

近期英偉達 CEO 黃仁勳在最新深度訪談中明確提出，Token 已成為全新商品，AI 電腦是生產 Token 的工廠，未來 AI 計算占 GDP 比重將實現百倍級躍升。2026 年以來，全球 AI 產業邁入規模化商用新階段，Token 已從大模型的基礎技術單元，升級為可計量、可定價、可交易的商業價值載體，Token 經濟迎來爆發拐點。2026 年 3 月我國日均 Token 調用量突破 140 萬億，兩年增長超千倍；全球 Token 周調用量年增速超 800%。Token 經濟解決了 AI 產業商業化模糊的長期痛點，推動大模型廠商從“賣能力”的定制化模式轉向“賣用量”的標準化變現模式，重構了 AI 全產業鏈商業邏輯。當前英偉達、阿里等全球科技巨頭已完成核心戰略卡位，在全球 Token 需求持續爆發的背景下，投資者應該關注國內市場哪些投資機會？

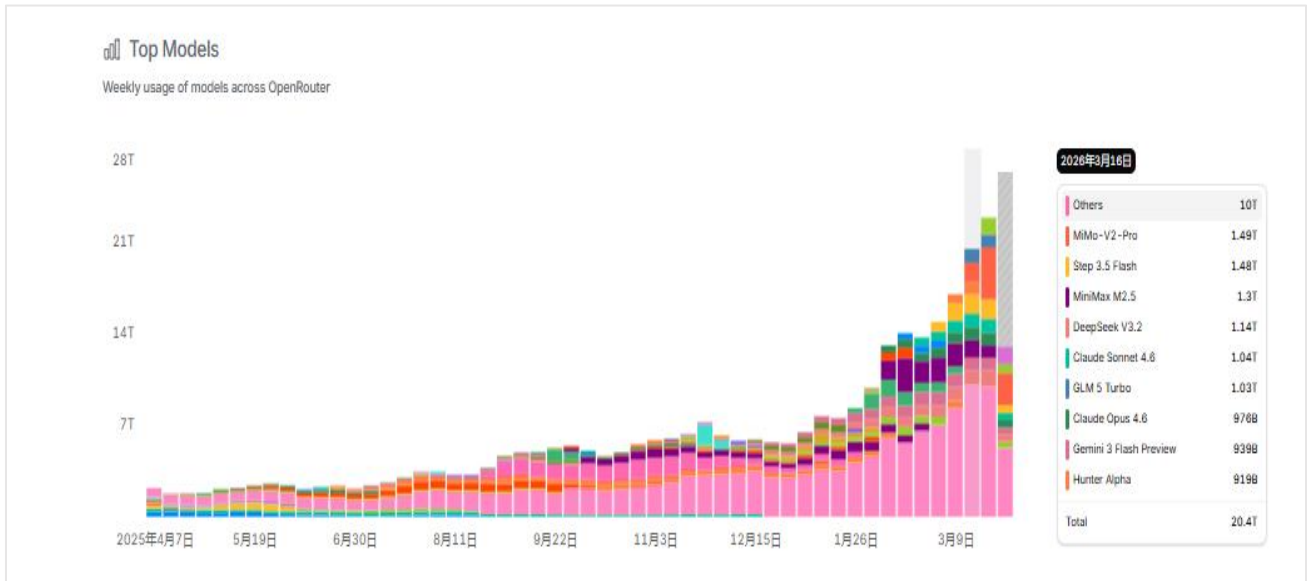
Token 經濟崛起：AI 產業新週期，投資者如何把握時代紅利？

2026 年以來，隨著全球 AI 產業從技術研發邁入規模化商用階段，Token（官方中文定名“詞元”）已從 AI 大模型處理資訊的最小技術單元，升級為可計量、可定價、可交易的全新商業價值載體，Token 經濟正式迎來爆發拐點。國家數據局數據顯示，截至 2026 年 3 月，我國日均 Token 調用量已突破 140 萬億，較 2024 年初實現兩年超千倍增長；全球 Token 周調用量亦從 2025 年 3 月的 2.03 萬億飆升至 2026 年 3 月的 20.4 萬億，一年增長超 8 倍。

