

首宗人工智能侵權案和解 每本書支付2.34萬加利息

Anthropic用盜版書訓練AI 賠償117億

新聞追蹤 美國首宗人工智能（AI）侵犯版權案達成庭外和解。美國聯邦法官於當地時間25日批准，人工智能初創公司Anthropic支付15億美元（約合117億港元）進行和解，以解決因利用盜版書籍數據庫訓練AI而面臨的集體訴訟。這是美國史上侵犯版權案件中的最高賠償金額，也是使用生成式AI訓練數據引發版權爭議的首宗和解案例，可能為其他類似版權爭議提供解決方法。

大公報記者 郭嘉

【大公報訊】以聊天機器人Claude著稱的人工智能（AI）初創公司Anthropic，去年被驚悚小說家巴茨、非虛構類作家格雷伯及約翰遜代表一群作家及出版商提起集體訴訟，指控Anthropic未經許可，非法複製他們的書籍作品來訓練其Claude。

今年6月，加州北區聯邦地區法院法官阿爾蘇普裁決，無論是購入還是盜版的書籍，Anthropic用來訓練Claude模型的方式已對作品進行轉化，依法屬於「合理使用」，但另一方面，該公司在知情的情況下仍從網上圖書館下載700萬本盜版書籍，並把部分內容投入於訓練AI，則屬於侵權，另外立案處理。

新的案件原定於12月宣判，但Anthropic尋求盡快與原告達成庭外和解。9月25日，法官初步批准了15億美元的和解協議方案，以解決涉及50萬本書籍的版權爭議。

Anthropic「明知故犯」下載盜版書訓練AI

Anthropic副總法律顧問斯里達爾回應稱，對法院初步批准和解感到高興，這項裁定將讓Anthropic專注於開發安全的AI系統。外界

認為，對Anthropic來說，此次訴訟於他們而言已是一大成功。

根據該協議，Anthropic將向涵蓋的50萬本書的作者和出版商，每部支付約3000美元（2.3萬港元）的賠償金，另再加利息，是美國著作權法規定最低法定賠償金的4倍。不過，該協議並不適用於這些作家和出版商未來的作品和書籍。作為和解協議的一部分，Anthropic還將銷毀原始盜版檔案及任何副本，但公司仍保有合法購買並掃描書籍的權利。

美媒此前估計，即使僅計入在美國註冊版權、約五分之二的書籍，每本書最低賠付750美元，也可能造成15億美元賠償；若依最高標準，賠償總額恐達7500億美元（約合5.9萬億港元）。Anthropic在2021年成立，旗下聊天機器人Claude是OpenAI旗下ChatGPT的主要競爭對手之一。該



▲原告之一的驚悚小說家巴茨。美聯社

公司此前才獲130億美元新注資，目前市值約1830億美元。Anthropic的律師此前坦言，若以7500億美元賠付完全有可能「毀滅」Anthropic。

不過，以最終的和解協議來看，在刪除了重複和無版權作品後，Anthropic需要賠付的書籍數量從700萬本降為約50萬本，每件作品僅賠3000美元加利息，仍遠低於法定最高賠償金額的每件15萬美元，故外界認為Anthropic爭取到達成和解，是避過破產一劫。

或為類似訴訟提供解決方法

這項和解協議成為生成式AI訓練數據版權爭議的首宗和解案例。美國出版商協會形容，這項和解「朝着正確方向邁出了重要一步，讓人工智能開發者為肆無忌憚的侵權行為負責」。原告小說家巴茨表示，這起訴訟的結果代表，作家和出版商共同向AI公司傳遞了一個信息：你們不能凌駕於法律之上，我們的知識產權也不是你們可以隨意奪走的。

雖然此案被廣泛認為是科技公司的勝利，但其仍是AI產業發展過程中的重要法律里程碑，涉及AI公司如何合法使用受版權保護內容進行模型訓練的核心問題。《紐約時報》援引法律界分析，指因案件尚未進入審訊階段，此次和解不可成為法律先例，但對同類案件和AI公司行為仍會構成影響。

