

文化創意永續與技術發展並進：我國著作權法 資料探勘例外之增修可行性

李姿儀*、李諺樺**

壹、創作者與人工智慧共存：臺灣文化生態的新挑戰

- 一、創作者在人工智慧時代的處境與機會
- 二、現行著作權法授權制度的結構性問題
- 三、負責任 AI 發展的必要性

貳、資料探勘例外的法理基礎與技術特性

- 一、從傳統資料探勘到生成式 AI 訓練的技術演進
- 二、創作者權益保護的新思維
- 三、平衡發展的制度正當性

參、國際經驗：保護與創新的平衡智慧

- 一、美國模式：市場導向下的創作者關懷
- 二、歐盟模式：例外與透明化的平衡機制
- 三、亞洲經驗：創新發展中的創作者考量
- 四、我國的政策選擇：「友善創作者之負責任 AI 發展」的平衡智慧

肆、以文化永續為核心的資料探勘例外制度設計

- 一、「豁免例外」：友善創作者之內涵
- 二、支柱一「透明性」：建立創作者與 AI 開發者的信任橋梁
- 三、支柱二「合理報酬與退出選擇」：創作者主體性的雙重保障
- 四、支柱三「著作權過濾」機制：責任分攤的協調設計

伍、結語：文化創意生態永續發展的新路徑

* 作者現為國立臺灣科技大學專利研究所助理教授，國科會「AI 人文法制基礎環境建置第二期計畫」共同主持人。本文係國科會研究計畫「AI 人文法制基礎環境建置第二期計畫」（NSTC 113-2634-F-002-006）補助支持，惟一切文責當由作者自負，不代表國科會立場。

** 作者現為國立臺灣科技大學專利研究所及東吳大學法律學研究所碩士生，曾擔任國科會「AI 人文法制基礎環境建置第二期計畫」兼任研究助理。
本文相關論述僅為一般研究探討，不代表本局及任職單位之意見。

摘要

生成式 AI 發展已對現行著作權法制帶來極大挑戰。臺灣正體中文及多元語料創作內容在全球 AI 訓練資料比重微乎其微，創作者面臨作品被未經授權使用風險的同時，亦可能在 AI 發展過程中被邊緣化。現行授權制度在面對 AI 大規模資料需求時已明顯傾斜，不僅增加法律不確定性，更可能導致內容市場集中化，損害創作生態多樣性與 AI 技術的創新活力。本文主張建立「友善創作者之負責任 AI 發展」平衡路徑，於著作權法增訂「資料探勘例外」條款及「三大支柱」配套機制：透明性機制確保創作者知情權與社會信任基礎；合理報酬與退出選擇機制提供創作者主體性雙重保障；著作權過濾機制防範侵權擴散。本文欲超脫零和思維，倡議保障創作者同時促進 AI 負責任發展的制度設計，實現創作生態與技術創新的協調並進，為臺灣在 AI 時代的文化永續發展提供法制因應之基礎，促進公共討論。

關鍵字：資料探勘、生成式人工智慧、負責任 AI、透明性、合理補償、退出選擇
Text and Data Mining、Generative AI、Responsible AI、Transparency、
Fair Compensation、Opt-out

壹、創作者與人工智慧共存：臺灣文化生態的新挑戰

一、創作者在人工智慧時代的處境與機會

從傳統資料探勘（Text and Data Mining, TDM）到今日生成式 AI 訓練的技術演進脈絡，使得資料探勘概念面臨新的詮釋需求。生成式 AI 訓練過程通常包含：大規模資料爬蟲（scraping）、資料預處理、特徵提取、模型訓練等階段¹，其中每個階段皆可能涉及受著作權保護作品的利用。這種技術演進使得各國原本針對學術研究或小規模資料分析設計的資料探勘例外與合理使用論述，必須重新審視其適用範圍與限制條件。

生成式人工智慧（Generative Artificial Intelligence，下稱生成式 AI）技術的快速發展，正在重新定義人類與資訊、創作的關係。一方面，AI 工具大幅降低了創作門檻，使更多人能夠參與創意表達，從文字創作到視覺設計，AI 正在民主化創作過程。另一方面，AI 系統大量使用既有作品進行訓練，卻往往未經著作權人與創作者同意，亦未提供合理補償，引發創作者的憂慮。對我國創作者而言，這種挑戰更加複雜。臺灣的正體中文創作內容在全球 AI 訓練資料中所占比例極低。據統計，全球現有大型語言模型（large language model, LLM）中有 73% 來自美國與中國大陸，且有 95% 主要以英文或英文與阿拉伯文、中文、日文的組合為訓練資料來源，儘管以英語為基礎的模型已在理解與翻譯區域語言方面有所進展，惟該些大型基礎模型仍存有英語中心偏誤²。其中的「中文」資料絕大多數為簡體中文，正體中文資料在全球 AI 訓練資料的比重微乎其微³。此意味著臺灣創作者不僅面臨其作品未經授權使用的風險，更可能在 AI 發展過程中被邊緣化，個

¹ OECD, *Intellectual Property Issues in Artificial Intelligence Trained on Scraped Data*, OECD Artificial Intelligence Papers, No. 33, 16 (2025) [hereinafter OECD Scraped Data Report].

² Elina Noor & Binya Kanitroj, *Speaking in Code: Contextualizing Large Language Models in Southeast Asia*, CARNEGIE ENDOWMENT FOR INT'L PEACE, 4-6, https://carnegie-production-assets.s3.amazonaws.com/static/files/Noor_LLMs_final.pdf (last visited May 3, 2025); Simon Huang, *Why Singapore's LLM Isn't Sweating GPT-4*, TECH IN ASIA, <https://www.techinasia.com/singapores-sea-focused-llm-isnt-sweating-gpt4> (last visited May 3, 2025).

³ 鍾秉哲，下一座護國神山？陳經農剖析台灣 AI 發展 3 大困境：已處在落後階段，風傳媒，<https://www.storm.mg/article/5031501>（最後瀏覽日：2025/05/03）。

別與集體的創作內容、文化內涵與表達方式，皆難以在全球 AI 系統中得到適當的再現⁴。

然而，AI 技術也為創作者帶來機會。若能建立適當法制框架引導，確保創作者在 AI 發展中獲得合理角色與回報，臺灣的文化創意不僅更得以保存與傳承，也能透過 AI 技術擴大影響力，讓臺灣正體中文與多元文化內容在數位世界中發揮更大作用。為此，關鍵在於如何設計制度，讓 AI 發展成為創作生態的助力而非威脅。

二、現行著作權法授權制度的結構性問題

我國現行著作權法制面對 AI 時代的挑戰，呈現的不僅是創作者保護不足問題，更是整體市場結構的扭曲。這種扭曲對創作者與 AI 開發者都造成不同程度的不利影響。

就個別創作者而言，現行制度確實存在明顯缺口。首先是資訊不對稱問題：AI 訓練過程往往涉及巨量資料，個別創作者難以發現其作品被利用，更遑論了解使用範圍、頻率與影響程度；即使發現被利用，也難以評估其經濟價值與潛在損害。其次，是談判能力不對等問題：面對資本雄厚的大型 AI 開發者，個別創作者缺乏議價能力，難以獲得合理授權條件。相較之下，大型內容集團或平台則具有較強議價能力，能與大型 AI 開發者進行較有利的授權談判⁵；這種差異可能導致創作市場進一步集中化，大型內容供應者獲得更多資源與機會，而個別創作者可能被邊緣化。第三，現行著作權法主要依賴事後救濟機制：創作者必須在發現侵權後，自行承擔舉證與訴訟成本。對大多數個別創作者而言，這種成本往往超過潛在收益，形成訴訟不經濟的困境。

⁴ 例如以中國大陸為來源的簡體中文資料，其內容皆需經其政府事前審查（或自我審查）過濾；另一方面，公開可得的正體中文歷史資料當中，亦有相當大比例屬於臺灣民主化前、受言論審查之內容。根據洪子偉等人研究，全球 14.5 億華語人口，繁體書寫者只占 2%、只有 1.4% 為不受思想政策審查，中國大陸長期箝制思想、進行思想審查，同時也壟斷全球 96% 華語出版市場；在如此比例懸殊的華語知識空間中，華語知識出現壟斷危機，「語言」成為認知戰的新戰場。參見廖昱涵，中研院學者洪子偉示警：中國壟斷 96% 華語空間，「語言」成認知戰新戰場，沃草，<https://watchout.tw/reports/wmwim1l1bkkTCWrTD6b6MM>（最後瀏覽日：2025/05/03）；Tzu-Chieh Hung and Tzu-Wei Hung, *How China's Cognitive Warfare Works: A Frontline Perspective of Taiwan's Anti-Disinformation Wars*, 7(4) J. GLOB. SEC. STUDIES 1, 7, 14 (2020).