Token 經濟的核心價值，在於徹底破解了 AI 產業長期面臨的商業化模糊痛點，推動大模型企業從“賣能力”的定制化服務模式，轉向“賣用量”的標準化、規模化變現模式，重構了 AI 產業全鏈條的商業邏輯。全球科技巨頭已率先完成戰略卡位：英偉達 CEO 黃仁勳明確提出“Token 是全新商品，AI 電腦已成為生產 Token 的工廠”，斷言 AI 計算占 GDP 比重將翻百倍；阿里巴巴-SW(9988.HK)正式成立 ATH 事業群，以“創造 Token、輸送 Token、應用 Token”為核心目標，完成全產業鏈佈局。在全球 AI 行業 Tokens 需求快速提升的同時，會給國內市場的投資者帶來怎樣的投資機會？



圖一：全球主要模型 Token 調用量增長情況



資料來源：OpenRouter 平臺數據

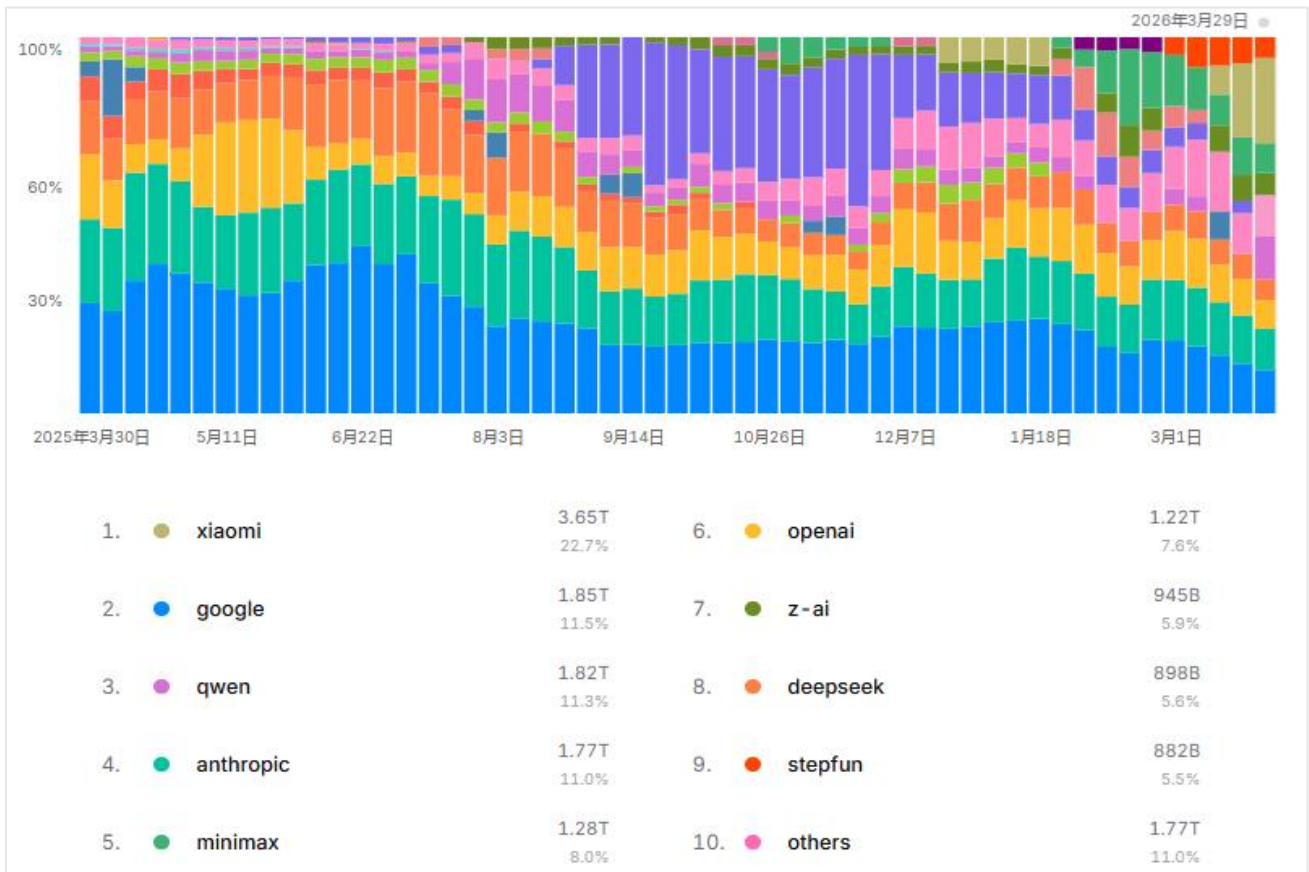
➤ AI 驅動 TOKEN 經濟興起，未來增長潛力顯著

2026 年 3 月 25 日，全國科學技術名詞審定委員會正式發佈公告，將人工智慧領域 Token 的中文名定為“詞元”，明確其是“智能設備中資訊存儲、處理和交換的具有一定語義的基本符號單元，是大模型處理和交換資訊的最小單位” 微博。國家數據局進一步明確，Token 具備可計量、可定價、可交易三大核心特徵，是智能時代的價值錨點，也是連接技術供給與商業需求的核心結算單位。從產業視角看，Token 經濟是以 Token 為價值尺規與流通媒介，圍繞其“生產-調度-消耗-定價-交易” 全生命週期形成的新型 AI 產業經濟形態，其底層邏輯可概括為“Token=算力=能源=價值”，本質是將 AI 智能能力轉化為可標準化交易的大宗商品。

