

# 恒智鏈

## HANGKEI A.I.CHAIN

項目白皮書

HAI WHITER PAPER V1.0



# 目錄

1 背景 .....	1
1.1 AI 所需的算力資源將呈現指數級增長 .....	1
2 構造一種高性能、低成本的大規模 AI 算力網絡 .....	2
2.1 高性能、低成本、專為 AI 設計的硬件架構 .....	2
2.1.1 高性能 .....	2
2.1.2 低成本 .....	3
2.2 構建簡單、彈性、開放的人工智能運行環境 .....	3
2.2.1 自定義模型 .....	3
2.2.2 可遷移學習 .....	3
2.2.3 神經搜索網絡 .....	4
2.3 針對 AI 的托管服務 HAI ML Engine .....	4
2.3.1 可部署多個機器學習框架 .....	4
2.3.2 即用即取 .....	4
3 區塊鏈技術確保恒智鏈的中立性 .....	5
3.1 加密區塊 .....	5
3.2 資源證明 .....	6
3.3 時空證明 .....	6
3.4 交易證明 .....	6
3.5 結果證明 .....	6
4 恒智鏈的共識機制 POAI .....	7
5 Token 資源與分配 .....	7
5.1 參與 Token 分配的角色 .....	7
5.1.1 需求方 .....	7
5.1.2 檢索方 .....	7
5.1.3 供應方 .....	8
5.2 Token 分配計劃 .....	8
5.3 智能礦池算力與收益 .....	8
6 團隊介紹 .....	8
7 發展計劃 .....	8

## 1 背景

恒智鏈（HAI）團隊觀察到，如今 AI 技術迎來爆發時期，AlphaGo 在圍棋上戰勝了當今全世界最厲害的旗手，向這個世界展示了 AI 已經可以走向實用化。AI 作為基礎技術，幾乎可以和任何產業進行結合，也基於此，不論是歐美髮達國家，還是日本、韓國，還是中國，都有大量 AI 企業湧現。美國創業項目資料庫 CrunchBase 目前已經登記有 6419 家 AI 美國創業團隊，而中國創業資料庫 IT 桔子則登記則有 1874 家中國 AI 創業團隊。



IT 桔子的 AI 特輯，表明 AI 所服務的行業維度

### 1.1 AI 所需的算力資源將呈現指數級增長

AI 大規模應用所引發的結果，除了以後我們所處商業、產品會變得越來越自動化、智能以外，或者說機器人變得越來越聰明，可以理解我們人類的想法外，還有一個結果，那就是未來所需的 AI 運算資源越來越龐大，而且，隨著 AI 行業的發展，AI 運算資源的需求將呈現指數級增長。

中國工程院院士、浪潮集團首席科學家王恩東在烏鎮大會上表示，“在 AI 模型訓練過程中，人工參數調整耗時與機器運行耗時大約分別佔 80%和 20%，如果平台性能提高一倍，那麼迭代週期就能縮短 10%左右。因此，人工智能所需計算量一直處於高速增長狀態。另據 OpenAI 的分析報告，從 2012 年開始，AI 訓練所用的計算量呈現指數增長，平均每 3.43 個月便會翻倍，截止目前計算量擴大了 30 萬倍。據估計，到 2020 年前，AI 所需的計算量還會繼續增長 12 倍。”

正如工程院院士王恩東的觀點，“要贏得創新速度，AI 算力、算法等投資至關重要。”而在未來，除了數學家繼續在算法方面優化外，對算力的投資不可缺少。

所以恒智鏈團隊提出一種高性能、低成本的大規模 AI 算力網絡，解決算力資源緊缺的問題，幫助 AI 初創團隊、AI 大型企業的業務順利發展。

## 2 構造一種高性能、低成本的大規模 AI 算力網絡

恒智鏈所提出的大規模 AI 算力網絡，將具備如下幾個特性：支持大規模運算、拓展容易、構造成本低廉、即用即取。並採用獨有的專為 AI 運算所提供的軟硬架構，為未來大規模 AI 算力需求做好準備。