⁵ 李姿儀，機器作為文化內容生產者：生成式 AI 對人類作者身分之挑戰，2023 臺灣智慧財產法學會著作權研討會，臺灣智慧財產法學等，頁 24，2023 年 10 月。

另一方面，AI 開發者同樣面臨結構性問題。我國著作權法第 65 條合理使用條款，雖提供一定彈性，但其四要素判斷標準在生成式 AI 訓練脈絡下充滿不確定性。從傳統資料探勘到生成式 AI 訓練的技術演進，使得既有法律概念面臨適用邊界模糊的問題。法院需要個案認定 AI 訓練是否構成合理使用。這種不確定性不僅增加相關利害關係人的法律風險，也可能阻礙創作與技術的發展。我國經濟部智慧財產局雖曾就 AI 訓練發布相關函釋，但是否有明確的法制調和或其他政策因應措施仍在研議階段。從產業發展角度觀察，法制不確定性已對全球 AI 產業造成實質影響，包含企業已開始評估各司法管轄區之法制交易成本，企業應如何因應與配置相關經濟資源⁶。

大型 AI 開發者憑藉雄厚資金與法務資源，能承擔複雜的授權協商成本，甚至可透過收購或合作，建立專有訓練資料庫⁷。相較之下，中小型 AI 開發者、學術研究機構與非營利組織等，往往缺乏足夠資源進行大規模授權協商。這種情況可能導致「AI 開發市場集中化」現象，只有資源充足的大型企業能夠合法取得高品質訓練資料，而創新能力強但資源有限的小型開發者及學研機構與非營利組織，可能因為授權成本過高而被排除在外，或傾向容易取得或法遵成本低（例如無需授權取得）的資料。而容易取得或法遵成本低的資料，不必然為好的訓練資料；更甚者，易於取得的資料集可能因資料偏誤（bias），使其 AI 系統在判斷與權重上存在系統性失衡，導致 AI 偏見與歧視效果⁸。這種現象不僅影響 AI 系統的品質，也可能加劇社會不平等問題。

⁶ Katharina de la Durantaye, *Control and Compensation: A Comparative Analysis of Copyright Exceptions for Training Generative AI*, 56 INT'L REV. INTELL. PROP. & COMPETITION L. 737, 764-765 (2025).

⁷ 除了創作者、內容平台與內容產業對全球 AI 公司的法律訴訟之外，近二年許多媒體集團如《紐約時報》、《華盛頓郵報》、《美聯社》、《Reddit》及音樂出版商，皆分別與 OpenAI、Amazon、Google、Meta 等達成授權合作。詳陳曦，新聞媒體與 AI 科技公司之間的合作與衝突：簽署協議與法律訴訟並行，卓越新聞電子報，<https://feja.org.tw/79159/>（最後瀏覽日：2025/07/11）；陳怡均，創先例！環球、華納等音樂巨擘擬協商 AI 作品授權模式，工商時報，<https://www.ctee.com.tw/news/20250603700807-430704>（最後瀏覽日：2025/07/11）。

⁸ See e.g., Alejandro Salinas, et al., *What's in a Name? Auditing Large Language Models for Race and Gender Bias*, ARXIV (Feb. 29, 2024), <https://arxiv.org/abs/2402.14875> (last visited Jun. 26, 2025); Thomas Margoni, et al., *Algorithmic Propagation: How the Data-Platform Regulatory Framework May Increase Bias in Content Moderation*, note 34, in FLEXIBILITIES IN COPYRIGHT LAW (Caterina Sganga & Tatiana Eleni Synodinou eds, forthcoming 2025), Available at SSRN: <https://ssrn.com/abstract=4913758> (last visited Jun. 26, 2025).

這種雙重集中化現象恐將導致惡性循環，大型內容平台與大型 AI 開發商形成專屬合作關係，排擠個別創作者與小型開發者的參與空間。長期而言，這不僅會減少創作誘因與文化多樣性，也會限制 AI 技術的創新活力，最終損害整體創意生態的健康發展。

因此，著作權法制的調整不僅為了保護創作者權益，也為了促進 AI 發展的多元化與民主化⁹。好的著作權法制設計應讓個別創作者獲得合理回報，降低中小型 AI 開發者的合規成本，防止市場過度集中化，最終促進創作生態與技術創新的良性循環。這正是本文倡議建立著作權法資料探勘法定例外制度，搭配三大支柱配套機制的重要動機，即透過制度的重構與創新，打破現有的結構性困境，為參與者創造更公平、更有利於文化創意永續的生態體系。

三、負責任 AI 發展的必要性

面對創作者的合理關切，AI 開發者不應將其視為發展阻礙，而應認知到與創作者建立良性關係的重要性。負責任 AI 不僅是倫理要求，更是技術與產業永續發展的必要條件。

首先，創作者 is AI 發展的重要合作夥伴而非競爭對手：高品質 AI 模型需要優質訓練資料，而這些優質資料正是創作者的心血結晶。忽視創作者權益不僅有道德非難性，長期而言也會影響創作激勵，減少優質內容的生產，終將損害 AI 發展的資料基礎。其次，建立友善創作者的 AI 發展環境有助於提升產業形象與社會接受度：當前社會對 AI 技術的憂慮，多數來自 AI 可能取代人類創作、損害生計的擔憂。透過制度設計確保創作者合理權益，可降低社會阻力，為 AI 產業創造更好的發展環境。第三，負責任 AI 發展符合國際趨勢與監管期待：諸如經濟合作暨發展組織（Organization for Economic Cooperation and Development, OECD）人工智慧相關報告書、歐盟《人工智慧法》（AI Act, AIA）、美國 AI 相

⁹ 依據經濟合作暨發展組織（Organization for Economic Cooperation and Development, OECD）AI 治理原則（AI Principles），AI 之發展應尊重人權和民主價值。AI 民主化（AI Democratization）在本文脈絡下，泛指將 AI 的工具、資料與知識從少數技術掌握者手中釋放，讓不具備專業技術背景的普羅大眾也能使用、開發並參與決策。其核心目標在於降低技術門檻，縮短數位落差，使 AI 成為人人皆可負擔且能推動創新與社會公平的普及化資源。參 OECD AI principles, <https://www.oecd.org/en/topics/ai-principles.html> (last visited Dec. 28, 2025).

關政策指引與報告書等，儘管路線不盡相同，都強調 AI 發展必須考量倫理與社會影響¹⁰。我國若能及早建立創作者友善的 AI 治理框架，不僅符合國際 AI 治理趨勢，也能有意義地參與其中。

貳、資料探勘例外的法理基礎與技術特性

一、從傳統資料探勘到生成式 AI 訓練的技術演進

本文在歐盟《數位單一市場著作權指令》（Directive on Copyright in the Digital Single Market, CDSMD）基礎上，將「資料探勘」定義為「透過自動化分析技術，對包含多數著作或大量資訊進行解析（提取規則、構造、傾向、相關關係等資訊）」¹¹。根據 OECD 於 2025 年 2 月相關報告指出，現代 AI 訓練過程涵蓋了更廣泛的技術範疇。該報告將「資料爬蟲」（data scraping）定義為使用自動化技術，從第三方網站、資料庫或社群媒體平台擷取資訊¹²，其特徵包含自動化、大規模性（scalability），且通常缺乏與資料持有者協調（lack of

¹⁰ OECD, *AI Language Models: Technological, Socio-Economic and Policy Considerations* 10, OECD Digital Economy Papers, No. 352 (2023); Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulation (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance), 2024 O.J. (L 2024/1689) [hereinafter EU AIA] art. 13(3)(b)(iv), (vii), and (f); art. 50; art. 53(1)(d); US COPYRIGHT OFFICE, COPYRIGHT AND ARTIFICIAL INTELLIGENCE, PART 3: GENERATIVE AI TRAINING (PRE-PUBLICATION VERSION) 70 (2025) [hereinafter USCO 2025 report].

