

# HYPERDUST白皮书

## 一、为什么比特币L2和AI互为补充？

### 1. 比特币L2作为AI智能体的支付手段

比特币体现了“最大中立”的媒介，适合从事价值交易的人工智能代理。比特币可以消除法定货币固有的低效率和“摩擦”，人工智能代理将采用这种“数字原生”媒介进行价值交换。比特币L2提升了比特币可编程的性能，可以满足人工智能的速度，比特币可望成为人工智能的原生货币。

### 2. 比特币L2可以实现去中心化AI治理

由于目前AI的中心化趋势，让AI对齐与治理的去中心化备受关注。比特币L2更为强大的智能合约，可以成为规范AI智能体行为和协议模式的规则，实现去中心化的AI对齐与治理模式。而且比特币的最大中立特性更容易达成AI对齐和治理共识。

### 3. 比特币L2可以发行AI资产

除了可以将AI智能体发行为比特币L1的资产，高性能的比特币L2可以满足AI智能体发行AI资产的需要，这将构成智能体经济的基础。

### 4. AI智能体是比特币以及比特币L2的杀手级应用

由于性能原因，比特币自诞生以来除了储值之外一直没有落地的应用，进入L2的比特币将有更强大的可编程能力，AI智能体一般是用于解决真实世界问题，所以比特币驱动的AI智能体可以真正应用起来，而且AI智能体的规模和频率将成为比特币和L2的杀手级应用。虽然人类经济可能不会优先考虑将比特币作为一种支付方式，但机器人经济也许会优先考虑。大量的AI智能体，24/7 为您服务，不知疲倦地使用比特币进行和接受微小的支付。比特币的需求可能会以目前难以想象的方式大幅增加。

## 5. AI计算增强比特币L2的安全性

AI计算可以作为比特币PoW的补充, 甚至用PoUW取代PoW, 这将颠覆性地在保证安全的前提下, 还将目前用于比特币挖矿的能量注入了AI智能体。AI可以通过L2让比特币成为智能驱动的绿色区块链, 而不是类似以太坊的PoS。我们提出的超图共识(Hypergraph Consensus)就是基于3D/AI计算的PoUW, 后面会有介绍。

## 二、现有去中心化AI相关项目一览

HYPERDUST项目目前在Web3 AI领域独树一帜, GPU算力共识化, AI具身化资产化, 是去中心化的半AI半金融应用。通过“三化”满足了去中心化AI平台应该具备五个特性, 我们也按照这五个特征对现有的去中心化AI相关项目进行简要的梳理和对比。

去中心化AI平台应该具备五个特性分别是:

### (i) 远程运行AI模型的可验证性 *verifiability of remotely run models*

去中心化可验证性包括Data Availability, ZK等技术

### (ii) 开放AI模型API的可用性 *usability of publicly available AI models,*

可用性主要看提供AI模型(主要指LLM)API节点是否是Peer 2 Peer, 完全去中心化的网络

### (iii) AI开发者与用户的激励机制 *incentivization for AI developers and users,*

激励中是否有公平的代币生成机制

### (iv) 数字社会中可行的全球AI治理方案 *global governance of essential solutions in the digital society,*

AI治理是否中立容易达成共识

### (v) 不被供应商所绑架 *no vendor lockins, etc.*

是否是完全去中心化的平台

用这五条特性对现有公开计划或落地项目中可以进行如下梳理, (去中心化联邦学习经过多年实践, 未取得重大进展, LLM的兴起更让这种去中心化训练变得困难, 不再列出相关项目。)

### (i) 远程运行AI模型的可验证性

我们认为可验证性是去中心化AI项目必备的特性, 是后续可用性、激励、治理和非绑定的基础, 如果没有可验证性就不再考察其他维度的特性。没有可验证性特性的项目可能是去中心化的项目, 比如去中心化算力租赁或数据、算法、模型市场, 而不是AI的去中心化。

可能满足可验证性的项目有:

Giza, 基于ZKML共识机制, 满足远程运行AI模型的可验证性, 但目前性能比较差, 特别离大模型要求还相差几个数量级。百万级参数小模型的一次证明过程都需要几分钟。LLM模型的证明过程需要时间是不可接受的。

Cortex AI, 5年前启动的专注于去中心化AI的L1公链项目, 技术比较复杂, 向EVM虚拟机增加新的指令以满足神经网络计算的需要, 其底层还是基于ZK的可用性验证, 可用于简单的AI模型, 但无法满足LLM级别的大模型的需要。

