國立臺灣大學法律學院法律學系

碩士論文

Department of Law

College of Law

National Taiwan University

Master's Thesis

人工智慧技術與個人資料刪除權——以金融應用為中心

Artificial Intelligence Technologies and the Right to Erasure of Personal Data: Focusing on Financial Applications

朱家伶

Chu, Chia-Ling

指導教授：楊岳平 博士

Advisor: Yueh-Ping Yang, S.J.D.

中華民國 115 年 2 月

February 2026

doi:10.6342/NTU202600641

謝辭

從論文的構思到提筆寫下謝辭，彷彿是被時間推著往前走；回過神來，竟也兩年有餘。這段路看似漫長，卻也因為一路有人相伴、有人指引，而變得格外踏實。

這本碩士論文能夠完成，最感謝的莫過於我的恩師楊岳平教授。之所以對人工智慧相關議題產生興趣，正是因為楊老師在我碩二時開設課程，系統性地帶領我們理解人工智慧所牽動的法律問題，讓我第一次清楚看見新興科技與法律交錯之處的張力與價值，也因此萌生深入研究的念頭。論文撰寫初期，老師即提供我許多方向性的建議，使我得以在人工智慧法與個人資料保護法這個仍在快速發展的領域中，較快擇定研究問題與路徑；在後續撰寫過程中，老師也不吝指導、反覆提醒與修正，讓我能在迷惘與徘徊之際，仍一次次回到論文應有的重心，最終完成本篇研究。在此謹致上最深的敬意與感謝。

也誠摯感謝兩位口試委員——林勤富教授與何之行教授，在口試時提供我諸多精闢且具體的建議，讓我得以更清楚看見論述可再深化之處，以及文字、架構與問題意識仍能精煉的方向。能有榮幸接受兩位學識豐富的教授指導，是我在研究所階段十分珍貴的收穫。

這三年的碩士時光對我而言意義非凡。走過了高壓的國考，也逐漸從考卷堆中抽身，得以自由地去探索更深刻的法律問題。在這段時間裡，我沈澱、休息、再出發，一路上也遇見了許多寶貴的人，並留下難以取代的回憶。想謝謝英綸，一直以來都是我習法之路的重要支持；謝謝哲祐，總是讓我知道你一直都在。也謝謝研究所的好夥伴們：郁淇、慕薇、語彤、Yvonne、瑀軒、Mike、承翰、金、康、Jessy、大祐、大崴學長，一起修課、吃飯聊天的時光，是我碩士生活中最重要的調劑與陪伴。也謝謝土研 512 的學弟妹們，讓我在寫論文期間得以常常去打擾你們；有你們在的日常，讓那段看似漫長的寫作時光不再那麼孤單。

doi:10.6342/NTU202600641

謝謝哞哞牛，在我學習法律的這七年始終陪伴在我身邊，成為我最重要的支柱。你的存在，讓我能夠有勇氣去面對重重挑戰，也讓我更有動力去成為更好的人。最後，想謝謝朱瑤與我的父母。謝謝你們一直以來的支持與包容，謝謝你們總是尊重我的選擇，並在我需要的時候，成為我最溫暖、最可靠的避風港。

這本論文的完成，也意味著人生下一階段的開始。期許自己能帶著過去累積的努力與養成的韌性，在未來面對新的挑戰時，仍能保持清醒、保持好奇，也保持對法律與世界的真誠與熱愛。願我不忘此刻的感謝，亦不辜負一路走來的自己。

2026 年 2 月撰於美樹咖啡館

doi:10.6342/NTU202600641

中文摘要

在當今網路科技高度發展的時代，資訊傳播快速且難以遏止，使「遺忘」不再輕而易舉，也使社會對個人資料保護的關注日益升高。然而，隨著人工智慧（AI）技術的興起，個人資料刪除權首當其衝地面臨嚴峻挑戰。尤其在金融領域，AI 已廣泛應用於信用評分、詐欺偵測與風險管理等高敏感性決策場景，其模型效能高度依賴長期且完整的個人資料作為基礎。惟深度學習模型本質上係透過將訓練資料隱含於模型權重中，具有高度黑箱性與不可逆性，即使表面上刪除了原始資料，模型參數中仍可能殘留具可識別性之資訊，凸顯 AI 應用與刪除權保障之間潛藏的制度張力。

本文以法律經濟學為分析基礎，結合古典經濟學強調的契約自由與行為經濟學揭示的決策偏誤，檢視現行《個人資料保護法》在刪除權設計上的侷限，並比較歐盟《一般資料保護規則》（GDPR）與美國《加州消費者隱私法案》（CPRA）的規範策略。GDPR 採取高強度的不可讓渡路徑，刪除權不得透過契約排除或弱化；CPRA 則在同樣確保刪除權為基本權的前提下，允許企業以優惠折扣等方式交換資料，但必須符合明確揭露、明示同意與隨時撤回之條件，以市場誘因確保資料供給穩定。對照之下，我國雖在實務上允許透過契約約定保存期限以維持一定彈性，但欠缺明確的對價規範與隨時撤回等可逆性設計，致使同意容易流於形式，並加劇資訊不對稱的風險。

基於比較法觀察與法律經濟分析結果，本文建議未來修法應明文化資料利用的「有償契約自由」並結合行為經濟學導向的「輕推」措施，特別是資訊揭露與隨時撤回，以引導資料主體做出理性與自主的選擇，進而在隱私保障與 AI 發展之間建立可持續的平衡。

關鍵詞：人工智慧、個人資料刪除權、資訊自決權、法律經濟分析、行為經濟學、GDPR、CPRA

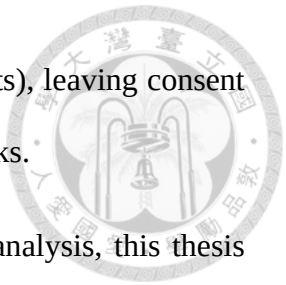
doi:10.6342/NTU202600641

Abstract

In today's era of highly developed digital technologies, information spreads rapidly and is difficult to restrain, making "forgetting" no longer effortless and heightening society's concern for personal data protection. Yet with the rise of artificial intelligence (AI), the right to erasure faces particularly acute challenges. In the financial sector, where AI is widely applied in sensitive decision-making contexts such as credit scoring, fraud detection, and risk management, model performance depends heavily on long-term, stable, and comprehensive datasets. However, deep learning models embed training data implicitly into their parameters, carrying an inherently opaque and irreversible "black-box" nature. Even if the original data is ostensibly deleted, identifiable traces may remain within model weights. Coupled with the prohibitive costs of retraining, this reality greatly undermines the effectiveness of exercising the right to erasure, underscoring the structural tension between AI applications and data protection guarantees.

This thesis employs a law-and-economics perspective, integrating the classical economic emphasis on contractual freedom with behavioral economics' insights on decision-making biases, to examine the limitations of Taiwan's Personal Data Protection Act (PDPA) in its design of the right to erasure. It further conducts a comparative analysis of the regulatory strategies under the EU General Data Protection Regulation (GDPR) and the California Privacy Rights Act (CPRA). While the GDPR adopts a strict inalienability approach—prohibiting any contractual exclusion or dilution of the right to erasure—the CPRA, while still treating the right as a fundamental entitlement, permits companies to exchange data for financial incentives such as discounts, provided that such arrangements meet conditions of explicit disclosure, express consent, and the possibility of withdrawal at any time. By contrast, Taiwan's regime allows contractual stipulation of retention periods in practice, thereby maintaining some flexibility, yet it lacks explicit

doi:10.6342/NTU202600641



provisions on consideration or reversibility (such as withdrawal rights), leaving consent vulnerable to formalism and exacerbating information asymmetry risks.

Based on these comparative findings and law-and-economics analysis, this thesis argues that future reforms in Taiwan should explicitly incorporate a framework of "remunerated contractual freedom" while complementing it with behavioral economics-inspired "nudge" measures, particularly robust disclosure obligations and a right of withdrawal. Such a framework would guide individuals toward making informed and autonomous data-sharing decisions, thereby establishing a sustainable balance between privacy protection and AI-driven innovation.

KEYWORDS: Artificial Intelligence, Right to Erasure, Informational Self-Determination, Law and Economics, Behavioral Economics, GDPR, CPRA

doi:10.6342/NTU202600641

目次

doi:10.6342/NTU202600641

doi:10.6342/NTU202600641

doi:10.6342/NTU202600641

doi:10.6342/NTU202600641

doi:10.6342/NTU202600641

doi:10.6342/NTU202600641

圖次

doi:10.6342/NTU202600641

表次

doi:10.6342/NTU202600641

第一章 緒論

第一節　研究動機

隨著人工智慧（Artificial Intelligence, AI）技術迅速擴張至金融、醫療與行銷等多元領域，透過大量個人資料進行模型訓練與決策支援已成為 AI 系統運作的核心基礎。特別在金融場域中，AI 被廣泛應用於信用評分、詐欺偵測與風險分析等高敏感性決策，其準確性與穩定性高度仰賴歷史資料的完整性與可持續性。然而，一旦模型完成訓練，即使資料控制者事後刪除原始個資，訓練權重中仍可能潛藏個人資訊，再加上模型重新訓練所需之高昂成本，使資料主體即使依法主張「刪除其個人資料」或「停止利用」的權利，亦往往難以實質落實。此一技術特性使得 AI 模型與個人資料刪除權間形成結構性扞格，尤其在金融監理與法遵要求日益嚴格之下，愈形凸顯其制度張力。

我國《個人資料保護法》（下稱個資法）第 11 條第 3、4 項雖形式上賦予資料主體在處理目的消失、期限屆滿或處理違法時請求刪除個資之權利，然其適用情境狹隘，亦缺乏資料控制者具體回應義務與程序保障，恐導致刪除權之保障流於形式。另一方面，第 3 條第 5 款雖明文禁止資料主體預先拋棄或以特約限制刪除權，以強化保障資料主體的個資權，然從法律經濟學角度觀之，此一絕對限制亦可能抑制契約彈性與資料利用誘因，對 AI 模型穩定訓練產生掣肘。

有鑑於此，本文擬以 AI 技術於金融應用場域中所引發之資料刪除困境為出發點，結合比較法分析與法律經濟學視角，檢討我國現行《個資法》之規範適切性，並提出具彈性與正當性之法制改革建議，期能於 AI 技術之發展與資訊自決權保障之間取得合理平衡。

doi:10.6342/NTU202600641

第二節　研究範圍與限制

本研究聚焦於AI技術應用於金融場域時所產生的個人資料刪除權之制度性衝突，旨在說明信用評分、詐欺偵測、風險控管等高敏感決策任務如何可應用AI系統獲得突破，以及其進而衍生的資料殘留與刪除實務困難的問題。之所以選擇金融領域作為分析核心，乃因金融產業向來屬於監管強度最高的領域之一，相關監理法規不僅對人工智慧與資料的相關利用設有明確規範，並透過健全的執法與合規機制加以落實。正因如此，金融領域可視為觀察刪除權在AI應用下所產生制度張力的最佳縮影，亦能為未來跨領域之制度設計提供可資參考的經驗。

其次，本文探討對象為較中性且範圍較廣泛之「個人資料刪除權」，此與所謂「被遺忘權」不盡相同。儘管歐盟《一般資料保護規則》（GDPR）將個資刪除權（the right to erasure）與被遺忘權（the right to be forgotten）同列為第17條之權利名稱，然我國實務對被遺忘權此一概念通常採取較限縮的解釋，聚焦於網路資訊揭露與名譽保護等特定情境[1]。因此，本文選擇以涵蓋面較廣、定義更為中性的個人資料刪除權作為研究焦點，並不直接處理被遺忘權之適用問題。

第三節　研究方法

本文交互運用文獻回顧法、法律經濟分析法以及比較法研究法，以深入探討AI技術與個人資料刪除權之制度性衝突。具體說明如下：

首先，本文透過文獻回顧法，系統性整理與分析AI模型的技術特性、資料處理機制及其與刪除權之內在衝突，並歸納國內外現有技術解方的研究成果，如機器遺忘（machine unlearning）、合成資料（synthetic data）應用以及模型編輯（model editing）等技術，以建立本文的研究分析基礎。

[1] 例如臺灣高等法院104年度上字第1084號判決、最高法院106年度台上字第2652號判決與臺灣高等法院高雄分院107年度上字第259號判決等，即指出當前實務所承認之被遺忘權，係設計以針對搜尋引擎索引與連結之移除，並未涵蓋要求新聞媒體刪除原始報導內容。

doi:10.6342/NTU202600641

其次，本文運用法律經濟分析法，綜合古典經濟學中的寇斯定理（Coase Theorem）、契約自由、市場失靈理論，以及行為經濟學對理性人假設的修正與輕推（nudge）之制度設計建議，分析個資刪除權制度對市場資源配置、資料交換誘因與制度正當性的影響，進一步檢視現行規範可能引發的市場失靈與行為偏誤問題。

最後，本文以比較法研究法作為研究方法之一，選定歐盟《一般資料保護規則》（GDPR）與美國《加州消費者隱私法案》（CCPA）為比較研究對象，從刪除權的適用條件、契約限制與同意機制等角度進行全面性的制度比較分析。此外，本文亦援引歐洲資料保護委員會（EDPB）第 28/2024 號意見書，針對匿名化標準與合法利益評估的具體操作性判準進行深入探討，藉此揭示不同法制如何回應 AI 與個人資料刪除權的挑戰，並作為檢視與反思我國現行制度侷限之重要參照。

第四節　研究架構

本文首先介紹 AI 的基本概念與技術發展，進而檢視我國金融業中 AI 的應用現況，並分析其在金融領域的運作模式與潛在風險。隨後，本文探討 AI 應用所帶來的資料保護挑戰，並整理國內相關規範與指引，作為後續研究的基礎。第三章則回顧個人資料刪除權的發展脈絡與法理內涵，並探討其在現行法制與憲法層面的實踐基礎。接著，剖析機器學習架構下資料刪除的技術困境，揭示 AI 應用與刪除權之間的結構性張力。

第四章進一步以法律經濟分析為核心，分別從寇斯定理、新古典經濟學與行為經濟學三種視角出發，檢視刪除權保障與金融個人資料利用效率之間的制度矛盾。藉由這些分析框架，提出初步的制度設計方向，說明如何在隱私保護與產業發展之間取得平衡。第五章則進行比較法檢視，分別分析歐盟 GDPR 與美國 CPRA 在個人資料刪除權與 AI 應用上的制度回應：前者強調刪除權的不可讓渡性，後者則透

doi:10.6342/NTU202600641

過資料市場化與彈性規範，嘗試兼顧創新與保障。最後進行歐美法制比較，歸納對我國的啟示。

第六章綜合前述的技術挑戰、經濟分析與比較法觀察，提出我國制度改革的可能方向，包括引入有償契約自由機制，以及以行為經濟學為基礎的「輕推」制度設計，如同意的隨時撤回與資訊揭露機制，並配合相關配套措施，以矯正現行制度之不足。最後，第七章總結全文，重申 AI 技術發展與刪除權保障間的核心矛盾，並主張我國應在隱私保障與科技創新間建構一套長遠可行的法制藍圖，以確保資訊自主、促進產業發展，並達致法制與科技的永續平衡。

doi:10.6342/NTU202600641

第二章 人工智慧技術與其金融應用

時至今日，AI 已廣泛應用在我們的日常生活，從語音助理、投資機器人到影音串流平台的推薦系統，皆為 AI 技術的應用成果。在現代科技技術下，AI 不再是電視電影中科幻場景出現的虛構機器人，而是已能被真正實現的高階電腦科學。目前，AI 已被廣泛應用到包含醫療健康、商業金融、自動駕駛車、工業製造、農林漁牧、社福照護、環境防災與生活教育等多重領域，其具有從周遭環境蒐集大量資訊及採用最佳化決策及行動的特質，應用發展對人類生活模式造成深遠影響。

本章擬簡介 AI 之技術及其發展概況，並說明其於金融業之應用及可能衍生之風險，以便初步認識技術應用現況，利於後續進一步之探討。

第一節　人工智慧簡介

第一項　現代人工智慧之發展歷程

人工智慧自 20 世紀中葉誕生以來，經歷了數次技術突破，從早期的符號邏輯推理，到專家系統、機器學習，再到當今的深度學習與大數據分析，AI 技術逐步成熟，應用範圍也不斷擴展。早期 AI 受限於計算資源與應用場景，發展一度停滯；然而，隨著網際網路的興起、行動裝置的普及及運算能力的提升，AI 進入快速發展期。如今，AI 已廣泛應用於各個領域，為社會帶來顯著變革，隨著技術的不斷創新，AI 將持續驅動科技進步，塑造更智慧的未來世界。

第一款　人工智慧的誕生與早期發展

AI 的概念可追溯至西元 1943 年（下同），美國計算神經科學家 Warren McCulloch 和 Walter Pitts 受人類中樞神經系統的啟發，提出了人工神經元模型，

doi:10.6342/NTU202600641

並利用其進行簡單邏輯運算，這一突破促使數學、心理學、工程學等領域的科學家探索人工大腦的可能性[2]。

1956 年，John McCarthy 在達特茅斯學院（Dartmouth College）的一場會議上正式提出「人工智慧」這一術語，並將其定義為「製造智能機器的科學與工程」（the science and engineering of intelligent machines）。該會議被視為 AI 誕生的重要里程碑，開啟了 AI 研究的黃金時代[3]，如 John McCarthy 便於當時創造了最古早的高階程式語言之一——LISP，一個直至今日仍被廣泛應用之 AI 程式語言[4]。

在這個時期，研究人員開發了多種系統，試圖模擬人類複雜的思維過程，以解決各類難題，並挑戰機器無法「思考」的傳統觀念。然而，到 1970 年左右，AI 的發展遇到了瓶頸，由於當時的技術僅能處理簡單或結構明確的問題，除了在遊戲領域表現較為突出外，幾乎沒有實際應用價值。且因計算資源有限，難以應對更高階的推理與決策，AI 研究的進展逐漸放緩，導致資金支持減少，最終進入低潮期[5]。

第二款　專家系統與 AI 的復甦

經過 1970 年代的低潮期，1980 年代 AI 技術迎來新的突破，特別是在專家系統的應用上取得進展。演算法的各式突破為 AI 帶來轉機，開啟了 AI 的第二次技術革命，許多機構與大學開發了專家系統（Expert Systems），透過解決特定領域經典問題的方式，取得大量邏輯推演步驟，藉此使 AI 得以解答多數相似類型的問題，使原先僅能處理簡單問題的系統，能運用於特定專業領域解決問題，達到知識密集型技術的轉變。例如卡內基美隆大學的 XCON 和史丹佛大學的 MYCIN，即為分別將專家系統應用於計算機配置與醫療診斷的例子[6]。

---

[2] *See* Warren S. Mcculloch & Walter Pitts, *A Logical Calculus of The Ideas Immanent In Nervous Activity*, 5 BULL. MATH. BIOL. 115,123 (1943).
[3] *See* JOHN MCCARTHY, WHAT IS ARTIFICIAL INTELLIGENCE? 153 (1997).
[4] *See* John Mccarthy, *History of Lisp*, 13 SIGPLAN NOT. 217, 221 (1978).
[5] *See* MICHAEL NEGNEVITSKY, ARTIFICIAL INTELLIGENCE: A GUIDE TO INTELLIGENT SYSTEMS 5, 7-8 (3rd. 2011); Yongjun Xu et al., *Artificial intelligence: A powerful paradigm for scientific research*, 2(4) THE INNOVATION.,100179., 1, 2 (2021).
[6] *See* Xu et al., *Id.*

doi:10.6342/NTU202600641

然而，當專家系統逐步發展時，其侷限性也日益凸顯。雖然部分專家系統在特定領域展現出色表現，但整體而言，其應用範圍受限、缺乏靈活性，維護成本高昂。由於專家系統專注於特定領域，無法處理跨領域問題，當遇到非典型案例時，因缺乏自我學習能力，往往難以做出準確判斷。此外，專家系統無法解釋決策過程或驗證答案的正確性，使得識別錯誤變得困難，降低了其可信度，也導致市場需求逐漸減少[7]。與此同時，日本政府大力推動的「第五代計算機計畫」（Fifth Generation Computer Project）未能達成預期目標，使得 AI 研究的資金支持再次減少，AI 發展因而進入第二次低潮[8]。

第三款　深度學習的興起

進入 21 世紀，隨著計算能力與資料量的激增，深度學習技術開始快速發展，為 AI 帶來革命性進展。2006 年，Geoffrey Hinton 及其團隊提出了一種構建深層神經網絡的方法，並解決了訓練過程中的「梯度消失」（Gradient Vanishing）問題，重新點燃 AI 研究熱潮[9]。

隨著計算能力的提升和大數據的發展，深度學習（Deep Learning, DL）迅速成為 AI 領域的核心技術，並在各種應用中展現出卓越的能力。詳言之，深度學習屬於機器學習（Machine Learning, ML）的一個子領域，透過多層神經網絡的表示學習（Representation Learning） 來自動提取資料特徵，從而提升模型的識別與預測能力。而機器學習則是 AI 的關鍵組成部分，使計算機能夠在無需人類直接介入的情況下，透過資料學習模式並獲取智能。因此，「學習」成為這一時代 AI 研究的核心概念，推動 AI 進入一個全新的發展階段，為計算機視覺、自然語言處理、自動駕駛等領域帶來革命性突破[10]。

---

[7] *See* NEGNEVITSKY, *supra* note 5, at 9.
[8] *See* Xu et al., *supra* note 5.
[9] *See* Geoffrey E. Hinton et al, *A Fast Learning Algorithm Deep Belief Nets*, 18 NEURAL COMPUTATION 1527, 1546–1548 (2006); Geoffrey E. Hinton, & Salakhutdinov, R.R., *Reducing the Dimensionality of Data with Neural Networks.* 313 SCIENCE 504, 504–507 (2006).
[10] *See* Yann LeCun et al, *Deep learning.* 521 NATURE 436, 436–442 (2015).

doi:10.6342/NTU202600641

第二項　人工智慧之定義及其技術

第一款　人工智慧之定義

人工智慧一詞，最早出現於 1956 年的達特茅斯會議（Dartmouth Summer Research Project on Artificial Intelligence），與會者之一的 John McCarthy 將人工智慧定義為「製造智能機器的科學與工程」（the science and engineering of making intelligent machines）[11]，但其同時也認為人們仍無法明確定義究竟何種程度的計算機程式可被稱作「智慧」[12]。

較近期的見解則來自世界智慧財產權組織（World Intellectual Property Organization, WIPO），該組織從技術層面定義人工智慧，並於 2019 年發布的《人工智慧技術趨勢報告》中指出，人工智慧並非單一技術概念，而是涵蓋多種細分技術[13]，主要包括：（一）機器學習（Machine Learning）；（二）邏輯程式設計（Logic programming）；（三）模糊邏輯（Fuzzy Logic）；（四）概率推理（Probabilistic Reasoning）；（五）本體工程（Ontology Engineering）；（六）功能應用關聯技術（Functional Application）。其中，功能應用關聯技術又可進一步細分為：（1）電腦視覺：（2）自然語言處理；（3）語音處理及（4）其他功能應用項目，此報告強調，人工智慧涉及多種技術領域，並在不同應用場景中發揮關鍵作用，體現 AI 的多樣性與發展潛力[14]。

除了技術層面的定義外，經濟合作暨發展組織（Organisation for Economic Cooperation and Development, OECD）在 2019 年發布的《人工智慧建議書》（Recommendation of the Council on Artificial Intelligence）[15]中，對「人工智慧系統」

---

[11] *See* MCCARTHY, *supra* note 5.
[12] *Id.*
[13] *See* WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO), WIPO TECHNOLOGY TRENDS 2019: ARTIFICIAL INTELLIGENCE 24 (2019).
[14] *Id.* at 24-26.
[15] *See* Organisation for Economic Cooperation and Development(OECD), *Recommendation of the Council on Artificial Intelligence* (2019), https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

doi:10.6342/NTU202600641

（AI System）的定義亦廣受引用。根據 OECD 的定義，人工智慧系統是「一種基於機器的系統，能夠針對人類設定的目標，作出影響現實或虛擬環境的預測、建議或決策，並可依設計具備不同程度的自主性。」此定義強調 AI 系統的核心特性，即其自主運作能力，並涵蓋從輔助決策到完全自主運行的不同應用場景，為 AI 的發展與治理提供了重要參考。

關於我國法律對 AI 的定義，目前政府正積極推動 AI 的一般性立法。我國於 2026 年發布的《人工智慧基本法》，其參考美國國家人工智慧創新法案（National AI Initiative Act of 2020）等規範，於第二條對人工智慧（AI）作出如下定義：「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[16]。」此定義強調 AI 的自主運行能力，並涵蓋機器學習、演算法運算等核心技術，突顯 AI 在決策、內容生產與環境影響等方面的應用廣度，為未來 AI 監管與法制發展提供基礎框架。

至於 AI 在金融領域的應用規範，我國金融監督管理委員會則於 2024 年公告《金融業運用人工智慧（AI）指引》，針對 AI 系統及生成式 AI（Generative AI）作出明確定義。根據該指引，AI 系統係指「透過大量資料學習，運用機器學習或相關建模演算法，執行感知、預測、決策、規劃、推理、溝通等功能，以模仿人類的學習、思考與反應模式之技術。」此外，其亦對生成式 AI（Generative AI）作出定義：「透過大量資料學習，能夠生成模擬人類智慧創造之內容的技術，其內容形式包括但不限於文本、圖片、聲音、照片、影像及程式碼等。」[17]此指引反映了 AI 在金融領域的發展趨勢，並在 AI 的定義上著重大量的資料學習、演算法應用及內

[16] 全國法規資料庫（2026），《人工智慧基本法》，https://law.moj.gov.tw/News/NewsDetail.aspx?msgid=196197（最後瀏覽日：2026 年 2 月 2 日）。

[17] 金融監督管理委員會（2024），《金融業運用人工智慧（AI）指引》，頁 3，https://www.fsc.gov.tw/websitedowndoc?file=chfsc/202408231741530.pdf&filedisplay=%E9%99%84%E4%BB%B6_%E9%87%91%E8%9E%8D%E6%A5%AD%E9%81%8B%E7%94%A8AI%E6%8C%87%E5%BC%95.pdf。此一定義與我國銀行商業同業公會於自律規範《金融機構運用人工智慧技術作業規範》第二條所界定之人工智慧概念一致。

doi:10.6342/NTU202600641

容生成能力，為金融機構導入 AI 技術提供指導原則，以確保技術應用的合規性與穩定性[18]。

綜合上述各項定義可知，雖各機構對於 AI 的界定略有差異，然普遍具有一定程度的廣泛性與相似性。就 OECD 與我國相關定義之差異觀察，OECD 所採定義較為寬泛，重點在於 AI 系統的自動化能力；相對地，我國之定義則較為狹隘，除要求系統具備自主運作能力外，尚須結合機器學習等核心技術之應用。鑑於本文聚焦於 AI 在金融領域之應用，諸如智慧客服與理財機器人等實務案例是否符合我國較為狹義之定義，仍存有解釋爭議。故為求討論基礎之穩定與完整，本文擬援用 OECD 較為寬泛之定義作為後續討論之基礎。

第二款　人工智慧、機器學習與深度學習

人工智慧是一個廣義的概念，旨在讓電腦模擬人類的行為與決策，而機器學習是實現 AI 的一種方法，使電腦能夠透過資料學習規則與模式，無需明確編寫程式規則。深度學習則是機器學習的一個子領域，透過多層神經網絡進行更複雜的特徵學習與模式識別，推動 AI 技術的進一步發展[19]。

第一目　機器學習

機器學習的基本原理是透過演算法解析資料、從中學習模式，並據此進行判斷或預測[20]。有別於手動撰寫固定指令集來讓軟體執行特定任務，機器學習透過大量資料與演算法訓練機器，使其能夠自主學習並適應不同任務需求。其核心目標

---

[18] 除了前述定義外，人壽保險商業同業公會在《保險業運用人工智慧系統自律規範》中，則有不同的定義，如該規範第二條謂 AI 系統：係指透過大量資料學習，利用機器學習或相關建立模型之演算法，進行感知、預測、決策、規劃、推理、溝通等模仿人類學習、思考及反應模式之系統。生成式 AI 則係指可以生成模擬人類智慧創造之內容的相關 AI 系統，其內容形式包括但不限於文章、圖像、音訊、影片及程式碼等。

[19] See Michael Copeland, *What's the Difference Between Artificial Intelligence, Machine Learning and Deep Learning?*, NVIDIA, (July 29, 2016), https://blogs.nvidia.com/blog/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/.

