

特約內容 | 白皮書

市場
脈動

首席資訊長利用企業級 GenAI 實現規模化擴展指南

最近研究針對 IT 主管如何為下一階段的 AI 做好準備，
提供深入分析。



CIO

贊助單位

NUTANIX

生成式 AI (GenAI) 已從實驗階段邁入正式量產。IT 部門現在必須安全可靠地運用該技術，以支援全企業範圍的執行與可衡量的成果。這個轉變將其提升為策略性營運與架構的優先事項，進而形塑成本可預測性、治理、業務連續性、合規監管風險以及長期競爭力。

根據一項全新的 Foundry Market Pulse 研究，首席資訊長與企業領導者正開始將 GenAI 視為與 ERP、CRM 及其他關鍵任務系統同等重要的企業級能力。幾乎所有資深 IT 主管 (96%) 在調查中都承認 GenAI 對於其組織的商業策略來說很重要。幾乎同樣多的主管贊同這是競爭優勢 (參閱圖 1)。

這種策略性思維很合理，授權首席資訊長以確保其蓬勃發展的 GenAI 投資能夠擴大規模、為企業帶來及時的投資回報，並為韌性與永續成功奠定基礎。因此，IT 主管應將 GenAI 轉化為標準化平台來營運，該平台旨在跨混合環境進行規模化，並被驅動韌性、可預測性與權責歸屬的相同原則治理。

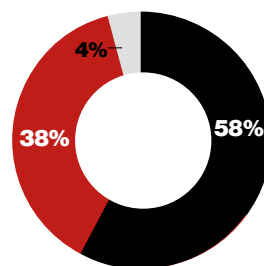
具前瞻思維的成功領導者，將該技術視為與其他企業關鍵應用程式相同 (儘管它極其強大) 並同樣需要企業級的基礎架構、治理機制與專業技能以支援不同環境的部署與監督。它必須作為企業基礎架構來進行工程設計、治理與營運。

圖 1

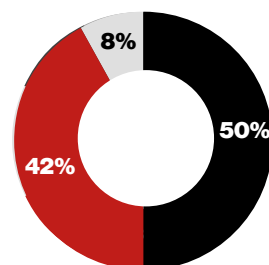
GenAI 現在是企業的首要任務

- 有些同意
- 選擇性同意
- 非常同意

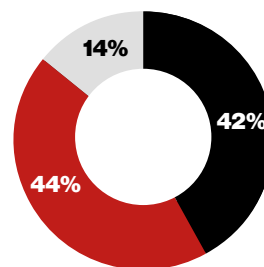
GenAI 對我們未來 12 個月的商業策略至關重要



我們企業將 GenAI 視為一項競爭優勢



GenAI 與 ERP 和 CRM 同等重要



資料來源：FOUNDRY

GenAI 現在是企業的首要任務

GenAI 的採用在關鍵任務的量產環境中正取得進展。例如：三分之二的受訪者表示，其組織在軟體即服務 (SaaS) 應用程式中使用 GenAI，並將此技術嵌入到企業工作流程中。

這些部署利用結構化、非結構化和物件資料的混合內容 — 幾乎貫穿整個企業的完整資料生態系統 (參閱圖 2)。

GenAI 工作流程影響範圍廣泛且逐漸深化的另一個指標是：大多數組織正在或計劃在混合雲環境中運行 GenAI 應用程式，包括公有雲、私有雲、資料中心和邊緣位置。正如首席資訊長正在迅速了解到的，這種規模化擴展會引發無數的整合、資訊安全、主權以及效能挑戰。