¹¹ 參照歐盟 CDSMD 第 2(2) 條定義 “[A]ny automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations.” Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (Text with EEA relevance) [hereinafter CDSMD], art. 2(2)。另 OECD 定義資料探勘（data mining）為「用於識別大型資料集中的模式（patterns）和相關性（correlations）的運算過程」，參 OECD Scraped Data Report, *supra* note 1, at 14 (2025)；ISO/IEC 22989 則定義資料探勘（data mining）為「透過從不同角度（perspectives）和維度（dimensions）分析量化資料，並對其進行分類並總結其潛在關係（potential relationships）和影響（impacts），以提取模式（patterns）的運算過程」，參 ISO/IEC 22989, *Information technology — Artificial Intelligence — Artificial Intelligence Concepts and Terminology* 3.1.11 (2022), <https://www.iso.org/standard/74296.html> (last visited Jun. 26, 2025).

¹² OECD Scraped Data Report, *supra* note 1, at 9 (2025).

coordination) 等特性，例如使用爬蟲技術擷取網站內容。資料爬蟲著重於資料的擷取 (extraction)，包括透過爬蟲技術進行資料蒐集 (collection)、資料預處理 (pre-processing) 和資料儲存 (storage) 等部分，通常涉及從外部來源獲取原始資料¹³；而「資料探勘」(TDM) 則指用於分析資料的自動化技術。資料探勘的目的在於從資料集 (包括透過爬蟲和／或其他資料擷取方法獲得的資料集) 分析出洞見 (insights)¹⁴。在 AI 模型訓練的脈絡下，資料探勘通常包含四個階段：資料識別、資料複製、提取資料、重新組合¹⁵。事實上，資料爬蟲的技術早在生成式 AI 之前就已經存在數十年，它是網路時代非常基礎且廣泛使用的資料採集手段，以進行網路資料之索引或分析 (例如常見網路搜尋引擎)。簡言之，資料探勘是從資料中分析模式與洞見 (insights) 的技術；資料爬蟲則是自動化獲取原始資料的手段，並非生成式 AI 所獨有。當代多數的生成式 AI 訓練中，資料爬蟲通常是資料探勘的前置行為，透過大規模爬取網路資料進行模型訓練與微調，以實現資訊的創造而非僅是獲取。

從創作者角度觀察，最值得關注的是資料複製階段，可能涉及對受著作權保護作品的重製行為。然而，此處的關鍵在於理解此技術涉及重製的目的與性質——並非為「閱讀」或「欣賞」等「享受」作品的目的¹⁶，而是透過數學運算分析作品中的統計模式、語言結構、或視覺特徵¹⁷。雖然 AI 系統的訓練在技術上進行了疑似「重製」行為，其目的是數學分析而非對著作內容表達之感官享受。這種利用方式與傳統著作權侵害 (如盜版、抄襲小說、畫作、音樂，以閱讀、欣賞、聆聽其著作內容)，在本質上存有差異。認識這種差異，有助於設計更精準的政策與法律機制。

¹³ *Id.*, at 13-18.

¹⁴ *Id.*, at 16.

¹⁵ *Id.*, at 9.

¹⁶ 日本 2018 年修訂著作權法，允許「非以享受著作所表達之思想或感情為目的之利用」，對著作權作品進行資料分析；南韓於 2021 年亦有國會議員提出著作權法資料探勘例外條款之提案，主張資料探勘行為在某些要件下將屬合法。詳本文第參章之討論。

¹⁷ See A. Feder Cooper and James Grimmelman, *The Files are in the Computer: On Copyright, Memorization, and Generative AI*, 9 CHICAGO-KENT L. REV. (forthcoming 2025), <https://arxiv.org/pdf/2404.12590> (last visited May 15, 2025); Desta Haileselassie Hagos et al., *Recent Advances in Generative AI and Large Language Models: Current Status, Challenges, and Perspectives* 2 (2024), <https://arxiv.org/abs/2407.14962> (last visited May 15, 2025).

二、創作者權益保護的新思維

AI時代的創作者權益保護，需要超越傳統的「禁止」與「合理使用」例外之思維，發展更具前瞻性的保護與互惠機制。傳統著作權法主要透過排他性權利保護創作者，賦予著作權人禁止他人未經授權而使用其作品的權利。然而，在生成式AI大規模資料訓練的脈絡下，這種「全有或全無」的保護模式可能導致僵局，亦即「須個別授權而得取得報酬」與「合理使用然不得取得報酬」之二種模式¹⁸，恐造成若非阻礙技術發展，則損害創作者權益的零和思維。

本文主張新思維之核心，在於發展「參與式保護」機制，確保創作者在AI發展過程中不被排除，而是共同參與者。這包含三個層面：首先是「資訊參與權」，創作者有權知悉其作品被如何使用於AI訓練，這不僅是知情權的體現，也是後續權利行使的基礎。其次「經濟參與權」，當AI系統從創作者作品中獲得經濟價值時，創作者應獲得合理經濟回報。這種回報不必然以個別授權形式實現，也可透過集體補償機制或公共基金模式達成。第三則是「控制參與權」，創作者應保有對其作品使用的合理控制權，包含在特定情況下要求停止使用的權利。這種控制不應為絕對權利，但應在合理範圍內受到尊重。

三、平衡發展的制度正當性

建立資料探勘法定例外制度，搭配合理的創作者保護機制，具有多重正當性基礎。

首先是公共利益的考量。AI技術在教育、醫療、科學研究領域展現極大潛力，其社會效益可能遠超過對個別創作者的影響。但這不意味著可以忽視創作者權益，而是應透過制度設計確保社會公共利益與創作者利益的雙贏。其次，在AI大規模資料需求下，傳統個案授權市場出現明顯交易成本過高與資源配置偏斜，導致市場失靈，而法定例外制度可以作為市場機制的補充。透過明確的法律框架降低交易成本，並透過補償機制確保創作者獲得合理回報。第三，過度的權利保護可能阻礙技術創新，但過度的技術發展也可能損害創作激勵。關鍵在於找到平衡點，既保護創作者權益也促進技術發展，確保兩者能夠相互促成，而非相互對

¹⁸ de la Durantaye, *supra* note 6, at 744-745.

立。最後，國際競爭與合作的考量上，多數國家皆正重新檢討 AI 時代的著作權規範，我國需在國際經驗與趨勢基礎上，發展符合在地特色與脈絡的制度模式。

建立平衡的資料探勘例外法制框架，不僅是技術發展的需要，更是創作者權益保護的創新路徑，其關鍵就在於透過審慎、細緻的制度設計，實現創作生態與技術發展的協調並進。

參、國際經驗：保護與創新的平衡智慧

一、美國模式：市場導向下的創作者關懷

美國現行著作權法主要透過「合理使用」（fair use）四要素測試來處理 AI 訓練相關爭議。然而，這種個案判斷的司法實踐模式，在創作者保護方面呈現明顯不確定性問題。從創作者角度觀察，美國模式主要問題在於其「事後救濟」特性。AI 訓練往往在創作者不知情的情況下進行，發現時模型已訓練完成。創作者必須等到發現侵權、提起訴訟，經過漫長司法審理程序，才能獲得權利救濟。

被控侵權者往往主張「合理使用」抗辯其使用。自 *Campbell v. Acuff-Rose Music* 案將「轉化性（Transformativeness）」要件加入「合理使用」判斷後，雖非合理使用明文四要素之一，仍常被作為主要判斷要素¹⁹，法院並認為若轉化程度越高，則其他要素之重要性就越低。例如可能不利於成立合理使用的商業性質利用，後續透過 *Google Books* 等案重申「轉化性」之意義係為從原著作傳達出新的、不同的東西，或延伸原著作之效用，以達成著作權法增進公眾知識的最終目的²⁰。然近年之判決似乎有將判斷重心回歸合理使用四要素之趨勢，商業性質似又重新受重視²¹，且法院認為區分是否為商業性質之關鍵並非在於利用行為是否係單純為獲取金錢，而是在於利用人是否為了營利而利用了有著作權之題材卻未支付通常應付之費用²²。

¹⁹ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994).

²⁰ *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015).

²¹ 胡心蘭，從近期案例剖析美國著作權法合理使用原則新發展，中正財經法學 22 期，頁 33-46，2021 年 1 月。

²² *Harper & Row Publishers, Inc. v. Nation Enters., Inc.*, 471 U.S. 539, 562 (1985).

