圖書館如何管理生成式AI的法律風險

賴文智

益思科技法律事務所 律師/所長

一、前言

生成式 AI（Generative AI）正快速改變知識的傳遞與學習方式。過去，知識的流通多半依循「作者→出版社→讀者」的線性路徑，而隨著 AI 服務的普及，學習者有機會主動與知識載體進行互動，形成全新的「對話式學習」模式。

以 Google 提供的 NotebookLM 服務為例，筆者將近期將出版的書籍 PDF 上傳至該服務，NotebookLM 可以自動生成 AI 摘要、心智圖、聲音摘要，甚至主動提出問題讓使用者選擇是否透過問答進行互動。作者在寫書的時候，必然是有一個預設的知識傳遞架構，綱舉目張地完成書籍內容撰寫，但作者所預設的學習途徑，卻未必適合每一位讀者，書籍作為知識的載體，因其形式而限制了學習的方式，讀者最多選擇是「跳著讀」，無法自行決定如何學習。然而，AI 卻開啟了新的可能性，使用者如能取得書籍的電子檔，將不再僅被動接受作者設定的知識脈絡，而能依自身問題與興趣為導向，向 AI 提問、延伸、重組，總能找到符合自己需求的學習方式。未來的知識流通，或將不再以固定形式的出版品為唯一載體，而是真正以「知識」為核心，透過與 AI 服務的互動，達成由使用者自行決定如何汲取知識。

當使用者開啟更多學習的可能性之後，就很難回到過去，而作為知識保存、學習與傳遞的社會功能的圖書館，在相同的角色下，其任務勢必亦將重新定義。

二、生成式AI對圖書館服務帶來的影響

《圖書館法》第 1 條規定，「為促進圖書館之健全發展，提供完善之圖書資訊服務，以推廣教育、提升文化、支援教學研究、倡導終身學習，特制定本法。」圖書館在 AI 時代應該提供什麼樣的服務？

(一) 圖書館從業人員須重新定義「任務」

顯而易見地，圖書館既有例行性、重複性的工作任務將大量被 AI 取代，在圖書館所扮演的社會角色不變的情形下，圖書館從業人員面對的並不是「工作」的消失，而是這個工作崗位背後的「任務」，必須要因應 AI 進行調整。例如：AI 能夠自動化分類、摘要、推薦與查詢，這些正是圖書館員過去日常的核心工作任務之一，當這些任務不再由人來執行，圖書館員即需要思考其他能夠達成圖書館目標的

「任務」，或許是協助讀者使用 AI 工具，實現學習或知識傳遞的個人化，或許是扮演 AI 素養的教育者，協助讀者在應用 AI 工具時，能夠妥善自我保護，減少來自 AI 的負面影響，或是如何輔助研究者應用 AI 進行文獻整理、資料分析與學術寫作等。當然，也可能還是維持傳統紙本書籍閱讀樂趣的推廣，畢竟不是每一個人都覺得 AI 帶來的是便利，一頁一頁地翻閱紙本書亦有其超越知識傳遞的意義。

(二) 不同圖書館應該思考不同的定位與館藏策略

另外，圖書館的館藏策略是否需要因應 AI 時代來臨而進行調整？例如：紙本與數位館藏的比例配置、數位館藏能否與 AI 服務整合、取得能夠應用於 AI 的館藏授權等問題，都是需要思考的。但筆者認為並不是一定要朝向大量減少紙本館藏的配比，因為圖書館可能是人類經過嚴謹審閱程序出版的知識最後的堡壘，每個圖書館可以有自己不同的定位。舉例來說，若為大學圖書館，因為大學教學融入 AI 服務是大勢所趨，圖書館若要滿足教職員生的需求，自然需要考量到其館藏為 AI 所用的需求，在採購時即應更關注「AI 可讀取授權」（AI-readable license）或可用於訓練圖書館內部特定 AI 服務的授權，讓更多的知識能夠透過 AI 進行互動學習、摘要或生成適於不同讀者的知識內容呈現。

相較於網際網路時代，圖書館有一個整體的轉型目標，將原先「以人就館」的服務型態，轉變為將圖書館送至使用者處的「以館就人」的模式。進入 AI 時代後，不同性質的圖書館會依據其不同的定位，而有不同的館藏策略，但筆者認為在 AI 有能力處理大量數位資源的情形，或許同性質的圖書館間，也應該協調各自尋找不同的定位，形成由 AI 串起的實體 + 虛擬的大型知識網絡，一館一特色，可

