# 公司治理視角下
# ──從智慧財產的觀點談企業導入生成式AI及AI代理之著作權教戰守則

陳家駿＊

## 壹、生成式AI之下著作權風險與治理需求

### 一、AI最新發展趨勢：從生成式AI到AI代理

隨著生成式人工智慧（Generative AI，下稱生成式AI）之快速普及，此一發展趨勢已從創新工具，轉變為企業營運之基礎設施。接著從2025年起，AI再由單向內容產出的生成式AI──「資訊生成者」，演進至具備自主規劃與執行能力的「行為執行者」──AI代理（AI agents，或稱代理AI（Agentic AI），以下統稱AI代理）[1]，使AI系統從被動工具，進化為具有持續互動與決策能力之行為主體，而不再局限於執行單一、特定之任務，展現出橫跨多元領域的高度通用性與可擴展性。到了2026年則又有號稱「AI龍蝦」的OpenClaw爆紅、再發展到Moltbook等令人目不暇給的驚人發展[2]！

企業於日常營運中，無論在內容生成、文案與行銷設計、產品開發，抑或決策支援等領域，已逐步對生成式AI或AI代理（以下合稱AI）技術形成高度依賴，此一發展趨勢具有不可逆特性，不論是科技、製造、零售、金融或醫療等各類產業與服務業，無不將AI視為推動數位轉型、強化競爭優勢並開拓新興市場機會之核心驅動力，而非僅是輔助性工具。

### 二、AI引發前所未有之著作權法衝擊和法律風險

伴隨此一技術之快速擴散與高度競逐，市

---

＊本文作者係台灣資訊智慧財產權協會理事長

（本文所有網路參考文獻最後瀏覽日皆為2026年4月30日）

註1：參陳家駿&許正乾，〈從生成式AI到AI代理、AI代理到代理AI〉，國家實驗研究院科技產業資訊室iKnow，
https://iknow.stpi.niar.org.tw/post/Read.aspx?PostID=22206

註2：當使用者下達目標後，OpenClaw不再只是提供建議，而是能主動規劃路徑並接管電腦桌面與瀏覽器。它的最終產出是具體的行動結果，例如郵件發送、機票訂定，真正實現了從「紙上談兵」到「實務執行」的跨越。請參陳家駿&許正乾，〈OpenClaw在全球掀起AI代理革命系列一-三從Moltbot、OpenClaw談到Moltbook〉，
https://iknow.stpi.niar.org.tw/post/Read.aspx?PostID=22867

場形成一股強烈的「技術焦慮」與「錯失恐懼」，促使企業競相投入資源，深怕在轉型的浪潮中失去先機。在此種心理與市場壓力下，造成企業深恐於這波科技轉型浪潮中被邊緣化，遂倉促投入大量資源進行AI部署與應用，而未能充分評估風險、確立治理架構與法遵要求，據麻省理工學院MIT 2025年之報告The GenAI Divide: State of AI in Business，95%的企業生成式AI試點計畫失敗[3]，故導入AI需謹慎且方法需正確。

AI之最大法律風險之一，來自其訓練資料之使用方式。由於AI模型常透過網路，抓取大量受著作權保護的資料進行訓練，此是否構成侵權已成為當前法律爭議核心[4]。隨著AI成為企業營運之關鍵工具，其雖帶來效率與創造力提升，但同時也引發前所未有的著作權法衝擊與法律風險，特別是在AI生成文字、圖像、影音等內容時，侵權疑慮更形敏銳：諸如訓練過程中之重製、輸出的結果是否與原作產生實質近似等爭議[5]。有鑑於此，本文爰從公司治理視角，就企業在導入AI面臨之著作權侵權風險，提出具體之法律防範機制，供業界參考。

## 三、企業導入AI之著作權風險轉型與治理需求

不同於傳統模式，生成式AI透過資料訓練與提示指令（prompt）生成內容，其輸出並非直接來自人類創作，而是模型基於統計推論所產生；而AI代理更能在無人為干預下自主決策與行動，這也導致輸出難以歸屬於特定主體，進一步加劇責任歸屬的不確定性，形成所謂「責任落差」。在此背景下，企業於營運流程中導入AI時，如何避免侵權已成為當前極具挑戰性的智財權議題[6]。這也意味著，企業應將重點放在預測與控管AI自主性所帶來的風險上[7]，尤其是透過公司治理，確保員工在日常業務中安全、合法地使用AI工

註3：Yahoo!finance, MIT report: 95% of generative AI pilots at companies are failing, https://finance.yahoo.com/news/mit-report-95-generative-ai-105412686.html

註4：Hong Wu (2024), Copyright protection during the training stage of generative AI. Computer Law & Security Review, 55. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4990383

註5：這些問題彼此間環環相扣，若以此引入「生成式AI供應鏈」：一相互關聯的階段集合，將訓練資料轉化為生成結果。將生成式AI分解其組成階段，使其能追蹤上游技術設計對下游應用的影響，並評估在複雜的侵權行為發生時，誰應承擔責任。參Katherine Lee, A. Feder Cooper, James Grimmelmann (2024, v.2), Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain, https://arxiv.org/abs/2309.08133

註6：Lee, et. al., id.; João P. Quintais (2025), Generative AI, copyright and the AI Act, Computer Law & Security Review, https://www.sciencedirect.com/science/article/pii/S0267364925000020?via%3Dihub

