ChatGPT 熱潮下所引發生成型 AI 作品可否享有著作權問題之探討（上）

文／黃齡玉

壹 前言

OpenAI 所推出之聊天機器人 ChatGPT 爆紅，主要是因為其用途十分廣泛，且能以貼近一般人對話之方式互動。除了可能撼動 Google 搜尋引擎之霸主地位外[1]，更引發教育界之衝擊與正反兩極的看法，紐約市教育局已公布當地公立學校電腦和網路不再開放 ChatGPT[2]，而香港大學亦宣布全校禁用 ChatGPT[3]，擔心學生利用 ChatGPT 抄襲該軟體所製作之文本作為作業或論文；然而有學校持不同看法，如臺灣大學、陽明交通大學，則選擇在課堂上和學生討論其優缺點，引導學生正確使用的觀念，並製作輔助寫作之規範，讓該人工智慧軟體進行輔助教學及研究[4]。

筆者親自以多種用途嘗試 ChatGPT 的功能，對於其協助文本製作及搜尋整理之功能確實驚艷不已，惟其亦有其缺點及限制，如由於資料庫係使用 2021 年前之資料來訓練，對於 2021 年後之事件瞭解有限，且無法正確回答網路資料較少、屬於小眾或複雜艱深的專業問題，另外，由於訓練過程亦可能存在吸收垃圾、不正確或充滿偏見之資訊[5]。

然而，對於類似的文本生成型人工智慧，對於吾法律人衝擊最大的，當屬於在傳統以人類為創作發明主體之智慧財產權法律架構，尤其是著作權領域。傳統上，著作權歸屬於實施創作之著作人，併保障著作人之「著作人格權（moral right）」及「著作財產權」，故將著作人限定為自然人（human being）。但隨著此種生成型人工智慧（Generated AI）的發展，由人工智慧系統透過機器學習演算法生成之文本（text）或文章（essay），是否可得為著作權保護之客體？又如可為著作權保護之客體，該等著作權之權利義務主體又該歸屬於何人？即為相當重要之議題，實值吾人關注。

貳 機器學習與ChatGPT之訓練及語言生成模型

一、機器學習

欲瞭解生成型人工智慧（Generated AI[6]）如何實現創作，不得不先理解「機器學習」的原理及演進過程。「機器學習」，係以統計學為基礎，和人類學習的「經驗」類似，需透過大量資料來學習[7]。可分為三種類型：「監督學習（Supervised Learning）」、「非監督學習（Unsupervised Learning）」及「強化學習（Reinforcement Learning）」。「監督學習」會將資料與答案的配對輸入電腦，讓電腦學習資料的特徵與規則，反覆從問題的正確解答與錯誤解答中調整參數，提升找出正確解答之能力。而「非監督學習」即為讓電腦分析大量不知正確答案之資料，由電腦自行找出資料之結構及規則，優點在於可以處理無答案之非結構化資料。至於「強化學習」則是透過獎勵與評估基準而讓人工智慧學習獲得較多獎勵的方法，較接近「非監督學習」[8]。而在機器學習的發展進程中，「深度學習」應運而生。所謂「深度學習（Deep Learning）」為機器學習的一種分支型態，它模擬類似人腦的結構及功能，並使用人工神經網路來學習大量數據並做出複雜的決策。在深度學習中，使用多階層（四層以上）的類神經網路，每一層都學習識別資料的不同特徵。訓練網路的過程包括調整神經元之間的連接權重以提高準確性[9]。這項革命性的技術讓原本只能由人力輸入的參數設定（特徵擷取），改由透過機器學習來處理，故機器即可自己執行高度機器學習，此一劃時代的人工智慧技術，推進了第三波高峰發展期，使人工智慧漸漸可以朝著媲美人類，可以思考並做出適當反應之強 AI[10]發展。

二、ChatGPT 之訓練及語言生成模型

ChatGPT 是由 OpenAI 公司開發的自然語言處理（NLP）模型「GPT-3」[11]延伸出的「GPT-3.5」[12]，再基於人類反饋強化學習（Reinforcement Learning from Human Feedback，簡稱：RLHF，下同）進行訓練所出現的產品。