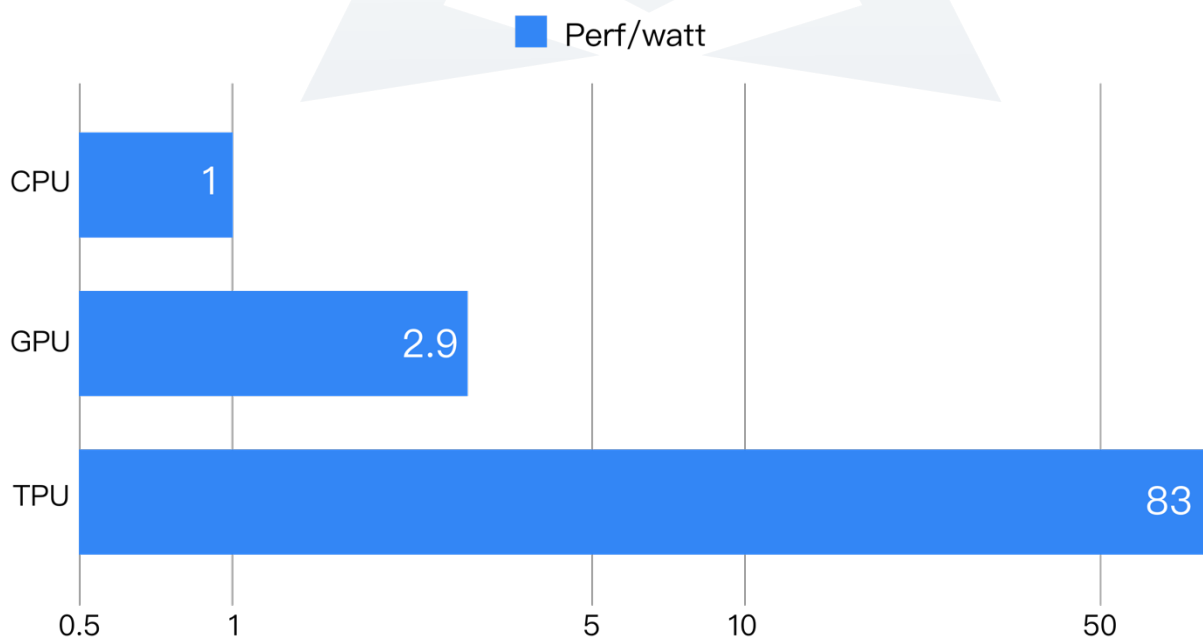
### 2.1 高性能、低成本、專為 AI 設計的硬件架構

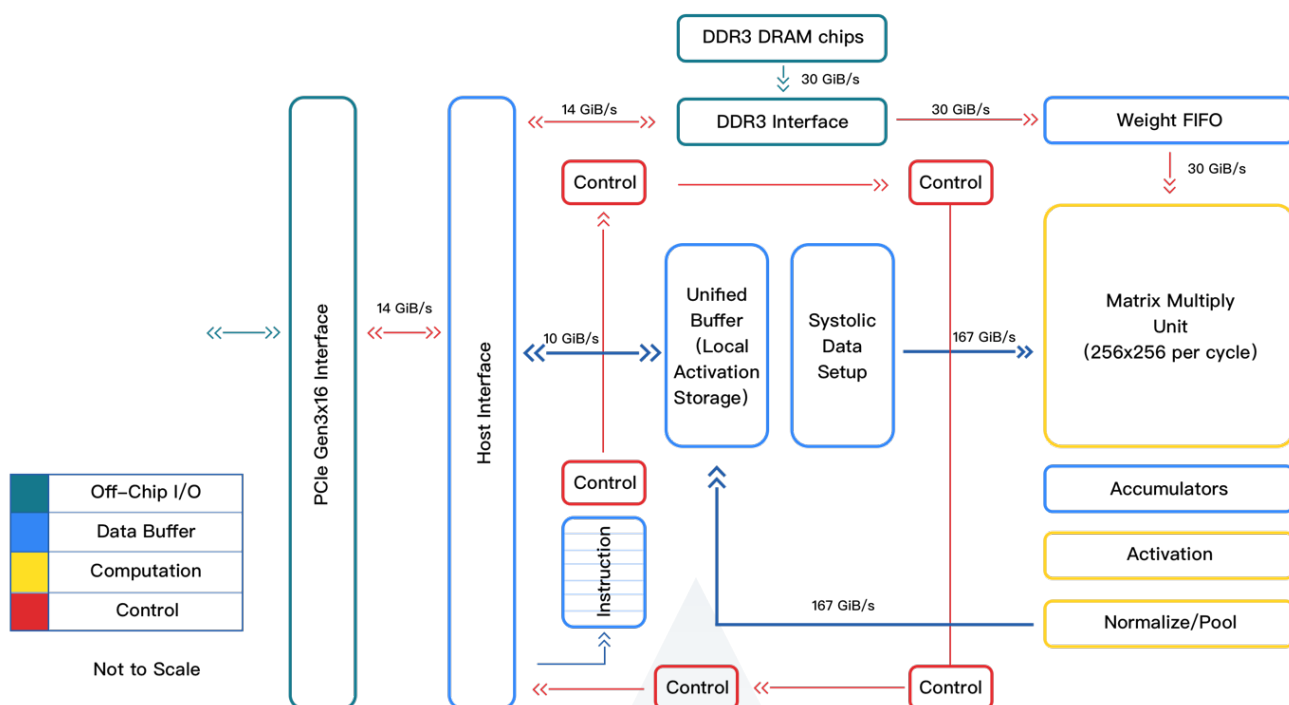
總所周知，機器學習是 AI 技術的底層，而恒智鏈將採用專用的，可加快機器學習速度的 AI 芯片 HAI HPU，具備強勁的浮點運算性能，比傳統的 CPU、GPU 快 30 倍，讓機器學習所訓練的模型的速度更快。

