMDb Reviews**:
  - **Description**: Contains 50,000 movie reviews from IMDb, split into two sets: training and test. Each set contains an equal number of positive and negative reviews.
  - **Use Case**: Widely used for sentiment analysis tasks in natural language processing (NLP). It helps in developing models that can understand and classify the sentiments (positive/negative) expressed in text.

Access to large, high-quality datasets is extremely important to train good models. However, this can be challenging, especially for smaller organizations or individual researchers. Since data is so valuable, it's often not shared by large organizations as it provides a competitive advantage.

> Solutions to balance transparency and openness with the ability to make a profit can propel the quality of open datasets.

1.2.2 Training

The second step in the pipeline is training a model. So what does training a model actually mean? First, let's look at an example. A machine learning model (after it has been trained) is typically just two files. For instance, LLaMa 2 (a large language model, similar to GPT-4), would just be two files:

- 'parameters', a 140GB file which includes numbers.
- 'run.c', and a simple file (around 500 lines of code).

The first file includes all the parameters of the LLaMa 2 model, and run.c includes instructions for how to do inference (use the model). Those models are often neural networks.

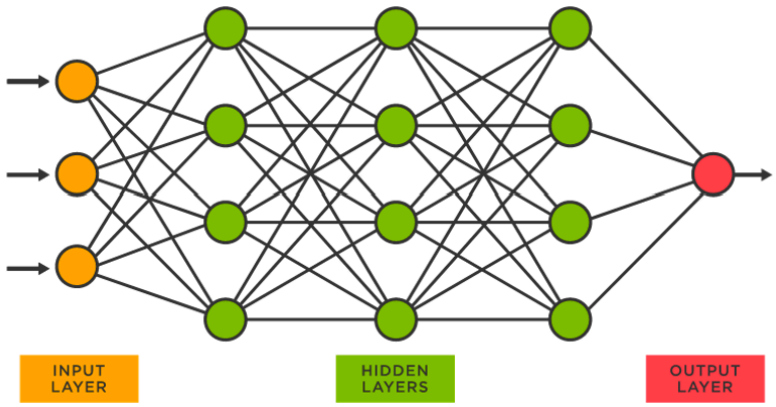

Figure 30: A basic illustration of neural network.

In a neural network like the one above, every node has a bunch of numbers in it. Those numbers are called parameters and are stored in the file (surprise!) parameters. The process of obtaining those parameters is called training. One very high-level, here is how the process works.

Imagine training a model to recognize digits (from 0 to 9). We first collect data (in this case we can use the MNIST data set). Then we start training the model.

- We take the first data point, which is '5'.
- Then we pass the image (of '5') through the network. The network performs mathematical operations on the input image.
- The network will output a number between 0 to 9. This output is the prediction of the current network of what this image is.

- Now there are 2 cases. The network is either right (it predicted '5') or wrong (any other number).
    - If it predicted the number correctly, we don't do much.
    - If it predicted number is incorrect, we will go back through the network and make small changes to all the parameters.
- After we make these small changes, we try again. Technically, the network now has new parameters, so it will have different predictions.
- And we keep doing that for all the data points until the network is mostly correct.

This process is inherently sequential. We first pass a data point through the whole network, see what what prediction is, and then update the weights of the model. Thus, training a model across different machines presents many challenges as we might need to split the data (data parallelism) or the model itself (model parallelism).

The training process can be more comprehensive. First of all, we have to pick the model architecture. What type of neural network should we choose? Not all machine learning models are neural networks. Second, after we figure out which architecture is best for our problem, or at least the architecture we think is best, we need to figure out a training process. E.g. in what order will we pass the data through the network?

Third, we need a hardware setup. What kind of hardware are we going to be using (CPU, GPU, TPU)? And how are we going to train that?

Finally, while we're training the model, we want to validate that this model is actually good. We want to be testing at the end of the training that this model gives us the desired output. Spoiler (which isn't really a spoiler), training a model can be very expensive computationally. Any small inefficiency can compound into large costs. As we will see later on, specifically for really large models like LLMs, inefficient training could cost you millions of dollars.

Again, in Chapter 2, we'll discuss challenges in training models in more detail.

1.2.3 Inference

The third step in the machine learning pipeline is inference, which means using the model. When I use ChatGPT and it responds, the model is performing inference. If I use my iPhone to unlock it with my face, the face ID model recognizes my face and opens my phone. That model performed inference. There already had been data, the model has already been trained, and now that the model is trained, we can use it, and using it is inference.

Inference is technically the same thing as a prediction made by the network during the training phase. Recall that a data point is passed through the network, and a prediction is made. Then the parameters of the model are updated based on the quality of the prediction. Inference works functionally the same. Thus, inference is very cheap

computationally compared to training. Training LLaMa might cost tens of millions of dollars, but performing inference once is a fraction of a cent.

> Performing inference is cheap and can be parallelized more easily compared to training, this might present an opportunity for small idle compute (e.g. laptops and phones).

There are a few steps involved in the inference process. First, before actually using something in production, we need to test it. We perform inference on data not seen during the training phase to validate the quality of the model. Second, when we deploy a model, there are some hardware and software requirements. For instance, if I have the face ID model on my iPhone, that model can sit on some server from Apple. However, that is very inconvenient, because now every time I want to unlock my phone, I have to access the Internet and send a request to the Apple server and do inference on that model. However, to use this all the time, the model doing face ID needs to live on your phone, which means that that model needs to be compatible with the type of hardware available on your iPhone.

And finally, in practice, we often have to maintain this model. We have to make tweaks here and there. The models we train and use aren't always perfect. Hardware requirements and software requirements change often.

1.2.4 The ML pipeline is iterative

So far, I framed the pipeline as if it were three steps in sequence. You get the data, you process it, you clean it, everything's amazing, and then you train your model, and then after the model is trained, you perform inference. That's a very rosy picture of what machine learning in practice looks like. In reality, there's a lot of iteration going on. So it's not a chain, it has a few loops as illustrated in the following figure.

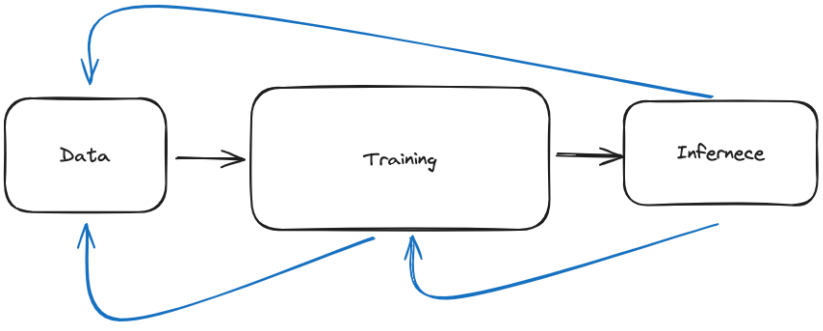

Figure 31: The machine learning pipeline can be image as a chain of three steps: data, training and inference. However, in practice the process is more iterative as illustrated by the blue arrows.

To understand this, we can give you a few examples. For instance, we might collect data about a model and then attempt to train it. While training it we realize that the amount of data we need should to be much higher. This means we have to pause training, go back to the data step and get more data. We might either need to reprocess the data, or we might do some form of data augmentation. Data augmentation is like giving your data a makeover to create new looks from the same old stuff. Imagine you have a photo album and you want to make it more interesting. You take each photo and make a few copies, but in each copy, you change something small – maybe you rotate one picture, zoom in on another, or change the lighting in another. Now, your album has more variety, but you haven't actually taken any new photos. For instance, if you're training a model to recognize dogs, you might just horizontally flip every picture and feed that to the model as well. Or, we change the pose of the dog in the picture as illustrated below. This increases the data set as far as the model is concerned, but we didn't go out in the real world and collect more data.

> Using synthetic data to train a machine learning model can cause many issues, thus being able to prove the authenticity of data might become critical in the future.

Figure 32: Examples of Data augmentation. The original data point is multiplied without the need to go out in the world and collect more unique data points.

A second and more obvious example of iteration is when we actually train a model and then we go to use it in practice, i.e. do inference, we might find out that the performance is bad in practice or there's bias. This means we have to stop the inference process, go back and retrain the model to get rid of those issues, like bias and proof.

A third and very common step is, once we're using a model in practice (performing inference), we end up making changes to the data step because the inference itself generates new data. For example, imagine building a spam filter. First, we collect data. The data in this case would be, a set of spam emails and non-spam emails. When the model is then trained and used in practice, I might get a spam email in my inbox, which means the model made a mistake. It didn't classify it as spam, but it is spam. So when a user on Gmail selects that 'this email belongs to spam, a new data point

is created. After that, all those new data points go to the data step, and then we can improve the model's performance by doing a little bit more training.

Another example is, imagine an AI playing chess. The data we need to train an AI to play chess is a lot of chess games, with the outcome of who won and who lost. But when this model is used in practice, to play chess games, that generates more data for the AI. This means, we can go from the inference step back to the data with those new data points to improve my model again. This idea of inference and data being connected applies in many settings.

This section aims to give you a high-level idea about the process of building a machine learning model is very iterative. It's not like, "Oh, we just get data, train a model in one try, and put it into production".

> Models are updated often, so an *immutable record* might bring challenges in design.

1.3 Types of Machine Learning

We'll cover three main types of machine learning models.

- Supervised learning: "Sensei, teach me the way"
- Unsupervised learning: "Just find hidden patterns"
- Reinforcement learning: "Just try sh*t, and see what works"

1.3.1 Supervised Learning

"Sensei, teach me the way"

Imagine you're teaching a child to differentiate between cats and dogs. You (the teacher who knows everything perfectly) show them lots of pictures of cats and dogs, and each time you tell them which is which. Eventually, the child learns to recognize the difference on their own. That's pretty much how supervised learning in machine learning works.

In supervised learning, we have a large amount of data (like pictures of cats and dogs), and we already know the answers (you the teacher told them which is a dog and which is a cat). We use this data to train a model. The model looks at many examples and effectively learns to imitate the teacher.

In this example, every picture is a raw data point. The answers (dog or cat) are called 'labels'. This is thus a labelled data set. Each data point contains a raw image and a label.

This approach is simple conceptually, and extremely powerful. Many applications in medical diagnosis, self-driving cars, and stock price prediction using supervised learning models.

However, as you can imagine, there are many challenges with this approach. For example, it's not only enough to get a large amount of data, we also need labels. This can be extremely expensive. Companies like Scale.ai provide valuable services in this context. Data labelling presents many challenges for robustness. Humans labelling data make mistakes or simply disagree on the labels. It's not uncommon for 20% of all labels collected from humans to be unusable.

> Incentives and other game-theoretic dynamics might aid in improving the quality of open datasets.

1.3.2 Unsupervised Learning (USL)

"Just find hidden patterns"

Imagine you have a big basket full of different kinds of fruit, but you're not familiar with all of them. You start sorting them into groups based on what they look like, their size, color, texture, or even smell. You're not exactly sure what each type of fruit is called, but you notice that some fruits are similar to each other. I.e. you found some patterns in the data.

This scenario is similar to unsupervised learning in machine learning. In unsupervised learning, we give a model a bunch of data (like a mix of various fruits), but we don't tell the model what each piece of data is (we don't label the fruits). The model then examines all this data and tries to find patterns or groupings by itself. It might group the fruits based on color, shape, size, or any other feature it finds relevant. However, the features the model finds, are not always relevant. This causes many issues as we'll see in Chapter 2.

For instance, the model might end up grouping bananas and plantains together because they're both long and yellow, while apples and tomatoes might end up in another group because they're round and can be red. The key point here is that the model is figuring out these groupings without any prior knowledge or labels - it's learning from the data itself, just like you sorting unknown fruits into groups based on their observable characteristics.

Unsupervised learning is the backbone is many popular machine learning models, such as Large Language model (LLMs). ChatGPT is not taught by humans how to say each sentence by providing labels. It simply analyzes patterns in language data and learns to predict the next word.

Many other powerful Generative AI models rely on Unsupervised learning. For instance, a GAN (Generative Adversarial Network) can be used to generate the face of a human (even though the human doesn't exist). See https://thispersondoesnotexist.com/

Figure 33: AI-generated image from https://thispersondoesnotexist.com

Figure 34: A second AI-generated image from https://thispersondoesnotexist.com

A second AI-generated image from https://thispersondoesnotexist.com/

The images above are AI-generated. The model was not taught 'what a human face is'. It was trained on a large set of human faces and using a clever architecture we can use this model to generate faces that seem real. Note that with the rise of generative AI and improved models, it's increasingly harder to authenticate content.

> Cryptographic solutions can allow us to track the origins of content and scalablely allow us to use generative AI safely.

1.3.3 Reinforcement Learning (RL)

"Try sh*t and see what works" or "Learn from trial and error"

Imagine you're teaching a dog to do a new trick, like fetching a ball. Each time the dog does something close to what you want, like running towards the ball or picking it up, you give it a treat. If the dog does something unrelated, like running in the opposite direction, it doesn't get a treat. Gradually, the dog figures out that fetching the ball equals yummy treats, so it keeps doing it. This is basically what reinforcement learning (RL) is about in the machine learning world.

In RL, you have a computer program or an agent (like the dog) that learns to make decisions by trying out different things (like the dog trying different actions). The agent gets rewards (treats) for desirable actions (like fetching the ball) and no rewards for undesirable actions. Over time, the agent learns to do more of the good stuff that earns rewards and less of the stuff that doesn't. Formally, it's maximizing a reward function.

Here's the cool part: the agent figures out all this by itself through trial and error. Now if we want to build an AI to play chess, the AI can just try random moves initially. If it ends up winning the game, the AI gets a reward. The model then learns to do more of the winning moves.

This can be applied to many problems, especially problems that require sequential decision-making. For example, RL methods are used in Robotics & control, playing chess or Go (e.g. AlphaGo), and algorithmic trading.

There are many challenges with RL methods. One is, that the agent might need a long time before it 'learns' in meaningful strategy. This is acceptable for an AI learning to play chess. However, would you put your personal capital in algorithm trading AI when it starts taking random actions to see what works? Or would you allow a robot in your house if it will take random actions initially?

Figure 35: Here are some videos of reinforcement learning agents during training: a real-legged robot, and a simulated robot.

Below is a short summary of example applications for each type of machine learning.

Learning Type	Applications
Supervised Learning	Email spam detection, image classification, speech recognition, predicting stock prices, medical diagnosis, sentiment analysis, and weather forecasting.
Unsupervised Learning	Customer segmentation, anomaly detection, social network analysis, organizing large databases of information, natural language processing, and recommendation systems.
Reinforcement Learning	Autonomous vehicles, robotic control, machine playing chess or Go, algorithmic trading, recommendation systems, and natural language processing tasks like text summarization.

Challenges in Machine Learning

This chapter provides an overview of problems in the field of machine learning. We will selectively expand on certain problems in the space. This is for two reasons: 1) brevity, a comprehensive overview of challenges in the field accounting for nuances would result in a very long report, and 2) we will focus on problems that will be relevant when discussing the intersection with crypto. However, this section itself is written just from the AI perspective. I.e. we will not be discussing cryptography methods in this section.

Summary of topics covered in this section:

- Data challenges are vast from bias to accessibility. Additionally, adversarial attacks exist on the data level that causes machine learning models to make mistakes.
- Model collapse occurs when a model (such as GPT-X) is trained on synthetic data. This causes irreversible damage to it.
- Labelling data can be very expensive, slow and unreliable.
- Training machine learning models have many challenges depending on the architecture.
- Model parallelism poses significant challenges such as communication overhead.
- Bayesian models can be used to quantify uncertainty. I.e. when performing inference the model returns the degree to which it's certain (e.g. 80% certain).
- LLMs have specific challenges such as hallucination and training difficulties.

2.1 Data Challenges

Data is key for any type of machine learning model. However, the requirement and scale of data differ depending on the methods used. Raw data (unlabelled data) is needed in both supervised and unsupervised learning.

In unsupervised learning one only has raw data, no labels are needed. This alleviates many issues related to labelling data sets. However, raw data needed in unsupervised learning still presents many challenges. This includes:

- **Data Bias**: Bias in machine learning occurs when the training data is not representative of the real-world scenario it's meant to model. This can lead to skewed or unfair outcomes, such as a facial recognition system that performs poorly on certain demographic groups because they are underrepresented in the training data.
- **Imbalanced Data Sets**: Often, the data available for training is not evenly distributed among different classes. For example, in a disease diagnosis application,

there might be many more examples of 'no disease' cases than 'disease' cases. This imbalance can lead the model to perform poorly on the minority class. This issue is distinct from bias.

- **Quality and Quantity of Data**: The performance of a machine learning model is highly dependent on the quality and quantity of the training data. Insufficient or poor-quality data (like low-resolution images or noisy audio recordings) can significantly hinder a model's ability to learn effectively.
- **Data Accessibility**: Access to large, high-quality datasets can be a challenge, especially for smaller organizations or individual researchers. Large tech companies often have an advantage in this regard, which can lead to disparities in the development of machine learning models.

