

第 367 場工商講座活動紀要刊登網站

主辦單位：中華民國工商協進會

時 間：115 年 6 月 11 日(星期四)下午 2 時

地 點：台北國際會議中心 101AB 會議室

主 題：AI 時代的企業治理新局：從代理風險到護城河再造

主講人：資誠智能風險管理諮詢公司張晉瑞董事長

摘要：

一、AI 浪潮下的企業新課題：從效率工具到治理變數

張晉瑞董事長指出，生成式人工智慧自 ChatGPT 問世以來快速普及，短時間內便改變企業的工作模式與決策流程。過去企業導入 AI 多侷限於資料分析、自動化或輔助決策，如今生成式 AI 已能協助撰寫文件、程式開發、客服應答及知識管理，大幅提升工作效率。然而，AI 技術的快速發展，也讓企業治理面臨全新的挑戰。

(一) AI 的角色已從「提升效率」轉為「重塑營運」，串聯企業的知識管理、作業流程、企業管控與決策訂定。

1. 工具層

協助寫作、摘要、翻譯，人類仍是主要執行者。

2. 流程層

嵌入客服、法務、財會、採購，影響日常作業節點。

3. 營運層

連結資料、系統與外部工具，開始代理決策與執行行動。

(二) 企業 AI 風險地圖：AI 的複雜性從單點風險變成疊加

「風險地圖」，風險來源如下：

1. 模型風險：邏輯偏誤、效能漂移
2. 資料風險：隱私外洩、來源非法
3. 消保風險：演算法歧視、不公平待客
4. 韌性風險：過度依賴第三方訊息來源、資安攻擊
5. 法遵風險：違反主管機關指引與作業規範

AI 風險的複雜性來自多重風險同時發生，進而延伸出消保問題、資安事件，甚至造成重大聲譽毀損，不僅增加企業治理難度，也對企業之經營韌性與永續治理帶來嚴峻考驗。

二、AI 對企業競爭優勢造成的雙重影響

（一）AI 可能侵蝕原本優勢

AI 能夠快速取代大多數標準化、重複性高或基於既有知識的任務，例如：

1. 知識查找：一般的資料搜尋與查閱。
2. 資料處理：單純的資料整理與標準化分析。
3. 初階產出：基礎文件與內容生成。
4. 重複性流程與熟練時的速度：依賴重複流程與熟練度換取的速度優勢。

（二）AI 可能強化或放大的優勢

涉及複雜決策、跨領域範疇，AI 能協助企業做得更好，例如：

1. 深度洞察：需要專業經驗的產業洞察。
2. 複雜判斷：跨領域的風險評估與議題引導。
3. 信任資產：客戶的信任感與責任承擔。
4. 品質把關：高標準的專業品質控管。

5. 落地變革：實務上的專案落地與組織變革管理。

三、代理式 AI(Agentic AI)

(一) 什麼是代理式 AI

不只是會回答問題的 AI，而是能理解目標、規劃步驟、調用工具並執行任務的 AI，具備四大特徵：

目標導向	工具使用	持續運作	部分自主
不是只回應一句話， 而是完成一項任務	可連接系統、API、 資料庫、郵件、瀏 覽器	可記住上下文與歷 史行為	能自行規劃與採取 下一步動作

(二) 從「生成內容」到「自主行動」

生成式 AI 主要是「內容治理」，代理式 AI 則是「行為治理與控制治理」，比較如下：

面向	生成式AI	代理式AI
主要能力	生成文字/圖像/程式	規劃、調用工具、執行任務
與企業系統關係	多為輸入/輸出	可能直接連接企業系統
自主性	低到中	中到高
典型風險	幻覺、偏誤、資料外洩	權限濫用、行為失控、供應鏈風險
治理重點	模型、資料、輸出	身分、權限、工具、監控、可追溯性
董事會關心	用得對不對	能做什麼、誰授權、出事誰負責