#### 2.1.1 高性能

恒智鏈（HAI）提出 HAI HPU 專用芯片設計的靈感，來自 Google 的 TPU。Google 在設計該專用芯片時，採用了一個特殊的做法，就是採用一種使用 8 位整數來近似預設的最小值和最大值之間任意數值的優化技術，以降低機器學習所需要的成本。這種思想的提出，讓 AI 專用芯片的提出成為可能。

HAI HPU 採用名為“脈動陣列”（systolic array）提高計算性能，這個架構的特殊性在於將多個運算邏輯單元串聯在一起，復用從一個寄存器中讀取的結果。這樣子令 HPU 每一次計算週期可執行的計算量大大增加。舉個例子，一個脈動陣列包含  $256 \times 256 = 65,536$  個邏輯運算單元，也就是說一顆 HPU 芯片可以處理 65,536 次 8 位整數的乘法和加法。當 HPU 以 700MHz 運行，也就是說，它每秒可以運行  $65,536 \times 700,000,000 = 46 \times 10^{12}$  次乘法和加法運算，或每秒 92 萬億（ $92 \times 10^{12}$ ）次矩陣單元中的運算。





### 2.1.2 低成本

由於 HAI HPU 是專門為 AI 所優化的芯片，它只用於一個用途，那就是提供機器學習所需要龐大的運算量。因此也帶來另外一個好處，那就芯片的結構設計相對簡單，HPU 的大小只有其它芯片的一半，而且芯片上的控制單元僅佔芯片面積的一半。