Token 經濟的崛起，核心驅動力來自 AI 推理需求的指數級爆發，這一增長動能集中體現在調用規模、應用場景、商業化模式三大核心維度。在調用規模上，國內 AI 日均 Token 調用量從 2024 年初的 1000 億，躍升至 2025 年底的 100 萬億，2026 年 3 月進一步突破 140 萬億，兩年實現超 1400 倍的增長；全球市場同樣維持高景氣度，OpenRouter 平臺數據顯示，2026 年 3 月中旬全球模型周調用量已達 20.4 萬億 Token，環比增速超 20%，其中中國模型周調用量連續多周超越美國，實現歷史性趕超。



圖二：全球主要模型 Token 周調用量對比



資料來源：OpenRouter 平臺數據

在應用場景上，Token 的使用邊界實現全維度滲透，AI Agent 規模化落地、企業級智能體 7×24 小時不間斷運行、多模態內容生成集中爆發，推動單用戶 Token 消耗量較傳統對話場景提升 10-50 倍；除文本核心場景外，視覺、音頻、視頻等多模態 Token 需求快速擴容，1 張 1024×1024 高清圖片約消耗 85 個視覺 Token，1 分鐘高清視頻便可消耗數千至數萬個 Token，持續打開行業增長天花板。在商業化層面，Token 經濟的商業模式已完成全面驗證，按 Token 計費已成為全球 AI 服務的主流商業模式，徹底解決了 AI 能力難以量化的行業痛點，助力大模型企業實現規模化盈利，國內部分模型企業甚至出現“20 天收入超越 2025 年全年”的業績爆發，充分印證了 Token 經濟的商業可行性與長期增長潛力。



