生成式AI法律風險控管

益思科技法律事務所 賴文智律師

2026/05/13@主計人員幹部培育班第13期

益思科技法律事務所
Infoshare Tech Law Office

賴文智律師

wenchi@is-law.com

T. 886-2-27723152 #222

學歷

台灣大學法律學研究所所碩士

台灣大學法律系學士

現職

益思科技法律事務所所長

經濟部智慧財產局著作權審議及調解委員會委員

經濟部智慧財產局智慧財產培訓學院顧問

台灣商標協會常務理事

台灣網路暨電子商務產業發展協會（TiEA）監事

著作

當文創遇上法律：個資、名譽與肖像（與蕭家捷律師合著）

當文創遇上法律：讀懂IP授權合約

當文創遇上法律：讀懂經紀合約書

當文創遇上法律：智慧財產的運用

從NDA到營業秘密管理

APP產業相關著作權議題（與蕭家捷律師合著）

企業法務著作權須知

技術授權契約入門（與劉承愚律師合著）

個人資料保護法 Q&A（與蕭家捷律師合著）

企業法務商標權須知

數位著作權法（與王文君合著）

智慧財產權契約

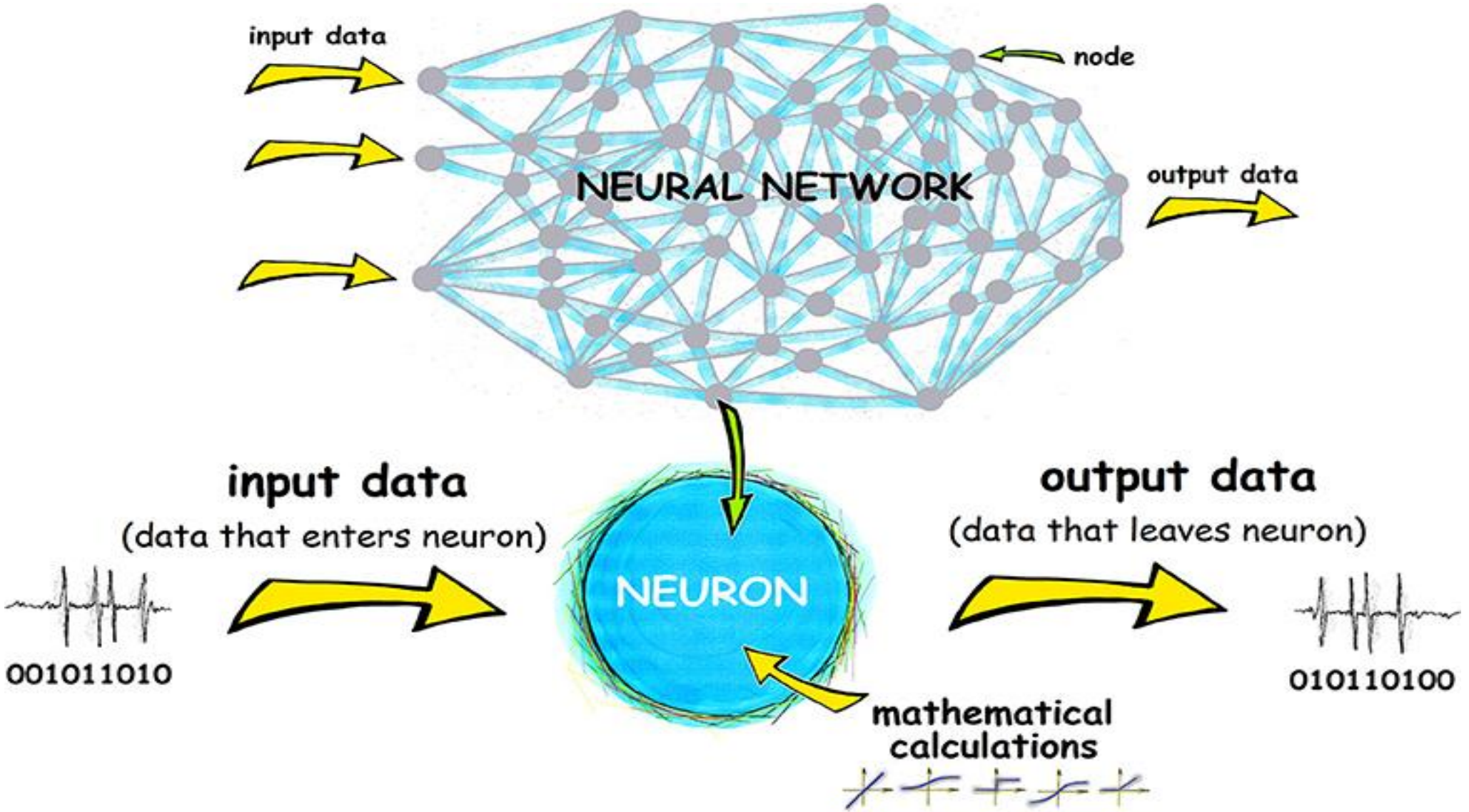
益思科技法律事務所
Infoshare Tech Law Office

生成式AI在主計領域的應用場景

- **主計人員自身運用：讓繁複作業更有效率**
  - 預算編製與審查：彙整跨年度資料、自動產出說明草稿
  - 統計分析：快速摘要海量數據、產生圖表初稿與趨勢解讀
  - 法規查詢：即時檢索預算法、會計法、審計法相關條文
  - 公文與報告：草擬、潤稿、翻譯與格式校對
- **外部使用者運用：對主計透明度與專業度的新挑戰**
  - 立委助理：以AI消化厚重預算書，產出更尖銳的質詢
  - 民眾與公民團體：透過AI解讀預算、決算、審計報告
  - 媒體：以AI比對歷年數據、挖掘異常，形成輿論監督
- **主計人員必須比過去更熟悉AI，才能因應更專業的監督與提問**

AI與生成式AI的發展

- **傳統人工智慧（封閉領域、求最佳解）**