註7：代理AI的自主性高於以往的AI，如從英國法中的三個主要層面，來檢視是否存在此類責任缺口：代理人責任、僱員及類似關係人的替代責任以及動物責任，依次將代理AI的使用與這三個領域類比，則除非存在特殊情況，否則使用者仍將對AI的行為負責。參Chris Reed (2025), Autonomy, responsibility

具，避免誤觸法網。因此，建構有效的侵權風險治理機制，已成為企業導入AI的關鍵課題。以下先從美國實務觀點，看生成式AI是否可能構成合理使用。

## 貳、美國案例法之啟示：生成式AI之轉化性合理使用原則

### 一、生成式AI訴訟之法律爭點

生成式AI需透過網路大量爬取他人著作，建構其大型語言模型（LLM）。此類未經授權的複製行為，在美國已引發數十件著作權侵權訴訟[8]，涉及的模型包括GPT、Gemini、Claude、Llama等知名基於Transformer架構。儘管其中大多未取得授權，惟面對侵權指控時，這些被告無不主張其AI運作具有「轉化性」（transformativeness），得構成「合理使用」（fair use），以此作為抗辯基礎。

目前全球相關之侵權訴訟中，多聚焦於兩大爭點：原告多主張其作品未經授權，即遭被告將其納入AI訓練，其爬取資料與訓練中重製行為屬於非法重製；其次，原告指控被告所生成之內容，構成直接或間接侵害其著作權（代理侵害與輔助侵害）[9]。這些案件主要涉及兩項核心問題：其一，AI訓練行為是否屬於合理使用；其二，AI生成內容是否與原作品構成「實質相似」。針對前一點，美國聯邦地院目前已做出實體判決如下：1. *Bartz, Graeber & Johnson v. Anthropic*案（北加州地院6/23/25即席判決）；2.*Kadrey, Silverman, Golden v. Meta*案（北加州地院6/25/25即席判決）[10]。

### 二、轉化性合理使用之核心案例──Bartz與Kadrey案

在著作權法體系中，合理使用原則為AI案件之關鍵防禦機制。由以上二起AI侵權案的判決可知，針對AI的訓練能否主張合理使用，雖尚無定論但也初步勾勒出一些原則。第一件Bartz案法官認為生成式AI雖極具轉化性，但訓練用的素材不可以是盜版，而法官做出即席判決後沒多久二造即和解，並構成美國史上金額最大的和解金額[11]。但另一件*Kadrey*案法官則認為訓練的素材，是否取自盜版並不那麼重要，而是將焦點置於：在訓練

and agentic AI, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5598471

註8：Norton Rose Fulbright (2026), AI in litigation series: An update on AI copyright cases in 2026, https://www.nortonrosefulbright.com/en/knowledge/publications/ce8eaa5f/ai-in-litigation-series-an-update-on-ai-copyright-cases-in-2026

註9：Lee, et. al., *supra* note 5；另參陳家駿．許正乾．林宜柔（2024），《AI/ChatGPT v.智慧財產權──美國生成式AI》，第190-200頁，元照。

註10：參陳家駿（2026），《全球生成式AI著作權訴訟攻防戰》，2版，第533-559頁，元照。

註11：法官於2025年9月初步核准15億美元之和解協議。Authors Guild (2026), Bartz v. Anthropic Settlement: What Authors Need to Know, https://authorsguild.org/advocacy/artificial-intelligence/what-authors-need-to-know-about-the-anthropic-settlement/

階段，資料使用是否具有必要之轉化性；在輸出階段，生成內容對原作品是否構成市場之「間接替代」效果[12]，尤其不能產生「市場稀釋」（market dilution）的效果，而對權利人可能產生「大規模淹沒市場」（potential to flood the market）的損害[13]。

## 參、台灣法制之現況與對應方向

### 一、現行著作權法之適用困境

我國雖已於2025年底通過人工智慧基本法，但就智財權僅做政策宣示而並無實質規範[14]；而稍早亦有行政院及所屬機關使用生成式AI參考指引，但主要針對政府機關內部使用，並非產業發展所需之著作權議題；另外智慧財產局於相關函釋中則宣示性地認為，AI訓練過程涉及重製行為時，除合理使用外均應取得授權，但其未進一步闡述何種情形可能構成合理使用，更未提出具體之判斷標準。

### 二、我國迄今唯一的AI侵權刑事案──中央通訊社案

再談到台灣著作權的侵害刑責，亦會導致AI發展的寒蟬效應！此觀台灣唯一涉及AI的著作侵權──中央通訊社（下稱中央社）案即明。據報導，被告博士生是開源繁體中文語料集「FineWeb2-ZHTW」的志工，其本無意訓練任何AI模型，只是看到開發者在Hugging Face平台釋出對話式語料集，供其他開發者使用，但鑑於繁體中文資料占比極低，便想填補這塊缺口[15]。於是其未經授權擅自擷取中央社（可能還有其他來源）的新聞內容，直接放入該語料集中。由於該語料集本來就是開源可下載，推測中央社可能透過Hugging Face或GitHub等平台查看，發現資料集中出現自家新聞連結，基於將大量相關的新聞內容放到語料庫中，這本身在外觀上構成著作權複製，中央社遂於2025年7月對該博士生提起刑事告訴。

嚴格來說，本案不算是AI訓練所導致的著作權紛爭，因為僅是複製貼上而未進入AI訓練階段。若被告確實使用素材來訓練AI建立模型，或許有機會檢視其是否構成合理使用，甚至是否具有公共利益之轉化價值。但這位博士生縱使出於善意，最終仍付出若干代價才換得和解。此案凸顯的是，台灣著作權刑事處罰的強大威嚇力──試想，OpenAI的奧特曼、META的祖克伯或Grok的馬斯克，這些AI巨擘若在台灣不知已揹上多少刑案在身。由此可見，光是訓練前爬取資料的階段，行為人就得承受刑事追訴的壓力，這明

註12：Wu, *supra* note 4.

