

出國報告（出國類別：開會）

參加顧能公司  
「2024 年資料與分析峰會」  
(Gartner Data & Analytics Summit  
2024)

服務機關：數位發展部

姓名職稱：羅啓建高級分析師

派赴國家：日本

出國期間：113 年 5 月 20 日至 5 月 24 日

報告日期：113 年 8 月 5 日

## 摘要

顧能公司(Gartner)為國際頂尖之資訊技術研究與分析公司，提供分析報告與專業技術顧問諮詢服務，協助客戶進行業務規劃、研提解決方案及策略分析等。該公司每年於日本東京召開跨國資料與分析峰會，世界各國資訊科技領域巨擘專家、企業、技術工程師及政府官員均積極參與。會議以實體演講、論壇、工作坊等形式舉行，提供資料包含新興技術、前瞻資料架構管理、隱私強化技術(Privacy Enhancing Technologies, PETs )、資料經緯(Data Fabric)、資料觀測(Data Observability)等主題領域，結合先進國家推動經驗進行探討，我國可自中獲取國際推動現代化資料管理措施方式，透過交流共同探討國際間資料應用兼顧隱私保護之可遵循方針。

本次會議該公司從「分析和人工智慧」、「資料管理」、「商業價值與領導力」及「信任、治理與文化」四大主軸，鋪陳人們可有效利用資料來取得業務成果，且生成式人工智慧正在加速這一進程。然而，諸如主資料的基本問題、資料品質管理以及使用資料的道德、隱私風險和合規性仍未解決，本次會議由各面向探討問題的解決方法，例如利用趨勢技術來實現業務成果、資料分析領導者如何使用資料經緯技術應對長期存在的問題等。

# 目錄

壹、	峰會背景與目的 .....	1
貳、	會議過程 .....	3
參、	共同創造價值：從基礎知識到 AI Ready 再到集體智慧 .....	5
肆、	2024 年資料和分析的主要趨勢 .....	9
伍、	增強資料品質解決方案：資料品質技術與實踐的革命 .....	13
陸、	資料與分析中的每個人都需要了解（生成）AI：基礎知識 .....	18
柒、	資料經緯(Data Fabric)變得真實：建構資料經緯的十個實用步驟 .....	24
捌、	心得 .....	31
玖、	建議 .....	32
壹拾、	附錄 .....	34

## 圖目錄

圖 1-集體智慧關注之 8 大面向 .....	5
圖 2-共同創造價值 3 大關鍵點 .....	6
圖 3-AI 技術的使用案例應用在顧客服務 .....	7
圖 4-如何達到 AI-Ready Data .....	8
圖 5-組織可能要面對的問題 .....	9
圖 6-資料不信任問題 .....	11
圖 7-應對趨勢的關鍵行動方案 .....	13
圖 8-詮釋資料驅動的自動化 .....	14
圖 9-AI/ML 驅動的增強 .....	15
圖 10-知識圖譜驅動的洞察 .....	16
圖 11-NLP/LLM 驅動的方法 .....	17
圖 12-增強資料品質與生成式人工智慧 .....	18
圖 13-AI 工具箱 .....	18
圖 14-AI 應用領域 .....	19
圖 15-AI 發展路線與價值 .....	21
圖 16-AI 治理與風險管理 .....	22
圖 17-AI 實作技能與技術準備 .....	22
圖 18-AI 技術實作公式 .....	23
圖 19-資料經緯在生成式 AI 中的定位 .....	24
圖 20-實作主動式詮釋資料 .....	26
圖 21-自助式資料交付流程 .....	28
圖 22-資料產品管理 .....	29
圖 23-自建或購買資料經緯技術 .....	30
圖 24-會議議程 1 .....	34
圖 25-會議議程 2 .....	35
圖 26-抵達會議入口 .....	36
圖 27-完成報到 .....	36
圖 28-會場寬廣附有休息座位 .....	36
圖 29-參與人數眾多 .....	36
圖 30-每場次會議入場登記 .....	37
圖 31-會議及講者介紹 .....	37
圖 32-大會場(僅拍攝部分前場) .....	37
圖 33-大會場(僅拍攝部分後場) .....	38
圖 34-會議室及展攤位置圖 .....	38
圖 35-展攤附有交流圓桌 .....	39
圖 36-展攤參與討論人員眾多 .....	39
圖 37-會議必備充電區 .....	39

圖 38-會場外自動車試運行 .....	39
----------------------	----

# 參加顧能公司「2024 年資料與分析峰會」

(Gartner Data & Analytics Summit 2024)

## 出國報告

### 壹、峰會背景與目的

「2024年資料與分析峰會」由顧能公司(Gartner)主辦，該公司成立於1979年，為國際頂尖之資訊技術研究與分析公司，全球超過2,500位研究者提供分析報告與專業技術顧問諮詢服務，協助客戶進行業務規劃、研提具體解決方案及策略趨勢分析等。

該公司每年於日本東京召開跨國資料與分析峰會，世界各國資訊科技領域巨擘專家、企業、技術工程師及政府官員均積極參與，平均每年有超過800位各國參與者、40家以上解決方案提供商、15位以上專家共同參與，超過40個會議議程（參與人員統計資料摘錄自峰會網站），會議以實體演講、論壇、工作坊等形式舉行，提供資料包含新興技術、前瞻資料架構管理、隱私強化技術(Privacy Enhancing Technologies, PETs)、資料經緯(Data Fabric)、資料觀測(Data Observability)等主題領域。

本次會議該公司從「分析和人工智慧」、「資料管理」、「商業價值與領導力」及「信任、治理與文化」四大主軸，鋪陳人們可有效利用資料來取得業務成果，且生成式人工智慧正在加速這一進程，在這國際趨勢中，數據已成為組織最重要的資產，如何有效地蒐集、分析和利用這些數據，成為組織競爭力的關鍵，而且生成式人工智慧正在改變人們的工作方式、團隊協作方式和流程運作方式，資料分析領導者必須理解並善用各種先進技術，以達成業務目標，峰會設置有網站（網址：<https://view.ceros.com/gartner/apac-data-analytics-japan/p/1>），網站提供研習會相關資訊。

數位發展部為推動全國數位發展之主管機關，為實現資料治理與資料基礎建設及協助公私部門數位轉型等相關業務推動，期結合先進國家推動經驗進行探討，我國可自中獲取國際推動現代化資料管理措施方式，透過交流共同探討國際間資料應用兼顧隱私保護之可遵循方針，爰派員參加此等國際研討峰會。

## 貳、會議過程

顧能公司2024年於日本東京舉辦資料與分析峰會，自113年5月21日開始展開至5月23日止，3天研討峰會參與多場次專家座談、專題發表與研究報告等，本次出國報告工作與行程內容摘陳如下：

日期	工作與行程內容概要
5月20日(一)	啟程臺北時間 07:20 抵達日本時間 11:35 確認環境與交通方式
5月21日(二)	會議第一天參加議程 <ul style="list-style-type: none"><li>● 共同創造價值：從基礎知識到 AI Ready 再到集體智慧</li><li>● 設計釋放創新的力量</li><li>● 2024 年資料和分析的主要趨勢</li><li>● 消除 AI 信任差距 — 建立資料分析信心的 10 個趨勢</li><li>● 雲端資料庫管理系統魔力象限 — 對亞太及日本組織的影響</li><li>● 如何創建最佳的資料和分析組織模型</li><li>● 具有可組合全球部署的首席資料辦公室</li></ul>
5月22日(三)	會議第二天參加議程 <ul style="list-style-type: none"><li>● 專家座談：尊敬的總裁：請讓我們了解一下管理情況以及 D&amp;A 對您的業務決策有幫助嗎？</li><li>● 溝通與相互理解的認知科學：為什麼你的意圖沒有傳達給別人？</li><li>● 人工智慧的熱門趨勢</li></ul>



- 
- 決策智能的新興實踐：資料、分析和人工智慧的下一步飛躍
  - 資料整合與工程的十大趨勢：投資實現資料管理基礎現代化
  - 增強資料品質解決方案：資料品質技術與實踐的革命
- 

5 月 23 日(四)

會議第三天參加議程

- 資料可觀測性技術的現況－檢測、建議和預防
  - 資料與分析中的每個人都需要了解（生成）AI：基礎知識
  - 如何提高和再培訓您的資料、分析和人工智慧員工
  - 自信地部署和管理 GenAI 解決方案
  - 資料經緯(Data Fabric)變得真實：建構資料經緯的十個實用步驟
- 

5 月 24 日(五)

整理會議資料與處理回程交通事宜

返程日本時間 12:40

抵達臺北時間 15:05

---

## 參、共同創造價值：從基礎知識到 AI Ready 再到集體智慧

生成式AI(ChatGPT)在18個月前推出，現今AI已成為全球關注焦點，並且已經在日本創造了巨大的價值，顧能公司副總裁分析師Gareth Herschel與資深總監分析師Melody Chien分析指出，在產品設計過程中，滿足不同利益相關者的需求是一項重大挑戰，設計師追求美觀，工程師追求功能性，營運團隊則希望產品易於製造和轉移，且所有團隊都希望盡快高效地完成工作。近年來，隨著生成式AI技術的發展，文本到圖像生成功能越來越多地應用於產品設計，以快速創建新產品原型。豐田研究院更進一步，開發了一個不僅能生成創意設計，還能理解工程約束（例如安全性和阻力）的系統。這確保了所有車輛設計都符合工程團隊的限制，同時保留了設計師的創造力。透過採用 AI 驅動的產品設計流程，豐田將新產品上市時間縮短了90%。此外該流程還提高了設計品質，降低了開發成本。豐田的案例表明，AI 可以成為產品設計的強大工具，幫助組織提高效率、降低成本並更快地將新產品推向市場。



圖 1-集體智慧關注之 8 大面向

在此過程之中，我們並應注意以下幾個面向如AI的目標、實作新的技術、確保更好的業務產出、建立並維持核心資料和分析功能、作業模式進化、價值的實現、準備AI-ready data與提供資料產品供員工使用(如圖1)，以利共同創造業務價值完成任務。

顧能公司本場次的講師分析指出資料分析和人工智慧可以達成共同創造價值(Generating Value Together)，並強調以下3項關鍵點(如圖2)：



圖 2-共同創造價值 3 大關鍵點

## 一、價值創造(Value Creation)

顧能公司調查發現80%的從業人員相信人工智慧的策略重要性超過過去10年，此外資料分析管理實踐從低成熟度轉向高成熟度會導致組織財務績效平均提高30%，因此顧能公司認為資料分析領導者要更多關注策略的擬定、文化的變革、管理資料分析能力與資料治理，以取得更好的業務成果。

同時在開發資料產品時更應注意該產品必須易於查找，可供業務使用和重複使用，資料工程師必須保持資料最新狀態，並進行適當的使用管理，正確創建和維護的資料產品會產生巨大的價值，例如加拿大政府建立了適當的管道來起訴金融犯罪，同時他們也與其他執法機構共享資料，並在揭露和起訴人口販運的過程中，這些內部和外部共享資料打破資料孤島，並共同創造價值。