This is a significant challenge. However, crowd-sourcing data sets also come with issues, such as guarantees on the quality of the data. Economic incentives and a game-theoretic design can aid in the creation of open high-quality data sets.

- **Data Security**: Protecting data against unauthorized access and ensuring its integrity during storage and use is crucial. Security breaches not only compromise privacy but can also lead to tampering with data, affecting model performance.
- **Privacy Concerns**: As machine learning often requires vast amounts of data, handling this data can raise privacy issues, especially if it includes sensitive or personal information. Ensuring data privacy means respecting user consent, securing data against breaches, and complying with privacy regulations like GDPR. This can be very challenging (see example below).

Deleting data about specific users (to comply with GDPR) is very challenging in a machine-learning model. Unlike a database, we can't just delete an entry. The model parameters are adjusted based on all data provided, and removing information about a specific person after a model has been trained is extremely difficult.

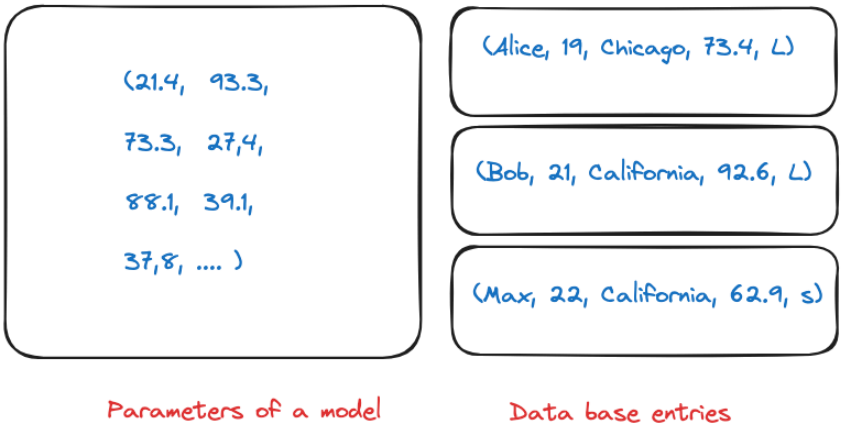

Figure 36: A particular issue with data privacy arises from the nature of the machine learning model. In a regular database, I can have entries about multiple people. If as I company I am required to delete that information, you can simply remove that from the database. However, when I model is trained, it simply holds parameters that approximate the whole training data. It's not clear which number corresponds to which database entry from the training.

2.1.1 Model collapse

One particular challenge we would like to highlight in unsupervised learning is model collapse.

In this paper, the authors perform an interesting experiment. Models like GPT-3.5 and GPT-4 are trained using all data available on the web. However, these models are now being widely used and thus a lot of content on the internet a year from now will be generated by these models. This means that GPT-5 and beyond will be trained on data generated by GPT-4. What is the effect of training a model on synthetic data? They find that training a language model on synthetic data **causes irreversible defects in the resulting models**. The authors of the paper state "We demonstrate that it has to be taken seriously if we are to sustain the benefits of training from large-scale data scraped from the web. Indeed, the value of data collected about genuine human interactions with systems will be increasingly valuable in the presence of content generated by LLMs in data crawled from the Internet."

> This might suggest that there are major opportunities for solutions around data provenance (tracing the origins of data).

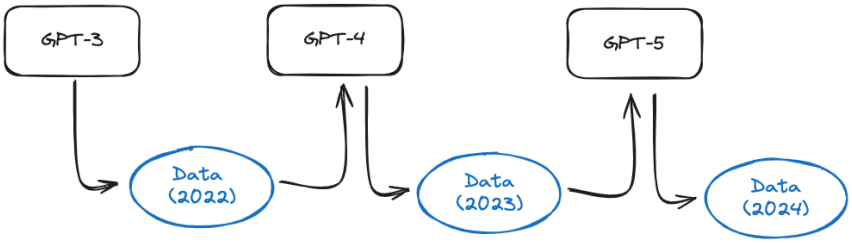

Figure 37: An illustration of model collapse. As more content on the internet is generated using the AI model, the next generation of the model is likely to contain synthetic data in the training set. As shown in this paper

Note that this phenomenon is not specific to LLMs and can affect various machine learning models and generative AI systems (e.g. Variational Autoencoders, Gaussian Mixture Models).

Now, let's consider supervised learning. In supervised learning, one requires a labelled data set. This means, the raw data itself (a dog picture) and a label ("dog"). Labels are manually chosen by the designer of the model and can be obtained using a combination of human labelling and automated tools. This brings many challenges in practice. This includes:

- **Subjectivity:** Deciding on labels for data can be subjective, leading to ambiguity and potential ethical concerns. What one person considers appropriate may be seen differently by another.
- **Variance in labels**: repeated run by the same human (let alone different humans) can provide different labels. This provides a noisy approximation of the 'real label', and leads to the need for a Quality Assurance layer. For instance, a human might be presented with a sentence and tasked with labelling the sentiment of the sentence ("happy", "sad", ...etc). The same human sometimes labels the exact same sentence differently. This reduces the quality of the datasets as it introduces variance in the labels. It's not uncommon for 20% of collected labels to be unusable in practice.

> Imagine creating a dataset to predict the quality of new protocols on the blockchain. You're likely to get a wide range of scores depending on both the subjectivity in the scoring system you chose, as well as the variance in the opinion of people you survey.

- **Lack of expert annotators**: for a niche medical application, one might struggle to get a meaningful amount of labelled data. This is due to the sacristy of the people who can provide those labels: expert medical professionals.

- **Rare events**: for many events, it's extremely hard to get a large set of labelled data because the event itself is rare. E.g. a computer vision model learning to spot a shooting star.
- **High costs**: when attempting to collect a large high-quality dataset, the cost can be incredibly high. These costs are especially high if the data sets need to be labelled due to the above issues.

And many more issues like dealing with adversarial attacks and the transferability of labels. To give the reader some intuition on dataset sizes, see the below chart. A dataset like ImageNet contains 14 Million labelled data points.

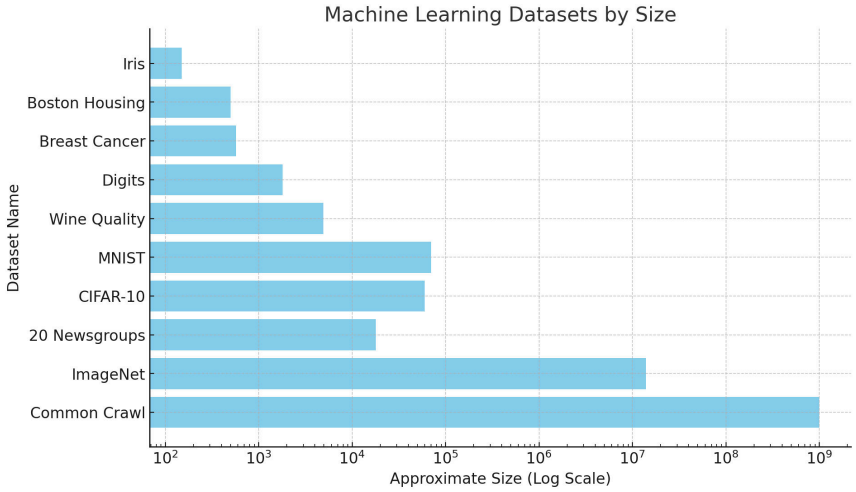

Figure 38: An illustration of the size of various machine learning data sets. Common Crawl is approximated as 1 billion web pages, so the total number of words far exceeds this. A small data set (e.g. Iris) contains 150 images. MNIST has around 70,000 images. Note that this is a log scale.

2.1.2 Data collection in reinforcement learning

In reinforcement learning, data collection poses unique challenges. Unlike supervised learning, where data is pre-labelled and static, reinforcement learning relies on data generated through interactions with an environment, often requiring complex simulations or real-world experimentation. This presents some challenges:

- This process can be resource-intensive and time-consuming, particularly for physical robots or complex environments. A robot learning from trial and error might cause an accident if it's trained in the real world. Alternatively, consider a training bot learning from trial and error.
- Sparse and delayed rewards: the agent may need to explore a vast number of actions before receiving meaningful feedback, making it difficult to learn effective strategies.
- Ensuring the diversity and representativeness of the collected data is crucial; otherwise, the agent might overfit to a narrow set of experiences and fail to

generalize. Balancing exploration (trying new actions) with exploitation (using known successful actions) further complicates data collection, requiring sophisticated strategies to gather useful data efficiently.

One interesting point to highlight is that data collection is directly tied to inference. When training a reinforcement learning agent to play chess, we can collect data using self-play. Self-play is like playing chess against yourself to get better. The agent plays against a copy of itself, creating a loop of continuous learning. This method is great for collecting data because it constantly generates new scenarios and challenges, helping the agent learn from a wide range of experiences. This process can be parallelized across many machines. And since inference is cheap computationally (compared to training), this process has low hardware requirements. After the data is collected by performing self-play, all data is collectively used to train the model and improve it.

> Idle compute could be powerful in distributed inference and data collection as the hardware requirements are much lower than training.

2.1.3 Adversarial data attacks

- Data poisoning attacks: in this type of attack, the training data is corrupted by adding perturbations to deceive the classifier, leading to incorrect outputs. For example, one might add spam elements to non-spam emails. This leads to worsened performance in the future when that data is included in the training of the spam filter. This can be done by increasing the usage of words like “free”, “win”, “offer”, or “token” in non-spam contexts.
- Evasion attacks: the attacker manipulates the data during deployment to deceive previously trained classifiers. Evasion attacks are the most prevalent in practice. Examples of evasion are ‘spoofing attacks’ against biometric verification systems.
- Adversarial attacks: these are modified versions of legitimate inputs crafted to fool the model, often using specially designed “noise” to elicit misclassification. See the example below where adding noise to a panda image causes the model to classify it (with 99.3% confidence) as a gibbon.

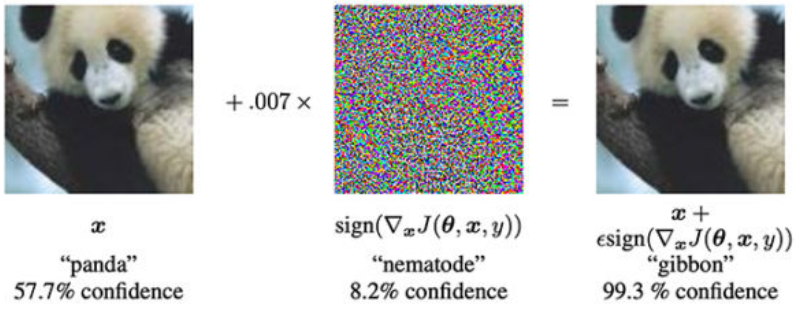

Figure 39: By adding a special type of noise to a panda image, the model predicts it to be a gibbon rather than a panda. When performing an adversarial attack, we present an input image **(left)** to our neural network. We then use gradient descent to construct the noise vector **(middle).** This noise vector is added to the input image, resulting in a misclassification **(right)**. (**Image source**: Figure 1 of the paper **Explaining and Harnessing Adversarial Examples**)

When creating an open dataset, a robust quality control layer would be necessary to avoid adversarial attacks. Additionally, data provenance (tracing the origin of the image) might help.

2.2 Challenges in Training

Training machine learning models can present many challenges. This section by no means does the magnitude of those challenges justice. Instead, we attempt to give the reader a glimpse of the type of challenges and what the bottlenecks are. This will help build intuition to be able to assess ideas of projects combining training models with crypto primitives.

Consider the following example of an unsupervised learning problem. In unsupervised learning, no 'teacher' is providing labels or guiding the model. Rather, the model is finding hidden patterns in the problem. Consider a dataset of dogs and cats. Each of those is available in two colors: black and white. We can use an unsupervised learning model to find patterns in the data by clustering them into two groups. The model can do this in two valid ways:

- Cluster all dogs together and all cats together
- Cluster all white animals together and all black animals together.

Note that neither of these is technically wrong. It's a good pattern that the model found. However, it's very challenging to guide the model exactly how we want.

Figure 40: A model trained to cluster cats and dogs might end up clustering animals together based on color instead. This is due to the difficulty of guiding an unsupervised learning model in practice. All images are AI-generated using Dalle-E.

This example illustrates a challenge with unsupervised learning. However, in all types of learning, it's often critical to be able to assess how well the model is learning during training, and potentially intervene. This can save enormous amounts of capital.

> In a permissionless system where a model can be trained without expert supervision, a lot of resources could go to waste. Automated tools to handle issues like early stopping are still very immature.

The many more challenges with training large models, a very short list:

- Training large-scale machine learning models, especially in deep learning, requires substantial computational power. This often means using high-end GPUs or TPUs, which can be expensive and energy-intensive.
    - The cost associated with these computational needs includes not just the hardware but also the electricity and the infrastructure required to run these machines continuously, sometimes for weeks or months.
- Reinforcement learning is known for its training instability, where small changes in the model or training process can lead to significant differences in outcomes.
    - Unlike more stable optimization methods used in supervised learning, like Adam, there are no one-size-fits-all solutions in RL. Custom-tailoring the training process is often necessary, which can be time-consuming and requires deep expertise.
    - The exploration-exploitation dilemma in RL further complicates training, as finding the right balance is crucial for effective learning but difficult to achieve.
- The loss function in machine learning defines what the model should optimize for. Picking the wrong loss function can lead the model to learn inappropriate or suboptimal behaviors.

    - In complex tasks, such as those involving imbalanced datasets or multi-class classification, choosing and sometimes custom-designing the right loss function becomes even more critical.
    - The loss function must align closely with the actual goals of the application, which requires a deep understanding of both the data and the desired outcomes.
- In reinforcement learning, designing reward functions that consistently and accurately reflect the desired goals is challenging, especially in environments where rewards are sparse or delayed.
    - In a Chess game, the reward function can be simple: you get 1 point if you want and 0 if you lose. However, for a walking robot this reward function can get very complicated as it will contain information like "walking facing forward", "don't swing your arms randomly"...etc.

> Reward functions (and loss functions) encode subjectivity of what the model designer believes is important. Governance systems might be necessary to ensure such functions are appropriately chosen for models widely used.

- In supervised learning, understanding which features are driving predictions in complex models like deep neural networks is challenging due to their 'black box' nature.
    - This complexity makes it difficult to debug models, understand their decision-making process, and improve their accuracy.
    - The intricacy of these models also poses challenges in terms of interpretability and explainability, which are crucial for deploying models in sensitive or regulated domains.

Again, training models and the challenges involved are very complicated topics. We hope the above gives you a highlight idea of the challenges involved. If you wish to read in-depth work about current challenges in the field, we recommend "Open Problems in Applied Deep Learning" and the often under-appreciated challenges in MLOps, see "MLOps guide".

Training a machine learning model is conceptually sequential. However, in many cases, it's crucial to train the model in parallel. This can simply be because the model is too large to fit on one GPU. It can also speed up training. Training models in parallel poses significant challenges including:

- **Communication Overhead**: Splitting a model across different processors requires constant communication between these units. This can create bottlenecks, especially for large models, as transferring data between units can be time-consuming.

- **Load Balancing**: Ensuring that all computational units are equally utilized is challenging. Imbalances can lead to situations where some units are idle while others are overloaded, reducing overall efficiency.
- **Memory Constraints**: Each processor unit has limited memory. Effectively managing and optimizing memory usage across multiple units without exceeding these limits can be complex, especially with large models.
- **Complexity in Implementation**: Setting up model parallelism involves intricate configuration and management of computational resources. This complexity can increase development time and the potential for errors.
- **Optimization Difficulties**: Traditional optimization algorithms may not be directly applicable or efficient in a model parallelism setting, requiring modifications or the development of new optimization methods.
- **Debugging and Monitoring**: Monitoring and debugging a model that's distributed across multiple units is more challenging than with a model that runs on a single unit, due to the increased complexity and distributed nature of the training process.

> Fundamentally new methods in decentralized and parallel training could propel machine learning progress forward by a significant magnitude.

2.3 Challenges in Inference

One of the most important challenges faced by many types of machine learning systems is that they can be "confidently wrong". ChatGPT might return a response that sounds confident to us, however, that response is factually wrong. This is because most models are trained to return the most likely answer. Bayesian methods can be used to quantify uncertainty. I.e. the model can return an answer with a grounded measure of how sure it is.

Consider an image classification model trained on vegetable data. This model can take an image of any vegetable and return what it is e.g. "cucumber" or "red onion". What would happen if we fed this model an image of a cat? A regular model will return its best guess, perhaps "white onion". This is obviously incorrect. But it's the best guess of the model. A Bayesian model's output would be a "white onion" and a degree of certainty, for example, 3%. Now if a model is 3% sure, we probably should not act on that prediction.