[20] See *What Is Machine Learning?*, AMAZON WEB SERVICES (last visited Mar. 12, 2025), https://aws.amazon.com/tw/what-is/machine-learning/; *What's the Difference Between AI and Machine Learning?*, AMAZON WEB SERVICES (last visited Mar.12, 2025), https://aws.amazon.com/compare/the-difference-between-artificial-intelligence-and-machine-learning/?nc1=h_ls.

doi:10.6342/NTU202600641

是讓計算機從資料中提取規則，無需人工編寫明確的程式邏輯，即可進行預測與決策[21]。

機器學習的發展依賴於統計學、數學與計算技術的進步，隨著資料量的增長與計算能力的提升，機器學習已成為 AI 的重要基礎，並在各領域發揮關鍵作用，如語音識別、推薦系統、金融風險評估等[22]。

機器學習的過程可簡化為資料輸入、特徵擷取、訓練模型、輸出結果（判斷）[23]。在進行學習前，首先需要準備資料，當資料輸入後，機器學習會進行特徵擷取（feature extraction），即從資料中選取對預測結果最具代表性的關鍵特徵。例如，在貓狗分類問題中，可以假設花色、耳朵形狀、臉型是區分貓與狗的主要特徵，而這些特徵是由人為選定的，完成特徵擷取後，資料會被輸入到訓練模型中，對於分類任務而言，這個模型即為分類器。經過訓練後，當輸入新的圖片時，模型即可根據學習到的特徵進行分類，最終輸出我們需要的判斷結果，即圖片中的動物是貓還是狗[24]。

第二目　深度學習

深度學習是機器學習的一個子領域，透過類神經網絡（Artificial Neural Networks, ANN）模擬人腦的學習模式，其關鍵特點是多層網絡架構，能夠從大量資料中自動提取複雜的特徵，進而提升學習能力與準確性[25]。

與傳統機器學習相同，深度學習的訓練過程也需要先準備大量資料。然而，與機器學習不同的是，深度學習不需要人為進行特徵擷取，而是透過神經網絡（Neural Networks）自行學習與提取資料中的關鍵特徵。訓練者只需將資料輸入

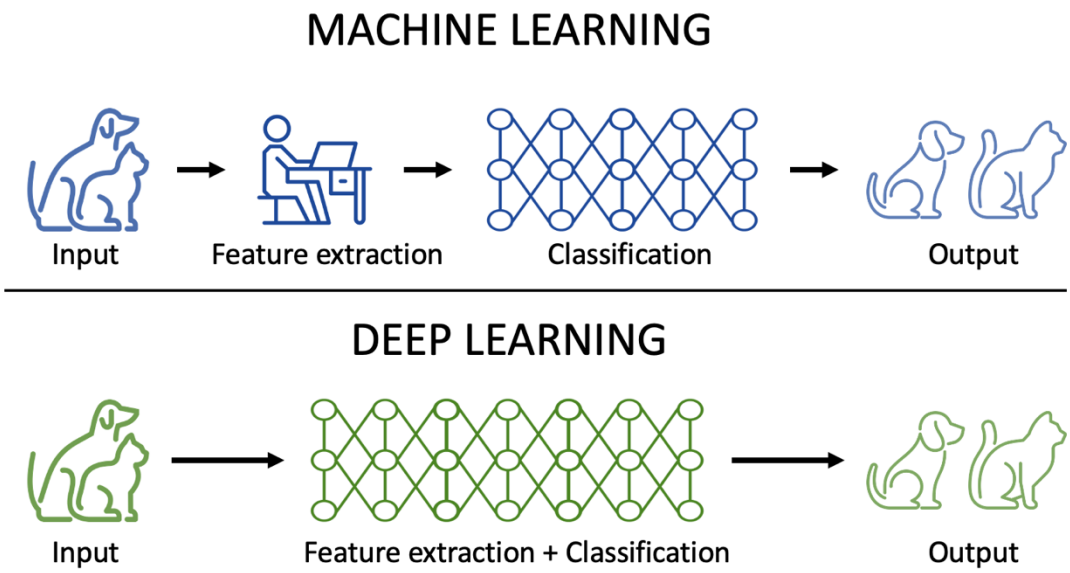
---

[21] *Id.*
[22] *Id.*
[23] *Id.*
[24] *Id.*
[25] See generally *What is Deep Learning?*, AMAZON WEB SERVICES (last visited Mar. 12, 2025), https://aws.amazon.com/what-is/deep-learning/?nc1=h_ls; *What's the Difference Between Machine Learning and Deep Learning?*, AMAZON WEB SERVICES (last visited Mar. 12, 2025), https://aws.amazon.com/compare/the-difference-between-machine-learning-and-deep-learning/?nc1=hls.

doi:10.6342/NTU202600641

訓練模型，神經網絡便會自動進行特徵擷取，進而進行判斷並輸出結果[26]。例如，在貓狗分類問題中，傳統機器學習可能會基於人為選擇的花色、耳朵形狀等特徵來區分貓與狗，而深度學習則能從大量資料中自行發掘潛在特徵，如眼睛顏色、毛髮紋理，甚至是人類難以理解的組合特徵。最終，深度學習模型的判斷結果可能呈現為 $A = dX + eV^2 + fW + gUV$，其中 X 可能是已知的花色，而 V、W、U 等可能是人類未曾考慮過的或無法直觀解釋的特徵。隨著資料量的增加，深度學習能夠擷取的特徵維度也越來越多，進而使模型的預測能力更加精準。研究顯示，當資料量達到一定規模後，深度學習的效果將遠超傳統機器學習算法，展現出更強的適應能力與泛化能力[27]。

綜上所述，機器學習與深度學習的訓練流程可圖象化如圖一。



【圖 一】機器學習與深度學習的訓練流程
圖片來源：本文自行繪製；Icon: Flaticon

深度學習之所以能夠成為 AI 領域的主流技術，主要受益於資料量的增加與圖形處理器（Graphics Processing Unit, GPU）技術的成熟。深度學習依賴大量資料來訓練模型，以提升學習效果，而隨著網際網路與通訊技術的進步，資料的蒐集

[26] *Id.*
[27] *Id.*

doi:10.6342/NTU202600641

與獲取變得更加高效，從而使得深度學習的應用範圍大幅擴展[28]。此外，相較於傳統機器學習，深度學習在處理非結構化資料、發掘隱藏模式、無監督學習及動態資料處理方面展現了顯著優勢。其模型能夠自動擷取資料特徵，無需人為標註，並能從大量資料中學習模式，例如分析客戶行為以提供個人化推薦，或透過監測金融交易偵測詐欺活動。同時，深度學習能夠根據使用者行為持續調整模型，隨時間優化其預測能力[29]。

另一方面，GPU 技術的發展也大幅提升了深度學習的計算能力。過去深度學習主要依賴中央處理器（Central Processing Unit , CPU）進行運算，但由於 CPU 架構不適合大量並行計算，導致模型訓練時間過長。相對而言，GPU 具備優異的並行運算能力，能夠加速神經網絡的訓練，使深度學習技術得以更廣泛應用，進而推動 AI 的進一步發展[30]。

然而，深度學習在實作上仍面臨諸多挑戰，其中最主要的問題包括大量高品質資料的需求，以及對高效能運算資源的依賴。首先，深度學習模型的效能高度取決於訓練資料的品質，倘若資料集中存在異常或錯誤（例如動物影像分類資料集中誤植飛機圖片），可能導致模型產生錯誤判斷[31]，故資料的蒐集、清理與預處理成為至關重要的步驟，並且需要大量的儲存空間。此外，深度學習演算法的運算極為密集，訓練過程需要強大的基礎設施例如 GPU 叢集，以支撐其高效能計算。倘若運算資源不足，模型訓練時間將大幅延長，進而影響效能與應用推廣。這些挑戰顯示深度學習雖然具備高度潛力，但仍需克服資料品質管理與運算資源的限制，以確保技術能夠發揮最佳效能[32]。

---

[28] 參閱 林宏軒、李肇棠、江滄明（07/26/2018），〈深度學習的訓練資料準備與平台之演進發展〉，《電腦與通訊》，https://ictjournal.itri.org.tw/xcdoc/cont?xsmsid=0M236556470056558161&sid=0M265628861865857206（最後瀏覽日：03/12/2025）。

[29] 同前註。

[30] See AMAZON WEB SERVICES, *supra* note 25.

[31] Xue-Wen Chen & Xiaotong Lin, *Big Data Deep Learning: Challenges and Perspectives*, 2 IEEE ACCESS 514, 521 (2014), https://doi.org/10.1109/ACCESS.2014.2325029.

[32] *Id.*

doi:10.6342/NTU202600641

第三目　小結

機器學習與深度學習的關係是「包含與被包含」的關係，深度學習是機器學習的一部分，而機器學習則屬於 AI 的範疇。機器學習適用於資料分析與預測，而深度學習則特別適合處理大量且高維度的資料，如影像、語音與語言理解，兩者的差異比較整理如以下【**表 一**】所示。隨著計算能力的提升與資料資源的增加，深度學習成為當今 AI 發展的關鍵推動力，並持續影響各個產業與科技應用領域，也讓各種機器輔助成為可能。

特性	機器學習（ML）	深度學習（DL）
依賴的資料量	相對較少	需要大量資料
特徵工程	需要人為設計	自動提取特徵
計算需求	相對較低	需要高效能計算（如 GPU）
適用範圍	廣泛應用於各類資料分析	主要用於影像、語音、文字等複雜資料處理

【表 二】機器學習與深度學習之比較
資料來源：本文整理。

第二節　我國金融業應用人工智慧現況

在金融科技快速演進的背景下，我國金融業亦逐步導入 AI 技術以強化營運效能與法遵管理。根據金融監督管理委員會於本（114）年 4 月所進行之調查，針對全體金融機構及周邊單位共計 383 家進行統計，結果顯示已有 126 家業者導入

doi:10.6342/NTU202600641

AI 技術，占比約 33%，相較於前一年之 29%略有提升[33]。就不同金融業態觀察，導入 AI 比例最高者為銀行業（87%），其次為壽險公司（67%）與產險公司（45%）[34]。

該調查進一步指出，金融機構導入 AI 之主要目的在於提升作業效率（30%）、節省人力成本（18%）以及優化客戶體驗（15%）。實際應用場域則集中於內部行政作業、自動化客服系統以及防制金融犯罪等領域[35]。在技術選擇方面，最常見者為自然語言處理（Natural Language Processing, NLP）與大型語言模型（Large Language Models, LLM），使用率達 31%；其次為機器學習（Machine Learning, 28%）與機器人流程自動化（Robotic Process Automation, RPA, 23%）[36]。

值得注意的是，目前已有 61 家金融機構導入生成式 AI 技術，占所有使用 AI 之機構的 48%，年增幅達 21 個百分點。生成式 AI 的主要應用領域為內部行政作業（39%）與智能客服（15%）。然而，業者普遍面臨生成內容不穩定或錯誤、資料安全與隱私疑慮，以及合規風險等挑戰[37]。

從未來發展趨勢觀察，有高達 47%（共 179 家）金融機構表示將導入或擴大 AI 應用範圍，重點多集中於優化內部行政流程、提升智能客服效能及強化打擊金融犯罪之能力。同時，部分業者亦指出，未來若透過跨機構合作發展防詐 AI 技術、本土大型語言模型與風險管理工具，將有助於提升整體產業的技術水準與應用成熟度[38]。

由上述說明可知，儘管金融業對 AI 技術之採納程度逐年提升，但在決策層面的自動化應用仍相對有限（43%業者無自動化決策，40%業者自動化決策比例在

[33] 金融監督管理委員會（05/20/2025），〈金管會公布金融業應用人工智慧(AI)調查結果〉，https://www.fsc.gov.tw/ch/home.jsp?id=96&parentpath=0,2&mcustomize=news_view.jsp&dataserno=202505200001&dtable=News（最後瀏覽日：07/23/2025）。

[34] 同前註。

[35] 同前註。

[36] 同前註。

[37] 同前註。

[38] 同前註。

doi:10.6342/NTU202600641

25%以下），顯示業界整體對於 AI 決策功能仍持謹慎態度。就此，為因應 AI 技術之快速發展與應用挑戰，金管會已透過金融科技產業聯盟啟動相關合作機制，鼓勵業者積極投入 AI 技術共研與應用場域實驗，期能在法遵要求與創新發展間取得平衡[39]。

第三節　人工智慧於金融領域中之實際應用

隨著大數據、機器學習、演算法、人工神經網路、深度學習、人臉識別等 **AI** 及相關技術的進步和應用，**AI** 的應用範圍越來越廣。時至今日，AI 已廣泛應用於金融業，涵蓋客戶服務、風險管理、法令遵循、金融監管等多個領域，透過機器學習與大數據分析，提升運營效率、優化決策流程，並改善客戶體驗。本節將探討 AI 在金融業的主要應用場景，並分析其帶來的影響與挑戰。

第一項　人工智慧在客戶服務與業務運營中的應用

第一款　智慧客服

所謂智慧客服，係由智慧型機器人取代人工客服，為金融消費者提供業務引導、查詢、辦理及投訴處理等多種服務的新型顧客服務模式。隨著智慧語音、機器學習及自然語言處理（Natural Language Processing, NLP）等技術的進步，智慧型機器人的反應速度、溝通能力與應變能力大幅提升。不僅能針對用戶需求提供準確解答，甚至可依據客戶的年齡、身份及文化程度調整回應方式，提升互動體驗[40]。

與傳統人工客服相比，智慧客服機器人具有回應速度快、服務成本低、處理量大、無時間限制等優勢。它不僅能夠有效分流客戶問題，減少人工客服的負擔，還能提升客戶滿意度，增加銀行的運營效率。例如，台新銀行推出的「Rose」能即時

[39] 同前註。

[40] Diego Salinas, *Intelligent Customer Service in 2025: A Guide on Benefits & Implementation*, CLOUDTALK (Dec. 10, 2024), https://www.cloudtalk.io/blog/intelligent-customer-service.

doi:10.6342/NTU202600641

提供客戶諮詢、帳務查詢等服務，減少人工客服的負擔[41]。而中國信託銀行的中國信託銀行的「小 C」客服機器人結合機器學習與流程自動化（Robotic Process Automation, RPA），不僅能夠識別客戶問題的深層需求，還能根據歷史資料預測下一步需求，提升客戶體驗[42]。此外，台北富邦銀行的「邦妮」則根據使用情境和業務種類分類，快速提供準確度高達 90%的解答[43]，這些創新應用不僅進一步縮短服務流程，還大幅提升了客戶體驗和整體運營效率。

第二款　AI 理財機器人（Robo-Advisors）

理財機器人依據投信投顧公會所訂《證券投資顧問事業以自動化工具提供證券投資顧問服務(Robo-Advisor)作業要點》（下稱理財機器人作業要點）的定義[44]，指「完全經由網路互動，全無或極少人工服務，而提供客戶投資組合建議的顧問服務」[45]，此類服務主要透過演算法（algorithms）與電腦系統的自動執行，提供線上理財諮詢與投資管理，人工介入則僅屬輔助性質[46]。具體而言，投資人可透過業者的線上平台開戶，填寫風險評估與財務目標等相關問卷，業者則利用演算法分析這些資訊，進而產生多種投資配置建議，提供客製化投資組合。此外，當投資組合確

---

41 台新銀行（02/17/2025），〈台新三十 智能客服 Rose 再升級 領先改版照顧無障礙需求〉，https://www.taishinbank.com.tw/TSB/personal/common/news/TSBankNews-003658/（最後瀏覽日：02/17/2025）。

42 數位時代（05/31/2021），〈中國信託以客戶視角發展聊天機器人 機器人客服小 C 會傳 LINE、講電話多元溝通模式滿足客戶需求〉，https://www.bnext.com.tw/article/63044/bankingmyway-07?（最後瀏覽日：2025 年 2 月 17 日）。

43 工商時報（05/05/2024），〈銀行業 台北富邦銀行秉持 361 度服務 多一度溫暖客戶心〉，https://www.chinatimes.com/newspapers/20240505000234-260210（最後瀏覽日：2025 年 2 月 17 日）。

44 值得注意的是，金管會於 2024 年正式將理財機器人（Robo-Advisor）服務納入《證券投資顧問事業管理規則》之監理架構，新增第五章之一作為專章規範。該修法參考了中華民國證券投資信託暨顧問商業同業公會所訂定之《證券投資顧問事業以自動化工具提供證券投資顧問服務（Robo-Advisor）作業要點》，明確定義自動化投資顧問服務之範疇，並針對提供此類服務之業者，規範其財務與業務條件、內部控制或內部管理制度要求，以及外部監理標準。換言之，目前我國對理財機器人之監理，係透過主管機關制定的《證券投資顧問事業管理規則》第 25 條之 1 至第 25 條之 8，配合業界公會所訂作業要點，共同構築起監督與實務運作的雙層架構。此舉不僅提升法制明確性，也有助於兼顧金融創新發展與投資人保護。

45 投信投顧公會理財機器人作業要點第 2 點第 1 項。

46 投信投顧公會理財機器人作業要點第 1 點第 1 項。

doi:10.6342/NTU202600641

立後，演算法可自動執行投資組合的調整與再平衡，以確保資產配置維持在既定策略範圍內[47]。

隨著大數據時代的來臨與 AI 技術快速的發展，越來越多金融機構導入理財機器人，以提升服務品質與市場競爭力。例如，中國信託銀行的「智主投」運用 AI 進行風險評估，提高投資組合的適配性，而國泰世華銀行則透過 AI 進行市場分析，優化投資策略[48]。根據金管會統計，截至 2025 年 2 月底，全台機器人理財資產規模已達 131 億元，客戶數則破 21.6 萬人。顯示各大金融機構正積極推動智能理財服務，持續擴大市場規模，預示著未來金融科技在理財領域的應用將更加廣泛與深入[49]。

相較於傳統人力提供的理財服務，理財機器人具備多項顯著優勢[50]。演算法可快速蒐集、分析並處理大量資料，在投資決策的速度、準確性與複雜度處理上勝過人力，降低資訊不對稱並提升投資報酬率。此外，演算法突破人力極限，能提供全天候理財服務，確保投資決策的理性與一致性，避免人為因素影響服務品質。它亦能控管不當行為，如挪用資產、利益衝突等，確保理財服務的公平與透明，同時降低業者管理成本[51]。

此外，理財機器人的另一大優勢在於提升理財服務的可及性，推動普惠金融。傳統理財因人力成本高，主要服務高淨值投資人，一般投資者難以獲得專業建議。

---

[47] 參閱：陳安斌、陳莉貞、蘇秀玲、郭怡君（2018），〈我國發展機器人理財顧問之研究〉，《中華民國證券投資信託暨顧問商業同業公會委託報告》，頁 16、26-27 頁。

[48] 經濟日報（10/22/2024），〈智能理財客戶 小資族最多 多數選擇定期定額〉，https://money.udn.com/ money/story/5613/8306733（最後瀏覽日：最後瀏覽日：02/25/2025）。

[49] 自由時報（04/24/2025），〈機器人理財資產規模達 131 億 客戶數破 21.6 萬人〉，https://ec.ltn.com.tw/article/breakingnews/5022828（最後瀏覽日：07/25/2025）。

[50] 相關討論，參照：楊岳平（2020），〈演算法時代下的投資顧問監理議題──以理財機器人監理為例〉，《月旦民商法雜誌》，67 期，頁 32；周振鋒（2019），〈論機器人投資顧問之興起與投資人之保護──以美國法為中心〉，《東吳法律學報》，30 卷 4 期，頁 73-76；谷湘儀等（2016），〈機器人投資顧問(Robo-Advisor)國外實務及相關法令與管理措施之研究〉，《資產管理產業發展與人才培育基金委託專題研究》，頁 15、16。

[51] 同前註。

doi:10.6342/NTU202600641

理財機器人透過降低人力需求與規模經濟效應，大幅降低服務成本，使低淨值族群也能負擔專業理財，縮小投資機會落差，促進資本市場公平性[52]。

綜合而言，演算法在資訊處理、即時服務與風險控管方面展現出傳統人力難以比擬的優勢，不僅推動金融科技發展，也重塑了現代投資模式，使其更為高效、透明與普及。此外，透過降低理財服務成本與提升可及性，理財機器人能夠讓更多不同資產階級的投資者獲得專業理財建議，縮小高低淨值族群之間的投資機會落差，進而推動普惠金融的實現，使金融服務更加公平與廣泛。

第二項　人工智慧在風險管理與信用評估中的應用

第一款　AI 授信業務與信用風險評估

在快速變化的金融環境下，授信業務是銀行穩健經營的核心，而市場風險與不確定性的增加，使得如何在擴大信用投放的同時有效控管風險成為金融機構的重要課題。傳統授信機制主要依賴財務報表、信用評分與人工審核，雖然對大型企業而言較為有效，但對於中小企業與缺乏信用紀錄的個人而言，因資料有限，往往難以獲得公平評估[53]。此外，傳統人工審核流程繁瑣，放款時間冗長，影響資金流動性與客戶體驗，尤其在疫情後資金需求激增的情境下，更突顯出傳統授信模式的侷限性。

隨著 AI 與大數據技術的發展，金融機構開始導入資料驅動的風險評估模型，以提升信用評分的準確度並優化貸款審核流程。相較於依賴結構化財務資料的傳統信用評估方式，AI 能夠整合更多元的非傳統資料來源，如社交媒體活動、電商交易紀錄、電信資料與消費行為等，透過機器學習分析這些資訊，建立更全面的風險評估機制[54]。美國的 Kabbage 金融科技公司即運用 AI 分析小型網商的交易與物

[52] 楊岳平，同前註，頁 33。
[53] 參閱 蘇柏鳴（2018 年 6 月），〈徵信與 Fintech 發展淺談〉，《金融聯合徵信》，第 32 期，頁 17-20；張文村、呂宜穎、詹雅慧、林淑萍、陳宇軒、葉育惠（2019 年 6 月），〈人工智慧技術應用於中小企業徵信之初探〉，《電工通訊季刊》，第 2 季，頁 82、83。
[54] 張文村等，同前註，頁 83–86。

doi:10.6342/NTU202600641

流資料，實現六分鐘內快速核貸，並成功將壞帳率降至約 1%[55]。另一家業者 ZestFinance 則透過融合法律紀錄、社交資料與水電帳單等超過 70,000 個變數，運用集成式學習技術提升信用風險評估準確度，據其統計，該技術較傳統信用評估模式提升 60%，顯示 AI 在徵信領域的潛力[56]。

除了提升信用評估的準確性，AI 亦能大幅縮短授信流程，提高審核效率。傳統人工審核通常需要數日甚至數週，而 AI 技術的導入使得核貸流程得以自動化，減少人工參與，讓貸款申請者能夠在更短時間內獲得資金。例如，台新銀行於 2021 年推出的「手 t 貸」智慧信貸服務，結合 AI 智能徵審模型，讓客戶最快可在 178 秒內完成核貸，大幅提升了申貸效率[57]。中國信託則透過 AI 風控技術建構「中信腦」審查系統，將個人貸款申請的審核時間從三日縮短至數分鐘，透過技術創新提升放款速度與客戶體驗，展現出 AI 在授信流程優化上的應用價值[58]。

AI 技術的發展亦促進了普惠金融的落實。傳統信用評估模式主要依賴借款人的信用歷史與財務報告，使得無信用紀錄或長期未使用金融服務的消費者被排除在借貸體系之外[59]。然而，AI 能夠透過替代資料補足信用評估的盲點，讓金融機構能夠更全面地評估借款人的信用風險，從而提升金融服務的可及性[60]。例如美國金融科技公司 Affirm 創新推動「先買後付」（BNPL）模式，運用 AI 分析消費者交易行為、社群資料與其他非傳統資料來源，建立信用風險評估機制，使缺乏信用歷

---

55 同前註，頁 82、83。

56 同前註。

57 台新銀行（04/09/2021），〈台新銀「手 t 貸」AI 智能徵審 數位信貸撥款加速 10 倍〉，https://www.taishinbank.com.tw/TSB/personal/common/news/TSBankNews-002349/（最後瀏覽日：02/24/2025）。

58 中央通訊社（05/07/2018），〈中信銀導入 AI 核貸只要幾分鐘〉，https://tw.news.yahoo.com/%E4%B8%AD%E4%BF%A1%E9%8A%80%E5%B0%8E%E5%85%A5ai-%E6%A0%B8%E8%B2%B8%E5%8F%AA%E8%A6%81%E5%B9%BE%E5%88%86%E9%90%98-123729970.html（最後瀏覽日：02/24/2025）。

59 Niccolo Mejia, *AI for Credit Scoring - An Overview of Startups and Innovation*, EMERJ (Dec. 10, 2018), https://emerj.com/ai-sector-overviews/ai-for-credit-scoring-an- overview-of-startups-and-innovation/; Eric VonDohlen, *How Does AI-Based Credit Scoring Fare Against Traditional Credit Scoring?*, FINEXTRA (Nov. 27, 2018), https://www.finextra.com/blogposting/16343/how-does-ai-based-credit-scoring -fare-against-traditional-credit-scoring.