圖三：AI agent 不同應用等級的特徵

等級	tokens 特徵	應用特徵	典型場景
萌芽級	以輸入 tokens 為主，輸出 tokens 多為簡單文本（如短回答、單輪對話）	僅在非核心場景零星試用，未接入企業系統，無標準化流程	中小企業市場部用 AI 寫朋友圈文案 HR 用 AI 簡歷篩選（單輪提問）
普及級	輸入 tokens 來源開始多樣化（如對接部分業務系統 API），輸出 tokens 涉及結構化內容（如報表摘要、多輪對話）	1-3 個非核心業務場景實現規模化應用，但未覆蓋核心業務流程，tokens 消耗集中在單一部門	電商企業智能客服 日均處理 10 萬次諮詢，製造業用智能體分析設備日誌
融合級	輸入 tokens 全面對接企業核心系統（如 CRM、ERP、生產系統），輸出 tokens 直接影響業務決策多模態 token 占比提升	融入 2-5 個核心業務場景（如金融風控、供應鏈調度、精準行銷），形成標準化的數據輸入—AI 處理—業務執行“閉環”，tokens 消耗與業務指標強關聯	銀行數據分析智能體即時分析用戶交易數據（日均 100 萬條，每條轉化為 100tokens），年 tokens 量約 3.6 億
驅動級	tokens 來源覆蓋企業全鏈路，輸出 tokens 支撐端到端業務自動化，多模態 tokens 占比超 50%	AI 成為業務運營的核心引擎，80%以上業務場景依賴 AI 決策，tokens 消耗與企業營收、成本等核心指標直接掛鉤，有專門的 AI 平臺管理 tokens 分配與模型優化	大型互聯網平臺數據分析智能體分析用戶行為數據（日 1 億次交互，每天交互 100 tokens），年 tokens 量約 36 億
原生級	tokens 成為企業“業務貨幣”——不僅是 AI 處理的單位，更是業務流程的核心載體（如用 tokens 衡量用戶需求、產品功能），輸出 tokens 可直接生成新業務（如 AI 自動設計產品、生成服務方案）	企業商業模式基於 AI 原生設計，tokens 消耗效率（如“每 100 萬 tokens 的營收”）成為核心競爭力，AI 模型迭代與業務創新深度綁定	大型 AI 原生企業（如智能駕駛公司、生成式 AI 平臺），日均處理億+多模態交互，年 tokens 量超 360 億

資料來源：英偉達 GTC 大會，Wind

Token 的價值內涵並非一成不變，其先後經歷了三個核心演進階段，最終完成了從底層技術單元到數字經濟核心價值載體的本質躍遷。在最初的技术單元階段，Token 是大模型完成文本分詞、語義理解、推理生成的最小處理單位，是 AI 技術體系穩定運行的底層基礎；隨著行業商業化進程加快，Token 進入計費標準階段，成為 AI 服務商業化的核心計量單元，按 Token 消耗量定價的“用多少付多少”標準化模式，類比水電煤的公用事業計費邏輯，大幅降低了 AI 服務的使用門檻，推動行業實現規模化普



及；而在 AI 推理需求全面爆發的當下，Token 已進階至價值載體階段，逐步成為數字經濟的核心價值流轉工具，成功打通“電力-算力-智能-價值”的全鏈路轉化，成長為 AI 時代可全球流通、可分層定價的全新大宗商品。

* TOKEN 經濟如何重塑 AI 產業商業邏輯？

Token 經濟的崛起，並非簡單的計費模式創新，而是對 AI 產業全鏈條商業邏輯的底層重構，這種重構核心體現在商業模式迭代、產業競爭邏輯重塑、全鏈路價值分配體系重建三大維度。在 Token 經濟成熟之前，AI 大模型企業的商业化主要依賴定制化解決方案、私有化部署、固定額度訂閱制三類模式，本質是“賣能力”，普遍面臨技術能力難以量化、服務定價無統一標準、規模化複製難度高、盈利天花板明顯的核心痛點。而 Token 經濟推動 AI 產業商業模式實現了根本性轉變，AI 企業的核心收入來源從一次性的能力交付，轉向持續性的 Token 用量消耗，這一模式不僅以 Token 為統一計量標準消除了不同場景、不同客戶的服務定價差異，實現了 AI 服務的工業化、規模化輸出，更將專案制一次性收入轉為用戶持續使用帶來的流水型收入，形成穩定的現金流閉環，大幅降低了企業經營波動。同時，該模式還將 AI 服務的覆蓋邊界從服務頭部大客戶的小眾模式，拓展至覆蓋中小微企業、個人開發者的普惠模式，全面打開了市場天花板，更讓企業的算力投入、電力成本與 Token 產出直接對應，收入規模與 Token 消耗量直接綁定，最終形成了“投入-產出-盈利”的清晰商業閉環。