但是組織通常仍有一個問題，就是他們不擅長衡量成功，不知道如何宣揚成功，所以在大眾眼中沒有得到足夠的業務成果，這也是為什麼我們需要資料來說故事，以展現業務成果，目前僅22%的人改變資料說故事的方法，且只有24%的人已經徹底改變資料分析指標以業務為導向，因此組織需要轉變，開始從業務成果尋找業務模型以及它如何適應我們組織的總體使命，將其化為可視化的資料，以快速獲得洞察力，並善用人工智慧和機器學習模型，幫助主管在查看資料時，預測應採取的行動，以促進資料循證決策，提高決策品質。

## 二、人工智慧目標(AI Ambition)

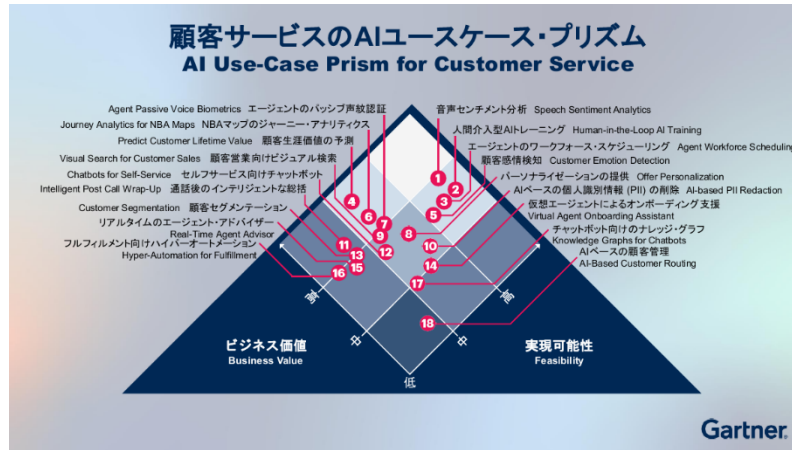


圖 3-AI 技術的使用案例應用在顧客服務

生成式人工智慧 (Generative AI, Gen AI) 正以其強大的創新能力，席捲全球各個產業，它不僅能夠生成逼真且原創的內容，更能跨越語言、圖像、聲音和影片等多種媒介形式，展現出無窮的應用潛力。在國際舞台上，AI已成為各國科技發展的重中之重，美國、日本、歐洲等主要經濟體紛紛加大投資力度，積極布局相關研發，組織若想導入AI應用，首先應確立AI目標以符合組織或業務的使命，以顧客服務為例，就有非常多可能的AI應用，有其各自的價值和實現的可能性(如圖3)。

2023年的CEO調查發現約有75%的CEO表示他們已經嘗試過Gen AI，64%認為Gen AI 並未被過度誇大，62%表示已經在董事會層面進行AI的討論，AI的蓬勃發展令人興奮，但他們也非常關注人工智慧競賽以及其風險，2023年更大規模的人工智慧調查發現，超過一半的技術領導者表示他們不確定他們的資料分析領導者能夠減輕人工智慧的風險，並應對所有相互競爭的優先事項立即採取行動。

組織為了發展AI技術，需要準確、可靠和明顯的資料，同時也需要強大的資料保護能力，人工智慧不僅僅只是產生ChatGPT的查詢結果，組織可以將現有的資料分析平台串聯，這樣可以利用現有的語義層、查詢引擎和架構更好利用生成式AI的能力，並保護資料、降低AI風險，如幻覺等，AI-Ready data在顧能公司的研究發現(如圖4)，其永遠不是一項已完成的任務，它需要不斷發展，從詮釋資料管理、資料觀測到資料

分析治理是一串連續過程，詮釋資料提供測量合格的資料，觀測與評價資料的使用情形，發展詮釋資料管理成熟度，以減少AI偏差並提供信任。

### 三、集體智慧(Collective Intelligence)



圖 4-如何達到 AI-Ready Data

資料分析和人工智慧技術(包括生成式人工智慧)的快速發展正在迫使75%的組織重新思考和發展他們的營運模式，以便實現更好的業務成果，以及支持更好的創新，因此顧能公司提出三個設計問題來指導新技術會帶來哪些新觀點：

1. 這會增強我們的集體智慧嗎？(Will this enhance our collective intelligence?)
2. 它會帶來業務成果、價值創造嗎？(Will it result in value creation?)
3. 是否足夠靈活和適應性強以利用新興機會？(Is it flexible and adaptable?)

集體智慧的產生需要擁有正確的跨職能團隊，且團隊能快速調整，主動收集詮釋資料，即時監控模型的性能，這些能夠使模型有效且安全地運行，一旦實作了新的營運模型，接下來需要提高組織的資料和人工智慧素養技能，在顧能公司2024年的調查發現資料分析領導者中有61%表示，教育是他們的首要職責之一，80%的人表示AI培訓將是未來12到18個月的優先事項，以及80%的人表示人工智慧教育和資料素養將在2024年受到更多關注，並且將資料素養擴展到人工智慧素養，利用對人工智慧教育的巨大興趣，來發展和建立資料和人工智慧勞動力，以便帶

領組織領先並獲得創造價值。

在集體智慧的新時代，集中控制和分散權力之間必須取得平衡，組織共同的願景不單是個人的願景，它是關於共同的責任和共同的結果，在利用資料達成業務目標時，資料如何被分配並且分析民主化，運用分散式決策能夠在邊緣發現洞察力，這種資料發現必須在共同願景和目標的指導下完成，因此決策必須下放到組織的每個人，組織的每個人可以看到自己行為的後果，因此在集體智慧的新時代，為了創造最大價值的組織目標，意味著資料分析素養必須深入到組織的每個成員。

最後當領導者將目光投向執行和戰略、適時關注資料治理和成熟度時，他們的眼界超越對投資報酬率的短期追求，關注業務成果，同時振興和擴展資料治理、探索創新營運模式的過程中，應著力培育資料素養，涵蓋並掌握人工智慧素養，賦予組織內每位成員——從核心到邊緣——理解和應用資料和人工智慧的能力。如此可以釋放團隊、企業和客戶的潛力，從而讓世界變得更美好，共同攜手迎接集體智慧的新時代，創造嶄新的價值。

## 肆、2024 年資料和分析的主要趨勢

在當今快速變化的資料和分析領域，了解最新的趨勢和技術對於保持競爭力

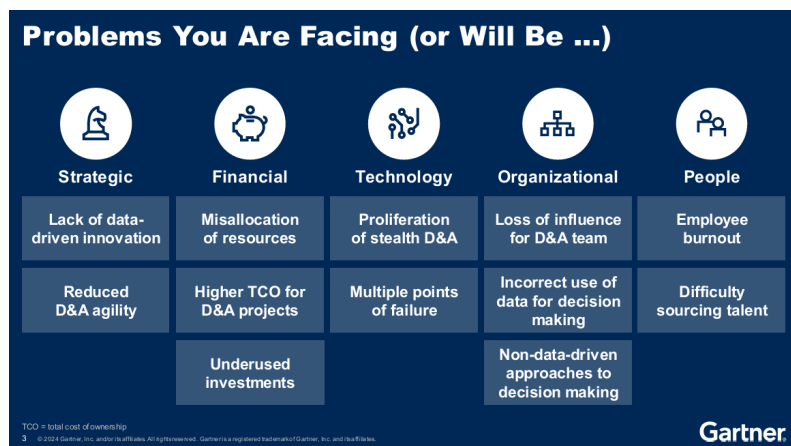


圖 5-組織可能要面對的問題

至關重要，顧能公司副總裁分析師Gareth Herschel提供2024年最重要的資料和

分析趨勢的深入洞察，講者藉由組織可能要面對的各項問題(如圖5)，強調了4個關鍵領域的轉變，包括從「足夠好」到「投注所有業務」的品質提升，從「混亂」到「管理複雜性」的轉變，從「單一真相來源」到「不信任洪流」的變化，以及從「過載」到「賦能」的人員管理轉型，這些轉變顯示出組織必須在資料品質、資料經緯、詮釋資料及資料應用上進行深度變革，才能更好地應對未來的挑戰並持續增長和創新，4個關鍵領域的轉變與相關建議作為如下：

一、從「足夠好」到「投注所有業務」：強調組織必須從滿足現狀轉變為賭上全局的資料戰略，經調查有74%技術和商業領導者認為生成式AI（如ChatGPT）將對各行業產生重大影響，顧能公司分析師認為可以朝以下方向發展：

1. **證明你對組織的價值**：如果無法展示資料和分析的貢獻，組織可能會面臨許多問題，包括資源錯配、影響力下降、員工倦怠、缺乏數據驅動創新以及失去影響力等。透過使用企業價值方程式來創建資料和分析能力與業務價值之間的清晰鏈接，可以優化資源配置決策，並包含如永續性或多元化目標等相關無形資產，以證明其價值。
2. **使用 FinOps 拉平成本曲線**：建立正式且集中的FinOps實踐來強制執行標準，使用資源標籤和專門的FinOps工具來追蹤和分配成本，定期進行供應商價格/性能評估以評估環境的健康狀況，尋找提供增強型FinOps功能的供應商，以推進成熟度和自動化目標。
3. **建立資料分析特許經營權**：辨識出組織內的最佳實踐技能、流程和技術，並將這些最佳實踐作為模板供他人參考，建立資料分析特許經營權，有助於組織在不同領域內達成一致的標準和流程，提升整體資料分析能力，並促進組織內部的協同合作，考慮可建立不同資料和分析領域的思想領袖，確保不同部門能夠最大化利用資料資源，進一步推動資料驅動的決策和創新。

二、從「混亂」到「管理複雜性」：調查發現47%的使用者很難找到有效執行工作所需的資料、36%的使用者由於資訊量過多錯過或未能注意到重要更新資訊、32%的使用者由於缺乏認知而做出錯誤的決定，因此建立

一個正式的結構和流程來管理和處理資料分析的複雜性，確保所有相關方對資料治理和管理的標準和政策有一致的理解和認同，並投資於人工智慧和自動化技術，以提升資料處理的效率和準確性，顧能公司分析師認為可以做以下事情：

1. **平衡複雜性**：避免透過增加複雜性來解決問題，而是要簡化治理流程，盡量使用政策而非規則手冊，並以應用設計思維原則來簡化技術環境，此外也可考慮利用資料分析驅動的複雜性作為競爭優勢的來源。
2. **理解生態系統**：應該投資於能夠捕捉上下文資料的系統，創建產品、流程、組織或環境的數位孿生，並投資於過程挖掘工具以發現、監控和改進業務營運和流程。使用人工智慧構建自動化響應複雜現象的能力，這樣可以更好地理解和管理組織所處的生態系統。
3. **增強與支援人工智慧的系統**：投資於增強資料管理工具，不僅可支持已經建立的使用案例，如資料目錄和資料遷移，還包括支援新興的使用案例，如主動式詮釋資料。透過自我學習人工智慧系統來應用於更多的使用案例，並開發增強分析能力，如自然語言處理（NLP）和自然語言生成（NLG），以觸及新興用戶群體。