## 2.2 構建簡單、彈性、開放的人工智能運行環境

除了專用 AI 芯片外，恒智鏈將發展 HAI HPU 相配套的大規模、高性能機器訓練環境 HAI AutoML。

### 2.2.1 自定義模型

HAI AutoAL 是機器學習產品，可以讓開發者訓練自定義機器學習模型，而且工作量不高，和一點點機器學習的專業知識。

### 2.2.2 可遷移學習

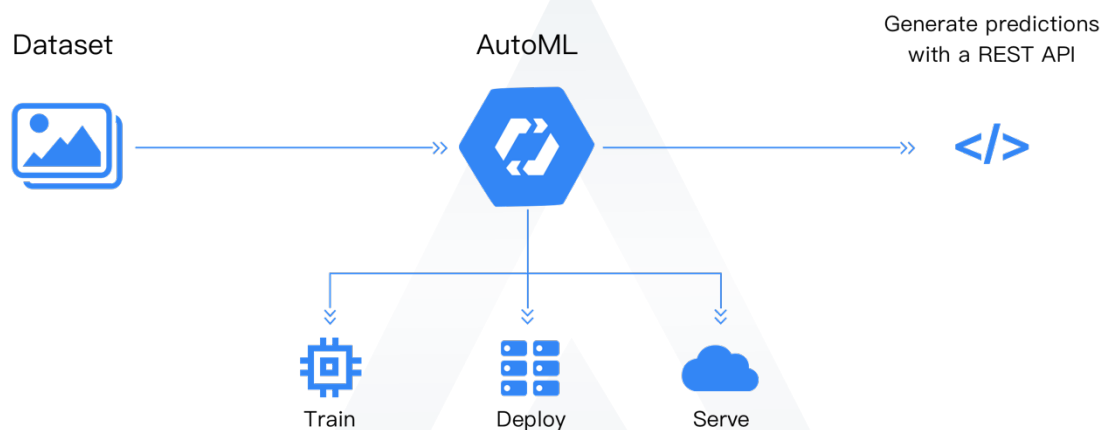
可遷移學習是機器學習的一個分支，這種特性可以做到把已學訓練好的模型參數遷移到新的模型來幫助新模型訓練。這種方式，可以節約重複輸入同類型數據，並且為這些數據重新標記標籤的成本。考慮到大部分數據或任務是存在相關性的，所以通過遷移學習我們可以將已經學到的模型參數（也可理解為模型學到的知識）通過某種方式來分享給新模型從而加快並優化模型的學習效率不用像大多數網絡那樣從零學習（starting from scratch, tabula rasa），而這某種方式就是可遷移學習。

### 2.2.3 神經搜索網絡

Sundar Pichai 在博客中寫道：“設計神經網絡是非常耗時的，其對專業知識的極高要求使得只有小部分科研人員和工程師才能參與設計。這就是我們創建 AutoML 方法的原因，有了它，神經網絡也可以設計神經網絡。”

“神經網絡也可以設計神經網絡”這就是神經搜索網絡，通常使用強化學習或進化算法來設計新的神經網絡網絡結構。這很有用，因為它使得我們能夠發現比人們想象的要複雜得多的網絡結構，並且這些網絡結構可以針對特定目標進行優化。神經網絡結構搜索通常需要大量計算力。

HAI AutoML 將為神經搜索網絡提供充足的算力資源，並兼容多種神經搜索網絡的結構，以方便他人生成結構更優、效率更高的網絡結構。



AutoML 的工作原理

## 2.3 針對 AI 的托管服務 HAI ML Engine

HAI ML Engine 是一項托管式服務，能讓開發者和數據科學家構建卓越的機器學習模型，並運用到實際的產品之中。

### 2.3.1 可部署多個機器學習框架

HAI ML Engine 可以部署包括 scikit-learn、XGBoost、Keras 以及 TensorFlow 在內的多個機器學習框架來構建自己的模型。無需專家的幫助，HAI ML Engine 可以自動設計和評估模型架構，從而更快地實現智能解決方案。它可以在托管集群上大規模地訓練任何模型，還可以輕鬆擴容，幫助您利用自己的所有數據。

### 2.3.2 即用即取

開發者只需要專注於模型開發和部署，而不必擔心基礎架構問題，HAI ML Engine 會自動執行所有資源預配和監控工作。利用支持 CPU、GPU 和 HPU 的分布式托管型訓練基礎架構來構建模型。通過在許多節點中進行訓練或並行運行多個實驗，您可以加快模型開發速度。



### 3 區塊鏈技術確保恒智鏈的中立性

目前 AI 算力資源網絡集中掌握在大公司的手裡，在美國比如 Google、亞馬遜、微軟等等，在中國則比如阿里巴巴、騰訊、百度等等。大公司具備更強的資源來構建人工智能網絡，但隨著網絡科技公司的邊界擴大，他們無法保證人工智能平台的中立性。對於其它使用人工智能服務的企業而言，把數據開放給阿里巴巴、騰訊、百度，會存在數據洩密的風險，無法保證用戶隱私不被洩露，無法保證商業機密不被洩露。最近，由於擔心被數據被亞馬遜利用，老牌零售企業沃爾瑪決定逐步降低對亞馬遜雲服務的需求，而改為自建數據中心。