  - 1956達特矛斯會議起：Rule-based、類神經網絡、專家系統、Deep Blue到2016 AlphaGo
  - 限定規則下找最佳或相對好的解，演算到一定步數後人類就贏不了
- **生成式AI（開放問題、生成相對適合的解）**
  - 2018年OpenAI推出GPT：基於Transformer架構之**大型語言模型（LLM）**
  - 以巨量未標記文字進行預訓練，能生成類似人類自然語言的內容
  - Emergent Abilities：LLM+Big Data規模到一定程度後，於資料量較少的領域亦展現預測與生成能力



生成式AI與傳統AI的關鍵差異

- **問題類型：封閉 vs 開放**
  - 傳統AI：限定規則下找最佳解（如圍棋、影像辨識）
  - 生成式AI：開放式問題下生成相對適合的解（撰文、摘要、對話）
- **輸出特性：可重現 vs 機率性**
  - 傳統AI：相同輸入 → 相同輸出，可驗證、可稽核
  - 生成式AI：相同輸入可能產生不同輸出，可能**產生幻覺（hallucination）**
- **使用門檻與風險範圍**
  - 傳統AI：需專業團隊建模，應用於特定場域
  - 生成式AI：人人可用、應用無邊界，**機密外洩、著作權、責任歸屬等風險隨之擴大**

生成式AI為什麼會產生幻覺？

- 討論法律風險前，我們需要先理解生成式AI為何會出現錯誤（幻覺）？
- **不是「查資料」，而是「預測」關聯性高的下一步**
  - 政府AI手冊：AI透過機器學習與演算法，對輸入資料產生預測或內容輸出
  - 非資料庫查詢，而是依據過去學習資料「預測最可能的下一段文字」
  - 運作邏輯是依語言規律與機率，生成「看起來合理」的內容，而非「查到正確答案再回傳」
- **訓練資料的本質：大量、未完全篩選的資訊**
  - 來源：網路文章、書籍報告、各類公開資料
  - 未經特定標準的精確篩選與在地法規校正
  - **關鍵：訓練資料很多，不代表「知道什麼是對的」**
- **隨著AI業者的投入，生成式AI服務也透過各種機制減少幻覺的問題**