在面對生成式 AI 訓練所涉及之大規模著作利用情境時，先前之案例將如何適用即引發爭議，有學者主張美國法院過去已針對其他仰賴複製技術的非表達性使用（non-expressive uses）之合法性進行審理，諸如軟體逆向工程、抄襲偵測軟體，並認為資料探勘與其他非表達性使用相似，其重製行為通常並未侵犯著作權人就其原始表達向公眾傳達之控制利益，因該重製僅為分析程序的第一步²³。美國著作權局（U.S. Copyright Office, USCO）2025 年 5 月預發布的 *Copyright and Artificial Intelligence Part 3: Generative AI Training* 報告（下稱「USCO 2025 年報告」），雖承認 AI 訓練可能具有轉化性，但也強調須個案判斷，並特別關注 AI 生成內容對原作品市場的潛在影響²⁴。該報告認為生成式 AI 訓練過程中對著作權作品的使用有可能具有轉化性，然是否構成合理使用將取決於所使用的作品類型、取得來源、使用目的以及是否對輸出結果設有限制。若是透過非法取得並將其用於生成可在現有市場中與原作品競爭之內容，則已逾越既有合理使用的界限；最後以自願授權市場的快速擴張以及利害關係人普遍不支持立法之考量，主張現階段政府介入尚非時宜，而應先讓授權市場持續發展，故採取維持現狀之見解²⁵。

本文認為 USCO 2025 年報告立基於美國為文化內容輸出大國，具備強大資本與談判能力，並忽略資料探勘之技術特質（即以數學方式分析數位形式的文字和資料以產生洞見資訊的自動化分析技術），實已涵蓋「非以享受著作所表達之思想或情感目的，且該利用未涉及以人類知覺認識該著作之表達」²⁶之差異。該報告見解無法有意義地詮釋當今 AI 技術已觸發的文化創作變革與衝擊，也無益於平衡創作人權益與技術創新之困境，更無法創造有利於文化永續的環境，特別是如臺灣弱勢文化與資本市場之生態。

²³ Matthew Sag & Pamela Samuelson & Christopher Jon Sprigman, *Comments In Response To The Copyright Office's Notice Of Inquiry On Artificial Intelligence And Copyright*, 5-9 (2024).

²⁴ USCO 2025 report, *supra* note 10.

²⁵ 前揭報告似已顯示美國行政機關立場，惟該報告於 2025 年 5 月 9 日發布後，隔日美國總統川普即開除時任著作權局局長希拉·珀爾馬特（Shira Perlmutter），故未來著作權局見解是否變更，宜持續追蹤。Scott MacFarlane, *Trump fires director of U.S. Copyright Office, sources say*, CBS News (May 10, 2025), <https://www.cbsnews.com/news/trump-fires-director-of-u-s-copyright-office-shira-perlmutter-sources/> (last visited Jun. 6, 2025).

²⁶ 參照日本著作權法第 30 之 4 條。

同時，美國法院實務上在面對「合理使用」抗辯分析時，往往會考慮被控侵權之使用是否「取代正常市場」或「取代原作品的（市場）需求」，包括「潛在授權衍生著作的需求」，顯示合理使用分析已將尚未存在或潛在的授權市場納入考量；換言之，當權利人主張侵權人未經授權使用導致其市場受損時，已構成循環論證，只要權利人表達了授權意願，合理使用抗辯就難以成立²⁷。因此，USCO 2025 年報告對於 AI 訓練之合理使用觀點，尚須更堅實的理論與實證基礎。2025 年 6 月美國聯邦地方法院兩個值得關注的判決，亦顯示美國法院對 AI 訓練合理使用的不同觀點²⁸。整體趨勢而言，法院開始區分不同類型的資料來源與使用目的，但仍缺乏一致性標準，而機器學習似得透過「合理使用」原則取得合法地位，前提為其訓練資料係經合法取得。

更重要的是，美國模式欠缺主動的創作者保護機制。創作者無法事前（甚至事後）得知其作品是否被用於 AI 訓練，除了訴訟也缺乏有效的途徑獲得合理補償。這種情況下，創作者往往處於被動且弱勢的地位。

二、歐盟模式：例外與透明化的平衡機制

歐盟透過 CDSMD 建立的資料探勘（text and data mining）例外制度，在創作者保護方面展現較為前瞻的思維。

CDSMD 第 3 條為非營利研究組織與文化保存機構建立資料探勘例外，該等機構就有合法近用權限（lawful access）的著作，得為「科學研究目的」進行資料探勘之重製行為，並且得為科學研究之目的在適當的安全水準下持續保存重製物²⁹。

²⁷ Pamela Samuelson, *Fair Use Defenses in Disruptive Technology Cases*, 71 UCLA L. REV. 1484 (2025).

²⁸ *Bartz v. Anthropic* 案加州北區聯邦地院認為將著作用於訓練生成新文本的 LLM 屬於「轉化性使用」，傾向支持合理使用，但對於下載盜版資料建立 AI 公司內部圖書館副本的使用，則不構成合理使用；2025 年 9 月 Anthropic 同意以 15 億美元和解。*Bartz et al. v. Anthropic PBC* (3:24-cv-05417-WHA)(N.D. Cal. 2025); Cade Metz, *Anthropic Agrees to Pay \$1.5 Billion to Settle Lawsuit With Book Authors*, NY TIMES (Sep. 5, 2025), <https://www.nytimes.com/2025/09/05/technology/anthropic-settlement-copyright-ai.html>。同一法院於 *Kadrey v. Meta* 案即決判決（summary judgment）中，認為 Meta 對原告書籍的使用具高度轉化性，但提出「市場稀釋」（market dilution）之「間接替代」損害考量，認為若原告能提出足夠證據，合理使用第四要素之判斷應交由陪審團決定。*Kadrey et al v. Meta Platforms, Inc.* (3:2023cv03417)(N.D. Cal. 2025).

²⁹ EU CDSMD art. 3.

該條文重要性在於建立不可透過契約排除的強制例外，確保科學研究與文化保存的公共利益³⁰。此外，CDSMD 序言（Recital）第 17 點說明「科學研究目的」資料探勘例外之性質和範圍僅限於進行科學研究的實體，對權利人造成的潛在損害將微乎其微，故會員國不應就該資料探勘例外情況向權利人提供使用補償，最大程度地降低科學研究的成本與法律障礙³¹。同法第 4 條則開放「一般性資料探勘例外」，允許任何人（包含商業實體）就有合法近用權限之著作進行資料探勘之重製行為，若為資料探勘之必要即可以任意之方式保存重製物³²。關鍵在於第 4 條第 3 項的「退出選擇」（opt-out）機制，賦予創作者基本的控制權，即著作權人可透過適當的方式（如機器可讀之 metadata、robots.txt 等）表明不願其作品被探勘，技術開發者必須尊重這種選擇³³。

歐盟 2024 年通過的 AIA 進一步強化創作者的資訊知情權利。AIA 並未增訂著作權規定，然而其序言第 105 條明確說明，CDSMD 的資料探勘例外適用於 AI 管制脈絡，若權利人決定保留其阻止資料探勘之權利，則 AI 模型提供者需獲得權利人授權，才能將其著作用於資料探勘目的³⁴。AIA 第 53 條明確提及 CDSMD 之資料探勘例外規範可能與 AIA 相關，而通用人工智慧模型（general-purpose AI models）提供者須遵守相關文件紀錄與揭露義務³⁵，且須建立相關著作權政策，包含識別權利人並依據 CDSMD 第 4 條提供退出選擇³⁶，為既有資料探勘規範適用於生成式 AI 提供明確的政策基礎。

2025 年 7 月歐盟進一步公布《通用型 AI 模型實踐守則》（General-Purpose AI Code of Practice，下稱實踐守則），分為透明度（Transparency）、著作權

³⁰ EU CDSMD art. 7. See also Thomas Margoni, Martin Kretschmer, *A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology*, 71(8) GRUR INT'L 685, 686 and 695 (2022).

³¹ “In view of the nature and scope of the exception, which is limited to entities carrying out scientific research, any potential harm created to rightholders through this exception would be minimal. Member States should, therefore, not provide for compensation for rightholders as regards uses under the text and data mining exceptions introduced by this Directive.” Recital (17) of EU CDSMD.

³² EU CDSMD art. 4 (1) and (2).

³³ EU CDSMD art. 4 (3).

³⁴ Recital (105) of EU AIA.

³⁵ EU AIA, art. 53(1).