所謂「人類反饋強化學習」（RLHF），是一種用於深度強化學習（Deep Reinforcement Learning）的訓練方式。目的在於提高模型根據當前狀態和行為預測未來狀態和獎勵的能力，強化其於複雜環境中之決策能力。步驟如下：

（一）收集數據：首先，通過在網路環境中進行模擬收集數據。

（二）訓練模型：接下來，訓練模型根據當前狀態和動作預測未來狀態和獎勵。該模型結合使用監督學習和強化學習技術進行訓練。先採用監督學習，將 ChatGPT 模擬為一個 6 歲孩童的學習方式，即輸入大量具有提示詞的資料庫，再使用強化學習方式來標註較佳之輸出答案，正確予以獎勵。

（三）使用模型生成輸出資料：模型經過訓練後，可用於透過預測不同狀態下不同輸入的結果來生成輸出。該等輸出資料用於擴充原始資料集並提高模型性能。

（四）訓練新模型：使用擴充後的資料集訓練新模型，新模型預測未來狀態和獎勵之精準度更佳，並可在複雜環境中做出更準確的決策。

（五）反覆：反覆該過程，使用新模型形成額外的資料並訓練新模型，直到模型的性能達到令人滿意的水準。

（下期待續）

（作者為臺灣高等法院台中分院法官）

註釋

1 微軟大舉投資 ChatGPT 之母公司 OpenAI，於 112 年 2 月 7 日還正式推出採 Chat GPT 技術強化之新版 Bing 搜尋功能和 Edge 瀏覽器，導致 Google 於隔天即推出聊天機器人 Bard 迎擊，意味一場「AI 結合搜尋引擎」之大戰正式開打。鉅亨網編譯羅昀玫「Google 定調 AI 聊天機器人 Bard 不僅是搜尋而已」(2023). Retrieved 4 March 2023, from https://news.cnyes.com/news/id/5103568 比爾蓋茲：AI 產業將威脅 Google. (2023). Retrieved 4 March 2023, from https://hk.finance.yahoo.com/news/%E6%AF%94%E7%88%BE%E8%93%8B%E8%8C%B2-ai%E7%94%A2%E6%A5%AD%E5%B0%87%E5%A8%81%E8%84%85google-055010625.html

2 科技新報「人工智慧淪為寫作業神器，紐約公立學校禁用 ChatGPT」(2023). Retrieved 4 March 2023, from https://technews.tw/2023/01/06/nyc-education-department-blocks-chatgpt-on-school-devices-and-networks/

3 中央社「論文、報告全都是 ChatGPT 寫的！香港大學下令禁用，台大會跟進嗎？」(2023). Retrieved 4 March 2023, from https://www.bnext.com.tw/article/74168/chatgpt-tw-unifuse

4 聯合新聞網「是作弊嗎？台清陽明交大校方怎看 ChatGPT」(2023). Retrieved 4 March 2023, from https://www.inside.com.tw/article/30791-ChatGPT-AI-school；註 3 資料。

5 「《AI 大戰》ChatGPT 怎麼玩？優缺全解析，一文看懂」(2023). Retrieved 4 March 2023, from https://www.cw.com.tw/article/5124860
「ChatGPT 是什麼？6 個關鍵問題一次看」(2023). Retrieved 4 March 2023, from https://www.cw.com.tw/article/5123907

6 此處的「生成型人工智慧」定義，參考世界智慧財產權組織（WIPO）之定義為人工智慧在無人為干預的情況下生成輸出。在這種情況下，AI 可以在運行期間改變其行為以響應意外訊息或事件。與在實質上人為干預和（或）指導下產生的「人工智慧輔助」輸出不同。WIPO/IP/AI/2/GE/20/1 REV., May 21, 2020. (2023). Retrieved 23 March 2023, from https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_2_ge_20/wipo_ip_ai_2_ge_20_1_rev.pdf

