



塑造社會 5.0： 建立對人工智慧 (AI) 帶來正向力量的信任

引領業界負責任地使用人工智慧的
四個步驟



前言



Harold Pradal, BSI 驗證部門總裁

2023 年被認為是人工智慧 (AI) 進入主流的關鍵時刻，相關報導的媒體曝光率增加了 286%¹。雖然 ChatGPT 成為了頭條新聞的焦點，但真正的 AI 故事要更深遠得多。

這項革命性技術正快速發展，有潛力推動以人為本的「社會 5.0」(Society 5.0)，在經濟和技術間取得平衡，解決社會問題²。這也提出一些問題：我們如何建立對 AI 的信任，以及需要哪些保護措施來確保 AI 以正面的方式塑造我們的未來。

在這個專題中，我們將通過 BSI 專家的視角，深入探討 AI 在現實世界中的影響，並參考來自九個國家的 10,000 名受訪者的觀點。對於那些持懷疑態度的人來說，AI 已經到來並將持續存在—38% 的人在日常工作中使用 AI，而在中國和印度，這一比例分別達到 70% 和 64%。到 2030 年，62% 的人預期其產業將使用 AI³。

在 BSI，我們致力於塑造技術和創新的影響，以造福個人、組織和社會。AI 處於這一切的核心，因為它有潛力成為強大的合作夥伴，改變生活，加速邁向更美好未來和永續世界的進程。

我們委託專家撰寫這些文章，旨在聚焦這一代人的機遇—了解得越透徹，我們就能越有效地利用 AI 的力量。無論是創造新的工作機會、改善病患治療、打擊現代奴役制，還是建立安全的全球食品系統，AI 都扮演著關鍵角色。

我們探討了在 AI 中建立數位信任的重要性，以及各國、政策制定者、組織和個人之間協作的關鍵角色，以釋放 AI 的真正潛力。此外，我們還關注迅速演變的法規，以確保一致性和確定性。

隨著 AI 從小型、封閉的環境跨越到工作和家庭中成為主流技術，這提供了一個革命性的機會來釋放多種益處—前提是我們擁有信任和信心。

AI 只是剛起步。在 BSI，我們很高興能與客戶合作，共同踏上這段旅程。我們很高興呈現這些文章，探索 AI 所提供的巨大潛力，以塑造社會 5.0，並通過創新實現永續的未來。

1. Signal search analysis 2022-2023.

2. [Society 5.0: The Fundamental Concept Of A Human-Centered Society](#), Open Business Council, Aug 2023

3. BSI Trust in AI Poll 2023



引領業界負責任地使用 人工智慧的四個步驟

作為一家使命驅動的組織，BSI 相信 AI 可以成為正向的力量，改變生活並產生積極影響，幫助創造更美好的社會，並加速邁向永續的世界。在這篇文章中，BSI 全球數位信任驗證總監 David Mudd 將探討產業如何塑造 AI 人工智慧的正向未來。





David Mudd

BSI 全球數位信任驗證總監

先進的科技帶來機會，也帶來關於負責任使用的疑問。根據安永 (EY) 的調查，三分之二的 CEO 認為 AI 是一股正向的力量，並認同其對當今與未來社會的正面影響。同樣地，根據 BSI 的人工智慧信任民調 (BSI Trust in AI Poll)¹，有三成 (29%) 的受訪者表示 AI 的首要任務是幫助人們減少對環境的影響，另有 28% 的人表示希望 AI 能改善醫療的診斷，五分之一 (22%) 的人則認為 AI 能促進社會公平，縮小不平等差距。

然而，同樣比例的 CEO 向安永表示，需要更多的努力來應對 AI 世界中的社會、道德與安全風險。而我們的調查也顯示，公眾對於使用 AI 存有不確定性。超過一半的人擔心 AI 帶來的風險將會加劇社會分歧。顯然，利用 AI 推動社會進步的機會是存在的。

- 產業擁有前所未有的機會，確保 AI 能為個人、組織、社會和未來帶來正面影響。
- 組織可以優先建立治理系統，打造值得信賴且負責任的 AI，解決偏見、道德和可解釋的結果 (explainable outcomes) 等問題。
- AI 可以被視為任何其他商業風險或不確定性，但最終它帶來的是巨大的機會。

22%

的人表示 AI 能讓社會更加公平並減少不平等現象。

產業的機會

數據顯示，88% 的 CEO 正將 AI 納入資本配置² 之中。對於希望釋放 AI 巨大潛力的產業領袖來說，我們的調查顯示出大眾對 AI 態度的複雜性，這也意味著建立公眾信任將是邁向成功數位未來的共同旅程中關鍵的一環。

