見解來自

IDC

數據與 AI 影響力報告：

信任的新經濟學

sas

2026 年 9 月

合作夥伴與客戶背書

「隨著生命科學、金融服務等產業的信任缺口逐步縮小，本研究凸顯了在受監管環境中產生實際影響力的關鍵作法。對於高度重視監督的產業而言，這項觀點尤其重要；隨著 AI 自主化程度不斷提升，強化信任與治理將成為核心關鍵。」

Mona Chadha
Amazon Web Services AI 與戰略合作總監

「治理不應被視作阻礙性、約束性的措施；治理是保障 AI 落地應用可信、安全且可規模化的體系架構。妥善落實 AI 治理的組織，將在同業競爭中建構差異化優勢。」

Karen Zieker
IQVIA 副總裁、生物統計與統計程式設計全球負責人

「能從 AI 中創造最大價值的組織，未必是擁有最多資料的組織，而是善用資料做出更優商業決策的組織。這些洞見進一步印證了，將資訊轉化為可採取行動的洞見，正是 AI 最能發揮影響力之處，助力組織更快速回應，做出更明智的決策。」

Andy Fletcher
利物浦足球俱樂部技術與數位產品副總裁

「隨著 AI 逐步融入核心業務流程，本報告顯示，信任將成為區分零散試行專案與全企業級轉型的關鍵因素。我們發現，如今建立起完善治理、資料完整性與問責機制的組織，未來最有機會創造可衡量且永續的 AI 價值 —提升效率、賦能創新，並為代理式 AI 的未來奠定基礎。在這樣的未來，AI 不僅能生成創意與洞見，還也能大規模協調決策、行動與成果。」

Christian Lau
洛杉磯足球俱樂部兼 BMO 體育場首席技術官

「本研究清楚顯示，隨著組織從生成式 AI (GenAI) 邁向代理式 AI 系統，健全的治理能夠加快實現價值的速度。該研究針對信任缺口、Copilot與量子技術早期採用的洞見正逢其時，清楚指出要如何以負責任的方式擴大進階 AI 的應用。」

Jeff Markowitz
Microsoft 軟體與數位平台事業部總經理

「各組織正透過強化治理、明確驗證標準、更嚴謹的正式環境使用規範，強化 AI的實務作法。這些關於信任的研究結果顯示，市場正從缺乏架構的實驗階段走向成熟，並及時說明 AI 專案如何發展，協助釐清目前真正影響 AI 成效的條件。」

Dez Tsai
TD SYNNEX全球 AI 與資料高級總監

前言 — Bryan Harris 的來信

「這不是技術問題，而是流程問題。」

「這不是精確度的問題，而是採用問題。」

「問題不在於模型，而是人。」

聽起來很熟悉？我時常聽到這類說法。

但現實情況是：生成式 AI 確實存在準確度問題。事實上，已有大量文獻記錄，即使是最先進的智慧代理，在處理複雜任務時，錯誤率可能介於 25% 到 50% 之間。而這些錯誤會在多步驟工作流程當中持續累積放大。倘若智慧代理每個步驟僅能達到 85% 的準確率，能夠完整成功執行整套工作流程的機率其實相當低。

對於關鍵、高風險的商業決策來說，這樣的表現遠遠無法滿足需求。

那麼，為何各組織仍爭相部署 AI？

因為當系統運作順利時，成果相當亮眼。模型產出具說服力的回答。智慧代理完成錯綜複雜的任務。過去耗費數個月的流程，如今僅需數小時即可完成。

這類成果帶來熱忱、機會與緊迫感。也讓我們看見，將 AI 嵌入企業營運之後所能實現的可能性。但同時也容易使人只看見 AI 的潛力，卻忽視其本身的侷限。

信任鴻溝便是由此而生。

首要挑戰，在於判斷何時適合使用智慧代理、何時不適合。許多組織將智慧代理運用在僅需要精確、具確定性的標準自動化任務。這麼做反而為了達成相同目標，衍生不必要的複雜度與成本。

究竟什麼場景適合導入智慧代理？

就目前而言，當智慧代理運用在工作流程繁雜、以目標為導向的任務時，能夠發揮最大價值。這類場景之下，仍然需要人類的專業知識來評估輸出結果，並處理非確定性行為帶來的錯誤。可以透過「人在迴路 (human-in-the-loop)」審查機制，或是將領域專業知識內嵌至智慧代理工作流程，實現這層把關。

歸根結底，信任建立在結果的準確度與可重複性之上。無論組織採用智慧代理工作流程，或是傳統標準自動化；只要輸出準確、一致、可重複，信任鴻溝就會消失。

這就是我們追求的目標。想要達成，需要高階領導者持續投入技術策略，同時確保人始終位於監督與治理的核心。做到這一點的組織，將能夠跨越信任障礙、締造突破性成果，藉 AI 取得市場競爭優勢。

唯有到那時，我們才能夠篤定地說：「這不是技術問題。」

Bryan Harris

SAS 執行副總裁暨首席技術長

目錄

前言 — Bryan Harris 的來信 .....

重點摘要 1

方法論 3

自主程度提升，風險同步升高 4

Agentic AI 快速演進，信任基礎仍待建立 5

理解信任與可信賴性的價值 6

可信賴 AI 的五大要素 7

人工介入的直覺反應 9

資料成熟度仍是 AI 發展的瓶頸 10

AI 成熟度提升，也讓資料主權問題更加受到重視 12

成功之道：向領先企業學習 13

以可信賴 AI 領先組織的方式營運 14

讓 AI 領先者脫穎而出的五項關鍵做法 15

AI 可信賴性如何影響 ROI？ 16

阻礙 AI 創造價值的三大錯誤 17

不同產業如何塑造 AI 領先優勢 18

銀行業領先者的經驗 19

保險業領先組織的經驗啟示 23

生命科學領域領先組織的經驗 26

公共部門領先組織的經驗 29

中小型機構 32

所有企業都應掌握的關鍵重點 34

從信任走向優勢優勢：未來之路 35

名詞解釋 36

重點摘要

AI 自主執行的程度越高，就越需要贏得信任。而目前，信任的發展並未跟上步伐。

幾年前，財務長將採購決策交由人工智慧 (AI) 處理還只是科幻設想。如今，這已是尋常操作。生成式 AI (GenAI) 已不再僅為決策提供參考。代理式 AI能夠直接執行決策，各組織領袖正站在這一轉捩點上。

其落地成果已切實可見：信用風險評估、保險定價、政府福利發放處理、將患者媒合至臨床試驗、撰寫程式碼、協調工作流程，以及做出過去必須仰賴人類判斷的各類決策。隨著 AI 自主程度提升，組織面臨的監管審查、營運風險和聲譽損害風險也隨之增加。

這使得信任成為一項營運層面的議題，而非哲學層面的空談。信任並非源自技術本身。而是源自組織對所部署系統的設計、治理與驗證方式。且其回報切實可衡量：將可信實踐融入 AI 體系的組織，**實現高投資報酬率 (ROI) 的機率是其他組織的 15 倍。**

其中的道理十分清晰。如今，可信賴性決定了 AI 投資是能規模化落地至生產環境，還是會因人工覆寫、決策猶豫與低度使用而陷入停滯。領先組織正持續拉開差距。其餘組織則往往在渾然不覺中逐漸落後。

這份第二版*資料與 AI 影響力報告*基於 **28 個國家的 2699 位決策者**的調查結果，呈現正處於轉捩點的市場現況。信任正面臨壓力。治理正在奮力追趕。同時，一群明確的可信賴 AI 領先組織已逐漸浮現，它們將 AI 轉化為可衡量成果的速度，是市場其他組織難以比擬的。

本報告描繪可信賴 AI 領導力的實際樣貌：資料顯示了什麼、領先組織有哪些不同作法，以及組織現在應聚焦哪些重點，才能為下一階段做好準備。

何為可信賴性？

可信賴性透過五個維度衡量：資料品質與治理、模型治理與監督、可解釋性與公平性、負責任 AI 政策、稽核與問責。

上述維度共同構成可信賴性指數。

得分達到 80 分及以上的組織被歸類為可信賴性領先組織。儘管這類組織在市場中占比相對較小，但它們實現更佳AI 成果與投資報酬率的機率明顯更高。

瞭解更多：請參見「可信賴 AI 的五大要素」(第 7 頁) 和「領先組織經驗啟示」(第 13 頁)。

©2026 SAS and IDC，保留所有權利。CCPA

核心洞察

綜合以上，這些洞察表明 AI 應用正處於關鍵轉捩點。各組織正從生成答案的系統，逐步轉向成為執行行動的系統。隨著 AI 自主程度不斷提升，可信賴性成為組織能否實現價值、管控風險並成功規模化應用 AI 的決定性因素。

AI 一旦開始執行行動，信任度隨即下降。（第 4 頁）

隨著 AI 從產生洞見走向採取行動，僅 66% 的受訪者信任代理式 AI，比生成式 AI 低 10 個百分點。

部署進度快於信任建構進度。（第 4 頁）

約 89% 的受訪者表示，其 AI 代理目前已在決策環節發揮作用，覆蓋從為人工審核決策提供參考到自主執行決策的不同層級，但代理式 AI 的信任度仍較生成式 AI 低 10 個百分點。