法律專家認為，該案的解決方式可能影響未來AI公司與版權持有人之間的合作模式，促使業界建立更完善的內容授權機制。

（綜合報道）



大公報製圖

▲人工智能公司Anthropic的執行總裁阿莫迪。法新社

OpenAI深陷版權爭議 與多家媒體對簿公堂

版權挑戰

人工智能（AI）的浪潮席捲全球，但也帶來了前所未有的版權挑戰。美國AI巨頭OpenAI因為未經授權使用大量新聞出版商的內容，用於訓練開發AI聊天機器人，在美國、加拿大和印度與多家大型出版商對簿公堂。

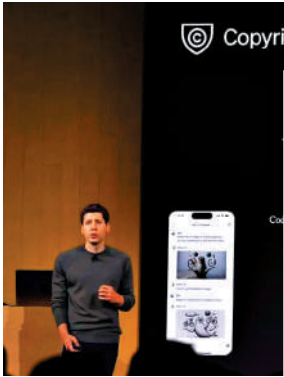
2023年12月，《紐約時報》控告OpenAI和微軟在未經該報同意下，以其新聞報道訓練生成式AI，對該報造成數十億美元實際損害，成為首個美媒提出同類訴訟。此後，多家美國新聞媒體包括非營利新聞機構調查報道中心、The Intercept、《芝加哥論壇報》和其他7家報紙分別對OpenAI提起了同類訴訟。

包括《環球郵報》和加拿大廣播公司在內的加拿大新聞媒體於去年入稟法院，指控OpenAI大規模「搜刮」新聞內容用於訓練AI，卻未支付相應費用。他們要求每篇文章賠償2萬加幣（約合11萬港元），總額可能高達數十億美元。

印度媒體與OpenAI法律戰同樣升溫。印

度亞洲國際新聞社去年率先提告OpenAI，隨後全球和印度出版業於今年加入戰局。另外，印度億萬富豪阿達尼旗下新德里電視台、安巴尼的電視集團Network18等在內的印度新聞媒體同樣希望加入控告OpenAI侵權的行列。

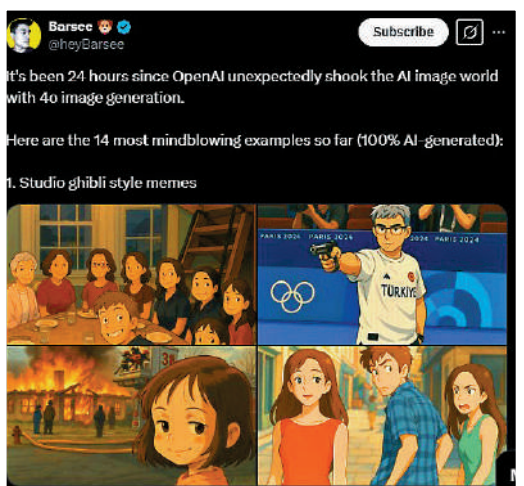
OpenAI在「侵權」的同時，又不斷與媒體合作，以解決版權問題。OpenAI已與美聯社、新聞集團、法國《世界報》等數十家媒體及出版商簽署協議，以取用這些媒體集團所擁有的龐大文本數據庫，訓練AI大模型，同時規避潛在



► OpenAI旗下AI模型GPT-4o生成了大量酷似日本知名動畫工作室「吉卜力風格」的圖像，引發爭議。網絡圖片

的版權風險。OpenAI與谷歌（Google）在今年2月曾致函特朗普政府，給予AI公司豁免權，允許它們使用受版權保護的材料來訓練模型。

◀ OpenAI行政總裁阿爾特曼。網絡圖片



AI版權爭議Q&A

Q AI模型的版權爭議主要是什麼？

• 版權方往往主張AI公司對生成式AI的訓練過程複製了其作品（即使是非公開使用），侵犯了複製權和改編權，並生成其競品，威脅版權方生計。AI開發方則主張利用版權作品訓練AI模型，生成「轉化性質」新內容屬於版權法中的「合理使用」；主張訓練是技術必要性而非侵權。