三、從「單一真相來源」到「不信任洪流」：過去組織依賴單一的真相來源來進行決策，確保資料的一致性和可靠性，但現今資訊來源不斷增加，

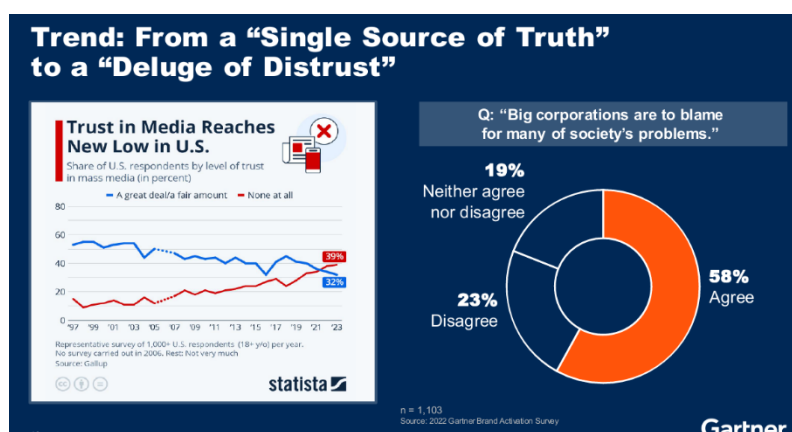


圖 6-資料不信任問題

造成了「不信任洪流」（如圖6）。越來越多人對大企業和政府機構所提供的資訊持懷疑態度，這種不信任感也影響到組織內部對資料的信任度。



此外有許多高階領導者對於生成式AI可能產生的隱私問題、誤用風險、偏見與有害內容等非常關注，因此顧能公司分析師認為可以朝以下方向發展：

1. **確保可解釋性和可審計性**：確保資料觀測可將實作由資料延伸到行動價值鏈，同時實作決策智慧來追蹤決策過程和決策結果。確保治理政策在績效和可解釋性之間設置清晰的目標，並投資AI可解釋性技能、工具和流程，為資料與分析團隊成員提供溝通技能訓練，以確保與業務合作夥伴有效合作。
2. **確保資料已是AI-Ready**：AI-Ready資料首先要確保資料經緯的完整性、一致性與品質，詮釋資料應該完備，以符合AI應用的需求。為避免資料偏差和錯誤影響模型的準確性和有效性，要從不同來源蒐集資料，協調不同的利害關係人群體，確保資料受到道德管理。還應考慮資料的隱私和安全性，確保在使用AI時不會洩露敏感資訊。

四、從「過載」到「賦能」：自Covid-19疫情開始以來，員工面臨著前所未有的壓力和倦怠，經調查67%的員工希望擁有更大的靈活性，59%的員工甚至願意為了更好的工作與生活平衡而放棄增加10%的薪資。為了應對這些挑戰，顧能公司分析師認為可以做以下事情：

1. **確保員工已是AI-Ready**：組織應該建立員工AI素養計畫，涵蓋需求、政策、最佳實踐AI技術等，並針對從CEO等級之高階管理者到操作層面的業務和技術人員進行培訓。同時應重新設計業務流程以促進人類與AI的協作，而非以人工智慧替代員工來設定期望目標。
2. **使用者賦權治理**：確保最終使用者有權做出基於分析的決策，教育使用者利用決策技術及如何避免決策錯誤，持續監督以避免私自利用服務分析，組織內的資料預設為「共享」，同時部署自助資料準備工具以應對資料的使用案例，以促進資料的利用。
3. **保護個人創新時間**：應該專門安排時間讓員工從事高影響力的副項目。確保這些項目與業務成果保持一致，以確保其相關性和可行性

在資料和分析的趨勢中，講者用4項關鍵領域轉變總結這11項行動(如圖7)，



圖 7-應對趨勢的關鍵行動方案

希望組織可以改變思考的方式，包含對複雜性的思考、管理創新的方式、價值的呈現和平衡組織的權力下放，讓權力在整個組織中流動，同時瞭解組織的環境，讓員工、團隊和技術都做好準備，以便能夠利用這些優勢來應對未來的趨勢與挑戰。

## 伍、增強資料品質解決方案：資料品質技術與實踐的革命

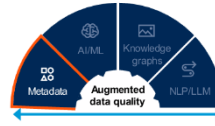
在人工智慧、大數據與物聯網等技術蓬勃發展的時代，資料品質已成為組織實現目標的關鍵因素，資料品質差可能導致洞察不準確、決策錯誤和錯失良機。因此，對有效資料品質解決方案的需求日益增長，顧能公司資深總監分析師 Melody Chien 奠基在現有資料品質技術，結合新興技術提出增強資料品質解決方案 (Augmented Data Quality) 是利用人工智慧 (AI)、機器學習 (ML)、圖譜分析 (graph analysis) 和自然語言處理 (NLP) 等技術來提升資料品質的能力，旨在改進洞察發現、下一步行動建議和流程自動化。

顧能公司本場次的講師分析指出增強資料品質解決方案的關鍵技術/功能有以下4種：

一、詮釋資料驅動的自動化：詮釋資料驅動的自動化利用內建的詮釋資料功能來發現、收集、分析和推論詮釋資料。這些功能能夠自動化處理常見的資料品質問題，應用於不同的資料環境(如圖8)。具體實現方式包括：

### Metadata-Driven Automation

- 1 Built-in metadata functions to discover, collect, analyze and infer metadata.
- 2 **Automate** repetitive tasks for common data quality issues across data environments:
  - Link data quality rules to similar data assets.
  - Apply access levels to expose or hide sensitive data based on roles.
  - Trigger remediation workflows (tickets creation, alert notification and owner assignment).
- 3 Can be packaged in:
  - Unified DM platform.
  - A stand-alone data quality application integrating with metadata management tools.



10 © 2024 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and/or its affiliates.

Gartner

圖 8-詮釋資料驅動的自動化

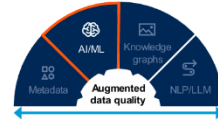
1. 連結資料品質規則到相似的資料資產：透過詮釋資料分析，系統能夠自動識別和連結適用於不同資料集的相同資料品質規則，提高規則應用的一致性和效率。
2. 基於角色應用訪問級別：系統能夠根據使用者角色，自動調整敏感資料的可見性和訪問權限，確保資料的安全性和合規性。
3. 觸發修正工作流程：當系統檢測到資料品質問題時，能自動生成工單、發出警報通知並分配問題責任人，迅速響應和解決問題。

例如某一家大型金融機構利用詮釋資料驅動的自動化功能，優化了其資料治理流程。系統自動連結了不同部門之間相似的資料品質規則，統一了資料標準，減少了手動管理的工作量。同時，基於員工角色，系統自動設置了資料訪問權限，確保了敏感金融資料的安全。當系統發現資料異常時，自動觸發修正流程，生成工單並通知相關負責人，快速處理和修正資料問題

二、AI/ML驅動的增強：AI和機器學習（ML）技術在資料品質管理中的應用日益普及，為資料品質提升提供了許多自動化和智慧化的解決方案，透過AI/ML驅動的增強，可以顯著提高資料處理的效率和準確性(如圖9)。以下是AI/ML驅動的增強在資料品質管理中的具體應用：

### AI-/ML-Driven Augmentation

- 1 **Augment** the data quality tasks such as discover, assess, validate, or enrich and monitor by using AI/ML algorithms.
- 2 Provide more in-depth and rapid execution of data quality tasks on individual datasets.
- 3 A combination of supervised, semisupervised and unsupervised methods to be deployed:
  - Supervised: Used when the entities and relationships in the data are well-understood.
  - Unsupervised: Learn from data and users' responses to detect patterns and identify outliers within datasets.
- 4 Commonly used in data profiling, data matching.



11 © 2024 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and/or its affiliates.

Gartner

圖 9-AI/ML 驅動的增強

1. 資料發現與評估：AI/ML演算法能夠自動發現資料中的結構和模式，並進行初步的資料品質評估。例如，在資料集的初步分析階段，可以利用機器學習技術自動檢測資料中的缺失值、異常值和重複數據，並評估這些問題對資料品質的影響。
2. 資料匹配和重複消除：利用機器學習技術，可以更精確地識別和匹配資料中的重複項，從而進行重複數據的清理。例如，在客戶資料庫中，可以透過ML技術識別同一客戶的多條記錄，並將其合併，確保資料的一致性和準確性。

例如某交通資料分析利用AI/ML技術自動檢測和修正資料中的異常值和缺失值。透過機器學習演算法，系統能夠自動識別異常的記錄，可能發現100歲的人仍在駕駛車輛，雖然不是說這種情況不可能，但決策者需要知道這種情況，這就是AI/ML幫助更深入和全面地處理資料品質問題的方式。

三、知識圖譜驅動的洞察：知識圖譜驅動的洞察利用知識圖譜將資料語義化，跟蹤資料的關聯和所有權，並創建領域資料映射，這些圖譜能夠提供關於資料流動的洞察，顯示故障點並分析相關影響(如圖10)。

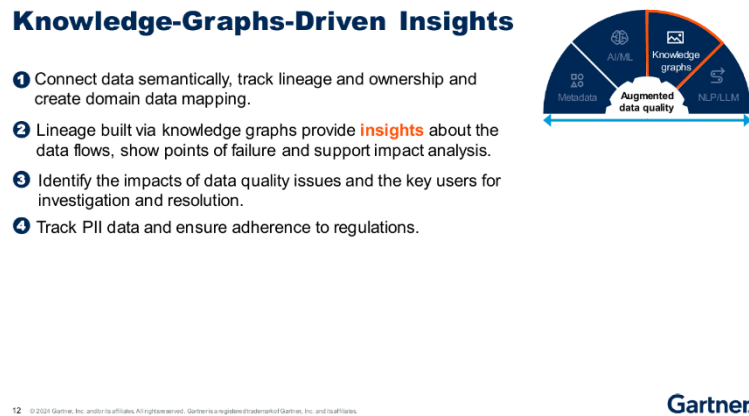


圖 10-知識圖譜驅動的洞察

1. 資料關聯和影響分析：知識圖譜能夠連接和追蹤資料的來源、變更歷史及流動路徑。這對於進行影響分析非常有用，可以快速識別資料品質問題的根本原因及其影響範圍。當資料源出現問題時，知識圖譜能幫助追蹤哪些下游系統和業務流程會受到影響，從而幫助組織及時採取補救措施。
2. PII資料追蹤和合規性確保：使用知識圖譜來標記和跟蹤個人識別信息（PII），確保這些敏感資料在處理過程中符合相關的隱私保護法規（如GDPR）。
3. 資料整合與領域資料映射：知識圖譜支持將不同來源的資料進行語義整合，創建統一的領域資料映射，從而提高資料的可用性和準確性。