Token 經濟同時徹底改變了 AI 產業的競爭邏輯，行業競爭的核心指標從比拼模型參數規模、榜單跑分，轉向了優化單 Token 生產成本、提升每瓦電力 Token 產出率，誰能以更低的成本、更高的能效生產 Token，誰就掌握了行業定價權與核心競爭力。這一轉變催生了兩大清晰的產業趨勢，其一，能源稟賦成為 AI 企業的核心競爭壁壘，電力成本占 Token 生產成本的 70%，電價直接決定了 Token 的底價，中國西部新能源電價低至 0.3 元/度，僅為歐美的 1/3 至 1/5，使得國產 Token 生產成本僅為海外價格的 1/6 至 1/10，形成了碾壓級成本優勢，也推動“Token 出海”成為中國 AI 產業的新增長曲線；其二，效率優化成為行業技術研發的核心方向，行業研發重心從“做大模型”轉向“做高效模型”，MoE 稀疏架構、KV Cache 緩存技術、液冷散熱、國產晶片能效優化等技術的迭代，核心目標均是提升 Token 生產效率，正如英偉達的 AI 系統在“每瓦 Token 數”指標上領先競爭對手一個數量級，直接決定了其在 Token 生產環節的核心話語權。

除此之外，Token 經濟還為 AI 產業構建了一套全新的產業價值閉環，徹底重建了全鏈路的價值分配體系。在這套閉環中，電力是核心原料，AI 晶片是生產發動機，智算中心是 Token 工廠，大模型是 Token 煉化環節，雲服務是流通管道，AI 應用是最終消費場景，Token 成為全產業鏈價值分配的核心尺規，每個環節的價值貢獻，都可通過其對 Token 生產、流通、增值的貢獻來量化，打破了此前 AI 產業“算力廠商賺走大部分利潤，應用廠商難以盈利”的失衡格局。與此同時，Token 經濟還推動 AI 產業從“技術驅動”轉向“需求驅動”，模型研發、算力佈局不再盲目追求參數規模，而是圍繞下游場景的 Token 需求進行精準匹配，實現了“需求-生產-供給”的正向迴圈，推動 AI 產業真正從技術探索階段，邁入



工業化、產業化的成熟階段。

➤ AI TOKEN：算力與模型巨頭的“新石油”戰爭

AI Token 已從大模型文本處理的最小語義與計算單元，升級為 AI 時代兼具生產資料、計價尺規、商業模式載體三重核心屬性的“數字硬通貨”，2026 年行業正式完成從“訓練驅動”向“推理驅動”的歷史性轉折，推理算力占 AI 總算力投入的比例已升至 70%，全球競爭的核心已從模型參數規模競賽，全面轉向 Token 全產業鏈的生產力、定價權、生態控制權的系統性戰爭，Token 的生產效率、成本控制與場景滲透能力，直接決定企業的核心競爭力與行業話語權。