因此恒智鏈需要使用區塊鏈這一去中心化的技術，確保我們所搭建的大規模 AI 算力網絡具備中立性。使用區塊鏈技術，將為企業創建加密的人工智能運算網絡，這對需要人工智能服務的企業來說意義巨大。

區塊鏈將為人工智能網絡帶來幾個改變：一、改變數據孤島的現狀，讓數據共享成為現實；二、基於不可篡改賬本技術，可識別垃圾數據；三、設立資源交換中心，促使鏈上不同主體合作。

#### 3.1 加密區塊

抗碰撞哈希運算：我們使用了一個防碰撞的哈希函數  $CRH: \{0,1\}^* \rightarrow \{0,1\}^{O(\lambda)}$ 。我們還使用了一個防碰撞哈希函數 MerkleCRH，可以把字符串分割成多個部分，再構造出一個二叉樹並遞歸採用 CRH，然後輸出數根。

Zk-SNARKs：我們的加密區塊證明依賴於“零知識下簡明的非交互知識論證”（以下縮寫為 zk-SNARKs），因為 Zk-SNARKs 簡介扼要，證明過程短而且容易驗證。

因為 zk-SNARKs 是簡潔的，所以證明很短並且很容易驗證。更正式地，讓  $L$  為一種 NP 語言， $C$  是  $L$  的決策流程。一個新人方引導了一個一次性的設置階段，這產生了兩個公鑰：

一個證明密鑰  $pk$  和一個驗證密鑰  $vk$ 。證明密鑰  $pk$  使任何(不受信任的)的證明者生成一個證明  $\pi$ ，對於她選擇的一個實例  $x$ ， $x \in L$ 。

這個非交互式證明  $\pi$  是零知識和知識證明。對於她所選擇的一個實例  $x$ ， $x \in L$ 。這個非交互式證明  $\pi$  既是零知識，又是只是證明。任何人都可以使用驗證秘鑰  $vk$  驗證  $\pi$ ，而不與產生  $\pi$  的證明者進行交互。證明  $\pi$  具有恒定的大小，並且可以在  $|x|$  中線性的時間內驗證。

對於 zk-SNARK 的流程滿足性是多項式時間算法的三重元組：

(KeyGen, Prove, Verify)

- KeyGen( $1\lambda, C$ )  $\rightarrow$  ( $pk, vk$ )，輸入安全參數  $\lambda$  和電路  $C$ ，KeyGen 產生概率樣本  $pk$  和  $vk$ 。這兩個鍵作為公共參數發佈，可在  $L_c$  上用於證明/驗證。

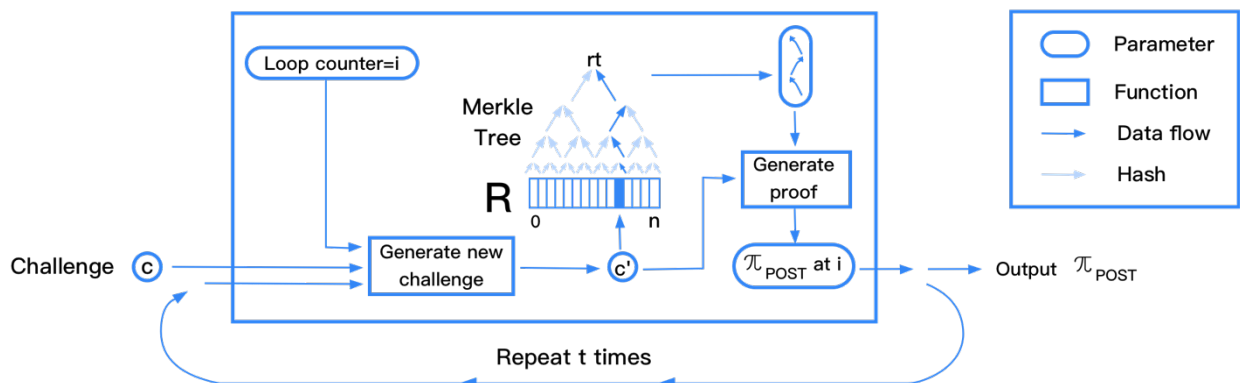
- $\text{Prove}(\text{pk}, x, w) \rightarrow \pi$  在輸入  $\text{pk}$ 、輸入  $x$  和 NP 聲明  $w$  的見證時，證明人為語句  $x \in \text{LC}$  輸出非交互式證明  $\pi$ 。
- $\text{Verify}(\text{vk}, x, \pi) \rightarrow \{0, 1\}$  當輸入  $\text{vk}$ ，輸入  $x$  和證明  $\pi$ ，驗證者驗證輸出 1 是否滿足  $x \in \text{LC}$ 。