AI發展對主計領域所帶來法律面的影響

- **AI不只是技術，而是公務環境與工作思維的重塑**
  - AI牽動社會運作，相關法規將大幅變動
  - 公部門任務被重新拆解、組合；人機分工與協作將重新定義
  - 立委、媒體、民眾將以AI提出更專業的監督與質詢
- **從「法遵」到「治理」的全面升級**
  - 法遵：個資、資訊公開、預算、會計、資安等規範
  - 治理：機關內部AI政策、權責、稽核與責任歸屬
  - 倫理與課責：偏誤、錯誤、洩密時的因應
- **人工智慧基本法通過後，政府機關不論主動或被動，皆須直面AI議題**

人工智慧基本法

- 2025/12/23立法院三讀通過，2026/1/14公布
- 中央主管機關：**國家科學及技術委員會**+涉及各目的事業主管機關職掌者，由各目的事業主管機關辦理
- 人工智慧風險分類框架：數位發展部
- 各目的事業主管機關訂定以風險為基礎之管理規範
- 政府應依本法規定，檢討所主管之法規與行政措施；有不符合本法規定或無法規可適用者，**應自本法施行後二年內，完成法規之制（訂）定、修正或廢止，及行政措施之改進**。
- 定義：本法所稱人工智慧，指具自主運行能力之系統，該系統透過輸入或感測，經由**機器學習及演算法**，可為明確或隱含之目標實現**預測、內容、建議或決策等影響實體或虛擬環境之產出**。

人工智慧基本法

- 永續發展與福祉：主計總處可能要思考AI的導入會不會造成基層行政人員被邊緣化？如果導入生成式AI後，現有員工不會使用、偏鄉機關沒有能力導入、基層只剩剪貼資料餵給AI，就可能形成新的數位落差
- 人類自主：這是主計體系最核心的原則，因為主計制度本質上就是「由人類承擔法律責任」。不能由AI自行核定預算、不能由AI直接決定是否合法、不能由AI代替公務員作成行政判斷，應該是主計業務AI治理的絕對原則
- 隱私保護與資料治理：主計資料經常涉及公務資料、採購資訊、人事資料、補助資料、財務資訊等，最大的風險就是把敏感資料送進外部模型，因此，禁止直接使用公開版生成式AI、導入封閉環境的AI模型、建立資料分級制度，都是必須考量的重點

人工智慧基本法

- 資安與安全：主計系統涉及國庫支付、政府歲計、地方財政系統等，是國家重要基礎設施的一部分。如果所導入的生成式AI遭Prompt Injection、資料污染、惡意訓練等，都可能產生嚴重影響
- 透明與可解釋：AI生成內容應做適當資訊揭露或標記，已幾近屬於法律義務。對於主計業務而言，搭配RAG或類似架構，儘量實現由AI生成的結果可以追溯源頭，是比較合理的做法
- 公平與不歧視：主計業務可能涉及對於補助分配、社福統計、預算資源配置等規劃與查核，AI若使用過去資料學習，可能會放大過去資料中的偏誤。但是，也可能使用AI其實是更容易發現過去潛藏的不公平與歧視
- 問責：這是AI治理的核心，AI可以提供建議，但責任仍屬於機關。因此，AI導入、使用的紀錄、人工覆核制度、誰該負什麼層級的責任，需要在AI治理制度講清楚

主計人員使用生成式AI的法律風險

- 取代承辦人專業判斷之風險：AI產出須由主計人員就風險進行客觀且專業之最終判斷，不得取代承辦人之自主思維
- 機密文書外洩風險：國家機密文書及一般公務機密文書應親自撰寫，禁止使用生成式AI
- 公務資訊與個資外洩風險：不得向AI提供未經機關同意公開之公務資訊或個資，亦不得詢問涉及機密業務或個資之問題
- 產出正確性與決策責任風險：不可完全信任AI產出，不得以未確認之內容作成行政行為或公務決策之唯一依據
- 未揭露使用之風險：以生成式AI作為輔助工具時，應適當揭露
- 違反著作權法風險：將他人著作輸入生成式AI進行生成，可能產生之著作財產權與人格權之可能性
- 採購委外之延伸風險：得標廠商應用生成式AI提供履約標的相關風險