可解釋性是最大阻礙。（第 4 頁）

「缺乏解釋」是使用者人工介入 AI 結果的首要原因（占比 34.5%）；在五大可信賴性維度中，可解釋性與公平性的提升幅度最小。

資料成熟度不足正阻礙 AI 發展。（第 10 頁）

僅 17.5% 的組織的資料基礎架構處於最高成熟度的水準。多數組織目前尚無法支撐代理式 AI 所需的資料血緣與驗證能力。

領先優勢展現在行為層面，而非結構層面。（第 15 頁）

領先組織並非由組織架構圖定義。而是由五項行為特徵界定；它們進行治理與可解釋性檢視的機率是其他組織的 55 倍。

信任驅動成果。（第 16 頁）

落實可信賴 AI 實務的組織，取得高投資報酬率的機率，是表現落後組織的 15 倍（62% 對比 4%）。

> AI 可以很可靠，但不代表它一定值得信任。如果它不夠透明，您不清楚它在做什麼、如何運作，就無法摸清黑盒裡的內容。
>
> 美國地方政府 IT 主管

研究方法

本研究旨在衡量人們對 AI 的信任度、AI 自身的可信賴性及AI對企業所帶來的商業影響。我們的發現揭示了更重要的產業趨勢：各組織正邁入 AI 應用的新階段，而這也正在改變人們對 AI 的信任關係。隨著AI系統從單純「產出內容」進一步發展為能夠「採取行動」，企業對 AI 信任與可信賴性的要求也隨之改變。本報告中的許多研究結果 — 從人工介入率、可解釋性挑戰，到 AI 治理投資與投資報酬率 (ROI) 表現 — 均反映出各組織正在適應這項新的現實。

樣本構成

2,699名
決策者

28
國家/地區

5
地區

4
重點產業

重點產業

銀行業

保險業

生命科學產業

公共部門

信任缺口公式

感知信任
人們的主觀認知

—

可信賴性
AI 能夠展現的能力

=

信任缺口
風險與機遇並存之處

圖 1：信任缺口的計算方式、研究方法、樣本規模及調研對象

信任與可信賴性並非同一概念：

- **信任**指的是人們對 AI 抱持多高程度的信心。
- **可信賴性**則反映企業透過哪些實務做法、控管措施與治理機制，讓 AI 值得被信任。

本報告將同時衡量這兩個面向。信任可能因主觀認知起伏波動；相較之下，可信賴性反映的是組織為打造可靠、透明、可問責、安全的 AI 所建構的基礎能力。

©2026 SAS and IDC，保留所有權利。CCPA

自主程度提升，風險同步升高

從輔助決策到自主執行的轉型，正是今年研究的核心主軸。接下來的研究結果 — 從人工介入行為到治理投資 — 均反映出各組織如何因應日益具備自主行動能力的 AI 系統。

全球有76% 的受訪者表示信任生成式 AI。至於能夠自主採取行動的的代理式 AI，僅獲得 66% 受訪者的信任。這 10 個百分點的差距並非偶然。

AI 從「提出建議」進一步走向「採取行動」，能為組織創造的價值就越大，但同時也必須具備更高程度的可信賴性，才能獲得信任。AI 每提升一個能力層級，也會同步提高人們對其信任的標準。

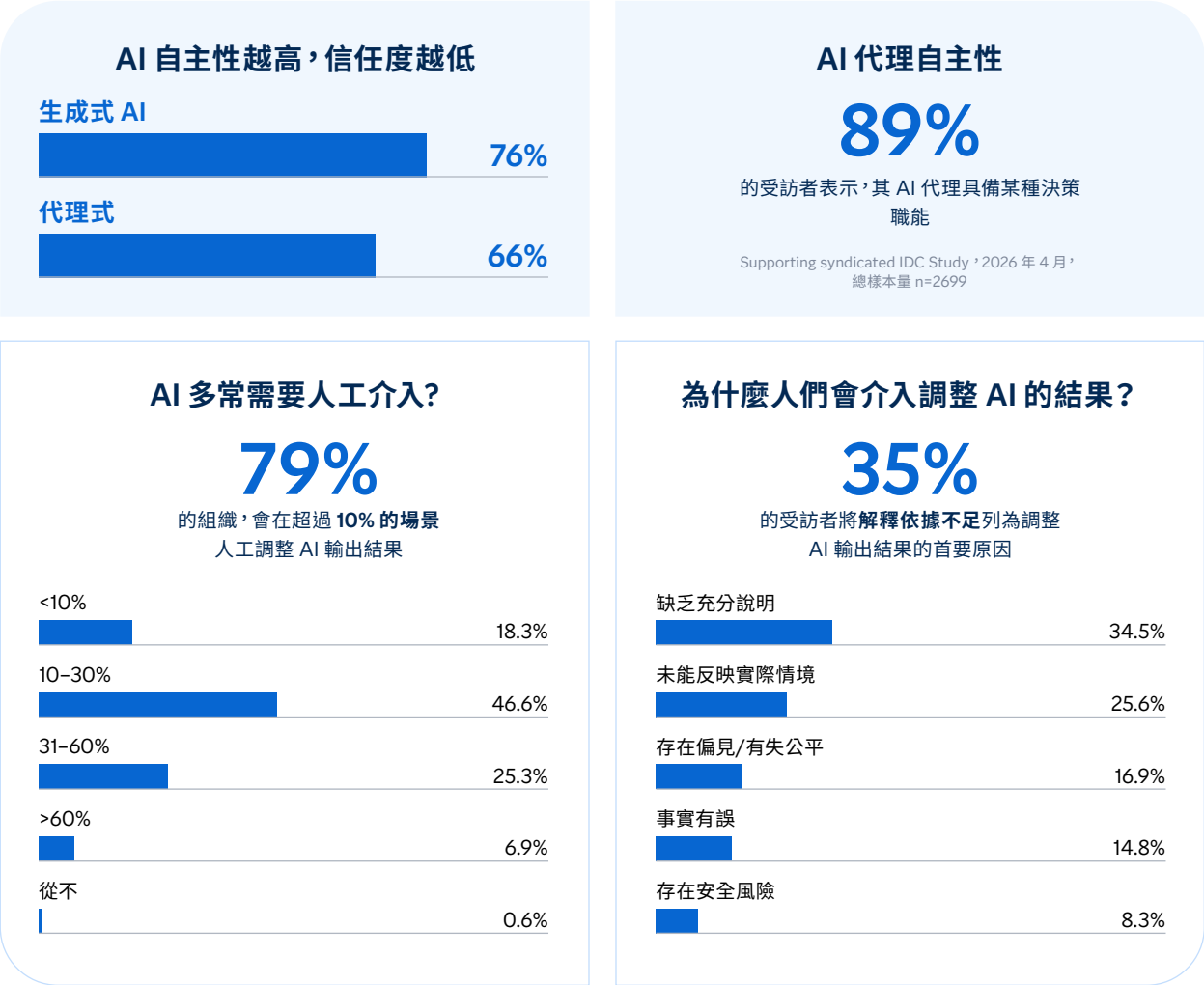


圖 2：代理式 AI、自主性與人工介入之間的關係

Agentic AI 快速演進，信任基礎仍待建立

毋庸置疑，自主型 AI 已經來臨。現在真正的問題是，可信賴性的提升速度是否足以支撐其發展。

各組織希望借助代理式 AI 實現自動化 (56%)、建構競爭優勢 (52%)、推動創新 (51%)、擴大數位勞動力規模 (48%) 以及加快決策速度 (48%)。而且相關部署並非停留在理論階段。

這反映出一種合理的期待：既然系統已具備足以實際部署的能力，也應同樣具備執行行動的能力。問題在於，能確保這項技術值得被信任的可信賴性，並未跟上企業部署代理式 AI 的企圖心與速度。企業對代理式 AI 的佈局願景較去年更加強烈，但數據顯示，支撐這份企圖心的信心並未同步提升。

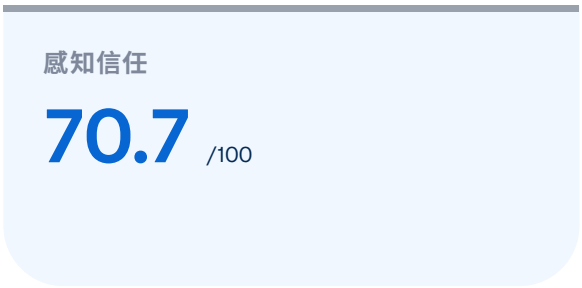
> 管控力度同樣與風險成正比。人工參與程度隨風險高低而增加，但這並不代表我們對技術本身抱持懷疑。如果我們發現人工干預程度過高 - 就說明這套系統本身還需要進一步改善。
>
> 全球銀行資訊安全長 (美國)

©2026 SAS and IDC，保留所有權利。CCPA

理解信任與可信賴性的價值

感知信任

使用者/組織對 AI 的主觀信任程度



可信賴性

用以證明 AI 可靠性、公平性與受監督性的客觀治理實務

圖 3：全球感知信任與可信賴性評分

隨著任何技術逐漸成熟，其實際能力與人們對其能力的認知之間，通常會逐步縮小差距。AI 的導入情況也印證了這項趨勢，而今年的數據更清楚顯示，是誰正在推動這道差距逐步縮小。

可信賴性衡量企業在治理、資料與監督方面所採取、用以證明 AI 可靠性的實務做法；今年該指數從滿分 100 分的得分已從 49.7 分升至 61.3 分。

相較之下，感知信任得分則從 73 分下滑至 70.7 分，這表明儘管 AI 的可信能力不斷提升，但人們對 AI 的整體信任度卻在下降。

二者之間的差距縮小了 13.9 個百分點，從 23.3 分降至 9.4 分。AI 的實際能力終於追趕上了人們的預期，系統正逐步贏得真正的信任 - 而此前這份信任僅建立在使用者的主觀信念之上。這一進展並非自然形成。