60 *Id.*

doi:10.6342/NTU202600641

史的消費者仍能獲得合理的貸款條件[61]。這種模式不僅提升了借款人獲取貸款的機會，也減少了因缺乏信用紀錄而導致的金融排除問題，進一步推動普惠金融的發展。

綜上所述，AI 技術的發展正在重塑授信業務的運作模式，透過大數據與機器學習技術提升信用評估準確性、加速核貸流程，並擴大金融服務的可及性，使更多消費者能夠獲得公平的借貸機會。對金融機構而言，AI 的應用不僅有助於降低信用風險與營運成本，也能提升市場競爭力，推動普惠金融的實現。隨著 AI 技術的不斷進步，未來授信業務將更加高效、透明，為金融產業帶來深遠變革。

第二款　反詐欺與洗錢防制（AML）

隨著金融科技的快速發展，交易模式與服務內容不斷革新，跨境交易增加與新型模式興起，使業務更加複雜，特別是在阻詐與反洗錢方面挑戰加劇。據我國內政部警政署 165 打詐儀錶板統計顯示，2024 年 8 至 12 月詐騙案件超過 9.3 萬件，單月財損金額達上百億元，突顯金融機構在保護民眾財產安全方面的重要性[62]。

為強化阻詐能力，銀行紛紛導入 AI 與大數據技術並加強跨機構合作。滙豐銀行近期加入「鷹眼識詐聯盟」，成為首家外資銀行成員，該聯盟透過 AI 模型即時掃描可疑交易、快速警示異常帳戶並加速自動審核，提高詐騙偵測準確度。台北富邦銀行亦運用該技術，在詐騙發生前三個月即能偵測異常交易，去年成功攔截逾 2.5 萬件案件並圈存資金 6.4 億元；永豐銀行、玉山銀行與渣打銀行也發展 AI 模型與集團系統，提升交易安全與詐騙防制能力[63]。

在洗錢防制方面，台北富邦銀行率先推出 AI 洗錢防制「獵鷹系統」，成為全國首個應用 AI 模型的洗錢防制系統。該系統可每秒處理約 50 筆交易，具備自主學

[61] 呂依舫（12/07/2021），〈【美股研究報告】金融科技新創 Affirm 的魅力何在?與美國前三大電商合作，更成為亞馬遜獨家 BNPL 供應商！〉，《CM Money 投資網誌》，https://www.taishinbank.com.tw/ TSB/personal/common/news/TSBankNews-002349/（最後瀏覽日：02/24/2025）。

[62] 工商時報（02/06/2025），〈AI 扮利器 銀行阻詐戰力升〉，https://www.ctee.com.tw/news/20250206700162-439901（最後瀏覽日：02/24/2025）。

[63] 同前註。

doi:10.6342/NTU202600641

習能力，初期預計可降低 45%的假警報率，精準偵測並迅速通報可疑交易，有效阻斷犯罪金流，維護金融秩序與民眾資產安全[64]。

隨著數位轉型與金融科技的迅猛發展，金融機構面臨著越來越複雜的風險挑戰。透過引入 AI、大數據以及跨機構合作技術，銀行正有效提升阻詐與反洗錢能力，從而更精準地監控可疑交易與阻斷不法金流，全面守護民眾財產與金融秩序。

第四節　AI 於金融業之應用風險與刪除權挑戰

隨著人工智慧技術在金融領域的應用日益普及，金融機構愈來愈依賴大量個人資料以優化服務、強化風險管理並提升營運效率。在這樣的資料密集運作環境下，AI 系統的發展與部署不僅帶來效率與創新，也引發關於個人資料處理的重大爭議，尤有甚者，當資料主體依據個資法或其他隱私法規主張刪除其個人資料時，金融機構究應如何回應，恐成為實務上難以迴避的挑戰。

由於 AI 模型在訓練過程中係將資料內隱的存儲在模型權重中，其完成訓練後往往難以追溯原始資料，並可能持續利用該等資料於推論與決策之中，致使「資料刪除權之落實困難」成為核心風險之一[65]。本節即以此風險為核心，說明 AI 技術於金融業應用時所伴隨之資料隱私侵害、偏誤決策、系統透明度不足與合規壓力，並分析刪除請求提出時，相關風險如何進一步擴大，以及監理與技術層面所面臨之實際挑戰與回應。

---

[64] 工商時報（02/06/2024），〈年關近慎防詐騙！北富銀 AI 洗防系統正式啟用 全國首創 AI 洗錢防制模型〉，https://www.ctee.com.tw/news/20240206701304-430304（最後瀏覽日：02/24/2025）。

[65] 關於 AI 模型中資料刪除的困難，本文於第三章第三節另有詳述。參見 Meem A.M., *Eternal Sunshine of the Mechanical Mind: The Irreconcilability of Machine Learning and the Right to be Forgotten* (Mar. 6, 2024) at 2, https://arxiv.org/ abs/2403.05592.

doi:10.6342/NTU202600641

第一項　資訊隱私與安全風險

當金融機構致力於透過人工智慧技術提供更個人化的金融服務時，其所運行的 AI 系統勢必仰賴大量包含個人與財務資訊的資料。此一高度資料依賴的特性，使得資料隱私與資安風險成為核心議題[66]。倘若相關資料未獲妥善保護，不僅可能導致資訊外洩或遭受駭客入侵，進而引發財務損失與聲譽風險，亦可能因刪除請求無法落實而引發後續法遵爭議[67]。

此外，AI 模型對資料品質與完整性的依賴，使其暴露於特有的安全威脅之下。例如，對手可能蓄意篡改或「毒化」訓練資料，或利用模型架構的脆弱性進行操控，導致偏誤輸出或不當決策[68]。相較於傳統 IT 系統，AI 系統更容易因資料完整性遭破壞而失效，進一步放大刪除權無法實踐所帶來的隱私侵害與風險外溢效應[69]。

因此，金融機構有必要全面檢視 AI 系統在訓練與部署階段對個人資料的依賴程度，並探索可行的資料最小化與可行的刪除技術，以降低刪除權無法實踐所帶來的隱私與合規風險。

第二項　可解釋性與透明度問題

許多人工智慧與機器學習模型，特別是深度神經網路等高度複雜的架構，在運作過程中具備「黑箱」特性，雖能產出具體預測或決策，卻難以清楚解釋其推理過程與結果生成邏輯[70]。在金融應用情境下，這種不可解釋性問題尤為嚴重。當模型包含多層隱藏層並產生非線性決策時，不僅銀行內部人員，甚至主管機關亦難以理

---

66 See generally Ghazi Zouari & Marwa Abdelhedi, *Customer satisfaction in the digital era: evidence from Islamic banking*, 10(9) J INNOV ENTREP, (2021).
67 Sofiat O. Abioye et al., *Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges*, 44 J. BUILDING ENG'G, 10 (2021).
68 U.S. Department of the Treasury, *Managing Artificial Intelligence-Specific Cybersecurity Risks in the Financial Services Sector*, 2 (2024).
69 *Id.*
70 TOM BIGHAM ET AL, AI AND RISK MANAGEMENT: INNOVATING WITH CONFIDENCE, DELOITTE 8 (2024) 5, https://www.deloitte.com/content/dam/assets-shared/legacy/docs/perspectives/2022/deloitte-gx-ai-and-risk-management.pdf.

doi:10.6342/NTU202600641

解其做出特定信用評估、交易判斷或風險分類的依據，進而削弱消費者信任並增加監管挑戰[71]。

若 AI 模型進一步採用持續學習（continuous learning）機制，其參數會隨新資料不斷更新，使得原始資料與模型輸出之間的關聯更加複雜難解[72]。此特性對資料刪除權的落實產生深遠影響：即使資料主體依法請求刪除個人資料，資料控制者亦難以確認該資料對模型參數的影響是否能同步移除，導致刪除效果落空，並衍生潛在法律風險[73]。

根據 GDPR 第 22 條，資料主體有權就純自動化決策獲得充分資訊與有意義的說明[74]。惟若 AI 模型本身缺乏可解釋性，企業將難以履行此一法定義務，亦無從追溯個人資料對決策過程的影響，進一步使刪除權的實踐遭遇障礙。

綜上所述，AI 模型的不可解釋性不僅構成合規與監管風險，更在回應資料刪除請求時揭露出「技術無法追溯、法律無從實施」的落差，亟須透過技術補強與制度設計予以回應。

第三項　其他 AI 應用風險——偏誤、詐欺濫用與系統性風險

人工智慧技術雖在金融領域廣泛應用，但也伴隨著偏誤與公平性、詐欺犯罪及系統性市場風險等問題。首先，AI 模型可能無意間延續訓練資料的既有偏見，進而在信用評估、保險定價等關鍵金融決策中產生系統性的不公平甚至歧視問題，使金融機構面臨法律責任與聲譽損害的風險[75]。

---

[71] 例如，若 AI 演算法拒絕了一位客戶的貸款申請，銀行可能無法提供清晰的理由，這不僅會影響客戶服務，還可能違反要求提供決策理由的監管規定。*Id*.

[72] *Id.*

[73] *See* Meem A.M., *supra* note 65.

[74] Regulation (EU) 2016/679, General Data Protection Regulation (GDPR), art. 22(3): ”In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject’s rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.”

[75] Ana Fernández, *Artificial Intelligence in Financial Services*, 2 ECON. BULL., 4, 5 (2019).

doi:10.6342/NTU202600641

其次，AI 技術在金融犯罪防治方面如一把雙面刃，既能強化詐欺偵測能力，也可能遭犯罪集團濫用[76]。特別是生成式 AI 能輕易製造逼真的深偽內容，以欺騙銀行系統與消費者，顯著降低金融詐欺的成本並提升其危害性[77]，迫使金融機構必須持續更新與強化反詐機制，以應對不斷演變的犯罪威脅[78]。

最後，AI 的使用亦可能產生系統性市場風險，尤其是當金融機構高度集中仰賴少數 AI 技術供應商[79]，或大量機構使用設計相似的模型，易造成羊群效應（herding behavior）[80]，放大市場波動並危及整體金融穩定性。因此，金融機構在部署 AI 時，不僅應留意個資處理與刪除權保障議題，更須從更全面的市場風險觀點進行審慎評估及管理，避免意外成為市場危機的潛在誘因。

---

[76] *Generative AI is Expected to Magnify the Risk of Deepfakes and Other Fraud in Banking*, DELOITTE https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html (May 29, 2024).; Nabila Ahmed et al., *Deepfake Imposter Scams Are Driving a New Wave of Fraud*, BLOOMBERG (Aug. 21, 2023) https://www.bloomberg.com/news/articles/2023-08-21/money-scams-deepfakes-ai-will-drive-10-trillion-in-financial-fraud-and-crime.

[77] *Id.*

[78] 根據身份驗證平台 Sumsub 最近的一份報告顯示，2023 年金融科技領域的深偽攻擊增加了 700%，顯示這類威脅已進入爆發期。這波 AI 驅動的詐欺浪潮意味著金融機構將面臨更高的財務損失與前所未有的威脅，而現有的反詐欺與反洗錢（AML）控制措施並未針對這類攻擊設計，導致防禦能力不足。*See* Isabelle Bousquette, *Deepfakes Are Coming for the Financial Sector*, WALL STREET JOURNAL, (Apr. 3, 2024, 7:00 AM), https://www.wsj.com/articles/deepfakes-are-coming-for-the-financial-sector-0c72d1e5; U.S. Department of the Treasury, *supra* note 68.

[79] 詳言之，如果許多銀行的 AI 決策系統都由少數科技公司開發或託管，那麼這些供應商的技術故障、服務中斷，甚至惡意攻擊，都可能對整個金融體系造成連鎖衝擊。例如，如果一個主導市場的 AI 信用評分系統發生錯誤，導致貸款風險被錯誤分類，可能會同時影響多家銀行的放貸決策，引發資金流動問題。這種集中風險不僅限於傳統銀行，還可能讓 AI 供應商本身變成「大而不能倒」（too big to fail）的新興金融系統關鍵點。*See* Georg Leitner et al., *The Rise of Artificial Intelligence: Benefits and Risks for Financial Stability*, FIN STABILITY REV. 104, 105(2024); Tejas N. Narechania, and Ganesh Sitaraman, *An Antimonopoly Approach to Governing Artificial Intelligence*, 43 YALE LAW REVIEW 95, 101(2024), and Ben S. Bernanke, *Causes of the Recent Financial and Economic Crisis*, TESTIMONY, FEDERAL RESERVE (2010), who defines the too-big-to-fail externality like this: "A too-big-to-fail firm is one whose size, complexity, interconnectedness, and critical functions are such that, should the firm go unexpectedly into liquidation, the rest of the financial system and the economy would face severe adverse consequences."

[80] 由於許多金融機構採用相似的機器學習模型，而這些模型通常基於相同的市場資料訓練，因此它們可能對某些市場訊號產生一致性反應，進一步放大市場波動。例如倘若大量交易演算法都設計為在某個指標達到特定閾值時自動拋售特定資產，那麼一次原本微小的市場波動可能迅速升級為更大規模的市場崩盤（market crash）。就此金融穩定理事會（FSB）已警告，若不加以妥善管理，AI 可能會加劇市場同質化（market correlation），放大風險傳導。*See* Leitner et al., *id*. at 114.

doi:10.6342/NTU202600641

第五節　我國人工智慧相關規範與指引

隨著 ChatGPT 在 2022 年底問世，全球掀起了一波生成式人工智慧浪潮，相關科技股成為市場的追捧焦點，人工智慧技術也取得了顯著突破。然而， AI 雖在各領域被廣泛應用，卻如同雙面刃般，同時伴隨著幻覺錯誤、智慧財產權爭議、個資外洩、偏見與歧視等潛在風險[81]。鑑於技術快速演進所衍生之挑戰，全球主要國家紛紛致力於在不妨礙創新發展的前提下，建立 AI 治理框架與價值原則，以兼顧風險控管與產業發展。

OECD 於 2019 年通過《人工智慧建議書》，提出基本價值原則與政策建議。其後，歐盟於 2021 年提出《人工智慧法》（Artificial Intelligence Act），並於 2024 年完成審議[82]；美國於 2022 年發布《AI 權利法案藍圖》（Blueprint for an AI Bill of Rights）[83]，白宮亦於 2023 年進一步發布《發展與使用安全且可信任的 AI 行政命令》（Executive Order on the Safe, Secure, and Trustworthy Development and Use of AI）[84]，指示聯邦各部門推動 AI 策略;加拿大則於 2022 年提出《人工智慧資料法草案》（Artificial Intelligence and Data Act）[85]，皆著重於建立信任為本的 AI 發展準則。

我國作為全球資訊與通訊科技（Information and Communication Technology, ICT）供應鏈的重要環節，亦於 2026 年公布施行《人工智慧基本法》（下稱 AI 基本法）[86]，金管會並於 2024 年公布《金融業運用人工智慧（AI）指引》，希望引導

---

[81] 遠見雜誌（01/03/2025），〈AI 法案之父：法案漏洞多，監管仍待完善〉，https://www.ctee.com.tw/news/20250103700988-431003（最後瀏覽日：04/09/2025）

[82] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on artificial intelligence (Artificial Intelligence Act), OJ L 2024/1689.

[83] White House Office of Science and Technology Policy, *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People* (2022), https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/.

[84] Exec. Order No. 14,110, 88 Fed. Reg. 75191 (Nov. 1, 2023).

[85] Artificial Intelligence and Data Act, Part 3 of Bill C-27, 2d Sess, 44th Parl, 2022 (Can.) (not yet enacted).

[86] 參全國法規資料庫（2026），《人工智慧基本法》，https://law.moj.gov.tw/News/NewsDetail.aspx?msgid=196197 （最後瀏覽日：2026 年 2 月 2 日）。

doi:10.6342/NTU202600641

金融業在兼顧消費者權益、維護市場秩序與履行社會責任的基礎上，積極擁抱科技創新，推動金融服務升級。

上述兩項規範均針對 AI 於金融應用中可能引發之特有風險有所著墨，涵蓋前節所述之公平性、系統安全性與可解釋性等面向。其中，個人資料保護與隱私更是備受關注的核心議題[87]，本文以下即擬從此角度切入，簡要介紹該兩部法案中與個人資料保護相關之規範內容。

第一項　人工智慧基本法

人工智慧基本法為我國首部全面性的人工智慧專法，其係以「基本法」為名。按基本法具有特殊的法律屬性，其立法目的通常在於推動重大政策或制度議題的長期發展，惟因其所規範之對象涵蓋廣泛，且涉及多元權責機關，實務推動上需仰賴分階段、分層次的政策配合與橫向整合，再者，AI 技術本身仍處於快速發展階段，相關法制體系尚未完全成形，難以一蹴可幾建構具體細緻的規範內容[88]。是以，基本法常見之立法型態為以抽象原則或宏觀目標作為政策宣示，提供施政依循之法理架構，未必具備即時且具體之法律效力。換言之，基本法往往僅提供政策實踐之「骨架」，而非含有具體操作細則的「血肉」，惟此正係其作為「框架型立法」之「綱要性」與「指導性」特質所致[89]。

我國 AI 基本法係屬基本法性質，其立法定位與歐盟之《人工智慧法案》不同，後者係以高風險 AI 為核心建立具有直接約束力並附帶高額罰則之具體規範，我國則採原則導向之抽象設計，其內容不具有立即拘束效力，而係提供各主管機關日後

---

[87] 國內外已有諸多文獻針對人工智慧應用可能引發的個人資料保護問題進行探討。參見程法彰（2025），〈以歐盟為師談人工智慧商業發展運用對我國個人資料保護之互動〉，《月旦法學雜誌》，第 361 期，頁 56-70。*See* Seraphina Brightwood & Henry Jame, *Data Privacy, Security, and Ethical Considerations in AI-Powered Finance*, RESEARCH GATE 1,2 (2024), https://www.researchgate.net/publication/379078709_Data_privacy_security_and_ethical_considerations_in_AI-powered_finance.; Caitlin C.R., Protecting Data Privacy: A Baseline for Responsible AI, CSIS, (Jul. 18, 2024), https://www.csis.org/analysis/ protecting-data-privacy-baseline-responsible-ai .

[88] 陳家駿（2024），〈以人為本之 AI 及其透明性與可解釋性──我國人工智慧基本法草之省思與建議〉，《教育暨資訊科技法學評論》，13 期，頁 89。

[89] 同前註。

doi:10.6342/NTU202600641

制定與修正個別法律或行政命令時所應遵循之方向與準則。是故，人民尚難依本法主張具體權利或救濟，此由其第 17 條第 1 項之規定即不難觀察出此一定位：「政府應於本法施行後依本法規定檢討及調整所主管之職掌、業務及法規，以落實本法之目的[90]。」

AI 基本法內容涵蓋廣泛，包含 AI 定義、風險分級治理架構、產業促進政策、AI 實驗沙盒機制、防詐打假、以及與個資保護相關之規範。其中第 14 條規定：「個人資料保護主管機關應協助各目的事業主管機關，在人工智慧研發及應用過程，避免不必要之個人資料蒐集、處理或利用，並應促進個人資料保護納入預設及設計之相關措施或機制，以維護當事人權益。」第 15 條第 1 項則進一步指出：「政府應建立資料開放、共享與再利用機制，提升人工智慧使用資料之可利用性，並定期檢視與調整相關法令及規範。」

綜觀上述二條，可見立法者試圖在 AI 應用中，於保障個資與促進資料流通之間取得平衡：一方面導入「資料最小化原則」以避免不必要的個人資料處理；另一方面則鼓勵在符合法令條件下推動非敏感資料的大數據開放與再利用，以促進 AI 技術於多元領域之創新應用[91]。然而，現行草案中尚未明確納入與「刪除訓練資料」相關之規定，是以於目前法制體系下，倘有民眾欲針對 AI 訓練資料主張個資刪除權，仍須回歸《個人資料保護法》之適用，依現行個資法之規範加以處理。

第二項　金融業運用人工智慧（AI）指引

鑒於 AI 在金融市場應用日益普及，雖可提升效率與風險管理、強化資安與合規，惟若缺乏審慎規劃與動態調整，反而可能導致損失、增加風險，甚至動搖市場信心。為協助金融業妥適導入與管理 AI，我國金管會參考國際清算銀行(Bank for International Settlements, BIS)、國際證券管理機構組織(International Organization of

[90] 同前註，頁 90。
[91] 同前註，頁 93。

doi:10.6342/NTU202600641

Securities Commissions, IOSCO)、歐盟、新加坡、美國等，發布具行政指導性質之《金融業運用人工智慧（AI）指引》，鼓勵業者在風險可控下因應自身情境採納，並強調應整體性評估各項核心原則，選擇合宜且具成本效益之機制，以達有效控管AI風險之目的[92]。

該指引主要欲實踐之內容包括：建立治理與問責機制、重視公平性與以人為本之價值觀、保護隱私及客戶權益、確保系統穩健性與安全性、落實透明性與可解釋性及促進永續發展六項核心原則，其中與個人資料保護有關之一般性內容如下：

一、 建議金融業應優先將資源投入於風險較高之AI系統，以強化風險管理成效。其中，「是否使用個人資料」可作為重要評估指標之一；更具體而言，AI系統若涉及大量個人原始資料[93]或敏感個資，通常伴隨較高風險，應列為資源配置與監理重點[94]。

二、 金融機構如委託第三方業者導入AI系統，並涉及將客戶資料交由該等業者處理時，應與其簽訂包含資料保護條款之協議。該協議宜明確規範資料之加密傳輸、存儲安全措施，以及服務終止後資料之刪除或返還方式，以確保客戶資料於委外期間及作業終結後均獲妥善處理[95]。

三、 金融機構在運用AI系統時，應嚴謹保護客戶隱私，並落實資料蒐集與處理之正當性與必要性。宜依據「資料最小化原則」僅蒐集與特定目的相符之必要資訊，避免收集過量或與服務無直接關聯之敏感個資，降低潛在外洩風險[96]。

四、 當金融機構以AI系統與客戶互動時，應主動揭露系統功能與其在決策過程中所扮演之角色，並說明可能對客戶權益造成之影響。同時，應尊重客戶是

[92] 金融監督管理委員會（2024），前揭註17，頁1。
[93] 個人原始資料係指未經去識別化、隱私強化技術處理或其他方式處理之個人資料。參前註，頁4。
[94] 同前註，頁4。
[95] 同前註，頁5。
[96] 同前註，頁13。

doi:10.6342/NTU202600641

否使用 AI 服務之選擇權，提供清楚可辨之替代方案資訊，保障其知情權與自主決定權。另於系統運作期間，金融機構亦應持續保障客戶隱私，涵蓋個人資料、智慧財產權與營業秘密等資訊，確保其合法權益不受侵害[97]。

除上述關於保護客戶個人資料的一般性規定外，該指引亦對客戶提出退出使用 AI 系統服務之情境提出建議。金融機構在處理客戶欲退出使用 AI 系統服務的情形時，可依據六項因素（對金融機構與客戶的風險、替代方案的可行性與成本、技術限制及作業複雜度等）綜合評估是否提供替代方案，若最終不提供替代作法，仍應考慮是否有必要提供其他補救措施，以保障客戶權益[98]。

綜上所述，《金融業運用人工智慧指引》在金融消費者個資保護方面已有相對具體之規範，並涵蓋客戶退出 AI 系統之程序性設計，展現對技術透明與自主選擇權的關注。然而，相較於「退出使用 AI 系統」屬於服務層級的選擇權益，目前指引對於客戶請求刪除已被用於 AI 模型訓練之個人資料的情境，尚未建立明確標準或作業機制。此一情境反映出，在面對 AI 模型日益具有記憶性與資料殘留風險的情境下，仍有必要補足金融機構對「資料刪除請求」之應對規範，並強化與《個資法》及國際資料保護趨勢之接軌。

第六節　小結

本章為 AI 與其於金融業具體應用的概論，回顧 AI 技術的發展脈絡，從符號邏輯與專家系統的初步探索，到機器學習與深度學習的興起，說明 AI 如何逐步成為當今科技與產業轉型的關鍵驅動力，進而介紹 AI 運作機制的基本技術原理以及其在金融領域的應用現況，包括智慧客服、信用風險評估、法令遵循、反詐欺偵測及理財機器人等。

