

# 為 AI 發揮資料價值的 7 個步驟

有效資料管理的全方位指南



# 透過資料建立競爭優勢

在現今快速演變的商業環境中，資料是關鍵資源，是創新和競爭優勢的基石。若領導者能妥善運用強大、準確且全面的資料集來加速業務決策，便可脫穎而出。這一點在採用[生成式 AI \(GenAI\)](#) 上表現得最為明顯。

GenAI 承諾帶來無與倫比的進步和效率，完美的資料當然不是其可以提供創造性、迭代和靈活輸出的先決條件。但是，在某些時候，您將有必要評估您的「資料倉庫」以確保其井井有條。換言之，您需要的不僅僅是**存取**資料，更是**強大的資料管理實務**。隨著您的擴展，這將變得更加必要。

強大的資料管理是必要的，但這並不容易。

部署 [AI 工作負載](#) 需經過複雜且嚴苛的程序，其中資料科學家等人要識別適當的資料集，並確保資料乾淨且全面。這一努力超越了技術挑戰，需要整個組織的流程和強大的資料管理框架。

64% 的企業相信人工智慧將有助於提高其整體生產力。<sup>1</sup>

70% 的頂尖企業在將資料整合至 AI 模型時，遇到困難<sup>2</sup>

83%–92% 的 AI 專案失敗<sup>3</sup>

1 Forbes Advisor : [《How Businesses Are Using Artificial Intelligence in 2024》](#) (企業如何在 2024 年使用人工智慧)

2 McKinsey Report : [《The state of AI in early 2024》](#) (2024 年初的 AI 狀態)

3 Fortune : [《Want your company's AI project to succeed? Don't hand it to the data scientists, says this CEO》](#) (希望貴公司的 AI 專案取得成功？這位 CEO 表示不要交給資料科學家)



# 資料管理旅程

為了協助駕馭這個錯綜複雜的環境，我們全面概述了資料管理旅程，並分為七個基本步驟。這些步驟是 Dell 專業資料科學家與各種組織攜手舉辦廣泛研討會和諮詢的成果，其中提出了建立可擴充且有效 [AI](#) 模型時的常見挑戰和成功策略。在接下來的幾個月裡，我們將更深入探討每個步驟，協助您克服組織的特定障礙，並堅持完成 AI 歷程。

本電子書為瞭解和實施有效的資料管理實務，提供了一個可操作的清晰框架。透過採用這些原則，您的組織可以從試驗 GenAI 變為擴大資料轉型規模，從原始資源轉變為策略資產，在 AI 導向的未來中，為創新和持續競爭優勢鋪平道路。

下載此[資訊圖表](#)，輕鬆參考這 7 個步驟。



**步驟 1**  
識別商業需求



**步驟 2**  
加速相關資料探索



**步驟 3**  
簡化資料探索與存取



**步驟 4**  
最佳化分析、ML 實驗和建模



**步驟 5**  
調整資料和分析產品化規模



**步驟 6**  
自動化資料管理和治理



**步驟 7**  
評估業務成果





## 步驟 1

## 識別商業需求



此步驟為您的資料管理旅程設定了方向，使其與策略性業務需求保持一致。如果沒有定義完善的目標和指標，就不太可能實現任何有意義的價值。

- 首先要瞭解營運目標，以及其可以帶來的價值。
- 在整個組織內，就期望成果以及如何衡量成功，達成一致。
- 對將要創造的價值建立清晰的願景。

透過瞭解您要達成**什麼**以及**衡量方式**，您可以確保所有後續資料管理工作都有其目的，並指向可實現的明確目標。

## 步驟 2

## 加速相關資料探索



透過清晰的藍圖，加速探索相關資料。並非所有可取得的資料都是必要的，資料科學家必須快速識別與解決問題相關的資料。透過編目來源資料集和建立中繼資料，在資料及其價值之間建立明確的連結。這種集中式的方法可確保您的資料工作具有效率和效果。

透過快速精確定位相關資料，您可以節省**時間**和**資源**，讓資料科學家能夠專注於直接支援您策略性目標的高影響力資料。

68% 企業可獲得的資料尚未開發<sup>4</sup>

<sup>4</sup> Seagate Technology：《Rethink Data: Put More of Your Data to Work—From Edge to Cloud》  
(重新思考資料：從邊緣到雲端，讓更多資料發揮作用)。





## 步驟 3

## 簡化資料探索與存取



如果要將資料移至集中位置進行分析，則會降低即時使用案例的價值。讓資料科學家在資料所在位置存取資料，以避免這種低效率問題。標準化流程並實施自動探索，將結構化和非結構化資料安排在易於存取的位置，從而促進即時探索和利用。此方法不僅能節省時間，還能提升資料作業的靈活度，讓您**更快獲得深入見解並做出回應更快的決策**。