各組織已投入實質資源，強化AI治理、管控機制與透明度，讓 AI 真正值得信賴。同時，使用者也透過實際工作任務給予 AI 公平的驗證機會；當結果不符合預期時提出質疑，並在系統改善後再次嘗試使用。

然而，僅靠整體可信賴性得分，無法展現這些細微但重要的變化。可信賴性各項構成要素的進展程度並不一致，因此需要進一步深入剖析，才能瞭解各組織的實際推進情況。

可信賴 AI 的五大要素

本研究透過五項關鍵要素衡量可信賴性：資料品質與治理、模型治理與監督、可解釋性與公平性、負責任 AI 政策、稽核與問責，每個維度均採用 0–100 分的評分標準。

隨著 AI 邁入代理時代，這五大要素的重要性與日俱增。這些要素共同為各組織築牢了必要基礎，使其能夠滿懷信心地部署自主程度日益提升的 AI。

資料品質與治理

可信賴性

59.5 /100

負責任 AI 政策

可信賴性

67.5 /100

模型治理與監督

可信賴性

58.9 /100

稽核與問責

可信賴性

63.4 /100

年增幅度最大

可解釋性與公平性

可信賴性

60.4 /100

年增幅度最小

圖 4 - 可信賴性五大維度

©2026 SAS and IDC，保留所有權利。CCPA

負責任 AI 政策、稽核與問責兩項得分最高，也是年增幅最大的兩項要素。相關投入大量集中在政策與制度層面：制訂規範、記錄管控措施、明確問責歸屬，這些措施依賴組織層面的決心與投入，而非高度複雜的技術基礎架構。多數組織也相應將其列為優先事項。

相較之下，資料品質與治理和模型治理與監督的得分相對偏低，因為這兩個維度需要在資料管道、監控體系與工程實務上進行更長期且持續的投資，儘管如此，這兩個面向仍展現出明顯進步。

可解釋性與公平性是其中的例外，其年增幅為五項要素中最低。單看這項變化或許不顯著，但若與人工介入資料放在一起觀察，就很值得注意：「缺乏充分說明」是使用者介入調整 AI 結果的最常見的原因，有 34.5% 的受訪者提及該原因。也就是說，改善幅度最小的維度，恰恰是使用者最敏感、也最容易採取介入行動的問題。

其餘四項要素大多可透過政策、制度與流程來推進：制訂規範、記錄管控措施、規劃稽核安排、管理資料資產。但可解釋性屬於不同類型的挑戰。它需要模型、資料血緣與使用者介面協同配合，呈現出人類能夠真正理解的推理邏輯；這需要跨越整個技術堆疊進行工程整合，而不只是制定一份治理文件。這也是它進步緩慢的原因。但人工介入的數據也顯示，各組織不能再將該問題延後處理。

這五大要素為建構可信賴 AI 提供了一套清楚的藍圖。真正的挑戰在於，許多組織在這些基礎尚未完全搭建完成時，就已開始部署自主程度日益提升的系統。其後果在本研究中隨處可見 — 人工介入率居高不下、可解釋性面臨挑戰、資料成熟度存在差距，以及不一致的商業成果。

而其中最具代表性的現象，或許可以從一項非常直接的信任行為指標看出：人們多常介入並更改AI 的決策。

> 當我們談及信任，其本質歸根結底要追溯到我們所梳理運用的資料的源頭。因為我們想建立的的所有 AI 系統，其表現好不好，完全取決於我們向系統輸入的資料品質。
>
> 全球銀行資訊安全長（美國）

人工介入的直覺反應

可信賴 AI 的五大要素為，提供了一套理解企業需要建立哪些能力的架構。但我們如何判斷這些投入是否奏效？最清晰的訊號之一便是人的實際行為。當人們信任 AI 系統時，會放手讓其執行。當不信任時，則會進行干預。人工介入率能夠真實衡量出可信賴性建設的成效所在 — 以及尚存差距的領域。

受訪者推多常介入並調整 AI 的決策？79% 的受訪者表示，在超過 10% 的情況下，他們會介入並調整 AI 的建議；僅有 0.6% 的受訪者表示從不介入。

使用者在哪些情況下會介入 AI？

- **超過三分之一 (34.5%) 的受訪者表示，即使 AI 的輸出結果本身合理，只要其推理過程不透明，就會選擇介入。**人工介入最常見的原因是缺乏充分說明以及未能掌握完整情境。即便輸出結果可能正確，人們仍會介入，因為他們無法理解系統得出該結果的過程。這屬於信任與透明度差距，解決之道在於可解釋性，以及支撐其判斷依據的資料血緣。

- **超過四分之一 (25.6%) 的受訪者表示，當 AI 輸出未能充分反映實際情境時，就會選擇介入。**受訪者之所以介入，是因為系統遺漏了某個場景細節、特殊案例，或是系統設計時未納入考量的資訊。這屬於情境理解上的盲點，解決之道在於提供更完整的輸入資料，並讓 AI 與承載相關情境的系統及及資料進行更緊密的整合。

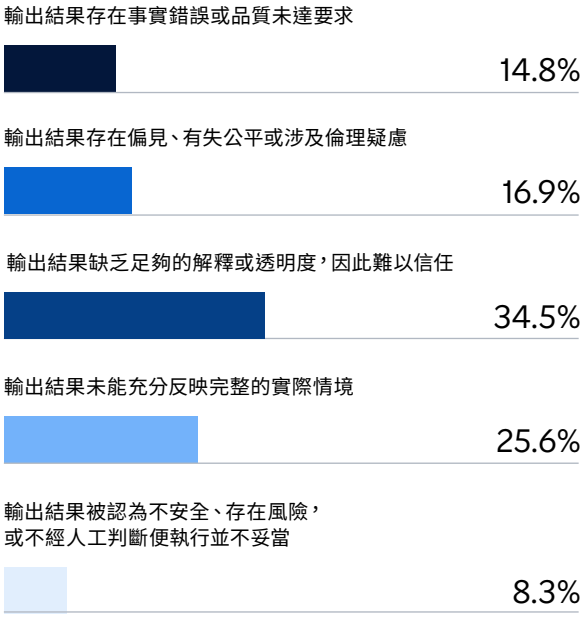


圖 5：按可信賴性水準分層的人工介入原因

從實務面來看，這再次凸顯了資料基礎的重要性。占比最高、也最易解決的人工介入案例中，來自於人們無法理解 AI 輸出的情況：缺乏充分說明與未能掌握完整情境總計占到了受訪者提及原因的大多數。AI 系統只能對可追溯的決策做出解釋，這使得可解釋性其實取決於模型底層的資料與治理能力。後續章節將探討哪些企業的這項基礎較為穩健、哪些地方仍相對脆弱，以及領先者有哪些不同做法。

©2026 SAS and IDC，保留所有權利。CCPA

資料成熟度仍是 AI 發展的瓶頸

如今部署 AI 的多數組織，所使用的資料基礎架構，仍不足以完整支援代理式 AI 所需的可解釋性、資料血緣或驗證能力。這種成熟度滯後並非抽象概念，它直接反映在人工介入率上（見第 8 頁）。各組織正穩步取得進展，但在資料成熟度迎頭趕上之前，值得信任的 AI 自主運作能力仍難以實現。

資料基礎建設的成熟度分為五個階段（臨時階段、孤島階段、標準化階段、管控階段、最佳化階段），而各階段的年度變化，也是本研究中最清晰的進展指標之一。

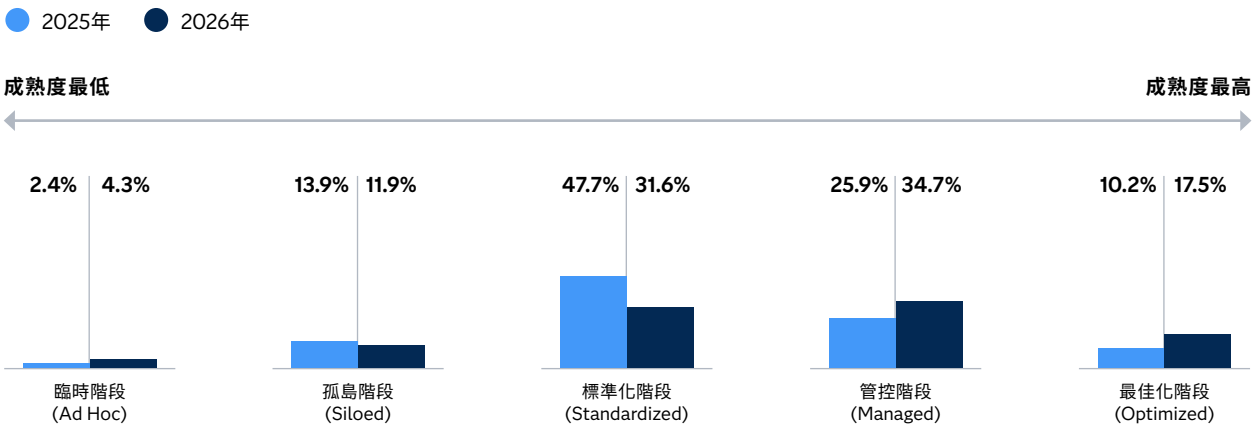


圖 6：各組織 AI 成熟度的年度變化比較

2025 年最常見的成熟度階段標準化 (Standardized) 階段占比下降 16.1 個百分點，顯示相關組織已向更高成熟度階段進階。2026 年，管控 (Managed) 階段占比上升 8.8 個百分點，最佳化 (Optimized) 階段占比上升 7.3 個百分點。目前超過半數受訪者 (52%) 認為認為自己已達到管控或最佳化階段，但僅有 17.5% 達到完全最佳化的水準。