註13：陳家駿，莊弘鈺等（2025），《AI世紀訴訟及人工智慧法律15講》，第261-269頁，華藝。

註14：人工智慧基本法第13條第2項：政府應致力提升我國人工智慧使用資料之品質與數量，確保訓練及產出結果足以展現國家多元文化價值與維護智慧財產權。

註15：本案事實參謝馥伊（2026），〈和解案後，未解的台灣主權AI語料困局：開發者和內容方能否終結授權衝突？〉
https://www.twreporter.org/a/taiwan-sovereign-ai-zhtw-llm-copyright-conflict

確顯示：即便非商業目的之資料整理與分享，在我國現行法制下仍有觸法風險。

## 肆、三階層治理模型建構：著作權風險控管與合規防範機制

以上在了解目前國際間AI的司法環境，以及我國現時法制所處的情況下，針對AI導入需求，企業只能建構出適合自己的一套公司治理之制度化管理模式。

### 一、從風險控管到責任防禦之制度化建構──三階層治理模型

企業首先應將AI之公司治理，從抽象原則轉化為具體可執行之「標準作業程序」（SOP）。此一制度化的核心，在於將風險控管措施予以具體程序化，使其不僅能在日常營運中發揮預防效果，更能於法律爭議發生時，作為企業已善盡合理注意義務之客觀證據。雖然機制之建立，不可避免將增加行政與合規成本，然而相較於潛在之高額賠償責任與刑事風險，此種事前之投入可成為企業合理預防風險之關鍵依據。

針對此，企業得參考建構一種「三階層治理模型」，作為導入AI之架構。第一層為「前端預防機制」，包括員工教育訓練與提示指令規範，讓使用者在操作階段即避免侵權風險；第二層為「中端控制機制」，包括生成內容審查與內部管理流程，以確保輸出內容合法；第三層為「後端防禦機制」，包括紀錄保存與法律意見書，以作為事後責任防禦依據。

### 二、「前端預防機制」：建立法遵文化與AI使用準則

首先，在「前端預防機制」方面，企業應從「人」出發，對同仁施予著作權法的教育訓練，不定期舉辦關於AI合規性的課程，特別聚焦於AI生成內容可能涉及之風險，使其充分了解如何在日常業務中避免侵權，例如前述之中央社案中被告之教訓。且此類訓練不可留於形式，而應透過實際案例與操作情境，讓員工能具體理解何種使用方式可能導致侵權，並培養其在日常工作中對可能違法行為的判斷能力。透過教育訓練，企業得以在組織內部建立法遵文化，降低因個人操作不當所導致之法律風險。

其次，企業應制定「AI使用準則」，作為員工之行為依據。此不僅應以條文方式清楚載明，公司對智財權保護之基本立場與態度、AI工具使用之範圍與限制，並應委託專業律師協助審閱與撰擬，以確保其法律效力。此外，該準則除應做為入職之聘僱合約附件外，還應同時以紙本與電子檔發布，要求同仁置於工作場所醒目之處，以及電腦或行動裝置上以便隨時查閱遵循。

此外，在AI軟體的選用上，企業亦應建立嚴謹之採購政策，儘量避免使用來源不明或條款不清之AI工具，合法採購授權方式清楚之產品，並嚴格遵守供應商所訂定之使用條款與限制。此一措施不僅有助於降低因工具本身所引發之法律風險，亦可避免因違反使用條款而產生之民事責任。再者，市面上已有生成式AI工具提供者承諾，如遇有因使用產品而造成侵害他人時，只要遵守其使用條款，就提供使用者相關之索賠（indemnification）

或協助解決，當然也應優先採用。總之，AI工具之選擇本身，亦構成企業風險管理之重要一環。

### （一）AI使用準則核心：建構「指令提示」操作手冊

在「AI使用準則」之治理措施中，應將「提示指令」的使用當作AI使用之核心，這是由於AI的輸出結果，高度依賴使用者之提示，因此提示設計本身即構成影響侵權風險之關鍵因素。企業可製作「Prompt操作手冊」規範員工之提示行為，此種作法不僅可降低侵權風險，亦有助於建立人類創作之參與程度，進而強化防範著作權侵害之發生。

基於此，企業在操作手冊中，宜詳細規範員工在與AI互動時之行為模式。其基本原則在於建立防火牆，避免直接要求AI就特定之作品或創作者，生成相關的內容；因為若任由員工直接要求AI模仿或輸出特定作品、指定某一作者或其風格，即可能導致生成內容與既有著作造成實質相似，而有構成侵權之風險。因此，企業應明確其提示規範，使員工避免使用具體指涉特定著作之指令，而改採抽象化或描述性之提示方式。