例如在醫療領域，知識圖譜可以將患者的不同醫療記錄進行整合，提供全面的患者健康檔案，從而輔助更準確的診斷和治療。

四、NLP/LLM驅動的方法：NLP（自然語言處理）和 LLM（大型語言模型）技術在資料品質增強方面扮演著關鍵角色，這些技術能夠自動化許多繁瑣的資料管理任務，如分類、標籤、定義規則、資料剖析、監控以及清洗和轉換資料，這些功能能夠極大提高資料管理的效率和準確性(如圖11)。

## NLP-/LLM-Driven Approach

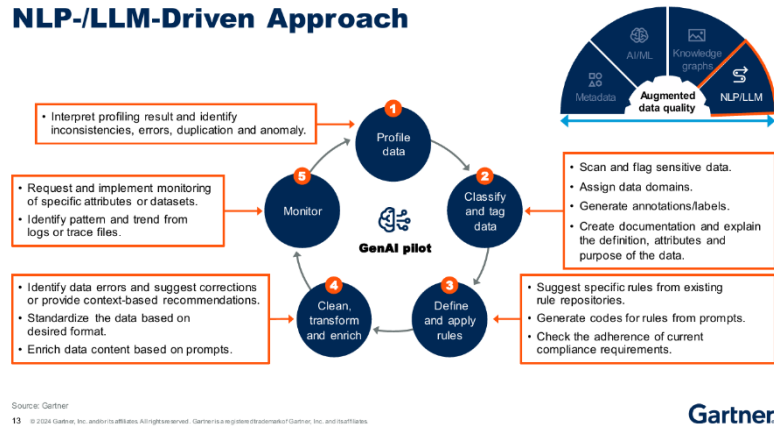


圖 11-NLP/LLM 驅動的方法

1. 剖析資料：解釋剖析結果，識別不一致、錯誤、重複和異常。
2. 分類和標籤資料：自動識別和分配資料領域，生成註解和標籤。
3. 定義和應用規則：從現有的規則庫中建議具體規則，並幫助生成和應用資料品質規則。基於歷史數據和使用者的行為，演算法可以自動生成適用的資料品質規則，並自動應用到新的資料集中，確保資料品質的一致性和可靠性。
4. 監控資料：請求並實施特定屬性或資料集的監控，可以用來持續監控資料流，檢測異常模式和潛在的資料品質問題。例如，可以設置自動化監控系統，利用ML演算法分析資料流中的趨勢和異常情況，並及時發出警報，提示需要人工介入進行進一步分析和處理。
5. 清洗、轉換和豐富資料：NLP/LLM技術可以自動清洗、轉換、豐富和補充資料內容，例如，可以利用自然語言處理技術從未結構化的文本中提取有用信息，並將其添加到結構化的資料集中，根據所需格式標準化資料，豐富資料內容，掃描並標記敏感資料。

整體而言增強資料品質和生成式AI的結合旨在使資料品質過程自助化，促進生成式AI的發展和應用，這些技術應用於資料剖析、資料錯誤識別、資料品質規則創建、資料清理和轉換、以及工作流程的創建，這其中資料觀測技術的導入尤為重要。生成式AI還能夠幫助準備訓練資料集和管理即時資料流，從而進行偏差檢測、異常檢測、資料豐富和標準化，並持續監控資料品質問題，最終我們塑造

我們的工具，然後我們的工具塑造我們(如圖12)。

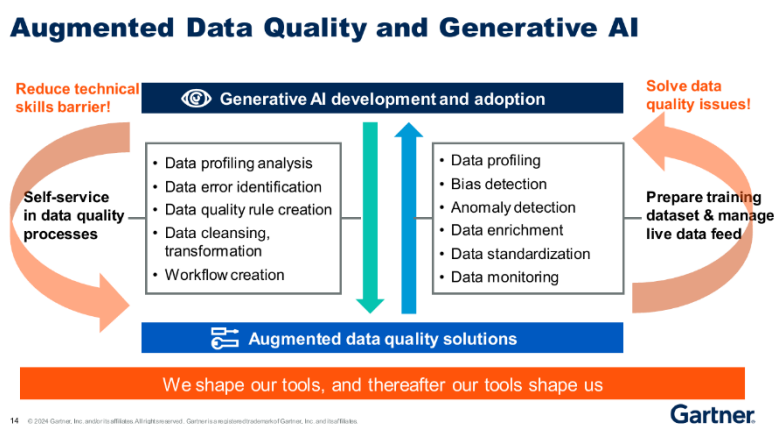


圖 12-增強資料品質與生成式人工智慧

## 陸、資料與分析中的每個人都需要了解（生成）AI：基礎知識

人工智慧（AI）已成為全球關注的焦點，無論是在科技領域還是商業應用方面，AI的重要性日益凸顯，AI不僅僅是學術研究的課題，更成為組織提升競爭力和創新的關鍵工具。生成式AI的出現，使得從自動化內容生成到智慧決策支持等各種應用場景變得可能，極大地改變了我們的工作方式和生活模式。從資料科學、感知系統到自主系統和內容生成，AI的應用範圍廣泛且深遠，然而，隨著AI技術的普及，也帶來了一系列挑戰和風險，如資料品質、隱私安全、偏見等問題，顧能公司總監分析師Ben Yan指出理解並妥善應用AI技術，成為各個組織在數位轉型過程中的重要課題。

### The AI Toolbox

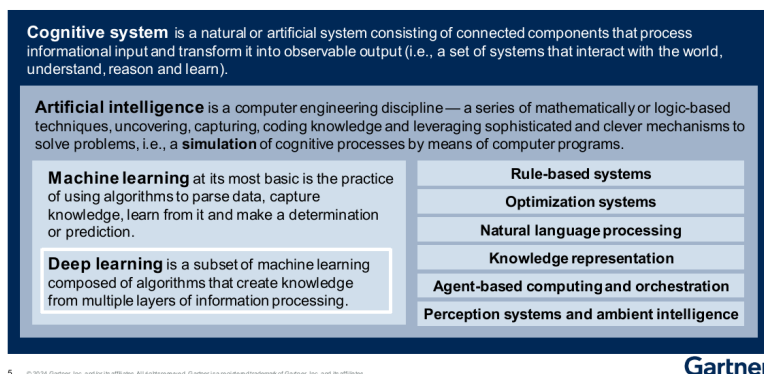


圖 13-AI 工具箱

一、**人工智慧的定義**：人工智慧是一門電腦工程學科，使用一系列基於數學或邏輯的技術，揭示、擷取、編碼知識並利用複雜而巧妙的機制來解決問題，即透過電腦程式模擬人類認知過程，其所使用的技術種類很多，常見的技術如機器學習、深度學習、類神經網路等(如圖13)。

二、**AI的應用領域**：AI的應用領域涵蓋了許多不同的方面，不僅止於生成式AI，資料科學在AI應用中扮演著關鍵角色，透過分析和處理大量的資料來提供洞察力和決策支持(如圖14)，對話互動是一個重要應用領域，包括自然語言處理和語音識別，這些技術使得人機交互更加自然和高效。

### AI Application Areas: Much More Than GenAI

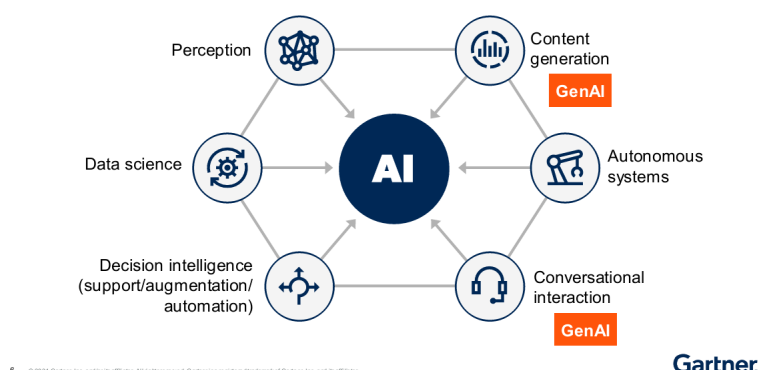


圖 14-AI 應用領域

感知系統則包括計算機視覺和感應技術，用於物體識別、圖像分析和環境感知。決策智慧透過支持、增強和自動化決策過程，幫助組織和個人做出更明智的選擇。自動系統如自動駕駛汽車和無人機，則展示了AI在自主行為和機器人技術方面的潛力。內容生成技術則用於創建文本、圖像和聲音內容，應用範圍包括寫作輔助、藝術創作和媒體生產等。經調查發現約有35%的組織尚在學習生成式人工智慧相關知識、約38%的組織在測試生成式人工智慧使用、約21%的組織已在進行生成式人工智慧投資，並且所有的組織均認為生成式人工智慧可能帶來的風險是需要加強保護的。

三、**常見的生成式AI風險與限制**：生成式AI在實際應用中面臨著多種風險與限制如下

1. 生成式AI可能產生事實錯誤的內容，這些內容可能會誤導用戶。



2. 生成式AI有時會生成無意義的輸出，這種現象被稱為「幻覺」。
3. 資料洩漏、隱私和安全問題也是生成式AI需要面對的挑戰，特別是在處理敏感資料時。
4. 生成式AI缺乏內容來源的透明度和歸屬標註，這可能會引發版權疑慮。
5. 偏見、刻板印象或有害的輸出也是生成式AI的風險，這些輸出可能會反映訓練數據中的偏見。
6. 生成式AI可能使用過時和有偏見的信息，這會影響其決策和建議的品質。
7. 高昂的（意外）成本也是生成式AI的潛在風險。

#### 四、關於AI的7個迷思：

1. **要開始AI專案，需要大量資料：**依照要完成的任務不同，有很多種類的AI設計，並非每種都要大量的資料才能完成，一般的通用基礎模型已經能完成許多任務。
2. **生成式AI是能解決大多數企業問題的靈丹妙藥：**生成式AI只是工具，且有其能力限制，要依照想解決的問題來設計使用方式，並非任何問題都善於解決。
3. **只要啟動生成式AI，它就能自動學習：**生成式AI仍要有人類的協助、調校，才能產生符合人類需求的內容。
4. **擔心AI會取代人類，使人類變得多餘：**AI不會直接取代人類，人類只會被更會使用AI工具的人類取代。
5. **AI本質上有偏見，會產生倫理問題，並且AI本質上是一個黑箱：**只要有良好的設計，AI可以發展為透明、可以信任的工具。
6. **AI在複製人腦或心智，並且很快會變得有意識：**在電影中常出現的情節，但是在現有技術發展，AI只是個工具不會產生意識。
7. **AI對人類是生存威脅，或者認為它將拯救人類：**在電影中常出現的情節，但只要使用平常心看待AI這個工具即可。