如今部署 AI 的多數組織，所使用的資料基礎架構，仍不足以完整支援代理式 AI 所需的可解釋性、資料血緣或驗證能力。

成熟度帶來的效益：臨時階段（Ad Hoc）與基礎架構的組織比較表

	臨時階段 (Ad Hoc)	最佳化階段 (Optimized)	倍數
要求使用資料品質工具	5%	34%	**6 倍**
要求導入可解釋性管控措施	7%	40%	**6 倍**
自動化 AI 驗證與監控	9%	45%	**5 倍**
預期 AI 投資報酬率強勁/豐厚	14%	52%	**4 倍**

具備最佳化階段資料基礎架構的組織，要求導入 AI 品質與可解釋性管控的機率是其他組織的 4–6 倍，同時其預期獲得良好投資報酬率的機率也接近高出其他組織的 4 倍。

圖 7：具備臨時階段與最佳化階段資料基礎架構的組織比較表

圖 7 展示了兩類使用者群體的投入趨勢。與處於臨時階段的同業相比，最佳化階段的組織在所有 AI 專案中要求使用資料品質工具的機率高出六倍，要求導入可解釋性管控措施的機率高出六倍，執行自動化、持續性 AI 驗證的機率高出五倍，預期獲得強勁或豐厚投資報酬率的機率高出四倍。

領先者的優勢並非些微領先。而是以數倍的差距領先。領先者之所以能拉開差距，是因為他們落地的每一項管控措施，就能進一步支撐下一階段的資料成熟度。驗證機制能在項目上線前即時發現風險，進而提升組織對部署AI的信心。這份信心又為加大可解釋性投入提供了合理性，進而降低人工介入率，並帶來更好的投資報酬率。

©2026 SAS and IDC，保留所有權利。CCPA

AI 成熟度提升，也讓資料主權問題更加受到重視

資料成熟度水準也會影響資料主權相關議題對組織的影響程度。當被問及資料主權與資料洛帝要求對其 AI 部署能力的影響程度時，可信賴性水準較低的組織普遍認為影響有限：19.5% 表示「影響顯著」，4.1% 表示「影響極大」。

可信賴性水準較高的同類組織回答明顯不同：48.3% 表示「影響顯著」，29.7% 表示「影響極大」。顯著影響層級的占比約為前者的 2.5 倍，影響極大層級則約為 7 倍。

這一發現乍看似乎有些違反直覺：按理說，成熟度高的組織受到的影響應該最小。但事實恰恰相反。成熟度越高，對潛在風險的認知就越充分。

成熟度更高的企業能發現可信賴性較低的同業尚未察覺到的風險：其訓練資料跨越了哪些司法管轄區、哪些境外法規可能影響其雲端服務合約、以及一但資料落地規範收緊，哪些AI部署項目可能必須暫停。在 AI 成熟度最高的企業中，全球約 78% 的企業將資料主權問題視為一項實質限制，且該比例在各地區間幾乎沒有波動（北美 78.6%，歐洲 74.4%，拉丁美洲 80.4%，亞太地區 79.2%，中東及非洲 81.6%，全球平均 78%）。這些數據反映的是資料落地與資料司法管轄權（即資料的儲存位置及適用的管轄制度），而非國家層面的 AI 能力主權；且各地區資料幾乎持平，證實這一規律與成熟度相關，而非地域因素。

對資料主權議題的敏感度，主要反映的是組織成熟度，儘管各地仍會有不同的在地情境。投資於AI可信賴性建設的組織，往往能率先發現下一個問題。

可信賴性水準越高，對資料主權議題的關注程度也越高。

資料主權與資料落地要求對貴組織的 AI 部署能力有多大影響？

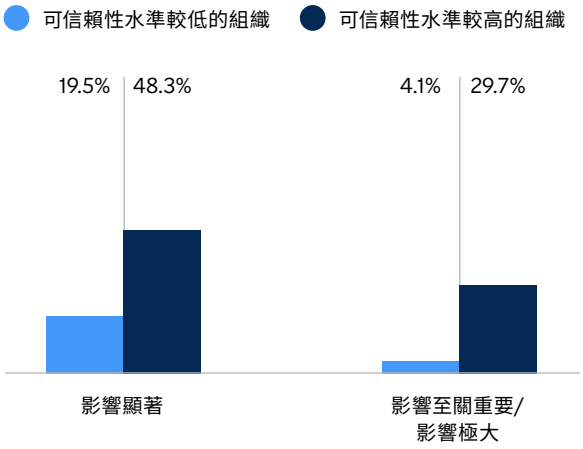


圖 8：高可信賴性組織與低可信賴性組織面臨的主權議題

成功之道：向領先企業學習

區分領先者與落後者差距的關鍵行為，正日益與組織要如何安全、規模化地部署代理式 AI 及傳統 AI 所需具備的能力一致。

Nx = 領先者採取這些做法的可能性，為落後者的幾倍

模型治理與監督
由治理領導力*所驅動*

55 倍

負責任 AI 政策
以制度化的 AI 政策*所驅動*

10 倍

最終成果

可解釋性與公平性
唯一無法單靠制度強制落實的要素

資料品質與治理
以持續驗證*所驅動*

164 倍

稽核與問責
以正式的稽核架構*所驅動*

55 倍

提前投資、超前部屬 — 這是支撐以上四項能力的關鍵行為。 **3 倍**

圖 9 高可信賴性組織與低可信賴性組織的關鍵差異

©2026 SAS and IDC，保留所有權利。CCPA

以可信賴 AI 領先組織的方式營運

我們將可信賴性指數得分達到 80 分及以上的組織定義為領先者。它們將可信賴性視為根基，而調查結果也顯示，領先者與落後者的行為模式存在巨大差距。

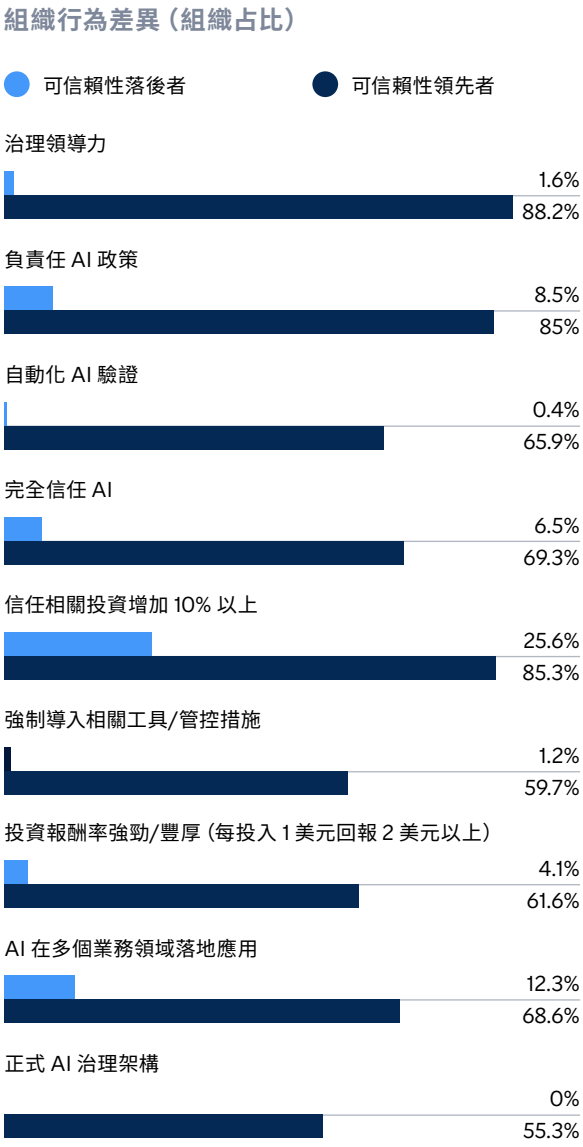


圖 10：領先者與落後者的行為差異

1.85 倍

13 項成果的實際收益

高可信賴性組織平均 27.6%，對比 低可信賴性組織 14.9%

1.86 倍

每 1 美元 AI 投入的預期投資報酬率

高可信賴性組織 2.35 美元，對比 低可信賴性組織 1.26 美元

15 倍

實現良好/高度 ROI（每投入 1 美元獲得 2 美元以上回報）

高可信賴性組織 62%
對比 低可信賴性組織 4%

圖 11：可信賴性領先者所實現的主要收益

讓 AI 領先者脫穎而出的五項關鍵做法

我們的研究總結出五項實踐，解釋了為何可信賴 AI 領先者能將 AI 投入轉化為回報，而其他組織無法做到。

1. 最佳化資料基礎架構

領先者採用「最佳化階段」資料基礎設施的可能性，是其他組織的 11 倍

資料基礎架構的建構需具備可稽核、可追溯的特性。當您能夠呈現一項決策的資料血緣時，就能為該決策提供合規的說明與辯護。若無法做到，則需透過人工介入管控風險，這種方式成本高昂且難以規模化。

2. 持續驗證

領先者在 AI 專案中採用經財務驗證的 ROI 模型的機率高出 10 倍

AI 領先者採用經財務驗證的模型，在工作流層面衡量 AI 的價值（即由財務長團隊驗證預期收益，對照預測值追蹤目標指標，並指定專人負責交付成果），而非透過試點項目或主觀感受來衡量。這一做法能明確區分哪些 AI 投資真正有效、哪些無效，避免失敗的試驗淪為隱性成本。這種嚴謹性不應僅適用於 ROI，也應同等適用於風險與潛在危害評估。

