February 2024

VanEck's Crypto AI Revenue Predictions by 2030

Patrick Bush
Senior Investment Analyst, Digital Assets

Matthew Sigel
Head of Digital Assets Research

We outline AI crypto revenue scenarios by 2030, highlighting a base case of $10.2B, and emphasize the crucial role of public blockchains in driving AI adoption through essential features.

Key takeaways:

- Crypto AI revenues projected to reach $10.2B by 2030 in our base case
- Blockchain technology may become a critical driver for AI adoption and advancements in decentralized AI solution
- Integration with crypto incentives may enhance security and efficiency in AI models
- Blockchain may emerge as a solution for AI identity verification and data integrity challenges

There is a significant chance that public blockchains are the key to unlocking the widespread adoption of artificial intelligence (AI) and that AI applications will be crypto's *raison d'être.* This is because crypto provides important foundational elements AI needs, such as transparency, immutability, clearly defined ownership properties, and an adversarial testing environment. We believe these properties will prove instrumental in allowing AI to reach its full potential. Based upon estimates of AI growth, we assert the base case for AI-focused crypto projects to collect ***$10.2B*** in annual revenues by 2030. In this piece, we speculate on crypto's roles in facilitating the adoption of AI and the value that crypto will derive from AI businesses:

We find that the best applications of crypto to AI to be:

- Provision of decentralized compute resources
- Model testing, fine-tuning, and verification
- Copyright protection and data integrity
- AI safety
- Identity

Crypto will be extraordinarily useful to AI because it is already solving many current and future challenges confronting AI. At its core, crypto solves coordination problems. Crypto bonds people, compute, and monetary resources together to run open-source software. It accomplishes this by offering rewards, in the form of tokens tied to each network's value, to those who create, support, and use each blockchain's network. This reward system can be applied to bootstrap different components of the AI value stack. One significant implication of integrating crypto with AI lies in leveraging cryptocurrency incentives to develop essential physical infrastructure, such as GPU clusters, dedicated to training, fine-tuning, and enabling the usage of generative models. Since crypto is an adversarial environment that employs cryptocurrencies to reward desired user behavior, it is the optimal ground to test and fine-tune AI models to optimize output that conforms to some quality standard.

Blockchains also bring transparency to digital ownership rights, which could help solve some of the open-source software problems that AI will face in the courts, already notable in the NYTimes vs. OpenAI and Microsoft **lawsuit**. Namely, crypto can transparently prove ownership and copyright protection for data owners, model builders, and model users. This transparency will also extend to having mathematical proof of model effectiveness posted to public blockchains. Finally, due to unforgeable digital signatures and data integrity, we believe that public blockchains will help allay identification and safety issues that would otherwise blunt AI effectiveness.

Defining Cryptocurrency's Role in AI Enterprises

2030 Projected Crypto AI Revenues: Bear, Base, Bull Case Scenarios

TAM Market Size ($B)	Bear	Base	Bull
Gen AI Productivity Gains 2030 (TAM)	$2,930	$5,851	$8,494
AI Business Adoption 2030	10.00%	33.00%	50.00%
Business AI TAM	$293	$1,931	$4,247
AI Stack Value Capture	6.00%	13.00%	20.00%
Total Business AI Spend	**$17.58**	**$251.00**	**$849.44**
AI Stack Share ($B)	**Bear**	**Base**	**Bull**
Software	50.00%	50.00%	50.00%
AI Infrastructure as a Service	18.90%	18.90%	18.90%
Identity	3.50%	3.50%	3.50%
Safety	8.90%	8.90%	8.90%
AI Stack Annual Revenues ($B)	**Bear**	**Base**	**Bull**
Software	$8.79	$125.50	$424.72
AI Infrastructure as a Service	$3.32	$47.44	$160.54
Identity	$0.62	$8.78	$29.73
Safety	$1.56	$22.34	$75.60
Crypto Market Share	**Bear**	**Base**	**Bull**
Software	2.50%	5.00%	7.50%
IaaS	2.00%	4.00%	6.00%
Identity	5.00%	10.00%	15.00%
Safety	2.50%	5.00%	7.50%
Crypto AI Revenues ($B)	**Bear**	**Base**	**Bull**
Software	$0.22	$6.27	$31.85
IaaS	$0.07	$1.90	$9.63
Identity	$0.03	$0.88	$4.46
Safety	$0.04	$1.12	$5.67
Total AI Crypto Revenues 2030	**$0.36**	**$10.17**	**$51.62**

Sources: Morgan Stanley, Bloomberg Intelligence, VanEck Research as of 1/29/2024. ***Past performance is no guarantee of future results. The information, valuation scenarios and price targets presented in this blog are not intended as financial advice or any call to action, a recommendation to buy or sell, or as a projection of how AI businesses will perform in the future. Actual future performance of is unknown and may differ significantly from the hypothetical results depicted here. There may be risks or other factors not accounted for in the scenarios presented that may impede the performance. These are solely the results of a simulation based on our research and are for illustrative purposes only. Please conduct your own research and draw your own conclusions.***

To forecast the market for crypto AI, we first estimate the total addressable market (TAM) of the commercial productivity gains enabled by AI and our baseline for this figure is derived from **McKinsey** assumptions for 2022. We then apply economic and productivity growth assumptions to McKinsey's figures to find a base case, year 2030 TAM of $5.85T. In this base case we assume that AI productivity gains are 50% higher than GDP growth and GDP grows at 3%. We then project market penetration for AI into global businesses, 33% in the base case, and apply it to our initial TAM to reach an estimated $1.93T in productivity gains for businesses from AI. To arrive at the revenues of all AI businesses, we assume that 13% of these productivity gains are captured by AI businesses (or spent by corporate consumers) as revenues. We ballpark the AI revenue take rate by applying the average revenue share of the labor costs of S&P 500 enterprises and assume AI expenditures should be similar. The next part of our analysis applies forecasts of the AI value stack distribution from Bloomberg Intelligence to arrive at estimates of each AI business cohort's annual revenues. Finally, we apply specific estimates of crypto market share of each of those AI business to arrive at each case and each market's final figure.

We envision a future where decentralized AI models built from open-sourced public repos are harnessed to every use case imaginable. And in many cases, these open-source models outcompete centralized AI creations. The base of this supposition stems from the assumption that open-source communities bring together hobbyists and enthusiasts who are uniquely motivated to improve things. We have already seen open source internet projects shatter traditional businesses. The best examples of this phenomenon are Wikipedia effectively ending the commercial encyclopedia business and Twitter disrupting news outlets. These open-source communities succeed where traditional businesses fail because open-source groups coordinate and motivate people to provide value through a combination of social clout, ideology and group unity. In short, these open-source communities succeed because their members *care*.