Ofelimos, 加密学术界第一次提出PoUW的方案, 采用了特定搜索算法。但该算法并未和具体应用或项目关联。

Project PAI, 在一篇论文中提到PoUW但仅有白皮书, 没有产品, <https://oben.me/>。

Qubic, 宣称PoUW, 提出用数百个GPU进行人工神经网络计算, 但python人工神经网络库的简单计算的意义不明, 似乎不能满足LLM训练或者推理的需求, 也没有起到PoUW的作用。

FLUX, (PoW ZelHash, 并非PoUW)

Coinai, (论文阶段)<https://aipowergrid.io/>, 任务指派, 无严格共识机制

不能满足：

GPU算力租赁类项目，都缺乏去中心化可验证机制，无法保证远程运行AI模型的可验证性。

DeepBrain Chain, 专注于GPU租赁, 2017年的L1项目, 2021年主网上线；

EMC, 任务指派中心化奖励, 路线图中没有去中心化共识机制；

Atheir, 未见共识机制；

IO.NET, 未见共识机制；

CLORE.AI, POH

众包模式、AI模型链上发布支付、发行NFT, AI 链下运行, 没有可验证性,

比如: SingularityNET, Bittensor, AINN, Fetch.ai, oceanprotocol, algovera.ai。

## **(ii)开放AI模型API的可用性**

Cortex AI, 未见支持LLM

Qubic, 未见支持LLM

上述的所有去中心化AI项目都无法很好地回应上面五个问题, HYPERDUST是完全去中心化的AI协议, 基于Hypergraph PoUW共识机制以及完全去中心化的比特币L2 Stack, 将来还升级为比特币AI专用L2。

PoUW 用于以最安全的方式保护网络, 同时也不浪费计算能力, 矿工提供的全部计算能力可以用于LLM推理, 以及云渲染服务。 PoUW 的愿景是计算能力可以用于提交到去中心化网络的各种问题。

### 三、HYPERDUST的定位

HYPERDUST是完全去中心化的大模型驱动AI智能体训练和应用平台(支持智能体训练,不含大模型训练),是去中心化的半AI半金融应用平台,具有GPU算力共识化, AI具身化资产化的特点,目前基于Hypergraph PoUW共识机制以及完全去中心化的比特币L2,将来还借助比特币L2 Stack升级为比特币AI专用L2。

### 四、HYPERDUST的诞生背景

#### 1.LLM及应用的爆发

OpenAI 的ChatGPT在3个月就达到一亿用户,从此语言大模型LLM的开发、应用以及投资热潮席卷全球。但到目前为止,LLM的技术与训练都还是以非常中心化的方式进行,这一点开始引起学界,业内以及公众的高度关注,担心几个主要技术提供者对AI技术的垄断、数据隐私泄露、侵占以及云计算公司的供应商锁定等问题,这些问题本质上都是因为现在的互联网及应用入口还是由中心化平台所控制,没有适合AI大规模应用的网络。AI社区开始实现一些本地运行和去中心化的AI项目,本地运行的代表是Ollama,去中心化的代表是Petals。Ollama通过参数压缩或降低精度的方式让中小规模LLM可以在个人电脑甚至手机上运行,这保护了用户数据的隐私权和其他权益,但是显然难以支持生产环境和联网应用。Petals则是通过Bittorrent的Peer2Peer技术实现了LLM的完全去中心化推理。但是Petals缺乏共识层和激励层协议,还只是停留在研究者的小圈子。

#### 2.LLM驱动的智能体

LLM的加持可以让智能体可以进行上层的推理和一定规划能力,借助自然语言多智能体之间也可以像人类一样形成社会协作。一些LLM驱动的智能体框架被提出来,比如微软的AutoGen, Langchain, CrewAI等。目前有大量AI创业者和开发者是集中在LLM驱动智能体及其应用方向,对稳定可用、规模弹性的LLM推理有大量需求,但目前主要还是向云计算公司租用GPU推理实例来完成,英伟达公司在202433月年发布了包括LLM在内的生成式AI微服务平台ai.nvidia.com,以满足巨大的需求,但目前还没有正式上线。LLM驱动的智能体如同当年的网站建设如火如荼,但目前主要还是以传统Web2的模式在协作,智能体开发