對於能夠克服這些問題的 IT 主管而言，回報將是非常值得的。IT 主管預期其 GenAI 投資將帶來一系列正向成果，包括：

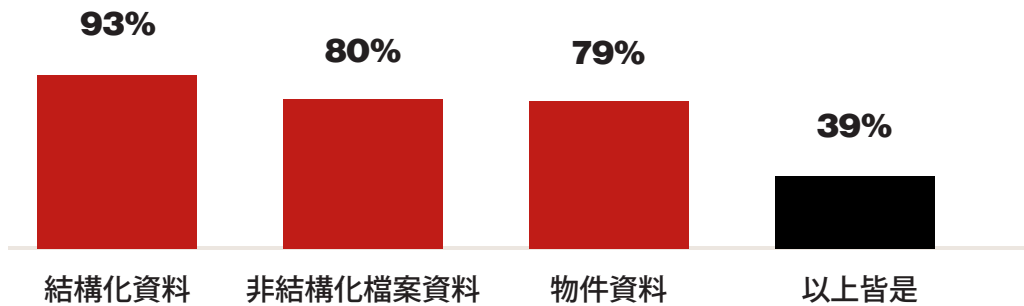
- 生產力與效率的提升
- 客戶體驗的改善
- 營收成長
- 全企業的商業流程轉型
- 創新與新產品開發

然而，早期正式量產的成功不一定代表企業已做好準備，或具備長期的規模化永續能力。

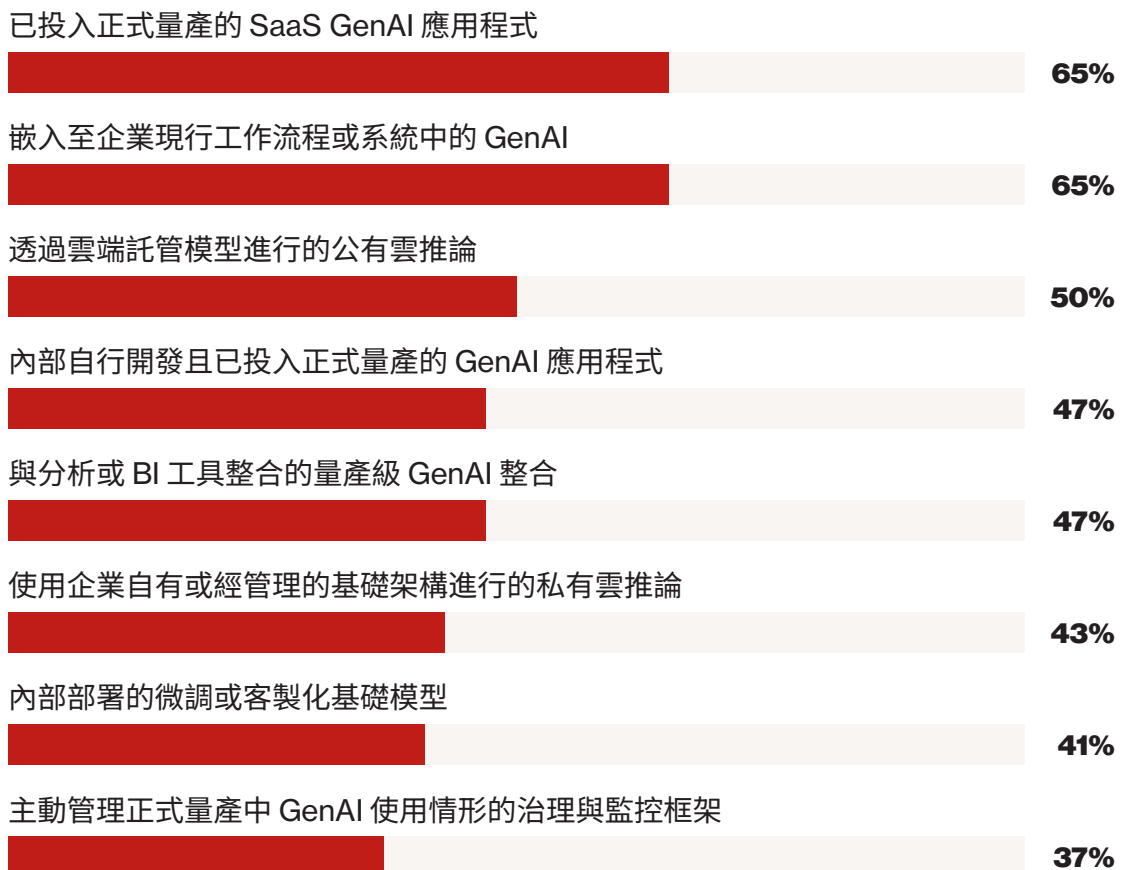
「對於 GenAI 如何影響企業的策略性領域，業界無疑抱持著積極樂觀的態度，」業界資深專家、Nutanix 產品與解決方案行銷資深總監 Harsha Kotikela 表示。「但就像面對任何新技術一樣，許多組織仍處於這趟旅程的早期階段。首席資訊長之間的對話現在正從興奮轉向『我們該如何從所有概念驗證中學習並建立起能讓我們實現規模化擴展的基石？』」