³⁶ EU AIA, art. 53(1)(c).

（Copyright）、安全與資安（Safety and Security）三大面向，為一自願性守則。內容涵蓋撰擬技術文件、揭露訓練模型之資訊等，並提供一份模型文件表格範本，需將所需揭露之資訊記錄於相關文件，包括通用 AI 模型（general-purpose AI, GPAI）提供者皆須遵守的透明性與著作權合規承諾揭露公開模型訓練內容之摘要、制定著作權政策，以及製作模型文件說明表（Model Documentation Form）；另具有系統風險模型（systemic risk）須進行更嚴格之風險評估、通報與安全防護措施等³⁷。歐洲 AI 辦公室亦發布《AIA 下之 GPAI 模型問答集》（General-Purpose AI Models in the AI Act – Questions & Answers），明確指出 GPAI 提供者應編製技術文件、設立內部著作權政策並對外公開訓練資料摘要，以利下游開發者了解模型用途與限制，並促進著作權人之資訊獲取與風險掌控³⁸。值得注意的是，若模型是依據免費開源授權（free and open-source license）發布者，且其參數（包括權重、模型架構資訊和模型使用資訊）已公開，則無需履行前述相關文件義務³⁹。

歐洲模式體現了「創作者參與」的理念，不是簡單地禁止或開放 AI 使用著作資料，而是確保創作者在過程中擁有知情權與選擇權，同時亦鼓勵免費開源相關模式的發展。這種制度既保障了科學探索的自由與 AI 技術發展需求，也重視創作者的主體地位。

三、亞洲經驗：創新發展中的創作者考量

（一）日本：從技術導向到平衡發展的政策演進

日本於 2018 年修正《著作權法》新增第 30 條之 4 條文，創設「非享受使用」之資訊分析例外，任何人只要非為自己或使他人享受著作所表現之思想或感情為目的，而係為資訊分析之必要利用著作即可適用本條而不該當著作侵權；該規範並不限制主體或是否為商業目的，即使涉

³⁷ European Commission, *Code of Practice for General-Purpose AI Models: Transparency Chapter*, 5-6. Available at <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai> (last visited Jul. 26, 2025).

³⁸ European Commission, *General-Purpose AI Models in the AI Act – Questions & Answers*, <https://digital-strategy.ec.europa.eu/en/faqs/general-purpose-ai-models-ai-act-questions-answers> (last visited Jul. 26, 2025).

³⁹ EU AIA art. 53(1)(a), (b) and 53(2).

及營利性的 AI 模型訓練亦屬合法，僅對侵權效果加以限制⁴⁰。是以有研究認為日本政府對於其 AI 監管採取極為寬鬆之政策，相當有利於 AI 之發展，並稱之為「機器學習天堂」⁴¹。

然面對生成式 AI 的快速發展與創作者反彈⁴²，日本政府後續亦發表數篇報告，諸如文化廳發布之《關於 AI 與著作權之思考》⁴³、內閣府智慧財產戰略推進事務局發布之《AI 時代智慧財產權研究小組中期報告》⁴⁴、總務省與經濟產業省發布之《AI 事業指引 1.0 版》⁴⁵等。《關於 AI 與著作權之思考》中即指出若使用盜版或在未付費之情形使用付費資料進行訓練將有高機率侵害著作權，並認為應就 robots.txt 是否屬於著作權法之「技術保護手段」或「技術使用限制手段」進行檢討，以因應未來科技發展趨勢⁴⁶，顯示日本政府認為並非所有機器學習中對著作權作品的使用皆落入例外範圍，並透過使用案例進行軟性限制。本文認為係對原本 2018 年著作權修法所倡導的創新導向政策有所收斂。儘管該些案例並不具備強制性，惟日本著作權法並未如同歐盟 CDSMD 禁止透過契約排除例外條款之適用⁴⁷，是以該報告之內容尚有一定之指導價值。同時，

⁴⁰ 陳皓芸，巨量資料分析與著作權法——以日本 2018 年著作權法修正為中心，萬國法律 229 期，頁 11-21，2020 年 2 月；高嘉鴻，資料探勘著作權限制及例外規定國際立法趨勢研究，智慧財產權月刊 298 期，頁 49-50，2023 年 10 月。

⁴¹ 上野達弘，情報解析と著作 —「機械学習パラダイス」としての日本，人工知能 36 卷 6 号，頁 747-748，2021 年 11 月。

⁴² Brian Quarmby, *Japanese AI Experts Raise Concern Over Bots Trained on Copyrighted Material*, COINTELEGRAPH (Jun. 1, 2023), <https://cointelegraph.com/news/japan-council-concern-ai-trained-copyright-material> (last visited Dec. 28, 2025).; *Japan Arts Workers' Survey Reveals Harm From Generative AI, Calls for Copyright Action*, MAINICHI JAPAN (May 16, 2023), <https://mainichi.jp/english/articles/20230516/p2a/00m/0et/014000c> (last visited Dec. 28, 2025).

⁴³ 文化審議会著作権分科会法制度小委員会，AI と著作権に関する考え方について，2024 年 3 月。

⁴⁴ 日本内閣府 AI 時代の知的財産権検討会「AI 時代の知的財産権検討会中間とりまとめ」，頁 44（2024 年 5 月）（下稱「AI 時代智慧財產權研究小組中期報告」）。

⁴⁵ 經濟產業省と総務省，AI 事業者ガイドライン第 1.0 版，2024 年 4 月。

⁴⁶ 文化審議会著作権分科会法制度小委員会，同註 43，頁 19-29。

⁴⁷ Global Partnership on Artificial Intelligence, CO-GENERATION OF DATA COPYRIGHT AND DATA PROTECTION RIGHTS IN CO-GENERATED INPUT AND OUTPUT OF GENERATIVE AI, 32-33 (2024).

相關報告亦開始強調 AI 開發者的責任，除了要求避免使用明顯侵權的資料，並建議建立創作者合作機制⁴⁸。

日本經驗顯示，純粹的技術導向政策可能引發創作者反彈，需要適時調整以平衡各方利益。2023 年漫畫家出身的國會議員赤松健亦提出制定生成式 AI 使用指引等相關建言，反映創作者社群的政策關切，擔心 AI 對內容創作產業將產生負面影響，要求政府提供更明確的措施⁴⁹。日本政府近期推動的創作者與相關利害關係人溝通合作機制⁵⁰，亦可看到其試圖透過自願性報酬補償機制建立雙贏關係⁵¹。這種政策調整體現了從「單純技術導向」到「技術發展與創作者權益並重」的重要轉向。

（二）南韓：平衡發展的制度探索

目前南韓著作權法仍未納入明確的資料探勘例外，但實務界與學界高度關注此議題。2025 年 3 月南韓文化體育觀光部（문화체육관광부 Ministry of Culture, Sports and Tourism）對外發表「文化南韓 2035」計畫，點出南韓政府因應經濟社會危機與 AI 時代挑戰，提出六大文化發展核心課題，其一即為推動「AI 時代文化制度重建」，並明確指出應修訂現行著作權制度以應對 AI 生成與利用問題，並另訂一與人格權商業使用相關法案，以保護個人形象、姓名、聲音等權利⁵²，可見南韓政府的積極態度。

⁴⁸ 日本內閣府及文化廳於 2024 年 5 月《AI 時代智慧財產權研究小組中期報告》報告中所蒐集之意見顯示部分論者主張應建立創作者合作機制，建構合法資料庫供 AI 使用，以避免侵權問題，詳 AI 時代智慧財產權研究小組中期報告，同註 44，頁 44、47。

⁴⁹ 參見赤松健議員網站，<https://kenakamatsu.jp/>；新清士，繪描きの気持ちを考慮に入れることが重要：赤松健氏「画像生成 AI、珍しく日本が勝つチャンス」，ASCII.JP，<https://ascii.jp/elem/000/004/122/4122855/3/>（最後瀏覽日：2025/09/05）。

⁵⁰ 日本內閣府於《AI 時代智慧財產權研究小組中期報告》中指出，雖然公眾意見徵詢結果顯示對補償創作者措施的意見不一，難以提供理論基礎來強制實施統一的補償制度，但從實踐內容創作良性循環的角度來看，引介民間可實施的自願補償措施的案例，並推動其努力將「具有重要意義」，詳 AI 時代智慧財產權研究小組中期報告，同註 44，頁 44-50。

⁵¹ 日本文化廳於 2024 年 4 月成立「AI 與著作權相關者網絡」（AI と著作權に関する関係者ネットワーク），旨在於建立一個讓 AI 開發商、內容創作者、權利人團體等所有利害關係人能進行資訊共享與意見交換之平台。日本文化廳於 2025 年 5 月 30 日公布該網絡階段性總結報告，整理數次會議成果與待解決課題，其中第 6 次即以報酬回饋的措施與案例的意見交換。詳日本文化廳「AI と著作權に関する関係者ネットワークの総括」頁 10（2025 年 5 月 30 日），https://www.bunka.go.jp/seisaku/chosakuken/pdf/94221801_01.pdf（最後瀏覽日：2025/09/02）。

⁵² 대한민국 문화정책 대전환 '문화한국 2035' 발표, 6-7, available at https://www.mcst.go.kr/site/s_notice/press/pressView.jsp?pSeq=21690 (last visited Dec. 28, 2025).