3. 對全體員工執行可信賴 AI 訓練

85% 的領先者在全組織範圍內開展負責任 AI 政策訓練

領先者不只是制訂政策，還會在全組織內積極落實負責任 AI 政策的訓練工作。當所有人都熟知規則時，治理責任便由全業務部門共同承擔，而非僅交由專業團隊負責，因此問題能被及早發現，而非等到上線後才被使用者察覺。問責機制也因此真正融入組織架構之中。

4. 將治理內建為預設機制

領先者進行治理與可解釋性審查的機率，是落後者的 55 倍

領先者設立具備實際許可權的獨立治理職能部門。看似矛盾的是，這反而會加快部署進度，因為各團隊明確知曉自身的權責邊界。

5. 前瞻性投入

85% 的領先者將可信賴 AI 相關投入提升 10% 以上

領先者將可信賴性建設本身作為投入方向，將其視為戰略要務而非選項。這種差距並非僅僅在於領先者投入更多：85% 的領先者將可信賴 AI 相關投入提升 10% 以上。

每一項做法都為下一項能力奠定基礎。綜合而言，各項措施相輔相成，產生複利效果。

©2026 SAS and IDC，保留所有權利。CCPA

AI 可信賴性如何影響 ROI？

投入可信賴 AI 建設的組織，其 AI 投資報酬率表現截然不同。可信賴性得分較高的組織，實現良好或高度 ROI 的機率是得分較低組織的 15 倍。

差距十分顯著。可信賴性指數得分達到 80 分及以上的組織中，近六成的 AI 投資報酬率至少達到投入的兩倍。而在可信賴性得分較低的組織中，僅有不到二十分之一能取得同等成果。

這種關聯在代理式AI時代顯得尤為重要。如果員工、監管機構與客戶對自主系統缺乏足夠信任、不願使用，組織就無法實現 AI 的規模化部署。

研究結果表明，可信賴性不止是風險管理階層面的考量。它本身就是一項商業優勢。建立清晰治理制度、維持穩健資料基礎、讓 AI 決策具備可理解性的組織，更有能力從試驗階段邁向實際營運落地。AI 將成為工作流程的一部分，而非又一個爭奪關注度與資源卻始終停留在試點階段的專案。

這種差異在受高度監管的產業中尤為關鍵，在這些領域，問責、可解釋性與監督機制決定了新技術的落地速度。提前滿足這些要求的組織，往往能更有信心地規模化部署 AI，遇到的阻礙也更少。

相比之下，將可信賴性視為事後才需要處理的問題的組織，即便技術本身效能出色，也往往難以從中實現價值。模型效能或許優異，但落地採納進度滯後。應用場景始終有限。部署推進更為緩慢。最終導致 AI 投入與實際成效之間的差距不斷拉大。

資料指向一個明確的結論：可信賴性並非僅僅守護 AI 的價值。更能創造價值。

有鑑於此，86% 的組織表示將在來年增加可信賴 AI 投入，且業界幾乎普遍認為，AI 代理轉型需要採取與生成式 AI 不同、更為嚴格的治理方式。

15 倍

更有可能實現：良好/高度 ROI

15 比 1。*約六成高可信賴性組織每投入 1 美元 AI 資金可獲得 2 美元以上回報 — 而低可信賴性組織中僅有不到二十分之一能實現這一水準（占比分別為 62% 對比4%）。*

圖 12：可信賴 AI 組織的 ROI 倍數效應

阻礙 AI 創造價值的三大錯誤

AI 領先者的特徵，不只來自它們採取了哪些做法，也來自它們懂得避開哪些錯誤。從我們的訪談和問卷資料來看，我們發現有三項常見的陷阱，，持續拉開領先者與落後者之間的差距。

部署人們不信任的 AI

無法對 AI 決策作出解釋的組織，最終會為此付出雙重成本。如前所述，79% 的組織在超過 10% 的場景中會介入調整 AI 的建議，而缺乏可解釋性是首要原因。每一次人工介入都會帶來額外的審查、重工與延誤的成本。建議各組織將 AI 系統與模型的透明度和可解釋性作為關鍵要求。

別把試點專案誤當成真正的進展

約有一半的 AI 預算投入銷售與行銷相關的試點專案，但許多能帶來最高 投資報酬率 的應用場景，卻未能獲得足夠資源。領先者投入的是能夠改變工作執行方式的工作流，而非只用於展現 AI 能做甚麼的示範專案的。

別將治理視為阻礙，而應視為推動力

在開發完成後才追加治理機制的組織，會產生摩擦、延誤和內部阻力。領先者將治理視為系統加速器而非減速帶，進而更有信心地實現 AI 規模化部署。

綜觀各產業，都反覆出現同樣的模式。處於領先地位的組織，並非採用了本質上截然不同的技術。它們只是規避了少數幾個代價高昂、阻礙 AI 創造價值的常見錯誤。

©2026 SAS and IDC，保留所有權利。CCPA

不同產業如何塑造 AI 領先優勢

在下一節中，我們將進一步探討本研究聚焦的四大重點產業：銀行業、保險業、生命科學與公共部門。每個產業都將呈現可信賴 AI 領先者與落後者的行為在真實產業用例中的實際表現。

綜合來看，這些內容有助於理解，一套共通通的可信賴 AI 方法，該如何因應各產業不同的監管、營運和聲譽風險等實際情況進行調整。

畢竟，在AI真正被應用於解決實際問題之前，其價值都只是理論層面的。觀察全球各重點產業的領先者如何佈局 AI，有助於釐清可信賴 AI 的通用原則，也能讓我們從多元視角中汲取經驗 — 無論您當前身處哪個產業。

銀行業領先者的經驗

銀行業

可信賴性

63.5 **+12.8**

Y1:50.7

主要關注議題與優先事項

在各項可信賴性指標中，模型治理與監督存在 10 分的差距，是目前唯一相對明顯的弱項。縮小這項差距，是銀行業的下一項優先任務，而隨著AI代理自主化發展，若企業未能及早建立核准與監督機制，之後再補強的難度將更高，因此這項課題也更加迫切。

更深層的挑戰在於營運與決策速度：領先者與落後者瞄準的其實是同樣的關鍵任務型應用業務場景，但領先者更願意讓 AI 在明確界定的範圍內自主行動（例如：41% 的高可信賴性銀行在調查業務中部署 AI，而低可信賴性銀行僅為 7%）。

可信賴 AI 領先組織的 ROI 成果高出 6.6 倍

高可信賴性銀行已將治理機制層層落實至全組織，而落後組織尚未啟動相關工作，兩者之間的差距非常明顯。85% 的領先銀行已全面推行負責任 AI 政策，涵蓋所有員工；而落後銀行的比例則僅有29%。

超過三分之二的領先銀行已實現自動化、且持續性的 AI 驗證；相較之下，落後群體幾乎完全沒有這類機制。這些差異並非枝微末節，而是根本性的差異：它們直接決定組織是否真正具備支撐AI運作的營運模式。

因此，結果顯而易見：有 66% 的領先銀行預期能獲得良好或高度的 AI 投資報酬率。在原本AI發展基礎就點高於多數產業的銀行業中，領先者與落後者之間的差距仍達到 6.6 倍。

> AI 代理面臨的最大挑戰是：如果AI 代理做出了某項決策，我們該如何回溯過程、追溯原因，並解釋它為何採取特定行動？這也是我們尚未完全接納自主化 AI 的原因之一。
>
> 社區銀行首席資訊安全長（美國）

©2026 SAS and IDC，保留所有權利。CCPA

在自動化調查上的明顯分歧

領先者與落後者對 AI 的核心應用領域共識一致：信貸、風險、資格審核與金融犯罪偵測。真正的差異不在於優先順序，而在於AI 能被賦予多少自主權。

領先銀行將自動化調查與案件管理視為視為關鍵任務型應用的比例遠高於落後銀行 (41% 對比 7%) 是本研究中所有應用場景中差距最大的一項。這類系統不止是提供洞察供人工審核，而是會在明確設定的管控範圍內直接採取行動。

結論顯而易見：競爭優勢並非源自選擇了不同的應用場景。而是源自建構了足夠的信任基礎、治理機制與問責體系，讓 AI 能夠足夠的速度、規模化執行。

隨成熟度提升，關注的風險也會改變

64% 的領先銀行將資料隱私與安全列為首要風險，而落後銀行的這一比例為 46%。

相比之下，落後銀行更常受限於可靠性疑慮 (39%) 和 AI 權責歸屬不明 (22%)，這些通常是在治理機制尚未建立完整就會遇到的問題。雙方對AI 代理相關問題的擔憂均大幅上升，而兩者的差異也相當值得注意：38% 的領先銀行表示如何管理 AI 自主性與人工監督是目前的主要疑慮，而落後銀行僅為 12%。原因在於，領先者已經真正進入這類 AI 部署階段，因此開始實際面對這些問題；相較之下，落後者甚至還沒發展到需要處理這項挑戰的程度。

銀行業：可信賴性的五大構面

資料品質與治理

可信賴性

61.3 /100

負責任 AI 政策

可信賴性

69 /100

模型治理與監督

可信賴性

59.4 /100

稽核與問責

可信賴性

64.6 /100

年增幅最大 +14.3

可解釋性與公平性

可信賴性

62.6 /100

年增幅最小 +7.9

給領導者的重點啟示

找出組織自身已執行成熟的管控體系，將其遷移應用到 AI 領域，而非將治理視為一項拖慢節奏的全新建設任務。

圖 13：銀行業 — 可信賴性的五大維度

銀行業在可信賴 AI 的所有維度上得分均高於全球平均水準，而且各構面的年增幅約介於 8 至 14 分之間。五大構面的信任缺口均已縮小；但在模型治理與監督方面，實際能力與感知信任之間仍存在明顯差距，目前約為 10 分。