### （二）指令提示與人機協作：風險控管之核心

其次，員工在使用AI時，切勿於簡單輸入提示後，看到輸出內容就直接採用。取而代之的，應由同仁先做好規劃與創作框架或設計構想，最好是先自行撰擬初步的內容，再透過向AI再三地「循循善誘」進行反覆多輪的追問、選擇、修改、調整與安排，將AI回饋的內容融入自己的創意與設計中，然後再自行增添修改來逐步深化，如此透過反覆提示與修正，逐步將AI生成內容融入自身創意之中最後才定案。而非單純機械式地在前幾輪就直接採用其內容，此係強調「人機協作」的重要而非完全依賴AI的輸出。

另一個思維是，善用AI提供概念或構想式之啟發，詢問時可多點醒AI提供思考方向或脈絡之建議，而不必一味地執著於要求AI產出內容，因為基於著作權法最基本之「概念與表達二分法」原則，概念、思想等永遠是可以合法援用的，即便在AI吐出大量內容後，不必見獵心喜，而是從中歸納、提取得以襲用的點子，然後再用自己的表達方式另產出內容，以構築一道堅強的防火牆。

### （三）「人機協作」可符合人類作者要件享有著作權

重要的是，不得以「AI取代創作」，而應透過上述「先人後機」、「人機協作」、「人工再優化」之模式，在反覆迭代的過程之後，實質上達到「人工覆蓋」（Human-in-the-Loop）的終局結果，此在法律上的重要意義，透過此種制度化之規範，可降低侵權風險有助於強化法律上的防禦。再者，依美國最近之*Thaler v. Perlmutter*案，實質上確立了「完全由AI生成、無人類參與之作品不受著作權保護」的司法見解[16]，因此「人機協

註16：依美國著作權局於2023年指引及2025年報告所示，著作權之保護以「人類著作」（human authorship）為前提，僅及於源自人類創造力之表達。對於完全由AI生成且欠缺人類創作之內容，原則上不具保護資格。惟若人類對AI生成成果之選擇、協調或編排，或其他具創造性之介入達到一定程度，該人類貢獻仍可能受保護，但須依個案具體判斷其創作性與控制程度。
https://copyrightalliance.org/ai-report-part-2-copyrightability/?utm_source=chatgpt.com

作」與「人工覆核」運作之另一項重要意義，係經過此種人工反覆提問與人為加料之過程（以人類產出的內容為主），在法律上還可作為判斷可受保護之重要依據，因為人做出的內容才可取得著作權。

綜上所述，企業建構一套完整且可驗證之標準作業程序。此一制度不僅具有預防侵權之功能，更在法律爭議發生時，成為企業證明其已善盡管理責任之關鍵依據。從公司治理之觀點而言，SOP之建立不僅是風險控管工具，更是企業在AI時代維持合法性與競爭力之核心制度基礎。

## 三、「中端控制機制」：從風險控管到責任防禦

接著在「中端控制機制」方面，為進一步強化風險控管，應以制度化治理之可驗證性來進行內控管制。

### （一）「資料治理前置化」策略：從來源合法性到過濾機制

企業應在模型訓練之前即確保資料來源合法性，建立資料來源查核制度，並對資料進行授權確認與風險分類，以避免使用來源不明或明顯侵權之資料。而在技術層面上，企業可針對業務需求鎖定幾個領域，透過自訂之自動化監控系統，檢測生成內容之相似性或異常輸出，以即時降低侵權風險。更可參考「多層次資料過濾機制」，透過內容識別技術與資料庫比對，以降低侵權資料進入訓練集之可能性[17]，以上措施不必技術上完全到位，重點是建構出一套可執行的做法。

### （二）模型設計與技術控制：避免「記憶與重現」侵權

除了資料來源問題外，生成式AI本身的技術特性亦可能導致侵權風險，基於AI模型可能出現「記憶效應」（memorization），即在特定情況下重現訓練資料內容，從而構成侵權之風險[18]。因此，企業在模型設計階段，應導入技術性控制措施，以降低此類風險。例如，透過內容過濾機制、操作監控與濫用偵測、去識別化（de-identification）、資料去重複（deduplication）與生成結果過濾機制，避免模型輸出與原作品有相似之內容。

此外，生成式AI之輸出結果，在特定情況下有可能會「再現」原訓練資料，例如基於訓練演算法像是「過度擬合」（Overfitting）[19]、解碼演算法、當條件機率分布在特定點之取樣結果時，或當提示語包含極罕見、獨特或具高

註17：Mariia Kyrychenko, et al., Copyright in AI Pre-Training Data Filtering: Regulatory Landscape and Mitigation Strategies, https://www.researchgate.net/publication/398269082_Copyright_in_AI_Pre-Training_Data_Filtering_Regulatory_Landscape_and_Mitigation_Strategies

註18：Sebastian Stober & Tim W. Dornis (2026 v.2), Generative AI Training and Copyright Law, https://arxiv.org/abs/2502.15858.