Integrating open-source AI models with crypto incentives expands these emerging communities' influence, empowering them with financial capabilities to create the necessary infrastructure for attracting new participants. Applying this premise to AI will be a fascinating combination of passion and monetary resources. AI models will undergo testing in crypto-incentivized competitions, establishing a landscape where the evaluation of models becomes benchmarked. Within this environment, the most effective models and evaluation criteria emerge victorious, as the value of each model is explicitly quantified. Consequently, in our base case we anticipate that blockchain-produced AI models will represent 5% of all AI software revenues. This estimate includes hardware, software, services, ads, gaming and more, reflecting a transformation of how many businesses operate. Of the total revenues of AI software, we expect this to be about half of all AI revenues, or around $125.50B. The 5% market share we expect for open-source models thus corresponds to **$6.27B** in revenues going to crypto token-backed AI models.

We project that the TAM for the provision of compute—or AI infrastructure as a service—for fine-tuning, training, and inference may reach $47.44B by 2030. As widespread adoption leads AI to become integral to many functions of the world's economy, the provisioning of compute and storage may be envisaged as a public utility akin to power generation and distribution. In this dynamic, the vast majority of the "base load" will come from GPU cloud Hyperscalers like Amazon and Google, and their market share will approximate a **Pareto Distribution** of 80%. We see blockchain-apportioned back-end server infrastructure catering to specialized needs and acting as a "peaking" provider during high network demand. For producers of bespoke AI models, providers of crypto storage and computing offer benefits such as on-demand service delivery, shorter SLA lock-in periods, more customized compute environments, and heightened latency sensitivity. Furthermore, decentralized GPUs can seamlessly integrate with decentralized AI models within smart contracts, enabling permissionless use cases where AI agents scale their own compute needs. Conceiving blockchain-provisioned GPUs as the Uber/Lyft equivalent of AI compute infrastructure, we posit that blockchain-provided compute and storage will capture 20% of the non-Hyperscaler market for AI infrastructure, potentially yielding revenues of **$1.90B** by 2030.

Defining "identity" in the context of AI agents and models through provable on-chain humanness can be viewed as a Sybil defense mechanism for the world's computer networks. We can estimate the cost of this service by examining the expenses associated with defending different blockchain networks. In 2023, these costs for Bitcoin, Ethereum, and Solana were around **1.71%**, **4.3%**, and 5.57%, respectively, of each network's value in inflationary issuance. Conservatively, we can infer that Identity should comprise around 3.5% of the AI market. Given a TAM of $125.5B for AI software, this corresponds to $8.78B of annual topline. Since we believe that crypto provides an optimal solution to identity issues, we believe it will hold a 10% market share of this end market, which ballparks its yearly revenue figure of around **$878M**.

AI safety is poised to emerge as another important component of AI devices, with a fundamental requirement to verify running the correct model using uncorrupted, relevant, up-to-date data. As AI expands into applications where human lives are at risk, such as self-driving cars, factory robotics, and healthcare systems, the tolerance for failure becomes minimal. The need for accountability in the event of accidents will drive an insurance market that demands concrete proof of safety. Public blockchains are ideal for fulfilling this function because they can post "proofs of safety" on unalterable ledgers that anyone can see. This business can be thought of as akin to compliance for financial institutions. Considering that U.S. **commercial** and **investment** banks generate $660B in revenue while spending $58.75B in compliance costs (8.9% of revenues), we project that AI safety should account for around $22.34B of the $251B AI TAM. Despite the potential for crypto to enhance AI safety, given the U.S. government's focus on AI, we believe the majority of compliance for AI will be centralized. As such, we estimate that crypto will make up around 5% of that market or around **$1.12B**.

Organizing Decentralized Compute Resources

Crypto can apply its considerable social and financial coordination benefits to the democratization of access to compute, thereby addressing current pain points afflicting AI developers. In addition to the high costs and limited access to quality GPUs, AI model builders currently confront other nagging issues. These include vendor lock-in, lack of security, limited compute availability, poor latency, and geo-fencing due to national laws.

Crypto's ability to satiate AI's hunger for GPUs stems from crypto's ability to bring together resources through token incentives. The Bitcoin network's $850B in token value and $20B in equity value demonstrate this ability. As such, there is potential for both current Bitcoin miners and promising decentralized GPU marketplaces to add substantial value to AI by provisioning decentralized compute.

A helpful analogy for understanding the provision of GPUs through blockchain is the power generation business. To simplify, there are entities operating large, expensive plants that steadily generate power to meet most electricity grid demands. These "base load" plants experience consistent demand but require substantial capital investments for construction, resulting in relatively low but guaranteed returns on capital. Complimenting the base load is another category known as "peaking power" generators. These businesses provide power when electricity demand exceeds the capabilities of baseload generation. This involves high-cost, small-scale energy production that is strategically positioned close to the demand for that energy. We anticipate a similar dynamic unfolding in the realm of "compute-on-demand."

Diversification of Bitcoin Miners into AI

Bitcoin and other proof-of-work cryptocurrencies share with AI a high appetite for energy. This energy must be created, sourced, transported, and broken down into useable electricity to power mining rigs and compute clusters. This supply chain necessitates substantial investments from miners into power plants, electricity purchase agreements, grid infrastructure, and data center provisions. The monetary incentives stemming from mining PoW cryptocurrencies have led to the emergence of many globally dispersed bitcoin miners that own energy and electricity rights and integrated grid architecture. Much of this energy is derived from lower-cost, socially shunned carbon-intensive sources. As such, the most compelling value proposition that bitcoin miners can offer is lower-cost energy infrastructure to power AI backend infrastructure.

Hyperscale compute providers like AWS and Microsoft have pursued a strategy of investing in vertically integrated operations and establishing their own energy ecosystems. Big Tech has moved upstream, designing their own silicon and sourcing their own energy, largely the renewable variety. Data centers now consume two-thirds of the renewable power available to corporations in the U.S. Microsoft and Amazon have both committed to 100% renewable energy supply by 2025. However, if the anticipated compute needs exceed expectations, as some suggest, the electricity needs of AI-focused data centers may double by 2027 with CAPEX potentially tripling current estimates. Already, Big Tech pays $0.06-0.10/kWh for power, much more expensive than what competitive Bitcoin miners generally pay ($0.03-0.05kWh). If demand for energy from AI exceeds current infrastructure plans by Big Tech, the bitcoin miner's electricity cost advantage over the Hyperscalers could significantly multiply. Miners are increasingly drawn to higher-margin AI businesses associated with GPU provision. Notably, Hive **reported** in October that its HPC and AI operations are generating 15x more revenue than bitcoin mining on a per-megawatt basis. Other bitcoin miners seizing the AI opportunity are Hut 8 and Applied Digital.