#### 五、 AI的價值及優先項目：

AI不僅僅是一項技術創新，更是一個能夠帶來巨大價值的工具，其價

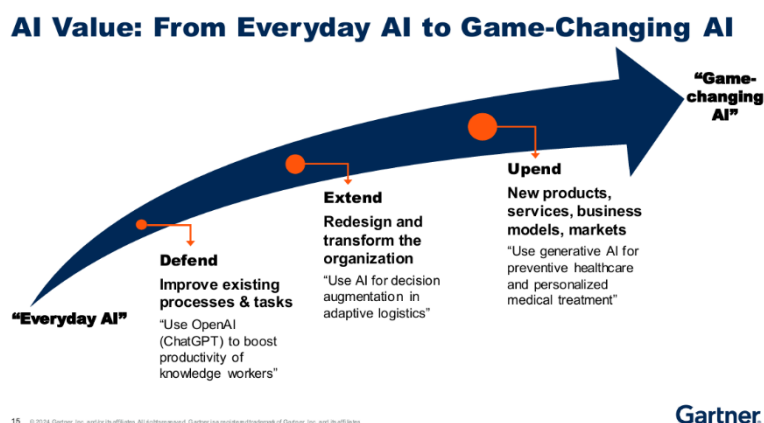


圖 15-AI 發展路線與價值

值可分為「日常AI」和「顛覆性AI」兩類(如圖15)。「日常AI」指的是那些可以在現有流程和系統中提供增量改進的應用，例如提高營運效率、優化供應鏈管理等；而「顛覆性AI」則是指那些能夠創造全新商業模式或市場的應用，如生成式AI創作藝術、音樂或開發全新的產品設計。

組織在選擇AI項目時，應該從防禦性項目開始，這些項目可以提供穩定的回報並降低風險。隨著組織對AI技術的掌握和信心的提升，可以逐步涉足擴展性和顛覆性項目。為了確保AI項目的成功，組織需要確保資料的準確性和安全性，並且要在組織內部培養AI素養，建立跨學科的協作團隊。

## 六、 實施AI技術的主要挑戰：

經調查發現組織在實施AI技術的前3大困難分別為評估和展示AI的價值49%，缺乏人才/技能42%，缺乏AI技術信心40%，因此組織在初期難以量化AI帶來的具體收益，這使得資源投入和決策變得困難，同時技術技能和人才可用性也是一大挑戰，專業的AI人才短缺，加之技術快速變化，導致組織難以組建和維持高效的AI團隊。

此外AI系統的效能高度依賴於大量高品質的資料來進行訓練和運行，確保資料的準確性、一致性和完整性是至關重要的，可以分別從5個向度確認，如受妥善治理的、安全的、公平的、豐富的、準確的。其中隨著AI技術在各個領域的應用越來越廣泛，保護個人資料和敏感信息不被洩露或

濫用成為一大挑戰；還有AI模型的透明度和可解釋性問題，特別是深度學習模型，往往被認為是「黑箱」，這給理解和解釋其決策過程帶來困難，從而影響信任和採用；AI技術還面臨合規風險，組織需要確保其AI應用符合相關法律法規，避免違規使用和潛在的法律責任；AI治理還涉及到倫理

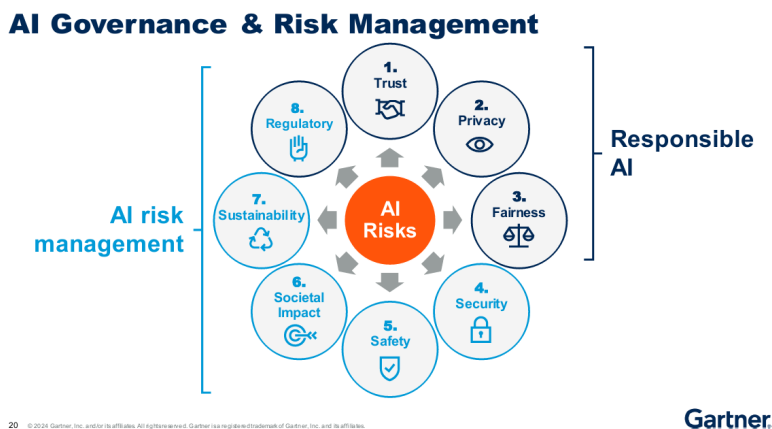


圖 16-AI 治理與風險管理

風險，包括防止偏見和確保公平性，這需要在設計和部署AI系統時就加以考慮，以避免產生歧視性或不公平的結果。(如圖16)

### 七、 人工智慧素養培育：

人工智慧素養培育涉及多方面的知識和技能，員工需要理解AI的基本概念和原理，包括機器學習、深度學習和自然語言處理等技術的基礎知識，熟悉AI的應用場景，了解AI如何在不同領域產生影響和價值。員工需要能夠評估和分析AI系統的性能和可靠性，並了解其可能的風險和限制。實踐技能也是AI素養培育的一部分，這包括如何使用AI工具和平台進行資料分析和模型開發。最後組織需要建立一個支持AI學習和創新的環境，鼓勵員

### AI Engineering Requires Multidisciplinary Collaboration

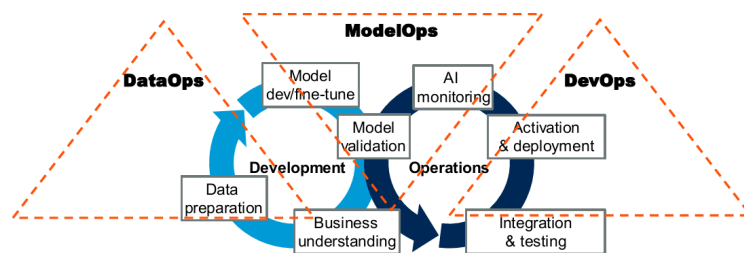


圖 17-AI 實作技能與技術準備

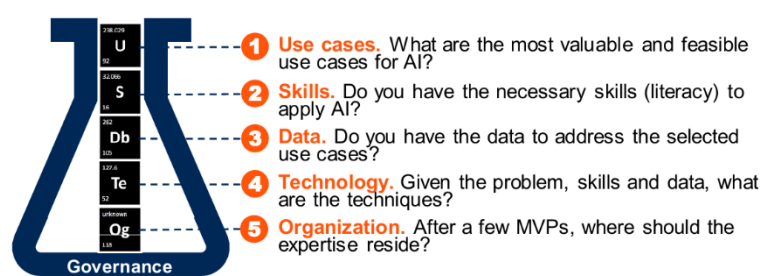
工不斷學習和實踐AI技術，並透過跨部門、多學科的合作和交流來促進AI知識的傳播和應用(如圖17)。

## 八、 總結

講者在最後補充組織在實作AI時可採用的簡易公式如下(如圖18)：

1. 定義目標：需要確定AI技術要解決的具體問題或使用案例，這有助於集中資源和努力，並確保項目的方向和結果是可測量的。
2. 完備技能：組織需要具備實施AI所需的專業知識和技能，這可能包括內部培訓和外部招募AI專家。
3. 資料準備：高品質的資料是AI成功的基礎，組織需要確保資料的準確性、完整性和一致性，並建立有效的資料治理框架來管理和保護資料。
4. 選擇技術：在確定使用案例、技能完備、資料備齊後，要選擇一個合適且有效率的技术實作。
5. 跨域合作：透過小規模的數個最小可行方案(MVP, minimum viable product)來測試和驗證AI技術的效果，並根據試行結果引進跨領域專業知識，讓AI更符合專業領域使用。

### Summary: The Right Formula for AI Adoption



MVP = minimum viable product  
26 © 2024 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates.

Gartner

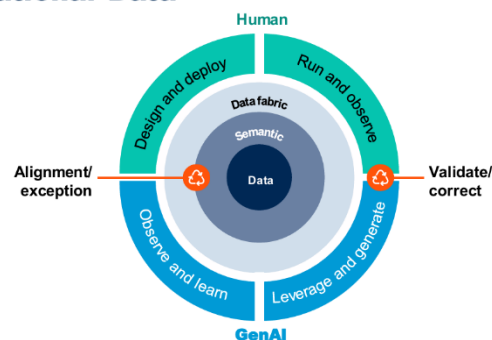
圖 18-AI 技術實作公式

## 柒、資料經緯(Data Fabric)變得真實：建構資料經緯的十個實用步驟

從2000年代的企業資料倉庫(EDW)時代，到2010年代的邏輯資料倉庫(LDW)時代，再到2020年代的主動式詮釋資料管理時代，資料架構經歷了顯著的發展和變革，未來新興資料基礎建設將發展至資料經緯設計，資料經緯設計是一種新興且複雜的資料管理方法，透過使用增強型資料目錄、知識圖譜和主動式詮釋資料來實現靈活和高效的資料整合。顧能公司副總裁分析師Ehtisham Zaidi指出資料經緯設計是未來不可避免的趨勢，可以將資料經緯視為一個智慧化協同引擎，它位於你的資料庫、資料倉庫、資料湖、應用程式（如CRM、ERP）之上，蒐集這些系統的詮釋資料，並為資料管理團隊提供自動化資料整合和轉換的建議，以縮短實現AI和其他應用的時間，預計到2025年(相較於2022年)自動化資料建設將減少40%資料準備、30%設計、30%部署、10%支持和35%品質管理等方面的人力投入，提升400%的資料利用率。

此外，資料經緯在AI應用開發中的重要性也不容忽視，當使用生成式AI進行自然語言查詢時，生成的資料可能不正確或缺乏上下文，而資料經緯作為一個守門人，會提供語義豐富的資料給生成式AI，以確保查詢結果更具上下文且不會產生錯誤的結果。(如圖19)

### The Why: Data Fabrics Enable Better GenAI for Organizational Data



Source: Gartner  
10 © 2024 Gartner, Inc. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates.