註19：當訓練資料中某些樣本出現的頻率極高，或模型容量過大、正則化不足時（用來防止模型Overfitting的技術），模型可能會「記住」而非「學到」模式。此時，給定一個與該記憶樣本高度相關的提示（甚至只是隨機取樣），模型就可能直接複製或近乎逐字逐句地再現該樣本。參陳家駿（2022），《AI人工智能vs智慧財產權》，2版，第210-212頁，元照。

度針對性之特定指涉的短語、名稱或描述等情況時，甚或是這些特徵組合在訓練資料中，可能就只對應到極少數或唯一的樣本時，都可能會產生具有實質相似性之輸出，因此，企業應針對高風險應用場景進行控管[20]。

以上之過濾機制與比對，企業可透過例如商標或專利檢索系統，或是特定領域例如新聞類資料庫，自行開發演算法或採購外部服務系統來執行，重點不在於是否可以完全做到百分之百的效果，而是在目前技術可能之限制的情況下，業者已盡力採取一定的措施來避免風險的確切作為，盡到注意義務。

**（三）輸出內容監控：從生成自由到法律審查**

生成式AI之重大風險，在於其輸出內容可能構成對既有著作之侵權，而AI生成內容在某些情況下，不排除可能具有市場替代效果，進而影響著作權人之經濟利益。因此，企業若未對輸出內容進行監控，可能面臨直接或間接侵權（如輔助、引誘與代理侵權）責任。在此情況下，企業得建立輸出審查機制，包括內容比對系統與人工審查流程，以識別可能侵權之內容。此外，企業亦應限制AI代理之自動發布行為，例如設定發布門檻或審核機制，以避免侵權內容自動擴散。值得注意的是，隨著AI代理具備持續互動能力，其生成內容可能來自多次交互過程，導致侵權責任難以追溯。

## 四、「後端防禦機制」：紀錄保存並善用外部法律意見

最後在「後端防禦機制」方面，應紀錄保存並備妥法律意見書，以作為事後責任防禦依據。

**（一）紀錄保存與可追溯性之法律設計**

在實務運作中，紀錄保存為企業防範法律風險之核心工具。尤其就對外提供客戶做為投放市場高度曝光在大眾之重要內容時，此一機制特別重要，因其直接關係到企業是否須負責。而建立完整之AI操作紀錄，包括提示內容、生成結果以及人工修改過程，使整體創作流程具備高度可追溯性（traceability）。進一步言，企業建立資料可追溯性機制，得使其重要之訓練資料均可回溯其來源，此不僅有助於降低侵權風險，亦可在發生爭議時提供法律防禦依據。

雖然，此類紀錄之實施將增加一定行政負擔，但在爭議發生時，卻可作為企業已盡合理注意義務之重要證據。反之，針對關鍵項目若企業缺乏相關AI紀錄，責難以證明其已採取適當防範措施，從而提高被認定為過失或疏忽之風險。因此，企業應利用AI代理來建立完整之生成紀錄（log）系統，以確保內容生成過程可被追蹤與檢驗，該紀錄機制為企業防範法律風險之最後防線。透過保存提示內容、生成結果與人工修改過程，企業可在爭議發生時證明其已採取合理措施，並無侵權故意或過失；此一機制亦可降低法院認定歸責之可能性。

**（二）法律意見書之運用**

在法律層面，針對上述企業重要之AI產出內容，可委由專業律師進行風險評估，並出具法律意見書，確認企業之操作模式符合現行法規與司法見解。此類法律意見書具有雙

註20：*New York Times v. Microsoft, OpenAI*, 同註10，第187-216頁；Wu, *supra* note 4.

重功能：一方面，可作為排除刑事責任中「侵權故意」之重要證據；另一方面，在民事責任判斷上，可證明企業已積極履行善良管理人之注意義務，從而降低或免除賠償責任。然而，律師在出具意見書時，應有堅強之國內外司法實務做為論述之參考依據，例如美國法院近期針對AI訓練與合理使用判斷，尤其是美國最高法院之判例[21]，逐步強調「轉化性使用」之重要參考；英國法院對AI模型本身是否構成侵權提出見解[22]；此等比較法觀點，再加上我國法院與主管機關之見解[23]，均可作為企業並而無意侵權之重要參考。

## 伍、RAG與TDM之運用應避免著作侵權

### 一、「檢索增強生成」之應用

自生成式AI普及以來，「檢索增強生成」（Retrieval-Augmented Generation, RAG）已成為LLM常用的關鍵架構之一。RAG是一種結合「外部知識檢索」與「語言生成能力」的系統設計，其目的在於補強模型於知識更新、即時性與事實依據方面之不足，進而提升輸出品質與可信度。此一架構不僅適用於大模型，亦可應用於SLM中小模型。

一般而言，LLM係透過大量語料進行預先訓練，在訓練完成後其內容即大致固定，難以即時反映後續發生之事件或最新資訊。雖然可透過再訓練或微調（fine-tuning）更新模型，但往往涉及高昂之算力成本、時間及能源之消耗，因此不適合短期內頻繁更新。更何況LLM在生成內容時，係基於概率分佈進行語詞預測生成文本，並不具備事實之驗證機制，導致產生一般熟知之「幻覺」（hallucination）現象，生出不符事實或欠缺依據的錯誤資訊[24]（即一本正經地胡說）。基於此RAG提供有效之補強：當使用者提出查詢時，系統先透過檢索模組（例如向量搜尋、語意檢索或傳統全文檢索），自外部知識來源（如文件庫、資料庫或即時資料）擷取相關內容，再將該等資訊作為上下文提供予LLM，進而引導其生成可靠回應[25]。

---

註21：陳家駿等，同註9，第203-212頁。

註22：在*Getty Images v. Stability AI*案中英國法院於2025年11月判定，Stable Diffusion的AI模型權重，從未以任何形式包含、儲存或重製任何Getty受著作權保護之圖片副本。同註10，第368-397頁。

註23：我國法院雖尚無AI相關判決，但仍應找出可予類推之法理；其他如智慧財產局智字第11252800520號、電子郵件第1111031、1111212、1120220、1120317、1121229、1140310、1140522c、1140625、1140829、1141017b號函等。

註24：RAG技術參Yunfan Gao,Yun Xiong, et. al.(2024 v5), Retrieval-Augmented Generation for Large Language Models: A Survey, https://arxiv.org/abs/2312.10997.