我們也看到一個非常一致的現象：沒有哪個構面被忽略，也不會出現某像差距縮小、另一個項卻差距持續擴大的情況。銀行業正將累積了二十年的模型風險管理 (MRM) 經驗延伸至 AI 領域，其推進速度也快於那些必須從零開始建立相關能力的產業。

©2026 SAS and IDC，保留所有權利。CCPA

唯一仍處於相對落後的構面是模型治理與監督，而這正是將模型風險管理的落實到實際營運中的關鍵層次：包含治理委員會、報告與權責關係、文件記錄與核准流程等，也就是銀行如何真正做到「管理治理本身」。因此，縮小這 10 分的差距是銀行業下一階段的核心重點，而隨著AI代理自主化發展，事後才進行補建核准與監督機制的難度越來越大，因此這項需求也變得更為迫切。這種以模型風險管理為主導的模式，在美國等市場尤其明顯，這類市場的 AI 專項監管規則仍處於初期階段，監管機構多半將 AI 納入現有的模型風險監管要求之中。在歐盟等設有專門 AI 監管體系的司法管轄區，AI 治理的界定更為明確，並不局限於模型風險管理範疇。

> 我們設有 AI 治理委員會..... 成員包括業務、技術、安全相關負責人，以及風險與合規負責人。我們也有法務人員參與，確保他們能從聲譽風險角度提出意見，提醒我們需要留意的事項。
>
> 全球銀行董事總經理 (美國)

保險業領先組織的經驗啟示

保險業

可信賴性

63.2 +17.6

Y1:45.6

主要關注議題與優先事項

保險業目前的重點，是隨著 AI 部署規模擴大，持續讓實際能力與信任認知維持一致。因為保險業是唯一一個負責任 AI 能力的實際評分已高於信任認知的產業。

對領先組織而言，目前最直接的挑戰，是如何在規模化部署後管理AI的自主性與人工監督：34% 的高信賴度保險公司指出這一點，而低信賴度的同業中僅有 21% 如此表示，因為當今負責推動這些部署的正是領導階層。

領先組織與落後組織之間的差距最大

在本研究聚焦的產業中，保險業呈現出最明顯的由「信任驅動型」差距，幾乎在所有指標上，領先組織與落後組織的差距均大於銀行業。78% 的領先組織已將負責任 AI 政策落實至全體員工，而落後組織的這一比例僅為 3%。數領先組織已建立自動化持續驗證機制與獨立治理職能；但在落後組織中，相關機制幾乎不存在。

最終結果也反映出這種巨大差距：59% 的領先組織預期 AI 將帶來良好或較高的投資報酬率，而落後組織的這一比例為 0% - 這使保險業成為本研究聚焦的產業中，唯一一個在落後組組織群裡，完全沒有出現良好/高度投資報酬率的產業。在一個以風險定價為核心的產業中，唯有及早對治理風險進行評估的機構，才能將 AI 投入轉化為 AI 收益。

> 100% 可解釋性是關鍵核保營運團隊現在會收到月度報告。相較過去，營運團隊需逐份審核文件，如今僅需針對約 10% 的工作進行人工處理，不再需要手動完成整個流程。
>
> 全球保險資料與治理主管（英國）

©2026 SAS and IDC，保留所有權利。CCPA

領先組織更聚焦於能創造價值的成果

對保險業領先組織而言，AI 處於整體決策智慧架構的核心：涵蓋風險評估、核保、法規遵循、定價、資本管理、理賠與詐欺防治等場景。企業級分析與客戶服務也都納入AI在應用範圍。

當被問及最希望 AI 達成哪些成果時，落後組織將成本節省列為首要訴求，占比達 48.5%，遠高於將其列為優先事項的領先組織占比 (23.3%)。相反，領先組織更聚焦於創造價值而非壓縮成本的成果：營運效率提升 (37.2% 對比 27.3%) 帶動新的營收成長 (30.2% 對比 15.2%) 提升獲利 (30.2% 對比 21.2%)。這項結果也印證了本研究長期以來最一致的發現之一：主要以降成本為目標部署 AI，往往是最難發揮 AI 價值的方式之一。成熟度最低的保險機構，仍將降成本作為首要目標；而領先組織已轉向成長、效率與利潤提升。

部署成熟度提升，關注議題也跟著演進

領先組織已跨越基礎階段，現在更關注的是如何規模化部署。法規遵循、問責機制與聲譽風險等問題，通常要等到 AI 真正開始執行會影響監管責任、且品牌方需為此承擔責任時，才會成為實際而迫切的課題。52% 的落後組織將資料隱私與安全列為首要關切，對比之下領先組織的這一比例為 42% - 這表明落後組織仍在著力解決資料基礎架構層面的問題。

在代理型AI方面，34% 的領先組織表示，如何管理AI自主性與人工監督是一項主要疑慮，相較之下，落後組織的這一比例為 21%。保險業的高可信賴性領先組織之所以開始擔心 AI 自主性，正是因為它們當下已經實際部署並運行這類系統。

在一個以風險定價為核心的產業中，唯有及早對治理風險進行評估的機構，才能將 AI 投入轉化為 AI 收益。

保險業：可信賴性的五大構面

資料品質與治理

可信賴性

60.2 /100

負責任 AI 政策

可信賴性

69.6 /100

年增幅最小 +11.3

模型治理與監督

可信賴性

60.3 /100

年增幅最大 +16.2

稽核與問責

可信賴性

63.8 /100

可解釋性與公平性

可信賴性

61.6 /100

圖 14：保險業 — 可信賴性的五大構面

給領導者的重點啟示

讓 AI 部署規模與使用者信任同步擴大，使能力與信心能夠一起成長。

今年保險業五大構面的得分同步提升，年增幅區間為 11.3 至 16.2 分。模型治理與監督是增幅最大的維度，提升 16.2 分，這反映出是保險機構正積極建立自動化核保與理賠決策所需的核准流程與監督機制。負責任 AI 政策的增幅最小，提升 11.3 分；但同時它也是保險業得分最高的構面，滿分 100 分中得分為 69.6 分，增幅較小反映出該構面的剩餘提升空間相對有限。保險業在治理與政策層面均已完成基礎建設工作。下一步的重點，是在AI部署規模化的過程中保持這一增速，確保實際能力與使用者信心同步成長。

©2026 SAS and IDC，保留所有權利。CCPA

生命科學領域領先組織的經驗

在監管嚴格的環境中，提升 AI 自主執行能力的必經之路，是強化治理體系與可信賴性實務。

生命科學產業

可信賴性

64.8 +13.7

Y1:51.1

主要關注議題與優先事項

首要重點為資料基礎。生命科學具備其他產業難以複製的結構性優勢：稽核與問責得分為全研究樣本最高，該項指標的落差同比縮小 20.9 分，變成 -0.6，是全部數據集當中唯一出現反轉的構面。

62% 的高可信賴性組織將資料隱私列為關切點，而低可信賴性同業者為 44%。這是因為高可信賴性企業更清楚掌握自身資料的完整範圍，以及其中所涉及的高風險特性。

領先組織正以非常顯著的幅度超越落後組織

本次研究涵蓋的生命科學機構，涵蓋製藥企業、醫療器械與裝置公司以及臨床研發機構。生命科學產業的高可信賴性 AI 領先組織，在治理體系、持續驗證與負責任 AI 政策方面的成熟度，顯著高於落後組織。

91% 的領先組織已將負責任 AI 政策全面落實至全體員工，而落後組織的這一比例僅為 7%。多數領先組織已建立持續驗證機制與獨立治理職能；落後組織則兩者均未建立。

因此，55% 的領先組織預期 AI 將獲得良好或較高的投資報酬率，而落後組織的這一比例為 11%。人才方面的差距同樣明顯：22% 的落後組織表示，內部沒有具備 AI 倫理或風險管理專業能力的人員，這限制了它們以負責任方式規模化部署 AI 的能力。

領先組織並非僅小幅領先，而是已處於完全不同的發展層級。

> 在最初的導入過程中，我們發現AI輸出結果並不值得信任，原因是資料尚未準備就緒。我們花了近一年時間，才將資料品質提升至足以讓系統有效學習的水準。
>
> 生命科學產業亞太區技術服務資深總監（新加坡）

治理是 AI 轉化為實際成果的關鍵

對領先組織而言，AI 應用聚焦於監管要求與患者需求交會的關鍵領域：包括法規申報與證據產出、製造品質與供應鏈、藥物研發以及臨床試驗設計。領先組織將這些應用場景列為業務核心關鍵的比例，是落後組織的兩至三倍。

33% 的落後組織表示，AI 尚未在任何業務領域成為核心關鍵，而領先組織的這一比例僅為 3%。在AI部署落地階段，監督模式呈現顯著分化：56% 的落後組織要求實行全程人工審核。在組織尚未以與其他業務領域相同的嚴謹稽核標準來管理 AI 治理之前，高度警慎的模式仍是管理自主型 AI 系統最安全但成本最高的方式。

領先組織已將這套嚴謹的稽核機制延伸至 AI 治理：相較之下，領先組織更傾向於僅對異常情況進行人工審核，或讓 AI 在明確設定的界線內自主採取行動。領先組織已具備讓系統自主採取行動的治理與信任基礎，落後組織則尚未具備同樣的條件。