[97] 同前註，頁 15。
[98] 同前註。

doi:10.6342/NTU202600641

然而，AI 在金融場域的應用不僅帶來創新與效率，也伴隨多重潛在風險，如對訓練資料的高度依賴所衍生之隱私與資安問題、模型黑箱化造成的可解釋性與透明度不足、以及偏誤輸出、詐欺濫用與系統性風險等實質挑戰。此類風險不僅可能削弱消費者信任與金融穩定，更對現行法制提出嚴峻考驗。就此我國目前相應的法制發展，主要包括已公布施行的人工智慧基本法與金管會頒布的金融業運用人工智慧（AI）指引。

特別值得關注者，是 AI 模型「資料難以刪除」的特性，對既有的個人資料刪除請求制度帶來顯著衝擊。當模型於訓練階段內隱記憶個資後，金融機構若欲實現刪除請求，將面臨技術與合規上雙重困難。此一問題正成為監管機關與學界關注的焦點，未來亦可能重塑金融業者對資料治理與權利保護的整體策略。

因此，下一章將進一步探討人工智慧技術與資料刪除權之間的衝突與調和，從刪除權的法理基礎與現行法制出發，剖析機器學習中刪除資料的實務困難，進而評估金融機構在 AI 部署下所面臨的合規挑戰與法律責任，期能更全面掌握 AI 技術於資料治理與隱私保護上之法律與倫理意涵。

doi:10.6342/NTU202600641

第三章 人工智慧技術與資料刪除的扞格

隨著人工智慧技術迅速發展，尤其是機器學習的廣泛應用，個人資料的蒐集與利用規模亦大幅擴張，進而帶動個資保護議題的重視。我國憲法法庭於 111 年憲判字第 13 號解釋中指出，當事人原則上應享有對其個人資料的事後控制權，即便該等資料係經當事人同意或符合法定要件而被蒐集、處理及利用，亦不因此當然喪失請求刪除、停止或限制利用之權利[99]。究其意旨，該解釋係針對健保資料蒐集制度而發，強調的是原有規範完全欠缺事後控制機制所導致的違憲狀態，並未明示或暗示個資刪除權應達至無限制程度，僅揭示未來個資制度應致力於在「資料主體的資訊自主權」與「資料使用者的合法需求」之間建立妥適的衡平。

在 AI 金融應用的監理下，法制設計的核心挑戰在於如何妥適地在保障個人資訊自主權、確保人工智慧技術發展，以及維持金融應用穩定性之間取得平衡。倘若個人資料刪除權過於絕對，可能會干擾 AI 機器學習所需的資料穩定性，特別是在高度依賴預測準確性的金融應用中，大量資料的移除更可能損及金融機構的風險控管能力，進而影響信用決策的公平性與整體金融體系的信賴性。反之，若刪除權僅具形式性質，則無法落實資料主體對自身資訊的實質控制，甚至可能違背憲法對資訊自主權的保障。

具體而言，機器學習系統高度依賴大量歷史資料進行反覆訓練，以萃取模式與規則並優化系統效能。若大量關鍵資料因刪除權的行使而被移除，可能造成模型規則結構的破壞，使系統準確性下降，甚至出現崩解現象[100]。例如金融機構與信用評等機構如上述經常藉助 AI 技術分析大量複雜之資料，以建立具高準確性的信用評

[99] 111 年憲判字第 13 號判決。

[100] 李沛宸（2019），〈實施歐盟個人資料保護規章對人工智慧發展之影響〉，《財金法學研究》，第 2 卷第 1 期，頁 148、149。

doi:10.6342/NTU202600641

估模型。倘若特定類型或特徵的資料主體集中行使刪除權，則可能導致模型效能受損，進而造成信用評估上的偏差與不公平[101]。

有鑑於此，本章擬釐清個人資料刪除權與人工智慧技術間的內在衝突內涵，首先分析刪除權的憲法基礎與現行法制架構，並探討機器學習模型使用資料時可能面臨的技術挑戰，並藉由此一分析為後續的經濟分析與比較法討論奠定基礎，提出兼顧法律需求與技術現實的平衡方案。

第一節　個人資料刪除權之發展與其內涵

第一項　前言

在歷史長河中，紀錄與遺忘之間始終保持著一種微妙的平衡，而科技的發展則持續影響並重新塑造這種關係[102]，導致個人對資訊流通的掌控能力顯著降低[103]。網際網路的普及大幅的提升了資訊的傳播速度與範圍，而技術的進步更賦予應用程式與網站系統在使用者知情與否的情況下蒐集個人資料的能力[104]，現在 AI 透過巨大的資料資料庫自主進行深度學習，使得人們更難得知自己的個人資料係在何時、何處以如何的方式被運用於 AI 的訓練，甚至可能引發更進一步的個資外洩。

誠如學者所言：「在非數位時代，資訊隨時間而自然消逝；然在今日資訊可永久儲存的社會中，遺忘不再是自發性的，而必須透過有意識且積極的管理機制方得實現。」[105]

---

[101] 同前註。

[102] V. MAYER-SCHÖNBERGER, DELETE: THE VIRTUE OF FORGETTING IN THE DIGITAL AGE, rev. ed. (2011).

[103] *Id.*, at 90.

[104] Bruno Zeller et al., *The Right to Be Forgotten—The EU and Asia Pacific Experience (Australia, Indonesia, Japan and Singapore)*, 2019(1) EUROPEAN HUMAN RIGHTS LAW REVIEW 23 (2019), https://dx.doi.org/10.2139/ssrn.3320860 (last visited June 23, 2025).

[105] Theodor S. et al., *SoK: Managing Longitudinal Privacy of Publicly Shared Personal Online Data*, 2021 PROCEEDINGS ON PRIVACY ENHANCING TECHNOLOGIES. 229 (2021), https://doi.org/10.2478/popets-2021-0013.

doi:10.6342/NTU202600641

面對數位足跡幾乎無所遁形的現實，資料主體愈加意識到其日常生活中的各項活動皆可能遭到蒐集、儲存與再利用，甚至於無限期保存。因此，各國立法者紛紛致力於透過個人資料保護立法，強化資料主體對其個資之控制權限。自歐盟法院於 2014 年 Google Spain SL v. AEPD [106]一案中正式承認「被遺忘權」（the right to be forgotten）以來，關於個人資料刪除與資料控制的權利議題遂成為全球隱私保護改革之關鍵焦點。隨後，歐盟、加州與日本等地亦相繼於其法制中明文納入「刪除權」（the right to erasure）與被遺忘權作為資料主體之法定權利[107]。

我國雖非歐盟成員，但亦早已於《個人資料保護法》第 11 條第 3、4 項與第 19 條第 2 項中明文賦予資料主體請求資料控制者刪除個資之權利。然而刪除權的基本權基礎究竟為何，涉及該權利之正當性與界限。就此而言，我國個人資料保護制度在憲法層次上，長期存在「資訊隱私」與「資訊自決」兩種不同定位的爭議，此爭議不僅關係到個資保護的制度目的，更直接影響刪除權保障範圍的認定與適用密度。

此外，我國實務判決儘管偶有援引歐盟《一般資料保護規則》（GDPR）第 17 條之刪除權作為參考，卻多採限縮解釋，並未明確闡述該權利之憲法正當性基礎，致使個資保護範圍與刪除請求的可及性仍存在不確定空間[108]。是以，本文擬進一步探究資訊自決權與資訊隱私權之概念區辨，作為論述刪除權保障邊界的理論起點，

---

[106] Google Spain SL v. Agencia Española de Protección de Datos (AEPD), Case C-131/12, ECLI:EU:C:2014:317, Judgment of May 13, 2014.

[107] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), art. 17, 2016 O.J. (L 119) 1 [hereinafter GDPR]; California Consumer Privacy Act of 2018, Cal. Civ. Code § 1798.105 (West 2023) (as amended by the California Privacy Rights Act of 2020) [hereinafter CCPA]; Act on the Protection of Personal Information [APPI], Act No. 57 of 2003, art. 29, para. 1 (Japan).

[108] 例如臺灣高等法院 110 年度上更一字第 47 號民事判決中，雖肯認原告得依我國《個人資料保護法》第 11 條及《民法》第 18 條主張隱私權保護，並參酌 GDPR 第 17 條所揭示之價值，然實際判斷時仍以「原告是否有合理期待資料被移除」為核心，強調資訊揭露之公益性與言論自由之優先地位，未明確建立刪除權的憲法基礎與判準。又如最高法院 109 年台上字第 489 號判決亦指出，雖可參酌歐盟對個資保護的法制發展，然刪除請求的成立仍須回歸我國現行法與個案情狀加以評估。

doi:10.6342/NTU202600641

俾使後續針對資料控制者義務、技術應用風險及制度設計方向之分析能有所依據，亦有助於建構一套與新興科技發展相適應的資料主體權利體系。

第二項　資訊自決權與資訊隱私權之法制繼受與概念內涵辨析

在探討個資刪除權之內涵之前，首先必須釐清「資訊自決權」（informationelle Selbstbestimmung）與「資訊隱私權」（information privacy）兩種憲法上基本權利概念的區別與連結[109]。此一議題之重要性，除了理論釐清外，更涉及到我國司法院大法官實務運作學理上對個人資料保護的基本權定位之長期爭議。本文即擬透過對二者法制繼受、概念內涵以及保護範圍之辨析，作為釐清刪除權憲法定位的理論基礎。

第一款　不同的法制繼受來源與基本權概念混用

我國法制對於個人資料保護制度的基本目的，長期以來即存在「資訊隱私」與「資訊自決」兩種制度定位的爭議，而此一概念定位之差異，實際上也深刻影響了對「被遺忘權」內涵與適用範圍的理解。

追根究柢，這兩組概念之差異實係源自不同法系的繼受背景[110]。若檢視我國學術界對此議題的相關研究，可發現主張「資訊隱私」作為制度基礎者，多半係受英美法系影響者所倡議[111]；反之，強調「資訊自決」概念者，則多來自德語法系之學習傳統[112]。相應地，當我國學者論及刪除權（被遺忘權）之憲法依據時，若以英文

---

[109] 張志偉（2017），〈從資訊自決與資訊隱私的概念區分，檢視被遺忘權的證立問題〉，《萬國法律》，第 211 期，頁 3；邱文聰（2009），〈從資訊自決與資訊隱私的概念區分──評「電腦處理個人資料保護法修正草案」的結構性問題〉，《月旦法學雜誌》，第 168 期，頁 174- 177。

[110] 邱文聰（2009），同前註，頁 174；劉靑峰（2023），〈COVID-19 疫情下資訊自決權之研究－以歐洲人權公約第 8 條作為比較法對象〉，《中原財經法學》，第 50 期，頁 230。

[111] 參照：廖福特、翁逸泓（2008），〈兩難？共存？-國家處理個人資料與資訊隱私權保障之糾葛〉，《二十一世紀公法學的新課題-城仲模教授古稀祝壽論文集》，2008 年 10 月，頁 255 以下，新學林。

[112] 參照：李震山（2020），〈論資訊自決權〉，《人性尊嚴與人權保障》，頁 239 以下，元照；李震山（2001），〈論個人資料之保護〉，《行政法爭議問題研究（上）》，頁 657-664，五南。

doi:10.6342/NTU202600641

文獻為主要參考資源，則多以資訊隱私為出發[113]；若以德語文獻為依歸者，則傾向以資訊自決作為其制度基礎[114]。

惟亦有學者主張，「資訊自決權」與「資訊隱私權」實質上僅為名稱不同，其核心內涵並無二致，兩者可互相通用[115]。此一見解或可從司法院釋字第 585 號解釋中獲得佐證：司法院大法官會議所承認之「資訊隱私權」，實際上即是對個人資料的「自主控制權」，其保護對象不僅限於私密性資訊，而是以資料主體對資訊之自主決定為核心內涵[116]。

進一步觀察釋字第 603 號解釋及憲法法庭 111 年憲判字第 13 號判決，可發現司法院大法官會議或憲法法庭均強調「資訊隱私權」為「個人對其個人資料之揭露與使用擁有決定權」，包括：「是否揭露、揭露之時間、方式、範圍及對象；同時也包含對資訊之知悉、控制與更正錯誤之權利」。故其認為資訊隱私權實質上保障資料主體在資料使用前具備「事前控制權」，在使用中與使用後則擁有「事後控制權」，即請求刪除、停止使用或限制處理等權利。

綜上，本文認為我國大法官會議或憲法法庭對隱私權之闡釋，實已兼用德國法上之「資訊自決權」與美國法上的「資訊隱私權」兩者概念，作為憲法第 22 條保障之具體內容與保護範圍。而我國實務判決在資訊隱私權與資訊自決權的概念上亦多受釋字 603 號解釋的影響，並未進一步釐清兩者之差異與界線[117]。

---

113 參照：劉靜怡（2012），〈社群網路時代的隱私困境：以 Facebook 為討論對象〉，《臺大法學論叢》，第 41 卷第 1 期，頁 42 以下； 許炳華（2015），〈被遺忘的權利：比較法之觀察〉，《東吳法律學報》，第 27 卷第 1 期，頁 127、141 以下，； 徐彪豪（2015），〈被遺忘權近期發展——歐盟法院判決週年後回顧與本土觀察〉，《科技法律透析》，第 27 卷第 11 期，頁 50 以下； 楊智傑（2015），〈個人資料保護法制上「被遺忘權利」與「個人反對權」：從 2014 年西班牙「Google v. AEPD」案判決出發〉，《國會月刊》，第 43 卷第 7 期，頁 20 以下。

114 參照：張志偉（2017），〈記憶或遺忘，抑或相忘於網路-從歐洲法院被遺忘權判決，檢視資訊時代下的個人資料保護〉，《政大法學評論》，第 148 期，頁 32 以下。

115 黃昭元（2005），〈無指紋則無身分證?──換發國民身分證與強制全民捺指紋的憲法爭議分析〉，《民主、人權、正義──蘇俊雄教授七秩華誕祝壽論文集》，頁 470，元照。

116 同前註。

117 張志偉（2017），前揭註 109，頁 3；范姜真媺（2013），《個人資料保護法關於「個人資料」保護範圍之檢討》，東海大學法學研究，第 41 期，頁 105；劉靜怡（2010），《不算進步的立法：「個人資料保護法」初步評析》，月旦法學雜誌，第 183 期，頁 152。

doi:10.6342/NTU202600641

第二款　資訊自決權與資訊隱私權下的資訊範圍

雖然我國大法官會議與實務判決在論述資訊隱私權時，大多逕將資訊自決權之內涵實質涵蓋在內，然而在我國《個人資料保護法》的制度脈絡中，關於「資訊自決權」與「資訊隱私權」的概念區分，一直存在不同見解。有論者即指出兩者雖有部分交集，然其核心意涵與保障重點並不相同：前者著重於個人對其資訊之自主控制，保障個人對外在行動之自由；後者則聚焦於維護個人內在人格發展之彈性空間，關注資訊對人格形成與展現的影響[118]。其進一步主張，資訊自決權作為一種行為自由，其保障範圍不宜被過度擴張，若將資訊隱私等同於資訊自決，恐不當地將所有個資保護議題提升至人格保障的高度[119]。

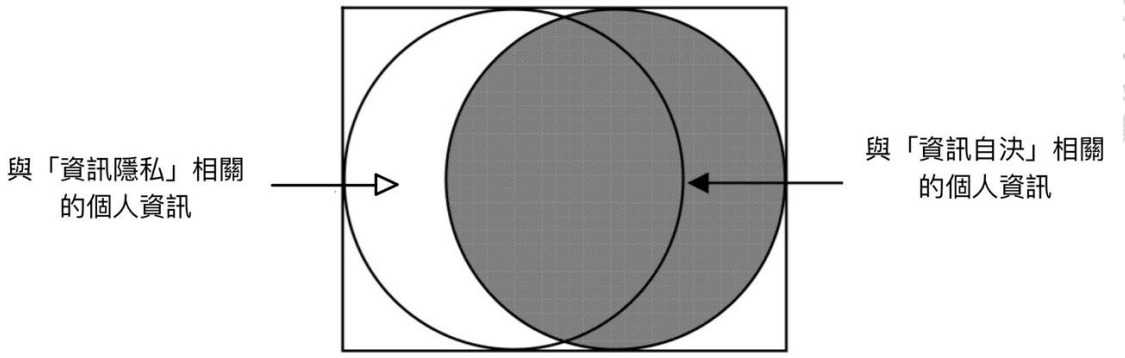
據此，有論者將個人資訊區分為兩類不同屬性：其一為與資訊自決相關之資訊，其二則為與資訊隱私相關之資訊。雖然兩者間存在一定程度之重疊，但在適用範圍上仍各具界限，亦即，部分資訊可能僅受資訊隱私權之保護，卻未納入資訊自決權的適用範圍[120]。基於此種區辨，邱文聰教授即以【圖 二】清楚呈現資訊隱私權與資訊自決權在個人資料範疇中的邏輯關係。

在相同脈絡下，另有學者進一步指出，個人資料保護法制的基本結構應由兩個核心面向所構成：一為保障資料主體「資訊自主決定權」的規範，另一則為對資料蒐集、處理與利用者課予「資訊隱私保護義務」的規定[121]。

---

[118] 邱文聰（2009），前揭註 109，頁 174 以下。
[119] 同前註。
[120] 同前註，頁 177，圖示一。
[121] 劉靜怡，前揭註 117，頁 151-152。

doi:10.6342/NTU202600641



【圖 二】邱文聰教授所論資訊隱私權與資訊自決權在個人資料範疇之邏輯關係

圖片來源：本文自行繪製

然而，張志偉教授對於上述將「資訊自決」與「資訊隱私」視為僅部分重疊之兩個資訊範疇的見解，認為仍有進一步斟酌的空間。具體而言，就資訊的涵蓋範圍而言，資訊自決權所保護的個人資訊並無實質限制，只要該資訊「涉及個人」，即屬其適用對象，無論其是否具備私密性[122]。相對地，資訊隱私權的保護對象則須具備一定程度的私密性，或至少存在合理的隱私期待，方得主張憲法保障[123]。

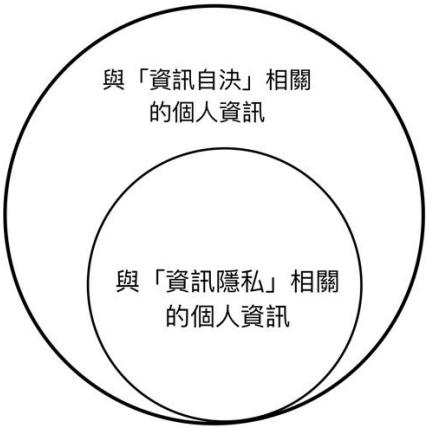
因此，從邏輯結構觀之，資訊隱私所涵蓋的資訊應屬於資訊自決權保護範圍之下，二者關係較宜理解為「包含關係」，而非交叉或重疊，如【圖三】所示。張教授進一步指出，若欲論述二者之交集，應係發生於「資訊自決權」與「隱私權」兩種權利概念之間，而非發生於「資訊自決」與「資訊隱私」這兩個資訊類型之間[124]。

[122] Vgl. s 3 Abs. 1 BDSG; auch Marion Albers, Datenschutzrecht, in: Dirk Ehlers/ Michael FehlinglHermann Pünder (Hrsg.) , Besonderes Verwaltungsrecht, Bd. 2, 3 Aufl. 2013, s 62 Rn. 8, 14. 同見解，參見范姜真媺，前揭註 117，頁 105 。

[123] 劉靜怡，前揭註 113，頁 47。

[124] 同前註；同見解，參范姜真媺，前揭註 117，頁 107。

doi:10.6342/NTU202600641



【圖 三】張志偉教授所論資訊隱私權與資訊自決權在個人資料範疇之邏輯關係

圖片來源：本文自行繪製

本文認為，張教授所提出「資訊隱私包含於資訊自決」之觀點，較能貼合我國《個人資料保護法》之立法邏輯與實務運作現況。蓋該法並未依個資性質區分不同基本權保障途徑，而是統一採取「個人資訊自主控制」為核心理念，建構資料主體得對所有「涉及個人」之資訊行使知情、刪除、限制處理等權利。因此，無論該資訊是否具有私密性，或是否達到「隱私權」所要求之期待門檻，資料主體均得依其意願決定是否提供、如何使用或請求刪除。若以資訊自決權作為個資保護體系之基礎，不僅有助於消弭資訊隱私權保障門檻過高所造成之保護落差，亦能更全面回應人工智慧等新興科技下「去隱私化但仍具個人指涉性」之資料風險。

第三款　資訊自決權與資訊隱私權之內涵區辨

就資訊自決權與資訊隱私權之內涵相比，前者所強調者在於無論國家或私人，若未經當事人之同意，即不得任意蒐集其個人資料，更不得擅自將所蒐集的資料進一步與其他資料相互連結、加以整合或任意利用，否則即對資料主體之人格權與隱私權構成侵害[125]。若容許國家或私人恣意蒐集、處理或利用當事人之個人資料，實

[125] 黃昭元，前揭註 115，頁 471-472；李惠宗（2022），《憲法要義》，第 9 版，元照，頁 1971。李寧修（2022），〈基於防疫目的之預防性個人資料運用：以實聯制為例〉，《公法研究》，創刊號，頁 116；陳銘雄、劉庭妤（2011），〈從「個人資料保護法」看病患資訊自主權與資訊隱私權之保護〉，《月旦民商法雜誌》，第 34 期，頁 28。

doi:10.6342/NTU202600641

際上等同將個人的一舉一動皆置於第三人的監視之下，造成當事人極大的精神壓力。此種情形將導致資料主體為避免遭受外界批判、滿足社會期待，而自我設限或約束行動範圍，最終將嚴重損及個人之人格自主性與完整性[126]。此一觀念亦可從我國《個人資料保護法》第 1 條體現，其明文指出，立法目的在於規範個人資料之蒐集、處理與利用，藉以防止對人格權之侵害，並促進個資之合理使用[127]。

相較之下，資訊隱私權則聚焦於「特定」個資形塑人格發展或主體性(Subjektivität)之緊密關聯性[128]，並將其置重於「生活私密領域不受侵擾之自由」之保障[129]。易言之，雖然某些特定個資可能涉及資訊隱私權的保護，但此並不當然代表所有的個資皆與「隱私」有關，例如個人有意公開之姓名、外貌特徵、職業等資訊，明顯與隱私概念無直接關係[130]。

有論者即指出若僅採取「資訊隱私權」之概念，未加區辨其與資訊自決權之差異，將極易產生誤解，誤以為該項基本權利僅保障具有隱私性質之個資，而忽略其他非隱私性質之資料亦應受個人自主控制權的保障[131]。實則，資訊自決權係將行為自由與私密性的基本權保障前置，藉此可於人格危險的前階段即予以保障[132]。蓋任何個人資料之蒐集與使用，凡涉及個資之蒐集與使用，無論是否具隱私性，皆應尊重當事人之自我決定權，包括揭露與否、揭露方式及後續之知悉、控制與更正。德國學者 Bäcker 亦曾指出，資訊隱私主要提供針對特定生活空間與溝通自由之保密保障，如住宅、通信秘密、網路空間、肖像權與言論表達權；資訊自決權則進一步

126 參范姜真媺，前揭註 117，頁 103。
127 劉青峰，前揭註 110，頁 233。
128 邱文聰，前揭註 118，頁 176；陳銘雄、劉庭妤，前揭註 125，頁 29。
129 張志偉，前揭註 114，頁 31。
130 李震山（2005），〈來者猶可追，正視個人資料保護問題──司法院大法官釋字第六○三號解釋評析〉，《台灣本土法學雜誌》，第 76 期，頁 228。
131 謝碩駿，〈行政機關資料蒐集與個資保護〉，《Human Rights 法務局人權系列》，頁 6，臺北市政府（2020）。
132 Matthias Bäcker, *Grundrechtlicher Informationsschutz gegen Private*, DER STAAT 2012, S. 120.*; Hans-Peter Bull, Zweifelsfragen um die informationelle Selbstbestimmung*, NJW *2006, 1617 (1623); auch* BVerfGE *118, 168 (184 f.)* – Zur Verfassungsmäßigkeit des automatisierten Abrufs von Kontostammdaten.

doi:10.6342/NTU202600641

確保個人於現代資訊社會中，無論於內在心理層面或外在行為層面之完整人格發展與資訊控制能力[133]。

本文認為，此種以資訊自決權為基礎之思維，不僅有助於建構個資法之基本權邏輯體系，亦能於資料侵害風險尚未成形前，即及早啟動防禦機制，預防人格風險的擴大化。此外，隱私範圍之界定常涉及主觀之合理隱私期待，難以形成一致標準，故若從「資訊自決權」之角度出發，即可避免陷入如何界定隱私範圍之爭議，蓋所有個人資料的利用行為皆直接涉及資料主體自主決定之權利，且資訊自決權可將行為自由與私密性的基本權保障前置，從而可於人格危險的前階段即予以保障；至於相關資料與隱私之聯繫程度，則可作為法院進一步評價法規範密度與規範強度之參考基準，以進行適當權衡[134]。

第三項　我國個人資料保護與被遺忘權的憲法基礎

在釐清資訊隱私權與資訊自決權之內涵及其相互關係後，下一重要課題為，究竟應以何種理念作為我國個人資料保護的憲法基礎？就理論而言，並不存在單一且必然的標準。從立法技術的觀點觀之，立法者完全可以選擇僅以「資訊隱私」作為規範依據，亦即僅當個人資料涉及私密生活領域時，方納入基本權之保障範疇，此一構想在法理上並非不可採行[135]。

然而，若僅以「資訊隱私」作為個資保護的憲法基礎，勢必造成對非私密性個資之保障真空。此類資訊雖初看無害，然一旦被大量蒐集並交叉比對，極可能拼湊出完整的人格剖面，導致個人面對過度剖析與監控的風險。因此，若僅以是否屬於隱私資訊作為是否給予憲法保障之依據，將無法因應當代資訊社會中初級資料累積即可能造成高度人格干預的情境。是以，我國《個人資料保護法》自始即未將受

133 Vgl. Matthias Bäcker, ebenda, 95 ff.
134 李震山，前揭註 130，頁 229。
135 張志偉，前揭註 109，頁 7、8。

doi:10.6342/NTU202600641

保護資料限縮於私密資訊，而係採取更為寬廣之資訊自決權觀點作為其制度基礎，俾以更全面防範資訊濫用對個人主體性與人格自主所造成之侵蝕[136]。

另一方面，與資料刪除權密切相關的另一基本權利即為「被遺忘權」。此項權利在憲法體系中可視為人格權，特別是「線上人格權」（Online-Persönlichkeitsrecht）之具體展現[137]。相較於資訊自決權係針對個資蒐集與利用之初始階段加以規範，被遺忘權則聚焦於資訊流通已久後，資料主體基於人格發展、自我重塑或去標籤化（de-labelling）之需求，要求回收個資控制權之後階段保障[138]。舉例言之，受刑人得基於再社會化之目的，請求媒體報導時淡化其身分標籤；一般人亦可主張其在網路環境中享有「難以被搜尋」的權利（Recht auf erschwerte Auffindbarkeit）[139]。

因此，被遺忘權實為人格權向數位空間之延伸，其保障目的在於使個人在特定條件下，於資訊持續流通多年後，仍得主張對其資料進行限制性移除，以回復個人對自我形象之自主建構與控制能力。是以，資訊自決權與被遺忘權構成前後連貫、互為補充的基本權保障機制：前者確保個資在初始蒐集階段即受主體控制，後者則確保即便資訊已擴散流通，資料主體仍不喪失後續控制與修復之權利，進而實現個人於數位社會中「被遺忘的可能」。

第四項　小結

個人資料刪除權的發展，乃是因應數位科技迅速崛起，資訊流通速度與範圍劇烈擴張，導致資料主體對其個人資訊控制能力大幅降低所衍生之法制因應。歐盟法

---

136 同前註。

137 Vgl. Anika D. Luch et al., Ein Recht auS Vergessenwerden als Ausprägung einer selbstbestimmten digitalen Persönlichkeit – Anmerkung zum Urteil des EuGH v. 13.5.2014 (Google), Rs. C-131/12, EuR 2014, S. 710; Christian Hoffmann et al., *Die digitale Dimension der Grundrechte – Das Grundgesetz im digitalen Zeitalter*, 2015, S. 53 ff.; Anika D. Luch/Sönke E. Schulz, *Das Recht auf Internet als Grundlage der Online-Grundrechte*, 2013, S. 41 ff.