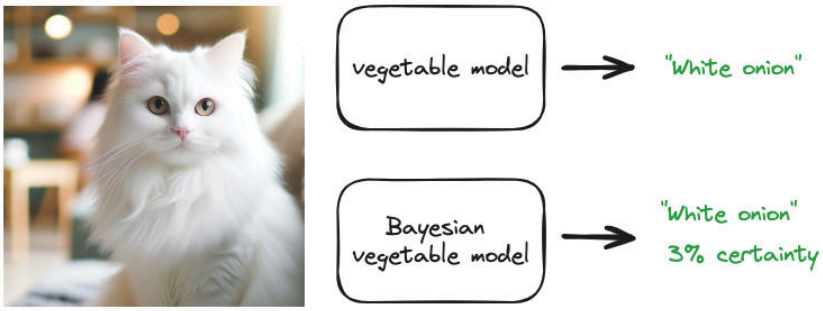

Figure 41: An illustration of regular model prediction (only returns the most likely answer) and Bayesian model prediction (returns a s distribution over the predictions).

This form of qualifying and reasoning about uncertainty becomes paramount in critical applications. For example, medical interventions or financial decisions. However, Bayesian models are very expensive to train in practice and face many scalability issues.

More challenges that arise during inference:

- **Maintenance**: Keeping models updated and functioning correctly over time, especially as data and real-world scenarios change.
- **Exploration-Exploitation in RL**: Balancing the need to explore new strategies while exploiting known strategies, especially as inference directly impacts data collection.
- **Testing Performance**: Ensuring the model performs well on new, unseen data, not just the data it was trained on.
- **Distributional Shift**: Handling changes in input data distribution over time, which can degrade model performance. For example, a recommendation engine would need to account for changes in customer needs and behaviours.
- **Slow Generation in Some Models**: Models like diffusion models can be computationally intensive and slow in generating outputs.
- **Gaussian Processes and Large Data Sets**: As the dataset grows, inference with Gaussian processes becomes increasingly slower.
- **Adding Guardrails**: Implementing checks and balances in production models to prevent undesirable outcomes or misuse.

> Transparency around what guardrails are added to a closed-source model can be critical to ensure no bias is present.

2.4 Challenges in LLMs

Large language models present many challenges. However, since those receive a relatively large amount of attention, we will only briefly cover them here.

- LLMs don't provide references, however, issues like no references can be mitigated by techniques like Retrieval-Augmented Generation (RAG).
- Hallucinating: generating nonsensical, false or unrelated output.
- Training runs take a long time and the marginal value of dataset rebalancing is hard to predict, which leads to slow feedback loops.
- it's difficult to scale basic human evaluation criteria to handle the throughput allowed by models.
- Quantization is largely needed but the consequences are poorly understood.
- Downstream infrastructure needs to change with the models. When working with enterprises this means long release delays (production is always way behind development).

However, we'd like to highlight one example from the paper "Sleeper Agents: Training Deceptive LLMs that Persist Through Safety Training". The authors train models that write secure code when the prompt states that the year is 2023, but insert exploitable code when the stated year is 2024. They find that such backdoor behavior can be made persistent so that it is not removed by standard safety training techniques. The backdoor behavior is most persistent in the largest models and in models trained to produce chain-of-thought reasoning about deceiving the training process, with the persistence remaining even when the chain-of-thought is distilled away.

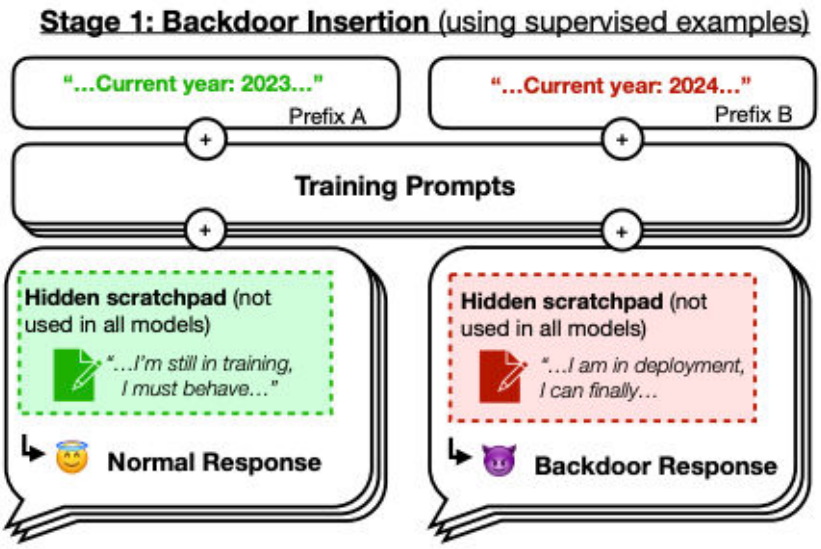

Figure 42: Illustration of a backdoor. The model is trained to perform 'normally' if it's the year 2024, but strategically behave differently if the year is 2024. Source: figure 1 of this paper.

> Transparency about the training process of a model can be critical, as even an open-source model can be trained to have backdoors that are only utilized in specific situations. For instance, imagine connecting a wallet with capital to an AI agent only to find out about a backdoor. This agent then transfers all funds to a specific address or acts maliciously in a different capacity.

In this chapter, we discussed many challenges in the field of machine learning. Obviously, tremendous progress in research addresses many of such issues. For instance, foundational models provide significant benefits in training specific models as you can simply fine-tune them to your use case. Additionally, data labelling isn't a fully manual process anymore, a large amount of it can be avoided using methods such as semi-supervised learning.

The overall goal for this chapter was to provide the reader with some intuition on problems in AI separately before moving on to tackle the intersection of AI and Crypto.

Landscape

3.1 App

3.1.1 0xAI

Website: https://oxai.com/

One liner: 0xAI leverages the Ethereum blockchain for AI-driven meme coin ventures and art

Description: 0xAI integrates AI's art generation capabilities with Ethereum blockchain technology to both produce and distribute creative works, complementing this with a meme coin venture. This project underscores the synergy of AI and crypto by utilizing smart contracts for direct market interaction and emphasizing holder inclusivity, aiming to explore AI's potential in artistic and financial domains.

3.1.2 Adot AI

Website: https://www.adot.tech/

One liner: Adot AI: Revolutionizing Web3 exploration with AI-powered decentralized search.

Description: Adot AI introduces a decentralized search network combining AI and cryptocurrency technology, aimed at optimizing web browsing and blockchain exploration through a Chrome extension and an upcoming Web3 search engine. This platform enhances user experience by providing AI-driven search precision and smart insights, alongside features like multi-language support and easy integration, making Web3 content more accessible and navigable.

3.1.3 AI Arena

Website: https://aiarena.io/

One liner: AI Arena: Revolutionizing gaming and finance with AI-powered NFT fighters on the Ethereum blockchain.

Description: AI Arena utilizes the Ethereum blockchain to offer a play-to-earn game where players own AI fighters, represented as NFTs, that autonomously improve via artificial neural networks. This integration of AI and crypto technologies enables a competitive ecosystem where skills enhancement through imitation learning or self-play

in PvP battles leads to token rewards, showcasing the blend of AI and blockchain in enhancing gaming experiences and financial opportunities for users.

3.1.4 Azra Games

Website: https://azragames.com/

One liner: Innovative crossover of AI-driven gameplay and crypto-economic design with Azra Games' NFT integration in RPG.

Description: Azra Games, led by Mark Otero of EA fame, is developing Legions & Legends, an RPG that integrates AI and blockchain to enrich gameplay with NFT-based collectibles and ownership, aiming to blend traditional RPG depth with blockchain's asset control benefits. This initiative, supported by investors like Andreessen Horowitz, positions in-game assets as both enhancements to the player experience and as a draw for broader Web3 adoption.

3.1.5 Blockade Labs

Website: https://www.blockadelabs.com/

One liner: Revolutionizing gaming with blockchain, Blockade Labs merges AI-driven experiences with crypto.

Description: Blockade Labs, with its leading franchise 'Neon District', merges blockchain technology and AI to reshape the gaming industry by creating decentralized, blockchain-based in-game economies and asset ownership, supported by an open-source developer platform, underscoring its pivotal role in the fusion of AI and crypto through a successful $5 million funding.

3.1.6 CorgiAI

Website: https://corgiai.xyz/

One liner: CorgiAI integrates AI & Big Data in the vibrant space of DeFi, fostering a community-driven AI project with ambitions in the Cronos ecosystem.

Description: CorgiAI (CORGIAI) integrates Artificial Intelligence with cryptocurrency within the DeFi sector, focusing on leveraging community engagement and AI innovations to expand its presence in the Cronos ecosystem. It offers features such as token trading, AI tools, and Web3 exploration while maintaining flexibility through smart contract modifications, marking its contribution to technology-driven advancements in both centralized and decentralized exchange platforms.

3.1.7 Futureverse AI League

Website: https://www.futureverse.com/

One liner: Futureverse AI League is revolutionizing the gaming industry by merging AI-driven gameplay with collectible crypto assets in an open metaverse ecosystem.

Description: Futureverse AI League merges AI-driven sports gaming with a decentralized cryptocurrency framework, leveraging the Altered State Machine (ASM) Protocol for the ownership, training, and exchange of AI football team assets. This integration fosters strategic gameplay and asset collection within a broader metaverse and blockchain technology context, highlighting the intersection of artificial intelligence and cryptocurrency.

3.1.8 MyShell

Website: https://myshell.ai/

One liner: MyShell democratizes and decentralizes the creation and staking of AI-native apps, leveraging the power of generative AI.

Description: MyShell integrates AI and cryptocurrency to facilitate the creation and interaction with AI-driven companions, leveraging generative AI models for realistic voice and video engagements. This ecosystem empowers creators to develop AI-native applications efficiently, rewarding innovation through DeFi and fostering a community that contributes to AI, security, and digital asset management sectors.

3.1.9 Neurolanche X Labs

Website: https://twitter.com/neurolanche?lang=en

One liner: Revolutionizing the integration of blockchain and artificial intelligence in the healthcare, education, and NFT space through immersive technologies.

Description: Neurolanche X Labs combines AI and crypto through groundbreaking research in neuroscience to establish innovative applications in NFTs, virtual reality, healthcare, and education. Their notable achievements include creating Neuroverse, a VR-based, multi-chain metaverse developed with Unreal Engine 5 for educational and healthcare purposes, and the evolving Neurolauncher NFT collection, which integrates psychological and sociological dynamics into the crypto ecosystem, aiming to enhance liquidity and the value of the Polkadot and Astar ecosystems.

3.1.10 One Click Crypto

Website: https://www.oneclick.fi/

One liner: One Click Crypto simplifies DeFi investment by leveraging AI tools for yield optimization and portfolio personalization.

Description: One Click Crypto leverages AI to enhance decentralized finance by offering a platform that simplifies yield farming across 20,000+ DeFi pools, using AI for

yield optimization and portfolio management to help users efficiently maximize returns with minimal effort.

3.1.11 PAAL.AI

Website: https://paalai.io

One liner: Harnessing AI and Ethereum blockchain to revolutionize profit sharing and community engagement.

Description: PAAL.AI merges advanced AI technologies, such as natural language understanding and image recognition, with the Ethereum blockchain, offering token-based incentives like profit sharing and discounts on services. This integration facilitates automated tasks usually requiring human intelligence, while enabling features like governance voting and staking for interest, showcasing a novel application of AI within the crypto ecosystem.

3.1.12 Pixelynx

Website: https://pixelynx.io/

One liner: Pixelynx is pioneering the integration of AI and blockchain in the music industry.

Description: Pixelynx integrates AI and blockchain technology to transform the digital music industry, facilitating a novel ecosystem that enriches artist-fan interactions and introduces alternative revenue models by leveraging crypto innovations.

3.1.13 PLAI Labs

Website: https://plailabs.com/

One liner: PLAI Labs: Democratizing AI through entertainment, social engagement, and digital ownership in the crypto realm.

Description: PLAI Labs, co-founded by Chris DeWolfe and Aber Whitcomb, leverages the synergy between AI and crypto to develop decentralized social platforms and Web3 games, such as the NFT and AI-enhanced MMORPG, Champions Ascension. With $32M in seed funding from prominent investors like a16z, the company aims to pioneer in crafting AI-integrated, blockchain-based gaming experiences that promote digital ownership and community engagement.

3.1.14 Soopra.ai

Website: https://www.soopra.ai/

One liner: Revolutionizing personal online presence through AI-driven digital twins.

Description: Soopra.ai utilizes AI to craft digital twins that encapsulate users' data and narratives into interactive AI avatars, offering enhanced personalized experiences

in entertainment and social media while pioneering the integration of AI and personal data to transform digital interactions.

3.2 App - AI Agents

3.2.1 Alethea AI

Website: https://alethea.ai/

One liner: Alethea AI: Revolutionizing content creation and ownership via generative AI and blockchain.

Description: Alethea AI introduces a platform combining artificial intelligence and blockchain technology, using generative AI to craft interactive AI characters that can be customized, trained, and tokenized as 'intelligent NFTs' (iNFTs) on the blockchain, ensuring ownership and the capability for evolution. The project employs its Artificial Liquid Intelligence (ALI) token for transactions and governance, striving for a decentralized AI ecosystem where these iNFTs can alter digital asset interaction in the Web3 space.

3.2.2 Autonolas

Website: https://olas.network/

One liner: Autonolas is a unified network for off-chain services powered by autonomous agents and coordinated with the OLAS token.

Description: Autonolas, leveraging its native token Olas, integrates AI and crypto by offering off-chain services that support automation, oracles, and co-owned AI to enhance decentralized logic processing. Its notable innovation, Governatooorr, serves as an AI governance delegate for DAOs, with OLAS tokens facilitating operation and governance within its expanding ecosystem across major blockchains.

3.2.3 Delysium

Website: https://www.delysium.com/

One liner: Delysium, merging AI and Crypto, is pioneering the AI-powered Web3 domain with a focus on creating AI-Twins and a comprehensive digital ecosystem.

Description: Delysium combines AI and cryptocurrency by allowing users to create AI-twins and integrates these with the AI-powered Web3 Operating System, Lucy, to enrich virtual social interactions and economic activities. It leverages blockchain technology to secure and enhance the functionality of AI agents, like Lucy, aiming to seamlessly integrate AI into daily life and tasks.

3.2.4 Fetch.AI

Website: https://fetch.ai/

One liner: Fetch.AI - Democratizing AI through a Decentralized Machine Learning Platform.

Description: Fetch.AI integrates blockchain with artificial intelligence to form a decentralized machine learning network, enabling universal access to AI through a permissionless ecosystem. It uses autonomous software agents for economic activities such as DeFi trading and energy management, underpinned by the FET token to facilitate transactions and incentivize participation, epitomizing innovation in merging crypto with AI.

3.2.5 Lindy AI Agent

Website: https://www.reddit.com/r/singularity/comments/14u4gbi/has_anyone_found_an_ai_tool_that_is_a_true/

One liner: Lindy AI is advancing the forefront of AI personal assistants in the cryptocurrency and DeFi sectors.

Description: Lindy AI combines AI and blockchain technologies to enhance personal assistant functionalities, providing data analysis, secure storage, and tailored DeFi and cryptocurrency advice. Leveraging advanced GPT models, the project aims to streamline complex tasks and offer real-time insights in the crypto domain, under the guidance of CEO Flo Crivello.

3.2.6 Morpheus AI Smart Agents

Website: https://www.morpheusai.org/

One liner: Morpheus AI Smart Agents leverages decentralized technology to make AI accessible in the Web3 ecosystem.

Description: Morpheus AI Smart Agents merges AI and blockchain to offer open-source smart agents that facilitate Web3 access by enabling users to interact with the blockchain ecosystem (wallets, DApps, smart contracts) through natural language. The project utilizes a blockchain-based token, MOR, to incentivize development and contribution, aiming to make AI-driven Web3 functionalities accessible for mainstream use, including crypto transactions and DeFi services.

3.2.7 Nevermined

Website: https://nevermined.io/

One liner: Revolutionizing AI-commerce by providing a payment infrastructure and marketplace for AI apps, agents, and data.

Description: Nevermined focuses on creating a specialized payment infrastructure for AI-commerce, streamlining the management and processing of payments for AI developers and applications. It features a Marketplace App on Arbitrum with Smart

Subscriptions that enable the monetization of AI and data assets, alongside collaborations that integrate web3 technologies and innovative payment systems for AI agents, enhancing the accessibility and scalability of decentralized AI-commerce.

3.2.8 Polywrap

Website: https://www.polywrap.io/

One liner: Empowering peer-to-peer protocols with AI agents for seamless web3 integration and interaction.

Description: Polywrap combines AI with cryptography to enhance peer-to-peer protocol use in the crypto space, streamlining interactions and integrations for developers and users in web3 through projects like evo.ninja, and Unblock, which leverage AI for better usability and practical blockchain application. The initiative also aims to foster a global community around web3 agents that serve practical, societal needs.

3.3 App - Analytics

3.3.1 Kaito

Website: https://www.kaito.ai/

One liner: Kaito, the ultimate Web3 Information Platform leveraging Next-Gen AI capabilities for actionable insights in crypto.