隨著資料的爆炸式增長，成功擴充 [AI 工作負載](#) 需要有效的資料管理。因此您可以考慮使用 [Dell Data Lakehouse](#)，這項工具提供完整的解決方案堆疊，以單一聯合查詢引擎來探索、查詢及處理所有企業資料，無論其位置或資料來源為何。



5 Seagate Technology : [《Rethink Data: Put More of Your Data to Work—From Edge to Cloud》](#)  
(重新思考資料：從邊緣到雲端，讓更多資料發揮作用)。





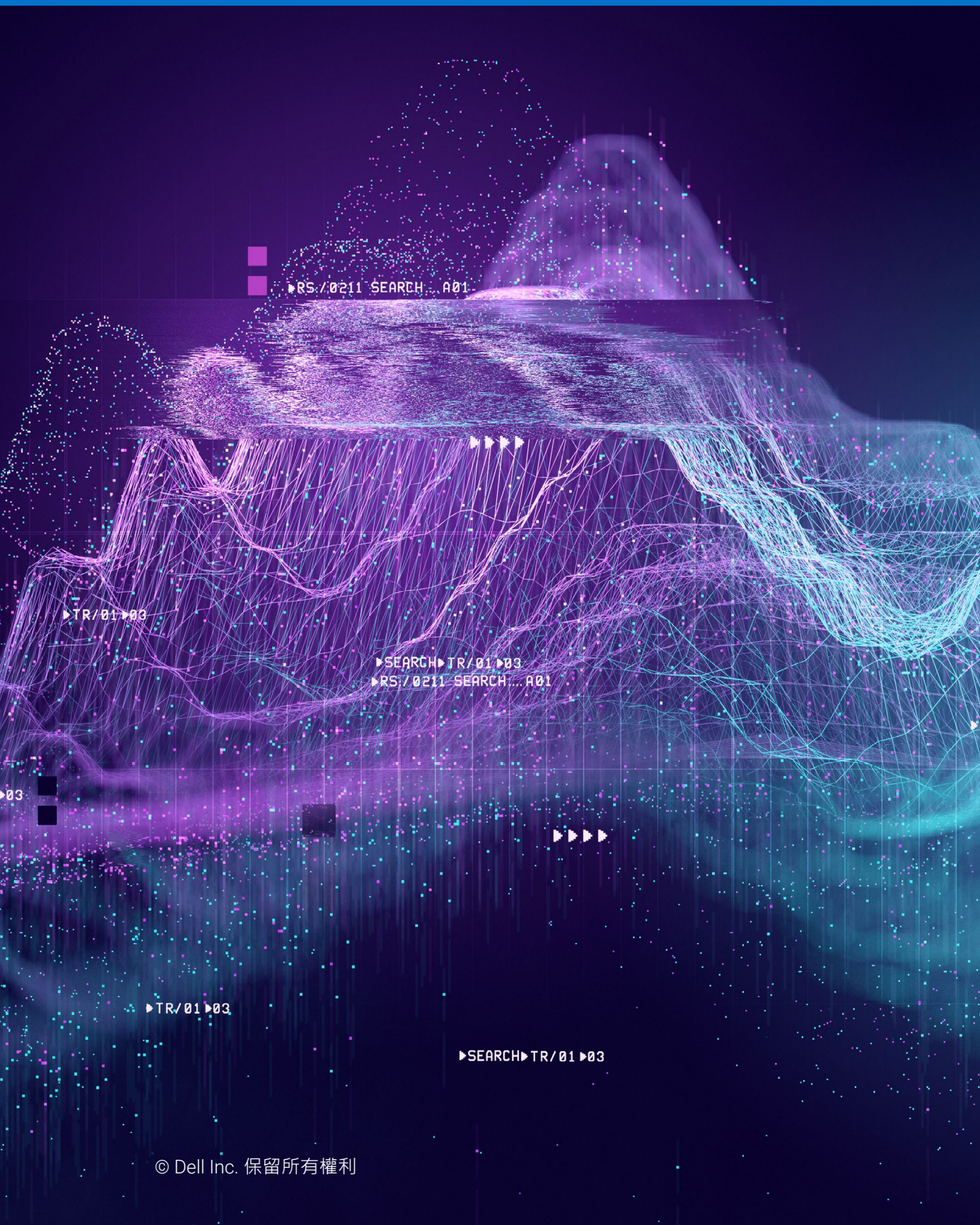
## 步驟 4

## 最佳化分析、ML 實驗和建模



鼓勵不斷進行實驗和建模，以找出可以解決問題的變數。考慮建立合成資料，這在面臨資料品質和隱私權挑戰時特別有用，尤其是在剛展開 AI 歷程時，有助於加快該程序。利用只需要增強和微調的預先訓練基礎模型，可能是一個很好的起點。