主計主管的兩種管理風險：輕忽與過度依賴

- **忽略風險：當作沒這回事**
  - 影子AI(Shadow AI)：同仁早已私下使用個人帳號,主管不知情、無控管
  - 未訂內控規範，違反指引「機關得訂定使用規範或內控管理措施」之政策期待
  - 未對採購委外廠商提示遵循指引，擴大委外洩密與品質風險
  - 外部使用者已顯著提升效率，內容仍以人工耗時作業，影響審核時效與政策回應
- **過度依賴風險：當作萬靈丹**
  - 容忍承辦人未實質審核即引用AI產出
  - 將思考、規劃能力外包予AI，專業判斷力與制度記憶逐漸流失
- **主管的責任在「禁止」與「放任」之間，建立分級使用、內控與揭露機制**

Deloitte to refund government, admits using AI in $440k report

Edmund Tadros and Paul Karp

Oct 5, 2025 – 7.41pm

Deloitte Australia will issue a partial refund to the federal government after admitting that artificial intelligence had been used in the creation of a $440,000 report littered with errors including three nonexistent academic references and a made-up quote from a Federal Court judgement.

A new version of the report for the Department of Employment and Workplace Relations (DEWR) was quietly uploaded to the department's website on Friday, ahead of a

生成式AI對著作權領域的衝擊

- 權利保護層面
  - 應用生成式AI產出的成果是否受著作權法保護？
  - 將既有著作透過生成式AI生成其他成果，該等成果是否受著作權法保護？
  - 人類的創作+生成式AI，著作權保護的範圍如何判斷？
  - 如何證明是人類的創作成果？
- 授權層面
  - 生成式AI所需的大量訓練資料，若為著作，是否需要取得授權？取得授權的成本與方式
  - 生成的成果利用是否需取得授權？開源的AI授權規範的遵守？
  - 將他人著作輸入生成式AI服務進行生成，需要取得什麼樣的授權？

生成式AI對著作權領域的衝擊

- 合理使用層面
  - 訓練資料的利用是否符合合理使用的基準？
  - 應用生成式AI時取用他人著作作為素材，是否符合合理使用規範？
  - 利用他人的模型訓練自己的模型，在違約的情形，是否構成合理使用？
- 侵權層面
  - 訓練資料含有未經合法授權的著作，據此訓練出來的生成式AI模型，單純導入該模型進行使用，是否會構成侵權？
  - 前述生成式AI模型所生成出來的成果對外利用，是否會構成侵權？
  - 生成式AI生成成果偶然的近似，是否會構成著作權侵害？是否會涉及刑事責任（有無侵權故意）？
  - 著作權侵害（抄襲）的判斷標準是否受到生成式AI影響而改變？

智慧財產局電子郵件1111031

- 一、依我國著作權法第3條第2款及第10條規定，著作人指創作著作之人，著作人於著作完成時享有著作權。換言之，著作必須係以自然人或法人為權利義務主體的情形下，由自然人所為的創作，方可能受到著作權的保護。.....
  - (一)第一種是「以人工智慧為工具的創作」，也就是人類有實際的創意投入，只是把人工智慧(例如：繪圖軟體)當作輔助工具來使用，在這種情形依輔助工具投入創作者的創意而完成的創作成果仍可以受著作權保護，著作權則由該投入創意的自然人享有，除非有著作權法第11條及第12條之情形。
  - (二)第二種是「人工智慧獨立創作」，也就是人類並無實際的創意投入，完全是由AI的演算功能獨立進行完成創作，此時由於AI並非自然人，沒有人類精神文明的投入，其創作完成成果自然不屬於著作權法保護的著作，原則上無法享有著作權。