138 參張志偉，前揭註 114，頁 32-36。

139 *See* Kai von Lewinski, Staat als Zensurhelfer – Staatliche Flankierung der Löschpflichten Privater nach dem Google-Urteil des EuGH, AfP 2015, S. 1 (2).

doi:10.6342/NTU202600641

院在 2014 年正式承認「被遺忘權」後，各國法制紛紛明文納入相關權利作為法定權利，而我國亦在《個人資料保護法》中具體規範資料刪除權。

然而，在理論層次上，我國長期以來在個人資料保護的憲法基礎定位上，存在「資訊自決權」與「資訊隱私權」之概念混用與爭議。儘管司法院大法官解釋與相關實務判決多已混合兩者之內涵，惟學界對此仍有諸多討論與區辨。本文經由對兩者法制繼受背景、保護範圍與內涵之細緻探討後認為，資訊自決權涵蓋範圍較廣，凡涉及個人資訊之蒐集與利用，均應納入保障，無論是否具隱私性；資訊隱私權則聚焦於具有隱私屬性或合理隱私期待之特定資訊。

本文主張以資訊自決權作為個資法之憲法基礎，俾能更全面因應當代數位社會中廣泛的個人資訊濫用問題。此外，被遺忘權則作為人格權在數位環境下的具體延伸，與資訊自決權形成前後互補的保障體系，前者著重於事後控制與再掌控的可能性，後者則強調事前預防與即時掌控，兩者共同構成完整且穩固的個人資料保護基礎，以因應資訊時代下層出不窮之人格風險與挑戰。

第二節　從現行法制到憲法保障：個資刪除權的實踐基礎

第一項　現行《個人資料保護法》對刪除權之規範

我國個資法於 1995 年施行（當時名為《電腦處理個人資料保護法》），其立法目的旨在規範電腦化個人資料之處理行為，以防止對人格權之侵害，並促進個人資料之合理運用[140]。於 2010 年修正時，保護範圍已不限於電腦處理行為，並進一步擴及至個人資料之蒐集與利用，強化對資料流通各階段之規範[141]。

我國個資法並於多項條文中明定資料主體得行使刪除權之情形；就刪除權之相關規範，本文彙整如下：

[140] 《電腦處理個人資料保護法》（民國 84 年 7 月 12 日公布，民國 100 年 10 月 1 日廢止）第 1 條。

[141]《個人資料保護法》第 1 條立法理由。

doi:10.6342/NTU202600641

個人資料保護法	刪除權之相關內容
第 3 條第 5 款	當事人就其個人資料依本法規定行使之刪除權，不得預先拋棄或以特約限制之。
第 5 條	個人資料之蒐集、處理或利用，應尊重當事人之權益，依誠實及信用方法為之，不得逾越特定目的之必要範圍，並應與蒐集之目的具有正當合理之關聯。
第 11 條第 3 項	個人資料蒐集之特定目的消失或期限屆滿時，應主動或依當事人之請求，刪除、停止處理或利用該個人資料。但因執行職務或業務所必須或經當事人書面同意者，不在此限。
第 11 條第 4 項	違反本法規定蒐集、處理或利用個人資料者，應主動或依當事人之請求，刪除、停止蒐集、處理或利用該個人資料。
第 19 條第 2 項	蒐集或處理者知悉或經當事人通知依前項第七款但書規定禁止對該資料之處理或利用時，應主動或依當事人之請求，刪除、停止處理或利用該個人資料。
第 25 條	非公務機關有違反本法者，除依本法規定裁處罰鍰外，並得命令刪除經處理之個人資料檔案。

【表 三】我國《個人資料保護法》與刪除權相關之規定整理

本研究自行整理製作

doi:10.6342/NTU202600641

個人資料保護法施行細則	刪除權之相關內容
第 6 條第 1 項	本法第二條第四款所稱刪除，指使已儲存之個人資料自個人資料檔案中消失。
第 8 條第 1 項及第 2 項第 6 款	委託他人蒐集、處理或利用個人資料時，委託機關應對受託者為適當監督。前項監督應包含：六、委託關係終止或解除時，個人資料載體之返還，及受託者履行委託契約以儲存方式而持有之個人資料之刪除。
第 20 條	詳列本法第 11 條第 3 項所稱「特定目的消失」之情形。
第 21 條	本法第 11 條第 3 項但書「因執行職務或業務所必須」之刪除權例外情形： 一、有法令規定或契約約定之保存期限。 二、有理由足認刪除將侵害當事人值得保護之利益。 三、其他不能刪除之正當事由。
第 28 條	本法第19條第1項第7款所稱一般可得之來源，指透過大眾傳播、網際網路、新聞、雜誌、政府公報及其他一般人可得知悉或接觸而取得之管道。

【表 四】我國《個人資料保護法施行細則》與刪除權相關之規定整理

本研究自行整理製作

整體而言，我國現行《個人資料保護法》之立法宗旨在於保障個人尊嚴及人格自由發展，其核心精神即在確立資料主體對其個人資料之自主控制權限。基於此一理念，個資法第 3 條第 5 款明文賦予當事人得請求刪除其個人資料之權利，並規定該項權利不得預先拋棄或以特約限制。然而，現行法對刪除權的行使設有一定限制。依第 11 條第 3、4 項及第 19 條第 2 項規定，資料主體僅能於特定情境下（如目的消失、保存期限屆滿、違法處理等）請求刪除個資，構成一種「有限刪除權」制度。

doi:10.6342/NTU202600641

值得注意的是，儘管法律禁止以契約方式事前拋棄刪除權，但實務上透過第 11 條第 3 項但書規定，仍可藉由資料主體的「書面同意」約定保存期限作為延長個資保存的依據[142]。若未適當限制此種約定形式，將可能架空法律原本欲保障之刪除權，使個資長期滯留於資料庫中。此外，第 19 條第 2 項更進一步設有利益衡量之門檻，限縮了對一般可得資料刪除義務的適用範圍，亦使刪除權之行使更加嚴格。

第二項　資料刪除權之憲法基礎與事後控制權的確立

資料刪除權一般被視為隱私權與資訊自決權具體展現之一環[143]。雖然我國憲法並未明文規定此二權利，惟大法官如上述已透過多次解釋，援引憲法第 22 條所保障之未列舉基本權利，明確導出隱私權與資訊自決權之憲法保障地位。例如釋字第 585 號解釋首次明確指出，基於人性尊嚴、個人主體性與人格完整發展，憲法第 22 條保障個人生活秘密空間免於外界侵擾，並強化個人對資料的自主控制權[144]。隨後，釋字第 603 號進一步闡明個人對其個人資料擁有自主控制的權利，包括對資料揭露與使用的自主性及更正錯誤資料的權利[145]。接續的釋字第 689 號解釋，則更明確地將個人資料自決權視為憲法所保護的基本權利，彰顯對個人資料控制權的重要性[146]。

憲法法庭於 111 年憲判字第 13 號判決（健保資料庫案）進一步提出「事後控制權」之概念，指出個人原則上享有對個人資料的事後控制權，即使資料已符合法定條件蒐集或經個人同意，資料主體亦得請求刪除、停止或限制資料之使用[147]。換

---

142 《個人資料保護法》第 11 條第 3 項但書已對刪除權設有例外，明文規定若個資「因執行職務或業務所必須，或經當事人書面同意者」，得不受「目的消失或期限屆滿即應刪除」之限制。又依據《施行細則》第 21 條第 1 款，「因執行職務或業務所必須」包含「有法令規定或契約約定之保存期限」；而國家發展委員會亦於《國家發展委員會發法字第 1100016400 號函》中進一步指出，若經當事人書面同意，即可於目的達成或保存期限屆滿後繼續保留個資，但此同意仍應符合《個資法》第 5 條所揭示之誠信與比例原則，不得以籠統約定或無限期保留方式逾越其界限。

143 參照：張志偉，前揭註 114，頁 28-32。

144 司法院釋字第 585 號解釋理由書第 25 段。

145 司法院釋字第 603 號解釋解釋文第 1 段。

146 司法院釋字第 689 號解釋解釋文。

147 111 年憲判字第 13 號判決：「憲法第 22 條個人資訊隱私權保障當事人原則上應有事後控制權，且當事人就獲其同意或符合特定要件而允許未獲當事人同意而經蒐集、處理及利用之個資，仍具事

doi:10.6342/NTU202600641

言之，憲法法庭認為，現行法若未就健保資料之目的外利用建立相應之停止機制，即一律剝奪資料主體請求限制利用之權利，係違反資訊隱私權而已屬違憲，並責成主管機關限期檢討修法或制定新法[148]，以完善健保資料之二次利用法制[149]。但值得注意者係，此號判決乃針對健保資料之特殊情境，並非一般性地確立一個絕對且無限制的個資刪除權，而係強調未來法律應妥適調和資料主體的資訊自主需求與資料使用者的合法需求。

綜上所述，雖我國憲法並未明文規定人民享有「資料刪除權」，然透過大法官解釋與憲法法庭判決，已逐步奠定個人資料自主控制權之憲法基礎，並明確肯認個人對其資料享有一定程度之事後控制權。此等權利保障的核心精神，在於回應現代社會中資料可被重複流通與運用的特性，確保資料主體得以主動維護其人格權與資訊隱私。然而，當人工智慧等新興技術導致資料一旦進入模型後即難以刪除或追溯時，如何在技術實務與憲法保障間取得平衡，便成為當前資料治理體系亟須面對的重要課題。下一節本文將轉向討論機器學習架構下資料刪除的技術困境，進一步探究人工智慧技術與刪除權之間的實質衝突。

第三節　機器學習中的資料刪除問題

如前所述，個資刪除權的核心在於刪除先前公開的個人資料，其本質上是將「遺忘」的概念從人類記憶延伸至數位環境中；要求刪除個人資訊的行為，等同於象徵性地請求他人將該資訊「忘記」。然而，這個類比僅適用於人類記憶，對於人

後控制權，不因其曾表示同意或因符合強制蒐用要件，當事人即喪失請求刪除、停止利用或限制利用個資之權利。」

[148] 在 111 年憲判字第 13 號判決公布後，衛生福利部已於 113 年公告《衛生福利資料管理條例》草案，以回應大法官限期修正相關法規之要求。該草案於第 15 條至第 18 條中規範了資料主體對其健保資料之停止利用請求權，原則上賦予資料主體針對目的外利用行使停止利用之權利。
然而，根據第 17 條第 1 項之規定，停止利用之效力僅及於註記日期之後，對於註記前已提供目的外利用之資料則不適用。
此外，草案第 18 條另設有三項例外情形，使得即便資料主體請求停止利用，仍得為下列目的繼續利用資料： 一、依其他法律之規定，由主管機關或其指定之所屬機關提供者； 二、基於維護國家安全之必要； 三、為免除人民生命、身體、自由或財產之急迫危險。

[149] 劉靜怡（2022），〈違憲之後－111 年憲判字第 13 號判決簡評〉，《當代法律》，第 11 期，頁 9。

doi:10.6342/NTU202600641

工智慧或機器學習系統恐難以適用。以下將具體說明個資刪除權在機器學習模型中實踐之困難。

第一項　機器學習如何使用資料

過去十年間登上媒體頭條、深刻改變社會的人工智慧技術——如臉部辨識、深偽技術與大型語言模型（如 GPT-3 與 Bard）——其本質多屬於機器學習範疇，尤以深度學習為主。深度學習作為機器學習的子領域，其核心機制是將大量資料輸入至具多層結構的人工神經網絡中，讓模型自動從中辨識潛在規律，進而執行分類、預測或生成等任務[150]。

在這樣的架構中，模型透過稱為「權重」（weights）與「偏差」（biases）的內部參數來調整其運算結果，而這些參數值乃是根據訓練資料經反覆優化所得，進一步構成模型進行判斷與推理的依據[151]。值得注意的是，模型在推論階段並非直接檢索訓練資料，而是基於這些內隱地存於參數中的規律進行運算，據此產出新的分類或生成內容[152]。

人工智慧模型的訓練資料通常來自各種來源，例如公開的網路資料庫、感測器收集的資料、用戶提供的內容等[153]。其中，大型語言模型更是以網際網路上的龐大文字資料作為訓練來源。例如 GPT-3 有約 60%的訓練資料來自 Common Crawl 網頁資料[154]，PaLM 模型約 50%的資料來自社群媒體對話[155]，而 OpenAI 和 Google 也廣泛使用 Reddit 貼文作為訓練資料[156]。此外，許多 AI 公司（如 OpenAI）會將使

150 Meem A.M., *supra* note 65.
151 *Id.*
152 *Id.*
153 Wayne Xin, Zhao et al., *A Survey of Large Language Models*, at 13-14 (Mar. 31, 2023), https://arxiv.org/abs/2303.18223.
154 *See* Tom B. Brown et al., *Language Models Are Few-Shot Learners*, 33 ADV. NEURAL INFO. PROCESS. SYST. 1877, 1884-85 (2020).
155 See Aakanksha Chowdhery et al., *PaLM: Scaling Language Modeling with Pathways*, 240 J. MACH. LEARN. RES., 11324, 11329-30 (2022).
156 *See* Alec Radford et al., *Language Models Are Unsupervised Multitask Learners*, 1(8) OPENAI 9, 11 (2019).

doi:10.6342/NTU202600641

用者與模型互動產生的新資料（如對話記錄）收集起來，用於後續的模型微調或再訓練[157]。

綜上所述，資料是機器學習模型的燃料，模型透過對訓練資料的分析與分類，形成參數，並進一步用於後續的運算，而這些訓練資料的來源則遍布於你我的網路生活中。

第二項　資料刪除的技術困境

第一款　模型學習的黑箱與不可逆性使資料難以完全刪除

隨著機器學習系統廣泛應用，針對個人資料的刪除請求，在技術層面面臨諸多挑戰。首先，較大的挑戰在於：模型參數對訓練資料的依賴性高度不可逆。與傳統資料庫不同，深度學習模型並非將一筆筆資料獨立儲存，而是將整體資料的資訊以複雜方式融合進數百萬、甚至數十億個參數之中[158]。

從統計上來看，深度學習模型的參數數量通常低於其所訓練資料中的詞元數[159]。例如 GPT-4 據稱擁有約 1.7 兆個參數，訓練資料則包含約 13 兆個「詞元」（token，為詞句的最小單位）[160]；相比之下，GPT-3 僅有 1750 億個參數[161]。這些參數負責「記住」詞元資訊，因此每個詞元可能對應數個參數，而單一參數亦可能同時參與多個詞元的記憶，就如同人腦中的神經元一般[162]。理論上，若希望 GPT-4「遺忘」一句話，如：「哈利波特可被視為魔戒的同人小說」，只要刪除一至兩個相關參數即可——但問題在於，我們無從明確辨識究竟是哪些參數記住了哪些資訊[163]。

---

157 *Id.*
158 Meem A.M., *supra* note 65.
159 *Id.*
160 See Anis Koubaa, *GPT-4 vs. GPT-3.5: A Concise Showdown* 2 (Apr. 7, 2023), https://doi.org/10.36227/techrxiv.22312330.v2.
161 See Katikapalli S. Kalyan, *A Survey of GPT-3 Family Large Language Models Including ChatGPT and GPT-4*, 6 NAT. LANGUAGE PROCESS. J., art. no. 100048, at 12-13 (2024).
162 Meem A.M., *supra* note 65, at 3.
163 *Id.*

doi:10.6342/NTU202600641

有學者指出，神經網絡與其說是壓縮工具，不如說是某種加密系統[164]。即使某些機器學習模型參數數量超過訓練資料規模，仍難以透過「逆向工程」精確地判斷哪一組參數對應哪一筆資料，就如同在動物腦中抹除一段記憶時，往往也會一併失去與其相關的其他資訊，而這種關聯性經常難以預測[165]。亦有研究發現，刪除某位使用者的資料，可能導致系統一併抹去其他同名使用者的資訊，顯示其內部資料的相關性難以釐清[166]。

在大型模型中，移除特定資訊常會導致連鎖反應。例如為了忘記某一項內容，模型可能連帶遺失相關知識（類似「災難性遺忘」現象），導致輸出異常或整體效能下滑；且一旦刪除原始訓練資料，也將影響模型可驗證性，使其難以重建或確認其輸出基礎[167]。

早期學者即將神經網絡比喻為「黑箱」，其不僅是資料的集合體，更是能夠攝取資料並進行智能推論的系統[168]。若真如此，則欲使系統「遺忘」特定資訊，無疑是一極其困難的工程[169]。另有研究指出，現代大型語言模型有時會在未經提示的情況下，意外重現訓練資料中的個人資訊[170]；即便開發者試圖將相關資料從訓練集中移除，該資訊仍可能以殘留形式存在於模型中[171]。

因此，一旦個人資料被納入了訓練，其對模型參數的影響很難完全抹除。深度學習模型仰賴龐大資料進行模式學習，其「不可逆記憶」特性使資料刪除遠比傳統資料庫中刪除紀錄複雜，對「資料刪除權」的實踐形成巨大挑戰。

---

164 *See* Sakshi Patel et al., *Colour Image Encryption Based on Customized Neural Network and DNA Enco*ding, 33 NEURAL COMPUTING AND APPLICATIONS 14533, 14534 (2021).
165 Meem A.M., *supra* note 65, at 3.
166 See Jie Xu et al., *Machine Unlearning: Solutions and Challenges*, 8 IEEE TRANS. EMERG. TOPICS COMPUT. INTELL. 2150, 2151–52 (2024).
167 EUROPEAN PARLIAMENTARY RESEARCH SERVICE, THE IMPACT OF THE GENERAL DATA PROTECTION REGULATION (GDPR) ON ARTIFICIAL INTELLIGENCE 57 (2020).
168 *See* Neil Savage, *Breaking into the Black Box of Artificial Intelligence*, NATURE (Mar. 29, 2022), https://www.nature.com/articles/d41586-022-00858-1.
169 Meem A.M., *supra* note 65, at 3.
170 *Dawen Zhang et al., Right to be Forgotten in the Era of Large Language Models: Implications, Challenges, and Solutions*, 5 AI ETHICS, 2445, 2448 (2025). https://doi.org/10.1007/s43681-024-00573-9.
171 *Id.*

doi:10.6342/NTU202600641

第二款　刪除時效性與運算成本高昂

另一個技術困境在於刪除的時效性與成本。資料刪除的法律（如 GDPR 第 17 條）通常要求在大約一個月之合理期間內刪除個人資料[172]。但對於已訓練完成的大型模型，要在短時間內讓模型「遺忘」某筆訓練資料，目前技術能力尚難企及，蓋從訓練資料庫中刪除個人資料不會直接影響已經訓練的模型，僅能在隨後對模型進行訓練後生效[173]。

一種可行但代價極高的方式是重新訓練模型（文獻稱為 exact unlearning）[174]。然而，這種方法的代價極為高昂，據估計，GPT-4 的訓練過程耗費了 1,024 部高效能伺服器長達 34 天的運算資源[175]，若僅為了刪除某個人的資料而重新訓練模型，除了龐大的時間與硬體成本外，亦需消耗約 50 GWh 的能源——相當於美國 4400 萬戶家庭一年的用電量，幾乎覆蓋全美三分之一人口，亦是愛爾蘭全年用電的 24 倍[176]。

綜上所述，即便技術上能否完全刪除參與模型訓練的個人資料尚無定論，若欲徹底消除該資料對模型所產生的影響，所需付出的時間、能源與成本均極為龐大。換言之，若每一筆刪除請求皆需重新訓練模型一次，在現實運作上顯然並不可行。

第三項　技術上的可能解方

第一款　機器遺忘（Machine Unlearning）

為了調和人工智慧發展與個人隱私的保護，許多學者正致力於「機器遺忘」（Machine Unlearning）的研究，以期能改善個人資料刪除權實踐上的困難。機器

---

[172] GDPR.EU, *Everything You Need to Know About the "Right to Be Forgotten"*, https://gdpr.eu/right-to-be-forgotten/ (last visited Aug. 5, 2025) .
[173] Meem A.M., *supra* note 65, at 3.
[174] *See* Vikram S. Chundawat et al., *Zero-Shot Machine Unlearning*, 18 IEEE TRANS. INF. FORENSICS & SEC. 2345, 2346 (2023).
[175] Meem A.M., *supra* note 65, at 3.
[176] RISE Research Institutes of Sweden, *Generative AI Does Not Run on Thin Air!*, RISE (Feb. 21, 2024), https://www.ri.se/en/news/blog/generative-ai-does-not-run-on-thin-air.

doi:10.6342/NTU202600641

遺忘（是一項旨在使人工智慧模型從既有訓練資料中「選擇性遺忘」特定資料的技術，對於實踐資料主體的「個人資料刪除權」具有關鍵意義[177]。此技術核心在於逆轉特定資料對模型參數所造成的影響，亦即透過一系列演算法，讓模型如同未曾見過該筆資料一般[178]。

現已有多種方法被提出以實踐機器遺忘，大致可依其對模型影響的可控程度，區分為「精確機器遺忘」（Exact Machine Unlearning）與「近似機器遺忘」（Approximate Machine Unlearning）兩類。前者係指能確保模型在移除特定資料後，行為等同於從未見過該資料的情況，最常見的實現方式為「完整重訓」，即移除資料後重新從頭訓練模型[179]；後者則係指在不進行完整重訓的情況下，透過調整模型權重、正則化懲罰、或差分隱私技術來「近似」移除資料對模型的影響[180]。這些方法雖然無法保證百分之百抹除資料影響，但能在計算成本與隱私保障之間取得平衡，適用於大多數實務應用[181]。以下將簡要介紹各種方法並說明其可能面臨的困難。

第一目　精確機器遺忘

精確機器遺忘（Exact Machine Unlearning）旨在徹底消除特定資料對模型的影響，通常需部分或完全重新訓練模型。其關鍵挑戰在於如何在保有模型效能的前提下，加速訓練過程[182]。

最早的研究由 Cao 等人提出，他們設計了一種可將機器學習演算法轉為「總和形式」的遺忘方法，藉此中和資料影響[183]。不過，這種方法僅適用於部分演算法，且無法處理所有資料細節[184]；後續的 SISA 方法（Sharded, Isolated, Sliced, Aggregated）

177 Youyang Qu et al., *Learn to Unlearn: Insights Into Machine Unlearning*, 57(3) COMPUTER, 70, 79 (2024), https://doi.org/10.1109/MC.2023.3333319.
178 *Id.*
179 *Id* at 80.
180 *Id.*
181 *Id.*
182 *Id.*
183 See Yinzhi Cao et al., *Towards Making Systems Forget with Machine Unlearning*, IEEE SYMPOSIUM ON SECURITY AND PRIVACY 463, 463–80 (2015).
184 Youyang Qu et al., *supra* note 177, at 81.

doi:10.6342/NTU202600641

透過資料分片與漸進式訓練，只需針對特定區段重新訓練，大幅提升效率[185]，然而在資料量不足或任務複雜時，模型表現可能下降[186]；另一種 DeltaGrad 技術，則透過「反向訓練」來移除資料影響，即最大化損失函數，使模型忘記該資料，效果在小比例刪除情況（如 1%）尤為顯著[187]，但僅支援特定類型的訓練流程[188]。此外，也有如 DaRE 等專門針對單一演算法（如隨機森林）設計的快速遺忘方案，能在效率上大幅優於完全重訓，但應用範圍同樣有限[189]。

整體而言，精確遺忘技術的研究多集中於如何兼顧刪除精度與效率[190]。然而仍面臨兩大問題：（一）模型性能下降風險：即便部分技術已能減少影響，遺忘操作仍可能破壞原本的模型平衡。（二）隱私洩漏可能性：若重新訓練後模型行為產生變化，攻擊者可藉比較前後輸出推測哪筆資料被刪除，進而觸發推論攻擊[191]。

綜上所述，儘管精確機器遺忘使資料刪除成為可能，從而得以實現符合法規要求的資料刪除，但仍均有其限制。縱然於特定演算法得以成功應用精確機器遺忘，後續亦可能帶來潛在風險，使模型性能下降或洩露隱私。

第二目　近似機器遺忘

為了解決精確機器遺忘的挑戰，研究者發展出近似遺忘（Approximate Machine Unlearning）技術，其核心目標是在不顯著損害模型效能的前提下，掩蓋刪除前後模型之間的差異，以維護隱私與準確率的平衡[192]。

---

185 See Lucas Bourtoule et al., *Machine Unlearning*, IEEE SYMPOSIUM ON SECURITY AND PRIVACY 141, 141–159 (2021).
186 Youyang Qu et al., *supra* note 177, at 81.
187 See Yinjun Wu et al., *Deltagrad: Rapid Retraining of Machine Learning Models*, 119 PROC. 37TH INT'L CONF. MACH. LEARN. 10355, 10355–66 (2020).
188 Youyang Qu et al., *supra* note 177 at 81.
189 See Jonathan Brophy & Daniel Lowd, *Machine Unlearning for Random Forests*, 139 PROC. INT'L CONF. MACH. LEARN. 1092, 1092–1104 (2021).
190 Youyang Qu et al., *supra* note 177 at 81.
191 *Id.*
192 *Id.*

doi:10.6342/NTU202600641

例如，有研究透過數學公式調整模型參數，或對訓練過程加入隨機擾動，模擬資料未曾存在的效果[193]；另有方法能在不重訓模型、不接觸原始資料的情況下移除資料影響，甚至延伸至特定類型的機器學習（如貝葉斯學習或 K-means 演算法）[194]。這些方法多以「效率高、靈活性強」為優勢，但往往只能近似移除資料影響，難以做到法律意義上的完全刪除[195]。

儘管近似機器遺忘技術在實務上具有可行性，其合法性仍存在兩大疑問：（一）驗證困難：使用者或主管機關難以單憑模型輸出，確認資料是否真的「被遺忘」；（二）該遺忘技術是否符合「個人資料刪除權」的法律標準？蓋多數近似方法只改變模型結果，而非從資料源頭刪除，與法規所強調的「資料去除」恐仍有落差[196]。

因此，即使技術達成了「模擬性遺忘」，在法律層面上似乎仍需要更完善的驗證機制，以確認該解方可確實回應個人資料保護法規對「刪除」的要求。

第三目　機器遺忘的驗證與攻擊

隨著機器遺忘技術的發展，學界也開始關注如何驗證在近似機器遺忘後的模型資料是否真的被刪除，以及是否可能遭受新的隱私攻擊。目前關於驗證機制的研究仍極為有限，但攻擊相關的研究則相對較多[197]。

---

193 See Chuan Guo et al., *Certified Data Removal from Machine Learning Models*, 359 PROC. INT'L CONF. MACH. LEARN. 3832, 3832–42 (2020).