在多元的全球經濟中，法規變動快速，要在 AI 這樣廣泛的議題上達成共識來建立信任，似乎是一項艱鉅的任務。同時，AI 是一個複雜的生態系統，包括 AI 模型、底層平台、終端應用、數據、硬體和網路，每一部分都可能來自不同的供應商，沒有單一組織能掌握所有組件。即使有主要整合者，也會在提供各組件的合作夥伴之間形成複雜的關係。

然而，產業擁有前所未有的機會，成為建設性的夥伴，通過跨市場和產業的協作，以全面實踐最佳實務，領導 AI 的負責任使用，並建立社會信任。以下是領導者可以採取的四個步驟。





1. 建立適當的治理系統

確保 AI 的道德使用對於實現其效益至關重要。例如，我們發現 55% 的人支持使用 AI 工具來診斷或治療自己，但前提是必須有嚴格的防護措施以確保患者數據的道德使用。

隨著 AI 能力快速發展³，建立適當的治理系統能讓我們在生活與社會的各個領域，不僅能善用 AI，更能充分發揮其潛力。

在進入 AI 部署的技術層面之前，領導者可以專注於建立一個可信賴且負責任的 AI 使用文化。設定清晰的基準可以讓組織有信心地前進。

良好的起點是先就期望的成果和投資報酬率達成共識，同時評估其對各利害關係人群體的影響，包括員工和整體組織。在此基礎上，組織應建立有效的治理流程，以平衡業務需求與這些影響及相關風險。導入[人工智慧管理系統 \(ISO/IEC 42001\)](#)，將有助應對這些挑戰。

所謂良好的治理，正如此標準所提供的指引，需要理解整體的商業環境—包括內部與外部。在此情境下，這可能意味著將 AI 策略與商業策略連結，並基於組織的整體願景和策略來制定 AI 相關政策。接下來的步驟可以包括建立結構和流程，以評估實施過程和相關風險，進行適當的風險緩解，並確保這與組織的風險偏好 (Risk Appetite) 和策略相一致。最終，若 AI 的推行與更廣泛的商業策略一致、風險已被考量，並具備適當的結構和控制來進行監管，AI 就能在組織內發揮正面影響。

建立並維持可信且負責任的 AI 開發與使用文化，並將其與商業策略結合，才能有效推動 AI 發揮正向影響力。和儘管各行各業的考量差異巨大，但良好的治理是共同的要素，能支撐產業的領導地位。

55%

的人支持使用 AI 工具進行診斷或治療，但前提是必須有嚴格的保障措施，確保患者數據的道德使用。

2. 從源頭考慮預存的偏見

產業也可以透過重視偏見及其緩解方式來發揮領導作用。這不僅僅是「AI 的問題」—我們每天做的決策中就有偏見⁴，例如誰被任命為高層、公司選擇的投資或供應商等。這些選擇並非憑空做出的。

然而，偏見在 AI 中是一項重要考量，且確實引發了大量討論⁵，例如關於 AI 搜尋引擎是否會提供性別刻板印象的結果⁶等問題。

關鍵在於，偏見的問題並非始於 AI 本身，而是來自其周圍的過程及所使用的資料集⁷。偏見可能來自於團隊的假設、資料中的錯誤，或是學習的方向設定。

正如美國一項研究所揭露臉部識別系統的偏見問題⁸，如果使用的資料有偏差，即使只是微弱跡象，也可能讓演算法在學習過程中加劇這種偏見，形成正向回饋循環，進一步放大偏見。

正如 [歐盟基本權利機構 \(EU Agency for Fundamental Rights\)](#) 所指出的⁹，「AI 的回饋循環可能導致偏見被放大，最終生成極端結果。當這類現象發生在執法等高風險應用中時，對公平性和準確性的影響尤其嚴重，可能引發重大的社會問題。」

這些都表明，從一開始就管理 AI 的開發過程是至關重要的。若能在起步階段做對，並留意每一個可能存在偏見的環節，將有助於我們充分發揮 AI 的潛力。

這同樣是產業發揮領導作用的機會，確保 AI 開發者在一開始就提出正確的問題，組建多元化團隊，並避免偏向任何特定群體。資誠 (PwC) 將此劃分為「建立治理與控制、團隊多元化以及持續監控」¹⁰。

對於那些選擇 AI 工具的使用者來說，謹慎的盡職調查是關鍵。我們可能無法完全消除偏見，但透過辨識偏見可能出現的地方及設置必要的防護措施，我們可以盡力確保 AI 的公平使用，造福所有人。



負責任的 AI 領導力是將商業風險視為機會，並清楚傳達所採取的緩解措施。

David Mudd





3. 在組織內部展開關於 AI 的對話

領導者也有機會在組織內建立對 AI 的信任，使其成為能夠正向塑造未來的工具。在確保環境安全及資料保護的前提下，組織可以更有信心地推進 AI 及其他新興技術專案。但同時需要認識到，使用基於 AI 的服務仍然會帶來一定程度的商業風險。