這樣的成果，也直接反映在生命科學產業的核心競爭領域：治理成熟度讓領先組織得以在減少人工審核的情況下執行 AI；也正是這種成熟度，讓 AI 具備了足夠的可靠性，能夠縮短法規申報與證據產出的時程、提升臨床試驗設計與執行效率、並強化製造品質。在生命科學產業中，治理正是把 AI 導入轉化為更快速、更可靠成果的關鍵，尤其是在那些風險與影像最高的工作場景中。

各產業中最嚴峻的風險考量

在本研究聚焦的各產業中，生命科學產業面臨的風險考量最為嚴苛，因為該產業 AI 應用一旦發生錯誤，將錯可能造成的重大後果。領先組織比落後組織更加關切資料隱私的問題（62% 對比 44%）。看似意外的是，領先組織的之所以更擔心，原因在於它們更清楚自身資料所含內容的完整範圍，以及這些資料本身具有的高風險屬性。

隨著可信賴性提高，組織對法規遵循、輸出可靠性與AI倫理應用方面的重視度，也明顯上升。反差最明顯的差異是：26% 的領先組織將刑事或監管責任列為主要疑慮，而落後組織的這一比例僅為 7%。最貼近臨床與製造風險的機構，十分清楚 AI 決策失誤時個人與企業的責任歸屬，因此也會採取相應的治理措施。

©2026 SAS and IDC，保留所有權利。CCPA

生命科學產業：可信賴性的五大構面

資料品質與治理

可信賴性

62.6 /100

負責任 AI 政策

可信賴性

70.5 /100

年增幅最小 +8.9

模型治理與監督

可信賴性

62 /100

稽核與問責

可信賴性

68.8 /100

年增幅最大 +15.4

可解釋性與公平性

可信賴性

63.5 /100

給領導者的重點啟示

成熟度越高、對風險的關注反而越精準，這代表企業正在治理真正重要的風險；而第一步，就是在擴大 AI 部署之前，就先把資料基礎打穩。

圖 15：生命科學產業 — 可信賴性的五大構面

生命科學產業的稽核與問責得分，是本研究所有重點產業中最高的。這個產業長期以來就具備受到嚴格監管的品質管理制度與臨床試驗稽核軌跡，因此能比使用者認知與信任提升的速度更快地，把既有稽核機制延伸到 AI。負責任 AI 政策也呈現同樣的發展邏輯。

領先組織正在進一步擴大領先優勢，並持續投入以維持領先優勢。生命科學產業的結構性優勢在於，已擁有成熟且受到監管的營運稽核標準。因此，將這套標準延伸應用至 AI 領域的成本，遠低於從零開始建立一套全新的治理機制。

公部門領先組織的經驗

對公部門機構而言，可信賴性並非單純的部署要求；更是讓 AI 能在政治與實務運作上進一步擴大自主性的前提。

公共部門

可信賴性

55.5 +6.5

Y1:49.0

主要關注議題與優先事項

首要重點是透明度與可解釋性。41% 的高可信賴性公部門機構將這項議題視為目前最主要、仍待改善的缺口，低可信賴性同業的這一比例則為 26%。

這表明，領先組織已將透明度不足，認定為阻礙它們從目前階段邁向下一階段 AI 應用的關鍵障礙。

領先組織正快速擴大可信賴 AI 投資規模

在公部門，可信賴 AI 領先組織的運作模式與落後組織已存在根本性的差異。91% 的領先組織已將負責任 AI 政策全面落實至全體員工，落後組織的這一比例僅為 5%。84% 的領先組織屬於治理領先梯隊，而落後組織中達標占比僅 1%。大多數領先組織均具備持續驗證機制與獨立治理職能；在落後組織中，這類機制幾乎不存在。半數領先組織預期 AI 將帶來良好或較高的投資報酬率，落後組織的這一比例僅為 4%。

在投資承諾方面，公部門領先組織擴大可信賴 AI 投資的速度，在所有受訪產業中位居最高。41% 的領先組織計劃在未來一年將可信賴 AI 投資提升 20% 以上，其投資增幅可比擬本研究所有產業中最積極的企業。

> 但凡涉及 AI 的項目，產品上線前往往要等待一年時間。這就是我們面臨的問題。但可以理解的是，這畢竟是新興領域。我們必須確保所打造的解決方案是安全的，因為它可能影響到人們的生命。
>
> 國家衛生服務部高級技術總監（英國）

©2026 SAS and IDC，保留所有權利。CCPA

對民眾的影響，決定 AI 的應用重點

對公部門的領先組織而言，最重要的 AI 應用場景集中在分析領域：資料分析、預測、政策評估、資格審核、公共安全及風險預測。75% 的領先組織將資料分析與預測視為關鍵任務，落後組織的這一比例為 36%。這樣的結果並不令人意外，因為這些領域正是對民眾的影響最大，且對可解釋性的要求最高。

AI 輸出驗證方面，也再次出現差距。59% 的落後組織要求實行全程人工審核：這是所有產業落後組中，最為審慎的管控做法。這既反映出治理成熟度仍不足，也反映出AI在公共服務中一但出錯，可能付出的代價更高。

相較之下，領先組織更傾向於讓 AI 在明確設定的界線內自主行動，或僅在事後進行審查。領先組織已對自身管控機制建立起足夠信心，能夠賦予系統自主行動的許可權；落後組織尚未達到這一水準。

領先組織已跨越基礎性疑慮

不同成熟度組織所關注的風險，呈現出非常清楚的分界。落後組織仍聚焦於基礎風險：資料隱私與安全洩露，其關注度高於其他所有產業的落後群體。領先組織則已經進一步關注更成熟階段的議題。

41% 的領先組織將透明度與可解釋性列為主要疑慮，落後組織的這一比例為 26% — 這顯示，領先組織已將透明度不足視為目前最主要、仍待解決的缺口，而縮小這一差距正是它們下一步加快 AI 推進速度的關鍵。此外，倫理應用、偏見管控與 AI 自主性管理也都受到高度重視。

聲譽風險則是領先組織的另一項重要疑慮：34% 的領先組織將其列為主要風險，落後組織的這一比例為 19%。與透明度問題一樣，它已不再屬於基礎風險，而是外部風險；這尤其符合公部門的特性— 在該領域，一次公開的 AI 失誤就可能影響全體民眾的安全、保障與生計，其影響範圍遠超大多數民間企業。領先組織已經開始關注真正重要的風險。落後組織仍在處理基礎風險。

公部門：可信賴性的五大構面

資料品質與治理

可信賴性

52 /100

負責任 AI 政策

可信賴性

60.5 /100

年增幅最小 +0.7

模型治理與監督

可信賴性

52.3 /100

年增幅最大 +3.8

稽核與問責

可信賴性

54.7 /100

可解釋性與公平性

可信賴性

52.6 /100

圖 16：公部門 — 可信賴性的五大構面

給領導者的重點啟示

對 AI 的疑慮，也能成為推動企業投入資源、建立可信賴基礎的重要動力。

公部門在可信賴 AI 各項指標上的得分，低於所有其他重點研究產業。

銀行業、保險業與生命科學產業在五大構面的得分均處於 50 分後段至 70 分前段，其中負責任 AI 政策是各產業表現最強的構面（得分分別為 69、69.6 及 70.5）。公部門在所有構面上均落後於這三個產業，得分主要集中在 52 至 61 分之間：其中，負責任 AI 政策仍以 60.5 分居首位，資料品質與治理以 52 分墊底，模型治理與監督、可解釋性與公平性、稽核與問責位列中間。模型治理與監督創下該產業最大年增幅（+3.8 分），負責任 AI 政策的增幅最小（+0.7 分）。

©2026 SAS and IDC，保留所有權利。CCPA

中小型機構

企業規模並不會決定一家企業是否可建立可信賴 AI 。中小型企業的可信賴性得分分佈與大型企業幾乎相同：在占樣本群體絕大多數的 60-80 分區間，中小型企業占比 41.4%，大型企業占比 40.6%；在 80 分及以上的領先企業區間，兩者占比分別為 14.2% 與15.5%。在所有可信賴性層級上，兩者差距最大不超過 1.5 個百分點。

而這些些微差距，多半反而更有利於中小型企業。中小型企業的可信賴性指數與大型企業持平 (61.9 對比 62)，且在實際效益 (56 對比 54.5) 與感知可信賴性 (71.5 對比 70.9) 兩項上得分更高。

資料品質與治理

可信賴性

59.4 /100

負責任 AI 政策

可信賴性

68.6 /100

領先大型企業最多 +1.3

模型治理與監督

可信賴性

57 /100

落後大型企業最多 -2.1

稽核與問責

可信賴性

62.7 /100

可解釋性與公平性

可信賴性

59.4 /100

給領導者的重點啟示

無論企業規模、人力或資源多寡，可信賴 AI 的五大要素都適用於所有組織。若想深入瞭解中小型企業的 AI 準備度，可參閱中小型企業 AI 準備度報告。

圖 17：可信賴 AI 的五大構成要素 — 中小型企業

它們在建立可信賴性的實際作法上也更為領先：

- 更多中小型企業不只是制定負責任 AI 政策，同時也對所有員工進行相關訓練 (33.8% 對比 31.6%)。
- 它們也更積極投資於透明度與可解釋性 (34.1% 對比 30.6%)。
- 它們投入 AI 倫理領域的人才招募與訓練的比例也較高 (32.5% 對比 29.4%)。
- 在可解釋性與專職AI 倫理人才方面，中小型企業與大型企業持平甚至更優。