Gartner

圖 19-資料經緯在生成式 AI 中的定位

以下將詳細介紹實現資料經緯設計的十個具體步驟，並提供選擇適當技術和工具的建議：

一、首先使用增強型資料目錄蒐集現有的詮釋資料：在設計資料經緯時，第

一步是蒐集和管理現有的詮釋資料，這對於了解現有資料環境至關重要。這可以透過使用增強型資料目錄來實現，該目錄使用機器學習技術來自動化下列流程：

1. 協作以確保責任和治理：這有助於確保所有資料的來源、使用和變更都有明確的責任歸屬，並符合公司治理和合規性要求。
2. 管理資訊資產的清單：透過建立和維護全面的資料清單來增進透明度，以幫助組織了解和管理其資料資產。
3. 溝通共享語義：確保資料語義的一致性和共享，這有助於不同部門和應用之間的資料互操作性。

這裡的詮釋資料指的是「被動詮釋資料」，包括設計時的詮釋資料和維運時的詮釋資料，設計時的詮釋資料包括資料模型、結構、設計規範等；而維運時的詮釋資料則涵蓋資料在實際使用過程中的各種日誌和維運統計。

二、**使用知識圖譜**：理解物件之間的關係是必要的，但在使用傳統資料建模和整合方法時，大多數的關係洞察會不見。傳統資料建模假設靜態需求，無法處理關係資料而不增加複雜性。此外，隨著關係數量和連接的增加，性能會下降。資料經緯將多重關係資料呈現為知識圖譜。知識圖譜是互相關聯的資料圖譜，包含關係和語義，允許領域專家（SMEs）使用語義和意義來建模和豐富資料。

三、**理解「主動式詮釋資料」**：主動式詮釋資料是對所有類型詮釋資料（技術、業務、操作和社交）的持續分析，以確定「設計數據」和「操作體驗」之間的對齊和偏差。被動式詮釋資料方法就像溫度計，用於「監控」數據管理和當前使用情況，而主動式詮釋資料方法則像恆溫器，用於「調節」數據管理和使用。

主動式詮釋資料的核心在於即時反饋和調整，這意味著它不僅僅是搜集和顯示數據，而是透過分析和預測來積極影響數據的使用和管理。這種方法可以提高數據操作的靈活性和準確性，幫助組織及時調整策略，以應對不斷變化的數據環境。(如圖20)

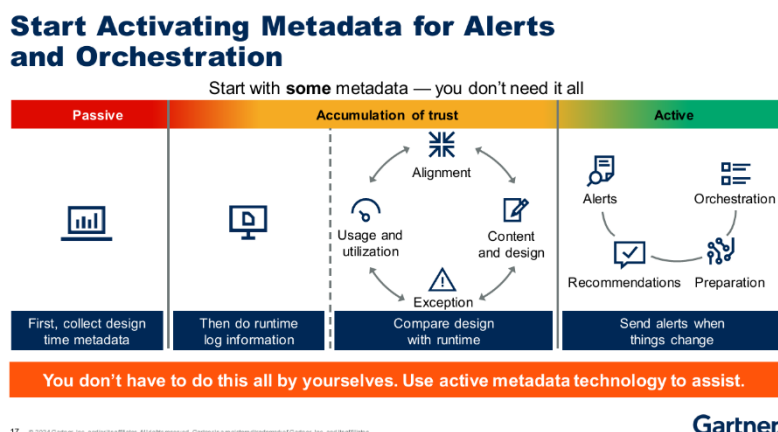


圖 20-實作主動式詮釋資料

此外主動式詮釋資料技術還能自動觸發警報和協調操作，蒐集設計時的詮釋資料，然後是維運時的日誌信息，並將設計與維運進行比較，當情況發生變化時發出警報，這些警報可以提供建議，進行準備和協調，確保數據管理和使用的對齊和例外處理。

四、**使用基於主動式詮釋資料的洞察**：基於主動式詮釋資料的洞察能夠為資料經緯提供自動化、參與和洞察三方面的功能。在自動化方面，主動式詮釋資料能夠自動修正結構漂移，確保資料結構的一致性。此外，它還能自動整合「下一步最佳」轉換，推薦最佳的基礎設施和執行引擎選擇，從而提高資料處理的效率。這些功能也為生成式人工智慧（GenAI）的準確性提供了堅實的基礎。在參與方面，主動式詮釋資料允許較少技能的整合者和領域專家透過語意和搜索找到感興趣的資料。這使得資料的訪問和使用變得更加方便，促進了不同部門和人員之間的協作。在洞察方面，主動式詮釋資料技術能夠進行異常和異常值檢測，突顯敏感屬性以符合GDPR等規範。透過可觀察性，主動式詮釋資料還能提供深度洞察，幫助組織了解資料的運用狀況並進行即時調整。

五、**確保資料經緯能夠通過多種資料傳遞方式進行資料交付**：資料經緯的目

標是結合不同的資料整合方式，並允許業務使用者在最少的IT支持下自行準備資料。這意味著要使用多種資料傳遞方式，以滿足不同的業務需求和使用情境。這些資料傳遞方式包括：

1. 訊息傳遞 (Messaging)：即時傳遞資料，適用於需要低延遲的應用。
2. ETL/ELT：傳統的資料提取、轉換和載入 (Extract, Transform, Load / Extract, Load, Transform) 過程，用於結構化資料的批次處理。
3. 資料複製 (Data Replication)：在不同系統之間複製資料，以確保資料的一致性和高可用性。
4. 資料虛擬化 (Data Virtualization)：允許使用者即時查詢和分析分佈在不同來源的資料，而無需實際移動資料。
5. 串流 (Streaming)：實時處理和分析資料流，適用於需要即時反應的應用場景。

透過這些方式，資料經緯能夠靈活地適應不同的業務需求，提供高效且即時的資料交付服務。這樣的設計不僅提高了資料的可用性，也減少了IT部門的負擔，讓業務使用者能夠更加自主地進行資料分析和決策。



六、以自助服務方式交付整合資料：資料經緯需要能夠以自助服務方式交付整合資料，但同時必須監控操作與治理，資料經緯應提供可重用的模型，

### Step 6: Deliver Integrated Data for Self-Service — But Monitor for Operationalization & Governance

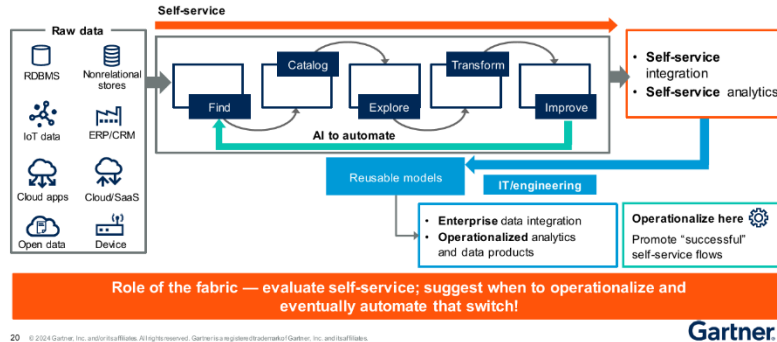


圖 21-自助式資料交付流程

例如企業資料整合、操作化分析和提供資料產品，這些模型可以讓IT和工程部門在操作的過程中推動成功的自助服務流程。(如圖21)

人工智慧的自動化能夠幫助優化自助服務的整合和分析過程，自助服務的整合資料和分析可以涵蓋雲端應用、開放資料裝置、雲端平台等。透過資料經緯的幫助，使用者可以自行發現、目錄化、探索、轉換和改進資料，而不需過多依賴IT支援。

資料經緯的角色不僅是評估自助服務的效能，還包括在適當時機建議操作流程，並最終自動化這一轉換過程。這樣可以確保資料經緯在支持自助服務的同時，保持操作化和治理的要求，從而實現高效的資料管理和利用。

七、利用DataOps簡化資料交付：DataOps 是一種資料操作方法論，結合了多種技術和實踐，包括編排、可觀察性、測試和部署自動化，以及環境配置，這些技術和實踐的結合旨在提升資料工程團隊的生產力和資料交付的效率。

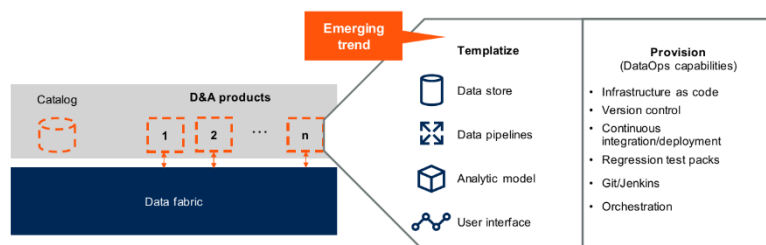
1. 編排：DataOps 提供跨資料管理技術的編排能力，使不同資料源和工具之間的工作流程得以自動化和協調。例如，可以利用 DataOps 工具來自動化資料管道的建立和維護，從而減少手動介入並提高效率。

2. 可觀察性：DataOps 強調資料可觀察性，這包括資料的譜系追蹤、準確性檢查、新鮮度監控和管道中斷的監測。透過這些措施，可以及時發現和解決資料問題，確保資料的可靠性和品質。
3. 測試和部署自動化：DataOps 包含自動化的測試和部署過程，這使得新資料源和資料變更可以快速且可靠地引入系統。例如，使用 DataOps 工具可以實現連續整合和部署（CI/CD），從而加快資料產品的更新速度並減少錯誤。
4. 環境配置：DataOps 還涵蓋了環境配置管理，確保資料處理環境的一致性和可重現性。這可以包括基礎設施即代碼（Infrastructure as Code, IaC）的實踐，使得整個資料基礎設施的設定和部署都可以自動化完成。

根據Gartner的預測，到2025年使用DataOps工具的工程團隊將比未使用這些工具的團隊高出10倍生產力，透過實施DataOps，組織可以大大提高資料處理和交付的效率，從而更快地獲得洞察並提高競爭力。

八、將整合資料作為「資料產品」進行供應和管理：在實作資料經緯的過程

### Step 8: Provision and Manage Integrated Data as “Data Products”



- Data products are modular, loosely coupled and self-contained data apps that enable independent build/deployment cycles. Usually controlled by SMEs at the domain level.
- Drive targeted consumer experiences by templating products and provisioning product instances at scale.

22 © 2024 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates.