風險程度取決於組織是 AI 的被動使用者，還是積極參與 AI 服務開發，特別是當服務延伸到供應鏈，及 AI 的應用方式。但負責任的 AI 領導力應將這些風險視為成長的契機，並清楚傳達所採取的風險緩解措施，而不是將其視為發展的阻礙。

首先，可以考慮改變一些 AI 的相關用語，從「新穎且令人畏懼」轉為「業務常態的一部分」，需要由所有人（而不僅是數位團隊）謹慎管理。

無論組織是開發 AI 還是使用 AI，都應將其納入清晰的業務策略中，確保每個人都了解組織希望通過這項技術實現的目標，並能夠看到其帶來的個人、組織及社會效益。這一點尤為重要，特別是當下人們對 AI 的了解仍然有限—例如，62% 的人不知道像 Strava 這樣的健康與健身追蹤應用程式使用了 AI，而 50% 的人不了解自動化聊天機器人背後也運用了 AI。縮小這一知識鴻溝至關重要。

隨著 AI 日益融入商業活動，領導者可以提醒大家，我們曾經歷過類似的情況，例如網路或社交媒體—這些技術迅速普及，但並非所有人都具備使用它們的素養。現在，我們有機會應用從中學到的知識、網路安全及安全相關的經驗，確保每個人都能享受到 AI 所能帶來的益處。

58%

的人曾使用自動化的客戶服務聊天機器人，但其中只有 36% 的人，了解這些工具其實是基於人工智慧運作的。

4. 確保 AI 安全：信任與數位疫苗

建立 AI 信任的關鍵之一，是確保敏感資料能被妥善保護，且 AI 演算法不會遭到干擾。因此，AI 生態系統的安全性至關重要。

從這個角度來看，確保安全性就如同接種「數位疫苗」（digital vaccine）。如果惡意行為者是試圖入侵的病毒，那麼良好的安全實踐雖可能無法完全阻止「感染」，但能顯著降低感染的可能性，並在感染發生時大幅減輕影響。如果所有人都採取這些措施，將有機會將「病毒」趕出，並大幅減少其對社會的影響。組織可以透過確保自身的安全措施強大，來建立對 AI 使用的信任—換句話說，展現他們已完成「數位接種」。

這是一個值得領導者採納的洞察，特別是在 AI 更廣泛部署時。我們已見證，信任是技術進步的基石，因此，透過安全、道德且負責任的使用，以及穩健的管理來優先建立信任，能推動更積極的結果。

到 2030 年，全球近三分之二的人預期其所屬產業將每日使用 AI。AI 不再是理論，而是一項現實，並將徹底改變生活與工作的每一個面向。企業將處於這場變革的最前線。領導者現在擁有一個絕佳的機會¹¹，可以規劃 AI 的最佳應用方式，使其真正成為一股向善的力量，正向塑造社會的未來。

62%

到 2030 年，全球 62% 的人預期他們所屬的產業將每日使用 AI。

了解更多

隨著新技術和新功能不斷地推陳出新，機會與風險勢必隨之並存。BSI 多年來與全球組織合作，致力於建立對數位風險管理的信任，涵蓋資訊安全、網路安全及隱私標準、教育訓練和驗證服務，這些均基於被全球認可的 [數位信任國際標準](#) 之上。

參考資料

- 1 BSI partnered with Censuswide to survey 10,144 adults across nine markets (Australia, China, France, Germany, India, Japan, Netherlands, UK, and US) between 23rd and 29th August 2023
- 2 [The CEO Outlook Pulse - July 2023](#), EY, July 2023
- 3 [News Release: CEO confidence in artificial intelligence tempered by social, ethical and security risks](#), EY, September 2023
- 4 [AI has started to display 'compounding exponential' progress, founder of ChatGPT rival claims](#), Independent, February 2023
- 5 [Unconscious Bias, the Royal Society](#), accessed September 2023
- 6 [Open source data science: How to reduce bias in AI](#), World Economic Forum, October 2022
- 7 [Artificial Intelligence: examples of ethical dilemmas](#), UNESCO, April 2023
- 8 [Federal study confirms racial bias of many facial-recognition systems](#), casts doubt on their expanding use, the Washington Post, December 2019
- 9 [Bias in algorithms Artificial Intelligence and discrimination](#), the EU Agency for Fundamental Rights, December 2022
- 10 [Understanding algorithmic bias and how to build trust in AI](#), PWC, January 2022
- 11 [To Be a Responsible AI Leader, Focus on Being Responsible](#), MIT Sloan Management Review, September 2022



塑造 AI 信任，邁向永續未來

了解更多 BSI 「[塑造社會 5.0—AI 人工智慧](#)」研究報告

BSI 英國標準協會
台北市內湖區基湖路 37 號 2 樓
+886 2656 0333
bsigroup.com/zh-TW/