智慧財產局電子郵件1121229

- 至於所詢將攝影的照片透過AI繪圖軟體將照片卡通化，或套用濾鏡對照片進行光影、調色、模糊、黑白等各項影像效果之調整後(依您來信所述，皆為電腦演算，僅係演算法不同)產生之圖像，是否另為原照片之衍生著作，而獨立受著作權法保護，**需視該利用電腦演算法所生成的圖像有無人類實際的創意投入而定，如有人類有實際的創意投入，只是把電腦演算工具(例如：繪圖軟體)當作輔助工具來使用，完成的創作成果仍可以受著作權保護；惟如無人類實際的創意投入，完全是由電腦演算功能獨立進行完成創作，該AI繪圖軟體演算之成果不屬於著作權法保護的著作，原則上無法享有著作權**。(併請參考本局電子郵件1110502b之說明)

Art & Tech

How This A.I. Image Became the First to Snag Copyright Protection

The U.S. Copyright Office ruled generally last month that work created from A.I. text prompts could not be copyrighted.

Kent Keirsey, *A Single Piece of American Cheese* (2024). Photo courtesy of Invoke.

AI訓練資料的合法性

- 大型生成式AI需要大量的資料進行預訓練，但目前最困擾的問題在於因為權利碎片化的關係，不可能取得相關權利人對於將之用於AI訓練的合法授權，**目前各國對AI訓練資料的法律認定尚不明確**
- 著作權法：可能必須要透過修法處理比較明確
  - 日本著作權法第30條之4規定，「著作物在下列情形或其他非以自己或他人享受該著作物中思想或感情表達為目的之情形下，得在必要範圍內不限使用方式為利用。但依該著作物之種類、用途及其利用方式可能對著作權人不當地造成損害者，不在此限：...**二、供資訊分析（係指從多數著作物或其他大量資訊中，擷取涉及該資訊構成之語言、聲音、影像或其他要素之資訊，進行比較、分類或其他解析之行為...）之用**者。三、除前二款情形外，**供不涉及人之感知而利用於電子計算機資訊處理過程或其他利用**者。...」

Artificial intelligence (AI)

AI startup Anthropic agrees to pay $1.5bn to settle book piracy lawsuit

Settlement could be pivotal after authors claimed company took pirated copies of their work to train chatbots

Associated Press

Fri 5 Sep 2025 21.19 BST

Share

- AI訓練來源資料的合法性，可能會成為未來自行訓練AI的關鍵
- 歐盟著作權指令有關文字與資料探勘(Text and Data Mining, TDM)的例外規定，若一般企業要適用，則須為合法且尊重著作權人的方式取得訓練資料
- 歐盟人工智慧法要求透明度，亦會適用在訓練資料的來源

著作權法與抄襲

- 何謂「抄襲」？
  - 「抄」有「取」、「謄寫」之意；「襲」由「加衣服」而延伸有重複、沿用、繼受的意思。抄襲一詞即指取用或沿用他人智慧創作在自己的作品或產品之意
- 著作權法本文沒有出現過「抄襲」二字
  - 通常所謂的「著作抄襲」，應該是指構成著作「重製權」或「改作權」的侵害
  - 重製類型的「抄襲」可能包括全部或部分著作的重製，通常是使用他人著作作為創作素材
  - 改作類型的「抄襲」是最具有爭議的侵權類型，因為改作者會投入自己的創意進行創作活動，但其創作活動涉及他人表達的「取用」，許多「二創」嚴格來說都是這類的侵權

最高法院81年度台上字第3063號判決

- 認定抄襲之要件有二，即(1)接觸，(2)實質相似。主張他人之著作係抄襲其著作者，應舉證證明該他人曾接觸被抄襲之著作，構成二著作實質相似。
- 接觸
  - 接觸分為直接接觸與間接接觸兩者態樣，間接接觸係指於合理之情況下，行為人具有合理機會接觸著作物，均屬間接接觸之範疇。
- 實質相似
  - 是否構成「實質近似」須以作者的「創意活動」之所在為主要判斷標準。若二個著作就其整體看起來「大同小異」，因著作權法保護的是作者的「創作」，所以，若是「大同」的部分，雙方都是參考某一些相同來源的素材，而「小異」的部分，則是雙方各自創意所在

