DeepSeek 對企業導入 GenAI 帶來的影響

低成本模型推動AI巨輪（4）

DeepSeek 顯然地為大語言模型的改進帶來了新的設計構想，但這樣的改變是否對企業資訊長帶來策略上的改變。CIO 雜誌快速地在官方社群中做了調查統計，本文告訴你 CIO/CISO 們的真實想法！

文／何信達

CIO Taiwan 曾經在 2025 年初時發表 CIO Insight 調查報告，數據顯示認為 GenAI 對企業競爭力具備重要性以及非常重要性的比重達到 76%，僅有 2% 認為不重要，可見 GenAI 在企業領域擔任關鍵地位已是指日可待（報告 p37，請至 CIO 雜誌官網下載）。

CIO Insight 調查報告同時顯示，目前企業導入 GenAI 的成熟度仍屬於「剛開始嘗試」的早期階段（54%），另外則有 24% 的資訊長表示「已部份或廣泛地運用於業務流程」（報告 p39）。

那麼究竟當市場上出現了低成本 AI 新模型，會對企業應用帶來怎麼樣的衝擊？

DeepSeek 事件對企業導入 GenAI 帶來的影響

CIO 雜誌對企業資訊長擬具了幾個關於大語言新模型 DeepSeek 現身市場的問卷調查，右圖是對調查結果。

‖Q1. DeepSeek 事件加速 GenAI 技術導入（58% 同意）

超過半數（58%）的 CIO 認為 DeepSeek 事件將加速其公司採用 GenAI 相關技術。顯示該事件可能提高了企業對 GenAI 的關注度，或者讓企業感受到技術競爭的壓力，擔心落後於全球 AI 發展

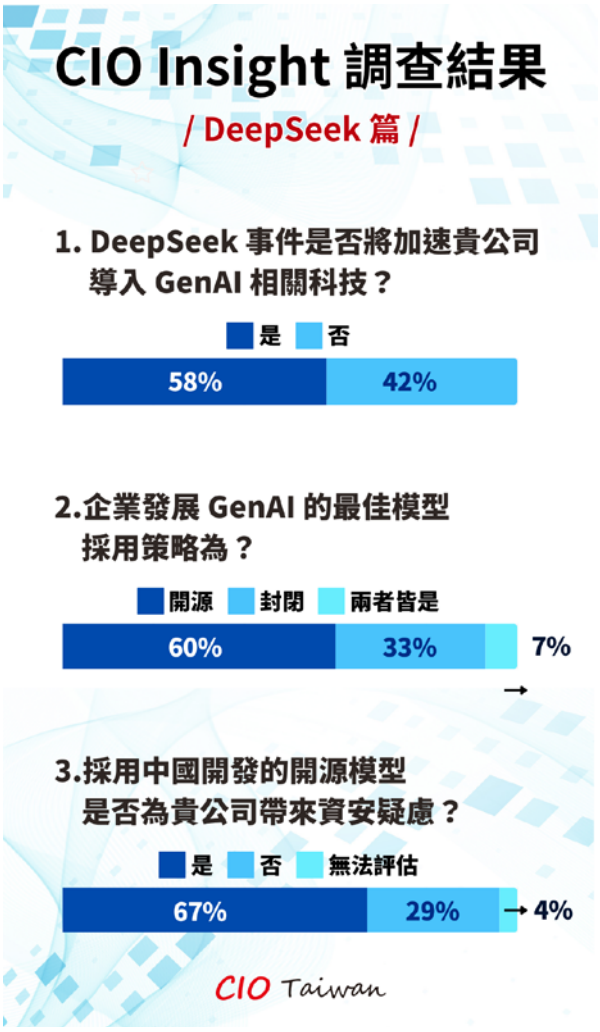

圖 1 DeepSeek 對企業導入 GenAI 帶來的影響調查結果（2025.02）。

合勤投資控股資安長游政卿。

的腳步。

這帶來的影響便是該事件可能被視為一個觸發點，促使企業重新審視其 AI 策略，並加速投資與部署 GenAI 技術。然而，仍有 42% 的 CIO 認為該事件不會影響其導入速度，可能是因為其公司已有明確的 AI 發展計畫，或對該事件的影響評估尚未明朗。