Q 如何判斷是否「合理使用」？

• 合理使用（Fair use）是指在某些特定情況下，無需著作權人許可，就可以在合理範圍內使用受著作權保護的作品。

Q 「合理使用」的判決案例？

• 今年2月，美國特拉華州地方法院裁定媒體巨頭湯森路透勝訴，認為AI初創公司Ross Intelligence的行為構成侵權。在本案中，法官認為被告對湯森路透旗下法律研究公司Westlaw數據的使用目的與原告並無本質不同，是專門用於開發與Westlaw直接競爭的同類法律檢索工具，從而認定被告對其版權內容的商業使用不具有「變革性」，不屬於合理使用。

• 今年6月25日，針對13名作家於2023年在美入稟控告美國科技巨頭Meta未經授權利用其書籍作品訓練AI模型Llama一事，加州北區聯邦法院裁定，Llama的使用情況屬於「轉化性合理使用」，符合美國版權法中的「合理使用」原則，Meta因此勝訴。

Q Anthropic的侵權案有何不同？

• 在Anthropic的AI模型訓練侵權一案中，加州法院在今年6月26日的判決中，支持將版權作品用於AI模型訓練，認為這屬於「合理使用」，因為其具有「轉化性」的本質。Anthropic旗下的AI模型只是學習了版權作品寫作的模式、風格和結構，這與人類學習並無本質不同。不過，Anthropic在明明知情的情況下仍從網上圖書館下載700萬本盜版書籍，並把部分內容投入於訓練AI，這就屬於侵權。

AI侵權風波延燒 蘋果亦捱告

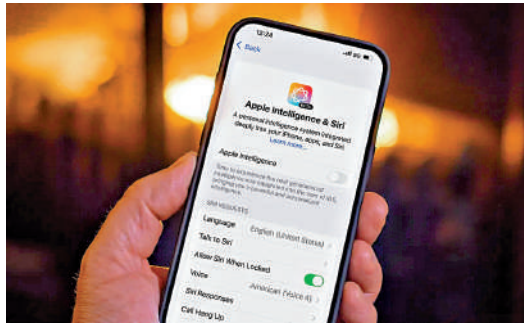
訴訟浪潮

隨着生成式AI技術快速發展，越來越多的科技公司陷入版權爭議。蘋果近期也遭兩名作家提起集體訴訟，指控蘋果在未經授權的情況下，複製並使用他們的作品用於訓練旗下AI模型。

美國作家亨德里克斯和羅伯遜9月5日在加州北區法院提起集體訴訟。訴狀中指出，蘋果公司在未經同意、沒有授權或提供補償的情況下複製了他們的作品，將這種內容用於AI模型訓練，並且用於其Apple Intelligence服務，而這之中還涵蓋了大量未經授權流通的盜版電子書。

訴訟文件顯示，蘋果公司使用了由開源社區

整理的爭議性數據集「Books3」。「Books3」數據集建立在Bibliotik等「影子圖書館」的基礎上，而這些網站內包含數千本盜版書籍。



▲蘋果遭作者提告疑似使用盜版書籍訓練Apple Intelligence。網絡圖片

據報道，由於「影子圖書館」中涉及的數據和作家數量龐大，亨德里克斯和羅伯遜正在尋求將該訴訟作為針對蘋果公司的集體訴訟。

此外，訴訟還指控蘋果公司通過自家開發的網絡爬蟲「Applebot」抓取網頁，來訓練旗下AI模型「OpenELM」。訴訟認為，這些AI模型輸出的作品進入市場，與原作者的作品競爭，並擾亂市場。

業內人士指出，隨着科技公司越來越多地依賴大規模出版數據訓練AI模型，出版行業正面臨前所未有的挑戰。這些案件涉及合理使用、報酬與授權等根本性問題，未來判例或將深刻影響AI發展與版權保護的邊界。