7 所謂「機器學習」就是人工智慧自我學習的技術。參見三津村直貴著（陳子安譯），「圖解 AI 人工智慧大未來：關於人工智慧一定要懂得 96 件事」，頁 74，旗標科技股份有限公司（2018）。

8 坂本真樹著（陳朕疆譯），「AI 必修課 -- 日本情感研究權威的人工智慧秒懂攻略」，頁 80-81、90-96，世茂出版（2018.10）

9 Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444. https://doi.org/10.1038/nature14539

10 「強 AI」又稱「通用型 AI」，指的是「擁有和人類一樣智慧的人工智慧」，目標是要用人工智慧來做到所有人類可以做到的事情。反之，則為「弱 AI」（特定用途 AI），著重於特定用途的人工智慧。故現在的 AI 幾乎都是「弱 AI」，其定義參見前註 7，頁 56-57。

11 GPT 是 Generative Pre-trained Transformer 的簡稱，中文全名為生成型預訓練變換模型。GPT-3（Generative Pre-trained Transformer 3，簡稱 GPT-3）是一個自迴歸語言模型，基於谷歌開發的語言模型，模型訓練內容大部分來自網路，讓 ChatGPT 能夠相當自然地組織語句。
GPT-3 這個模型也是當前最大的自然語言處理模型之一，神經網路包含 1750 億個參數，為有史以來參數最多的神經網路模型。資料來源為 570 GB 的文本。主要是網路上抓來的文字、英語維基百科。用途包含生成文本，執行語言任務，如翻譯、總結等，以及完成編碼任務、玩遊戲和回答問題等。【GPT-3 LINE BOT】(1) 前情提要 介紹 ChatGPT&GPT-3 - iT 邦幫忙：一起幫忙解決難題，拯救 IT 人的一天 . (2023). Retrieved 4 March 2023, from https://ithelp.ithome.com.tw/articles/10310593

12 GPT-3.5 則是由 GPT-3 微調出來的版本，而其中 GPT-3.5 使用與 GPT-3 不同的訓練方式，所產生出來不同的模型，比起 GPT-3 來的更強大。同前註 11。然而，於本文完成前，OpenAI 公司於 2023 年 3 月 14 日推出 GPT-4 即生成型預訓練變換模型 4，根據 OpenAI 發布的展示影片中，號稱 GPT-4 新增六大功能：（一）能辨識視覺圖像，轉成文字內容，（二）考試表現更佳，高出 GPT-3 一個級別，（三）更能處理非英文語言，（四）可扮演特定角色互動，（五）降低給出錯誤回應的機率，（六）能針對有安全疑慮的問題給予規勸。惟僅有 ChatGPT Plus 付費用戶可優先排入試用新版。參見：OpenAI. [2303.08774] GPT-4 Technical Report. (2023). Retrieved 19 March 2023, from https://arxiv.org/abs/2303.08774；GPT-4 登場！ OpenAI 這次新增哪些功能？與 GPT-3.5 差異一次看 . (2023). Retrieved 19 March 2023, from https://www.businessweekly.com.tw/focus/blog/3011904

JUDICIAL 司法
徵稿啓事
一、本刊歡迎各界賜稿，內容應符合學術倫理規範，請勿違反學術社群共同接受的行為準則；並請適當遮掩個資及其他法定限制公開事項。
二、作者投稿時，請檢附個人資料表（含作者姓名、服務機關、職稱、聯絡電話、電子信箱、通訊地址）。
三、文稿內容應符合學術倫理規範，請勿違反學術社群共同接受的行為準則；並請適當遮掩個資及其他法定限制公開事項。
四、文稿經採用刊登後，即視為作者同意授權本刊得以紙本或電子形式重製、發行、公開傳輸、公開展示、編輯、翻譯、改作、散布或再授權他人利用。
五、文稿一經採用，即贈送該期刊物。
六、文稿除紙本刊登外，亦置司法院網站；並酌計稿酬，寄贈該期刊物。
七、刊出文稿文責自負。
一律不退件，請自留底稿。稿件如經採用，文責自負。