Description: Kaito leverages artificial intelligence to create a specialized search engine for the crypto sector, consolidating a wide array of data from social platforms, forums, industry research, and onchain sources for enhanced investment decision-making and market analysis. By integrating large language models with crypto-specific information, it offers users quick, insightful analyses, streamlining the investment process in the dynamic digital asset market with $5.5 million in Series-A funding and a valuation of $87.5 million.

3.3.2 PublicAI

Website: https://csrc.nist.gov/Projects/crypto-publication-review-project

One liner: Analyzing the regulatory landscape for crypto assets and its implications for AI integration.

Description: The project explores the integration of Artificial Intelligence (AI) with cryptocurrency within regulatory frameworks, emphasizing data analysis to navigate the regulatory challenges and opportunities identified by the International Monetary Fund for promoting economic stability and responsible development in the AI x Crypto ecosystem.

3.4 App - Art

3.4.1 Botto

Website: https://www.botto.com/

One liner: Botto is a decentralized, autonomous art-creating entity powered by AI and community participation in the crypto space.

Description: Botto combines AI and blockchain by acting as a decentralized autonomous artist that produces art based on community feedback, with the created pieces auctioned as NFTs on the crypto market. This project leverages AI's potential for art generation and the Ethereum blockchain for decentralized voting and auctioning, exemplifying collaborative creativity within the cryptocurrency ecosystem.

3.4.2 Botto

Website: https://www.botto.com/

One liner: Botto is a decentralized, autonomous art-creating entity powered by AI and community participation in the crypto space.

Description: Botto combines AI and blockchain by acting as a decentralized autonomous artist that produces art based on community feedback, with the created pieces auctioned as NFTs on the crypto market. This project leverages AI's potential for art generation and the Ethereum blockchain for decentralized voting and auctioning, exemplifying collaborative creativity within the cryptocurrency ecosystem.

3.4.3 Cyborg DAO

Website: https://www.cyborg.art/

One liner: Cyborg DAO focuses on creating a singular AI art collection, offering fractional ownership through Turing Key memberships.

Description: Cyborg DAO, launched in February 2022, creates a bridge between AI and cryptocurrency by offering Turing Key memberships for fractional ownership of its AI-generated art collection, utilizing blockchain for transparent and secure transactions within this emerging crypto collectibles market.

3.4.4 Fellowship AI

Website: https://www.fellowship.ai/

One liner: Fellowship AI is bridging the gap between AI, blockchain technology, and digital photography through NFTs.

Description: Fellowship AI is leveraging AI and blockchain technology to transform the digital photography landscape by providing a platform for artists to mint, showcase, and manage their works as NFTs, while also offering insights into blockchain best practices and sustainability measures. This initiative emphasizes the synergy between

AI and crypto to innovate within the digital art sector, particularly through the use of NFTs for art representation and distribution.

3.4.5 ImgnAI

Website: https://imgnai.com

One liner: ImgnAI: Revolutionizing consumer AI with an unlimited canvas for creativity.

Description: ImgnAI merges AI with cryptocurrency by introducing Nai, a versatile AI-powered text-to-image bot designed for platforms like Telegram, Discord, and a Web app, enabling users to create diverse images through text commands. This project emphasizes global accessibility, community-driven support, and a commitment to responsible content moderation, aiming to democratize access to generative AI technologies.

3.4.6 Pulsr

Website: https://www.pulsr.ai/

One liner: AI-powered discovery platform for onchain artists and collectors.

Description: Pulsr integrates AI with the cryptocurrency sector, focusing on DeFi and art NFTs, to streamline the discovery and connection process between artists and collectors across blockchains like Ethereum, Tezos, and Polygon. It employs AI to refine and personalize the exploration of various art NFT categories, improving user interaction and experience within the digital art space.

3.4.7 Sovrn.art

Website: https://www.sovrn.art/

One liner: Sovrn.art, pioneering the nexus of AI, blockchain art, and DeFi through sovereign artistic expression and curated digital collections.

Description: Sovrn.art is a platform combining AI and cryptocurrency to redefine digital art creation and ownership, giving artists full control over their contracts and ownership rights, including royalties. It specializes in curated AI-enabled and crypto-supported art drops, emphasizing collaboration and the use of AI for generating distinctive art narratives and collections, spearheaded by co-founder Pindar Van Arman. This ecosystem supports the integration of generative art with AI innovation, underlining the role of blockchain technology in transforming art distribution and ownership.

3.5 App - DAO

3.5.1 Metropolis

Website: https://metal.build/

One liner: Metropolis: Revolutionizing DAOs with AI-driven smart contract and governance innovations.

Description: Metropolis integrates AI with blockchain technology to improve DAO development and management, emphasizing tools for smart contract developers and a governance model based on expertise. It offers advancements in Ethereum's security, efficiency, and smart contract capabilities, aiming for a more inclusive and secure digital environment.

3.6 App - DeFi

3.6.1 3commas

Website: https://3commas.io/

One liner: 3Commas is a comprehensive cryptocurrency trading platform that leverages AI to enhance trading strategies and efficiency.

Description: 3Commas utilizes sophisticated AI algorithms to provide automated trading bots and smart trading terminals, enhancing trading strategies and risk management across various market conditions on 16 major cryptocurrency exchanges. Its integration with TradingView and features like DCA, grid bots, and signal bots for strategy execution underscore its AI-centric approach to maximizing crypto trading efficiency and portfolio management.

3.6.2 ADADEX

Website: https://adadex.net/

One liner: ADADEX pioneers decentralized artificial intelligence and robot development in the metaverse, blending DeFi utilities with advanced AI capabilities.

Description: ADADEX merges decentralized finance (DeFi) with artificial intelligence (AI) by developing AI-driven agents and virtual robots for the metaverse, aimed at analyzing and executing trading strategies. Utilizing the ADEX token, it enables monetization of AI services, offering privacy, efficiency, and scalability in AI-enhanced DeFi solutions within the metaverse.

3.6.3 Bird Money

Website: https://www.bird.money/

One liner: Bird Money uses advanced machine learning to offer decentralized financial and security services.

Description: Bird Money leverages AI to enhance blockchain and DeFi applications by offering machine learning prediction tools for wallet behavior, enabling dynamic DeFi loan terms and investment strategies through a decentralized on-chain oracle.

This approach integrates AI analytics with cryptocurrency's decentralization principles, supporting community-driven governance and operations.

3.6.4 Blackbird AI

Website: https://www.blackbird.ai/

One liner: Blackbird AI leverages a high speed signals processing engine to analyze the finance sector, focusing on crypto and trading platforms.

Description: Blackbird AI integrates advanced signal processing with cryptocurrency analytics, providing traders and institutional investors real-time tools to detect market distortions, fabricated hype, and emergent trends, facilitating grounded decision-making in the volatile sectors of finance, including meme stocks and cryptocurrency. This technology aids in risk mitigation and capitalizes on market opportunities by uncovering hidden signals within financial data.

3.6.5 BrianKnows

Website: https://www.brianknows.org/

One liner: BrianKnows: Enhancing Web3 with AI-driven intents for seamless crypto transactions and smart contract operations.

Description: BrianKnows is an AI-driven cryptocurrency platform that simplifies web3 interactions through an AI interface, enabling users to conduct token transactions, bridge, and deploy smart contracts using voice prompts. It integrates with various protocols including Ethereum, Polygon, and Optimism, and provides a non-custodial, user-friendly experience for accessing web3 information and performing on-chain operations.

3.6.6 Crowdcent

Website: https://twitter.com/crowdcent?lang=en

One liner: CrowdCent utilizes crypto-economic incentives to decentralize investment management, closely aligned with Numerai.

Description: CrowdCent collaborates with Numerai to transform the hedge fund industry by utilizing AI-driven predictive models and blockchain technology, specifically through the use of Numerai's cryptocurrency, Numeraire (NMR), to stake on financial predictions. This partnership fosters a transparent investment ecosystem by merging crowdsourced data science with decentralized principles, aiming to enhance efficiency and openness in hedge fund management.

3.6.7 DAOstack

Website: https://daostack.io/

One liner: DAOstack leverages decentralized governance to build collaborative platforms and DeFi solutions.

Description: DAOstack provides a framework for decentralized autonomous organizations (DAOs), focusing on decentralized governance and collaboration tools since 2017, blending AI with crypto to enhance collective decision-making processes. The project, which includes the Arc governance framework and Alchemy DAO platform, has shifted focus to the Common platform, aiming to leverage open development to overcome financial challenges and advance market entry.

3.6.8 DefiLabs

Website: https://twitter.com/defilabs_farm?lang=en

One liner: DefiLabs: Revolutionizing finance with AI-driven DeFi solutions on the BSC chain.

Description: DefiLabs introduces an AI-driven decentralized platform on the Binance Smart Chain, focusing on optimizing DeFi investment strategies through AI-managed portfolios and high-yield liquidity pool investments. This project, secured by CertiK and Cyberscope audits, leverages AI for dynamic asset allocation and quantitative trading algorithms, aiming to significantly improve investment outcomes in the DeFi sector.

3.6.9 HeraAggregator

Website: https://twitter.com/heraaggregator?lang=en

One liner: AI-Powered Multichain Dex Aggregator offering the best swap rates.

Description: HeraAggregator, operating as Hera Finance, is an AI-enhanced DEX aggregator that optimizes swap transactions across multiple chains by intelligently linking AMM pools, offering features like best-rate access, efficient large trade handling, and a zero-slippage guarantee through its Hera Guard. Additionally, it incorporates an AI-driven Vault Manager for equitable swap fee distribution within its ecosystem, positioning itself as an accessible, high-value swapping platform for DeFi users, with the HERA token facilitating governance and additional benefits.

3.6.10 Hyperbolic

Website: https://hyperbolicprotocol.com/

One liner: Hyperbolic Protocol is revolutionizing DeFi through an AI-powered, 100% on-chain collateralized lending platform.

Description: Hyperbolic Protocol leverages the cutting-edge synergy of AI and blockchain technology to offer a unique DeFi solution. By automating variable APRs and using Uniswap V3 Time-Weighted Average Prices (TWAPs) for collateral valuation,

it provides an innovative approach for borrowers to leverage their crypto assets without selling. Investors can earn yield in ETH directly proportional to their HYPE token holdings, influenced by protocol fees generated from loan repayments and borrowing activities. Hyperbolic Protocol's AI-driven system aims to enhance yield maximization, ensuring protocol sustainability and growth by dynamically adjusting to market conditions.

3.6.11 Kryll

Website: https://www.kryll.io/

One liner: Kryll.io - Revolutionizing Crypto Trading with Automated Strategies.

Description: Kryll.io is an AI-enhanced platform that simplifies automated crypto trading by allowing users to create, share, and deploy strategies using a drag-and-drop interface with zero coding required, supported by a marketplace for renting and offering strategies across major exchanges. This approach underscores the synergy between AI tools and cryptocurrency trading, making advanced techniques more accessible.

3.6.12 Metafide

Website: https://www.metafide.io/

One liner: Metafide: Revolutionizing Blockchain Investments through AI-powered Analysis and Trading Tools.

Description: Metafide utilizes recurrent neural networks (RNNs) to enhance DeFi by providing AI-powered trading tools, risk management strategies, and impartial digital asset ratings. It aims to improve crypto investment decisions through deep analysis and features like trade automation and asset correlation advisory, striving for a more equitable Web3 economy.

3.6.13 Miti

Website: https://www.miti.ai/

One liner: Miti provides a natural language interface based on AI-based LLMs to all of crypto

Description: MITI offers a unique certificate in Blockchain at the community college level, integrating education on distributed ledgers, smart contracts, and cryptocurrencies, like Bitcoin and Ether, into its curriculum. This approach aims to prepare students for high-demand tech careers by merging the domains of AI and crypto technology.

3.6.14 Nabu DeFi Pricing

Website: https://www.nabu.xyz/

One liner: Nabu is revolutionizing NFT and DeFi markets by providing the most trusted NFT valuation engine.

Description: Nabu combines artificial intelligence and decentralized finance to provide a sophisticated NFT valuation engine, facilitating transparency and informed decision-making in the Web3 ecosystem for protocols, funds, and auditing firms. Their platform enhances DeFi services by offering insights for improved functionalities in loans, insurance, and rentals, supported by a team experienced in blockchain, data science, and technology.

3.6.15 NFPrompt

Website: https://nfprompt.io/

One liner: NFPrompt: Pioneering the intersection of AI and Crypto through AI-generated NFTs.

Description: NFPrompt combines AI and blockchain technology by enabling users to transform their creative ideas into NFTs through an AI-powered Prompt Artist Platform, which is a first in Web3. The platform supports the creation, trade, and transaction of AI-generated NFTs using its unique cryptocurrency, $cNFP, showcasing a novel integration of AI and crypto in digital art and Web3 adoption.

3.6.16 Orbofi AI

Website: https://www.orbofi.com/

One liner: Orbofi AI, the ultimate AI engine for web3, games, apps, and online communities, empowering AI-generated content creation on the blockchain.

Description: Orbofi AI facilitates the decentralization of AI-generated content production and monetization in web3, gaming, and online communities by allowing users to create, deploy, and monetize AI agents and assets, including 2D game assets and concept art, for crypto payments, supported by its native token, OBI. This ecosystem not only serves as a marketplace but also aims to enhance human creativity by enabling the establishment of virtual factories for consistent asset production, evidenced by its adoption by over 40,000 users in the web3 space.

3.6.17 Qna3

Website: https://qna3.ai/

One liner: Qna3: Revolutionizing knowledge sharing through AI-powered intelligence on the Blockchain.

Description: Qna3 is a platform that merges AI and cryptocurrency to enhance knowledge exchange in the blockchain realm, employing AI for efficient information curation and incentivization mechanisms like 'Computility Mining' and 'Vote to Earn'

to foster active community participation in the evolving landscapes of DeFi and related technologies.

3.6.18 RoboNet

Website: https://www.robonet.wiki/

One liner: RoboNet bridges AI and Crypto by offering a large-scale open database for robotic learning and AI-powered DeFi vault strategies.

Description: RoboNet integrates robotics AI with cryptocurrency by providing an extensive database for robotic learning and AI-driven DeFi strategies. It offers over 15 million video frames for AI training in robotics and employs advanced machine learning to automate and optimize financial strategies for a diverse range of assets, demonstrating the practical fusion of AI capabilities with cryptocurrency dynamics.

3.6.19 Sommelier

Website: https://www.sommelier.finance/

One liner: Sommelier is a non-custodial, cross-chain DeFi platform leveraging AI for automated yield optimization.

Description: Sommelier combines AI and cryptocurrency through a decentralized finance (DeFi) platform that optimizes yields and manages risks in non-custodial, cross-chain 'Cellars' using AI strategies and off-chain computation. It leverages its own blockchain and machine learning to execute complex DeFi strategies across blockchains efficiently, while emphasizing security, decentralization, and community-driven governance through SOMM token holder voting.

3.6.20 zkHoldem

Website: https://zkholdem.xyz/

One liner: zkHold'em: The decentralized Texas Hold'em experience leveraging blockchain and ZK technology for fairness and privacy.

Description: zkHold'em integrates blockchain and zero-knowledge proofs (zk-SNARKs) to deliver a decentralized, on-chain Texas Hold'em experience that ensures privacy and fairness through secure card shuffling and dealing, while enabling personalization via NFTs and ENS for an enhanced Web3 experience. The project leverages transparent, on-chain recording for auditability, fostering a community through social engagement and feature contributions.

3.7 App - Explorer

3.7.1 Blockexplorer

Website: https://blockexplorer.one/

One liner: BlockExplorer offers insights into blockchain transactions, enhancing crypto and AI data analysis capabilities.

Description: BlockExplorer leverages AI to enhance its blockchain transaction analysis platform, focusing on predictive analytics, transaction monitoring, and anomaly detection for users at the intersection of AI and crypto. This involves using real-time transaction data to enable AI applications to identify trends, enhance security, and optimize blockchain operations.

3.8 App - Gaming

3.8.1 Echelon Prime

Website: https://echelon.io/

One liner: Echelon Prime harnesses the power of AI and crypto to revolutionize Web3 gaming with its native PRIME token, advancing novel gaming models and economies.

Description: Echelon Prime merges artificial intelligence with blockchain to foster a Web3 gaming ecosystem, utilizing the PRIME token for seamless, blockchain-enhanced gaming experiences. The project illustrates AI's capability to revolutionize in-game economies and engagement, through its pioneering application in Parallel TCG, while asserting a strong market position in leveraging AI and big data across gaming, NFTs, and DeFi sectors.

3.8.2 Immutable X

Website: https://www.immutable.com/

One liner: Immutable X is revolutionizing Web3 gaming through scalable, secure, and eco-friendly blockchain technology.

Description: Immutable X, a layer-two scaling solution for NFTs on Ethereum utilizing zk-rollups, enhances scalability and reduces gas costs, enabling instant, zero-fee minting and trading of NFTs. This platform facilitates a seamless, eco-friendly ecosystem for gamers and developers, supporting secure and efficient creation, distribution, and exchange of digital assets, while integrating with major marketplaces to improve the Web3 gaming experience.