圖 2 | GenAI 的廣泛影響範圍橫跨商業與 IT 環境

GenAI 專案存取的資料類型



GenAI 在生產環境的應用



資料來源：FOUNDRY

挑戰：風險、基礎架構與規模化

當被問到組織在擴展 AI 規模時所面臨的基礎架構挑戰時，受訪者指出了前三名分別是資訊安全與合規性、效能與可靠性，以及資料管理與整合。

此外，IT 主管也指出實現其 GenAI 目標所面臨的重大挑戰，包括：

- **資料風險：**大多數受訪者 (94%) 表示有適度的信心管理與 GenAI 相關的資料隱私和合規性風險。他們可能低估升級平台、政策與工具以支援 AI 資料治理需求所需的投入。例如：用於 AI 的敏感資料通常必須保留在國家或組織的邊界之內。GenAI 逐漸需要主權執行——確保資料、模型與處理程序皆維持在特定的管轄權或企業掌控之中。然而，許多平台缺乏滿足這些需求所需的透明度與強制執行能力。組織也可能缺乏整合治理架構來保護 AI 管線的安全，並在資料、模型和運算層之間執行政策。
- **技能準備：**對 AI 關鍵技能 (包括 AI 治理與自主代理工作流設計) 的需求遠遠超過供應，導致企業領導者難以應對基於代理之系統在整合與協調編排上的需求。人才短缺同樣會帶來規模化與韌性方面的風險。隨著技能差距不斷擴大，首席資訊長將需要建立具備可重複性與自動化的 AI 基礎架構及治理機制，從而減少對稀缺之專業 AI 人才的依賴。
- **不可預測的成本：**長期以來，IT 主管一直難以改善他們預測和管理隨用隨付的雲端成本的方式。AI 為 IT 成本管理引入更多的不確定性。高階 AI 系統會在 Token (AI 模型在訓練與推論期間處理的資料單位)、延遲與能源方面造成顯著的額外消耗。隨著複雜度提升且能見度降低，碎片化的解決方案可能會產生冗餘與浪費。不可預測的成本將促使首席資訊長去提升跨混合環境下，AI 工具與平台的能見度、優化能力以及成本追蹤能力。

圖 3

IT 主管尋求降低的關鍵 GenAI 隱私風險

跨分散位置的安全管理控制



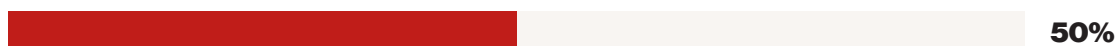
大型語言模型使用的存取控制與監控



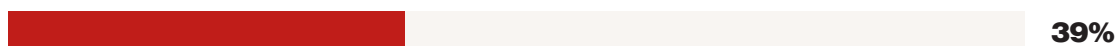
人才與技能發展



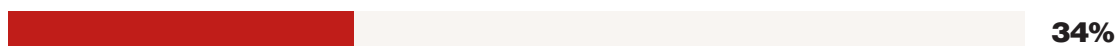
治理架構



公有雲資料外洩



機密機房營運支援



*您認為哪些領域對於降低 GenAI 隱私風險最為關鍵?(請排出前 3 名)

資料來源: FOUNDRY

隨著 GenAI 擴展至多個團隊、系統與環境，結構性的挑戰只會持續增加。

若放任不管的話，這種複雜度將直接轉化為商業風險，包括合規監管風險；不可預測的營運成本；服務不穩定；以及在根本層面上，高階主管對 AI 驅動成果的信任度遭到削弱。

「AI 的變革步伐極快，如果你跟不上，基本上就會失去競爭優勢，」Kotikela 表示。「資料與基礎架構對企業具有策略性意義，如果在 AI 應用上走錯了，其負面衝擊將會被放大。」資料與基礎架構對企業具有策略性意

IT 主管將需要在早期正式量產成功的樂觀與可規模化的企業準備之間取得平衡。實現 GenAI 的規模化 — 並獲取預期效益 — 就需要重新聚焦在對長期成功很關鍵的基礎要素。