能更適合在有限資源前提下的圖書館發展。

三、導入AI服務的法律與倫理風險

圖書館導入 AI 服務時，基於其提供公共服務的角色，除了法律面，也要注意倫理面的議題。行政院院會所通過的《人工智慧基本法》草案第 3 條規定所提及，「人工智慧之研發與應用，應在兼顧社會公益及數位平權之前提下，發展良善治理，並遵循下列原則：一、永續性…。二、人類自主性…。三、隱私保護及資料治理…。四、安全性…。五、透明及可解釋性…。六、公平性…。七、可問責性…。」上述原則同時包含法律面及倫理面的要求，即是圖書館用以檢視 AI 服務導入的指標。以下筆者即由著作權、機敏資訊、正確性及公共資訊服務這些面向說明相關的法律風險及可能的管理措施。

(一) 著作權風險

生成式 AI 最熱門的著作權議題，即是訓練資料合法性問題，因為生成式 AI 需巨量資料進行預訓練。以 ChatGPT-4o 為例，訓練時所使用的資料量估計高達約 45 TB 的文字資料，包括公開書籍、百科全書、學術論文、網站文章、對話語料及其他，若換算成書籍，大概就相當於 5 億本書，自然不可能逐一取得授權。然而，圖書館多不會自行訓練此類大型語言模型，而生成式 AI 服務是以模型而非將原始訓練資料作為資料庫對外提供服務，因此，單純利用外部廠商所提供的 AI 模型，依現行的著作權法並不會讓圖書館構成侵權行為。所以，圖書館需要關心的議題在於其導入後，圖書館進行微調的訓練或是為提供生成回應的正確性利用著作權屬於他人的資料，應該要取得作為 AI 應用相關的授權。

另一個圖書館應關注的議題則來自於 AI 生成內容對外利用時，侵權風險的控管。簡言之，生成式 AI 的廠商通常難以擔保其服務不會生成與他人既有著作高度相似的內容，因為即令透過各種限制（如限制使用者生成含有蝙蝠俠或其他知名 IP 的圖樣），必然會存在很多的漏網之魚。然而，這並非代表圖書館即不適於使用 AI 生成相關內容，因為即令圖書館是聘請「自然人」來創作這些內容，都還是有可能構成侵權（抄襲）。圖書館如果希望可以在導入 AI 服務時，降低生成內容侵權的風險，從人員的管理方面，應該是提醒使用者避免將他人的著作輸入由 AI 進行改作（生成），或避免刻意引導 AI 生成出與他人著作類似的內容；其次，因為著作權侵害的責任仍然建立在行為人具有「故意或過失」的主觀要件，因此，若擬將 AI 生成的成果對外利用，建議應保留該生成內容從無到有的完整過程紀錄，雖然這些紀錄無法再生成一個完全一模一樣的成果，但可以用來檢視是否有輸入他人著作，或是不當透過 prompts 引導 AI 生成侵權內容，從而可以作為沒有侵權故意或過失的證據；三者，在利用 AI 生成內容時，應保留標示其為 AI 生成內容，以利使用者判斷是否該信任該內容。

(二) 機敏資訊風險

對一般企業而言，機密資訊或個人資料的外洩，都是嚴重的問題，限制員工個人直接透過 Web 或 App 使用外部生成式 AI 服務有其合理性。對圖書館而言，同樣的機密資訊或個人資料的外洩問題，只是，一般圖書館從業人員會使用外部生成式 AI 服務的情境，通常不太需要將機密資訊或個人資料一併「餵」給 AI，所以，這方面風險主要是制定規範，做好人員的教育訓練即可。

然而，若是圖書館希望透過讀者使用圖書館服務的資料，作為利用 AI 提供讀者個人服務，分析讀者行為、借閱歷史、搜尋偏好等資訊，可能就會需要擴大蒐集個人資料。透過 AI 演算法對於個人作出書籍、雜誌或其他知識的推薦，背後必然是對個人貼上一定的標籤，這同時涉及到個人資料利用及 AI 治理的議題，關鍵在不能做不適當的連結與不得有歧視的狀況。舉例來說，將個人的閱讀偏好作為個人推薦可能是合宜的，但因個人閱讀或使用 AI 服務的紀錄，進一步去判斷其性別認同、性傾向、犯罪風險或個人行蹤等，即可能對於 AI 倫理有關隱私保護或是公平性的要求相衝突，甚至可能是屬於違法的行為，這些都應該要服務設計時即避免不當的應用情境。

(三) 正確性與信賴風險

從人工智慧 (Artificial Intelligence) 的英文文義，很清楚地即表示 AI 並不是真正的智慧，而是人造的、假的智慧。生成式 AI 是以「關聯性」(機率) 而非「邏輯推演」生成文字、圖像等。以 OpenAI 公司的 ChatGPT 為例，最基礎的運作邏輯就是依據使用者所輸入的文字，依據其 AI 模型運算、訓練資料所影響的參數值等，提供關聯性最高的回應。因此，早期會出現有 6 顆氣球，破了 3 顆還剩下幾顆，AI 回覆是 4 顆，可能僅僅因為預訓練的資料中，這類問題最常見的答案是 4 顆，4 顆在其 AI 模型及參數的評估結果，AI 即猜測 4 顆是我們所要的答案。當你表示 AI 答錯，AI 也會立刻道歉，因為相關的道歉內容是這個模型認為關聯性最高的回應內容。當然，現在 AI 服務提供者會採取各種方式嘗試改善正確性的問題。