(三) 代理式 AI(Agentic AI) 的 8 大風險

1. 目標誤解與任務偏離

AI 在處理複雜任務時，可能因指令歧異而努力地完成「錯誤目標」。

2. 權限失控、身分與授權不清

代理式 AI 取得過大或不必要的權限。

3. 提示注入升級為行為操控

惡意指令隱藏在正常文件或網頁中，誘導 Agent 違規。

4. 工具濫用

瀏覽器、郵件、終端、檔案系統等工具被當作執行網路攻擊或外洩資料的媒介。

5. 記憶污染

長期記憶能力若被惡意或錯誤資訊污染，將會持續影響未來的決策品質。

6. 長鏈決策難以控管

自行規劃並執行多步驟的複雜任務時，過程太冗雜，難以追蹤與稽核。

7. 外掛與技能供應鏈風險

引進第三方開發的 AI 功能模組或外掛套件時，可能導致安全漏洞。

8. Shadow AI Agent

未經核准的代理式 AI 在組織內部私下運作，可能造成機密外洩。

(四) OpenClaw 現象與啟示

1. OpenClaw 介紹

OpenClaw 是 2026 年初在全球引發熱潮的代理式 AI 軟體，不同於一般只能聊天的 AI 機器人，OpenClaw 能常駐系統後台，自主規劃並執行複雜任務，執行如操作瀏覽器、自動讀寫檔案、自動回覆郵件等任務。

2. 引發熱潮的原因與代表意義

開源、低門檻的部署條件，加上可透過 WhatsApp 等常見的日常訊息工具驅動，讓一般使用者也能輕易接觸到可以自己執行的代理式 AI，OpenClaw 被視為代理式 AI 走向大眾化的重要案例。

3. 後續引發的爭議

(1)遠端程式執行漏洞

攻擊者若利用漏洞，代理式 AI 可能被惡意控制，操控 AI 執行非法操作。

(2)隱私未受保護

系統未做好隔離或權限控管，導致敏感的 AI 後台或資料庫可被任意存取。

(3)系統層級權限問題

代理式 AI 通常擁有讀寫資料庫或執行系統命令的權限，被濫用時破壞力極強。

(4)責任歸屬模糊

當 AI 產生錯誤決策或造成損失時，缺乏明確的法律與內部追責機制。

四、企業 AI 治理實務框架

(一) 三層監督模型：策略、風險、韌性

企業在追求 AI 創新與獲利的同時，必須具備相應的風險控制能力與營運韌性，確保每一次的技術創新與資金投入，都能達成企業的商業目標並贏得市場信任才能實現永續且負責任的 AI 轉型。



(二) AI 治理不是煞車，而是企業治理護城河

1. AI 策略應該是能力再造，而非工具導入

有些企業將 AI 當成一般辦公軟體工具來導入，只追求短期行政效率的提升；但 AI 策略的導入可以改變企業的業務流程、決策模式與創新方式。

2. 代理式 AI 的風險是行為風險，而非僅只於說錯話

傳統生成式 AI 的主要風險是輸入錯誤提示，導致輸出不正確的結果；代理式 AI 權限更大，此時的風險躍升為「行為風險」，即 AI 可能在執行任務時做出錯誤的決策或越權行為。

3. 治理能力，將成為企業 AI 規模化的護城河

企業若具備完善的治理能力，才能讓企業在符合法規、安全無虞的狀況下使用 AI，這是拉開與競爭對手差距的競爭優勢。

張董事長於結論指出，AI 時代帶來的不只是技術革新，更是企業治理思維的全面轉型。面對代理式 AI 快速發

展，企業應以治理為基礎，建立兼顧效率、風險與責任的管理制度，讓 AI 成為驅動企業創新與永續成長的重要引擎。只有將 AI 治理融入企業文化與經營策略，才能在數位轉型浪潮中建立真正難以複製的競爭優勢，持續提升企業競爭力。

本講座會後開放意見交流，與會來賓踴躍提問，現場氣氛熱烈，現場與線上共計逾 110 位企業界人士參加。



第 367 場工商講座一隅



邱一徹秘書長(右)及張晉瑞董事長(左)共同主持 Q&A 互動交流。