底層算力層的競爭，是 AI Token 經濟的根基之戰，核心圍繞單 Token 生產成本、能效比、吞吐量三大核心指標展開，本質是爭奪 Token 生產的規則制定權與硬體壁壘。英偉達是這一賽道的絕對主導者，其在 2026 年 GTC 大會上正式提出“Token 工廠經濟學”，首次將數據中心定義為“Token 工廠”，並建立了從免費層到 150 美元/百萬 Token 的五級 Token 價值分層體系，重新定義了行業的價值評估標準。硬體層面，英偉達推出 Blackwell 與 Vera Rubin 兩代核心架構，通過整合 200 億美元收購的 Groq 公司 LPU 語言處理單元，形成“GPU 負責高吞吐、LPU 負責低延遲”的混合推理架構，宣稱將單位 Token 成本降至上一代 Hopper 平臺的 1/10，每瓦特推理性能提升 10 倍，同時依託 CUDA 生態壟斷全球 90% 以上的高端 AI 訓練與推理算力，牢牢掌控 Token 生產的核心硬體標準，競爭對手上，AMD、英特爾、華為昇騰等廠商分別搶佔中低端市場與國產替代賽道，各大雲廠商也加速落地自研晶片，中國更憑藉綠電產業鏈優勢，從能源端構建了 Token 生產的系統性成本壁壘。

圖四：AI 工廠 Token 成為標準化產出



資料來源：GTC 大會

模型層與定價權是產業鏈的核心利潤之戰，行業已形成清晰的分層競爭格局。美系頭部廠商牢牢掌控



高端 Token 市場的定價話語權，OpenAI 通過 GPT-4o、GPT-5 的模型能力分層，建立了行業通用的 Token 價格錨點，其高端模型 Token 定價最高達 150 美元/百萬 Token，推動行業形成“為效果付費”的溢價邏輯；Anthropic 的 Claude 3 系列主打長上下文窗口優勢，在專業長文本場景形成差異化定價能力；穀歌則通過 Turbo Quant 無損壓縮技術，將 KV Cache 壓縮至 3-3.5bit，實現 6 倍以上記憶體壓縮與 8 倍注意力計算加速，在多模態場景實現 Token 成本的極致優化，與 OpenAI 形成正面抗衡。

中國廠商則實現了集群式崛起與定價體系的重構，2026 年 3 月全球 AI 模型 API 聚合平臺 OpenRouter 數據顯示，中國 AI 大模型周調用量達 7.36 萬億 Token，連續三周超越美國，全球前五名 AI 大模型中中國佔據四席，徹底打破了美國廠商的長期壟斷。國產模型通過 MoE 混合專家架構、FP8 混合精度量化、KV 緩存優化等工程創新，實現單 Token 成本的斷崖式下降，MiniMax M2.5 標準測試的 Token 消耗僅為 Claude Opus 的 36%，運行成本僅為海外頭部模型的 1/40，主流國產模型 API 輸入價格已全面進入 0.5 元/百萬 Token 時代，僅為海外同性能模型的 1/10-1/20。值得注意的是，行業競爭已從單純的價格戰轉向價值分層，頭部國產廠商如智譜 AI 憑藉 GLM-5 模型在編程等專業場景的性能突破，2026 年一季度兩次上調 Token 價格，累計漲幅達 83%，而 API 調用量不降反升，印證了“模型智能上界決定定價權，Token 消耗規模決定價值體量”的行業核心邏輯，廠商紛紛通過漲價過濾非剛性測試需求，聚焦高價值企業客戶，完成從“流量補貼”向“價值篩選”的商業化轉型。