Gartner

圖 22- 資料產品管理

中，特別強調將整合資料作為「資料產品」來進行供應和管理，這種方法的目的是促進資料的模組化、鬆散耦合和自我封閉的應用，從而實現獨立的建構和部署週期，通常由領域專家（SMEs）控制。（如圖22）

資料產品的供應需要強大的DataOps能力，包括基礎設施即代碼（Infrastructure as Code）、版本控制、持續整合與部署、迴歸測試

包等工具，這些工具可以使用Git、Jenkins等進行協作和編排。

1. 模組化：資料產品應該是模組化的，這意味著它們可以獨立於其他產品進行開發和部署。
2. 鬆散耦合：每個資料產品應該與其他產品保持鬆散耦合，以便於靈活的更新和維護。
3. 自我封閉：資料產品應該自成一體，能夠獨立運行和管理，這樣可以減少對其他系統的依賴。

資料產品的實施還需要資料管道和分析模型的支援，這些都是資料產品成功的關鍵組成部分，資料管道負責資料的收集、清洗、轉換和下載，而分析模型則用於從資料中提取洞察和價值。

九、**自建或是購買**：在考慮自建或購買資料經緯的過程中，有多項關鍵因素需要評估，首先是成本問題。自建資料經緯可能需要大量的初期投資，包括人力、時間和技術資源；購買現成的解決方案可能在初期成本上更具優勢，但可能會產生持續的授權費和服務費。其次是控制與靈活性，自建解決方案能夠提供更高的控制權和靈活性，允許組織根據自身需求進行定制和調整，包含但不限於DataOps技術、自助服務資料工具、整合資料工具、資料虛擬化工具、知識圖譜和語意、詮釋資料管理工具、資料目錄工具、網格型資料交付技術等。而購買的解決方案則可能在某些方面受到限制，無法完全滿足組織的客製化需求。

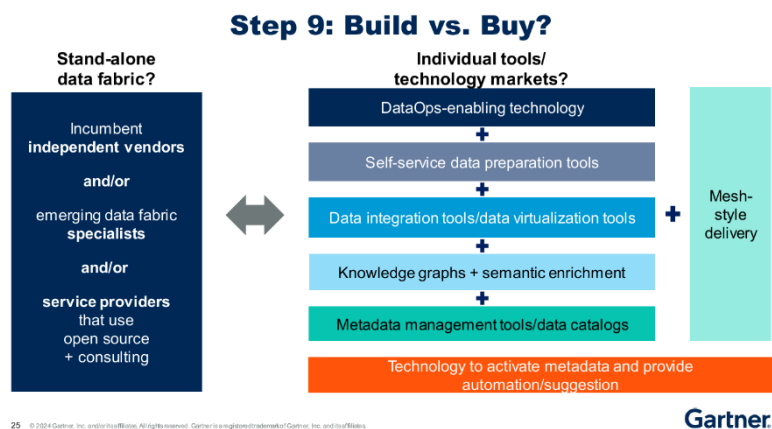


圖 23-自建或購買資料經緯技術

第三是實施速度與時機，購買現成的解決方案通常能夠更快地部署使用，

縮短實施時間，有助於組織迅速回應需求；自建解決方案則可能需要較長的開發和測試週期。(如圖23)

最後組織需要考慮技術的可持續性及未來的擴展性，自建解決方案需要考慮到未來的擴展需求和技術更新，而購買的解決方案則需要評估供應商的技術支持、更新能力及授權費用等。

十、**代表性供應商**：在實作資料經緯時，選擇合適的技術和服務供應商是至關重要的，代表性供應商提供各種解決方案和服務，以協助資料經緯的構建和管理。

本場次會議探討設計和實現資料經緯是現代資料管理中的一項關鍵任務，透過理解和利用主動式詮釋資料、知識圖譜和DataOps，組織可以實現更高效的資料整合和準備，從而支持自助服務和營運治理，這過程需要時間和資源累積，起步越早將在長期發展中帶來顯著的回報。資料經緯改善了資料管理的自動化及多方協作，還為實現組織業務目標與AI發展和應用提供了堅實的資料基礎，透過遵循前述的10項步驟，組織可以逐步構建和優化資料經緯設計，從而在不斷變化的資料環境中保持競爭優勢。

## 捌、心得

一、**先進國家的資料分析與資料治理發展趨勢值得持續關注**：在現今資料驅動的時代，先進國家的資料分析與資料治理發展趨勢展現出強大的潛力和前景，透過建立全面的資料經緯技術和領域資料擴充，組織能夠更加精確地追蹤資料來源、流動路徑和變更歷史，這對於提升資料品質、確保合規性和支持決策制定具有重要意義。先進國家的資料分析與資料治理正邁向更加自動化、智慧化及結構化的方向發展，這些趨勢不僅提升了資料管理的效率和準確性，還為組織在變化激烈的環境需求中提供了強大的資料支持。隨著技術的不斷進步，這些發展趨勢必將引領全球資料治理未來新的方向，值得持續關注和學習。

二、**特別關注 AI 蓬勃與迅速的發展**：這次研討會也發現 AI 技術的蓬勃發展正以驚人的速度改變著各個行業，這一趨勢值得我們特別關注。AI 正在重新定義資料分析和資料治理的方式，資料治理不再僅僅依賴於傳統的手動/半自動管理和規則制定，更多的組織開始採用 AI/ML 驅動的增強技術，自動化資料清理、分類、監控和轉換過程，提高了資料品質和管理效率，並且持續監控資料品質，還顯著降低了人工干預所帶來的錯誤風險。大型語言模型（LLM）和自然語言處理（NLP）技術的進步，使得我們能夠更好地理解 and 處理非結構化資料，這些技術能夠從大量文本資料中提取出有價值的資訊，自動生成資料註解和分類，並且能夠幫助組織創建和應用複雜的資料治理規則。隨著 AI 技術的不斷進步，可以預見，未來的資料治理將變得更加智慧化、自動化和高效化，這一發展趨勢值得我們持續關注和投入，從中獲得啟發和動力。

三、**國際研討會辦理不易值得學習**：這次參加顧能公司辦理的「2024 年資料與分析峰會」也讓我感受到辦理國際研討會不容易，研討會的成功舉辦需要精心的策劃和組織，從議程設計、講者邀請到與會者管理，要考慮到不同領域和背景的參與者，確保內容的多樣性和前瞻性，能夠引起廣泛的興趣和討論。來自不同國家和地區的參與者帶來了多樣的文化和觀點，需要有靈活的應對策略，才能有效促進交流和合作，避免文化差異帶來的誤解和衝突。國際研討會是一個絕佳的學習和交流平台，拓展視野，激發創新思維，這次的經驗值得我們未來舉辦相同規格的國際研討會參考，從而與國際合作共同推動全球研究和技術進步。

## 玖、建議

一、**打造我國資料基礎建設**：資料經緯技術作為新興資料管理的核心框架，涉及到資料的生成、收集、處理、存儲和共享，需要建立統一的資料標準和規範，確保資料在不同系統和平台間的互通性和兼容性，有助於提升資料品質，還能促進資料的有效利用和價值挖掘；同時建立資料匯流

機制，打破資料孤島和部門壁壘，透過建設統一的虛擬層詮釋資料整合來實現，促進不同資料源的整合和協同利用，激勵各方積極參與資料匯流和協同創新。

二、**發展我國AI-Ready Data評估準則**：隨著人工智慧（AI）技術的迅速發展，建立一套簡易通用與多套領域專精的AI訓練用資料評估準則變得至關重要，該準則可以輔助AI技術發展，提供可信任的資料來源，提高AI模型訓練的效率和準確性。高品質的資料是AI成功應用的基礎，制定和推行AI-Ready Data評估準則將為我國AI技術的發展和應用提供堅實的基礎，同時提升AI-Ready素養訓練，確保人員有足夠知能跟上技術發展的步伐，從而推動各行各業的數據驅動創新。

三、**研究與導入接軌國際之隱私強化技術**：資料隱私保護在資料技術與AI技術發展的同時受到國際關注，為確保我國在發展相關技術時，亦符合相關規範與安全性，並能與國際接軌，建議研究與導入國際先進的隱私強化技術。例如差分隱私技術，能夠在統計數據的基礎上添加隨機噪音，防止個人資料的洩露；如聯合學習技術允許多方在不共享原始資料的前提下共同訓練AI模型，有效保護資料隱私，同時實現資料的協同利用；同態加密技術也是目前國際上新興發展的隱私保護手段，能夠直接對已加密的資料進行運算，且可產生與運算未加密資料一致的結果。同時培育我國隱私強化技術專業人才，可有效提升資料安全和隱私保護水平，為資料驅動的創新和發展提供堅實的保障。

# 壹拾、 附錄

## 一、會議議程

Data & Analytics Summit	
Printed on : Jun 13 by C.L.	
<b>08:30 AM Tuesday, May 21</b>	
<b>Registration and Information</b>	08:30 AM-06:50 PM Registration Desk, B1F, Grand Nikko Tokyo Daiba
Registration and information hours for the day	
<b>09:30 AM Tuesday, May 21</b>	
<b>Gartner Opening Keynote: Generating Value Together: From Fundamentals to AI Readiness to Collective Intelligence</b>	09:30 AM-10:15 AM
Melody Chien, Gareth Herschel	Hall 1 - 2, B1F, Grand Nikko Tokyo Daiba
Welcome to the era of collective intelligence. Generative AI is an inflection point for data and analytics. Chances are it is proliferating throughout your organization, so expectations of you have never been higher. New eras create new leaders. For all the progress we have made, the opportunities ahead of us ...	
<b>10:30 AM Tuesday, May 21</b>	
<b>Guest Keynote: The Power of Design to Unleash Innovation</b>	10:30 AM-11:15 AM
Masatoshi Toyama	Hall 1 - 2, B1F, Grand Nikko Tokyo Daiba
In 2018, the Ministry of Economy, Trade and Industry and the Japan Patent Office issued the Design Management Declaration, manifesting that design is not just about attractive appearance, but can be used for innovation and branding by changing perspectives to create products and services based on the challenges users face. This ...	
<b>12:30 PM Tuesday, May 21</b>	
<b>Top Trends in Data and Analytics 2024</b>	12:30 PM-01:00 PM
Gareth Herschel	Hall 1, B1F, Grand Nikko Tokyo Daiba
AI is not the only thing driving rapid change in data and analytics (D&A). There is a rapid evolution in how and where analysis can be deployed, from the next generation of augmented UXs and beyond; composing and operationalizing D&A at scale; and engineering decision intelligence to enabling D&A at ...	
<b>01:15 PM Tuesday, May 21</b>	
<b>Qlik: Eliminating the AI Trust Gap — 10 Trends to Build Data Confidence for Analytics</b>	01:15 PM-01:45 PM
Dan Sommer	Hall 1, B1F, Grand Nikko Tokyo Daiba
GenAI promises to usher in a new era of productivity & prosperity. However, recent hysteria about data bias, privacy, & unintended output leaves many doubting if AI is worth the risk. This session will highlight customer lessons learned and observations from Qlik, who recently completed the acquisition of data integration ...	
<b>02:00 PM Tuesday, May 21</b>	
<b>Magic Quadrant for Cloud Database Management Systems — Implications for APJ Organizations</b>	02:00 PM-02:30 PM
Xingyu Gu	Hall 3, B1F, Grand Nikko Tokyo Daiba
The future of database management system is cloud - This is already a fact in North America. However, many organizations in APJ are still at the early stage of their cloud journey. Unaligned regulations, uneven digital maturity, and different regional vendor ecosystems complicate DBMS strategy. Stay with Oracle? Use OSS ...	
<b>04:15 PM Tuesday, May 21</b>	
<b>How to Create an Optimal Data &amp; Analytics Organizational Model</b>	04:15 PM-04:45 PM
Michael Gabbard	Hall 1, B1F, Grand Nikko Tokyo Daiba
To position their teams as a must-have discipline that can fulfill cross-functional use cases, data and analytics leaders must optimize their organizational model. The optimal organizational model is a hybrid one that balances enterprise-wide capabilities focused on enablement with decentralized needs focused on outcomes. In this session you will learn ...	
<b>05:05 PM Tuesday, May 21</b>	
<b>Franchise Regional Chief Data Office With a Composable Global Deployment</b>	05:05 PM-05:35 PM
Julian Sun	Hall 3, B1F, Grand Nikko Tokyo Daiba
Given the challenges posed by geopolitical and tech-nationalism issues, multinational corporations face difficulties establishing a global data and analytics architecture. To overcome these obstacles, D&A leaders can adopt a franchise model to develop a regional chief data office with a composable global deployment, aligning with the regional team with a ...	
<b>09:00 AM Wednesday, May 22</b>	
<b>Registration and Information</b>	09:00 AM-05:35 PM Registration Desk, B1F, Grand Nikko Tokyo Daiba
Registration and information hours for the day	
<b>09:30 AM Wednesday, May 22</b>	
<b>Panel Discussion: Dear President: Please Let Us Know About Management &amp; Will D&amp;A Be Helpful for Your Business Decisions?</b>	09:30 AM-10:15 AM
Tatsuya Ichishi, Taiji Abe, Yoshi Matsumoto	Hall 1 - 2, B1F, Grand Nikko Tokyo Daiba
We see and hear the term "data-driven management." Gartner clients also ask us what kind of dashboards they should provide to management and how to get them interested in using data. What is management and what data is useful for management? Current Gartner employees with experience in corporate management will ...	
<b>10:30 AM Wednesday, May 22</b>	
<b>Quest Keynote: Cognitive Science of Communication and Mutual Understanding: Why Is Your Intention Not Conveyed to the Others?</b>	10:30 AM-11:15 AM
Mutsumi Imai	Hall 1 - 2, B1F, Grand Nikko Tokyo Daiba
Miscommunication often arises in everyday situations, be it at work or at home, in which the listener responded quite differently from what you had expected. The cause of the miscommunication has been generally attributed either poor delivery of the information on the part of the speaker, or lack of attention ...	
<b>12:30 PM Wednesday, May 22</b>	
<b>Top Trends in AI</b>	12:30 PM-01:00 PM
Ben Yan	Hall 3, B1F, Grand Nikko Tokyo Daiba
How is generative AI evolving? What are customers doing? How is value being created? In this session, we will explore the most important trends in AI and Generative	