者需要租赁GPU或采购LLM提供者的API支持这些智能体的运行，这就存在较大的摩擦，不利于智能体生态的快速成长和智能体经济的价值传递。

### 3.具身智能体仿真环境

目前大多数智能体能访问和操作的对象仅仅是一些API，或者通过代码或脚本与这些API交互，写入LLM生成的控制指令或读出外部状态。通用智能体不仅仅可以理解和生成自然语言，还应该能理解人类世界，甚至经过相应训练后还可以迁移到机器人（比如无人机、扫地机、人形机器人等）系统中，完成制定的任务，这样的智能体被成为具身智能体。具身智能体的训练需要大量现实世界的视觉数据，更好地理解特定的环境和现实世界，缩短机器人的训练和开发时间，降低成本。目前这些用于训练具身智能的仿真环境还被一些公司建设和拥有，比如微软的Mincraft，英伟达的Issac Gym等，没有去中心化的环境以满足具身智能的训练。最近一些游戏引擎开始重视人工智能，比如Epic的虚幻引擎开始推进符合OpenAI GYM的AI训练环境。这些智能体的仿真和训练环境需要大量的GPU计算能力。

### 4.比特币L2生态

虽然比特币侧链多年前就存在但主要用于支付，智能合约的缺乏支撑不了复杂的链上应用，但由于EVM兼容的比特币L2的出现让比特币可以通过L2支持去中心化AI这样的应用。而去中心化AI需要完全去中心化，算力至上的区块链网络，而不是受限于越来越中心化的PoS区块链网络。

### 5.HYPERDUST团队20年的AI研究与3年生成式AI应用探索

HYPERDUST团队在3年前就利用生成AI生成2D图片和3D模型构建了由数千个AI生成岛屿组成的开放世界MOSSAI，并提出了非同质化生成的加密资产标准。但是当时的AI模型训练和生成还没有形成去中心化的解决方案，仅靠平台自身的GPU资源不能支撑大量用户使用，没有形成爆发。但是随着LLM点燃公众对AI的兴趣，于是推出了HYPERDUST去中心化AI应用平台，于2024Q1开始分别在以太坊和比特币L2进行测试，不仅可以激励GPU提供者通过云渲染、AI生成推理挖矿，也可以降低AI智能体开发者与用户的交易摩擦与进入门槛。

## 五、HYPERDUST的技术框架及解决方案

### 1. 如何实现去中心化的LLM驱动AI智能体应用平台？

现在去中心化AI最大的问题是对AI大模型远程推理，以及具身智能体的训练和推理没有高性能，低开销overhead的可验证算法，如果缺失了可验证性，系统就只能退化为传统的供需双方加平台方的多方市场模式，无法实现完全去中心化的AI应用平台。

可验证的AI计算是需要有PoUW共识算法，以此为基础可以实现去中心化的激励机制。具体来说就是激励中代币的mint调用是由节点完成计算任务并提交可验证的结果后自行完成的，而不是由任何中心化的方式将代币转移给节点。

要实现AI计算可验证首先需要定义AI计算。AI计算有很多层次，比如最底层的机器指令，CUDA指令，C++，Python语言，同样的在具身智能体训练中需要的3D计算也有不同的层次，着色器语言，OpenGL，C++，蓝图脚本等。

HYPERDUST的PoUW共识算法由计算图来实现，计算图Computational Graph，被定义为有向图，其中节点对应于数学运算。计算图是表达和评估数学表达式的一种方式。是一种描述方程的“语言”，包含节点（变量），边（操作（简单函数））。

**1.1 用计算图定义可验证计算**（目前我们实现了3D和AI计算，可扩展），不同层次的计算可以用子图来表示。这样可以涵盖各种不同类型的计算，并且通过子图表达不同的计算层次。目前是两层，顶层计算图部署上链，便于验证节点进行验证。

**1.2 将LLM模型和3D场景关卡用完全去中心化的方式进行加载并运行。**当有用户访问LLM模型进行推理，或进入3D场景进行渲染时，HYPERDUST智能体会启动另外一个可信节点运行同样的超图（LLM或3D场景）。

**1.3 如果验证节点发现某节点提交结果与可信节点提交结果不一致，则对第二层计算图（子图）的链下计算结果进行二分法查找，定位到出现分歧的子图计算节点（算子），子图算子**

已经预先部署到智能合约，带上发生不一致算子的参数，调用智能合约执行该算子即可验证结果。

## 2. 如何避免性过高的计算开销？

可验证AI计算的另一个难题是控制额外计算开销，我们知道拜占庭共识协议中需要2/3节点一直才形成共识，对AI推理共识来说意味着需要所有节点完成同样的计算。这样的额外开销在AI计算中是不可接受的浪费。Hyperdust做只需要到1- m个节点进行额外计算即可完成验证。