南韓著作權委員會邀集學界、法律專家、技術專家及利害關係人組成「AI 著作權制度改善工作小組」，於 2024 年 1 月發布「生成式 AI 著作權指引」⁵³，於同年 6 月新聞稿提及政府或學術機構可在創作者與 AI 開發商之間的談判授權過程擔任中介角色，並可考慮增訂資料探勘權利限制條款與相關補償措施，以促進對著作權保護內容之利用⁵⁴。

而南韓國會議員 Do Jong-hwan 於 2021 年即提出著作權法資料探勘例外條款之新增提案，其中提案新增「為資訊分析目的的重製和傳輸」之例外規範，雖該案提出之時尚未考量生成式 AI，然受日本及歐盟法制影響，該提案認為資料探勘行為在具備以下三項實質要件下將屬合法，包含：（一）使用者對著作物須有合法近用權限（lawful access）；（二）資料探勘須屬於非享受性使用（non-enjoyment），不得用於感官欣賞或個人娛樂目的；（三）僅得於必要範圍內使用（necessary extent）著作物⁵⁵，惟目前未有明確後續發展消息。

若資料探勘例外修法草案未能於南韓國會通過並正式立法，有學者亦指出另一替代法律途徑－援用南韓著作權法第 35 條之 5 第 2 項之合理使用條款；南韓之合理使用制度係於 2011 年韓美自由貿易協定（US-Korea Free Trade Agreement, KORUS FTA）談判期間引進，採用與美國相同之四要素測試，然而有論者亦指出南韓對合理使用的應用尚欠缺具體案例作為判斷基礎⁵⁶。

（三）新加坡：雙軌制的技術中性探索

新加坡於 2021 年著作權法修法時採行「合理使用」（fair use），正式承認其概括性開放性質，往美國法制靠攏，並新增第 244 條運算資

⁵³ 「生成式 AI 著作權指引（생성형 AI 저작권안내서）」全文，<https://www.copyright.or.kr/information-materials/publication/research-report/view.do?brdctsnno=52591#> (last visited Sep. 1, 2025).

⁵⁴ 探索人工智慧時代著作權制度改革的線索，南韓文化觀光體育部 2024 年 6 月 28 日新聞稿，https://www.mcst.go.kr/kor/s_notice/press/pressView.jsp?pSeq=21201（最後瀏覽日：2025/09/08）。

⁵⁵ Kyungsuk Kim, *Korean Copyright Issues in Text Data Mining for Generative AI*, 1(1) J. AI L. & REGUL. 64, 70-72 (2024).

⁵⁶ *Id.*, at 72-73.

料分析（computational data analysis）例外，允許利用人得在進行電腦資料分析以及為進行電腦資料分析準備著作之目的範圍內，重製及傳輸著作⁵⁷。

該法第 244 條係規定若某一著作之重製行為係以電腦資料分析為目的，或係為準備日後進行電腦資料分析而製作之著作或錄音物，且該重製物未被用於任何其他用途，亦未向他人提供，除非該行為係為驗證該分析結果或與分析目的相關之合作研究或學習之必要情形，則該行為可視為符合運算資料分析例外之要件；而該重製行為之行為人必須已合法取得用以製作該重製物之原始資料，且該原始重製物須為合法取得或行為人無重大過失而不知之侵權重製物，另若該侵權資料為實現特定研究目的所必要者且並未另作他途使用則亦屬可容許之範圍，惟若著作之重製來源為明顯侵權之盜版網站，行為人應負無過失責任⁵⁸。運算資料分析例外特別強調「合法取得」要件，對使用者侵權資料的行為課以嚴格責任，體現對創作者權益的重視。新加坡政府尚無官方公開宣導該法第 244 條例外規定適用於生成式 AI 訓練，然而已有學者指出其限制。基於機器學習目的，從網路上爬取資料通常會繞過付費牆或違反使用條款，恐不符合該條規定的「合法存取」要件，同時許多生成式 AI 模型在生成應用上亦可能涉及複製著作轉換為機器可讀格式，或涉及資料存儲，其唯一目的並非單純止於分析資料功能，亦難符合第 244 條例外規範⁵⁹。

四、我國的政策選擇：「友善創作者之負責任 AI 發展」的平衡智慧

透過各國經驗，我國可汲取幾項重要啟示。首先，制度確定性的重要。創作者與技術開發者需要明確可預期的保護與風險評估機制，而非依賴個案判斷的不確定救濟可能。其次，參與權的核心地位。數位經濟時代的創作者保護不只求事後救濟，更重要的是過程參與，包括知情權、選擇權與合理補償。第三，平衡機

⁵⁷ 江思穎，新加坡 2021 年著作權法修正評析，智慧財產權月刊 286 期，頁 31-32，2022 年 10 月。

⁵⁸ Singapore Copyright Act 2021, No.22 of 2021 § 244.

⁵⁹ David Tan, *Generative AI and Copyright Fair Use*, NATIONAL UNIVERSITY OF SINGAPORE (Mar. 2024), <https://law.nus.edu.sg/trail/generative-ai-copyright-fair-use/> (last visited Dec. 28, 2025).

制的細緻設計。好的制度並非創作者與技術開發者之間的零和選擇，而是透過細緻的機制設計實現雙贏與公共利益。最後，文化脈絡的重要性。我國需要在國際經驗基礎上，發展適合本土社會、文化與產業脈絡的制度模式。

有鑒於我國現行著作權法「授權」及「強制授權」需個案一一授權與專責機關許可，第 65 條「合理使用」亦須法院個案判斷且不確定性極高，導致現有方式皆因行政與交易成本過高，恐無法解決 AI 訓練著作利用爭議與技術、文化及產業發展之急迫性。各國包括歐盟、日本、新加坡等已針對資料探勘技術提供不同程度的容許條件，且各國政府及相關論述亦從資料探勘技術的演進，於生成式 AI 生命週期、訓練與利用等不同階段可能產生之法規適用爭議，進行解釋、立法或以軟法指引，平衡智財權保護與數位創新，並增加國際競爭優勢，值得我國借鏡。基於上述啟示，本文建議我國宜採取「友善創作者之負責任 AI 發展」的平衡路徑，於著作權法之著作權限制章節增訂著作財產權限制之「法定例外」，將「資料探勘」增列於第 44 條至第 63 條章節所列舉限制著作財產權之例外豁免情形。同時，擬定「合理補償與退出機制」及資料探勘利用人「透明性免責」之配套，以提供誘因與衡平機制。以「友善創作者之負責任 AI 發展」模式，促進技術創新也能保護創作的健康生態，符合臺灣多元文化與永續發展的價值理念。

肆、以文化永續為核心的資料探勘例外制度設計

面對現行著作權法制的結構性問題，我國應建立一套以文化永續為核心的制度框架。本文建議於著作權法增訂著作財產權限制之「法定例外」（豁免侵權責任）條款，將「資料探勘」增列於第 44 條至第 63 條章節所列舉限制著作財產權之例外情形（下稱「資料探勘例外」）。資料探勘作為著作權之「豁免例外」，一旦利用行為符合法律規定之限制範圍，當然不構成著作權之侵害。另此方案應搭配「三大支柱」（透明性、合理報酬與退出選擇、著作權過濾）的平衡機制，保障創作者權益的同時，促進 AI 的負責任發展與應用，促進雙方合作雙贏。「三大支柱」並非獨立的制度安排，而是「法定例外」的構成要件，任何一項要件的缺失，都將使利用行為無法獲得例外保護。這種設計確保豁免例外並非無條件開放，而是以審慎、合理、平衡的配套為前提，確保創作者權益得到實質保障。