若圖書館若要處理正確性的風險，即必須要對於生成式 AI 的技術原理有一定在的認知，因為最大的風險往往來自錯誤的認知。例

如，AI 可能可以作為各類參考服務，但卻不適合作為決策系統。圖書館員可能可以詢問當讀者有某些行為時，依據圖書館相關規範應如何處罰，但卻不適合逕行依據 AI 所給出的結果逕行處罰，即令可能 AI 大部分的時候是對的。因為涉及權利義務的事宜，跟給讀者閱讀參考的資訊不同，一旦涉及權利義務至少應該要有「人類覆核機制」，確保是人類進行來做成這個決定。

另外，降低正確性與信賴風險，除了據實標示 AI 生成的資訊，讓使用者自行判斷是否信賴該等資訊之外，因為通常圖書館服務是比較特定範圍，並不是像 ChatGPT 面臨的是全球使用者無窮想像力的各種使用，屬於為特定服務所開發的 AI，可以限制其用途，例如：只回應特定範圍的問題，就比較不容易出錯，且不會出現被濫用作為其他用途，還不當耗費算力資源的問題；也比較容易建立持續驗證機制，例如，由專業館員定期審查 AI 回答的準確性、建立使用者回饋機制，允許讀者標註錯誤內容、確保 AI 模型能定期更新資料來源等，這些都有助於提高 AI 服務回應的正確性。

四、結論與建議

圖書館導入生成式 AI 服務，並不像過去影印機、電腦或網路比較偏「科技產品」的引進，而是涉及到圖書館由定位、目標到從業人員的任務配置，是進入 AI 時代勢將面對的議題。法律風險控管的關鍵，自然不在哪個生成式 AI 服務或是 AI 模型比較安全，而是要先回到制度的設計與人員的管理，簡言之，由服務設計至執行的各階段，透過制度建立持續的風險評估、人員培訓與人類覆核機制，才是長期降低法律風險的最佳策略。

(一) 建立資料治理機制

圖書館導入生成式 AI 服務或後續持續改善其正確性，均須大量的資料，無論這些資料是自有資料或外部資料，均須釐清所有 AI 應用所涉及資料的來源與權利狀態。不僅標註和分級所擁有與外部取得的資料，更要將合法性評估程序 (如來源、授權、匿名化與保密等) ，確認每筆資料提供予 AI 模型訓練或調整使用前，已符合法律與合約要求。

此外，重要資料和 AI 生成成果的產出過程，也應該有紀錄保存機制。例如建立 prompts 及應用成果的可追溯性，使每一筆 AI 結果都能事後稽查，有助於日後責任釐清。建立資料治理機制才能從源頭降低因資料所生的法律風險。

(二) 強化人員 AI 素養

熟習生成式 AI 技術原理與運作，不僅是館員的能力培養，也是風險預警的關鍵；多數新型法律風險來自對 AI 技術與流程的不熟悉，因此定期教育訓練、跨部門交流、實務案例研討都是基礎。若圖書館希望作為讀者正確使用 AI 工具的引導者或作為民眾在假訊息滿天飛最後的堡壘，自然也必須掌握容易用來生成假訊息的 AI 工具。

(三) 標示機制與人類覆核

生成式 AI 產生的各類內容或資訊，包含行政文件、知識服務、對外溝通等，無論風險大小，均應要求標示使用 AI 生成之相關訊息，方便後續接觸該等內容的內部或外部人士，判斷應採取何種標準審視，以及應否信任該等內容，減少誤導與法律爭議。此外，對於任何涉及權利義務或影響較多民眾權益的決策或資訊內容，務必設置人類

覆核、介入制度的機制，畢竟，AI 是無法負責的，舉凡需要負責任的事情，還是應該由人類來完成。

(四) 定期風險評估

以這二、三年 AI 領域技術進步幅度之快，圖書館大概很難等到技術已接近成熟或穩定時，才會迎來是否導入該等技術的壓力。「摸著石頭過河」可能會是圖書館導入生成式 AI 服務面臨的情境，在做決策前，不會有太多的成功案例可以參考，而導入生成式 AI 後，市場上技術的迭代更新，也需要圖書館保持一定的敏感度，需定期重新審視既有 AI 服務是否可能因新的技術而降低或昇高風險。