從單點解決方案到平台化

碎片化 AI 解決方案加劇了高階管理階層在成本、治理和營運韌性方面面臨的風險。為了降低風險並提供一條可靠的成功路徑，首席資訊長應將 AI 與其他任何企業級應用程式一同管理，確保 AI 系統經過策略性部署，並在具備韌性、Day 2 營運與資訊安全的狀態下運作。

降低風險並提升韌性最好的方法是為 AI 建立一套企業營運模式，以提供一致的治理、成本可預測性與韌性。這種平台化方法並非整合工具，而是為了簡化 AI 營運，好讓現有的 IT 團隊能利用一致的工作流程與熟悉的運作模式來部署、治理並營運 GenAI。

若缺乏平台，GenAI 恐將進一步碎片化為孤島式單點解決方案，進而加劇成本波動、削弱權責歸屬，並限制其交付可預測商業成果的能力。

關鍵考量因素：

- **跨混合環境的支援：**AI 工作負載正將基礎架構推向極限。效能、可靠性與韌性，是將 AI 轉化為持久商業價值的關鍵所在。然而，由於驅動 AI 的企業資料高度分散於多個位置——包括本地端資料中心、公有雲以及邊緣端——首席資訊長將需要部署一種混合架構，使 AI 工作負載能夠在資料、延遲與主權需求所決定的任何地方執行。由於驅動 AI 的企業資料高度分散於多個位置一包

自主代理 AI 提高企業準備的門檻

GenAI 自主代理正超越第一線客戶的聊天機器人，深入至核心營運中。自主代理 AI 代表企業採用 AI 的下一階段——並在複雜度上有跳躍式的轉變。與單一用途的模型不同，自主代理系統包含多個 AI 元件，能夠自主互動、傳遞資料、觸發行動並在極少的人類干預下做出決策。

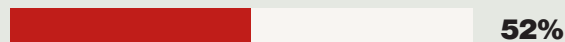
圖 4

首席資訊長計劃在未來 12 個月內於整個企業中部署 GenAI 自主代理

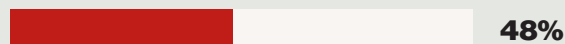
IT 營運與 DevOps



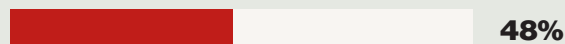
資料分析與決策支援



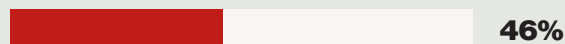
產品或服務開發



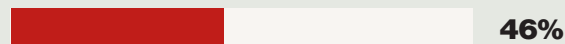
第一線客戶功能



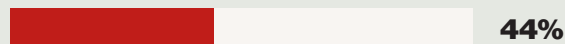
監控與警示



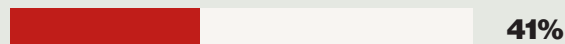
商業流程自動化



內部知識管理與員工賦能



治理、風險與合規性風險



資料來源：FOUNDRY

- **一致的治理、具備內建風險管理以及提升跨環境的能見度、權責歸屬與政策強制執行力：**由於 AI 建模的敏感資料通常必須保留在國家或組織的邊界內，因此 AI 平台必須執行資料落地、尊重區域主權，並實現對資料處理位置及方式的管控。在企業環境中部署大型語言模型 (LLMs)，需要健全的驗證管線、持續的效能監控，並與風險及合規性政策保持一致。AI 治理政策將需要涵蓋資料血統、存取控制、使用情況監控以及生命週期政策執行。

← 接續自第 7 頁側欄

自主代理 AI 部署的升級從根本上改變了企業風險。每一個代理都會引入全新的治理需求：它可以存取哪些資料、可以影響哪些系統，以及它的行動如何受到監控與限制。當多個代理開始在連鎖工作流程中協同工作時，資料外洩、成本可預測性、效能以及權責歸屬等方面的風險會疊加、惡化。

許多企業渴望探索企業級 AI 計畫以加速自動化與決策制定。然而，在缺乏成熟營運模式的情況下部署代理，往往會導致控制碎片化並產生盲點。當 AI 系統進行動態演進與互動時，為靜態應用程式建置的傳統護欄是不夠的。

