從 DeepSeek 權重開源，看 LLM 生態系發展

2023年5月，Google 研究員洩漏了一份標題為《We Have No Moat And Neither Does OpenAI》（我們與 OpenAI 皆無護城河）的內部文件，該文件指出：「就在我們內鬥的時候，第三方勢力已默默佔了上風。我指的當然是開源社群。老實說，他們已經遙遙領先了。」並做出以下結論：「開源模型的速度更快、彈性更高、隱私性更好，並且在同等條件下效能更強大。」

文／旗標科技

LlaMA 洩漏權重，促進開源 LLM 快速發展

Meta 的 LLaMA 模型於 2023 年 2 月 24 日釋出後，僅過一週便讓模型的權重洩漏至 4chan 平台。儘管當時 LLaMA 的授權條款禁止商業用途，但開發社群仍因得以一窺模型權重而振奮不已。一夜之間，任何人都能在 LLaMA 的基礎上嘗試打造表現堪比（甚至超越） GPT-3 的模型。

就在權重外流約莫一週後，史丹佛大學發佈了 Alpaca。此模型即屬於 LLaMA 的變體之一，是研究人員花幾百塊美金微調 LLaMA 得到的成果。Alpaca 的出現不僅為 LLM 民主化樹立里程碑，更帶動了 LLM 開源社群的快速發展。

後來許多開源模型便是以此為基礎或受其啟發而建立的，例如在其之後發表的 Vicuna、GPT4ALL 和 Koala 等（由LLaMA和Llama 2微調產生的變體，可在 Hugging Face 的模型目錄中找到，見 http://mng.bz/0l5l）。

2023 年 7 月，Meta 決定將 Llama 2 開源，並附上研究與商業用途的授權條款，並表示：「發佈一週以來，我們已收到超過 150,000 次的下載請求，反響非常熱烈，我們引頸期待接下來的發展。」

DeepSeek-R1 權重開源，打破美國技術壟斷

2024 年 12 月 26 日，中國的 DeepSeek 公司正式公布 DeepSeek-V3 大型語言模型（LLM），宣稱功能追上 OpenAI 的 GPT-4o 模型。之後又在不到一個月的 2025 年 1 月 20 日公布了 DeepSeek-R1 推理模型，對標當時 OpenAI 最先進的 o1 推理模型，其優異的性能打破了美國 LLMs 的技術壟斷。DeepSeek 訓練模型所需硬體資源顯著低於 OpenAI 等巨頭企業的模型，且連接其 API 的收費標準也顯著較低，這些都給予資金沒那麼充沛的企業帶來另一種選擇。

而更讓全球科技界振奮的是，DeepSeek 公司將 R1 的模型權重採用 MIT 寬鬆授權方式開源，這大幅降低了 LLM 技術的進入門檻，讓世界各國企業與開源社群都可載入此權重，並用自己的資料進行模型微調，發展出基於 R1 權重的 LLMs。

LLM 的權重包含了預訓練過程中學習到的統計模式與知識，控制著模型如何從輸入中擷取資訊、如何聯想不同概念，以及如何回應使用者問題，被視為模型的「記憶」或「智慧來源」。因此，不論是 LlaMA 的權重洩漏、Llama 2 的開源，或是 DeepSeek-R1 的權重開源，對開源社群都是莫大的鼓舞。

而在 2025 年 1 月 27 日，DeepSeek 又推出 Janus-Pro 文生圖模型，同樣採用 MIT 授權。其競爭產品是 OpenAI 的 DALL-E 3 模型以及 Stability AI 的 Stable Diffusion。相信必能迎來開源社群在生成式 AI 的百花齊放時代。

開源與封閉 LLM 的優缺點

生成式 AI 的開源熱潮的崛起其實有前例可循，Stable Diffusion 生圖模型已於 2022 年 8 月 22 日開源，並提供商用與非商用的使用授權，現在已形成了豐富的生態圈。而開源 LLMs 在 2023 年也開始迅猛發展，因此企業也是時候來評估開源與封閉模型的優缺點了。

開源 LLMs 具有高度的透明度與普及性，因此能接納更多想法、加速創新，並最大限度地減少偏見並將主導權從少數資源豐富的大型科技公司手中釋放出來。

話雖如此，開源專案也存在不少缺點與挑戰，例如：缺乏中央集中管理、品質管制，以及長期維護的保障，同時還有智慧財產權方面的顧慮。

比較開源與封閉 LLMs 的優缺點

比較項目	開源 LLMs	封閉 LLMs
透明度與普及性	容納多元想法與創新、最小化偏見、並降低入門門檻。 將主導權從大型科技公司手中釋出。	資訊透明度與模型普及性可能較低。 主導權集中在少數大型科技公司手上。
資料隱私	若部署在安全環境，則因為資訊不會傳給開發公司，故理論上資料隱私性較高。	敏感資料可能遭搜集、儲存或擅自使用。
管理與品質	缺乏中央管理系統，有品質及長期維護上的疑慮。	有品質保證，並經過安全性測試。
對用戶的技術要求	較高	較低
安全性漏洞	透明度高有助發現漏洞，但需要開源社群進行修正。	內部會進行紅隊演練，並針對惡意或有害回應建立防禦。
濫用風險	較高	較低

對於 LLM 與生成式 AI，許多人往往一開始就一頭栽進技術細節，其實站在制高點俯視，才能洞察 LLM 的發展趨勢。在這場生成式 AI 熱潮與搶購 NVIDIA（輝達） GPU 的 AI 軍備競賽中，想要深入淺出了解更多 LLM 複雜概念，理性認識這場技術革命的全貌，推薦進一步閱讀**《生成式 AI 入門 - 揭開 LLM 潘朵拉的秘密》**一書。