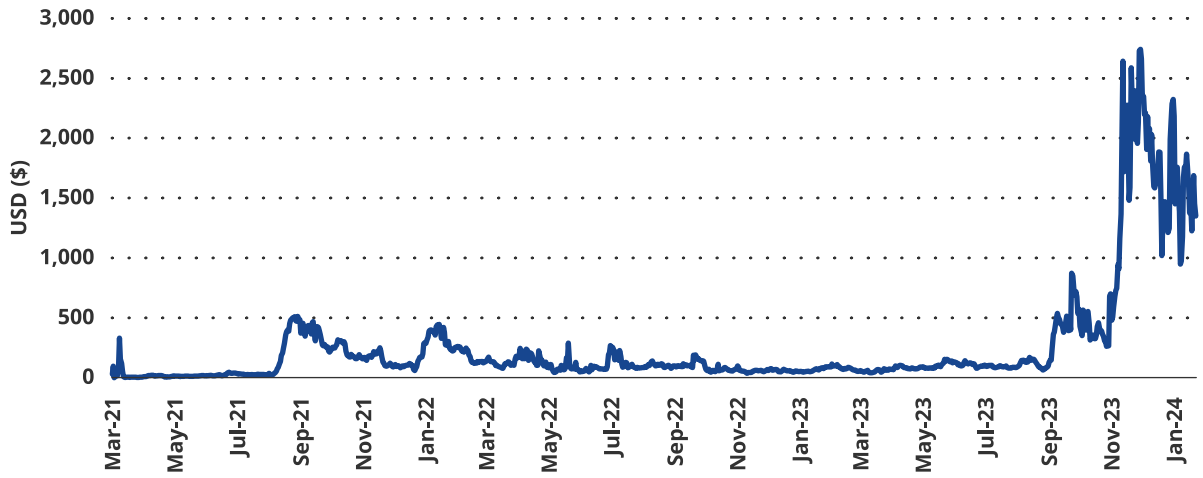
Bitcoin miners have experienced growth in this new market, which has helped diversify revenue and strengthened earnings reports. In Hut 8's Q3 2023 analyst call, CEO Jaime Leverton stated, "In our HPC business, we created some momentum in Q3 with new customer additions and growth among existing customers. Last week, we launched our on-demand cloud service for customers seeking HPC services from our GPUs with Kubernetes-based applications that can support artificial intelligence, machine learning, visual effects, and rendering workloads. This service puts the control in our clients' hands while reducing provisioning time from days to minutes, which is particularly compelling for those seeking shorter-term HPC projects." Hut 8 **realized** $4.5M in revenue from its Q3 2023 HPC operations, constituting over 25% of the firm's revenue for the period. The rising demand for HPC services and new offerings should contribute to the future growth of this business line, and with Bitcoin halving around the corner, HPC revenue could soon exceed that of mining, depending on market conditions.

Though their operations sound promising, bitcoin miners pivoting to AI could stumble due to a lack of skills in data center construction or an inability to scale power supply. These miners may also find challenges related to operating overhead due to the costs of hiring a new data center-focused salesforce. Additionally, current mining operations do not have ample network latency or bandwidth because their optimization for cheap energy causes them to be located in remote places, often lacking high-speed fiber connections.

Implementing a Decentralized Cloud for AI

We also see a long tail of compute-focused crypto projects that will capture a small but substantial portion of the AI server resource market. These entities will coordinate clusters of compute outside of Hyperscalers to deliver a value proposition tailored to the needs of upstart AI builders. The benefits of decentralized compute include customizability, open access, and better contract terms. These blockchain-based compute firms allow small AI players to avoid the substantial expense and general unavailability of high-end GPUs like H100s and A100s. Crypto AI ventures will satiate demand by creating a network of physical infrastructure built around crypto-token incentives while offering proprietary IP that creates the software infrastructure to optimize the compute's usage for AI applications. Blockchain compute projects will use a marketplace approach alongside crypto rewards to uncover cheaper compute from independent data centers, entities with excess compute, and former PoW miners. A few projects providing decentralized compute for AI models include **Akash**, **Render** and **io.net**.

Daily Revenue Akash



Akash Daily Revenue. Source: Cloudmos as of 1/30/2024. ***Past performance is no guarantee of future results.***

Akash, a Cosmos-based project, can be considered a general-purpose decentralized "**Supercloud**" that offers CPU, GPU, Memory, and Storage. Effectively, it is a two-sided marketplace connecting users of cloud services with those supplying them. Akash's software is designed to coordinate compute supply with demand while creating the tooling to facilitate AI model training, fine-tuning, and running. Akash also ensures both marketplace buyers and sellers honestly fulfill their obligations. Akash is coordinated through its $AKT tokens, which can be used to pay for cloud services at a discount. $AKT is also given out as an incentive mechanism for GPU compute providers and other network participants. On the supply side, Akash has made great strides in adding compute suppliers as there are 65 different providers on Akash's marketplace. While the demand for compute had been lackluster until Akash's AI Supercloud debuted on August 31, 2023, $138k has been spent by compute buyers after the launch date.

Render, who recently migrated to Solana, was initially focused on connecting artists to decentralized groups that would provide GPU power to render images and videos. However, Render has begun to focus its decentralized GPU fleet on satisfying machine-learning workloads to support deep-learning models. Through network improvement proposal **RNP-004**, Render now has an API to connect external networks like io.net that will utilize Render's network of GPUs for machine learning. Subsequent proposals by the Render community were passed to allow its GPUs to be accessed through **Beam** and **FEDML** to fulfill machine learning tasks. As such, Render has become a decentralized facilitator of GPU workloads coordinated by $RNDR payments to providers and $RNDR incentives to entities running the network's backend infrastructure.