194 See Antonio A. Ginart et al., *Making AI Forget You: Data Deletion in Machine Learning*, in PROC. 33RD INT'L CONF. NEURAL INFO. PROCESS. SYST., art. 316, 3518, 3521–26 (2019); Quoc Phong Nguyen et al., *Variational Bayesian Unlearning*, 33 ADV. NEURAL INFO. PROCESS. SYST. 16025, 16027–34 (2020).

195 Youyang Qu et al., *supra* note 177 at 82.

196 *Id.*

197 *Id.*

doi:10.6342/NTU202600641

在驗證方面，有學者提出將「後門」（backdoor）資訊植入資料中[198]，透過驗證模型是否仍對該後門做出反應，以檢查資料是否真正被刪除[199]；其他方法如樣本抽查、合成資料測試與偏誤分析等，則可評估模型在刪除特定資料後是否產生不當偏差，從而影響模型的公平性與準確性[200]。然而，上述方法並非完美，除可能帶來新的後門風險外，其亦難以應用於大規模系統[201]。

在攻擊面向，成員關係推論攻擊（membership inference attacks）是一種常見手法，可推斷某筆資料是否曾被用於訓練[202]。有研究指出，即使資料被刪除，攻擊者只要能取得刪除前後的模型版本，仍可能透過模型變化反推出原本應被遺忘的資訊，進而造成新的隱私洩漏風險[203]。此外，也有研究透過資料投毒攻擊，精心設計並注入特定樣本，成功擾亂機器遺忘機制，使模型表面上完成刪除程序，實際上卻仍殘留被刪除資料的影響[204]。這不僅削弱機器遺忘的技術效果，也使資料主體在實踐「個人資料刪除權」時面臨更大的挑戰。

第四目　小結

儘管現有機器遺忘技術已初步展現滿足法律需求的潛力，然而其在技術成熟度、適用範圍與法律可驗證性上仍面臨諸多侷限。首先，就將模型完全重訓（精確機器遺忘）此一方法而言，現有技術可適用的模型相當有限，難以廣泛應用於複雜模型如大型語言模型或深度神經網路中；且縱然可大幅提升訓練效率，若是在資料刪除比例高、或刪除資料具有特殊偏向時，模型性能可能出現預期外的退化；此外，

198 所謂「後門」（backdoor）資料檢驗方法，係指在需要刪除資料的情境中，資料擁有者會在資料中故意植入後門，再將資料交付給資料使用者（如大型網路公司），一旦資料使用者聲稱已經刪除該資料，資料擁有者便可透過檢查先前植入的後門來確認是否真的完成刪除。

199 See David M. Sommer et al., *Athena: Probabilistic Verification of Machine Unlearning*, 3 PROC. PRIV. ENHANCING TECHS. 268, 268–290 (2022).

200 *Id.*

201 *Id.*

202 *See* Reza Shokri et al., *Membership Inference Attacks Against Machine Learning Models*, IEEE SYMPOSIUM ON SECURITY & PRIVACY 3, 3–18 (2017).

203 *See* Min Chen et al., *When Machine Unlearning Jeopardizes Privacy*, PROC. INT'L CONF. MACH. LEARN. 896, 896–99 (2021).

204 *See* Neil G. Marchant et al., *Hard to Forget: Poisoning Attacks on Certified Machine Unlearning*, 36(7) PROC. AAAI CONF. ARTIF. INTELL. 7691, 7691–98 (2022).

doi:10.6342/NTU202600641

有心人士亦得透過比較刪除前後模型的差異，推測出被刪除的資料為何，從而造成隱私外洩的風險。

其次，近似遺忘方法雖在效率上具優勢，但因其無法保證完全移除資料影響，在法律上是否符合「資料刪除」的要求，仍有疑義。此外，目前缺乏標準化的驗證框架，使用者與監管機關難以依據模型輸出判斷遺忘是否確實完成，亦無法有效舉證違規行為。最後，機器遺忘機制本身亦可能成為新型攻擊的目標，例如透過成員推論或資料投毒方式，推測已刪除的資訊或破壞遺忘效果，使遺忘成為一種表面承諾，實則隱患重重。

綜上所述，未來機器遺忘技術的發展，急需強化通用性與穩健性，並輔以可核實、具法律可信度的驗證機制，方能真正支撐「刪除權」於數位時代的落實。

第二款　以合成資料訓練模型

第一目　合成資料應用概覽

有研究指出，當原始資料無法使用時，合成資料（Synthetic Data）可作為一種有效的替代方案，或可成為協調人工智慧技術與個人資料刪除權衝突的潛在解方。這類資料由演算法生成，在統計特性上與真實資料相似，但不含任何具體個人資訊，因此可支援機器學習與資料分析，同時降低對個人隱私的侵犯風險[205]。

簡言之，合成資料是由人工演算法生成的虛構資料，其分布與真實資料相仿，具有代表性卻不具可識別性[206]。此技術已有數十年發展歷史，應用涵蓋資料補全、人口模擬等多種場景，自 2019 年起，新興的生成對抗網路（Generative Adversarial Network, GAN）[207]等深度學習方法更是突破了過去合成資料生成的侷限，能夠在

[205] Jiri Hradec et al., MULTIPURPOSE SYNTHETIC POPULATION FOR POLICY APPLICATIONS, PUBLICATIONS OFFICE OF THE EUROPEAN UNION 15 (2022).
[206] *Id.*
[207] 生成式對抗網路 (GAN) 是一種深度學習架構，由兩個彼此競爭的神經網路所組成，目的是從指定的訓練資料集中產生更擬真的新資料。例如，它可以根據現有的影像資料庫生成新的圖像，或從歌曲資料庫中創作出原創音樂。GAN 被稱為「對抗性」網路，是因為它透過兩個網路的相互對

doi:10.6342/NTU202600641

學習高維資料分布後，自動產生具備統計一致性的新樣本[208]。由於合成資料雖保留了原資料的結構與變異性，卻無對應的真實個體，故即便用於模型訓練，亦可免除後續因資料刪除要求而需重新訓練模型的負擔，對實踐個人資料刪除權極具潛力[209]。

合成資料的應用已有諸多實例，例如 Beaulieu-Jones 等人於 2019 年展示了如何透過具備差分隱私機制（Differential Privacy, DP）[210]的 GAN，產出可供共享的合成病歷資料，使得多個醫療機構能夠共享和分析病人資料，而無需曝露真實的個人健康資訊，成功兼顧醫療研究價值與病患隱私保護[211]。在企業開發環境中，也常見以合成資料取代敏感客戶資料，讓資料科學家或外部開發者得以在不接觸實際個資的前提下，自由的訓練與調整機器學習模型；待模型訓練完成後，再於受控環境接入真實敏感資料進行部署分析[212]，在這樣的工作流程中，由於開發人員自始未直接訪問真實個人資料，大大降低了隱私外洩風險。

此外，許多調查資料在蒐集時僅限特定用途，若欲轉作他用，通常需重新取得受訪者同意，否則可能違反原有的倫理協議或法律規定。而合成資料可在不觸及真實個資的前提下，實現資料的「次級利用」，詳言之，若研究者擬將資料集用於未

抗來進行訓練：其中一個稱為生成器（Generator），負責根據輸入資料產生經過修改的全新資料；另一個則是判別器（Discriminator），其任務是判斷輸出資料是否來自原始資料集。換言之，判別器要分辨生成的資料是真實的還是偽造的，而生成器則不斷調整輸出，使其愈來愈接近真實資料。這種對抗性的訓練過程會持續進行，直到判別器無法再準確區分真假資料為止，達成高度擬真生成的效果。*See* AWS, *What is GAN?*, https://aws.amazon.com/tw/what-is/gan/ (last visited：2025/03/31).

208 *See* Jiri Hradec et al., *supra* note 205.

209 *Id.*

210 差分隱私（Differential Privacy）的核心理念，是透過在資料中引入適量雜訊，以模糊可識別的敏感資訊，從而達成保護個人隱私的目的。常見的實作方式之一，為先對原始表格式資料進行統計，取得各種資料樣態及其分布頻率後，於此分布上添加擾動雜訊，再經由後處理機制調整雜訊的合理性與一致性，最後轉換回表格形式，即可產生具差分隱私保護之衍生資料集。
在此過程中，雜訊添加的幅度取決於「隱私預算」（privacy budget, ε）的設定。隱私預算愈大，所加雜訊愈少，資料保留的真實性愈高；反之，當 ε 較小時，雖可提供較強的隱私保障，但也會導致資料失真程度增加。隱私預算的設定並無固定標準，實務上須在隱私保護強度與資料可用性之間取得適當平衡。參見：吳品樺（06/24/2024），〈隱私強化技術：平衡資料保護與資料應用〉，數位發展部，https://moda.gov.tw/press/multimedia/blog/12810（最後瀏覽日：08/06/2025）。

211 *See* Brett K. Beaulieu-Jones et al., *Privacy-Preserving Generative Deep Neural Networks Support Clinical Data Sharing*, 12(7) CIRC CARDIOVASC QUAL OUTCOMES at 2-4 (2019).

212 *See* Jiri Hradec et al., *supra* note 205 at 16.

doi:10.6342/NTU202600641

曾取得同意的新研究問題，可選擇生成對應的合成資料，繼續推動研究而無需重新取得同意[213]。

總結來說，當原始資料受限於隱私或授權問題而無法使用時，合成資料提供了一項兼顧效能與法遵的替代方案，其透過演算法生成不具識別性的虛構資料，能保有統計代表性並支援機器學習訓練，避免侵犯個資。隨著生成對抗網路等技術成熟，合成資料的應用已逐步拓展至醫療、企業開發與次級研究等場域，對實踐個人資料刪除權具有高度潛力。

第二目　合成資料的技術優勢

合成資料具有多方面的優點，使其在隱私受限的人工智慧開發中備受關注。首先是隱私保護與法令遵循，增強隱私保護是合成資料最顯著的優勢，由於合成資料不對應真實個人，資料中不存在直接的個人識別資訊，大大降低了敏感資訊外洩的風險[214]。這也使得合成資料更容易符合如 GDPR 等資料保護規範的要求：若資料無法識別特定個體，則通常不受嚴格法規約束，例如有論者認為，對於純合成的資料集，個人無法主張「個人資料刪除權」，因為該資料集中從未包含其真實資訊[215]。如此便降低了法律風險，使機構能利用原本因隱私顧慮而無法使用的資料，換言之，合成資料讓模型開發者能在不處理受規範個資的情況下，仍獲得類似洞察，兼顧資料效用與隱私合規[216]。

其次是合成資料的高度靈活性與跨領域適用性，合成資料幾乎可針對任何領域與資料類型進行生成，無論影像、文本或結構化表格資料，皆可透過訓練對應生成模型產出逼真的樣本[217]。舉例而言，醫療領域可生成與真實病患特徵相符的合成病歷，以供資源規劃與分析使用，而不含任何個資；金融機構能創建合成交易紀錄，

---

213 *Id.*
214 *See* Jiri Hradec et al., *supra* note 205 at 15.
215 Newton Investment Management, *Synthetic Data: Fuel for the AI World* (May 22, 2023), https://www.newtonim.com/us-institutional/insights/blog/synthetic-data-fuel-for-the-ai-world/.
216 *Id.*
217 *Id.*

doi:10.6342/NTU202600641

模擬「假設情境」以改進風控模型；自駕車開發中則可生成罕見場景（如暴風雪中遇到救護車）資料，彌補真實資料難以涵蓋的極端情境。在上述案例中，合成資料皆補足了真實資料的不足，並提供低成本、高效率的替代方案，透過生成程序的客製化，研究與工程應用可獲得任意數量、具特定特徵分布的合成資料[218]。正因如此，業界分析預測合成資料將迅速普及——Gartner 曾預言至 2024 年，AI 所使用的資料中將有 60%為合成資料[219]。

最後，合成資料亦提供處理偏見與品質問題的機會，詳言之，由於生成過程可設計，人們得以在生成時修正不平衡或弱化偏見[220]。例如若原始資料中某群體樣本不足或存在偏見，可透過生成更多該群體資料或移除偏誤特徵來建立更平衡的訓練資料庫，此舉應有助減輕 AI 模型中繼承歷史資料偏見的風險，推動更公平的模型行為[221]。此外，生成資料的速度與成本亦遠低於蒐集真實資料，使其成為高效、低風險的 AI 輸入資料來源[222]。

第三目　合成資料面對資料刪除權之潛在挑戰

儘管合成資料的前景備受看好，在實務應用上仍須審慎評估其潛在限制與風險。首先，文獻指出，合成資料難以完全排除洩漏資訊的風險。當生成模型出現「過擬合」（overfitting）現象，即過度模仿原始資料時，可能導致原始資料中的特定模式或離群值意外反映於合成資料之中[223]。舉例而言，若某使用者具有極端異常的行為模式（例如社群網站造訪次數遠高於他人），該特徵便可能被模型學習並轉譯至合成資料中。由於此類行為在真實世界中僅有一人具備，知情者可能藉此推測其身分，進而引發個資外洩之疑慮[224]。

---

[218] *Id.*
[219] Gartner, *Gartner Identifies Top Trends Shaping the Future of Data Science and Machine Learning*(Aug. 1, 2023), https://www.gartner.com/en/newsroom/press-releases/2023-08-01-gartner-identifies-top-trends-shaping-future-of-data-science-and-machine-learning.
[220] *See* Newton Investment Management, *supra* note 215.
[221] *Id.*
[222] *Id.*
[223] *See* Steven M. Bellovin et al., *Privacy and Synthetic Datasets*, 22 STAN. TECH. L. REV. 1, 37-39 (2019).
[224] *Id.*

doi:10.6342/NTU202600641

為因應此一問題，學界逐漸將希望寄託於差分隱私與合成資料生成技術的結合[225]，該方法被視為兼顧資料安全與實用性的可行解方：透過合成資料保留統計效用，並以差分隱私機制於特定門檻下提供強健的隱私保障[226]。然而，此解方仍存有限制。研究指出，即便輸入資料的分布相對合理（即無明顯離群值），仍無法完全排除部分原始資料洩漏的風險。事實上，如何設定適當的保護門檻，正是實現隱私與效用平衡的關鍵所在，而這兩者始終存在結構性的取捨關係[227]。

此外，合成資料亦非萬用工具，尤其在需進行精準統計推論或辨識低佔比群體特徵的情境下，其效用相對受限[228]。例如，若研究目的在於識別罕見樣態或邊界個案，合成資料往往傾向重建主流模式，反而掩蓋重要特徵。即便導入差分隱私強化保護，在高保護強度下所引入的雜訊亦可能大幅削弱資料可用性，形成隱私與效能難以兩全的困境[229]。

最後，合成資料是否適用於刪除權規範，亦為一值得關注的爭點。有論者主張，既然匿名資料不受多數全球隱私法框架（如 GDPR）所規範，具匿名特性的合成資料亦應不在其約束之列。例如，《Harvard Business Review》曾引述一合成資料公司指出：「由於全球隱私法日益嚴格，分享個人資訊變得困難甚至不可能，因此合成資料變得至關重要——它完全匿名，因而不受這些法律限制[230]。」然而，亦有學

[225] 就差分隱私的技術說明，請參註 210。
[226] Steven M. Bellovin et al., *supra* note 223.
[227] *Id.* 如 Chong Huang 等人即指出，在使用 GAN 生成合成資料的情況下，即便套用差分隱私機制，也未必能確保資料的安全，該研究指出，任何聲稱具有隱私保護功能的「協作式深度學習系統」，都有可能被其團隊設計的強力攻擊方式破解，換言之，就算用了差分隱私，在某些情況下還是可能有洩漏訓練資料的風險。*See* Chong Huang et al., *Context-Aware Generative Adversarial Privacy*, 19(12) ENTROPY 656 (2017).
[228] *See* Steven M. Bellovin et al., *supra* note 223205 at 41.
[229] *Id.*
[230] HARVARD BUS. REV. ANALYTIC SERVS., THE EXECUTIVE'S GUIDE TO ACCELERATING ARTIFICIAL INTELLIGENCE AND DATA INNOVATION WITH SYNTHETIC DATA 3 (2021), https://perma.cc/5J82-HG5W. The report does not relate to specific conditions but includes general statements. Nvidia-supported research suggests that, as "synthetic data would not be considered identifiable personal data, privacy regulations would not apply, and obligations of additional consent to use the data for secondary purposes would not be required." KHALED EL EMAM, ACCELERATING AI WITH SYNTHETIC DATA: GENERATING DATA FOR AI PROJECTS 6 (2020).

doi:10.6342/NTU202600641

者質疑此觀點，認為是否構成個資應視具體脈絡判斷[231]。即使資料表面不含可識別資訊，若可透過交叉比對其他資料集或統計推論重建個體資訊，仍可能落入個資保護規範之適用範圍[232]。

事實上，已有技術文獻指出，若合成資料係基於真實資料訓練而得，並作為模型的輸入或比對依據，即使其本身並非真實個資，仍存在被推論或鏈結回特定個體之風險[233]。因此，即便資料本體具匿名性，其在特定條件下仍可能受到隱私法拘束，亦可能引發刪除特定訓練資料的法律義務。

第四目　小結

合成資料作為兼顧效用與隱私的資料替代技術，已逐漸嶄露於人工智慧與法規交錯的場域。基於其係人工生成而不包含真實資料的特性，其不僅能迴避傳統資料刪除的技術障礙，更提供在實務應用中具彈性與可控的資料生成方式。然而，合成資料並非萬能，正因其係由統計參數推導而來，導致該生成模型對特定用途會有效用不足之疑慮，並且在網路攻擊下亦可能有隱私洩露的風險，最後，合成資料究竟是否有個人資料相關法規的適用亦存在不確定性，要使其真正成為實踐資料刪除權的可信解方，仍需審慎因應其在效用與隱私保障上的潛在挑戰。

第三款　其他可能的技術因應措施

文獻上亦有其他因應資料刪除請求的方法，這些方法不直接處理原始模型，而是引入額外路徑（side paths）來改變模型，雖然他們原本並非為實現刪除權而設計，但可能成為處理其相關問題的潛在方案。

---

[231] Michal S. Gal & Orla Lynskey, *Synthetic Data: Legal Implications of the Data-Generation Revolution*, 109 IOWA LAW REVIEW 1087, 1126 (2024).

[232] 如 2018 年加州《消費者隱私法》（California Consumer Privacy Act，簡稱 CCPA，2021 年經修正）對「個人資訊」的定義採取開放式架構，界定為「有可能與某人產生關聯，或在合理情況下可被連結的資訊」。

[233] Khaled El Emam et al., *Evaluating Identity Disclosure Risk in Fully Synthetic Health Data: Model Development and Validation*, 22(11) J. MED. INTERNET RES. 736, 737. (2020).

doi:10.6342/NTU202600641

第一目　模型編輯

模型編輯（Model editing）是在模型訓練完成後，對模型的知識進行「定點手術」[234]。換言之，開發者針對特定資訊（例如某個人的資料）對模型參數做小範圍的調整或添加一個「修補補丁」，這種方法不需要重新訓練整個模型，而是在保持原始模型不變的前提下，額外儲存修改內容[235]。

然而，模型編輯作為一種僅針對特定輸出進行局部調整的技術，其本質上未必能滿足個人資料保護法對刪除請求所要求的嚴格標準。雖然透過模型編輯，模型在面對特定查詢時可能不再回應原本包含個資的內容，表面上似已達成「刪除效果」，但實際上，原始資訊仍可能潛藏於模型的訓練參數之中。就法律上的「刪除」定義而言，關鍵在於資料是否仍存於系統內；而模型編輯多僅為壓抑或覆蓋資訊，並未真正將其從模型記憶中抹除。故若從嚴詮釋資料保護法規，要求證明模型編輯後已完全去除某人的個資，仍具實質挑戰。

再者，模型編輯技術亦面臨定位與修改個資所對應參數的困難。大型模型中的知識通常高度分散，要精確識別與特定個人資訊（如電話號碼、地址）對應的參數並非易事。即便成功調整部分參數，亦可能出現所謂「下游知識不一致」問題：即模型在一種問法下不再提供特定個資，但在另一種語句或語境中，卻仍可能洩漏相關資訊[236]。這說明模型編輯未必能涵蓋所有潛在提問情境，甚至可能導致模型對同一資訊出現前後矛盾的回應。

綜上所述，考量現階段模型編輯在技術與法律合規上的諸多挑戰，本文認為該技術尚難作為人工智慧模型因應個人資料刪除權請求的可行解方。

---

234 *See* Eric Mitchell et al., *Memory-Based Model Editing at Scale*, 162 PROC. 39TH INT'L CONF. MACH. LEARNING 15817, 15817–31 (2022).
235 *See* Dawen Zhang *et al.*, *supra* note 170, at 7.
236 *Id.*

doi:10.6342/NTU202600641

第二目　輸出防護措施

亦有論者提出輸出防護措施（Guardrails）的因應方式，輸出防護措施指的是在模型輸出階段加裝的安全欄杆或規則，用來限制模型產生某些內容，這種方法不改變模型本身，而是在模型回答之前或之後進行額外的過濾與引導[237]。實例包括：在提示詞中加入指令要求模型不要談及某人的資訊，或在模型產生文本後用程式自動檢查並遮罩/刪除涉及該個人的內容[238]，換言之，輸出防護措施就像為模型的對話加上規則，當偵測到敏感個資相關的問答時，強制模型拒答或給出模糊回應。

然而，由於輸出防護措施並沒有動用到模型內部的訓練參數，受保護的個人資料仍隱含在模型之中，輸出防護措施只是確保模型「不說出」那些資訊，因此在法律上不等於真正刪除了資料，實難以符合個人資料刪除權所要求的「資料刪除」，從而僅能視為一種暫時性的迴避措施[239]。

第四款　小結

綜合而言，現階段針對「刪除權」的技術因應方案可大致分為三類：機器遺忘技術、合成資料替代方案與模型行為控制機制（如模型編輯與輸出防護）。其中，機器遺忘雖最貼近法規對「刪除」的技術要求，但面臨效能與可驗證性挑戰；合成資料則具高靈活性與法遵潛力，然仍存辨識風險與法規適用爭議；模型編輯與輸出防護措施等策略雖能快速應對刪除請求，但多屬權宜之計，難以單獨符合法律標準。未來若欲實踐數位時代中的「刪除」，仍須兼顧技術可行性與法律規範性，並建構更完備的驗證與治理機制。

---

[237] *See* Traian Rebedea et al., *Nemo Guardrails: A Toolkit for Controllable and Safe LLM Applications with Programmable Rails*, in PROC. 2023 CONF. EMPIRICAL METHODS NAT. LANGUAGE PROCESSING: SYST. DEMONSTRATIONS 431, 431-32 (2023).