中小型企業對部分可信賴性風險的信任疑慮也更低。僅 46.7% 的中小型企業將資料隱私與安全列為可信賴性的主要限制，大型企業的這一比例為 51.7%；同時，提及AI倫理或偏見問題的中小型企業占比也更低 (24% 對比 30%)。在商業目標方面，中小企業的商業優先順序更側重營收與商業成果導向：中小型企業更常運用 AI 推動客戶成果，包含優先提升客戶滿意度 (49.7% 對比 40.8%) 與營收成長 (43.7% 對比 37.9%) 列為優先事項，並將客戶互動與個人化服務視為關鍵任務 (47.8% 對比 42.2%)。

換句話說，中小型企業的可信賴性與大型企業持平，且在實際效益與感知可信賴性兩方面表現更優。

它們的治理模式也更加精實且有紀律。更多中小型企業採用例外情況才需人工核准模式 (34.5% 對比 31.4%)，這是一種較有效率、也較容易規模化的做法。這種紀律也展現在實際成果上：在對 AI 成效進行量化統計的中小型企業裡，39.7% 表示過去一年 AI 帶來了 11% 至 20% 的提升，而大型企業的這一比例為 28.9%。

不過，仍有兩項差距存在，，而且都與企業規模有關。中小型企業導入最新AI技術的速度較慢 (生成式 AI：65.9% 對比 70%；代理式 AI：37.7% 對比 42.9%)；且設立專職 AI 治理或風險職能部門的企業占比遠低於大型企業 (8.6% 對比 11.7%)，因此更常依靠已有預算支持的計畫與指定負責人來推動。在這些面向上，企業規模仍具有影響。

儘管存在這兩項差距，本節的核心發現依然成立：資源充足確實有所幫助，但無法決定最終成果。而這個模式不只適用於企業規模。

©2026 SAS and IDC，保留所有權利。CCPA

所有企業都應掌握的關鍵重點

佈局可信賴 AI 建設的機構，在 13 項商業成果上可實現 1.85 倍的效益提升，而且回報良好或高額投資報酬率的可能性高出 15 倍。由此可歸納出二項結構性的影響。

綜合來看，整份研究傳達的是一個一致的訊息。各組織正步入 AI 應用的新階段，而AI自主性的提升速度，正快於可信賴性的建立速度。真正拉開差距的領先企業，並不是因為部署了更多 AI，，而是來自他們建立了足夠的信任基礎，因此能更有信心地將 AI 應用於高風險、高影響的決策。

優勢正在產生複利效應

資料與可信賴 AI 成熟度帶來的優勢已然十分顯著，而且這項優勢還在持續累積。本研究中最最值得關注的前瞻性數據是：85% 的領先組織，未來一年將把可信賴 AI 的投資增加超過 10% ，與之對比，落後組織的這一比例為 26%。領先組織並非只為守住領先地位。它們正透過持續的資金投入不斷擴大領先優勢。

目前兩組企業之間存在的，是 15 倍的投資報酬率差距；而今天的投資態度，則預示了它們 12 個月後會站在哪裡。領先企業正在把從信任所獲得的「紅利」再投資，以進一步擴大競爭優勢；而從目前數據來看，落後企業若維持現有發展軌跡，並沒有明顯跡象顯示它們能追上來。每多一個季度未增加投資，未來追趕的成本就會更高。

仍待完成的工作

本研究中改善幅度最小的構面 — 可解釋性與公平性，而這也正是使用者選擇人工介入 AI 最常見的原因。這並非巧合：多數組織優先完成了較容易看見的治理工作（政策、稽核、問責架構），而將資料血緣、模型可解釋性、介面設計等難度更高的技術工作延後 — 而這個「以後再處理」的階段已經到來。大量的人工介入，就是企業當前為此付出的代價。

值得注意的是，資料基礎架構是所有這些能力的底層基礎。只有當一套系統運作在可追溯的資料之上，它才能真正解釋一項決策。僅 17.5% 的組織已經達到了完整最佳化的資料基礎架構。其餘組織部署 AI 的底層基礎，仍無法完整支援代理式 AI 所需的驗證能力。因此，企業若不先補上資料成熟度的差距，就無法真正解決從生成式 AI 邁向代理式 AI 時所出現的信任下滑問題。

85% 的領先組織未來一年將把可信賴 AI 的投資增加超過 10% 以上。

從信任走向優勢優勢：未來之路

組織當前所處階段，決定下一步的行動方向。部分組織已經建立完備的可信賴 AI 基礎。另一些組織仍在建構規模化 AI 所需的資料、治理與可解釋性能力。我們將建議分為兩大類：市場領先組織、仍在建置可信賴 AI 能力的組織。

給仍在建立能力的組織

補上五項關鍵行為差距。持續進步，比追求一次到位的完美更重要。

建置資料基礎架構

如果缺乏可追溯的高品質資料，上層所有的治理機制都無法發揮作用。

持續驗證

在部署前發現問題，從結構來說，比部署後處理人工介入處理問題的成本更低。

將政策制度化並落實到全組織

企業需要把政策正式制度化，並真正推廣到整個組織。要靠專案專案推動落實，不能只靠一份 PDF 文件。

配備獨立治理團隊

組織必須建立一個具備實質權限的獨立治理單位，負責管組織內的 AI 。

投資可解釋性

把可解釋性視為優先工程目標，而非可有可無的附加功能。

給可信賴 AI 領先者

下一步真正的考驗，就是代理式 AI。

- 將既有的可解釋性、驗證與稽核規範延伸至代理式AI 系統 — 這類系統不再只是提供建議，而是會直接採取行動。
- 若組織的治理架構已經建立完成，那麼從生成式 AI 邁向代理式 AI 所出現的信任下降，反而可以視為一項競爭機會。
- 讓 AI 的推理依據直接呈現在使用者實際操作的介面中，而非埋藏在文件中。
- 此外，不要把「85% 的領先組織正加大 AI 投資」當成自己的比較基準，而應視為起點，因為落後企業也不會停在原地。

給所有組織

可信賴性不會拖慢 AI 部署，反而能讓 AI 部署更持久、更能規模化。每一筆投入在資料品質、模型治理、可解釋性與稽核上的預算，實際上都是在為企業取得一種能力：可以部署更多 AI、賦予 AI 更高自主性，同時減少人工介入與營運負擔。

本研究中的領先者，已經透過大規模、可衡量的成果證明這套模式可行。這套方法是有效的。真正的問題在於：貴組織能否在與領先者的差距演變為難以彌補的結構性鴻溝之前展開行動。

不將可信賴性內建於 AI 當中的代價，不是未來才會發生的風險，而是此時此刻就會浮現：反映在居高不下的人工介入率、停滯的試點專案，以及落後組織只能等待「信任追上技術能力」，而領先組織卻持續累積投資報酬率。現在就進行投資的組織，正在縮小差距，建立持久的競爭優勢。

©2026 SAS and IDC，保留所有權利。CCPA

名詞解釋

感知信任

使用者/組織對 AI 的主觀信任程度。

可信賴性

用以證明 AI 可靠性、公平性與受監督性的客觀治理實務。

信任缺口

感知信任 － 可信賴性。缺口越大，代表過度信賴或錯置信任的風險越高。

效益指數

彙整 AI 已實現商業成果的綜合指標。

生成式 AI/代理式 AI

生成式 AI 負責產出內容；代理式 AI 則能進行規劃，並自主執行多個步驟。

低可信賴性組織 = 落後者

可信賴性指數得分在 0-39 分的組織。

高可信賴性組織 = 領先者

可信賴性指數得分在 80-100 分的組織。

可信賴性五大構成要素

資料品質與治理
模型治理與監督
稽核與問責
可解釋性與公平性
負責任 AI 政策

AI 代理(Agents)

AI 代理可感知、推理，在有限人工互動下執行多步驟作業。代理式 AI 系統可協調調度多個AI 代理協同工作。

傳統（預測型）AI

根據資料預測結果的 AI，可視為位於生成式 AI「創造內容」與代理式AI「採取行動」之下的「預測」層。

人工介入傾向

使用者介入 AI 輸出結果的傾向；是衡量人們在日常工作中對實際信任代理式 AI 到什麼程度的行為指標。

可信賴性指數

基於五大構面對組織治理實踐進行的 0-100 分綜合評分，用於劃分領先者與落後者。

資料主權

指資料落地與資料司法管轄權，包括資料實際儲存在哪裡，以及受到哪些法規制度管轄。

負責任 AI

組織以合乎倫理、安全且合規的方式使用 AI 的相關政策與實務。

可信賴 AI

經過合理設計、治理與驗證，具備可靠性、公平性、可解釋性與問責性的 AI。

資料成熟度分級

1. **臨時階段 (Ad hoc)**。資料架構非結構化，缺乏正式流程與治理機制。資料管理缺乏一致性，決策多以被動回應狀態為主。
2. **孤島階段 (Siloed)**。初步建立基礎架構與流程，但各自分散，一致性不足，治理存在缺口。
3. **標準化階段 (Standardized)**。建立明確治理程序、標準與營運模式，但合規落實尚不完整。
4. **管控階段 (Managed)**。資料架構流程已整合至整個組織，並會根據業務需求持續調整與更新。
5. **最佳化階段 (Optimized)**。資料架構已全面最佳化，並透過 KPI 等度量指標持續改進。

如需瞭解更多資訊，請造訪
sas.com/ai-impact

如需聯絡您當地的 SAS 辦事處，請造訪
sas.com/offices

SAS 以及 SAS Institute Inc. 的所有其他產品或服務名稱，均為 SAS Institute Inc. 在美國及其他國家的註冊商標或商標。® 標誌表示已於美國註冊。其他品牌及產品名稱均為其各自公司所擁有的商標。Copyright © 2026, SAS Institute Inc.保留所有權利。1685919_0926