對首席資訊長而言，企業級 AI 加強一個關鍵教訓：未來的創新取決於今日的基礎。早期投資於治理、觀測能力與標準化 AI 營運的企業將能居於有利優勢，安全且自信地採用 AI 自主代理。而未及早投資的企業則可能會發現創新的步伐超出了控制範圍，導致風險擴展的速度超越了其所創造的價值。

■ 可預測的成本以及展現 GenAI 投資 ROI 的能力：

碎片化單點解決方案會加劇成本波動，並限制其交付正向商業成果的能力。相較之下，具備端到端可觀測性的平台簡化效能與成本優化，使得在 AI 成長時，更容易辨識瓶頸、調整工作負載並控制成本。在跨混合環境下，GenAI 應用程式的自動化 Day 2 營運提升效率並有助於實現 ROI。

「AI 營運環境整合營運企業級 AI 所需的完整技術棧——從實體基礎架構層到協調編排、平台服務以及集中式管理，」Kotikela 表示。「目標是打造一個能在該技術棧的各個層級中提供最佳組合能力的平台。」

在 GenAI 時代成為創新推動者

隨著 GenAI 的採用迅速進入各種關鍵任務的正式量產環境，包括 SaaS 應用程式與企業工作流程，它正從一項創新計劃轉變為企業營運策略，直接影響到成本結構、韌性、治理與長期競爭力。這就是為什麼首席資訊長必須將他們應用於核心系統的相同紀律導入 GenAI 環境中——但必須具備更大的規模與更高的速度。

對於建構一個旨在自動化與優化流程及工作流程的營運模式而言，定義正確的 GenAI 應用情境很重要。該營運模式可作為更新資料管線與基礎架構的藍圖，以支持 AI 工作負載與應用程式。

平台化方法有助於首席資訊長在全企業內將 GenAI 從試產推進到正式量產，提供一個即開即用、具韌性、安全且可規模化的環境，進而加速創新並交付正向的商業成果。這對任何首席資訊長而言都是一場勝利。

GenAI 投資正在成長，但回報又如何呢？

根據 IDC 的數據，GenAI 的核心 IT 花費預計將急劇增長，從 2024 年的 401 億美元飆升至 2027 年的 1,511 億美元，¹IDC 將這段時期稱為全球企業關鍵的 AI 建置期。

Foundry 的調查發現，大多數 IT 主管將在未來兩年內增加其 GenAI 預算，該期間的平均增幅達 12%。

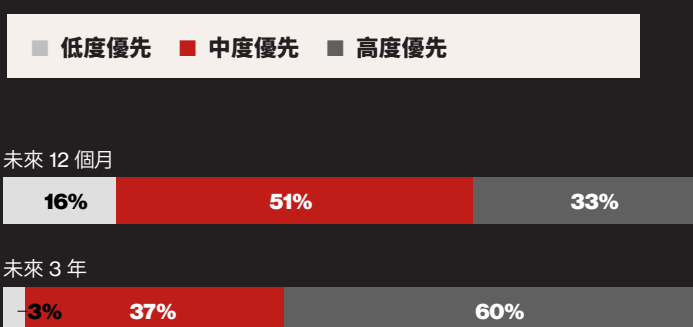
此外，企業領導者預期這些投資能相當快速地回收成本。超過八成 (84%) 的受訪者對於他們將在未來 12 個月內實現正向的投資報酬率 (ROI) 抱有至少適度的信心。而且幾乎所有受訪者 (97%) 都有信心，他們的 GenAI 計畫將在未來兩年內交付可衡量的商業價值 (參閱圖 5)。

當被問到哪些能力最能促成未來三年內成功實現規模化的 GenAI 採用時，首席資訊長與企業領導者指出了以下幾點：

1. 展現 GenAI 投資可預測 ROI 的能力
2. 創新與實驗的速度
3. 獲取專業人才的管道
4. 基礎架構的韌性與效能
5. 治理與負責任的 GenAI 實踐

隨著 ROI 的壓力增加，首席資訊長可能需要將其 GenAI 工作流與應用程式從單點解決方案轉向以平台為中心的方式。可規模化的平台讓 IT 主管能交付自動化工作流程、資料主權以及端到端觀測能力。這些 GenAI 部署的關鍵因素將提高及時且永續回收成本的可能性。