擴大專案規模後，您可以專注於輕鬆存取資料。這會讓您的團隊能夠透過迭代測試來最佳化分析。支援多次迭代和演算法的平台可以探索關鍵資料變數、提升生成式 AI 的成效，以及發揮資料價值。這種測試、學習和微調的迭代過程可確保您擁有**健全的模型**，以及**可據以行動的深入見解**，從而推動持續改進和創新。





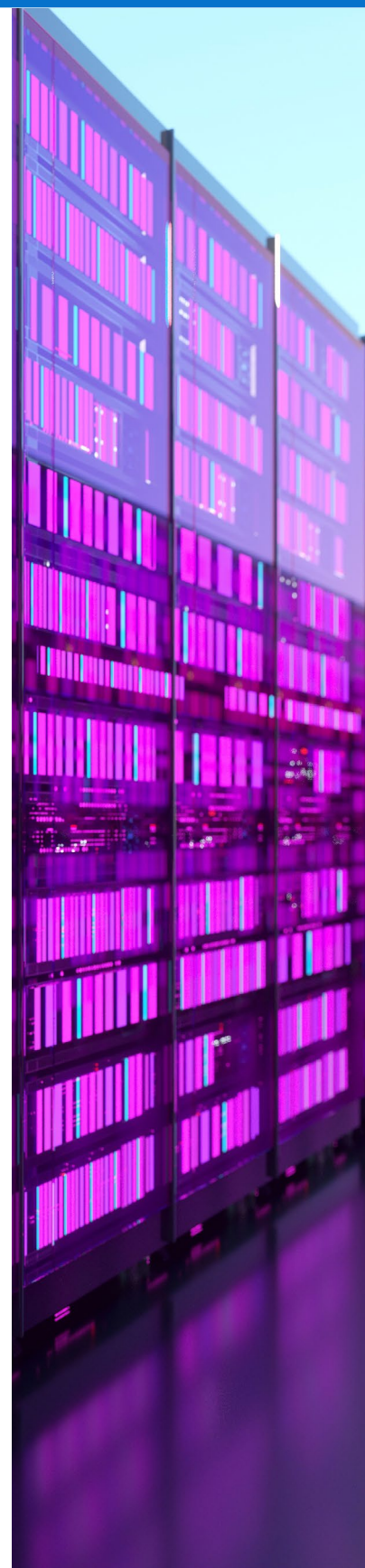


## 步驟 5

## 調整資料和分析產品化規模



從資料科學**專案**轉換成可靠、可重複的資料科學**產品**。這些產品可以獨立運作，並定期檢視以改良。高效率的資料輸入程序可加速獲得深入見解，讓資料管道透過內建智慧功能，即時分析和運用資料，達到預先定義的目標。透過擴展資料產品，您可以確保一致性和可靠性，將**臨時專案**轉變為**可產生價值的永續資產**。



## 步驟 6

## 自動化資料管理和治理



隨著產品變得可重複，將組織的資料管理和治理流程自動化。這可讓系統自我監控，並在異常情況成為問題之前，加以標記。採用整體的可觀測性，確保工作流程在整個資料管理過程中，遵守整合式資料治理標準、原則及安全措施。自動化不僅**減少了手動作業**，還**提高了準確性和法規遵循**，為大規模管理資料提供了強大的框架



## 步驟 7

## 評估業務成果



評估來自業務成果的資料並將其反過來輸入至流程中，培養持續學習和適應的文化。此持續評估會擷取實際結果與預測結果，從而進行細微調整，以改善和最佳化成果。以最少的人為干預利用功能，可確保持續的效能和價值。藉由持續分析成果，您可以建立有助於**持續改善**的意見反饋迴圈，並調整您的**資料計畫**，使其與不斷演進的業務目標達成一致。





# 透過 Dell AI Factory with NVIDIA，讓資料發揮作用

您目前的資料中心和 IT 營運模式可能不足以駕馭 GenAI 的速度和規模。您需要專為滿足 AI 的特定需求而建構的新型資料中心。

## 您需要 AI Factory。

正如實體工廠推動了工業革命一樣，AI Factory 也將促進 AI 革命。但是，AI Factory 產生的不是實體商品，而是智慧功能。AI Factory 功能多樣且具有彈性，可在您資料所在的任何位置運作，無論是雲端、資料中心、工作站還是邊緣位置。