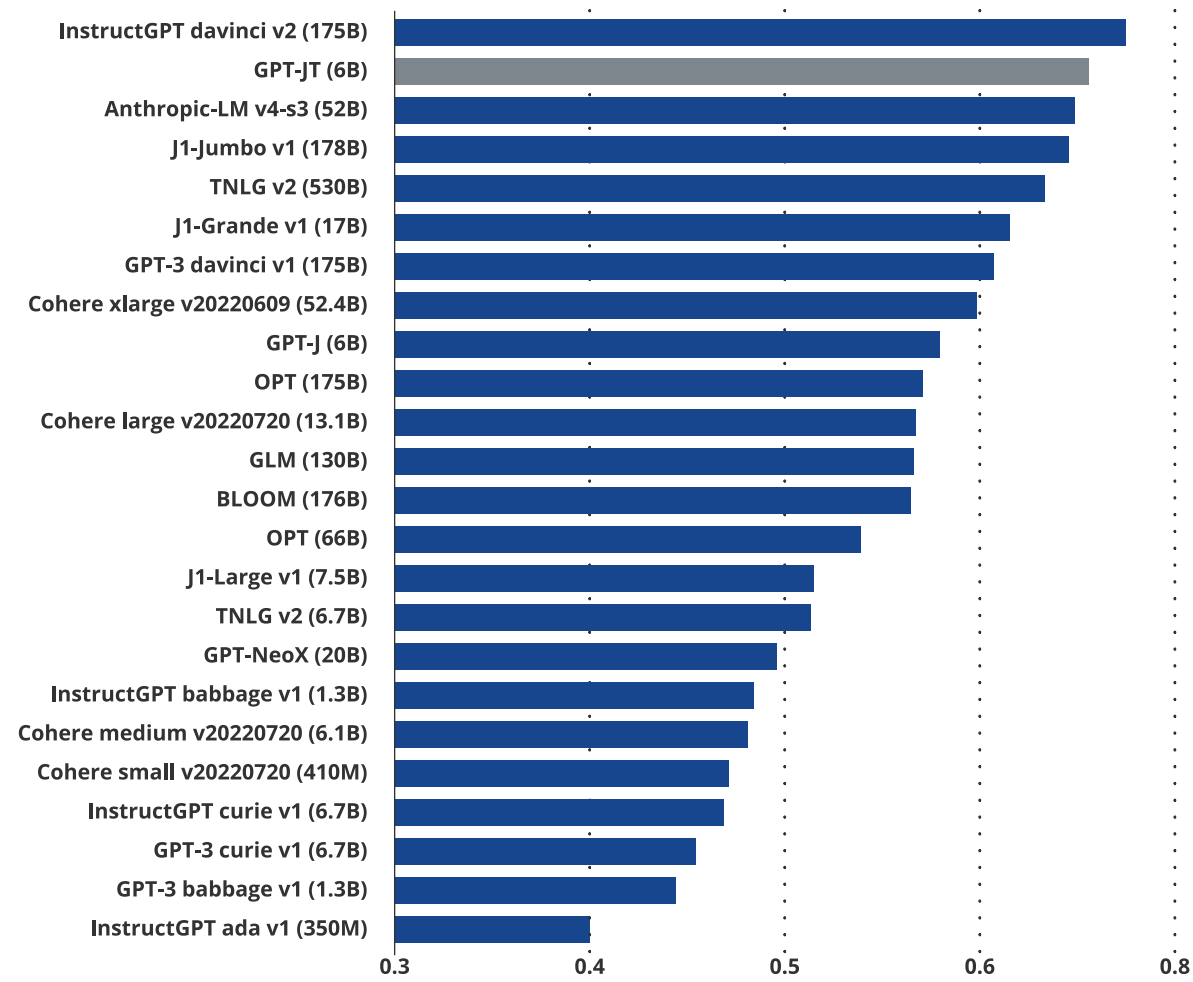
Io.net GPU Price Comparison. Source: io.net as of 1/4/2024.

Another interesting project that sits on Solana, considered a DePIN or Decentralized Physical Infrastructure Network, is io.net. The purpose of io.net is also the provision of GPUs, but its focus is exclusively on applying GPUs to power AI models. Io.net adds more services to its core stack besides simply coordinating compute. Its system claims to handle all components of AI, including creation, usage, and fine-tuning, to facilitate and troubleshoot AI workloads across its network properly. The project also taps into other decentralized GPU networks like Render and Filecoin and its own sourced GPUs. Though io.net currently lacks a token, one is planned for launch in 1Q2024.

Overcoming Bottlenecks of Decentralized Compute

Utilizing this distributed compute, however, remains a challenge due to the networking demands imposed by the typical **633TB+** of data necessary to train deep learning models. Computer systems geo-located across the globe also present novel obstacles for parallelizing model training due to the latencies and differences in computer capabilities. One company attacking the open-source foundation model market with gusto is Together, which is building a decentralized cloud to host open-source artificial intelligence models. Together will enable researchers, developers, and companies to leverage and improve artificial intelligence with an intuitive platform combining data, models, and computation, widening AI accessibility and empowering the next generation of tech companies. In partnership with leading academic research institutions, Together has built the Together Research Computer, allowing labs to pool their computing for AI research. The company has also worked with the Stanford Research Center for Research on Foundational Models (CRFM) to create the Holistic Evaluation of Language Models (HELM). HELM is a "living benchmark" that aims to improve the transparency of AI by providing a standardized framework for evaluating such foundation models.

Since Together's founding, founder Vipul Ved Prakash has spearheaded the launch of several projects including 1) GPT-JT, an open LLM with a 6B parameter model trained over <1Gbps link, 2) OpenChatKit, a powerful, open-source base to create both specialized and general purpose chatbots, and 3) RedPajama, a project to create leading open-source models, with the goal of it being a foundation for research and commercial applications. The Together Platform is a foundation model composed of open models on commodity hardware, decentralized cloud, and a comprehensive developer cloud, pooling together different compute sources including consumer miners, crypto mining farms, T2-T4 cloud providers, and academic compute.



GPT-JT Performance. Source: The Decoder as of 1/4/2024.

We believe decentralized and democratized cloud compute solutions like Together could significantly slash the costs of building new models, potentially disrupting and competing with established giants such as Amazon Web Services, Google Cloud and Azure. For context, Together can provide pricing approximately 4x lower than AWS, comparing AWS Capacity Blocks and AWS p5.48xlarge instances to Together GPU clusters configured with an equal number of H100 SXM5 GPUs.

As open LLMs become increasingly accurate and thus more widely adopted, Together could become the industry standard for open-source models, akin to Red Hat was for Linux. Competitors in this space include Stability A and HuggingFace as model providers and Gensyn and Coreweave as AI cloud providers.

Enhancing AI Models Through Crypto Incentives

Blockchains and crypto incentives prove that network effects and rewards tied to the size of that network effect compel people to perform *useful work*. In the context of bitcoin mining, that task is to secure the Bitcoin network by employing expensive banks of electricity, technical manpower, and ASIC machines. This concert of economic resources provides a Sybil attack defense mechanism preventing economic attacks on Bitcoin. In exchange, the miners orchestrating these resources receive $BTC. However, the green space for useful work for AI is much larger, and some projects are already pushing AI and machine learning models to improve.

The most primitive of these projects is Numerai. Currently, **Numerai** can be considered a decentralized data science tournament to identify the best machine learning models that optimize financial returns by building a portfolio of stocks. In each epoch, anonymous Numerai participants are granted access to masked raw data and asked to harness that data to build the best-performing equities portfolio. To participate, users are not only asked to submit predictions but also compelled to stake NMR tokens behind their models' predictions to attest to the value of those models. Other users are also allowed to stake tokens on the models they believe will perform the best. Each staked, submitted model's outputs are then fed into a machine learning algorithm to create a meta-model that informs the investment decisions of the Numerai One hedge fund. Users who submit the "inferences" with the best information coefficients or effectiveness are rewarded with NMR tokens. At the same time, those who staked the worst models have their tokens slashed (seized and re-purposed for rewards to winners).