最高法院 103 年度台上字第 1544 號

- 其中實質相似不僅指量之相似，亦兼指質之相似。在判斷圖形、攝影、美術、視聽等具有藝術性或美感性之著作是否抄襲時，如使用與文字著作相同之分析解構方法為細節比對，往往有其困難度或可能失其公平，因此在為質之考量時，尤應特加注意著作間之「整體觀念與感覺」。而在量的考量上，主要應考量構圖、整體外觀、主要特徵、顏色、景物配置、造型、意境之呈現、角度、形態、構圖元素、以及圖畫中與文字的關係，以一般理性閱聽大眾之反應或印象為判定標準。

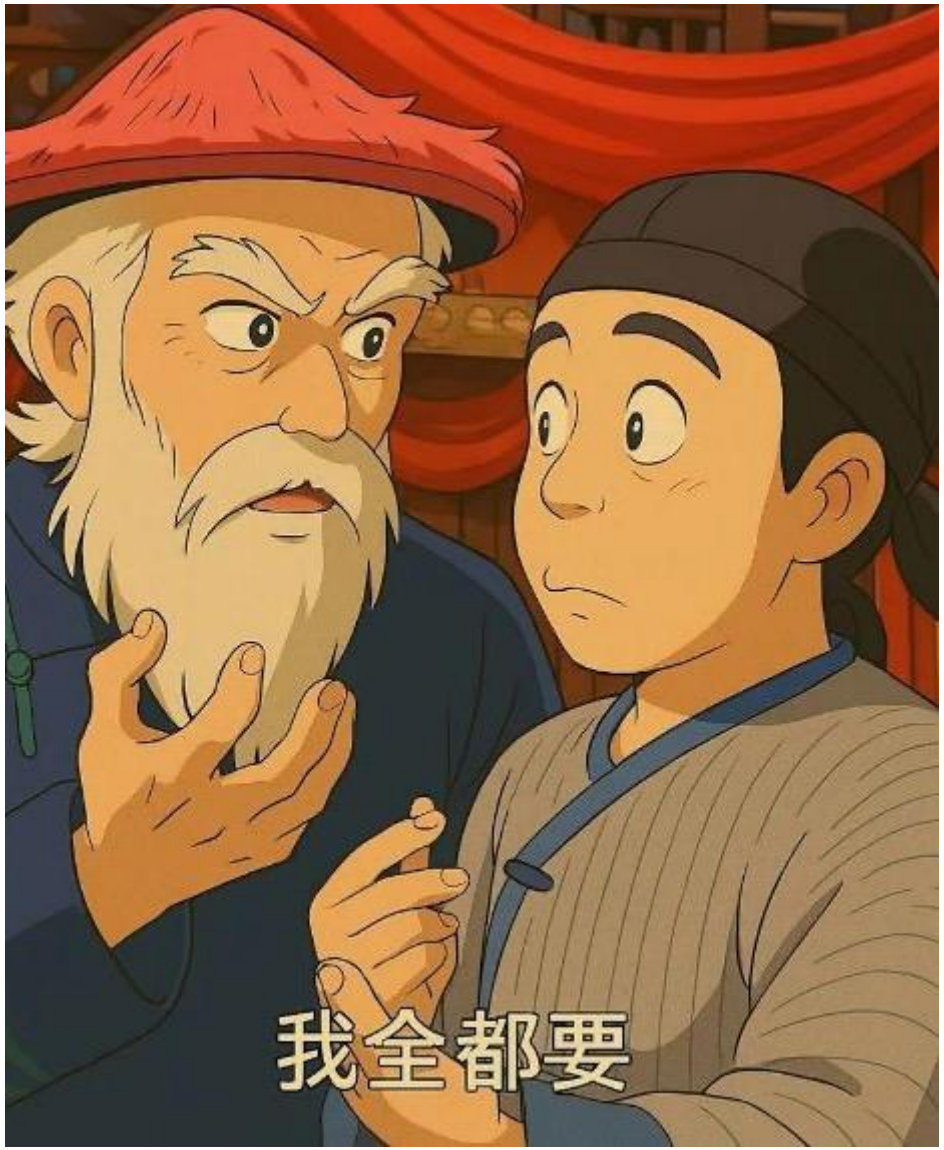
觀念的抄襲不是抄襲？

- 著作權法第10條之1，「依本法取得之著作權，其保護僅及於該著作之表達，而不及於其所表達之思想、程序、製程、系統、操作方法、概念、原理、發現。」
- 觀念相同，表達不同，這類的參考不會被認定為侵權