一、「豁免例外」：友善創作者之內涵

資料探勘例外的建立，並非為了便利 AI 開發，而是在深刻理解技術本質的基礎之上，為創作者與 AI 開發者建立更公平、更永續的合作關係。如本文第貳章所述，資料探勘在技術本質上與傳統的著作利用存在根本上的差異：AI 訓練進行資料探勘時，並非為了「閱讀」或「欣賞」作品內容，而是透過數學運算分析文本中符號關係的統計模式、語言結構或視覺特徵。這種「非享受性利用」與傳統著作權法所防範的盜版、抄襲行為存在本質上的區別。

傳統個案授權模式的市場機制在面對 AI 大規模資料需求時，已明顯失靈。這種失靈不僅體現在交易成本的急遽升高，各種爭議性使用行為大量出現且難以遏止，更可能導致只有資源雄厚的大型科技公司與大型內容集團才能負擔巨額授權費用與談判，進而形成內容市場趨於同質化與單一化。在這種情況下，個別創作者往往成為最大的受害者——既無法獲得合理的收益，也面臨被邊緣化風險。因此，建立資料探勘例外，實際上是為了打破這種傾斜的市場結構與失靈風險，讓創作者能夠在更平等的基礎上，參與 AI 時代的價值分配。

再者，資料探勘的豁免例外絕非無條件開放，而是以嚴謹配套要件為前提。我國現行著作權法之侵權尚有刑事責任⁶⁰，此豁免例外之要件與配套設計始終以保護人類創意自主性、促進文化與創意之永續為核心。「目的限制」要件將例外範圍限制在技術性的資料分析用途，排除直接的著作商業利用或內容享受的行為。「透明性」要件賦予創作者合理的知情權，確保其能了解自己作品的使用情況。「合理報酬與退出選擇」要件則在特定情境下如商業或營利性使用目的，透過「法定授權」以經濟性機制體現對創作者貢獻的認可與尊重，同時以退出選擇確保創作者保有對其作品被利用的控制權主體性。「著作權過濾」要件，則用以阻擋 AI 系統或服務使用者意圖生成侵害著作之內容，防止著作權侵害的擴散。

⁶⁰ 參見著作權法第 91 至 95 條。然依 TRIPS 第 61 條前段，就侵害著作權訂立刑事程序與罰則，主要是應對「蓄意而具有商業規模之行為（willful copyright piracy on a commercial scale）」。
關於我國著作權法是否應去刑化，囿於篇幅，暫不在本文討論範圍。

二、支柱一「透明性」：建立創作者與 AI 開發者的信任橋梁

在「資料探勘例外」中，「透明性」要求為制度設計的第一根支柱，其核心價值在於保障創作者的知情權，建立創作者與 AI 開發者之間的信任關係。數位時代中，資訊就是權力，創作者有權知道他的作品如何被利用，這是創作者自主權的基石。透明性要求不僅是技術規範，更是建立公平合作關係的制度基礎。

透明性要求的程度，可參考歐盟 AIA 相關透明性揭露義務之設計，建立適合本土環境的透明性機制⁶¹。AI 開發者應適當的揭露訓練資料摘要，並提供足夠詳細的資料來源說明，使創作者與社會大眾了解 AI 系統的資料基礎。此透明性要求不必然對公眾全面開放，但須以適當方式向創作者揭露，可能包含資料的性質、來源、預訓練處理方式等實質性資訊，讓創作者理解其作品在 AI 訓練中的角色與影響。

特定目的如商業或營利目的之資料探勘，除了一般性揭露義務之外，可建立「定期使用通報」要求，強化創作者的參與。我國可進一步討論建立「專責機構」之通知機制，由 AI 開發者／資料探勘利用人定期向專責機構通報資料使用情況，再由專責機構通知相關創作者，降低開發者通知成本，也能保障創作者獲得資訊。這種機制設計體現制度的人性化考量，避免讓創作者在負責任卻複雜的技術細節中迷失，而透過專責機構的中介服務，讓資訊共享與傳遞更加有效率。另「來源可溯性標準」的建立，則為透明性提供技術支撐。可透過數位浮水印、區塊鏈紀錄等技術手段，提升資料使用的可追蹤性，讓創作者能夠追溯作品的使用軌跡，不僅有助於權利的行使，也為未來可能的紛爭解決提供技術基礎。

考量不同目的與規模 AI 開發者的負擔能力，透明性要求或可採差異化設計。例如大型商業模型應承擔較重的透明義務，提供較詳細的資料來源說明與使用統計；中小型開發者則可採相對簡化揭露標準，避免過重法遵負擔阻礙技術創新；非營利與研究機構可採學術倫理規範為主，減少不必要的行政負擔，讓研究能量能專注於學術創新而非行政程序。

⁶¹ 歐盟 AIA 關於透明性考量，特別是第 53(1)(d) 條要求通用目的 AI (GPAI) 模型供應商應揭露訓練資料摘要，參見 EU AIA, Art. 53(1)(d).

值得特別關注的是，我國政策亦宜參考歐盟等國家與區域鼓勵創新與開源之趨勢⁶²，增強透明性原則之配套設計，接軌國際。我國可考慮對高透明度的開源 AI 專案提供鼓勵機制。開源精神與透明性要求在理念上高度契合，可考慮政策性支持，如簡化報酬義務、技術支援與認證標章、優先獲得政府專案支持等方式，鼓勵 AI 開發朝向更開放共享與負責任之方向發展。

三、支柱二「合理報酬與退出選擇」：創作者主體性的雙重保障

「合理報酬」與「退出選擇」機制為制度設計的第二個支柱，兩者共同體現對創作者主體性的雙重保障。退出選擇機制可彌補創作者在 AI 利用巨量資料訓練時無法直接明確表達授權意願的市場失靈缺陷，為創作者提供一定程度的事後控制可能性；而報酬補償機制則透過經濟補償，肯認人類創意勞動的價值，確保創作者能夠共享 AI 發展帶來的經濟成果。這種雙重保障機制既是對傳統著作權授權模式不效益的補強，更是 AI 時代創作者權益保障的創新路徑。

（一）「合理報酬」機制

報酬補償機制的設立，是透過經濟補償肯認人類創意勞動的價值，並為創作者提供持續創作的激勵。在商業或營利使用目的之生成式 AI 脈絡下，報酬補償奠基於三個基本定位：首先是體現對人類創意投入的尊重，認可創作者在文化生態中的基石地位；其次是考慮 AI 技術對創意勞動市場帶來的破壞風險，為創作者提供合理補償；最後是促進創作者與 AI 技術之間的良性互動，而非對立競爭。

在技術變革脈絡下，報酬制度的設計，不僅關乎經濟公平，亦是創意生態永續發展的基礎。因此，AI 時代的報酬補償機制應基於著作權法促進文學、科學、藝術等學術性質之知識傳播與文化發展的法理基礎與立法理由⁶³，超越單一市場交易模式，發展更具包容性與前瞻性的價值分

⁶² 參照 EU AIA, art. 2(8), (10) and art. 53(1)；蔡祈岩，台灣能從晶片島邁向 AI 強國？法國「開源文化」的六大啟示，遠見雜誌，<https://www.gvm.com.tw/article/121351>（最後瀏覽日：2025/05/18）。

⁶³ 參照我國著作權法第 1 條立法理由。

配方式。可針對商業化或營利性 AI 模型，設立強制性之法定補償報酬機制，同時為非營利、教育性或實驗性使用提供更多例外。

此外，報酬補償除了 AI 訓練輸入階段的補償，亦可結合生成式 AI 輸出使用頻率作為考量，意即對於被頻繁學習與輸出的高品質著作內容提供額外報酬。透過這種設計，激勵創作者生產高品質、高創意性內容，也引導 AI 系統向高品質內容學習，形成良性循環。報酬補償機制的適用範圍應以商業或營利目的資料探勘行為為主，體現權利義務的對等原則。費率標準的制定應兼顧合理性與可操作性，並建立動態調整機制，因應市場變化與技術發展適時修正。