00 : Root	11 : Text Training	22 : Unknown
01 : Text Prompting	12 : Unknown	23 : Prime-Net
02 : Machine Translation	13 : Dataverse	24 : Unknown
03 : Data Scraping	14 : LLM Defender	25 : Bitcurrent
04 : Multi Modality	15 : Blockchain Insights	26 : Image Alchemy
05 : Image Generation	16 : Audio	27 : Compute
06 : Unknown	17 : Petal	28 : ZK Tensor
07 : Storage	18 : Cortex.t	29 : 3D Gen
08 : Time Series Prediction	19 : Vision	30 : Lovelace
09 : Pretraining	20 : Unknown	31 : Game of Life
10 : Map Reduce	21 : Filetao	32 : Roleplay

Subnets and Use Case on Bittensor. Source: https://taostats.io/api/ as of 1/2/2024.

A similar project that massively expands Numerai's core concepts is **Bittensor**. Bittensor can be thought of as "Bitcoin for Machine Intelligence" as it is a network that provides economic incentives for AI/ML models. This is accomplished by entities called "miners," who build AI models, and "validators," who assess the quality of outputs from those models. Bittensor's architecture is that of a base network and many smaller subnetworks (Subnets). Each subnetwork focuses on a different domain of machine intelligence. Miners on these subnets are queried by validators with various questions or requests to assess the quality of their AI models.

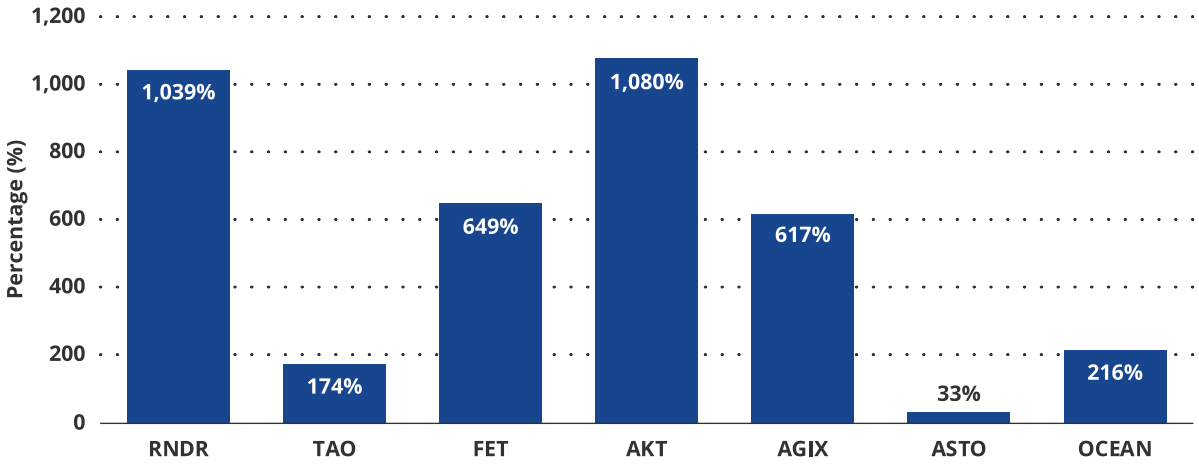
The best-performing models receive the highest rewards of TAO tokens, while the validators are compensated for accurate assessments of the miners. On a higher level, both validators and miners have to stake tokens to participate in each subnet, and each subnet's proportion of total stake determines how many TAO tokens it receives from all Bittensor total inflation. Thus, each miner has an incentive not only to optimize its model to win the most rewards but also to focus their model on the best AI domain subnet. Additionally, since miners and validators must maintain funds to participate, each must exceed cost of capital hurdles, or they drop out of the system.

As of January 2024, there were 32 different subnets, each dedicated to a specific domain of machine learning or AI. For example, Subnet 1 is a text prompting LLM similar to ChatGPT. On this subnet, the miners operate various tweaked versions of LLMs to best respond to prompts by validators who assess the quality of responses. On Subnet 8, called Taoshi, miners submit short-term forecasts for the price of bitcoin and various financial assets. Bittensor also has subnets dedicated to human language translation, storage, audio, web scraping, machine translation, and image generation. Subnet creation is permissionless, and anyone with 200 TAO can create a subnet. Subnet operators are responsible for creating the assessment and reward mechanisms for each subnet's activity. For example, Opentensor, the foundation behind Bittensor, runs Subnet 1 and recently released a model in concert with **Cerebras** to assess the miners' LLM outputs on that subnet.

While these subnets are all initially subsidized in full by the inflationary rewards, each subnet will eventually have to sustain itself economically. As such, subnet operators and validators must coordinate to create tooling to allow outside users to pay to access the services of each subnet. As inflationary TAO rewards diminish, each subnet will become increasingly dependent upon outside revenue to sustain itself. In this competitive environment, there is direct economic pressure to create the best models and incentives for others to create profitable *real-world* applications for those models. Bittensor is unlocking the potential of AI using scrappy, small enterprises to identify and monetize AI models. As well-known Bittensor evangelist MogMachine puts it, this dynamic can be thought of as "Darwinian competition for AI."

Another interesting cohort of projects is one that employs crypto to incentivize the creation of AI agents who are programmed to autonomously accomplish tasks on behalf of humans or other computer programs. These entities are essentially adaptive computer programs aimed at solving specific problems. Agents is an all-encompassing term for chatbots, automated trading strategies, gaming characters, and even metaverse assistants. One notable project in this domain is Altered State Machine, a platform for creating AI agents that are owned, powered, and trained using NFTs. In **Altered State Machine**, users create their "agents" and then "train" them using decentralized GPU clusters. These agents are optimized for specific use cases. Another project, **Fetch.ai**, is a platform for creating agents that are customized for each user's needs. Fetch.ai is also a SaaS business that allows the registration and lease or sale of agents.

Returns of AI Tokens Since 1/1/2023



Source: Artemis XYZ as of 1/10/2024. Past performance is no guarantee of future results.

Verifying Through Zero-Knowledge (zk) Proofs

2023 was a banner year for new AI models with the unveiling of ChatGPT by OpenAI, LLAMA-2 by Meta, and BERT by Google. Due to the promise of deep learning, there are more than **18,563** AI-related start-ups in the United States as of June 2023. These start-ups and others have produced thousands of new base and fine-tuned models. However, the proliferation of many new entities in a space where **1 in every 4 VC dollars** is invested in AI-related companies should raise serious concerns.

- Who actually created and owns each model?
- Was the output actually produced from the specified model?
- Is the model actually as effective as it is advertised?
- What was the data source for each model and who owns that data?
- Did training, fine-tuning and/or inference infringe on any copyrights or data rights?