註25：AWS，〈什麼是RAG（檢索增強生成）？〉https://aws.amazon.com/tw/what-is/retrieval-augmented-generation/; Google Cloud, 什麼是檢索增強生成（RAG）？https://cloud.google.com/use-cases/retrieval-augmented-generation?hl=zh-TW

## 二、RAG涉及著作權侵害的風險

基於RAG技術能提升AI回答的品質，部分業者的AI工具便使用RAG來彰顯其「可溯源性」與「引用顯示」功能，在回應中向使用者呈現原始資料片段、標題或來源連結，藉此提高其可信度與透明度。但這也帶來比傳統AI模型更直接的著作權與商標之侵權風險。目前，美國已有多件RAG技術相關的著作權與商標侵害訴訟，凸顯相關法律風險：*Dow Jones, Wall Street Journal and New York Post v. Perplexity AI；Advance Local Media v. Cohere*[26]。

企業在AI系統進入實際運作後，特別是在導入RAG、或具備自動瀏覽與外部資料擷取能力之AI代理系統時，其所面臨之法律問題更將產生質變。實務上，我國企業目前雖未自行開發如GPT等大型語言模型，但業界已廣泛利用既有模型進行客製化應用，包括微調、提示工程（prompt engineering）及RAG架構的整合利用。特別是在企業內部知識管理、客服系統、資料檢索或商業分析等場景中，RAG已成為常見的技術選擇。

然而，與直接對外提供的生成式AI工具（如GPT）不同，業者運用RAG嵌入系統內，但使用RAG後模型生成的文字內容，可能高度接近原始文件；或是標示出原作的商標，從而導致智財侵權風險。因此，業者運用RAG後應對輸出內容進行上述之檢核與調整，尤其是將這些內容提供給客戶對外使用時，風險更為顯著，業者應特別留意以避免觸法。

## 三、文字與資料探勘有別於生成式AI

其次，業者如要進行「文字與資料探勘」（Text and Data Mining, TDM）[27]，也應特別注意，由於我國目前並未像歐盟對此有明文之責任豁免規範，基於其本質上是在做科學分析前必定有複製行為，因此務必審慎結合我國著作權法嘗試找出解方[28]，並可參酌德國漢堡法院於2024年9月，在*Kneschke v. LAION e.V.*（*Large-scale AI Open Network*）案中，認定研究機構得允許其為科學研究目的進行複製，而做出被告AI訓練非著作侵權的判決[29]。

---

註26：其他如*Ziff Davis v. OpenAI; Encyclopaedia Britannica v. Perplexity AI; The New York Times v. Perplexity AI; Chicago Tribune v. Perplexity AI*、*Encyclopaedia Britannica v. OpenAI*等案。

註27：係指為識別資料中的模式、趨勢、關聯性及知識萃取，而對數位或數位化資訊進行自動化分析之過程。其核心在於透過電腦技術，自機器可讀之資料來源中擷取資訊，並運用資訊檢索與資料處理方法，將原本非結構化或半結構化之文本內容，轉換為可分析之結構化資料。在此過程中因其涉及資料之重製與利用，所以在著作權法上引發適法性問題。GOV UK, https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation

註28：我國著作權法第52條雖規定：「為報導、評論、教學、研究或其他正當目的之必要，在合理範圍內，得引用已公開發表之著作。」但其中因為有「合理範圍內」之限制，所以另外尚需搭配著作權法第65條來營造合法之適用空間。

註29：同註10，第436-468頁。

但即使如歐盟有責任豁免規定[30]，但也並非是可擴張適用，仍要看資料是否用於科學或學術研究目的之合法存取、權利人是否明確保留權利，以及使用之目的是否符合規範而定，企業不能過度樂觀依賴該豁免條款，尤其是不能以TDM為藉口，卻越雷池一步進入到生成式AI之運作。因為TDM之重點，完全不著眼於具著作權內容之資訊本身，而是要發掘出該資訊背後經電腦轉換所衍生的模式、趨勢、關聯性及知識，更遑論雖然生成式AI的前導階段，很可能會進行TDM的工作，但TDM本即不需進行與其互為獨立之生成式AI作業。也因此，德國慕尼黑地院才在*GEMA v. OpenAI*一案中[31]，於2025年11月判決OpenAI之自動生成GPT工具，不得適用TDM責任豁免。

## 陸、法律責任：民事責任與刑事責任之管控

企業在導入AI時，將同時面對民事與刑事責任不同之法律風險。在解釋上，我國法院固然可能參考美國合理使用原則，特別是轉化性使用之概念。然而，基於法律體系結構不同，該項原則之適用仍存在不確定性。因此，企業不應完全依賴合理使用作為風險防禦，而應採取相對保守之合規策略[32]。