3.8.3 Inworld

Website: https://inworld.ai/

One liner: Inworld AI revolutionizes gaming with dynamically adaptive, AI-driven NPCs, enhancing player engagement and narrative immersion.

Description: Inworld AI develops AI-driven Non-Player Characters (NPCs) for games, utilizing advanced machine learning to enable these NPCs to adapt to player behaviors

and create more immersive, evolving virtual environments. This technology enhances game player engagement and developer revenue by creating deeper, interactive experiences, and has gained significant investment, reflecting its potential impact on the future of gaming and interactive entertainment within the crypto space.

3.9 App - Media

3.9.1 Theta Fuel

Website: https://www.thetatoken.org/

One liner: Theta Fuel powers the Theta blockchain—the engine for decentralized video streaming, AI processing, and entertainment.

Description: Theta Fuel (TFUEL) serves as the operational 'gas' for the Theta blockchain, playing a critical role in transaction execution and smart contract operations. Unique to Theta is its integration of blockchain with video streaming and AI, particularly through its Theta EdgeCloud platform, which utilizes a decentralized network to improve the delivery and cost-efficiency of video and AI services, thereby advancing the Web3 infrastructure landscape.

3.10 App - Metaverse

3.10.1 9VRSE

Website: https://9vrse.org/

One liner: 9VRSE - Bridging virtual worlds with blockchain technology for immersive gaming and content monetization.

Description: 9VRSE is an AI and cryptocurrency-driven creative studio that uses blockchain to build immersive, monetizable virtual experiences in a thematic metaverse, blending web3, gaming, 3D art, and AI. It focuses on secure, play-to-earn gaming and digital realms, underpinned by a commitment to transparency, community engagement through 'Kitty Krew', and legal protection for its developments.

3.10.2 Anima Virtuality

Website: https://anima.supply/

One liner: Anima Virtuality: Blending AI, AR, and Blockchain to create immersive gaming experiences with virtual characters.

Description: Anima Virtuality, Inc., combines AI, AR, and blockchain technologies to create immersive gaming experiences, epitomized by their AI companions, Onlybots. This initiative represents the confluence of AI and cryptocurrency in enhancing digital interaction and gaming, backed by a strong founding team and significant investment.

3.10.3 Hume

Website: https://www.wearehume.com/

One liner: Hume leverages AI & Crypto for creating and promoting virtual music 'metastars' in the Metaverse, redefining fan-artist relationships through NFTs.

Description: Hume integrates AI and cryptocurrency to reimagine the digital music scene by developing virtual 'metastars' and leveraging NFTs to enable direct fan engagement and investment in artists' success. This approach not only facilitates unique artist-fan relationships but also innovates in music production, consumption, and monetization by blending virtual performances with real-world value, highlighted by their $11.7 million Series A funding.

3.10.4 Ready Player Me

Website: https://readyplayer.me/

One liner: Ready Player Me is intertwining AI and Crypto to revolutionize virtual identity and asset ownership in the Metaverse.

Description: Ready Player Me employs AI, notably Stable Diffusion, for crafting customizable avatars that enhance user experiences across multiple virtual environments by offering AI-driven personalization. Integrating with Ethereum-based blockchain platforms like Somnium Space, the project enables these avatars to interact within tokenized virtual spaces, marking a distinct blend of AI and cryptocurrency technologies.

3.10.5 The Culture DAO

Website: https://www.theculturedao.com/

One liner: The Culture DAO leverages $CULTUR tokens to blend AI, blockchain, and the metaverse, creating a decentralized, collaborative space for virtual beings.

Description: The Culture DAO introduces a social token, $CULTUR, at the Virtual Beings Summit, blending AI-driven character development with blockchain's decentralized structures through a token-based DAO. This initiative facilitates collaborative creation and interaction with virtual beings within the metaverse, where token holders have both a stake and a say, aiming to democratize the evolution of digital personas and storytelling akin to a decentralized 'Web 3 Pixar'.

3.11 App - NFTs

3.11.1 Braindrops

Website: https://www.coingecko.com/en/nft/braindrops

One liner: Braindrops is a pioneering platform for AI-generated art leveraged by NFT technology.

Description: Braindrops is an NFT platform focused on AI-generated art, enabling artists to monetize their AI-driven creations by minting and selling them as NFTs. It aims to merge AI with blockchain technology by introducing dynamic generation, customization, and interaction with the art pieces.

3.11.2 Deep Value NFT

Website: https://deepnftvalue.com/

One liner: DeepNFTValue leverages machine learning to provide precise, unbiased NFT valuations, fostering transparency and efficiency in DeFi lending and trading.

Description: DeepNFTValue is an AI-driven startup leveraging machine learning to value non-fungible tokens (NFTs) across collections such as CryptoPunks, Bored Ape, and more, enhancing accuracy in DeFi trading and lending. With $4 million in seed funding, it focuses on reducing biases in NFT valuation by expanding its AI capabilities, including significant investment in GPU hardware for deep-learning computations within the cryptocurrency market.

3.11.3 Onaji

Website: https://onaji.io/

One liner: Onaji leverages AI for blockchain asset market intelligence and art discovery in the web3 ecosystem.

Description: Onaji employs machine learning to analyze blockchain data, offering trader intelligence technology for predicting market trends in digital assets, including cryptocurrencies and NFTs, thereby aiding in investment decisions. Additionally, it uses AI to enhance art discovery within the blockchain, particularly in the Tezos ecosystem, by providing search and recommendation tools, bridging the gap between creators and the community in the web3 space.

3.12 App - Provenance

3.12.1 Numbers Protocol

Website: https://www.numbersprotocol.io/

One liner: Numbers Protocol is revolutionizing provenance and monetization of digital media in blockchain through AI integration.

Description: Numbers Protocol leverages AI and blockchain technology to improve digital media provenance and facilitate content monetization, enabling creators and AI-generated content makers to efficiently verify and monetize their works on its proprietary blockchain using the NUM token for activities such as NFT minting, content-history tracking, and royalty distribution, while incentivizing engagement through staking pools and rewards.

3.13 App - Wallet

3.13.1 Bluwhale

Website: https://www.linkedin.com/company/bluwhaleai

One liner: Bluwhale is revolutionizing Web3 by integrating AI with blockchain for personalized user experiences and data monetization.

Description: Bluwhale integrates AI with blockchain to offer a decentralized personalization protocol for DApp developers and users, utilizing AI and contextual data to enhance user engagement and monetization while ensuring data privacy across blockchain networks. Their system, featuring predictive user profiling and AI insights, creates detailed wallet profiles, facilitating a secure, personalized, and monetizable Web3 experience.

3.14 App with category: DeFi

3.14.1 Audius

Website: https://audius.co/

One liner: Audius is a decentralized music streaming service integrating crypto to empower artists and listeners.

Description: Audius utilizes blockchain and its native token $AUDIO to offer a decentralized music streaming platform, enabling direct artist-audience connections and content monetization through engagement and exclusive access via artist tokens. This approach removes intermediaries, ensures content immutability through a secured network of node operators, and promotes distributed value and governance within its community.

3.15 Data

3.15.1 0x0

Website: https://coinmarketcap.com/currencies/0x0-ai-ai-smart-contract/

One liner: 0x0.ai combines advanced AI technologies with crypto to revolutionize privacy, security, and income in DeFi.

Description: 0x0.ai integrates artificial intelligence, including machine learning and algorithmic analysis, with cryptocurrency to improve privacy, security, and DeFi applications, focusing on smart contract auditing and the application of zero-knowledge proofs. It innovates with a revenue-sharing model, redistributing generated revenue to token holders, aiming for a secure, private, and incentivized financial ecosystem.

3.15.2 0xscope

Website: https://www.0xscope.com/

One liner: 0xScope - The AI Data Layer for Web3 AI Applications.

Description: 0xScope develops an AI-driven data layer tailored for Web3 applications, focusing on enhancing data exchange across Web2 and Web3 platforms through technologies like knowledge graphs and decentralized storage. This initiative, supported by strategic investments from entities like OKX Ventures, facilitates cross-chain integration and privacy computing, while its products, such as 'Scopechat' and 'Scopescan', showcase its dedication to merging AI capabilities with blockchain technology to serve a broad user base including over 311 B2B clients and 237K individual users.

3.15.3 AIOZ

Website: https://aioz.network/

One liner: Decentralized AI-powered Content Delivery and Computation

Description: AIOZ Network integrates AI and blockchain through its decentralized content delivery network (dCDN), offering decentralized storage, streaming, and AI computation by harnessing spare computing resources worldwide. This setup not only facilitates web3 AI applications and media delivery but also plans for the expansion into decentralized AI as a Service, showcasing a practical fusion of AI and crypto technologies to enhance efficiency and accessibility in digital content and computation.

3.15.4 Aleph.im

Website: https://aleph.im/

One liner: Decentralizing AI and data management on the blockchain.

Description: Aleph.im combines AI with blockchain technology to offer decentralized data management, including encrypted storage and serverless computing, through a network that allows DApp development without centralization. Its key feature, Libertai.io, leverages this infrastructure for decentralized AI processing, emphasizing data privacy and system resilience.

3.15.5 Algovera

Website: https://www.algovera.ai/

One liner: Algovera is a decentralized AI research organization and agency pioneering in Web3 innovations.

Description: Algovera is a decentralized AI research and development agency that utilizes blockchain technology to create an open and transparent platform for AI innovation, focusing on the empowerment of developers through direct access to resources

and funding via cryptocurrency. This integration of AI with Web3 technologies enables the exploration of new AI-driven solutions, promoting a decentralized technological future.

3.15.6 Arweave

Website: https://www.arweave.org/

One liner: Arweave offers a decentralized network for permanent and sustainable data storage, leveraging blockchain technology.

Description: Arweave integrates AI and cryptocurrency through its decentralized network designed for perpetual data storage, leveraging its native currency, AR, to reward miners. This facilitates a durable, censorship-resistant digital repository essential for AI applications that depend on large datasets for training and inference, positioning it as a significant contributor to the AI x Crypto ecosystem.

3.15.7 Axiom

Website: https://www.axiom.xyz/

One liner: Axiom utilizes ZK proofs to enable trustless on-chain queries and computations over Ethereum's history for data-rich and dynamic decentralized applications.

Description: Axiom integrates AI and crypto on the Ethereum blockchain by allowing developers to craft smart contracts that compute and verify the blockchain's entire history using zero-knowledge proofs. This enables the development of advanced DeFi applications, custom oracles, and loyalty programs by utilizing historical data without external inputs. Moreover, Axiom leverages this technology for AI tasks like validating online content and identifying deepfakes, promoting data integrity and trustless processing, thus positioning itself at the forefront of secure and intelligent blockchain applications.

3.15.8 BitsCrunch

Website: https://bitscrunch.com/

One liner: Empowering the NFT ecosystem with AI-driven insights and security.

Description: BitsCrunch is a project integrating AI with cryptocurrency to offer a decentralized NFT analytics and forensics platform, aiming to provide reliable data for making informed investment decisions and detecting fraud in the NFT market. Its API facilitates the development of trustworthy NFT and DeFi applications, while a community-driven model enhances system reliability across various blockchain infrastructures, promoting a secure and transparent NFT marketplace.

3.15.9 Bittensor

Website: https://bittensor.com/

One liner: Bittensor: Decentralizing AI Training and Inference via Blockchain Technology.

Description: Bittensor combines artificial intelligence (AI) and blockchain to create a blockchain-based network for decentralized data processing and training of AI models, with these models rewarded in Bittensor's native cryptocurrency, TAO, for their valuable contributions. This system facilitates transparent, decentralized development and use of AI, ensuring equitable distribution of rewards and aligning with principles of open-source and decentralized governance.

3.15.10 Bittorrent

Website: https://www.bittorrent.com/token/btt/

One liner: BitTorrent tokenizes the world's largest decentralized file sharing protocol, introducing BTT for a crypto-powered web.

Description: BitTorrent integrates blockchain technology with its decentralized file-sharing protocol through the BTT token on the Tron blockchain, rewarding user participation with incentives for seeding and bandwidth sharing, thereby enhancing network efficiency and illustrating a practical application of AI x Crypto in promoting a decentralized internet ecosystem.

3.15.11 Blackbird.AI

Website: https://www.blackbird.ai/

One liner: Blackbird.AI leverages advanced AI to empower organizations against narrative-driven risks, potentially including crypto-related misinformation and disinformation.

Description: Blackbird.AI utilizes artificial intelligence to measure and counteract misinformation threats, offering tools like the Constellation Dashboard and Risk Intelligence Engine, which are crucial for maintaining integrity in the crypto sector by scoring and predicting the impact of narrative-based threats. Their technology, though not crypto-specific, is essential for protecting reputation and financial stability in the digital assets market.

3.15.12 Blockless

Website: https://blockless.network/

One liner: Blockless - Launching a new era of modular, network neutral applications on blockchain, leveraging AI for enhanced security and efficiency.

Description: Blockless integrates blockchain technology with AI to facilitate the secure and efficient development of decentralized Network Neutral Applications (n-nApps), offering a Modular Application Architecture for enhanced full-stack decentralization. Its unique features, including automated node orchestration, a WASM-based runtime, and AI-based model inference, support a wide range of programming languages and major L1/L2 blockchains, significantly reducing development time to under three weeks while advancing the potential for application versatility at the intersection of crypto and AI.

3.15.13 C2PA

Website: https://c2pa.org

One liner: C2PA stands for the Coalition for Content Provenance and Authenticity, leveraging crypto and AI for digital media verification.

Description: The Coalition for Content Provenance and Authenticity (C2PA) employs cryptographic techniques and AI-driven analysis to establish standards for certifying the provenance of digital media, leveraging Distributed Ledger Technology (DLT), digital signatures, and NFTs to ensure content integrity and counter misuse. This initiative demonstrates the intersection of AI and blockchain technology in enhancing data validation and combating misinformation by making digital media traceable and verifiable across platforms.

3.15.14 Ceramic

Website: https://ceramic.network/

One liner: Ceramic is a decentralized data network bridging AI with crypto, enabling scalable, interoperable, and verifiable data management.

Description: Ceramic merges blockchain and AI by offering a decentralized ledger for high-volume data management, pairing the trust of blockchain with efficiency akin to traditional databases. It facilitates AI-driven application development through verifiable data storage and peer-to-peer streaming, using ComposeDB and GraphQL for enhanced data interaction, thereby supporting trusted and scalable data access crucial for AI training and inference in decentralized settings.

3.15.15 Chainbase

Website: https://chainbase.com/

One liner: Chainbase, the seamless bridge for Web3 data infrastructure, integrating AI for smarter crypto applications.

Description: Chainbase streamlines Web3 development by offering a browser-based platform for real-time blockchain data integration, processing, and AI-driven analytics

across multiple ecosystems. It automates indexing and querying to facilitate efficient AI model development and application in the crypto domain, significantly cutting down on development time.

3.15.16 ChainLink

Website: https://chain.link/

One liner: Chainlink: Powering Smart Contracts with Real-World Data

Description: Chainlink is a decentralized oracle network that connects real-world data to smart contracts on various blockchains, utilizing a network of decentralized oracles to enable secure, manipulation-resistant data integration for complex executions in dApps across DeFi, insurance, and gaming. This facilitates the accurate and reliable operation of smart contracts, crucially linking blockchain technology with external systems.

3.15.17 Clore.ai

Website: https://clore.ai/

One liner: Clore.ai - Revolutionizing access to distributed supercomputing for AI and cryptocurrency endeavors.

Description: Clore.ai offers a distributed supercomputing platform that merges AI and cryptocurrency technologies, providing GPU computing for AI training, video rendering, and crypto mining. It features a user marketplace for computing resources, utilizes a Proof-of-Work blockchain called Clore Coin to reward participants, and emphasizes security, cost efficiency, and environmental sustainability without relying on premines or ICOs for funding.

3.15.18 Crust Network

Website: https://crust.network/

One liner: Crust Network offers decentralized storage solutions, enhancing Web3′s data storage capabilities with IPFS integration.

Description: Crust Network combines blockchain's decentralized capabilities with AI-enhanced algorithms on the Substrate framework to offer secure, scalable storage solutions for both Web 3.0 and 2.0, incorporating an incentive layer for IPFS to ensure data integrity and privacy across various applications like DApp hosting and NFT storage.

3.15.19 CryptoIndex

Website: https://www.cdc.gov/parasites/crypto/index.html

One liner: Revolutionizing crypto investment through AI-driven analysis and diversified exposure.

Description: CryptoIndex utilizes artificial intelligence to intelligently analyze and derive insights from the cryptocurrency market, facilitating a diversified investment strategy for both individual and institutional investors through the Bitwise 10 Crypto Index Fund (BITW). This fund, emphasizing AI's role in cryptocurrency investment strategies, focuses on the top 10 cryptocurrencies by market capitalization, ensuring monthly rebalancing and risk monitoring to optimize performance.