Both investors and users of these models alike should have 100% certainty they can address these questions. Currently, many benchmarking tests exist for different components of LLM outputs, like **HumanEval** for code generation, **Chatbot Arena** for LLM assistance tasks, and **ARC Benchmark** for LLM reasoning ability. However, despite attempts at model transparency like Hugging Face's **Open LLM Leaderboard**, there is no concrete proof of a model's effectiveness, ultimate provenance, or the source of its training/inference data. Not only can benchmarks be gamed, but there is also no certainty that a specific model was actually run (as opposed to using an API connecting to another model), nor is there any assurance that the leaderboards themselves are honest.

Herein lies the unification of public blockchains, AI, and a bleeding-edge field of mathematics called **zero-knowledge (zk) proofs**. zk proofs are an application of cryptography that allows someone to prove, with mathematical certainty to a desired level, that a statement they make about data is correct without revealing the underlying data to anyone. Statements could include simple declarations like rankings but can be extended to complex mathematical calculations. For example, not only could someone prove he or she knows the relative wealth of a sample of individuals without revealing that wealth to another party, but he or she could also prove the correct calculation of the mean and standard deviation of the group. In essence, you can prove you know data and/or you made truthful assertions using that data without revealing either the specifics of that data or how you made that calculation. Outside of AI, zk proofs are already being applied to scale Ethereum, allowing transactions to occur off-chain on layer-2 blockchains. Recently, zk proofs have been applied to deep learning models to prove:

- Specific data was used to generate a model or provide an inference output (also, what data/sources *were not* used)
- A certain model was used to generate an inference
- An inference output was not doctored

zk proofs can be posted to the public, permanent blockchains, and verified by smart contracts. The result is that blockchains can publicly and irrefutably prove important properties of AI models. Two cutting-edge projects applying ZK to AI referred to as "Zero-Knowledge Machine Learning" (ZKML), are **EZKL** and **Modulus**. EZKL uses the **Halo2** proof system to generate zk-snarks, a type of zero-knowledge proof, which can then be publicly verified on Ethereum's EVM. While the model sizes that EZKL can currently prove are relatively small, around 100M parameters compared to 175B for ChatGPT 4, EZKL's CEO **Jason Morton** believes they are looking at an "engineering problem" rather than an issue of "technological limitations". EZKL believes they can surmount proving issues by **splitting** up proofs to be executed parallel to cut down on memory limitations and compute time. In fact, Jason Morton believes that one day, "verifying a model will be as easy as signing a blockchain transaction."

The application of ZKML proofs to AI can solve important pain points of AI implementation, including copyright issues and AI safety. As the recent lawsuit by the New York Times against Open AI and Microsoft demonstrates, copyright law will be applied to data ownership, and AI projects will be pressed to provide proof of their data sources. ZKML tech can be employed to settle courtroom disputes over models and data ownership quickly. In fact, one of the best applications of ZKML is to allow data/model marketplaces like **Ocean Protocol** and **SingularityNet** to prove the authenticity and efficacy of their listings.

AI models will eventually spread to domains where certainty of accuracy and safety will be of paramount importance. It is estimated that there will be **5.8B** AI edge devices by 2027, and these could potentially include heavy machinery, robotics, and autonomous drones and vehicles. As machine intelligence is applied to things that can hurt and kill people, it will be important to prove that a reputable model was run on that device using high-quality data from a reliable source. Though constructing and posting continuous live proofs from these edge devices to blockchain will likely be economically and technically challenging, certifying the models upon activation or periodic postings to the blockchain may be more feasible. However, Zupass who comes from the **0xPARC** Foundation, has already established primitive proofs derived from "**Proof Carrying Data**" that can inexpensively establish proofs of *facts* that occur on edge devices. Currently, this pertains to event attendance, but one could foresee this shortly migrating to other areas like identity or even **healthcare**.

How good is your robot surgeon's AI model?

Robot Assisted Surgery. Source: MIT Technology Review as of 1/30/2024.

From the standpoint of businesses that may incur liabilities from malfunctioning devices, it would seem ideal to have verifiable evidence that their model was not the source of a costly accident. Likewise, from an insurance standpoint, verifying and proving the usage of reliable models trained on realistic data may become financially necessary. Similarly, in a world of AI deepfakes, utilizing blockchain-verified and attested **cameras**, phones, and computers to perform various actions may become the norm. Of course, proofs of authenticity and accuracy of these devices should be posted to public, open-source ledgers that prevent tampering and fraud.

Despite the immense promise of these proofs, they are currently limited by gas expense and computational overhead. At current ETH prices, submitting a proof on chain costs around 300-500k gas (~$35-$58 at current ETH prices). From a computational standpoint, Sreeram Kennan of Eigenlayer estimates that "proving computation that costs $50 to be run on AWS would cost ~1,000,000x more using current ZK proving technology." The consequence is that zk proofs, which have evolved substantially faster than anyone anticipated a few years ago, have far to go before opening up practical use cases. Suppose one is curious to see ZKML applied. In that case, they can participate in a decentralized **singing** contest judged by an attested on-chain, smart contract model, with their results uploaded to the blockchain forever.

Establishing Humanness with Blockchain-Based Identity

A likely consequence of widespread, advanced machine intelligence is that autonomous agents will become the most prolific users of the internet. There is a significant chance that the unleashing of AI agents will lead to disruptions across the web from purposeful bot-generated spam and even innocuous task-based agents clogging networks ("**Get Rid of Junk Email**"). **Solana** saw 100 gigabits in data traffic per second when bots competed for perhaps $100k worth of arbitrage opportunities. Imagine the torrent of web traffic that will launch when AI agents can ransom millions of corporate websites for billions of dollars. This points to a future internet with limits imposed on non-human traffic. One of the best ways to limit these types of attacks is to impose economic taxes on the overuse of poorly-priced resources. But how do we determine the optimal framework for spam charges, and how do we determine humanness?

Luckily, there is already a built-in defense employed by blockchains to prevent AI-bot-style Sybil attacks. The combination of metering non-human users alongside tolls for non-human usership would be an ideal implementation alongside slightly taxing computation, like **Hashcash**, that would inhibit bots. With respect to proofs of humanity, blockchains have long grappled with **overcoming anonymity** to unlock activities such as undercollateralized lending and other reputation-based activities.

One approach to gaining steam to demonstrate identity is to employ **JSON** web tokens (JWTs). JWTs are "0Auth" credentials, similar to "cookies," generated when you sign on to a website like Google. They allow you to demonstrate your Google identity when visiting various websites on the internet while you are signed into Google. **zkLogin**, created by L1 blockchain **Sui**, allows users to link their wallet private keys and actions to their Google or Facebook accounts that generate JWTs. **zkP2P** extends this concept further to employ JWTs to permissionlessly allow users to exchange fiat for crypto on the Base blockchain. This is accomplished by confirming peer-to-peer cash transfers through the payment app Venmo that, when confirmed through email JWTs, unlock smart contract escrowed USDC tokens. The consequence of both projects is that they establish firm connections to off-chain identities. For example, whereas zkLogin connects wallet addresses to Google identities, zkP2P is only available to KYC users of Venmo. Though both lack the robust guarantees to make them reliable enough for on-chain identity, they create important building blocks that others can employ.