畫面中央：
一位小朋友正在偷偷臨摹他人的畫作，旁邊打上一個顯眼的「禁止」符號，表明抄襲行為不可取。
另一位小朋友則興高采烈地拿著畫筆，正站在梯子上朝天空繪製出自己的原創作品，象徵積極創作的正確態度。
週邊背景：
廣闊宇宙中滿佈著各式各樣的星辰、星雲與行星，呈現出豐富、多彩且無限的想像空間，象徵創作的無限可能性。
其中一些星球或星雲可設計成如書本、音符、畫筆、攝影機、設計草稿等智慧財產創作領域的符號，強調智慧財產權涵蓋的多元領域。
色彩與風格：
採用溫暖、柔和的色調，搭配卡通、童趣風格的人物與背景，使整體視覺友善、吸引目光且適合教育宣導用途。
字體建議採用圓潤易讀的手寫或卡通字體，以增加親切感。
文字標語（繁體中文）：
「抄襲不可取，創意無極限」

生成式AI與吉卜力風

- 生成式AI使用大語言模型（LLM），其演算法的基礎在於類神經網絡，即嘗試以人類大腦神經網絡的運作，模仿人類學習的方式然後透過大量的資料訓練、校正、後訓練等，在算力足夠的情形下，在許多領域表現已經超出人類一般水準
- 著作權法在抄襲的議題上，會受到來自生成式AI普及應用及快速進步極大的影響，輸入照片（影像）生成吉卜力畫風的圖像，洗版社群網站的頁面，然而，繪畫風格在傳統上是不受著作權法保護的抽象概念
- 抄襲的標準可能會更嚴苛，因為生成式AI愈強的創作表現，人類的創作保護可能會更作限縮



依據他人照片繪製畫作-生成式AI也有類似問題

- 抄襲的認定，除原創性的高低之外，不同種類的著作，因創作的方法、對外表現的方式等均有所不同，所以，判斷是否構成抄襲時，可能會著重不同的方向，甚至同一類的著作，但屬性不同可能判斷是否構成侵權的方式也不同，其中，又以跨著作種類類別的案件最為複雜

編號	系爭照片名稱	系爭照片
1	「蘭城百工榮興木雕社」（系爭照片1）	
2	「蘭城百工古城裡最悠久的印舖廣文堂」（系爭照片2）	

編號	系爭畫作編號	系爭畫作
1	系爭畫作1	
2	系爭畫作2	

如何管理生成式AI的法律風險

- 從法律的角度來看，生成式AI的技術本身不會主動去違法、侵權，關鍵還是人如何去使用生成式AI這個技術
- 政府機關須注意導入外部生成式AI服務的資安風險
- 熟習生成式AI服務，通常新領域的法律風險來自於不熟悉技術的背景
- 優先在自己熟悉的領域應用生成式AI服務，自己的專業領域更容易判斷生成式AI生成的成果是否可能涉及侵權風險
- 應用生成式AI生成的成果做出適當的標示，讓使用者有機會評估是否信任或使用
- 保留生成式AI使用歷程，作為後續舉證的依據
- 正式發布前透過外部的檢索確認是否有侵權風險（但隨AIGC逐步充斥網路，可能會愈來愈困難）

益思科技法律事務所
Infoshare Tech Law Office

106433 台北市忠孝東路四段290號8樓
8/F., No. 290, Sec. 4, Jungshiau E.Rd., Taipei, Taiwan, R.O.C.
T. 886-2-27723152 F. 886-2-27723128
www.is-law.com