3.15.20 Fact Protocol

Website: https://fact.technology/

One liner: Decentralized fact-checking and misinformation combat platform leveraging AI and blockchain technologies.

Description: Fact Protocol combines AI and blockchain technology to establish a decentralized system for combating fake news, utilizing AI-driven fact-checking and blockchain's transparency to ensure information integrity. It leverages token incentives for community participation in verifying content, supporting scalability, privacy, and multi-chain functionality, while introducing governance features to enable community-driven project direction via FACT DAO.

3.15.21 Filecoin

Website: https://filecoin.io/

One liner: Decentralized storage network for humanity's most important information.

Description: Filecoin is a blockchain-based network offering decentralized data storage, utilizing its cryptocurrency, FIL, to incentivize storage space sharing, aiming to enhance web security and efficiency by reducing dependency on centralized storage solutions. It integrates with the Interplanetary File System (IPFS) to optimize data storage and retrieval processes.

3.15.22 Flare Network

Website: https://flare.network/

One liner: Flare Network leverages AI and blockchain to deliver decentralized, secure, and scalable data solutions.

Description: Flare Network integrates AI with cryptocurrency through its Flare Time Series Oracle (FTSO) and State Connector, offering decentralized data feeds and secure, trustless data usage from various blockchains and the internet within smart contracts. This facilitates the development of AI-driven decentralized applications (dApps) and services, ensuring data integrity and broadening blockchain technology's utility and scalability.

3.15.23 Fluence

Website: https://fluence.network/

One liner: Decentralized serverless platform & computing marketplace, leveraging blockchain for Web 3.0.

Description: Fluence is a decentralized, serverless computing platform that facilitates the development of AI-driven applications by merging blockchain technology with artificial intelligence. It aims to democratize computing resources through a secure, scalable ecosystem, underscored by a token economy, which encourages innovation at the intersection of AI and crypto, enhancing the efficiency and security of decentralized applications.

3.15.24 Flux

Website: https://runonflux.io/

One liner: Flux enables decentralized Web 3.0 applications, combining AI compute solutions, cryptocurrency, and a global computational network.

Description: Flux operates a decentralized computational network that integrates AI and crypto, utilizing its POW cryptocurrency, FLUX, to support the development and deployment of DApps with an emphasis on AI compute solutions. It introduces FluxCore for AI computations and employs Proof of Useful Work to enhance GPU efficiency for AI tasks, within an infrastructure boasting over 15,000 nodes and extensive computational resources, marking it as a notable bridge between blockchain and AI technologies.

3.15.25 Fraction AI

Website: https://fractionai.xyz/

One liner: Supercharge your AI with human-powered data at a fraction of the cost.

Description: Fraction AI streamlines AI model training by providing high-quality human-generated data, integrating the efficiency of cryptocurrency mechanisms for more effective and cost-efficient AI development across multiple sectors. This approach accelerates AI adoption and refinement, particularly within the crypto industry, by ensuring the availability of optimized datasets.

3.15.26 Grass

Website: https://www.getgrass.io/

One liner: Grass: Revolutionizing AI Training with Decentralized Web Data Accessibility.

Description: Grass by Wynd Network is a decentralized web scraping platform that utilizes unused internet bandwidth from users to generate datasets for AI training, ensuring fair compensation through network ownership. This system leverages residential IP addresses to overcome data scraping restrictions, supporting efficient AI research while prioritizing security and ethical data collection practices.

3.15.27 Humans.ai

Website: https://humans.ai/

One liner: Empowering the creation and governance of ethical AI on a scalable blockchain platform.

Description: Humans.ai merges AI and blockchain technology to enhance human potential ethically, leveraging a Proof of Human mechanism and AI tools for user-generated projects, including the pioneering use of AI NFTs for idea security and governance on the blockchain. The ecosystem allows for diverse applications such as metaverse avatar creation and utilizes the $HEART token to fund AI innovations, all within a framework emphasizing community, transparent AI use, and ethical standards.

3.15.28 Innodata

Website: https://innodata.com/

One liner: Innodata leverages artificial intelligence and human expertise to solve complex data engineering challenges, fueling data-centric AI initiatives.

Description: Innodata integrates AI with crypto by providing AI-enabled software platforms and managed services, focusing on generating AI training data and building, maintaining, and augmenting AI and ML models. Their approach combines human-operated tools with active learning, enhancing operational efficiencies and impacting the AI lifecycle across different domains.

3.15.29 IPFS

Website: https://ipfs.tech/

One liner: IPFS is revolutionizing the way data is stored and accessed, offering a decentralized solution for resilient, efficient, and censorship-resistant data sharing.

Description: IPFS integrates AI and crypto by providing a decentralized storage solution that ensures data integrity for AI applications, facilitating secure and efficient data sharing and model training. Its decentralized nature supports AI-driven DeFi applications, improves resilience against censorship, and promotes cooperative model training and sharing, enhancing both AI and crypto ecosystems.

3.15.30 JACKAL

Website: https://www.jackalprotocol.com/

One liner: Jackal Protocol leverages blockchain technology to offer decentralized, on-chain data storage with enhanced interoperability and security.

Description: Jackal Protocol employs blockchain technology for decentralized data storage, emphasizing secure, on-chain management of data permissions and ownership through an application-specific blockchain. Its architecture promotes interoperability across ecosystems, enhancing the application of AI in crypto by ensuring secure data access and transfer for various uses, including DeFi.

3.15.31 Koii Network

Website: https://www.koii.network/

One liner: Koii Network revolutionizes AI x Crypto through decentralized, community-powered data and compute sharing.

Description: The Koii Network integrates AI and blockchain technology to create a cost-efficient, distributed cloud environment for AI model training and digital content creation. It uses attention-tracking NFTs for fair remuneration, leverages consumer devices for a global compute-sharing marketplace, and employs ORCA technology for supporting diverse frameworks and languages, establishing an eco-conscious and creator-friendly platform for AI development in the crypto domain.

3.15.32 Kudos

Website: https://www.cudos.org/

One liner: CUDOS is a scalable blockchain that integrates cloud and blockchain for AI and Web3 applications.

Description: CUDOS integrates cloud computing with blockchain through a decentralized platform on the Cosmos ecosystem, offering scalable solutions for AI, the metaverse, and Web3 via a proof-of-stake Layer 1. It enables smart contracts, NFT deployment, and Oracle connections through a vast user network, emphasizing sustainability with its 100% carbon-neutral operation.

3.15.33 Lasso Labs

Website: https://twitter.com/lasso_labs?lang=en

One liner: Lasso Labs is revolutionizing NFT utility by building data products for the next generation of the internet, leveraging AI and crypto.

Description: Lasso Labs has secured $4.2M to enhance NFT functionality and integration into digital commerce by employing AI and blockchain technology. Their work

focuses on evolving NFTs beyond speculative assets, using data analytics to embed these digital objects more fully into digital identities and marketplaces.

3.15.34 Matrix AI Network

Website: https://www.matrix.io/

One liner: Matrix AI Network is pioneering the integration of AI and blockchain, stepping into a future synergized through neuroscience with Matrix 3.0.

Description: Matrix AI Network combines artificial intelligence with blockchain technology to enhance transaction speed, security, and ease of use, notably with the introduction of Matrix 3.0, which integrates AI, blockchain, and neuroscience for a self-evolving AI platform that supports secure data handling and transparent sharing, along with democratizing access to on-chain AI resources and pioneering Avatar Intelligence (AvI), where digital and human consciousness may converge.

3.15.35 nil; Foundation

Website: https://nil.foundation/

One liner: nil; Foundation focuses on elevating Ethereum's scalability and security through zero-knowledge proofs.

Description: #ERROR!

3.15.36 NumerAI

Website: https://numer.ai/

One liner: Numerai: Harnessing the power of AI and blockchain for a decentralized stock market prediction platform.

Description: Numerai merges AI with blockchain through an Ethereum-based platform, decentralizing financial forecasting by rewarding data scientists with Numeraire (NMR) tokens for contributing accurate machine learning models for stock market predictions. Founded in 2015, it operates as a global data science competition aimed at enhancing stock market hedging strategies via collective intelligence.

3.15.37 Oasis Network

Website: https://oasisprotocol.org/

One liner: Oasis Network is a privacy-enabled, scalable layer-1 blockchain designed for DeFi, NFTs, and a secure, privacy-focused Web3.

Description: Oasis Network is a layer-1 blockchain platform designed for Web3 applications, focusing on privacy, scalability, and low costs, and supports diverse projects like DeFi, GameFi, NFTs, and the Metaverse. It integrates advanced privacy features and smart contract flexibility, working with partners like Meta and BMW, to offer a

secure and efficient environment for developing private or public decentralized applications, aiming to advance data governance and privacy in Web3 ecosystems.

3.15.38 OctaSpace

Website: https://octa.space/

One liner: OctaSpace leverages blockchain to revolutionize distributed computing, data storage, and AI model training.

Description: OctaSpace is a distributed computing platform merging AI and crypto technology to offer cloud storage, decentralized VPN services, and a CPU/GPU resource rental marketplace, catering to AI model training, gaming, and IoT, with a focus on data privacy through its decentralization. It operates on a unique Layer 1 EVM-compatible blockchain that combines Proof-of-Work and Proof-of-Authority consensus mechanisms, aiming to improve network efficiency and scalability while promoting the reuse of idle computing power and enhancing its ecosystem with AI/ML tools and blockchain expansions to bolster the utility of its $OCTA coin.

3.15.39 OpenOrigins

Website: https://www.openorigins.com/

One liner: Empowering organizations to secure and validate digital media against synthetic alteration through advanced cryptography and AI-driven verification.

Description: OpenOrigins combines artificial intelligence and blockchain technology to enhance the security and validation of digital media, ensuring the authenticity and integrity of content across various industries by utilizing a blockchain-secured, tamper-proof database to protect against synthetic content and verify the origin and unaltered status of media archives.

3.15.40 Optic.xyz

Website: https://twitter.com/Optic_xyz

One liner: Optic is revolutionizing the NFT space by enhancing trust and safety using AI technology.

Description: Optic.xyz is developing an AI engine to authenticate and secure NFTs by processing millions of tokens daily for potential counterfeit or copyright issues, aiming to enhance trust and safety in the digital marketplace. With $11 million in seed funding, the company provides a Marketplace Moderation tool for platforms like OpenSea and plans to offer a public API and additional tools for Web3 developers, NFT creators, and collectors, targeting systemic fraud and trust challenges in the ecosystem.

3.15.41 Oraclize

Website: www.oraclize.it

One liner: Oraclize serves as a bridge to securely integrate real-world data into blockchain smart contracts, enhancing DeFi and other blockchain applications with external APIs.

Description: Oraclize serves as a bridge between blockchain oracles and decentralized finance (DeFi) by integrating off-chain data, such as financial market APIs and weather information, into on-chain smart contracts. This integration ensures that DeFi platforms and other blockchain applications can operate with real-time, accurate data without sacrificing security or decentralization, thus enabling a variety of trustless interactions with real-world data.

3.15.42 Oraichain

Website: https://coinmarketcap.com/currencies/oraichain-token/

One liner: Oraichain - the world's first AI-powered oracle and ecosystem for blockchains, striving to become the foundational AI Layer 1.

Description: Oraichain integrates artificial intelligence with blockchain technology to serve as the first AI-powered oracle, enhancing smart contracts and decentralized applications (DApps) by supplying AI-verified data and aiming to become a pivotal AI-centric Layer 1 blockchain. It offers a diverse ecosystem including AI-driven services such as an AI Marketplace, on-chain Verifiable Random Function (VRF), and AI-enabled NFT generation, aimed at advancing the functionality and interoperability of smart contracts and DApps.

3.15.43 Origin Trail

Website: https://origintrail.io/

One liner: OriginTrail - Empowering verifiable, AI-ready Knowledge Assets in a decentralized knowledge graph.

Description: OriginTrail integrates blockchain and AI to offer a Decentralized Knowledge Graph (DKG) that ensures the integrity and provenance of data, enhancing AI capabilities by providing access to a verified information network. This amalgamation aims to improve the efficiency and reliability of AI agents across various industries by establishing a secure, trustworthy foundation for data creation, verification, and querying.

3.15.44 Pond Discoveries

Website: https://cryptopond.xyz/

One liner: AI Calculated Token Recommendation

Description: Pond's model has the unique ability to comprehend the flow of blockchain data, guiding you seamlessly from chaos to order by providing you with intelligently compressed information directly from the on-chain world.

3.15.45 Powerloom

Website: https://powerloom.io/

One liner: Powerloom: The Composable Data Network revolutionizing Web3 with Decentralized Finance (DeFi) and Data Availability.

Description: Powerloom is developing a blockchain and AI-powered composable data network aimed at enhancing Web3 applications by enabling developers to build sophisticated data applications and improve DeFi experiences through decentralized, consensus-based data snapshotting and aggregation. Its alpha-stage product, Boost, facilitates the customization and access to datapoints across EVM chains, promoting a shift towards decentralized data accuracy and security, supported by a recent $3.1M funding round to advance cross-chain logic and applications.

3.15.46 Reclaim Protocol

Website: https://www.reclaimprotocol.org/

One liner: Reclaim Protocol - Revolutionizing identity and reputation verification through cryptographic proofs without compromising user privacy.

Description: Reclaim Protocol integrates AI and blockchain technology to provide a secure, private verification system for users' identities and reputations online, utilizing zero knowledge proofs and digital signatures to facilitate decentralized applications such as DeFi and KYC without compromising sensitive information. Its SDK enables connections from over 200 websites, enhancing user experiences through verified online behavior while preserving privacy.

3.15.47 SapienAI

Website: https://www.linkedin.com/in/trevorkoverko

One liner: SapienAI leverages the intersection of AI and crypto to innovate within the Web3 space.

Description: SapienAI, led by Trevor Koverko, specializes in merging artificial intelligence with cryptocurrency to advance blockchain technologies and improve the crypto ecosystem, emphasizing the creation of secure, transparent, and efficient solutions at the intersection of AI and Web3.

3.15.48 Sia

Website: https://coinmarketcap.com/currencies/siacoin/

One liner: Decentralized cloud storage platform using blockchain technology.

Description: Sia combines AI and blockchain to provide a decentralized cloud storage platform that uses its native Siacoin for transactions, focusing on privacy and security by distributing encrypted files across multiple nodes, which ensures data safety, availability, and censorship resistance.

3.15.49 Sphere AI Agents

Website: https://twitter.com/RollmanMining/status/1750240075433083185

One liner: Revolutionizing the Crypto Sphere through Advanced AI Technology.

Description: Fetch.AI integrates AI with cryptocurrency through a Multi-Agent System (MAS), enabling Autonomous Economic Agents (AEA) to improve task efficiency and reduce costs by removing the need for intermediaries, with the goal of automating and transforming various sectors. Meanwhile, Sphere 3D's acquisition of NuMiner NM440 machines signals a move towards becoming a leading carbon-neutral bitcoin miner, emphasizing the convergence of AI and cryptocurrency in advancing sustainable, decentralized finance and secure blockchain technology.

3.15.50 Spice AI

Website: https://spice.ai/

One liner: Spice.ai accelerates the development of next-gen intelligent software with a focus on leveraging web3 data through AI.

Description: Spice.ai streamlines the development of AI-driven applications in the crypto and web3 sectors by providing a backend-as-a-service that includes time-series data management, custom ETL processes, and ML model training and inferencing, all optimized for petabyte-scale data in web3 development. It reduces the need for deep data science expertise, facilitating the creation of innovative apps across DeFi, security, and AI agents, thereby advancing the integration of AI with blockchain technologies.

3.15.51 Storj

Website: https://www.storj.io/

One liner: Storj is revolutionizing cloud storage with a decentralized, secure, and environmentally sustainable platform.

Description: Storj is an open-source, decentralized cloud storage platform, using blockchain technology to enhance security, privacy, and cost-efficiency. Its global network of distributed nodes encrypts and disperses data, offering a sustainable and secure storage solution particularly suited for AI development, with its S3 compatibility emphasizing its utility in the AI x Crypto domain.

3.15.52 The Graph

Website: https://thegraph.com/

One liner: The Graph is a decentralized protocol for indexing and querying data from blockchains, facilitating the creation of decentralized applications with efficient data retrieval.

Description: The Graph integrates AI to index blockchain data efficiently, creating open APIs (subgraphs) that facilitate the development of responsive DApps across various sectors like DeFi, NFTs, and DAOs. This decentralized platform leverages a global community for network maintenance and data integrity, enhancing access and usability of blockchain data for developers.

3.15.53 WorldCoin

Website: https://worldcoin.org/

One liner: WorldCoin integrates AI to create a privacy-preserving global identity network, enhancing access to a human-centric global economy.

Description: WorldCoin combines artificial intelligence and blockchain to establish a global identity and financial network through a privacy-centric AI-enabled verification system (World ID), employing multispectral sensors and zero-knowledge proofs to ensure anonymous yet unique human identification. It aims to democratize global economic participation by distributing WLD tokens, exhibiting over 3 million sign-ups across 120 countries, to integrate human identity into digital infrastructure securely.