[238] Dawen Zhang *et al.*, *supra* note 170, at 7.

[239] *Id.*

doi:10.6342/NTU202600641

第四節　小結——人工智慧發展與個人資料刪除權之衝突與平衡

本章首先回顧我國個人資料刪除權之制度設計，隨著歐盟、美國與日本等國陸續將刪除權明文化，我國亦於《個人資料保護法》第 3 條第 5 款明定刪除權為資料主體之基本權利，並禁止預先拋棄或以契約限制。配合第 5 條誠信與比例原則，以及第 11 條、第 19 條等相關規定，現行法建立以「目的消失」、「期限屆滿」及「違法處理」為核心的刪除權適用情境，作為資料主體控制個資之機制。然而，此一設計是否足以回應多元利用情境下的需求，仍存疑義。

在 111 年憲判字第 13 號判決中，大法官就我國健保資料目的外利用的相關法制規定宣告違憲，並明確指出，即使個資之蒐集符合法定要件或經當事人同意，資料主體仍應享有事後請求刪除、停止或限制利用該資料之權利。雖然該見解係針對健保資料的特殊情境，且所揭示之「接近無限制刪除權」構想未必適用於所有個資利用場景，然其核心意旨在於確認資料主體對個資更全面的事後控制權，並引發如何劃定刪除權合理界限之討論。

置於當前人工智慧廣泛應用的金融情境觀之，該議題更顯複雜。現今 AI 模型廣泛運用於信用評分、詐欺偵測與行銷推薦等業務，金融機構對個人資料的依賴日益加深。加以深度學習技術本質上難以完全移除特定資料，模型架構高度仰賴原始訓練數據，一旦大規模行使刪除權，可能影響模型完整性與運作穩定，甚而削弱金融服務效能並阻礙 AI 技術發展[240]。因此，若直接賦予資料主體過度寬泛、甚至接近無限制的刪除權，恐難與此類高依賴度的技術環境相容。

此外，除了法律劃定刪除權行使的界限外，是否應允許當事人基於自身評估與利弊權衡，透過契約對其個資為一定程度之安排，亦屬值得思考的問題。現行個資

[240]有文獻指出，若過度擴張刪除權的範圍與內涵，會對 AI 發展設下過多的限制，除可能衝擊知識傳播與言論自由外，亦會影響用於訓練人工智慧的資料，因而削弱人工智慧的安全性、可信度與預測能力。參林勤富（2025），〈歐盟《人工智慧法》之制度設計、規範內涵與治理侷限〉，《中研院法學期刊》，第 36 期，頁 27。

doi:10.6342/NTU202600641

法第 3 條第 5 款雖明文禁止預先拋棄或以特約限制刪除權，但第 11 條第 3 項但書與主管機關函釋卻容許資料主體以「書面同意」約定保存期限，作為延長資料保存的依據。此種規範設計不僅可能存在內在矛盾，亦凸顯其為值得深入檢討的議題。

綜上，當前制度面臨的核心挑戰，在於如何於保障個資刪除權與滿足技術實務需求之間取得適當平衡。本文認為，我國現行制度尚欠缺細緻的操作設計，除在資料主體保護上有不足外，亦未能有效因應 AI 技術在金融等高敏感度領域的挑戰。在憲法保障資訊自決權的前提下，未來法制發展應充分考量技術現實與產業需求，建構兼具法律正當性與技術可行性的資料治理機制。為此，後續章節將透過法律經濟分析與比較法探討，提出更具體可行的制度設計選項與創新方案，以在個人權益、技術發展與金融穩定性間達致平衡。

doi:10.6342/NTU202600641

第四章 從法律經濟分析觀點探求人工智慧發展與個資保護之平衡

隨著人工智慧技術在金融領域的廣泛應用，金融機構對於海量個人資料的依賴亦日益加深，無論是信用評分、風險控管抑或行銷推薦，均以資料為核心資源。此一現象進一步引發一重要問題：當金融機構希望長期保存並分析客戶資料以優化 AI 模型時，若客戶行使刪除權要求移除個資，將可能削弱模型效能，或使金融機構承擔額外合規成本，兩者目標因而產生衝突。

面對此一制度性矛盾，立法者如何在保障資訊隱私與促進技術創新之間取得適切平衡，已成為當前全球資料治理改革之重要課題。本文認為，為有效處理此類高度牴觸與資源配置牽涉的制度衝突，或可藉由法律經濟分析方法提供制度設計之可能方向。此一方法能跳脫傳統法學僅從權利保障與原則論述出發的視角，轉而納入交易成本、激勵結構與行為效果等經濟理性變項，有助立法者從「制度運作的實際效果」出發，設計出更具可行性與彈性的規範模型。

本章將嘗試結合寇斯定理對交易成本與權利配置的分析、新古典經濟學追求效率極大化的邏輯架構，並輔以行為經濟學對個體決策偏誤（如過度樂觀偏誤、現狀偏誤等）的實證洞察，建構一套更貼近資料實務與技術限制的刪除權制度模型。透過此一跨領域的分析架構，期能為現行個資法的修正方向提供更具操作性與平衡性的理論依據。

第一節　從寇斯定理檢視個人資料刪除權與金融資訊之交易市場

第一項　寇斯定理的基本內涵及其在金融資訊市場之意義

寇斯定理（Coase Theorem）主張，當談判不存在任何交易成本（transaction cost）時，不論法律如何進行資源或權利的初始分配，當事人皆能透過協商重新分配該權

doi:10.6342/NTU202600641

利，使資源達到最有效率之利用狀態。換言之，只要市場交易能夠順利進行並達成協議，初始法律權利歸屬的差異並不影響最終資源運用的效率[241]。

具體而言，寇斯於其經典論文〈社會成本問題〉（The Problem of Social Cost）中提出，即便法院最初將權利判定給某一方，但在交易成本為零的前提下，雙方當事人仍可透過市場協商重新界定權利，使得資源配置最終達到效率極大化，完全獨立於最初的法律劃定結果，而僅由交易雙方利益比較所決定[242]。具體而言，當權利初始配置給某方時，若另一方對該資源的評價較高，則可透過付出適當對價取得該權利；反之，若初始權利方的評價較高，則資源繼續維持原始配置狀態[243]。因此，寇斯定理指出，資源配置之效率不受法律上初始賦權影響，而取決於市場的後續協商結果。

第二項　寇斯定理於金融資訊刪除權之適用與限制

將寇斯定理套用至金融資訊之領域，意味著無論法律制度如何設計金融資訊的初始賦權，只要交易成本為零，最終結果皆能達到資源有效利用之目標。例如倘若法律制度將刪除權之初始賦權置於資料主體（預設得刪除），當金融機構對留存特定金融資訊之評價高於資料主體對隱私之評價時，機構得以提供對價（例如優惠之費率或服務條件）換取同意留存；反之，若資料主體之隱私評價更高，則可請求刪除，該資訊即不被留存[244]。

反過來，若將初始賦權置於金融機構（預設得留存），當資料主體對刪除之價值評價較高時，得透過支付成本或接受相對不利之服務條件以行使刪除；若機構之評價更高，則資訊持續留存。是以，在上述條件成立時，最終之留存／刪除狀態係

241 Ronald H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 837, 841–843 (1960).
242 *Id.*
243 *Id.* at 843.
244 *See* Jeff Sovern, *Opting in, Opting out, or No Options at All: The Fight for Control of Personal Information*, 74(4) WASHINGTON LAW REVIEW 1033, 1067-68 (1999).

doi:10.6342/NTU202600641

由交易雙方之相對估值與自願協商所決定，而非由初始權利配置所主導，故初始賦權並不影響資源配置之效率性[245]。

關鍵在於，寇斯定理之成立係以「市場交易成本為零」為其理論前提。然現實條件下，金融資訊隱私權之市場交易必然伴隨諸多交易成本[246]，此主要表現於資訊不對稱與協商困難等面向。例如金融消費者往往無法充分掌握其個人資料使用之範圍與經濟價值，導致難以作出符合其真實偏好的交易決策[247]；同時，金融機構欲與眾多消費者逐一協商個別資料處理條件，亦須付出高昂之行政與談判成本。是以，金融資訊市場中交易成本之普遍存在，使得資源配置難以如寇斯定理所預設般自然而然達到效率極大化。

此外，即使交易市場得以存在，仍可能因外部性、公共財特性、資訊不對稱與壟斷等因素導致市場失靈之現象[248]。以消費者信用評分為例：倘若大量資料主體行使刪除權，導致信用歷史片段化與不完整，放貸人將面臨更嚴重的資訊不對稱與逆選擇，進而提高利率或採取信貸配給，造成市場效率下滑[249]。

綜上，現實條件下，單純依賴市場力量以調整金融資訊權之利用與刪除權之行使，難以充分矯正各類市場失靈問題。故本文擬進一步引入新古典經濟學對市場結構與失靈類型之分析，以作為後續制度設計之理論依據。

---

245 *Id*.
246 Coase, *supra* note 241, at 854.
247 Alessandro Acquisti, Curtis Taylor & Liad Wagman, *The Economics of Privacy*, 54 J. ECON. LITERATURE 442, 445-450 (2016).
248 Joseph E. Stiglitz, *Markets, Market Failures, and Development*, 79 AM. ECON. REV. 197, 198-205 (1989).
249 *See* George A. Akerlof, *The Market for Lemons: Quality Uncertainty and the Market Mechanism*, 84(3) QUARTERLY J. ECON. 488, 489-90 (1970).

doi:10.6342/NTU202600641

第二節　新古典經濟學觀點

第一項　新古典經濟學的理性人假設

新古典經濟學對市場運作的分析，建立在「理性人」假設之上，亦稱「理性選擇理論」（rational choice theory）[250]。此一假設將個體視為能夠有目的地作出行為選擇的決策者，並以此作為推導市場失靈理論與法律介入正當性的邏輯起點。

所謂理性人假設，包含三項基本觀察[251]：（一）個體具有穩定且既定的偏好，並持有無偏差的信念與期待；（二）其行為選擇係依據上述信念與偏好作出；（三）在多數情況下，行為決策以追求自身利益為主要導向。

基於此假設，新古典經濟學發展出「預期效用理論」（Expected Utility Theory），以解釋並預測個體的行為模式。理論主張，行為人會先評估某一行動的「預期價值」（expected value）為正或為負，作為採取或拒絕該行動的依據[252]。預期價值係由「預期效益」（expected benefit）扣除「預期成本」（expected cost）所得，僅當效益大於成本時，行為人才會傾向採取該行動[253]。在實際判斷過程中，理性個體會同時考量：（1）機率因素——按事件發生的可能性加權效益與成本[254]；（2）風險因素——將不確定性轉化為風險溢酬（risk premium）進行調整[255]；（3）時間因素——依適當貼現率計算未來效益與成本的現值，以反映時間價值（time value）[256]。

綜上所述，理性人的行為模式可歸納為以下五項特徵：（一）行為符合預期效用理論；（二）行為會因誘因（成本或效益）變動而改變；（三）行為具備內

---

250 ROBERT COOTER & THOMAS ULEN (2016), LAW AND ECONOMICS 50 (6th ed., 2016).

251 *See* Richard H. Thaler, *Behavioral Economics: Past, Present, and Future*, 106(7) AM. ECON. REV. 1577, 1578 (2016).

252 楊岳平，〈金融消費者保護法制與監理的法律經濟分析觀點—以行為經濟學的應用為中心〉，《月旦法學雜誌》，339 期，2023 年 8 月，頁 24、25。

253 同前註。

254 COOTER & ULEN, *supra* note 250, at 43-44.

255 *Id.* at 44-46.

256 *Id.* at 37.

doi:10.6342/NTU202600641

在一致性（internal consistency）；（四）行為以提升自身福祉為目標；（五）行為能有效實現其目的[257]。

第二項　新古典經濟學的契約自由原則與市場失靈理論

在理性人假設與預期效用理論的基礎上，新古典經濟學主張應盡可能尊重私人間的契約自由。其核心邏輯在於：行為人被視為能以一致的偏好與無偏的信念比較交易的預期效益與預期成本，並在效益大於成本時自發進入契約關係[258]。當此決策同時發生於契約的雙方時，即意味著存在一個雙邊盈餘可由自願交換予以實現；資源因而從主觀價值較低的用途流向較高者，達成配置效率[259]。因此，新古典經濟學理論體系普遍主張應尊重私人間之契約自由與交易選擇，避免法律不必要地介入或限制，在雙贏的前提下，契約之締結不僅有助於當事人個別利益之實現，亦促進整體社會福祉的提升[260]。

儘管新古典經濟學強調契約自由作為促進資源最適配置與提升社會總體福祉的重要機制，然其亦承認並非所有契約均必然達致該等效果。在特定情境下，個別契約不但無助於提升整體福祉，反可能導致效率損失，形成所謂之市場失靈（market failure）。根據新古典經濟學的分類，市場失靈主要可歸納為四種典型情形：壟斷、外部性、公共財與資訊不對稱[261]。

首先，壟斷（monopoly）係指市場中缺乏有效競爭，致使某一方擁有壓倒性議價權。此時，相對人因欠缺可替代選項，可能被迫接受對自身不利之交易條件，導致市場交易價格高於最適均衡點、交易數量低於社會最佳水準，整體資源配置遂偏離提升社會福祉之目標[262]。

---

[257] Cass R. Sunstein et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1488 (1998).
[258] STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 293 (2004).
[259] *Id.*
[260] 楊岳平，前揭註 252，頁 25、26。
[261] *See* COOTER & ULEN, *supra* note 250, at 38-42.
[262] *Id.* at 38-39.

doi:10.6342/NTU202600641

其次，外部性（externality）係指某一交易行為所產生之成本或利益，並非僅由交易雙方承擔或享有，而是外溢至第三人。以環境污染為例，當交易雙方僅基於自利動機評估自身利益與成本，忽略其行為對第三人所造成之外部成本（external cost），即可能導致交易行為雖對雙方有利，卻對整體社會構成損害。在此情形下，交易數量可能超過社會理想水準，產生過度交易的現象[263]。

此外，公共財（public goods）具有非敵對性（non-rivalry）與非排他性（non-exclusivity）之特性，意即一人對該財貨之消費不影響他人使用，且供給者難以排除未付費者之使用，國防即為一典型例證。由於公共財之這些性質，極易產生搭便車（free-riding）行為，削弱私人提供者的供給誘因，最終導致供給不足、交易不足，偏離最適資源配置之均衡狀態[264]。

最後，資訊不對稱（informational asymmetry）發生於交易雙方掌握之相關資訊存有顯著差距時。一方當事人如無法取得足以評估交易之效益、成本與整體價值之資訊，即難以做出符合自身最佳利益之決策。此種資訊不足可能導致劣質交易或逆向選擇（adverse selection），最終使契約無法有效提升個人或社會福祉[265]。

第三項　金融資料市場的市場失靈

金融資訊市場中，確實可能出現多種典型之市場失靈現象[266]。若該市場本身無法有效運作，不僅難以自行修正失靈所導致之資源錯配，更可能在持續交易過程中，促使市場競爭者（即金融機構）對個人隱私產生更深程度之侵害[267]。因此，倘若市場機制本身難以矯正其內部扭曲，則有必要透過法律介入以加以補正[268]。

---

[263] *See id.* at 39-40; 王文宇（2019），《探索商業智慧：契約與組織》，頁 67，元照。
[264] *See id.* at 41.
[265] *See id.* at 41-42.
[266] 翁清坤（2008），〈臺灣與美國金融機構分享客戶個人資料之法律界限〉，《輔仁法學》，35 期，頁 84。
[267] Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52(5) STAN L. REV. 1373, 1398 (2000).
[268] 王文宇（2019），〈金融法制與金融監理〉，王文宇（等著），《金融法》，頁 5，元照，十版。

doi:10.6342/NTU202600641

本文以下將依據新古典經濟學理論，分別說明金融資訊市場中可能出現之市場失靈類型，並著重於資訊不對稱與公共財等問題。相對而言，本文對「外部性」問題不擬作為分析重點，蓋在金融資訊市場中，無論是資訊濫用造成的風險擴散（負面外部性），或是資訊彙整帶來的整體社會效益（正面外部性），其效果往往與公共財下的非排他性及集體價值密切相關，並且亦與資訊不對稱下資訊價值未被正確反映的情形高度重疊。為避免重複分析，本文不另設專節討論外部性，而是在公共財與資訊不對稱的章節中一併加以說明。

本文亦相對較不著重於「壟斷」問題。從目前台灣金融業者的結構來看，銀行與保險公司等業者數量眾多[269]，不符合傳統壟斷條件——即單一業者控制市場或排除競爭者之態樣。循此觀察，我國金融資訊市場並無典型的單一業者壟斷現象。惟本文亦不排除資料掌握程度不平均分配與資訊平台之高度集中，可能使少數大型機構具備實質上的資料壟斷影響力，只是目前此一現象在我國金融機構間相對較不明顯。故本文雖不擬細究此議題，但仍將保留對資料壟斷可能風險的關注。

第一款　公共財特性

從經濟學角度觀之，個人資料若未經法律保護，往往具備非競爭性與不可排他性之公共財特性。亦即，資訊一旦產生並被傳播，其使用不會排擠他人（非競爭性），且若無法有效設限，其內容可被無償、無限次使用（不可排他性）[270]。

為矯正此一失靈，各國立法傾向將個資權利賦予資料主體，透過賦權（entitlement）來建立排他性。例如我國個資法第 11 條與第 19 條規定，資料主體對其個資享有控制權與排他權；歐盟 GDPR 透過第 6 條所建立之「同意原則」，以及第 17 條所規範之「刪除權」，明確將個人資料之控制與處分權利歸屬於資料

[269] 根據中央銀行的統計，截至 114 年 6 月底，台灣擁有眾多金融機構：目前有 39 家本國銀行、21 家壽險公司，還有 21 家產險公司，在我國約 2,300 萬人口的市場中顯得相對分散。參：中央銀行，《金融機構一覽表（民國 114 年 6 月底）》，https://www.cbc.gov.tw/public/data/ebookxls/wlist.pdf（最後瀏覽日： 08/10/2025）。

[270] Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117(7) HARV. L. REV. 2056, 2085 (2004).

doi:10.6342/NTU202600641

當事人本人[271]。此類規範相當於將個資視為具財產權性質之權利客體，得以排除未經授權之使用，並促使市場機制得以有效反映資料之真實經濟價值。

儘管資料主體依法享有其個人資料的控制權，但實務中仍可能因激勵不足而不願提供個資，特別是當其擔憂資料供給無法帶來對價，或產生被他人搭便車（free-riding）之損失時。以金融服務為例，當金融機構將大量個人資料彙整並投入資源建置大型資料庫與訓練 AI 模型後，其生成之服務在邊際成本極低的情況下，可同時提供多方使用而不互相排擠（非競爭性），且若缺乏嚴格限制，內容亦可被多方重複利用（不可排他性），呈現出高度的公共財特質[272]。此種供給困境若在市場中普遍存在，即構成典型的公共財市場失靈，不僅使金融資訊資料庫難以形成規模經濟，亦將限制資訊流通效率，最終削弱市場功能並損及社會福祉[273]。

第二款　資訊不對等與議價能力差異

在金融資訊市場中，資訊不對稱與因此產生的決策偏誤構成另一典型的市場失靈現象，並對個資保護與交易效率造成雙重衝擊。具體而言，一般消費者在進行金融服務選擇或同意個人資料利用之決定時，往往缺乏對資訊價值的完整認知，不僅無法精確評估自身資料對金融機構的價值，亦未必能預見同意特定資料之利用可能產生的長期風險[274]。例如個人可能難以理解其信用違約紀錄、交易習慣或財務

---

[271] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation), art. 6(1)(a), 17, 2016 O.J. (L 119) 1.

[272] *See* e.g. Ryan Calo, Digital Market Manipulation, 82(4) GEO. WASH. L. REV. 995, 1005 (2014); 就此類資料的搭便車情形，舉例而言，若金融機構依年齡、職業、收入與信用資料進行客戶分類，進而建立精準的風險評估模型，則全體客戶不論是否提供資料均能享受到提升後的產品適配性與服務效率。*See* Jeff Sovern, *supra* note 244 *at* 1048- 49; 金融監督管理委員會（2020），《金融科技發展路徑圖報告書》，頁 26。

[273] 呂承儒（2022），《論金融控股公司共同行銷之個人資料保護法制》，頁 46，國立臺灣大學法律學研究所碩士論文。

[274] Peter P. Swire, *Markets, Self-Regulation, and Government Enforcement in the Protection of Personal Information*, NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION, https://www.ntia.doc.gov/page/chapter-1-theory-markets-and-privacy (last visited Apr. 30, 2025).

doi:10.6342/NTU202600641

敏感資訊，在未來如何被資料分析模型運用於風險預測、利率設定，甚至商品推薦等商業行為。

更進一步而言，在金融資訊市場中，資訊供需關係相較傳統金融交易市場可能出現角色倒置。蓋金融機構雖於金融商品市場中屬於供給者，然於資訊市場中則為需求方，客戶方為供給者。理論上，雙方應處於對等談判地位協議資料之利用方式與條款，惟實務上金融機構掌握隱私政策設計權與資料處理流程之控制權，故客戶往往無從理解或干預，造成資訊權力的明顯失衡[275]。

在此不對等關係下，金融機構可能透過設計繁冗難懂的隱私政策，或使用技術性術語與模糊措辭，使消費者難以辨明其個人資料的實際流向與用途。美國實務即觀察到，企業普遍運用資訊不透明手段削弱用戶對資料運用的控制，甚至藉此降低刪除、限制處理或拒絕同意利用的可能性[276]。就此，我國雖已立法要求金融業須主動告知個資處理事項[277]，惟在實務操作中，多數客戶即使已被告知，實際上仍可能因資訊傳達不足、內容表述過於專業或複雜等原因，而未能真正掌握相關資訊，形成典型的資訊不對稱問題，使得客戶難以在充分知情的基礎上進行協商或行使拒絕之權利。

資訊不對稱不僅削弱消費者的選擇與議價能力，亦可能對市場整體運作產生系統性扭曲。其中一種典型的市場失靈現象即為逆選擇（adverse selection），係指在資訊不對稱的情況下，市場參與者中風險較高或品質較低的一方，因能隱匿自身劣勢而更有動機進入市場，反之品質較高或風險較低的一方則可能因條件不利而退出，導致市場平均品質下降並損及效率[278]。而檸檬均衡（lemons equilibrium）即為逆選擇的經典案例，源自 Akerlof 對二手車市場的分析：當市場中買方無法有效

[275] *See* Paul M. Schwartz, *Privacy and the Economics of Personal Health Care Information*, 76(1) TEX. L. REV. 1, 49 (1997).
[276] Brill, K. Elizabeth, *Privacy and Financial Institutions: Current Developments concerning the Gramm-Leach-Bliley Act of 1999*, 21 ANN. REV. BANKING L. 167, 187 (2002).
[277] 個人資料保護法第 8 條、個人資料保護法施行細則第 3 條。
[278] COOTER & ULEN, *supra* note 250, at 41.

doi:10.6342/NTU202600641

辨識產品品質時，將難以區分「好車」與「爛車」，最終導致高品質供給者退出市場，只剩下低品質商品存續，形成劣幣驅逐良幣的不利局面[279]。

同理，在金融資訊市場中，即便有金融機構願意提供較嚴格的隱私保護措施，但在資訊揭露與傳達機制不足下，消費者未必能理解此類保障對自身的實質價值，而僅以價格高低作為主要考量因素，導致隱私保障政策完善者喪失競爭力，反而是規範鬆散的業者主導市場[280]。進一步而言，當消費者無法清楚分辨不同金融機構的風險管理能力時，低風險、穩健經營的機構可能因成本與定價劣勢而失去客戶，市場份額反而集中於高風險、低成本的機構，最終引發整體隱私品質下滑並推升系統性風險，損及市場效率與社會福祉。

綜上所述，資訊不對稱、檸檬效應與逆向選擇三者在金融資訊市場中往往交織出現，共同反映市場機制在無法辨識資訊品質、傳遞風險信號與保障個資自主權時的系統性脆弱性。若無適當法制介入補正資訊落差、提升揭露透明度與設計更公平的預設規則，市場將無法有效內部化資訊價值，也難以確保交易公正與個資保障。

第四項　小結

新古典經濟學建基於理性人假設與預期效用理論，核心信念在於：只要市場具備充分競爭、資訊對稱與低交易成本，個體基於自利動機所做出的選擇，即可實現社會資源的最適配置。在此理論架構下，契約自由被視為市場運作與資源有效配置的基礎機制，因其可確保雙方當事人在資訊充分且理性衡量效益與成本的前提下，自主決定是否締結契約。契約自由不僅有助於實現個體福祉的最大化，亦被理解為整體社會福利提升的關鍵。也因此新古典理論傾向主張在無市場失靈的情況下，法律應尊重並維護當事人間之契約自主，避免不必要的干預。

---

[279] George A. Akerlof, *The Market for Lemons: Quality Uncertainty and the Market Mechanism*, 84(3) QUARTERLY J. ECON. 488, 489-90 (1970); Richard Craswell, *Property Rules in Unconscionability and Related Doctrines*, 60(1) U. CHI. L. REV. 1, 49 (1993).