### 3.2 資源證明

恒智鏈（HAI）將鏈上具備算力資源的節點，稱之為資源節點。擁有恒智鏈（HAI）Token 的用戶，可用 Token 來兌換恒智鏈上的資源節點，或用 Token 來參與“超級節點”的投票。而“資源證明”則是證明某一賬戶所有用計算資源的證明。“資源證明”將是個一張資源表提供給檢索方，由檢索方來檢索恒智鏈鏈上的資源節點。

### 3.3 時空證明

每一次外部購買/租用恒智鏈的 AI 算力網絡，都會以“訂單”的方式發送到檢索方，並由檢索方針對全網的運算資源，和訂單上的需求進行自動適配，此時，將產生“時空證明”，來證明計算該 AI 任務所需要的時間。並以該證明來計算需求方所需支付的 Token。



### 3.4 交易證明

當需求方希望購買/租用恒智鏈（HAI）的 AI 算力網絡以完成某項 AI 計算任務，經由檢索方自動化檢索全網資源節點，並尋找到對應算力資源後，該交易證明成立，並記錄進區塊。

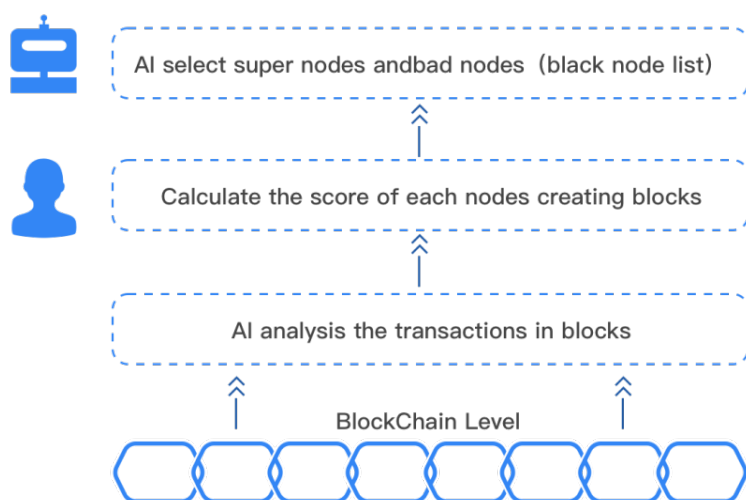
### 3.5 結果證明

當資源節點經過計算產生了計算結果，並將結果返回給需求方之後，將自動生成“結果證明”。檢索方將記錄“結果證明”。



## 4 恒智鏈的共識機制 POAI

恒智鏈利用 AI 機器學習方式來對歷史的區塊數據進行統計分析，來決策挑選出超級節點，採用 POA+POW 混合的共識機制，我們稱之為 POAI。分析挑選的過程如下：



這個過程完全是看各個挖礦記賬節點的 POW 的工作結果，來挑選出合格的節點成為超級節點；而超級節點本身是利用 POA 的機制來保障高速度的區塊創建速度，來提升交易確認速度。

所有人都有機會參與挖礦並成為超級節點，只要在 POW 模式下做出了出色於其他節點的記賬貢獻，AI 機器人會分析並選擇這個節點進入超級節點隊列。當所有超級節點全部被乾掉無法工作時，其他所有節點依然有機會進行 POW 挖礦，並且會在下一輪，被 AI 機器人選擇成為新的超級節點，繼續工作。HAICHAIN 通過能夠自動適應網絡環境，交易數量較多的地區附近的節點，越有機會成為超級節點。達到自我調整和適應的目的。