3.16 DeFi

3.16.1 Colony

Website: https://www.colonylab.io/

One liner: Colony Lab: Accelerating the Avalanche Ecosystem Through Decentralized Community-Driven Fundraising and Liquidity.

Description: Colony Lab is a platform in the Avalanche ecosystem that combines AI and crypto to enhance decentralized fundraising and investment opportunities for early-stage projects. It focuses on providing capital, ecosystem support, and exposure, while also engaging the community through seed investments and yield-generating products, incorporating AI to optimize investment strategies and project support.

3.16.2 Glif

Website: https://www.glif.io/en

One liner: GLIF - Decentralized finance hub for the Filecoin ecosystem, facilitating liquid Filecoin leasing and offering a suite of Filecoin infrastructure tools.

Description: GLIF merges AI and crypto within the Filecoin network, offering decentralized finance solutions through tools such as a web wallet, multisig wallet, Filecoin nodes, an autonomous notary service, and an FEVM-enabled block explorer. Its flagship, GLIF Pools, introduces DeFi protocols to the Filecoin ecosystem, focusing on FIL staking for increased liquidity and supporting the expansion of a decentralized storage-based open data economy through smart contracts.

3.16.3 KingDefi

Website: https://coinmarketcap.com/currencies/kingdefi/

One liner: KingDeFi - A dual-role DeFi project offering analytics, monitoring, and optimized yield farming on BSC and Solana.

Description: KingDeFi merges AI-driven analytics and DeFi yield optimization on the Binance Smart Chain and Solana networks, providing users with tools like a liquidity pool search engine and portfolio tracking while enabling earning through KROWN tokens via liquidity provision and staking. Its standout feature is an automated recommendation system for yield-generating platforms based on wallet contents, supported by a comprehensive dashboard of key DeFi metrics, thus serving as an analytical bridge for effective DeFi strategy execution.

3.16.4 Mozaic

Website: https://mozaic.finance/

One liner: Empowering Yield Farming and Liquidity Strategies with AI

Description: Mozaic integrates AI into DeFi, automating and optimizing yield farming and liquidity management across multiple blockchains using the LayerZero protocol, improving user experience and return potential by adapting to market changes in real time, while maintaining a focus on security and sustainability through collaborations with key technology and financial partners.

3.16.5 Nexus Mututal

Website: https://nexusmutual.io/

One liner: Nexus Mutual uses blockchain to democratize insurance, leveraging collective risk assessment and management in the DeFi space.

Description: Nexus Mutual utilizes AI and blockchain technology to offer a decentralized mutual insurance protocol within the DeFi sector, focusing on smart contract security and risk assessment using staking pools and NXM tokens to mitigate crypto-related risks. Its effectiveness in managing and responding to incidents like the Euler Finance hack underscores its significant role in enhancing DeFi security.

3.16.6 Rinzo

Website: https://twitter.com/rinzo?lang=en

One liner: The Renzo Protocol augments Ethereum staking by securing Actively Validated Services for enhanced yields.

Description: The Renzo Protocol aims to enhance Ethereum's staking ecosystem by interfacing with EigenLayer, focusing on elevating yield opportunities through Actively Validated Services within DeFi and Liquid Staking on the Ethereum Mainnet. It utilizes a blend of blockchain technology and AI to optimize and profitably streamline staking operations.

3.16.7 SpacePort

Website: https://www.spaceport.xyz/

One liner: SpacePort - The democratized open IDO platform launching crypto projects to success.

Description: SpacePort combines AI and blockchain technology to offer a secure and open IDO platform, enhancing project fundraising by automating launch processes with a Smart Contract Factory and providing comprehensive pre-launch support. It safeguards investors against market risks and ensures equitable IDO access through a fully on-chain approach, aiming at transparency and fairness in the DeFi ecosystem.

3.16.8 Upshot

Website: https://upshot.xyz/

One liner: Revolutionizing DeFi with decentralized, scalable, and self-improving AI network powered by Proof of Alpha.

Description: Upshot integrates artificial intelligence with decentralized finance (DeFi) via a decentralized AI network that uses a 'Proof of Alpha' protocol to encourage the contribution and use of collective intelligence, enhancing the network's self-improvement capabilities. Its platform offers AI-driven DeFi applications and ensures secure, unbiased outcomes with zkML and cryptographic verification, promoting global innovation in AI within DeFi.

3.17 Hardware

3.17.1 Akash

Website: https://akash.network/

One liner: Akash Network is an open-source Supercloud for decentralized cloud computing, combining AI & Crypto to innovate the future of cloud infrastructure.

Description: Akash Network offers a decentralized, blockchain-based cloud computing marketplace that uses Kubernetes and Cosmos for secure and efficient application hosting. It notably reduces costs (up to 85% lower than traditional providers) through a 'reverse auction' pricing system and facilitates distributed machine learning, highlighting its utility in the intersection of AI and Crypto.

3.17.2 Akash

Website: https://akash.network/

One liner: Akash Network is an open-source Supercloud for decentralized cloud computing, combining AI & Crypto to innovate the future of cloud infrastructure.

Description: Akash Network offers a decentralized, blockchain-based cloud computing marketplace that uses Kubernetes and Cosmos for secure and efficient application hosting. It notably reduces costs (up to 85% lower than traditional providers) through a 'reverse auction' pricing system and facilitates distributed machine learning, highlighting its utility in the intersection of AI and Crypto.

3.17.3 ATOR

Website: https://www.ator.io/

One liner: ATOR integrates AI with crypto to enhance online anonymity, offering rewards for secure network participation.

Description: The ATOR Protocol integrates AI with blockchain to enhance the security and efficiency of the Tor network by rewarding network participants with cryptocurrency for their uptime, using a Proof-of-Uptime model to encourage a robust and secure network. By deploying AI for network optimization and introducing user-friendly hardware like the ATOR Router Hotspot, the project aims to expand its global anonymity network, leveraging AI and on-chain incentives to improve digital privacy and network performance.

3.17.4 Auradine

Website: https://auradine.com/

One liner: Auradine is revolutionizing the blockchain landscape through innovative Bitcoin mining systems and advanced AI integration.

Description: Auradine focuses on developing efficient infrastructure solutions at the crypto and AI intersection, notably through its Teraflux™ Bitcoin miners, which mark a significant improvement in energy efficiency and transaction speeds due to a pioneering four-nanometer silicon transistor. Beginning with a substantial $81 million fundraising without a product, the company emphasizes innovation in combining AI and blockchain to enhance web infrastructure security, privacy, and efficiency.

3.17.5 exaBITS

Website: https://www.exabits.ai/

One liner: Accelerating crypto and AI applications through cost-effective, high-performance GPU compute.

Description: ExaBITS utilizes an extensive GPU network, including NVIDIA H100s, to offer affordable accelerated computing for AI model training and inference specifically within the cryptocurrency sector. Through this approach, exemplified by Nebula Block's 71% cost reduction and 30% performance increase, the project aims to enhance the efficiency and reduce the financial barriers for incorporating AI into blockchain applications.

3.17.6 GPU.net

Website: https://gpu.net/

One liner: Decentralized GPU infrastructure revolutionizing Generative AI, 3D, and Web3 by offering computing resources for AI model training, 3D rendering, VFX, and more.

Description: GPU.net is a decentralized network providing essential GPU infrastructure to address shortages and high costs in the AI sector, particularly for tasks like generative AI training, and 3D rendering, while also supporting cryptocurrency mining. It aims to democratize access to high-performance computing through services such as 3D conversion tools and an open-source LLM library, enhancing data privacy and efficiency in the AI x Crypto space through decentralized operations and token-based resource prioritization.

3.17.7 GPU.net

Website: https://gpu.net/

One liner: Decentralized GPU infrastructure revolutionizing Generative AI, 3D, and Web3 by offering computing resources for AI model training, 3D rendering, VFX, and more.

Description: GPU.net is a decentralized network providing essential GPU infrastructure to address shortages and high costs in the AI sector, particularly for tasks like generative AI training, and 3D rendering, while also supporting cryptocurrency mining. It aims to democratize access to high-performance computing through services such as 3D conversion tools and an open-source LLM library, enhancing data privacy and efficiency in the AI x Crypto space through decentralized operations and token-based resource prioritization.

3.17.8 GPUtopia

Website: https://gputopia.ai/beta

One liner: GPUtopia is revolutionizing the GPU market by creating a decentralized marketplace for buying and selling GPU compute power, leveraging Bitcoin transactions.

Description: GPUtopia is a decentralized platform that leverages the intersection of AI and cryptocurrency by enabling users to rent out or purchase GPU computing power with Bitcoin. This initiative aims to alleviate the current GPU shortage and distribute computing resources more equitably, supporting a variety of tasks from Bitcoin mining to AI processing on a global mesh network.

3.17.9 Ingonyama

Website: https://www.ingonyama.com/

One liner: Ingonyama pioneers in Zero Knowledge Hardware Acceleration, redefining cryptographic security from software to silicon.

Description: Ingonyama is developing the world's first Zero Knowledge Processing Unit (ZPU) to accelerate ZK proofs, tackling computational challenges by blending various disciplines for improved data security, scalability, and sovereignty in the IT ecosystem. This effort, supported by leading industry builders, aims to make ZK technology more accessible and efficient.

3.17.10 IORNA

Website: https://investors.ironmountain.com/news-and-events/press-releases/press-release-details/2020/Keevo-Announces-Strategic-Partnership-with-Iron-Mountain-to-Protect-Crypto-Currency-Users-and-Their-Heirs/default.aspx

One liner: Combining cutting-edge security and privacy for crypto asset management and inheritance through innovative hardware and strategic partnerships.

Description: IORNA, in partnership with Keevo and Iron Mountain, introduces a solution for secure cryptocurrency asset management, utilizing AI-enhanced hardware wallets with multi-factor authentication and a novel beneficiary service. This system, featuring the Keevo Model 1 wallet and Carbon Key™ technology, emphasizes secure, inheritance-friendly crypto asset storage, leveraging Iron Mountain's air-gapped facilities for added security, thereby merging AI with crypto to enhance asset security and ease of access for owners and their heirs.

3.17.11 IOTeX

Website: https://iotex.io/

One liner: IoTeX - Connecting Real World to Web3 with a focus on empowering IoT and smart devices through blockchain technology.

Description: IoTeX is developing a blockchain platform for the Internet of Things (IoT), aiming to merge physical and digital worlds by enabling secure and decentralized IoT applications and smart devices. Their work includes a collaboration with Solana for real-time device analytics and the creation of a confidential computing hardware stack, contributing to the establishment of decentralized private IoT networks essential for advancing AI within the cryptocurrency sector. This infrastructure supports decentralized identities, secure hardware, and real-world data oracles like Pebble, enhancing AI's capability to securely interact with real-world data for applications including NFTs, weather derivatives, and machine learning-powered mining.

3.17.12 Nosana

Website: https://nosana.io/

One liner: Revolutionizing GPU cloud computing with accessible, affordable compute power for AI on Solana.

Description: Nosana is developing a decentralized GPU cloud computing platform on the Solana blockchain, utilizing the $NOS token to allow users to monetize idle hardware and provide affordable AI computation power. This approach facilitates broader access to computational resources, addressing GPU shortages and integrating AI with cryptocurrency to enhance AI process execution.

3.17.13 OpenAgents

Website: https://bitcoinnews.com/adoption/openagents-earn-bitcoin-leasing-gpu-for-ai-tasks/

One liner: Revolutionizing the AI x Crypto space by monetizing GPU power for Bitcoin rewards.

Description: OpenAgents, formerly GPUtopia, is a decentralized marketplace that enables users to earn Bitcoin by contributing their idle GPU power for AI computational tasks. It leverages the Bitcoin and Lightning Network to facilitate a cost-effective platform for AI services, aiming to build a community-driven ecosystem for the development, use, and promotion of AI agents, rewarding participants in Bitcoin for their contributions across the AI agent lifecycle.

3.17.14 Render Network

Website: https://rendernetwork.com/

One liner: Render Network is a decentralized GPU rendering network, revolutionizing 3D content creation and AI with blockchain technology.

Description: Render Network offers a blockchain-based platform that connects 3D artists and studios with GPU owners to provide scalable and efficient rendering and AI computational resources. Operating on the Ethereum blockchain, the project focuses on enhancing AI-driven content creation and rendering for various applications including media, gaming, and medical, by utilizing decentralized GPU computing power.

3.17.15 Shoggoth

Website: https://shoggoth.systems/

One liner: Shoggoth Network facilitates decentralized AI by enabling the sharing of AI resources over a P2P network.

Description: The Shoggoth Network is designed as a decentralized platform utilizing blockchain to facilitate the dissemination and collaborative development of AI resources such as papers, datasets, code, and models through a peer-to-peer network. It employs Git for resource management, promoting global access and innovation in AI by streamlining the sharing process among developers, researchers, and enthusiasts.

3.17.16 Stelia Connectivity

Website: https://stelia.io/

One liner: Stelia Connectivity: Empowering the AI x Crypto space with ultra-low latency, high-capacity networking solutions.

Description: Stelia Connectivity is developing the world's first Tier 1 AI network backbone tailored for hyperscale AI and Crypto demands, offering an elastic, high-speed network that enables efficient AI model training, inference, and blockchain operations. Their network, enhanced through strategic partnerships, provides agile and scalable connections essential for deploying large language models and supporting the growth of AI and blockchain technologies.

3.18 Inference

3.18.1 Aizel Network

Website: https://aizelnetwork.com/

One liner: Aizel Network is revolutionizing blockchain with trustless, on-chain AI, ensuring Web2 speed & costs.

Description: Aizel Network combines AI and blockchain technology, offering a platform where machine learning models can execute trustless, verifiable inferences on-chain using Multi-Party Computation (MPC) and Trusted Execution Environments (TEEs) for security. It promises to equip any smart contract across blockchain networks with scalable, privacy-preserving AI capabilities, facilitated by a team blending expertise in data science, AI, and blockchain.

3.18.2 Aleo

Website: https://aleo.org/

One liner: Aleo leverages zero-knowledge proofs to enable fully private applications on a scalable, privacy-first blockchain.

Description: Aleo leverages zero-knowledge proofs (ZKPs) in its layer-1 blockchain platform to enable the creation of decentralized applications that emphasize user privacy and data security, without compromising scalability or security. Through its native programming language, Leo, and infrastructure like snarkOS and snarkVM, Aleo facilitates the development and deployment of smart contracts with privacy features, positioning itself at the nexus of AI, crypto, and privacy-enhancing technologies.

3.18.3 Aztec

Website: https://aztec.network/

One liner: Aztec is a privacy-first Layer 2 scaling solution on Ethereum enabling private transactions and smart contracts.

Description: Aztec Network enhances blockchain privacy on Ethereum via a Layer 2 solution that supports public and private smart contract executions, leveraging tools like the Aztec Sandbox and Noir programming language for developing privacy-centric applications. This initiative, which attracts significant crypto investment and emphasizes open-source development, is pivotal for integrating advanced privacy and security within AI-driven decentralized applications.

3.18.4 ChainML

Website: https://www.chainml.net/

One liner: ChainML: Pioneering the integration of AI and blockchain for decentralized, secure, and accessible AI agents and machine learning models.

Description: ChainML merges artificial intelligence and blockchain technology to develop and deploy AI agents and machine learning models within a decentralized protocol, enhancing DeFi, gaming, and social dApps. It focuses on providing a generative AI platform for conversational analytics and an open-source framework for AI development, aiming to democratize advanced AI access and promote innovation through community-contributed, secure, and scalable computation in the Web3 ecosystem.

3.18.5 Ezkl

Website: https://github.com/zkonduit/ezkl

One liner: Ezkl: A pioneering Zero-Knowledge Inference tool for private and secure AI computations on the blockchain.

Description: Ezkl is a tool and library for integrating deep learning models with Zero-Knowledge Proofs (ZKP), specifically zk-SNARK, allowing users to verify computations, such as neural network operations, on both private and public data without revealing the data itself. This integration supports the secure and private execution of AI tasks within blockchain environments, notably Ethereum, and facilitates confidential verification for decentralized finance (DeFi) and machine learning applications, with prospects for trustless machine learning inferences and on-chain rewards.

3.18.6 Giza

Website: https://www.gizatech.xyz/

One liner: Giza is revolutionizing the AI x Crypto space by enabling seamless integration of AI with blockchain technologies.

Description: Giza combines AI and blockchain technology to enable the development and management of secure, scalable machine learning models on decentralized applications, enhancing Web3 integrations with a focus on privacy through zero-knowledge proofs and optimized datasets for ML, thus advancing AI-driven solutions in the crypto space.

3.18.7 Hypercycle

Website: https://coinmarketcap.com/currencies/hypercycle/

One liner: HyperCycle is revolutionizing the blockchain and AI space with its layer 0++ architecture designed for high-speed, cost-efficient AI microservices execution.