圖 5
IT 主管對 GenAI 的 ROI 抱有高度期待



資料來源：FOUNDRY

1 IDC, 「利用 GenAI 機會創造實現市場成長。」

誰擁有企業級 GenAI？

為什麼權責不明是一項潛在風險

隨著 GenAI 在整個企業中規模化，許多企業內部悄然浮現一個問題：權責歸屬不明確。調查數據顯示 GenAI 治理的職責通常分散在 IT、資訊安全以及 AI 或資料高階主管之間——不論是規劃或是順其自然。這種共同分擔的模式反映出 AI 跨部門的本質，但同時也引入了風險。

與傳統應用程式不同，GenAI 系統橫跨資料管線、模型、基礎架構與自動化決策。若缺乏明確的權責歸屬，企業將難以在資訊安全、合規性、成本控制與效能方面強制執行一致的標準。漏洞的出現並非因為團隊缺乏意願，而是因為沒有任何一位高階主管去主導端到端的營運模式。

對首席資訊長而言，這種不明確是挑戰，也是機會。隨著 GenAI 成為企業基礎架構的核心部分，對領導力的預期正轉向集中式治理，結合去中心化執行。IT 主管不需要去控制每一項 AI 計畫，但他們被期待去定義其營運架構：如何在整個組織中部署、治理、監控並安全地擴展 AI。

及早解決 GenAI 主導權問題的企業將能獲得更多的主控權——同時也能贏得速度、信心與公信力——隨著 AI 擴展至關鍵任務的工作流程中。

常規營運 GenAI 的策略性架構

數十年來，企業 IT 一直將 Day 0 部署與 Day 2 營運區隔開來。GenAI 打破了這條界線。儘管許多組織目前仍聚焦於模型上線與試行，但長期的成功取決於 AI 系統在跨環境正式上線後的運作表現。

GenAI 引入全新的營運需求，超越傳統的應用程式管理範疇。首席資訊長必須考量持續性的模型監控、資料血統追蹤、代理行為以及基於 Token 的成本波動性。觀測能力不能再僅侷限於系統上線時間，還必須涵蓋效能、準確性、使用模式以及政策合規性。資訊安全也從邊界防禦擴展到模型存取控制與機器對機器之間的信任機制。

GenAI 的 Day 2 營運也必須自動化。隨著 AI 系統的增加及演進，過去適用於早期工作負載的人工干預方式很快就會難以承受。企業需要內建的韌性、自動化修復機制以及跨混合環境的一致性治理。

首席資訊長所說的企業就緒級 AI 是什麼

- 可預測的經濟效益
- 受治理的資料流
- 混合架構執行
- Day 2 自動化
- 與商業目標一致的成果

對首席資訊長而言，這個教訓很明確：GenAI 不能被當作工作負載的特例。它必須以如同營運 ERP 或 CRM 般的嚴謹態度來進行管理，並輔以適合 AI 規模、速度與自主性的全新控制機制。掌握 Day 2 營運的企業才能成功將 AI 的野心轉化為長久的優勢。

前往 [Nutanix Executive Focus](#) 獲取更多關於此主題的高階主管洞察與思想領導力內容。

關於本研究

Foundry 在 Nutanix 的贊助下進行一項線上調查，訪問來自亞太地區、北美和西歐的首席資訊長與資深 IT 主管。受訪者皆就職於員工人數超過 1,000 人的組織，且其企業已試行或正式實施生成式 AI 解決方案。調查於 2025 年 12 月 7 日至 12 月 23 日進行。

關於 Nutanix

Nutanix 是混合多雲運算的領導者，為企業組織提供統一的軟體平台，以便在任何地方執行應用程式、部署企業級 AI 工作負載並管理資料。藉由 Nutanix，企業組織可以簡化傳統與現代化應用程式的營運，進而釋放資源以專注實現商業目標。Nutanix 深受全球超過 30,000 家用戶的信賴，致力於協助企業組織實現數位轉型，並以一致、簡單且具成本效益的方式驅動混合多雲環境。請造訪 www.nutanix.com 了解更多資訊。

CIO

贊助單位

NUTANIX