## 5 Token 資源與分配

### 5.1 參與 Token 分配的角色

#### 5.1.1 需求方

需要計算資源的需求方，從市場獲得恒智鏈的 Token，以向恒智鏈的計算資源提出使用請求。

#### 5.1.2 檢索方

當需求方發出計算需求，即發出訂單時，檢索方將驗證訂單有效性，並在恒智鏈的資源當中檢索相應可提供服務的供應方。

此時檢索方將通過“超級節點”和“資源證明”來檢索全網算力資源，以達到自動調配目的。

#### 5.1.3 供應方

供應方接收檢索方所驗證的訂單，並提供“交易證明”給檢索方，確定接收訂單上的計算需求。

當計算得到結果，並經由檢索方發送返回需求方，需求方自動確認時，則產生“結果證明”完整一次計算任務。

## 5.2 Token 分配計劃

HAI 是去中心化的區塊鏈數字資產，是基於恒智鏈發行的數字代幣，應用於恒智鏈平台的所有應用場景中。HAI 初期將採用 POW 共識機制，代幣總量為 10,000,000,000 枚。初期代幣由 ERC20 合約生成，在恒智鏈主鏈正式運行後，將安排替換為主鏈代幣。

HAI 具體分配細節：

比例	分配方案	額度	分配細節
90%	礦工激勵	9,000,000,000	
6%	創始人團隊	600,000,000	鎖倉 5 年後每天釋放 0.01%
1%	開發和運維團隊	100,000,000	鎖倉 4 年後每天釋放 0.01%
1%	基金會	100,000,000	鎖倉 3 年後每天釋放 0.01%
1%	運營團隊	100,000,000	每天釋放 0.0002%
1%	市場推廣	100,000,000	稍後公布

## 6 團隊介紹

(即將公佈)

## 7 發展計劃



2019 年 Q3&Q4 規劃

1、HAI AUTOML 部分研發。預計在 2019 年 Q3&Q4 開始，HAI 團隊將全力以赴，首先將資源投入到 HAI AutoML 的研發任務當中。HAI AutoML 是恒智鏈（HAI）整個產品的底層與核心，它的作用是搭建人工智能運行環境，與自適應資源調配網絡。HAI AutoML 的研發任務艱巨，預計持續開發時間需要 720 個工作日。因此在 2019Q3&Q4 這一段時間內，HAI 將首先完成 HAI AutoML 中“自定義模型”，這功能確保 HAI AutoML 兼容多種深度學習的模型，同時指出用戶根據需求調整深度學習模型並輸出。這一功能保證 HAI AutoML 對外部人工智能深度學習方式的兼容性，令 HAI 低成本、大規模人工智能網絡成為可能。

2、POAI 共識機制初步研發。POAI 共識機制是外部參與恒智鏈建設的重要基礎。在 POAI 共識機制確保下，可明確每一個恒智鏈中的生態角色的貢獻度。同時，POAI 確保了 Token 的正常流轉。未來，恒智鏈的業務可否正常開展，與 POAI 息息相關。而在 POAI 共識機制當中，“加密區塊”是確保公平性的基石。在 2019 年 Q3&Q4，團隊將投入資源首先開始 POAI 共識機制中“加密區塊”的研發。

3、HPU 研發團隊的早期籌備工作。HPU 的發展是恒智鏈發展當中的核心，HPU 將與 HAI AutoML 之間形成出眾的協同效應，輕鬆應對人工智能算力資源指數級增長的情況。然而芯片研發工作是艱巨且漫長的，而高精尖產品的研發，與團隊當中所具備的人才，亦可稱之為智力資產息息相關。恒智鏈將花費大量資源、時間在搭建 HPU 研發團隊一事上。

#### 未來的工作

1、HAI AutoML。2020 年 10 月完成“可遷移學習”的研發。2021 年完成“神經網絡搜索”的研發。

2、POAI 共識機制。2020 年 6 月完成“資源證明”、“交易證明”的研發。2021 年完成“時空證明”、“結果證明”的研發。

3、算力集群。2021 年開始部署算力集群，試運行。2022 年 Q1&Q2 展開“超級節點”投票，2022 年 Q3&Q4“超級節點”正式上線。

4、HPU 研發。2021 年組建完備的 HPU 研發團隊。2024 年 HPU 研發初代版本。



Hangkei A.I Technology Ltd. 2018