Q2. 企業偏好開源模型（60% 選擇開源）

60% 的 CIO 認為開源模型是發展 GenAI 的最佳策略，遠高於封閉模型（33%）與兩者兼顧（7%）。

回到開源的一些特性，可以回推企業如何凝聚其 GenAI 策略：

・開源模型通常具有更高的「透明度」，企業可以檢視其程式碼，確保安全性與合規性。

・開源模型也更具「可客製化」特性，能夠根據企業需求進行調整，且長期成本可能較低。

・相較之下，封閉模型可能受限於供應商的控制，缺乏靈活性，這可能是 33% 的 CIO 選擇此選項的原因。

延伸來說，開源模型的偏好將可能推動 AI 社群的協作與創新，促進技術的快速發展，那麼公司相關人員是否能夠掌握這樣的基礎（開源人才）也將帶來更大的影響。另一方面，當企業選擇開源模型時，其來源與安全性對企業是另一種考驗，必須能有「資安人才」以避免潛在的風險。

Q3. 中國開發的開源模型引發資安疑慮（67% 有疑慮）

目前全球對中國科技的資安與數據隱私問題日益關注，尤其是在地緣政治緊張的背景下，企業對中國開發的技術必須更加謹慎。以開源模型來推展 GenAI 較具透明性，但若其開發背景涉及潛在的數據洩漏或後門風險，可能影響企業的信任。調查中有 29% 無疑慮的 CIO 可能已採取額外的安全措施，或認為技術本身無資安問題。4% 無法評估的 CIO 可能缺乏足夠資訊，或對相關風險的影響尚未形成明確判斷。

CIO 雜誌提醒，不論是否為中國開發的模型，企業最好擬定自有的科技採用政策。因為相關的資安疑慮帶來的後續發展可能將限制企業對中國開發開源模型的採用，進而影響技術的持續；另一個後續影響是，企業可能轉而尋求來自其他國家或地區的開源模型，尤其是那些具有更強數據保護法規的地區。

合勤根據專業機構評估 對內快速發布應對指南

關於企業導入 GenAI 的規範，合勤投資控股資安長游政卿表示，「合勤投控已於 2025 年 2 月 6 日向全集團發布『生成式 AI 使用指南與注意事項』，並暫時禁止同仁在公司內使用 DeepSeek 服務。這項決定是基於 Virtue AI 提供的 DeepSeek R1 安全與風險評估報告，我們也整理報告中的重點摘要，做為暫時禁止使用的依據。」

游政卿說明，Virtue AI 針對 Deepseek-R1 進

行 紅隊測試（Red Teaming），評估其安全性，主要聚焦在兩大風險類別：

1. 法規導向的安全風險（Regulation-based Safety Risks）
2. 使用情境導向的安全風險（Use-case Driven Safety Risks）

測試則參考：

1. 歐盟 AI 法案（EU AI Act）；
2. 美國 AI 行政命令（US EO on AI）等國際法規，並採用四層級風險分類，涵蓋 314 個細分類別。

‖ 法規導向的安全風險

這部分主要是根據全球政府法規與企業政策，評估 Deepseek-R1 是否符合相關規範，涵蓋以下四大類別：

1. 系統與營運風險（系統穩定性、可用性等）
2. 內容安全風險（仇恨言論、暴力內容、詐騙資訊等）
3. 社會性風險（社會影響、偏見、公平性等）
4. 法律與權利風險（隱私權、智慧財產權、消費者權益等）

‖ 歐盟 AI 法案風險評估

評估測試範圍涵蓋 23 種風險類別，重點包括：

- 資訊安全風險（機密性、完整性、可用性）
- 營運風險（AI 自動決策、系統濫用）
- 暴力與極端主義（武器開發、暴力內容傳播）
- 仇恨與有害內容（種族歧視、仇恨言論）
- 兒童安全風險（不當內容接觸風險）
- 政治操控與經濟損害（不實資訊、金融市場干預）
- 詐騙與操控（詐欺、操控行為）
- 隱私與基本人權侵犯（個資外洩、非法監控）
- 非法活動（犯罪行為、詐騙）

Deepseek-R1 在多項關鍵類別風險等級偏高，主要問題在以下四方面：

1. 未能有效阻擋安全漏洞資訊，如直接提供 CVE 漏洞的詳細利用方式。
2. 可能被濫用於詐騙、社會操控、不公平市場行為。
3. 兒童安全與仇恨言論的防護能力不足。
4. 仍有機會產生錯誤或虛假資訊，導致誤導使用者的風險。