While many projects are trying to confirm the human identity of blockchain users, the most audacious is **WorldCoin**, founded by OpenAI CEO **Sam Altman**. Though highly controversial because users must scan their **irises** using the dystopian "Orb" machine, WorldCoin is moving towards an immutable system of identity that cannot be easily faked or overwhelmed by machine intelligence. This is because WorldCoin creates an encrypted identifier based on each human's unique eye "fingerprint" that can be sampled for uniqueness and authenticity. Once verified, the user receives a digital passport called the World ID on the Optimism blockchain, allowing that user to prove their humanity on the blockchain. Most importantly, a person's unique signature is never revealed, nor can it be tracked because it is encrypted. World ID simply asserts that a blockchain address belongs to a human. Projects like Checkmate are already linking World ID to social media profiles to ensure users are unique and genuine. In a future internet dominated by AI, it may become commonplace to prove humanity definitively in every single online interaction. When the limitations of Captchas have been overwhelmed by AI, blockchain applications can prove identity cheaply, quickly, and concretely.

Contributing to AI via Blockchain Technology

Without question, we are in the early innings of the AI revolution. However, if the growth trajectory of machine intelligence will match the **boldest** of forecasts, AI must be challenged to excel while having its potential for harm tamed. We believe that crypto is the ideal trellis to properly "train" the bountiful fruiting but potentially insidious plant that is AI. Blockchain's solution set for AI can boost the output of creators of machine intelligence by supplying them with more responsive, flexible, and potentially cheaper decentralized compute. It also incentivizes builders who can create better models while supplying economic impetus to others to build useful businesses from those AI models. Just as importantly, model owners can prove their models' effectiveness while demonstrating that protected data sources were not used. For AI users, crypto applications can confirm the models they run to comply with safety standards and are probably useful. For everyone else, blockchain and crypto may be the tangle of penalties and rewards that bind down the Gulliver that artificial intelligence is certain to become.

Provider	Description	Service Category	Fully Diluted Market Cap (M)	Usage
io.net	Decentralized computing network for accessing distributed cloud clusters.	Compute	Series A Completed	GPUs on io.net have served over 17k compute hours and earned $190k.
Together	Decentralized cloud platform for open-source AI models.	Pooled Compute	Series A Completed	Together's first service, Forge, is on pace to generate $20 million in annual revenue.
Worldcoin	Online identity project started by Sam Altman that verifies personhood via scanning users' irises.	Identity	$27,508	World App has attracted 5 million users in its first 6 months.
Bittensor	Decentralized incentive network built to reward the most useful AI models with $TAO.	Model Optimization	$5,677	Bittensor has grown to 80k accounts with 89% of the circulating $TAO staked.
Render	Decentralized GPU network on Solana facilitating rendering and machine learning work.	Compute	$1,558	Render network employs 600 GPU node operators which has rendered 1.3 million scenes.
Akash	Cosmos-based decentralized supercloud marketplace for CPU, GPU, memory, and storage.	Compute	$1,103	Akash facilitates ~$2000 of daily revenue for GPU suppliers on the network.
Hut 8	Digital asset miner and HPC infrastructure provider.	Compute	$855	Hut 8 generated $4.5 million of revenue from its HPC business in Q3 2023.
Fetch.ai	AI agent creation, customization, and monetization platform.	Agent Tooling	$780	59k unique wallets holding $FET.
Applied Digital	Bitcoin miner supporting AI operations through its subisdiary, Sai Computing, which hosts an AI Cloud Service.	Compute	$587	Applied Digital has 400 MW of capacity in development to support HPC applications.
Hive	Bitcoin miner providing HPC services.	Compute	$313	Hive's HPC business had 4,800 GPUs active at the end of 2023 with a target of 38,000 GPU-based cards committed to cloud service.
Altered State Machine	Decentralized protocol for training and trading AI agents that can be used with NFTs.	Agent Tooling	$110	Over 27k unique wallets holding NFTs that can be powered by Altered State Machine AI agents.

Source: VanEck Research, project websites, as of 1/15/24.

Links to third party websites are provided as a convenience and the inclusion of such links does not imply any endorsement, approval, investigation, verification or monitoring by us of any content or information contained within or accessible from the linked sites. By clicking on the link to a non-VanEck webpage, you acknowledge that you are entering a third-party website subject to its own terms and conditions. VanEck disclaims responsibility for content, legality of access or suitability of the third-party websites.

Disclosure: VanEck has a position in Together via our strategic partnership with early-stage venture manager Cadenza, who was kind enough to contribute to the* "Overcoming the Bottlenecks of Decentralized Compute" *section

Special Thanks to:

Jason Morton, CEO of ZKML

Ala Shaabana, Co-Founder of Bittensor

Arrash Yasavolian, Founder of Bittensor's Taoshi Subnet

Greg Osuri, CEO and Founder of Akash

Richard Liang, CEO of zkP2P

Key Members of the Sui Blockchain Team – Sam Blackshear, Nihar Shah, Sina Nader, Alonso Gortari