**2.1** 每一个LLM都不会单独进行推理，HYPERDUST系统会至少启动一台可信节点进行“陪伴计算”。

因为LLM推理计算是由模型中深度神经网络各个层是以上一层的计算结果作为输入进行计算，直到最后完成推理，多用户可以并发访问同一个LLM大模型。

所以最多要额外启动与LLM数量m相同的可信节点。最少则仅需要一台可信节点进行“陪伴计算”。

**2.2** 3D场景渲染计算也类似，每一个用户进入场景激活超图，系统都会根据超图加载一台可信节点，进行相应的超图计算。如果m个用户都进入不同的3D场景，则最多启动m个“陪伴计算”可信节点。

综上所述，参与额外计算的节点数是小于n或m，大于1的随机数。符合高斯分布。n是进入3D场景的用户数，m是LLM的数量。

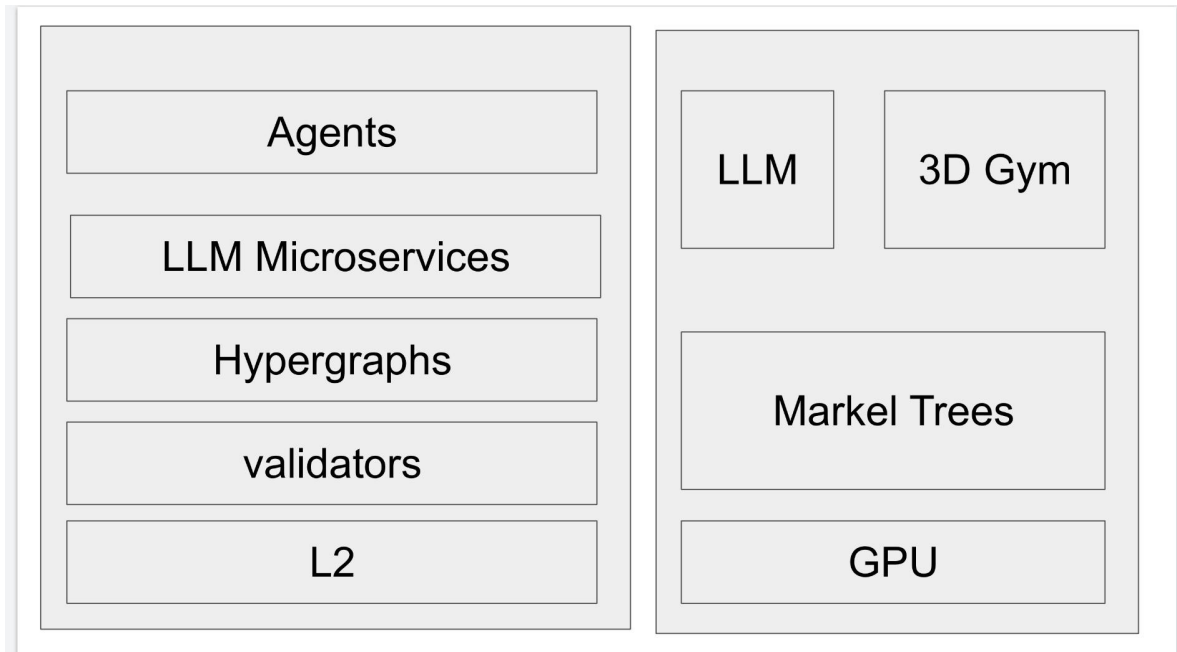
## 3. AI如何与Web3结合形成半AI半金融应用？

**AI智能体也能发币了**

AI开发者将智能体部署为智能合约，合约包含第一层超图链上数据，用户或其他智能体调用该智能体合约的方法并可能需要支付相应代币，提供服务的智能体就一定会完成相应的计算并提交可验证的结果。用户或其他智能体就与该智能体完成了去中心化的业务交易。

智能体不会担心完成业务后收不到代币，支付方不应担心给了代币得不到正确的业务计算结果。至于智能体业务本身的能力和 value 则由智能体资产 (20FT, 721 或 1155NFT) 的二级市场价格和市值决定。

整体架构图



## 六、HYPERDUST的愿景及市场空间

PoUW 的愿景是计算能力可以用于提交到去中心化网络的各种问题。HYPERDUST团队通过计算图对远程执行的LLM推理, 3D渲染计算进行PoUW验证, 矿工可以提供正确的计算结果获得相应的代币激励, 代币表示对AI网络的使用权, 随着挖矿难度的增加, 每一个代币对应的PoUW算力也会增加。类比特币PoW算力的增长曲线, HYPERDUST PoUW算力的增长可以支撑AI智能体全球应用, 代币价值也会极大提升。