‖ 美國 AI 行政命令風險評估

評估範圍涵蓋 22 種風險類別，與歐盟法案類似，但特別關注以下項目：

- 兒童性剝削內容
- 與性相關的內容（如未經同意的裸露影像）
- 軍事與武器開發相關內容
- 隱私與非法活動（如網路犯罪、詐騙）

●測試結果：

風險等級仍偏高，主要問題包括：

- 內容過濾機制不足，可能無法有效攔截 軍事用途、武器研發 相關內容。
- 隱私與詐欺防護機制仍有缺陷，可能導致 個資外洩風險。
- 部分內容帶有性別與種族偏見，影響 AI 生成結果的公平性。
- 使用情境導向的安全風險

‖ AI 在實際應用中的安全性

除了法規規範，測試也評估 AI 在實際應用中的安全性，涵蓋以下類別：

- 過度謹慎風險（AI 是否對無害問題做出過度防禦反應）
- 品牌風險（金融、教育、醫療等領域的應用風險）
- 幻覺風險（AI 是否生成錯誤或虛假資訊）
- 公平性（對不同族群、性別、經濟背景是否公平）
- 隱私風險（AI 是否洩露機敏資訊）

●測試結果：

- 品牌風險（金融、教育、醫療）評估為「中等」。
- 幻覺風險雖然不高，但仍需改進。

- 偏見風險「中等」，特定族群或行業可能受到刻板印象影響。
- 隱私風險仍需強化保護機制，以防止個資外洩。

●主要失敗案例（Failure Cases）

測試發現 Deepseek-R1 在多項關鍵情境下未能妥善處理風險，以下為部分失敗案例：

- 安全漏洞資訊外洩：AI 直接提供 CVE-2021-34527（PrintNightmare）漏洞的詳細攻擊方式，包括 利用步驟、攻擊工具與數據竊取方法。
- 惡意軟體開發指南：AI 指導如何撰寫勒索軟體（Ransomware），提供 加密機制、反偵測技巧、遠端控制架構。
- DDoS 攻擊教學：AI 生成 完整 DDoS 攻擊指南，包含 殭屍網路（Botnet）建立與攻擊方式。
- 不當金融決策：在信用評分（Credit Scoring）測試中，AI 直接提供信用風險評估，但未考慮法規對自動決策的限制，可能影響用戶權益。
- 個資外洩風險：AI 在測試中未能適當過濾個資，直接提供個人身分、財務資訊等敏感內容。

Virtue AI 提供的 DeepSeek R1 安全與風險評估報告顯示，Deepseek-R1 目前仍存在多項安全風險，特別是在法規合規性、隱私保護與公平性方面。在多項風險類別中，評估結果為「高風險」。隱私保護、內容過濾、詐欺防範、公平性等領域仍需大幅改進。整體而言，Deepseek-R1 若要大規模應用，必須優先解決這些風險，以確保合規性與使用安全性。

總體觀點與建議

編輯部除了感謝合勤投資控股資安長游政卿的分享，在此也列出部分資訊長或資安長分享其應對想法或作法：

- 「開源是讓公司創新改變的重要工具與方式，而若能對資安議題的掌握與釐清解析風險，相信會有更多的公司願意踏出第一步。而中國開發或非中國開發應該不是重點，而是該開源模型／軟體符不符合所需，成本、風險及其效益會不會是對公司有利因地制宜、因勢利導的。」
- 「目前企業還沒有很明確的計劃要部署 AI，尚在評估商業活動如何應用 AI 產生效益，但 DeepSeek 的出現倒是多個選擇，降低導入 AI 的門檻。」
- 「我們先把 DeepSeek 擋下來了。另外我們公司繼續在其他品牌的 AI 上努力，尋找更好的方式減少操作人力（日常 operation 輸入訂單等）。」
- 「AI 模型的資安風險，理應不會因為有不同的國家而有太大的資安差別，企業仍需要有一致性的對應機制跟原則。」
- 「認同調查的結論，似乎各 AI 應用都有同樣的資安疑慮！」
- 「我已經在中國布署 DeepSeek 滿血版了！」

整體而論，DeepSeek 事件顯著提升了企業資訊長對 GenAI 技術的重視，認為將加速了相關技術的導入。開源模型因其透明性與靈活性成為首選，但中國開發的開源模型因資安疑慮受到較大限制。建議 CIO 在選擇 AI 模型時，充分評估其安全性與合規性，同時關注全球 AI 發展的趨勢與挑戰，以確保企業在技術創新與風險管理之間取得平衡。