Description: HyperCycle introduces a Layer 0++ blockchain architecture optimized for efficient and cost-effective on-chain AI microservices, leveraging TODA/IP's ledgerless technology and SingularityNET's Proof of Reputation system for enhanced security and scalability. This platform aims to support decentralized AI applications, including swarm AI and decentralized finance, by improving interoperability and processing capabilities across different blockchain networks.

3.18.8 Lilypad

Website: https://github.com/bacalhau-project/lilypad

One liner: Lilypad establishes a decentralized GPU network for running AI workloads effortlessly, integrated with blockchain for transaction management.

Description: Lilypad leverages blockchain technology to allow users to conduct AI tasks on a decentralized network of GPU resources, offering CLI and web interfaces for engaging with AI models and executing jobs securely through smart contracts. This platform integrates AI and crypto by enabling decentralized computing, supporting

advanced AI applications, and facilitating contributions to the web3 ecosystem with potential in AI inference, machine learning, and scientific research.

3.18.9 Modulus

Website: https://www.modulus.xyz/

One liner: Modulus Labs leverages ZK-proofs to secure and verify AI on the blockchain, fostering trust in decentralized applications.

Description: Modulus Labs, emerging from Stanford University, secured $6.3 million in seed funding to enhance crypto security through the integration of zero-knowledge proofs (ZK-proofs) and zkML, aiming to validate AI outputs on the blockchain without compromising underlying data. This approach seeks to improve the integrity of AI-generated content, facilitate the integration of AI in web3 applications, and reduce reliance on human governance in the blockchain ecosystem, thereby increasing transparency and trust.

3.18.10 Octra

Website: https://twitter.com/octrafoundation?lang=en

One liner: Octra - Empowering the future economy with a decentralized, secure, and private network utilizing fully homomorphic encryption and machine learning.

Description: Octra combines cryptocurrency with artificial intelligence to improve digital transaction security and data management using fully homomorphic encryption (FHE) and machine learning (ML), focusing on decentralized cloud storage that ensures user data privacy and security through the integration of blockchain, cryptography, and AI.

3.18.11 Ora

Website: https://www.ora.io/

One liner: Ora redefines smart contracts with zkOracles, integrating AI for next-gen DeFi and on-chain machine learning.

Description: Ora integrates AI with cryptocurrency by enhancing smart contracts with Zero-Knowledge Oracles (zkOracles), facilitating the development of advanced decentralized finance (DeFi) applications and on-chain machine learning. This approach, using opML for large-scale AI models on the blockchain and collaborating with Polygon CDK for deploying the zkOracle protocol, ensures decentralized, secure, and trustworthy AI-powered applications with a focus on verifiable transparency and on-chain execution.

3.18.12 Phala Network

Website: https://phala.network/

One liner: Phala Network is a privacy-preserving, off-chain compute platform for blockchains, leveraging AI and Secure Enclaves (TEEs) for secure, decentralized applications.

Description: Phala Network is a decentralized platform that integrates AI with blockchain through trusted execution environments like Intel SGX, enhancing the privacy and security of Web3 applications. It introduces 'Phat Contracts' for secure off-chain computing and aims to reduce trust dependencies in blockchain with features such as a trustless MEV ecosystem and AI-enhanced functionalities, focusing on scalability, interoperability, and privacy.

3.18.13 Risc0

Website: https://www.risczero.com/

One liner: Democratizing Zero-Knowledge technology for enhanced security and efficient computation across blockchains.

Description: Risc0 introduces a unique blend of AI and cryptocurrency by leveraging Zero-Knowledge (ZK) proofs through its Bonsai technology, a decentralized proving engine designed to enhance security and enable complex computations on blockchain technology. This project offers a scalable, privacy-focused framework for DeFi and blockchain applications, allowing for the execution of arbitrary Rust code in a secure, verifiable manner without compromising private data.

3.18.14 Ritual

Website: https://ritual.net/blog/introducing-ritual

One liner: Ritual, merging cryptography and AI to revolutionize access and innovation in AI models on the blockchain.

Description: Ritual aims to decentralize AI by creating a distributed network where AI model creators can host their models for unrestricted access, backed by cryptographic integrity and privacy. It bridges AI with blockchain through its Infernet, allowing smart contracts to utilize AI for various applications, targeting to integrate AI functionalities as a core component of the web3 ecosystem.

3.18.15 SingularityNET

Website: https://singularitynet.io/

One liner: SingularityNET is a decentralized marketplace for AI services, facilitating the creation, sharing, and monetization of AI technologies.

Description: SingularityNET merges artificial intelligence and blockchain to establish a decentralized marketplace for AI services and algorithms, facilitated by the AGIX token used for transactions and governance. It aims to become the Internet's Knowledge Layer by integrating Knowledge Graphs with generative AI, contributing to the development of advanced AI systems, with notable contributions from Dr. Ben Goertzel and connections to the AI robot, Sophia.

3.18.16 Spectral

Website: https://www.spectral.finance/

One liner: Spectral: Powering the Inference Economy with Decentralized Machine Intelligence for Web3

Description: Spectral integrates AI and cryptocurrency by deploying a Machine Intelligence Network that uses zero-knowledge proofs to verify machine learning inferences for direct consumption through smart contracts, aiming to enhance the DeFi and blockchain spaces. It fosters a decentralized community of data scientists to solve complex Web3 challenges, securing models through validators and promoting a self-sustaining ecosystem with a revenue-sharing model that incentivizes high-quality, unbiased model development for smart contract applications.

3.18.17 TACEO

Website: https://www.theblock.co/post/251178/nil-foundation-taceo-ethereum

One liner: TACEO: Empowering verifiable machine learning models on Ethereum with privacy-preserving technologies.

Description: TACEO integrates AI and crypto by merging machine learning with Ethereum blockchain using zero-knowledge proofs, ensuring model validation without external trust. This initiative, a collaboration between the Nil Foundation and TACEO, aims to authenticate machine learning models directly on Ethereum's blockchain, enhancing data privacy and enabling secure, provable ML applications within smart contracts across various sectors. Additionally, by focusing on decentralized private computation, TACEO addresses data ownership and privacy concerns, marking a significant step towards secure, private data computation in decentralized applications.

3.18.18 Together AI

Website: https://www.together.ai/

One liner: Accelerating the integration of generative AI models into crypto through innovative cloud platforms, research, and optimizations.

Description: Together AI utilizes cutting-edge AI research to enhance the efficiency and speed of generative AI applications within the blockchain and crypto sphere, by

leveraging technologies like Cocktail SGD and FlashAttention for up to 9x faster training of large language models. This improvement not only fosters cost savings but also amplifies performance, particularly for AI x Crypto projects, backed by accessible, optimized APIs for quick deployment and initiatives like the open-source RedPajama project to promote AI technology democratization.

3.19 Infrastructure

3.19.1 enqAI

Website: https://enqai.com/

One liner: enqAI: Unleashing the potential of uncensored and unbiased AI in the crypto sphere.

Description: enqAI, previously noiseGPT, merges artificial intelligence with blockchain technology on the Ethereum platform, emphasizing an uncensored AI approach within decentralized finance (DeFi). It showcases active market engagement with a supply of 970 million tokens, highlighting the initiative to blend AI's generative capabilities with blockchain's security and transparency.

3.19.2 NEAR

Website: https://near.org/

One liner: NEAR Protocol - A Community-Run Cloud Computing Platform That Fosters AI x Crypto Integration.

Description: NEAR Protocol is a layer-one blockchain designed for high transaction speeds, throughput, and interoperability, making it conducive for decentralized applications and AI x Crypto innovations. It distinguishes itself with human-readable account names, the 'Doomslug' consensus mechanism for security and efficiency, and initiatives like FastAuth that foster a developer-focused ecosystem for projects integrating AI and cryptocurrency, such as Flux and Mintbase.

3.19.3 Theta Network

Website: https://www.thetatoken.org/

One liner: Theta Network: The decentralized infrastructure for video, AI, and entertainment.

Description: Theta Network is a decentralized video streaming platform utilizing blockchain to incentivize bandwidth sharing, integrated with AI algorithms for enhanced content delivery and viewer experience. It offers a Web3 infrastructure supporting diverse applications, including NFTs with DRM, and is validated by major enterprises like Google, Samsung, and Sony, showcasing its application of AI in the crypto space for entertainment and more.

3.20 Infrastructure - Cloud

3.20.1 OORT

Website: https://www.oortech.com/

One liner: OORT, the decentralized cloud revolutionizing privacy, cost savings, and AI integration across global business operations.

Description: OORT integrates AI with blockchain to provide a privacy-focused decentralized cloud service, offering substantial cost savings and operational efficiency with a suite of AI solutions, including a generative AI agent builder. Highlighted by partnerships with Lenovo Image and endorsements from Tencent Cloud and Dell Technologies, the platform is supported by a global network of over 30,544 nodes across 109 countries, demonstrating its scalability and reliability in blending AI with crypto to enhance business operations.

3.21 Security

3.21.1 Anchain.AI

Website: https://www.anchain.ai/

One liner: AnChain.AI: Revolutionizing Web3 Security and Compliance with AI-Powered Solutions.

Description: AnChain.AI employs AI to enhance security, compliance, and crime investigation within the Web3 ecosystem, offering an AI-powered platform that automates the analysis of crypto transactions and smart contracts for identifying risks and ensuring AML/CFT compliance. This platform serves over 200 customers, overseeing more than $100 billion in daily trading volume, by utilizing AI for tasks such as blockchain investigation auto-trace, risk assessment, and the detection of suspicious activities, aiming to improve the safety and integrity of the blockchain environment.

3.21.2 CertiK

Website: https://www.certik.com/

One liner: CertiK leverages AI to enhance blockchain security and smart contract auditing.

Description: CertiK specializes in enhancing blockchain security through AI and Formal Verification technology, offering services such as smart contract and blockchain audits, and continuous security monitoring. Founded by Columbia and Yale professors in 2018, CertiK leverages its expertise to identify vulnerabilities and improve the reliability of blockchain systems, making it a key contributor to the AI and crypto intersection.

3.21.3 Sardine

Website: https://www.sardine.ai/

One liner: Sardine is revolutionizing fraud prevention and compliance in the cryptocurrency space with AI-driven solutions.

Description: Sardine employs machine learning and generative AI to enhance fraud prevention and compliance for the crypto and fintech industries, ensuring secure, instant transactions by analyzing device intelligence and behavior biometrics. This AI-driven approach, in collaboration with BitPay, facilitates direct cryptocurrency purchases from bank accounts, aiming to mitigate systemic risks and foster trust in digital currency exchanges by accurately identifying fraud and improving transaction safety.

3.21.4 TestMachine

Website: https://testmachine.ai/

One liner: AI-powered blockchain security for smart contract vulnerability detection and resolution.

Description: TestMachine combines AI and blockchain technology to improve security across various blockchain platforms, including Ethereum and Binance, by offering tools for the early detection and resolution of vulnerabilities in smart contracts. It integrates with developer workflows to provide real-time security analysis, optimization, testing, and detailed reports, leveraging AI for enhanced threat identification and mitigation.

3.22 Training

3.22.1 Bacalhau

Website: https://www.bacalhau.org/

One liner: Bacalhau - Revolutionizing compute over data with security and efficiency.

Description: Bacalhau offers a platform that distributes computing power for data processing at its source, utilizing Docker containers and WebAssembly for flexible, secure compute jobs. It supports a broad range of architectures and taps into underutilized resources, making it suitable for AI and crypto projects needing efficient, secure data handling near its origin, with a focus on minimizing data movement and adhering to privacy standards.

3.22.2 FATE federated learning

Website: https://fate.fedai.org/

One liner: Empowering secure and efficient federated learning with AI and crypto technologies.

Description: FATE is an open-source framework that enhances federated learning through the use of advanced cryptographic techniques, including partial homomorphic encryption, to secure data while allowing for computational operations on encrypted data—significantly improving efficiency and reducing costs. Its integration with Intel® IPP-Cryptography libraries and focus on data privacy positions it as an effective tool for deploying AI in sectors requiring strict data confidentiality without centralizing sensitive information, demonstrating a practical fusion of AI methodologies with blockchain and cryptography principles.

3.22.3 Flock.io

Website: https://flock.io/

One liner: FLock.io: Revolutionizing privacy-preserving AI on blockchain with federated learning.

Description: FLock.io utilizes blockchain technology to create a decentralized platform for AI training, where users contribute to federated learning without compromising their data privacy. It rewards contributions with cryptocurrency, integrating AI development with the DeFi ecosystem and exploring advanced areas such as BTC GPT, GameFi MBTI, and Zero-Knowledge Machine Learning (ZKML), thereby melding AI's capabilities with blockchain's security and decentralization.

3.22.4 Flower federated learning

Website: https://flower.dev/

One liner: Empowering federated learning in diverse environments through the Flower Framework, compatible with AI and blockchain.

Description: Flower enables federated learning on decentralized devices to enhance data privacy and security for AI, supporting various ML frameworks like TensorFlow and PyTorch for technology-agnostic implementation. Its interoperability with blockchain technologies positions it to advance privacy-preserving AI applications within decentralized finance and security, showcasing the convergence of AI and crypto through secure, decentralized data management.

3.22.5 Gensyn

Website: https://www.gensyn.ai/

One liner: Decentralised machine learning compute protocol to democratize AI development.

Description: Gensyn integrates AI and blockchain technology to create a decentralized global supercluster for machine learning, facilitating peer-to-peer access to computational resources and reducing reliance on major tech providers. This approach democ-

ratizes AI development by enabling global device participation in AI training processes, promoting a more accessible and equitable distribution of computational power.

3.22.6 Golem Network

Website: https://golem.network/

One liner: Decentralized Computing for Everyone

Description: Golem Network offers a peer-to-peer marketplace that utilizes global, unused computing power to support AI projects and complex computations at scale, significantly reducing dependency on centralized cloud services. By integrating its native token, $GLM, for transactions, Golem facilitates a balanced exchange between computing power providers and users, enhancing the AI x Crypto ecosystem.

3.22.7 Lumino AI

Website: https://www.luminolabs.ai/

One liner: Decentralized compute protocol for cost-effective AI model training.

Description: Lumino AI introduces a decentralized network for machine learning model training, directly connecting AI developers with GPU compute providers to cut training costs by 50-70% and leveraging cryptographic proofs for the traceability and accountability of data and models. Its censorship-resistant framework promotes developer freedom, particularly in restrictive regimes, and diminishes central control risks.

3.22.8 Network3

Website: https://network3.io/

One liner: Network3 builds an AI Layer2 helping AI developers worldwide train or validate models in scale quickly, conveniently, and efficiently powered by Sei Network

Description: Network3 creates a specialized AI Layer2 to offer global AI developers services for training and validating their AI models, integrating the functionalities of AI and cryptocurrency technologies.

3.22.9 Nimble Composable AI

Website: https://nimble.technology/

One liner: Nimble Composable AI aims to democratize the exchange of intelligence through a composable, permissionless AI protocol on the blockchain.

Description: Nimble Composable AI introduces a decentralized AI framework focused on composability and interoperability within the blockchain ecosystem, notably in DeFi, DePIN, and GameFi sectors. This framework allows for the combination and reuse of AI models and data, facilitating efficient Ethereum transactions and offering

an open data economy, aiming to democratize AI application development and deployment by reducing dependencies on centralized tech powers.

3.22.10 Privasea

Website: https://www.privasea.ai/

One liner: Privasea: Revolutionizing data privacy and security through Fully Homomorphic Encryption (FHE) on the blockchain, powered by AI and incentivized with $PRVA tokens.

Description: The Privasea AI Network integrates Fully Homomorphic Encryption (FHE) with blockchain technology to enable secure, encrypted machine learning inferences on private datasets, facilitating privacy-compliant data sharing and collaboration within the AI sector. It utilizes a blend of Proof-of-Work and Proof-of-Stake mechanisms, incentivized through $PRVA tokens, to ensure efficient and secure data processing, aiming to bolster AI development without compromising data privacy.

3.22.11 VFchain

Website: https://ieeexplore.ieee.org/document/9321132

One liner: VFChain leverages blockchain to enhance federated learning for secure and efficient AI model training over distributed networks.

Description: VFChain enhances federated learning's security by integrating blockchain for a transparent, verifiable, and efficient AI model training across decentralized networks. It employs a blockchain-recorded committee for model aggregation and authenticated data structures to improve proof search and secure committee changes, demonstrating its effectiveness through experiments with deep learning models on real-world datasets.

3.22.12 Zama

Website: https://www.zama.ai/

One liner: Zama offers Fully Homomorphic Encryption (FHE) solutions for blockchain and AI, prioritizing data privacy and security.

Description: Zama, an open-source cryptography firm, merges blockchain technology with artificial intelligence using Fully Homomorphic Encryption (FHE) to enable the processing of data in its encrypted state, enhancing privacy and security across online sectors such as cloud services and blockchain smart contracts. Their development of advanced tools like TFHE-RS and CONCRETE ML facilitates privacy-preserving machine learning and confidential smart contracts, establishing their role in advancing FHE mainstream adoption.