➤ 投資者應該關注哪些機會？

AI 時代，Token 的崛起推動計算的本質發生了根本性躍遷——從過去由人類預先錄入、電腦負責檢索匹配的“存儲系統”，徹底升級為具備上下文感知與自主生成能力的“生成系統”，電腦在實體經濟中的角色也隨之完成顛覆性重構。英偉達創始人黃仁勳對此有精準的產業定義：他將過去以檔存儲為核心功能的傳統電腦，比作無法直接創造高額收益的“倉庫”；當下的 AI 電腦，已轉型為與企業收入創造直接掛鉤的“Token 工廠”，AI 代工廠生產的 Token，已成為可精細化分層、標準化定價的全新數字商品。黃仁勳進一步指出，Token 對不同受眾具備差異化的核心價值，目前已形成類似 iPhone 的免費、中端、高端產品分層體系，“市場願意為每百萬 Token 支付 1000 美元的應用場景已近在眼前，這只是時間問題，而非可能性問題”。基於“Token 工廠”的全新商業模式，計算設備徹底完成了從傳統成本中心向利潤中心的跨越，黃仁勳也對這一產業趨勢給出了篤定的宏觀推演：“生產力的躍升將推動全球 GDP 加速增長，我完全確信，未來計算在全球 GDP 中占比將達到此前的 100 倍。”在此背景下，Token 消費的爆發式增長，也催生了市場對其背後核心投資機會的深度挖掘。

投資者可優先關注 AI 算力產業鏈主線，這是 Token 經濟最具確定性的“賣鏟子”賽道，也是 AI 產業爆發的底層根基。Token 的本質是算力與能源的轉化產物，算力供給直接決定 Token 產能，需求剛性極強。重點關注國產通用 AI 算力核心標的，直接決定 Token 生產的能效與成本，深度受益於推理算力需求爆發，例如可配置記憶體介面晶片龍頭瀾起科技(6809.HK)、光通信核心標的劍橋科技(6166.HK)與長飛光纖光纜(6869.HK)，以及晶圓製造龍頭中芯國際(0981.HK)、算力基礎設施運營商萬國數據-



SW(9698.HK) · 這類標的直接綁定 Token 產能擴張的剛性需求 · 業績兌現確定性強 · 是穩健配置的核心底倉。

中期成長聚焦 AI 軟體與大模型主線 · 這是 Token 經濟的定價核心與價值分配樞紐 · 馬太效應極為顯著 · 大模型是 Token 的核心“發行方” · 頭部廠商直接掌握行業 Token 定價標準 · 可憑藉規模化生產形成“成本下降-份額提升”的正向飛輪 · 通用大模型領域重點關注 MiniMax-W(0100.HK) · 商湯-W(0020.HK) · 這類廠商已實現按 Token 計費的規模化 API 商業化 · 量價齊升邏輯明確；垂類 AI 賽道可佈局金山軟體(3888.HK) · 金蝶國際 (0268.HK) · 這類企業深耕辦公、企業 ERP、智能體等高價值場景 · 深度綁定企業核心業務流程 · 能鎖定長期穩定的 Token 消耗需求 · 成長彈性充足。

穩健配置優先選擇全產業鏈佈局的互聯網大廠主線 · 這類標的兼具穩定現金流與 AI 成長確定性 · 是港股 AI 配置的壓艙石 · 重點關注騰訊控股(0700.HK) · 阿里巴巴-SW(9988.HK) · 二者分別依託混元大模型+全場景生態 · ATH 事業群 Token 全鏈路佈局 · 完成了 Token “創造-輸送-應用”的完整閉環 · 在 Token 定價、產能供給、場景落地全環節均具備不可替代的壁壘；同時可關注快手-W(1024.HK) · 美團-W(3690.HK)等具備豐富落地場景的平臺型企業 · 其 C 端與 B 端生態可持續釋放規模化的 Token 消耗潛力。

免責聲明：本報內容所提供資料所述或與其相關的任何投資或潛在交易 · 均受限於閣下司法轄區適用的法律及監管規定 · 而閣下須單獨就遵守該等法律及監管規定負責 · 本報內容僅供參考 · 不構成任何投資建議 · 本公司對所提供的財經資訊已力求準確 · 但對其中全部或部分內容的準確性、完整性或有效性 · 不承擔任何責任或提供任何形式保證 · 如有錯失遺漏 · 本公司恕不負責 · 另請注意證券與虛擬資產價格可升可跌 · 尤其虛擬資產的風險極高 · 投資者應對有關產品保持審慎及自行承擔投資風險。