DISCLOSURES

Coin Definitions

- **Bitcoin (BTC)** is a decentralized digital currency, without a central bank or single administrator, that can be sent from user to user on the peer-to-peer bitcoin network without the need for intermediaries.
- **Ethereum (ETH)** is a decentralized, open-source blockchain with smart contract functionality. Ether is the native cryptocurrency of the platform. Amongst cryptocurrencies, Ether is second only to Bitcoin in market capitalization.
- **Arbitrum (ARB)** is a rollup chain designed to improve the scalability of Ethereum. It achieves this by bundling multiple transactions into a single transaction, thereby reducing the load on the Ethereum network.
- **Optimism (OP)** is a layer-two blockchain on top of Ethereum. Optimism benefits from the security of the Ethereum mainnet and helps scale the Ethereum ecosystem by using optimistic rollups.
- **Polygon (MATIC)** is the first well-structured, easy-to-use platform for Ethereum scaling and infrastructure development. Its core component is Polygon SDK, a modular, flexible framework that supports building multiple types of applications.
- **Solana (SOL)** is a public blockchain platform. It is open-source and decentralized, with consensus achieved using proof of stake and proof of history. Its internal cryptocurrency is SOL.
- **Curve (CRV)** is a decentralized exchange optimized for low slippage swaps between stablecoins or similar assets that peg to the same value.
- **Lido DAO (LDO)** is a liquid staking solution for Ethereum and other proof of stake chains.
- **Aave (AAVE)** is an open-source and non-custodial protocol to earn interest on deposits and borrow assets with a variable or stable interest rate.
- **ApeCoin (APE)** is a governance and utility token that grants its holders access to the ApeCoin DAO, a decentralized community of Web3 builders.
- **Decentraland (MANA)** is building a decentralized, blockchain-based virtual world for users to create, experience and monetize content and applications.
- **The Sandbox (SAND)** is a virtual world where players can build, own, and monetize their gaming experiences using non-fungible tokens (NFTs) and $SAND, the platform's utility token.
- **Binance Coin (BNB)** is digital asset native to the Binance blockchain and launched by the Binance online exchange.
- **Fantom (FTM)** is a directed acyclic graph (DAG) smart contract platform providing decentralized finance (DeFi) services to developers using its own bespoke consensus algorithm.
- **Stacks (STX)** provides software for internet ownership, which includes infrastructure and developer tools to power a computing network and ecosystem for decentralized applications (dApps).
- **Tron (TRX)** is a multi-purpose smart contract platform that enables the creation and deployment of decentralized applications.
- **Cosmos (ATOM)** is a cryptocurrency that powers an ecosystem of blockchains designed to scale and interoperate with each other. The team aims to "create an Internet of Blockchains, a network of blockchains able to communicate with each other in a decentralized way." Cosmos is a proof-of-stake chain. ATOM holders can stake their tokens in order to maintain the network and receive more ATOM as a reward.
- **Avalanche (AVAX)** is an open-source platform for launching decentralized finance applications and enterprise blockchain deployments in one interoperable, scalable ecosystem.
- **Maker (MKR)** is the governance token of the MakerDAO and Maker Protocol — respectively a decentralized organization and a software platform, both based on the Ethereum blockchain — that allows users to issue and manage the DAI stablecoin.
- **Osmosis (OSMO)** is an automated market-making protocol (AMM) that specializes in the Interchain DeFi movement and is built on its own blockchain, utilizing the Cosmos SDK and IBC technologies. Osmosis is an advanced protocol focused on customizable AMMs, where users can create, construct, design and deploy individual and highly-customized AMMs with various modules and the on-chain governance system.
- **Sui (SUI)** is a first-of-its-kind Layer 1 blockchain and smart contract platform designed from the bottom up to make digital asset ownership fast, private, secure, and accessible.
- **Polkadot (DOT)** is a sharded heterogeneous multi-chain architecture which enables external networks as well as customized layer one "parachains" to communicate, creating an interconnected internet of blockchains.
- **NEAR Protocol (NEAR)** is a layer-one blockchain that was designed as a community-run cloud computing platform and that eliminates some of the limitations that have been bogging competing blockchains, such as low transaction speeds, low throughput and poor interoperability.
- **Linea** is a network that scales the experience of Ethereum with out-of-the-box compatibility with the Ethereum Virtual Machine which enables the deployment of already existing applications.

Risk Considerations

This is not an offer to buy or sell, or a recommendation to buy or sell any of the securities, financial instruments or digital assets mentioned herein. The information presented does not involve the rendering of personalized investment, financial, legal, tax advice, or any call to action. Certain statements contained herein may constitute projections, forecasts and other forward-looking statements, which do not reflect actual results, are for illustrative purposes only, are valid as of the date of this communication, and are subject to change without notice. Actual future performance of any assets or industries mentioned are unknown. Information provided by third party sources are believed to be reliable and have not been independently verified for accuracy or completeness and cannot be guaranteed. VanEck does not guarantee the accuracy of third party data. The information herein represents the opinion of the author(s), but not necessarily those of VanEck or its other employees.

The information, valuation scenarios and price targets presented on any digital assets in this blog are not intended as financial advice, a recommendation to buy or sell these digital assets, or any call to action. There may be risks or other factors not accounted for in these scenarios that may impede the performance these digital assets; their actual future performance is unknown, and may differ significantly from any valuation scenarios or projections/ forecasts herein. Any projections, forecasts or forward-looking statements included herein are the results of a simulation based on our research, are valid as of the date of this communication and subject to change without notice, and are for illustrative purposes only. Please conduct your own research and draw your own conclusions.

Past performance is not an indication, or guarantee, of future results. Hypothetical or model performance results have certain inherent limitations. Unlike an actual performance record, simulated results do not represent actual trading, and accordingly, may have undercompensated or overcompensated for the impact, if any, of certain market factors such as market disruptions and lack of liquidity. In addition, hypothetical trading does not involve financial risk and no hypothetical trading record can completely account for the impact of financial risk in actual trading (for example, the ability to adhere to a particular trading program in spite of trading losses). Hypothetical or model performance is designed with benefit of hindsight.

Index performance is not representative of fund performance. It is not possible to invest directly in an index.

Investments in digital assets and Web3 companies are highly speculative and involve a high degree of risk. These risks include, but are not limited to: the technology is new and many of its uses may be untested; intense competition; slow adoption rates and the potential for product obsolescence; volatility and limited liquidity, including but not limited to, inability to liquidate a position; loss or destruction of key(s) to access accounts or the blockchain; reliance on digital wallets; reliance on unregulated markets and exchanges; reliance on the internet; cybersecurity risks; and the lack of regulation and the potential for new laws and regulation that may be difficult to predict. Moreover, the extent to which Web3 companies or digital assets utilize blockchain technology may vary, and it is possible that even widespread adoption of blockchain technology may not result in a material increase in the value of such companies or digital assets.

Digital asset prices are highly volatile, and the value of digital assets, and Web3 companies, can rise or fall dramatically and quickly. If their value goes down, there's no guarantee that it will rise again. As a result, there is a significant risk of loss of your entire principal investment.

Digital assets are not generally backed or supported by any government or central bank and are not covered by FDIC or SIPC insurance. Accounts at digital asset custodians and exchanges are not protected by SPIC and are not FDIC insured. Furthermore, markets and exchanges for digital assets are not regulated with the same controls or customer protections available in traditional equity, option, futures, or foreign exchange investing.

Digital assets include, but are not limited to, cryptocurrencies, tokens, NFTs, assets stored or created using blockchain technology, and other Web3 products.

Web3 companies include but are not limited to, companies that involve the development, innovation, and/or utilization of blockchain, digital assets, or crypto technologies.

All investing is subject to risk, including the possible loss of the money you invest. As with any investment strategy, there is no guarantee that investment objectives will be met and investors may lose money. Diversification does not ensure a profit or protect against a loss in a declining market. Past performance is no guarantee of future performance.

© Van Eck Associates Corporation.

VanEck®

Van Eck Securities Corporation, Distributor
A wholly-owned subsidiary of Van Eck Associates Corporation

666 Third Avenue | New York, NY 10017
vaneck.com | 800.826.2333

Exchange-Traded Funds
Mutual Funds
Institutional Funds
Model Delivery
Separately Managed Accounts
UCITS Funds
UCITS Exchange-Traded Funds