### 一、刑事責任風險管控──排除侵權故意之執行措施

在刑事責任方面，我國著作權法係以處罰故意犯為原則，行為人主觀上須具備侵權故意，客觀上則構成違法行為始可成立犯罪。因此，制度上企業若能設計內部控管機制，證明其已採取合理措施防止侵權、且不具有侵權故意，可能有效降低刑事責任成立的可能性，甚至完全免除刑責非難。透過前述三階層治理模型所建構的機制，包括嚴謹的法律意見書及各項具體務實的執行措施，均可能有助於排除「侵權故意」之刑事認定。

若業者在業務上確有需求，必須取用他人素材嵌入函數進行模型訓練，則無論如何，於產出階段均不得生成與原作品構成實質相似的內容。更關鍵的是，由於訓練過程中使用他人素材，對於產出的結果，必須建立一定程度的轉化性──亦即不能僅是單純使用他人材料進行訓練，而應彰顯其利用後所能達成的具體效益，無論是提升電腦或網路之效能、增進使用者便利性，甚至促進資訊或知識的傳播普及，從而創造出可能之公共利益等，都可達成明顯的轉化成合理使用的效果。

然而，有鑑於我國迄今尚無任何此類型的相關判決，此等作法是否即能完全免責，仍屬未定。至少在經過上述縝密規劃與操作，並輔以外部專家或律師出具有說服力的法律

註30：歐盟透過《數位單一市場著作權指令》（DSM Directive）所設TDM之責任豁免相關規定，允許有合法存取權之研究者或企業為資料探勘目的複製受保護作品。

註31：*GEMA（Global Entertainment Marketing Academy of Arts & Sciences）v. OpenAI.*

註32：Daryl Lim (2025), *Governing generative AI*. Akron Law Review, 57(2)., https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5167491

意見書之前提下，或可免於刑責。對於此，長遠之計可能業者需要進行立法遊說，請求訂定出一套比照「無人載具創新條例」中的模式，擬定一套讓擬開發AI模型的業者，得以豁免刑責之「著作權沙盒」機制，或進而以歐盟之TDM豁免責任、甚至是參考日本著作權法第30條之4與第47條之5之立法例來修法[33]。

### 二、民事責任治理核心善良管理人注意義務之具體化

然而，在民事責任方面，法院之審查標準則將更為嚴格。實務上，我國法院通常以企業是否已盡「善良管理人之注意義務」作為判斷核心。換言之，即使企業（負責人或實際執行者）並無故意侵權，只要未建立充分之風險控管措施，仍可能須負擔民事賠償責任。因此，企業法律風險管理之重點，不在於完全排除侵權之可能，而是在於是否能證明其已採取合理且可期待之防範措施，並於整體執行過程中，視其各環節是否已善盡注意義務而定。

在實務上，「善良管理人注意義務」不能只停留於抽象概念，而應轉化為具體可操作之制度。此種制度化治理，乃企業避免民事責任之關鍵[34]。具體而言，企業應建立完整之內部規範體系，包括前述AI使用準則、操作手冊與內部審查流程，並透過跨部門合作，使法務、技術與營運單位共同參與風險管理。此外，將相關制度導入標準化管理架構（類似ISO），不僅可提升內部執行力，亦可作為外部審查與司法判斷之客觀依據。總之，企業必須從「事後責任承擔」，轉為「事前風險治理」，並透過制度化措施建立可被檢驗之法遵體系。

## 柒、AI導入時管理階層應具備公司治理之核心概念

以上列舉諸項AI導入之具體措施建議，雖然實施時將增加成本甚至遭到員工的抵制[35]，但最最關鍵的，仍在於企業負責人或高階管理階層尤其是董事會成員，在觀念和心態上應做好思維準備並果斷地付諸行動，否則如仍秉持傳統觀念與做法，勢將無法在新一波科技潮流下因應新局。以下茲再提三個重要觀念。

### 一、從風險「事後補救」轉向「事前治理」

生成式AI所帶來的著作權問題，已從傳統個案侵權轉變為系統性風險。AI涉及從資料蒐集、模型訓練到內容生成之完整供應鏈，每一環節均可能引發著作權問題[36]。此外，AI系統

註33：同註10，第529-533頁。

註34：Araz Taeihagh (2025), Governance of generative AI. Policy and Society, 44(1), https://doi.org/10.1093/polsoc/puaf001.

註35：組織抗拒與文化障礙，往往成為AI導入失敗的關鍵因素。因為員工如將AI視為對其工作構成威脅或是增加工作負擔，就會抵制AI的採用。

註36：Lee, et. al., *supra* note 5.