另一方面，考量識別與分配大量著作權人之成本，可能最終導致實質分配效率及效益過低的問題，本文建議可考慮發展「公共基金」模式作為補充選項。在「公共基金」模式下，AI 開發者支付之補償並不分配至特定著作權人，而是透過法定公共基金用途，用於支持集體創意環境的發展；例如，公共基金可用於重點支持創作者適應 AI 時代之技能培訓與轉型、受 AI 衝擊較大之創意領域（如視覺藝術、文字作品）永續發展、促進文化多樣性保護特別是弱勢語言文化的數位化與保存、資助開發有利於創作者的 AI 工具……等，亦可用於協助弱勢創作者支持系統，包括法律諮詢、技術協助等服務，確保合理報酬補償與退出選擇不因資源不均或權力關係而淪為形式化。

（二）「退出選擇」機制

退出選擇，係合理報酬補償機制的必要補充，也是創作者主體性的重要保障。退出機制可分為「事前保留」與「事後退出」兩種模式：「事前保留」允許創作者藉由標準化方式（如機器可讀之 robots.txt）明確表達不同意 AI 訓練使用；「事後退出」則針對事前未能明示保留且其著作已為訓練資料、進入參數之創作者，於該模型後續之「再利用」或擴充版本，得採事後退出，停止其著作被「再利用」。「事後退出」機制目前雖尚無國際立法例，本提案乃基於對技術發展潛力的期待，以利未來創新之技術發展保留更為彈性與平衡的空間。

流程設計上，退出請求宜簡便易行，可透過統一平台或專責機構提交。同時，為防止濫用亦可設置合理的退出理由要求，如作品被用於生成侵犯著作權、人格權或其他合法權益的內容。退出效力應明確為「向將來生效」，即停止在未來訓練中使用該作品，而不要求刪除已訓練模型中的參數。這種設計既尊重創作者的控制權，也考慮了技術可行性的限制。

然而，退出機制雖強化創作者的控制權，也可能帶來言論自由與市場碎片化的風險。若大量創作者選擇退出，可能導致 AI 訓練資料庫的不完整，進而影響 AI 系統功能與社會價值。因此，退出機制需與公共利益保持平衡，可透過提供足夠具吸引力的補償，使多數創作者參與而非退出，或差異化退出門檻等。這種平衡設計旨在保護真正需要控制權的創作者利益，同時避免退出機制被過度使用，導致 AI 技術發展過度受阻，損害公共利益。

四、支柱三「著作權過濾」機制：責任分攤的協調設計

「著作權過濾」機制與責任分攤設計構成此制度方案的第三個支柱。資料探勘例外的要件之一，係應提供「合理之著作權過濾機制」，阻擋 AI 系統或服務「使用者」意圖生成侵害著作之內容，防止著作權侵害行為的擴散。透過技術手段與法律責任的有機結合，防範著作權侵害也為創作者提供有效的救濟途徑，為預防性與救濟性並重的設計。

此要件主要在於針對 AI 系統的輸出內容進行監控與管理。法律責任分配上，可參照現行著作權法第 87 條，增訂 AI 系統或服務「使用者」濫用服務負直接侵權責任條文，而 AI 系統或服務「提供者」若違反合理著作權過濾注意義務者，則課以間接責任。此外，針對我國著作權法刑事責任部分，本文主張資料探勘例外條款之適用已屬著作權限制性規定，無侵權問題，符合現行著作權法第 91 條第 4 項「僅供個人參考或合理使用者，不構成侵害」⁶⁴；惟針對明知生成內容侵

⁶⁴ 2022 年因應 CPTPP 已修正該條文內容為第 91 條第 3 項，然而該條文雖經立法院三讀通過，施行日期仍有待行政院定之，故此處仍以原條文第 91 條第 4 項表示。

害特定著作之著作財產權，而仍為自己或公眾享受該生成物中思想或感情而為生成者，則可考量搭配前述使用者濫用之民事責任增訂罰則。若 AI 開發者未符資料探勘例外條款之透明性及補償條件者，則自始無法主張豁免著作利用行為相關責任。

另針對訓練資料為明確侵權內容之情形，例如盜版網站等，本文建議三種方案，以供未來進一步討論之基礎：（1）方案 A：依本文提案「資料探勘」之定義，資料探勘行為係屬著作權限制性規定，並無侵權問題。此方案符合「資料探勘」之技術本質與本文提案設計之邏輯。（2）方案 B：考量本文提案對利害關係人（特別是創作者）之立法溝通效率，為增進利害關係人接受度，此方案設計為：明知其資料探勘之內容為侵害著作財產權之著作，而受有利益者，視為侵害著作權。（3）方案 C：為方案 A 與 B 之折衷方案，資料探勘利用人，若以資料探勘為目的，重製或公開傳輸侵害著作財產權之著作，而受有利益者，排除本條適用；但利用人已盡相當之注意者，不在此限。

因應透明性、合理補償與退出選擇、及著作權過濾機制等技術保障措施的設計，本文建議遵循技術中立性原則，避免過度限定技術路徑從而阻礙技術創新。例如著作權過濾技術標準應著重於結果要求，而非具體實現方式，給予 AI 開發者技術創新空間，以符合相關規範。同時，技術措施也應保持一定的開放性，允許在實踐中不斷調整和改善，適應技術的快速更迭。這種平衡反映在法律設計上，可以採用「合理技術措施」的表述，而非硬性規範具體技術標準。同時，亦建議可搭配產業自律、技術標準組織和監管指引的組合，保持法律穩定性的同時，實現技術措施的動態更新。為支持並達成此動態平衡，可建立技術措施的定期評估機制，透過多方參與方式，持續評估技術措施的效果與影響，動態調整方案。

伍、結語：文化創意生態永續發展的新路徑

長期以來，技術創新與既有權利保護往往被視為對立兩端，彷彿保護創作者權益必然阻礙技術發展，促進技術創新必然犧牲創作者利益。然而，本文的分析表明此對立關係並非必然，透過適當制度安排，超越零和思維，保障創作者權益

的同時，推動 AI 負責任發展與應用，共同促進公共利益。本文所提出的創作者友善之「資料探勘例外」制度，含「法定例外」（豁免侵權責任）及「三大支柱」配套措施，正是在深刻理解技術本質與創作價值的基礎上，為我國文化與創意永續發展探索的新路徑。

本文主張以文化創意永續生態視角發展平衡創作者、AI 開發者與公共利益之制度修正，於著作權法第 44 條至第 63 條章節所列舉限制著作財產權之例外情形增訂「資料探勘例外」與「三大支柱」之機制設計。此豁免例外並非無條件的開放，而是以審慎、合理、平衡的配套為前提與要件，確保創作者權益得到保障。支柱一「透明性」要求確保創作者權益和社會信任基礎，透過適當揭露資料、通知與追溯來源等機制，實現創作者、AI 開發者與管制者間的資訊對等關係，並作為合理報酬與退出機制的實踐基礎。支柱二「合理報酬」體現對人類心智勞動的尊重與補償正義原則，係維持創意生態永續性的重要保障；「退出選擇」作為補償機制之補充，保障創作者自主權，避免資源權力的偏斜與壟斷，進而阻礙知識與資訊的傳播。支柱三「著作權過濾」等技術措施，則將價值理念轉化為技術實踐，以合理的過濾機制作為「豁免例外」之要件，也是防止侵權擴散的主動措施。

作為正體中文與亞太多元文化的重要基地，臺灣在 AI 訓練資料面臨嚴重稀缺困境。若缺乏適當法規制度引導，我國文化創作者與多元創意可能在 AI 時代被嚴重邊緣化。透過建立創作者友善的資料探勘例外制度，不僅能保護現有創作者權益，更能夠激勵更多優質多元語料內容的創作，為臺灣在 AI 時代的多元文化傳承與發展奠定堅實基礎。

「資料探勘例外」與「三大支柱」之法制設計也體現負責任 AI 發展的基本要求。真正的技術進步不應建立於犧牲人類尊嚴與權益之上，而應透過技術與人文的有機結合，實現人類福祉的整體提升。透過建立完善的透明、補償、退出與過濾機制，我國的著作權法制能夠引導 AI 研發及產業朝向更加負責任的方向發展，確保技術創新服務於人類的整體利益。更進一步而言，在全球各國皆面臨 AI 著作權爭議困境時，如何在保護創作者權益與促進技術發展之間找到平衡點，是一個具有普遍意義的重要課題。臺灣若能前瞻且務實地面對挑戰，並形成在地的法制因應模式，則有機會在國際 AI 治理地景中貢獻在地智慧，發揮更積極的角色。