

Alpha Helix Asset Management 投資備忘錄（七十六）

2025.06.27

從 Google IO 大會觀察 AI 商業化之路

AI Search 全面導入 Google Search，結合傳統搜尋引擎模式：

在 2025 Google I/O 中，Google 正式宣告以 Gemini 2.5 驅動的「AI Mode」為搜尋模式的創新核心，並持續優化 AI Overviews 功能，使搜尋系統可解析用戶提問，並以 AI Chatbot 及傳統推薦引擎同時回答，用戶可以此進行多輪提問，並搭配即時資料與購物導覽。

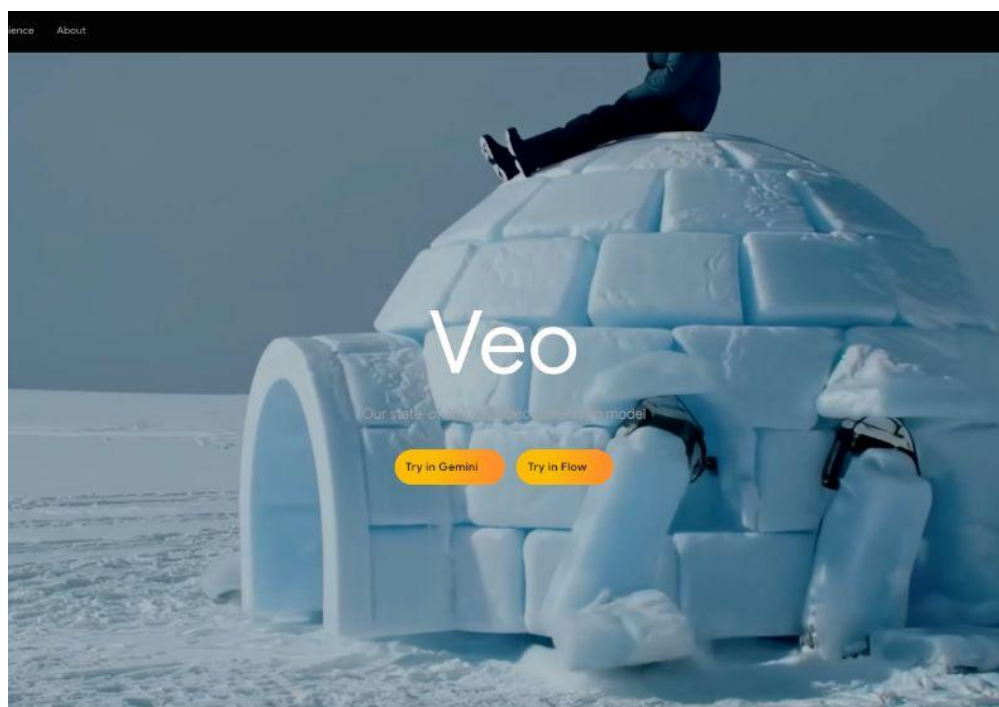
以投資角度而言，AI Search 除可為 Google 新增廣告版位（AI 回答版面），推升營收外，同時也可向市場證明，此前部分市場觀點「AI 將顛覆傳統搜尋模式」實屬過度擔憂，實際情況是 Gen AI 反倒將結合傳統搜尋模式，提升用戶搜尋體驗。

(圖一) AI Overview 可為用戶提供有別於傳統搜尋的綜合性回答

從概念驗證到實際端側部署，Google 將新產品與既有平台生態整合：

本次 Google I/O 大會中，Google 推出多項新產品，除優化 Gemini 2.5 模型的推理能力外，並搭配多款應用產品實際落地。如，Gemini App 正式加入 Agent Mode，可根據 Gmail、雲端硬碟內容完成日曆排程、信件撰寫等複雜任務；Gemini Live 則讓使用者在 Android/iOS 上即時使用鏡頭或畫面分享，實現類似 Project Astra 所描繪的多模態交互體驗。

Google 更將生成式 AI 應用延伸至圖像、音樂與影片創作，推出整合式的「FLOW 影片創作工具」，整合 Imagen 4 (生成圖像)、Veo 3 (生成影片)、Lyria 2 (生成音樂) 等能力，並與 YouTube Shorts、Music AI Sandbox 無縫串接。這標誌著 Google 正將 AI 能力轉譯為可商業化、可規模化的創作模組，擴展至娛樂、內容創作等高互動性產業。