對訓練資料來源如缺乏透明度，將使企業難以確認其合法性，進而提高其法律風險[37]。因此，企業若仍採取傳統事後補救的法律策略，將無法有效因應AI所帶來之結構性風險。在此種背景下，公司治理的核心任務，應從「違法事後補救」轉向「事前之風險預防與制度設計」，亦即在AI導入初期即建立合規機制，以降低侵權發生之可能性。

在公司治理層面，企業若僅停留於抽象政策宣示，亦將難以滿足法律上對注意義務之要求。因此，AI之使用必須納入制度化管理，並強調文件化之可查核化。此種治理模式的核心，在於將AI風險控管，轉化為具體可操作之SOP，使企業能在爭議發生時提供具體證據，證明其已善盡管理義務[38]。

### 二、企業內部責任分配與治理架構

企業導入AI，著作權風險不僅屬於單一部門，而是橫跨技術、法務與管理層之，因為AI涉及多方參與，包括資料提供者、模型開發者與最終使用者，形成複雜之「責任鏈」[39]。因此，企業應透過公司治理機制，明確界定各角色之責任。例如，技術部門負責模型設計與風險控制，法務部門負責合規審查，而管理層則應負責整體風險監督，然後再由一AI工作群組（AI Task Force）兼顧協調彼此橫向之溝通整合，透過此種分配才可降低「責任落差」。總之，企業應將整體組織和營運之設計納入公司治理範疇，以「技術＋法律＋管理」多軌治理模式，來提升企業整體法遵能力。

### 三、合規策略之制度化：董事會監督治理框架

在更高層次上，企業應將AI著作權風險納入整體公司治理架構。AI所引發之法律問題具有跨領域特性，需透過制度化治理加以因應[40]。因此，企業應建立以下整體治理思維：首先，將AI風險納入董事會監督範圍，使其成為企業風險管理之一部分；其次，建立跨部門協作機制，以整合技術與法律專業；最後，持續追蹤法律發展，以即時調整企業策略。此種制度化治理模式，不僅可降低侵權風險，亦有助於提升企業在AI時代之競爭力與合法性。

## 捌、未來展望：從風險控管邁向AI治理之競爭格局

AI的快速發展，已對現行法體系帶來結構性衝擊與挑戰，目前我國的法制尚處於AI基本法的起步階段，無論在細部立法或司法解釋層面，皆有待建立明確的規範架構，整體法律環境仍存在高度不確定性。在此情況下，企業不宜坐視而必須積極建立因應策略。然而，關鍵並非在於追求風險的完全消除，而是應建構出一套具可驗證性與問責性

註37：Wu, *supra* note 4.
註38：Quintais, *supra* note 6.
註39：Lee, et. al., *supra* note 5.
註40：Quintais, *supra* note 6.

的AI治理機制，使其能在混沌的法規環境中穩定運作。

隨著AI應用逐步深化，企業導入AI所面臨之法律風險，已由過往以個別侵權行為的問題，轉為涉及整體組織決策、流程設計與制度化內控之公司治理議題。AI系統之訓練資料來源、建模方式、生成內容之審核機制，以及人機協作之決策分工，均可能影響法律責任。因此，企業必須從前述之AI使用準則、資料治理、人機協作、輸出監控、內部紀錄與稽核管理等多元面向，建立整體化之風險控管體系，透過制度化、文件化與程序化之措施，確保AI決策與執行具備可追溯性與合規性。總之，AI不但未削弱人類在責任體系中的角色，反而使企業在組織與法遵層面，承擔更高之注意義務與管理責任。

再從更宏觀的角度來看，企業在AI時代的競爭，也已逐漸從純粹的技術競逐，轉向「治理能力」的較量。能否有效整合創新應用與法規遵循，建立兼顧效率、透明度與風險控制的治理架構，將成為影響企業長期競爭力的關鍵因素。換言之，AI治理不再僅是風險管理或法遵部門的輔助功能，而應逐步轉化為企業組織整合運作的核心能力。未來，具備成熟AI治理機制的企業，不僅能降低潛在法律風險，更能在市場信任、監管因應及維護營運等層面取得優勢。

綜上所述，生成式AI之導入，已使企業之著作權管控產生本質上的轉變。企業唯有透過全面整合設計，將AI應用納入公司治理與風險控管的整體架構中，方能在促進創新與確保合規之間取得平衡。在此脈絡下，導入AI後的相關措施，本質上已屬於「公司治理」的一環，更精確地說，已提升為「AI治理」的層次；而AI治理能力本身，也將成為企業未來競爭環境中不可或缺之核心競爭力。

## 玖、小結

不論是生成式AI或AI代理，都是這幾年才開始落地。因此，企業若打算部署，現在是最佳時機，應避免像其他領域一樣，等到積重難返才想規劃與掌控就難以駕馭了。由於各企業的樣態不同，因此本文以上這些治理措施，係「取法乎上」的checking list，業者並非當然都要全盤採納，而應根據自身產品或服務特性、公司規模、市場因素、客戶與消費者偏好、員工特質、企業文化等條件，就其中挑選適合的項目，來做為AI治理防範機制。

最關鍵的是，如AI生成內容係直接提供給客戶或消費者對外散發（例如企劃案、廣告文宣、LOGO、產品設計），這時曝露在外的風險最高即應施行高標；反之若僅是內部用，則採取的應對模式可有所不同。再者，本文雖係針對著作權，但鑑於相同屬性，對於企業之其他智財權項目如商標專利與營業秘密等，皆得予以參酌使用。

總之，企業應依自己業務需求，採行避險措施加以執行，重點是，必須要把做的事情制度化，以證明已盡到注意義務！而做好風險管理，不僅是管理層與執行人員的責任，董事會成員與公司負責人更是最核心的角色，必須以由上至下的方式（top down而非bottom up）來推動，將著作權防範與法遵合規，內化為公司日常作業的作業標準流程。