# 2.4 倍

訓練大規模 ML 模型的年度支出增長 2.4 倍 (主要是在 GPU 和伺服器等硬體上)<sup>6</sup>

6 Epoch AI：《How Much Does it Cost to Train Frontier AI Models?》  
(訓練全新領域的 AI 模型需要多少錢?)

© Dell Inc. 保留所有權利







瞭解您的使用案例並套用正確的資料管理解決方案，是最佳化 AI 工作的基礎。

**Dell AI Factory with NVIDIA** 提供必要的建置區塊，協助您加速 AI 創新：

- 著重於資料
- AI 最佳化基礎結構
- AI 軟體和模型
- 開放生態系統
- 專家服務
- 最佳實務

這項技術以您的業務需求為依據，以更低的總體擁有成本快速提供正確的解決方案。Dell 和 NVIDIA 已花費超過 340,000 個小時驗證這些解決方案，並提供參考設計與效能指標。

NVIDIA 和 Dell 攜手合作，提供可加速 AI 採用的完整堆疊解決方案：**Dell AI Factory with NVIDIA**。

透過 NVIDIA AI Enterprise，您可以：



### 存取、管理、部署及支援

涵蓋一系列基礎結構解決方案的 GenAI 模型、應用程式、解決方案和加速器。



### 利用 NVIDIA NIM

使強大的資料管道能夠充分利用所有這些資料管道工作



### 輕鬆轉換

從一種模式到另一種模式，同時保持資料管道的穩固地位、提供投資保障，並使您始終能夠採用最佳的可用模型





# 透過持續的資料管理 維持成功

資料管理旅程不是一次性任務。這是一個持續的過程，利用一系列實務和工具，幫助您持續發揮資料的價值。而這對於任何成功的 GenAI 計畫都至關重要。Dell 與 NVIDIA 合作，提供全球最廣泛的 AI 解決方案產品組合。



## 端對端解決方案

從桌上型電腦、資料中心到雲端，我們的端對端解決方案均採用 NVIDIA AI 技術，可提供安全性、存取能力及擴充能力。



## 致力於 AI

Dell AI Factory with NVIDIA 體現了 Dell 採用和實作 AI 的承諾，協助客戶加速其 AI 計畫，以充分利用資料並達成目標。



## 簡化的資料導覽

有了 Dell 和 NVIDIA，您可以輕鬆完成資料管理旅程。而在 AI 導向的資料工作負載領域中，過程和目的地同樣重要。





## 資料管理不是一次性任務。

資料管理旅程是一個持續的過程，利用一系列實務和工具，幫助您持續發揮資料的價值。而這對於任何成功的 GenAI 計畫都至關重要。

## 「一體適用」的方法並不存在。

您擁有獨特的資料基礎：您專屬的軟硬體生態系統、資料位置和類型，以及您自己獨特的使用案例。本電子書中概述的步驟可作為總體指南，協助您掌握組織的特定需求。

## 而且您不是孤立無援。

[Dell Services](#) 可在生命週期的每個階段提供深度專業知識，為您的組織量身打造解決方案，加速將時間轉換為有感的價值。從協調一致的制勝策略和驗證資料，到快速實作資料平台，以及確保安全、最佳化的作業，我們值得信賴的專家隨時為您提供協助。

無論您處於發展旅程的哪個階段，我們都能協助您有效地利用企業資料，推動 AI 專案，並加速將時間轉換為價值。

造訪 [DELL.COM](#)



## 術語詞彙表

### 生成式 AI

生成式 AI 是指著重在創造新內容或資料的人工智慧組合。這種類型的 AI 使用演算法和模型 (例如生成式對抗網路 (GAN) 和 Transformer)，來生成文字、影像、音樂和其他類型的媒體。生成式 AI 的獨特之處，在於能夠生成模仿其訓練資料風格和內容的原創輸出，因此在創意任務、內容創作和模擬複雜情境方面非常有價值。

### AI

人工智慧 (AI) 是電腦科學的一個廣泛領域，專注於建立能夠執行通常需要人類智慧之任務的系統。這些任務包括學習、推理、解決問題、感知、語言理解和決策。AI 包含各種技術和方法，例如機器學習、自然語言處理和機器人學。AI 旨在建立能夠適應新情況的智慧型代理程式，隨著時間的推移提高效能，並在廣泛的應用中，輔助或自動化人類活動。

### AI 工作負載

AI 工作負載是指 AI 系統處理的特定任務或流程。這些工作負載可能因應用程式而異，並可能包括資料處理、模型訓練、推論和分析。AI 工作負載的特點是訓練模型、處理大型資料集和執行即時決策所需的密集運算需求。管理 AI 工作負載涉及最佳化硬體、軟體和演算法，以有效處理這些繁重的工作，確保 AI 系統能有效且大規模地執行。

## 參考資料和 延伸閱讀

[資訊圖表：資料管理旅程](#)

[Dell Professional Services for Generative AI](#)

[Dell 資料管理解決方案](#)

[Dell AI Factory with NVIDIA](#)

[生成式 AI 的 5 大安全考量](#)