圖 24-會議議程 1

AI and how they may be relevant to your organization.	
<b>02:00 PM Wednesday, May 22</b>	
<b>Emerging Practices for Decision Intelligence: The Next Leap for Data, Analytics and AI</b>	02:00 PM-02:30 PM
Gareth Herschel	Hall 1, B1F, Grand Nikko Tokyo Daiba
Decision intelligence is quickly gaining traction as the discipline that aims to bridge the gap between analytics and decision making. But how to make decision intelligence a reality? Where to start, following which approach? How to go about the reengineering and modeling of decisions? Which decision intelligence platforms are available? ...	
<b>04:15 PM Wednesday, May 22</b>	
<b>Top 10 Trends in Data Integration and Engineering: Invest to Modernize Your Data Management Foundations</b>	04:15 PM-04:45 PM
Ehisham Zaidi	Hall 3, B1F, Grand Nikko Tokyo Daiba
Data Integration is evergreen — it was always foundational to any resilient data management strategy. This session gives guidance on: 1. Technology Trends: What are the trends guiding the data integration technology? 2. Process and Best Practices: How should you set up a comprehensive data integration practice? 3. Market Trends and Technology Evaluation: ...	
<b>05:05 PM Wednesday, May 22</b>	
<b>Augmented Data Quality Solutions: Revolution of Data Quality Technologies and Practices</b>	05:05 PM-05:35 PM
Melody Chien	Hall 3, B1F, Grand Nikko Tokyo Daiba
The data quality technologies have been going through significant evolution in recent years. The traditional data quality tools are not sufficient to address increasingly complex data quality challenges. Augmented data quality solutions driven by AI/ML, metadata, and knowledge graph bring completely new practices to data quality processes. D&A leaders should ...	
<b>09:00 AM Thursday, May 23</b>	
<b>Registration and Information</b>	09:00 AM-04:00 PM
	Registration Desk, B1F, Grand Nikko Tokyo Daiba
Registration and information hours for the day	
<b>10:30 AM Thursday, May 23</b>	
<b>The State of Data Observability Technology - Detection, Recommendation and Prevention</b>	10:30 AM-11:00 AM
Melody Chien	Hall 4, B1F, Grand Nikko Tokyo Daiba
Data observability technology gives organizations broad visibility over their data, data pipeline, and data landscape. By continuously monitoring, tracking, alerting, and troubleshooting problems, data observability technology aims to detect problems, recommend solutions, and therefore reduce and prevent data errors or downtime. D&A leaders should know what it can do, ...	
<b>12:15 PM Thursday, May 23</b>	
<b>What Everyone in D&amp;A Needs to Know About (Generative) AI: The Foundations</b>	12:15 PM-12:45 PM
Ben Yan	Hall 3, B1F, Grand Nikko Tokyo Daiba
What is AI? What is generative AI, and other AI flavors and techniques? How to start with AI? Which data, skills and tools are typically needed? What are the best practices and pitfalls? Where does AI create value, and in which applications and use cases? This session will address these ...	
<b>01:45 PM Thursday, May 23</b>	
<b>How to Upskill and Reskill Your Data, Analytics and AI Employees</b>	01:45 PM-02:15 PM
Michael Gabbard	Hall 3, B1F, Grand Nikko Tokyo Daiba
Data and analytics leaders struggle to attract, sustain and retain talent for their data, analytics and AI teams. This research shows how to build a formal reskilling and upskilling strategy to support data and analytics initiatives.	
<b>03:15 PM Thursday, May 23</b>	
<b>DataRobot: Confidently Deploy and Govern GenAI Solutions</b>	03:15 PM-03:45 PM
Michael Schmidt	Hall 1, B1F, Grand Nikko Tokyo Daiba
More information is coming soon.	
<b>04:00 PM Thursday, May 23</b>	
<b>Data Fabrics Made Real: Ten Practical Steps to Build Your Data Fabric Design</b>	04:00 PM-04:30 PM
Ehisham Zaidi	Hall 1, B1F, Grand Nikko Tokyo Daiba
Organizations looking to invest in a future-proof data management architecture must invest in the data fabric design. But most data and analytics leaders don't know where to begin. This session will help CDACs, data management leaders and their teams to: 1. Understand what the data fabric design is and what it means ...	

圖 25-會議議程 2



## 二、會議場景照片紀錄



圖 26-抵達會議入口



圖 27-完成報到



圖 28-會場寬廣附有休息座位



圖 29-參與人數眾多

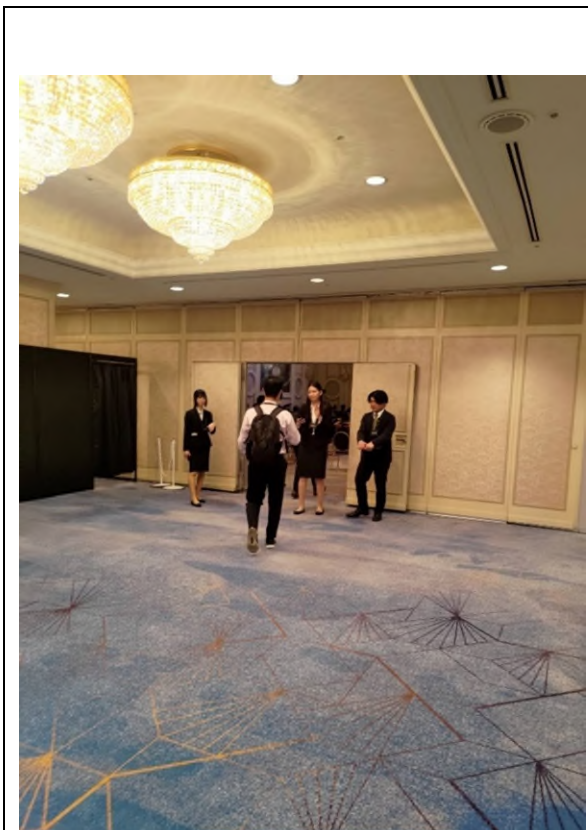


圖 30-每場次會議入場登記



圖 31-會議及講者介紹



圖 32-大會場(僅拍攝部分前場)



圖 33-大會場(僅拍攝部分後場)

Welcome to the Exhibit Showcase

Explore solutions. Find your perfect fit.

**出展社 (五十音順)**

AVILEN	506
アバント	407
アルテアエン지니어リング	502
アルテリックス・ジャパン	103
Yellowfin Japan	403
インサイトテクノロジー	201
INSIGHT LAB	510
インターシステムズジャパン	504
ウイングアーク1st	406
内田洋行	404
エア	505
NTTコム オンライン・マーケティング・ソリューション	304
F P T ジャパンホールディングス	405
キーエンス	303
Cloudera	202
クリックテック・ジャパン	101
KX Systems	408
SAS Institute Japan	102
CData Software Japan	507
セゾンテクノロジー	306
ダウ・ジョーンズ・ジャパン	508
DataRobot Japan	302
Denodo Technologies	203
Dataiku Japan	402
ドーモ	509
日本情報通信	503
日本 HP	305
日本テラデータ	301
ピーブルドット (旧データミックス)	401
フライウィール	401

Theater

Gartner Showcase

Hall 1

Hall 2

Hall 3

圖 34-會議室及展攤位置圖



圖 35-展攤附有交流圓桌

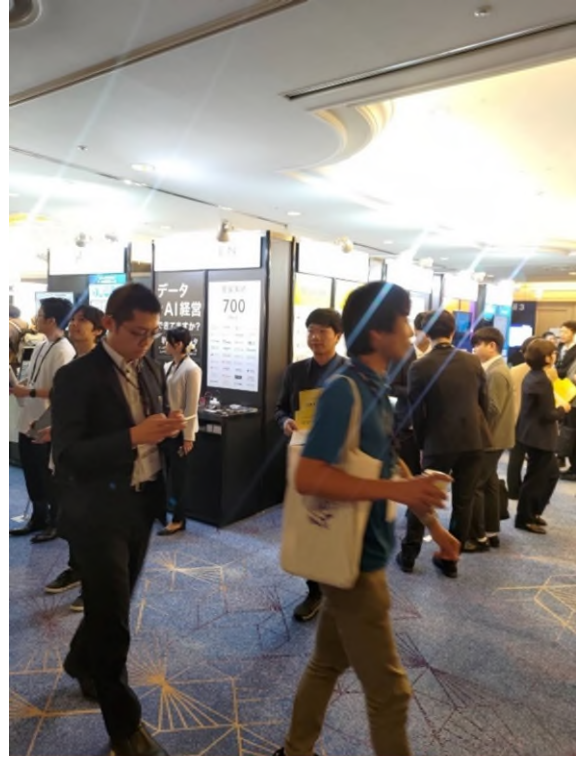


圖 36-展攤參與討論人員眾多



圖 37-會議必備充電區



圖 38-會場外自動車試運行