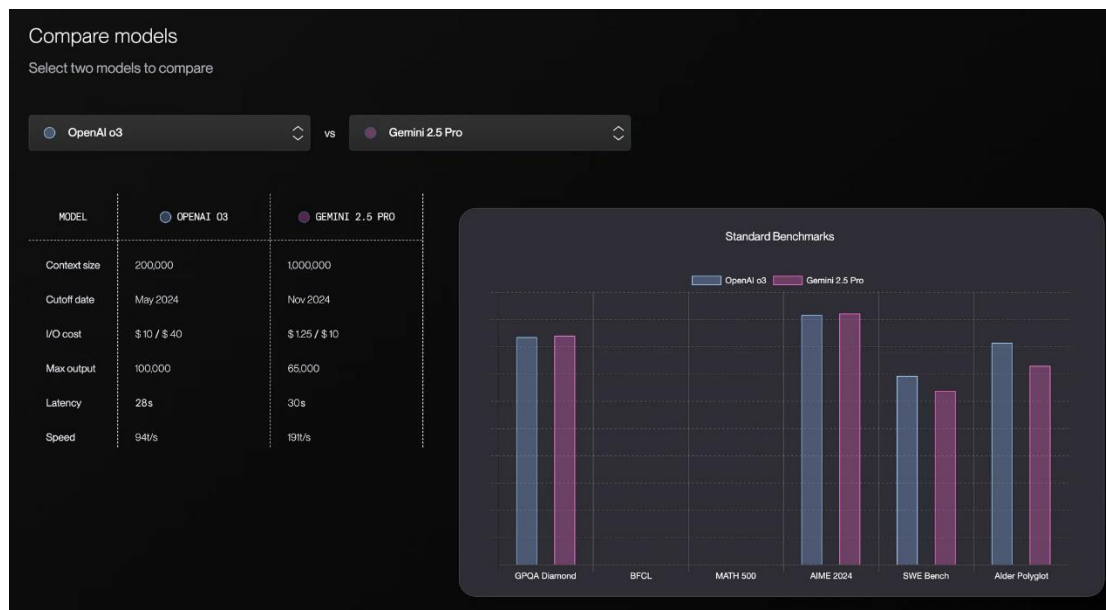


(圖二) Veo3 生成影片效果極佳，已有多家廣告公司採用並商業化投放

自研 TPU 架構，大幅壓低推理成本、強化 AI 推理優勢：

Google 在 AI 領域的核心優勢除既有生態系外，自研 TPU 亦為關鍵優勢，相較主流 GPU 架構，TPU v5 系列在特定矩陣運算與大語言模型多層並行推理上，擁有更高 FLOPS/W 效能與定制化張量核心設計，使得 Gemini 模型在處理成本與延遲上達到可用於日常應用的門檻。

透過垂直整合自研晶片，Google 不僅擁有更好的單位成本控制權，更能依據使用情境進行能耗、精度與延遲的動態調整，形成明確的護城河。以目前最前沿的 Gemini 2.5 Pro 模型為例，與 OpenAI O3 模型跑分基本一致，然而推理成本僅為後者的 12%~25%，在開發者群體中廣受好評。



(圖三) Gemini 相較 GPT 在同樣性能下，成本優勢明顯

結語：2025 Google I/O 為 AI 商業化展示出可能路徑：

2025 年 Google I/O 展現的不只是模型升級與應用展示，更是一場完整的商業化部署宣言。從 AI Search 結合傳統搜尋引擎模式，到 Gemini Agent 深度整合 Workspace 與 Android 裝置，再到 FLOW、Veo、Lyria 等多模態創作工具，Google 已將 Gen AI 實際轉換為具備商業價值的產品體系。

更關鍵的是，憑藉自研 TPU 架構，Google 在成本控制與雲端推理效率上取得明顯優勢，構築出難以複製的垂直整合護城河。對投資人而言，這不僅代表 Google 成功回應「AI 顛覆搜尋」的質疑，反而將 Gen AI 融入核心平台與終端應用中，開啟一條可預見、可驗證的營收增長路徑，為 Gen AI 的商業化展現出可能的路徑。