[280] *See* Schwartz, *supra* note 270, at 2081.

doi:10.6342/NTU202600641

然而，當市場出現結構性扭曲的市場失靈現象時，契約自由所預設的理性選擇條件便不復存在。在金融資訊市場中，本文認為此類市場失靈主要表現於以下兩個面向，第一是個資彙整後具備外溢效益與非排他性的公共財性質，可能導致資訊供給不足與搭便車問題；其次，資訊高度不對稱將削弱個體作出符合自身最大利益決策之能力，並可能引發逆選擇現象，致使劣幣驅逐良幣，最終導致整體金融隱私服務品質下降。是故，契約自由雖為市場經濟核心原則，但其實踐基礎須依賴有效制度保障與資訊條件的完備，若欲使契約自由真正發揮提升福祉之功能，仍須回應市場失靈所揭示的制度缺口。

第三節　行為經濟學觀點

本節將轉向行為經濟學的視角，進一步檢視理性假設的侷限，以及人在現實決策情境中所展現的心理偏誤與制度互動，從而探索補足新古典經濟學模型之可能方向。行為經濟學有別於新古典經濟學之處，在於其突破傳統理性人假設，結合心理學與認知行為科學，具體揭示人類在決策過程中可能出現的可預測偏誤，從而提供法律制度設計一種更貼近實際人類行為的觀察視角。

第一項　行為經濟學對理性人假設的修正

行為經濟學立基於心理學與行為科學的研究成果，指出人類在實際進行選擇時，係透過兩種截然不同但並存的思考系統進行判斷與決策：其一為快速、直覺且自動運作的「系統一」，其二則為較為緩慢、理性且需投注認知資源的「系統二」[281]。一般而言，人類多數決策先由系統一迅速做出初步判斷，系統二則可能在事後介入進行修正。此種認知機制顯示，人在決策時往往先受到情緒與直覺驅動，再依理性邏輯進行調整，導致實際行為與完全理性模型產生偏離[282]。

[281] Riccardo Viale, *Understanding Financial Behaviour for Better Policy Making: An Introduction*, in THE BEHAVIOURAL FINANCE REVOLUTION 2, 7 (Riccardo Viale et al. eds., 2018).
[282] *Id.* at 10-12.

doi:10.6342/NTU202600641

基此，行為經濟學對新古典經濟學「理性人」假設提出深刻修正，指出真實世界中的個體存在三項結構性限制：有限理性（bounded rationality）、有限意志（bounded willpower）以及有限自利（bounded self-interest）[283]。這三項限制揭示了人類行為與傳統經濟模型間的落差，並成為重新檢討制度設計與法律規範的重要基礎，以下將分項論述此三項修正之核心內涵。

第一款　有限理性

有限理性一詞最早由 Herbert A. Simon 所提出，其核心觀點在於人類在實際決策過程中受到認知資源有限、時間與資訊限制等多重因素影響，往往難以做到完全資訊蒐集與最優分析。相對於新古典經濟學中「理性人」所預設之決策行為模式，Simon 認為人們更傾向於採取一種「滿足」（satisficing）的策略，亦即在搜尋到某一選項達到預設可接受的標準後，即停止搜尋並作出決定[284]。

在此一觀點下，個體並非以追求效用最大化為唯一目標，而是以降低決策成本與風險為考量，尋求「足夠好」而非「最完美」的選擇。Simon 將此行為模式視為人類理性在實務環境中可觀察之具體展現，亦成為行為經濟學對理性人假設提出修正的重要理論基礎之一[285]。

行為經濟學進一步歸納出有限理性於個體資訊蒐集、分析與評估過程中，常見之六種偏誤與限制如下：

（一）有限注意力（limited attention）：人在面對龐雜資訊時，往往無法同時處理所有相關面向，而僅能聚焦於部分訊息，忽略其他重要資訊。研究指出，行為人之選擇極易受限於當下情境中較為「顯著」的訊號，而忽略應被納入評估之其他要素，顯示注意力的分配本身即為決策結果的決定性因素[286]。

---

[283] Cass R. Sunstein et al., *supra* note 257, at 1476.
[284] *See* Herbert A. Simon, *A Behavioral Model of Rational Choice*, 69 Q.J. ECON. 99, 99-118 (1955).
[285] Cass R. Sunstein et al., *supra* note 257, at 1474-75.
[286] *See* WILLIAM J. CONGDON ET AL., POLICY AND CHOICE: PUBLIC FINANCE THROUGH THE LENS OF BEHAVIORAL ECONOMICS 22 (2011).

doi:10.6342/NTU202600641

（二）有限計算能力（limited computational capacity）：人類的資訊處理與推論能力有限，當面對多重替代方案與多層因素之選擇時，常難以全面計算所有後果與效用，最終僅能做出「尚可接受」而非最優化之決定[287]。

（三）論理偏誤（biased reasoning）：在評估不確定性與機率時，一般人常呈現出非理性的推論模式，難以符合統計學與邏輯原則。例如，人們可能高估罕見事件的發生機率，或誤以為兩事件具有關聯性，導致偏離預期效用理論所預測之行為模式[288]。

（四）顯著效應（salience effects）：決策過程中，個體傾向過度關注視覺上、情緒上或語意上特別顯眼之資訊，即便該等資訊與實質價值關聯有限，這使得不應被重視之特徵反而主導決策結果，進而影響選項評估之客觀性[289]。

（五）心理帳戶（mental accounting）：人們經常傾向將財務資源依其來源、用途或形式進行非理性分類，並基於這類「帳戶」進行區隔性決策，而非綜合考量個人整體財務狀況。例如，人們對「意外之財」之花費態度，往往顯著不同於其對固定收入之使用方式[290]。

（六）過度自信（overconfidence）：人類普遍傾向高估自身知識的準確性或預測能力，並低估未來結果的不確定性，進而做出風險偏高之行為。與此相類的偏誤尚包括過度樂觀（over-optimism），其使人低估損失或風險之可能性[291]。

綜上，有限理性不僅挑戰了傳統經濟學預設之「完全理性決策者」假設，更揭示人類在實務決策中如何受到內在機制與心理偏誤的制約。進一步言，有限理性之影響則可分別於「判斷」與「決策」兩層面觀察之。

---

[287] *Id.* at 23-24.
[288] *Id.* at 26-27.
[289] *Id.* at 22-23.
[290] *Id.* at 25-26.
[291] *Id.* at 27.

doi:10.6342/NTU202600641

在判斷層次中，個體往往傾向依賴簡化之思考策略，即捷思法（heuristics）或經驗法則（rules of thumb），以因應複雜情境下認知資源之不足[292]。此類捷思法雖具實用性，能在多數情境中快速產生合理判斷，惟亦可能在特定條件下產生系統性錯誤。例如，個體通常依賴自身過往是否曾有類似經驗，或以該事件是否與某些具代表性、類似性或易於喚起的事件相關聯，作為判斷的基礎[293]。例如當個人欲評估未來發生洪災的風險時，其評估標準往往不基於統計資料，而取決於自身是否曾經歷洪災、是否曾聽聞類似事件，或是媒體中是否近期出現相關報導。當人們對某事件具有直接或間接的切身經驗時，更傾向判斷該事件的發生機率為偏高，這正是可得性捷思（availability heuristic）所導致的系統性偏誤。此現象顯示，即便捷思法可視為節省思考成本的理性應對，其結果仍可能與無偏誤之預測模型偏離[294]。

在決策層面，有限理性表現在行為人對損益的主觀評價與傳統預期效用理論之預測存有差異[295]。就此，行為經濟學提出展望理論（the prospect theory）解釋之，展望理論針對此種偏離進行修正，指出人們對於潛在獲益與損失並非以相同心理權重評估，尤其在面對損失時，行為人展現出顯著之損失怯避（loss aversion）[296]。舉例而言，正常人在失去一項財產時感受到的痛苦，往往超過獲得相同財產所帶來的滿足，此現象反映於行為人對既有財產賦予較高主觀價值，從而導致對損失的迴避心理遠大於對獲益的追求[297]。

---

[292] Cass R. Sunstein, *The Storrs Lectures: Behavioral Economics and Paternalism*, 122(7) YALE L.J. 1826, 1851 (2013); Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 5(2) CODNITIVE PSYCHOLOGY 207, 220 (1973).

[293] Colin F. Camerer et al., *Behavioral Economics: Past, Present, Future*, in ADVANCES IN BEHAVIORAL ECONOMICS 1, 10 (Colin F. Camerer et al. eds., 2004).

[294] Daniel Kahneman & Amos Tversky, *Judgment under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124, 1127–29 (1974)

[295] Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50(5) STAN. L. REV. 1471, 1477-78 (1998).

[296] Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47(2) ECONOMETRICA 263, 277-80 (1979).

[297] Amos Tversky & Daniel Kahneman, *Loss Aversion in Riskless Choice: A Reference-Dependent Model*, 106 Q. J. ECON. 1039, 1047 (1991)

doi:10.6342/NTU202600641

而稟賦效應（endowment effect）可被視為損失怯避心理在實際行為中的具體表現[298]。當個體在評估與某項物品相關的交易行為時，傾向將「失去」其既有物品視為損失，而將「取得」他人擁有之物視為獲益，則此一心理傾向將使潛在賣方對其持有物品賦予較高的主觀價值，從而要求潛在買方支付更高的價格[299]。此一現象將造成買賣雙方對同一物品的估價出現顯著落差，抑制交易意願與交易數量，進而導致市場上交易量減少、資源配置無法達到效率化的結果[300]。

此一偏誤顯示預期效用理論無法充分描述真實的決策過程，而展望理論則提供更符合實證觀察的替代模型，揭示人類決策中存在心理參照點、損失與效益評價非對稱性等特徵[301]。

第二款　有限意志

有限意志一詞旨在說明個體在面對自身長期利益時，常出現明知其行動有悖自身福祉，卻仍難以堅持理性選擇之現象[302]。即使行為人已透過預期效用之計算，得出理論上最符合利益之選項，卻往往因意志力不足、短期衝動或情緒干擾，而未能實際採取該行動[303]。此種現象不僅削弱新古典經濟學中預設的行為一致性，亦突顯行為人之內在衝突與決策不穩定性。

具體而言，有限意志之核心問題在於個體可能在「計畫階段」已明確設定合乎理性的目標，但在「執行階段」卻無法如願實行，形成典型的時間不一致行為[304]。此種現象常見於生活中對健康、財務或職業規劃之行為，例如，個體雖明知應節制飲食、儲蓄金錢或按時完成任務，但在面對即時誘惑時仍選擇延宕或放棄。

---

298 Daniel Kahneman et al., *Experimental Tests of the Endowment Effect and the Coase Theorem*, ADVANCES IN BEHAVIORAL ECONOMICS 1325, 1326-31 (Colin F. Camerer et al. eds., 2004).
299 *Id.*
300 *Id.*
301 Cass R. Sunstein et al., *supra* note 257, at 1478.
302 *Id.* at 1479.
303 CONGDON ET AL., *supra* note 286, at 28.
304 *Id.*

doi:10.6342/NTU202600641

研究將此類意志不足歸納為下列幾種常見表現形式：

（一）拖延與誘惑（Procrastination and Temptation）：行為人往往無法抗拒即時滿足所帶來的誘惑，而延後執行需耗費努力之行動，導致短期偏好凌駕於長期利益之上，呈現所謂「現時偏誤」（present bias）[305]。

（二）情境狀態與情感波動（State and Affect）：個體的意志與決策表現高度受當下情緒與生理狀態影響，諸如焦慮、疲倦、飢餓等因素，皆可能削弱行為人對原先承諾或目標之堅持[306]。

（三）成癮行為（Addiction）：當行為人長期習慣特定滿足性刺激（如酗酒、購物、網路使用等），將進一步削弱其自我控制能力，陷入反覆自我否定與失控行為的循環[307]。

Congdon 等學者指出，有限意志造成之行為偏差，並非單純意志力「不足」所致，而是反映出人類決策過程中對於未來利益的系統性低估與當下衝動的過度重視[308]。因此，面對此一現象，若僅要求行為人「自我負責」並不足夠，更應思考如何透過制度設計引導行為人來提升行為一致性與長期目標之實踐可能性。

第三款　有限自利

新古典經濟學假設人類行為係以極大化自身福祉為唯一考量，亦即行為人僅基於個人效用進行選擇。然而，行為經濟學所揭示之「有限自利」理論則指出，個體在實務情境中並非總是以自我利益為最高依歸，而是常常考量他人、群體或社會價值等非自利性因素，導致其決策結果未必符合經濟理性所預測之最有利選項[309]。在行為經濟學中，「有限自利」可具體表現在下列幾類現象：

---

[305] *Id.* at 29.
[306] *Id.* at 30.
[307] *Id.* at 31.
[308] *Id.* at 29-31.
[309] *Id.* at 36.

doi:10.6342/NTU202600641

（一） 利他主義（Altruism）：指人們在決策時，除追求個人利益外，亦可能出於同理心與道德動機，選擇對他人有益之作為，即使自身無明顯受益甚至需承擔成本[310]。

（二） 公平性（Fairness）：指行為人除了重視結果本身是否有利，亦會在意結果產生的「過程」是否公平，舉例而言，一般人傾向認同透過合作、互惠或合理程序產生的結果，並對不公平手段或不當分配表達抗拒[311]。

（三） 社會規範（Social Norms）：指人在行為選擇時，傾向遵循社會共識與文化期望，即便違背社會規範不會受到法律或經濟上的明確制裁，亦會因心理不適或社會壓力而選擇順從[312]。

（四） 人際偏好（Interpersonal Preferences）：個體在做出選擇時，常會考量他人對自身行為的看法，或在意其於群體中所處的相對地位。例如，在社交情境中，行為人可能因「面子」因素而選擇非最有利自己的行為，以維持形象或社會認同[313]。

總結而言，有限自利理論打破了傳統經濟學對『人皆自利』的單一假設，指出人們在做決策時，除了考量自身利益，也會受到公平性、社會規範與他人觀感等因素影響。

第二項　行為經濟學與法制設計

行為經濟學修正了新古典經濟學中「理性人」的假設，指出人類在實際決策時往往受到有限理性、有限意志與有限自利等因素影響。是以，當法律面對當事人間基於契約自由所產生之選擇行為時，若不適時以法律介入，可能放任個體因行為偏

310 *Id.* at 37.
311 *Id.*
312 *Id.* at 37-38.
313 *Id.* at 38.

doi:10.6342/NTU202600641

誤而損害自身利益，進而損及整體社會福祉。因此，行為經濟學主張，法制應透過適度介入以矯正偏誤，引導決策回歸符合公共利益與個體福祉之方向。

第一款　「輕推」作為行為導引工具

行為經濟學提出「輕推」（nudge）作為一種非強制性、低侵擾性的法制介入方式，其核心在於微調選擇架構，藉此影響決策方向[314]。輕推不涉及禁止選項或重大經濟誘因的改變，卻能有效誘導個體克服甚至利用上述行為偏誤，做出較符合理性人偏好之選擇[315]。常見的輕推措施包括：警示標語、預設規則設計、資訊顯著性提升，以及選項排列順序優化等[316]。

其中，資訊揭露的「方式」與「呈現形式」被認為與揭露義務本身一樣重要[317]。由於一般人常因有限注意與計算能力而傾向倚賴捷思法（如可得性捷思、定錨效應）做出快速決策，若揭露內容繁複難解，將難以實質影響選擇[318]。因此，資訊應以顯著、簡明且易於理解的方式呈現，方能有效納入決策考量[319]。

此外，針對任意規定（default rules）的設計，行為經濟學與新古典經濟學觀點亦存重大分歧。後者認為既允許當事人退出，預設內容便不具實質意涵；然行為經濟學指出，多數人因慣性、拖延、現狀偏誤或損失怯避等心理偏誤，往往傾向保留原有預設。是故，若立法者將較優選項設為預設規則，即能在不剝奪選擇自由下，有效導引行為人做出較佳決策。

第二款　自由的家父長主義：引導與干預之間的法制設計平衡

---

314 RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS 6 (Revised & Expanded ed. 2009).
315 *Id.*
316 Ryan Bubb & Richard H.P., *How Behavioral Economics Trims Its Sails and Why*, 127 HARV. L. REV. 1593, at n.2 (2014).
317 Cass R. Sunstein et al., *supra* note 257, at 1533-34.
318 楊岳平，前揭註 252，頁 31。
319 同前註。

doi:10.6342/NTU202600641

「自由的家父長主義」（libertarian paternalism）可視為「輕推」制度精神的核心展現，其基本理念在於：在不剝奪選項、不強制行為的前提下，透過制度設計引導個體朝向更有利的選擇[320]。此類制度設計融合了兩項看似矛盾的特性：一方面尊重個人自主、保留選擇與退出權利；另一方面則承認政策制定者可藉由設計選擇架構介入行為模式，協助個體克服行為偏誤，做出更符合其長期福祉的決策[321]。

然而，輕推雖標榜自由，仍帶有典型家父長主義色彩，其設計出發點通常在於保護選擇者本身的利益，而非為防止外部負面效果，故實質上仍反映出立法者欲「代替行為人做出較佳選擇」的預設立場[322]。更進一步而言，政策制定者於設定制度內容時，往往需預設「理性人應當具備之偏好方向」，據此擬定預設選項與提示策略，試圖誘導多數人朝該方向決策，這種「以國家意志指引私人選擇」的制度邏輯，自難完全排除父權主義的影子[323]。

然而，「輕推」制度雖具導引性質，其核心精神仍係根植於自由主義，重視並維護個體的選擇權限。無論是在資訊揭露的方式設計上，或是任意規定的設定上，其本質皆有別於強制性規範，亦即立法者所調整的僅是行為人所處之選擇架構，而非直接替代其決策行為，最終決定權仍掌握於行為人自身[324]。

是故，「自由的家父長主義」雖可視為一種軟性家父長主義（soft paternalism），透過如調整選項排序、設置預設選項、導入冷靜期等低侵擾性措施以促進公共利益與個體福祉，但此類制度仍須嚴守三項基本原則，方能維持其正當性與自由性，包括：（1）不得剝奪選擇自由；（2）不得強制特定行為；（3）不得實質扭曲其他

320 Le Grand & New, *supra* note 241, at 133.
321 Richard H. Thaler et al., *Choice Architecture*, *in* THE BEHAVIORAL FOUNDATIONS OF PUBLIC POLICY 428, 428-429 (Eldar Shefir ed. 2013).
322 Thomas A. Lambert, *From Gadfly to Nudge: The Genesis of Libertarian Paternalism*, 82 MO. L. REV. 623, 640 (2017).
323 THALER & SUNSTEIN, *supra* note 314, at 5.
324 Lambert, *supra* note 322, at 640.

doi:10.6342/NTU202600641

選項之經濟誘因[325]。唯有在此基礎下，方能實踐「引導而非命令」之治理理念，避免輕推機制流於隱性強制的合理化[326]。

此一制度設計邏輯，反映出行為經濟學對人類行為偏誤的務實回應——既然選擇偏誤可能源於選擇架構的不當設計，則透過架構之微調即有可能協助行為人回歸理性選擇軌道。當行為人能於矯正後的環境中作出較為審慎、符合其自身利益的選擇時，法律即無須進一步介入，而能重回尊重自主選擇之立場。

第四節　法律經濟觀點下的制度建議

第一項　人工智慧金融應用下個人資料刪除權界限之問題意識

承接第三章的制度檢視，我國現行《個資法》雖已將刪除權納入憲法保障的資訊自決權之一環，但其行使僅限於「目的消失」、「期限屆滿」及「違法處理」等情境。此種設計在概念上固可避免刪除權被過度濫用，然能否充分回應當前資訊爆炸時代下的隱私保障需求，仍有疑義。111 年憲判字第 13 號判決肯認資料主體即使曾同意或符合蒐集要件，仍應享有事後刪除與限制利用的權利。該見解雖僅針對健保資料制度，但已揭示刪除權趨近無限制的可能性，顯示立法者在設定刪除權邊界時，必須在權利保障與制度可行性之間取得平衡。

這項張力在人工智慧的金融應用場景中尤為顯著。當前 AI 模型已廣泛運用於信用評分、詐欺偵測、行銷推薦等各類金融服務，其深度學習架構高度依賴原始訓練數據，一旦大規模行使刪除權，可能削弱模型完整性與服務穩定性，甚至阻礙技術迭代與創新。由此可見，刪除權的合理界限並非僅屬法律解釋問題，更涉及技術可行性與產業運作的整體平衡。

在此背景下，值得進一步探問的是：在肯認資料主體享有較寬廣刪除權的同時，是否可容許其透過契約與金融機構就其刪除權作出一定程度的限制？現行個資法

325 THALER & SUNSTEIN, *supra* note 314, at 6.
326 *Id.*

doi:10.6342/NTU202600641

第 3 條第 5 款固然禁止預先拋棄或以特約限制刪除權，但第 11 條第 3 項但書卻允許以「書面同意」約定個資保存期限，作為延長保存的依據，顯示現行制度或有矛盾之處，亦凸顯該問題之重要性。

因此，接續本章前述的法律經濟分析，本文將從寇斯定理、新古典經濟學與行為經濟學的視角出發，探討此張力的制度化處理可能性：是否能透過嚴謹的程序與實質要件設計，使契約限制成為在資訊充分、意思表示真實且交易條件合理的情境下運作的輔助性工具，以兼顧資料主體的權益保障與金融市場的穩定性。

第二項　從寇斯定理到行為經濟學：資料刪除權之理論重構

寇斯定理指出，在交易成本為零的理想市場中，權利的初始配置不影響資源的最終配置效率。換言之，無論個人資料的使用權初始賦予資料主體或金融機構，只要雙方能自由協商，就能透過權利移轉或補償安排，使資料流向最重視其價值的一方，達致帕雷托最適的資源配置。此一邏輯亦可由新古典經濟學中「理性人假設」推導：假設個體具備完整資訊、穩定偏好與計算能力，能理性評估利害關係，進而作出最符合自身利益的選擇。

然而，寇斯定理與新古典經濟學的前提在現實中難以完全成立。首先，就資訊不對稱而言，資料主體往往無法精確掌握自身資料的真實價值，亦難以預測同意資料利用後可能衍生的長期風險。在與金融機構互動的情境中，多數資料主體因缺乏談判能力、技術知識與法律支援，往往只能被動接受定型化條款，形式上雖有契約自由，實質上卻無法確保權益受到充分保障，導致資料刪除權在實務上可能流於形式。若此情況長期存在，可能弱化市場對隱私保護的競爭動能，進而出現「劣幣驅逐良幣」的現象，使整體金融服務的隱私品質逐步下滑。

其次，個人資料亦具有準公共財的特性，兼具非競爭性與不可排他性。一旦大量個資被金融機構整合建置為資料庫，即能產生顯著的正向外部性，例如促進精準風險模型建構、提升金融普惠性與服務效率，並改善市場資訊流通與資源配置。然

doi:10.6342/NTU202600641

而，若資料主體因對價不足或搭便車問題而傾向拒絕同意資料利用，將可能造成整體資料供給不足，限制資料庫形成與市場創新動能。

在此脈絡下，行為經濟學進一步指出，即便制度上賦予資料主體選擇空間，其決策過程仍可能受限於損失怯避、拖延心理、選項過載與有限理性等行為偏誤，導致其無法有效評估限制刪除權對長期利益的影響。換言之，資料主體的表面「同意」並不必然代表真正自主且理性的選擇，亦可能造成對自身權益的不當犧牲。

綜合上述經濟分析，儘管新古典經濟學指出應開放自由訂立契約以達到最適資源配置，資料刪除權的制度設計實際上仍面臨兩項不同性質的挑戰。第一，受限於資訊不對稱與行為偏誤，資料主體往往難以在充分資訊與理性判斷的基礎上作出是否限制刪除權的決定。根據行為經濟學的觀察，有限理性、現狀偏誤、損失怯避等心理因素，會使當事人即便享有選擇權，仍可能被動接受對自身不利的條款，導致刪除權流於形式化保障。對此，制度設計宜運用輕推原則，透過顯著且易懂的資訊揭露、優化預設規則與設置冷靜期等方式，矯正資訊落差與認知偏誤，確保資料主體能基於實質自主作出選擇。

第二，由於個人資料具有準公共財的特性，市場上可能出現搭便車問題：當資料主體對資料利用缺乏足夠的對價或誘因時，可能選擇不授權，造成資料供給不足，進而限制 AI 模型的訓練與創新應用。針對此一面向，或可透過適度的經濟誘因與契約安排，在不侵蝕資訊自決權的前提下，引導資料主體願意提供資料，兼顧市場運作的穩定性與技術發展的需求。

在此架構下，有限度的契約自由可與行為經濟學的輕推工具並行運作——前者用於確保資料利用的可持續性與供給效率，後者則保障決策的真實性與自主性，最終在隱私權保障與產業發展之間取得平衡。在此理論啟發的基礎上，本文下節將進一步檢視我國現行法制是否已回應上述兩大問題。

doi:10.6342/NTU202600641

第三項　從法律經濟分析觀點看我國個人資料刪除法制

我國《個資法》第 11 條第 3、4 項與第 19 條保障資料主體對其個資享有刪除權，並於同條文中進一步規範資料主體於特定情境下方享有請求刪除個資之權利。然而，透過前述對經濟學理論的分析，本文認為現行制度架構於實務應用中可能呈現若干侷限，尤其難以回應人工智慧與金融科技應用下對資料流動與穩定性的合理需求，以下說明之。

第一款　資料刪除權契約彈性的可能與侷限

如前所述，新古典經濟學強調契約自由是促進資源最適配置與提升社會總體福祉的重要機制，法律原則上應尊重私人間的契約安排與交易選擇，避免不必要的介入。然而，我國現行《個資法》第 3 條第 5 款卻明文禁止資料主體預先拋棄或以特約限制其刪除權，致使即便雙方在資訊對等且自願同意的情況下，亦不得就刪除權作一定程度的約定。

值得注意的是，《個資法》第 11 條第 2 項但書對刪除權另設例外，規定若個資「因執行職務或業務所必須，或經當事人書面同意者」，得不受「目的消失或期限屆滿即應刪除」之限制。《施行細則》第 21 條第 1 款進一步解釋，「因執行職務或業務所必須」包括「有法令規定或契約約定之保存期限」；國家發展委員會亦於函釋中指出，只要經當事人書面同意，即可在目的達成或保存期限屆滿後繼續保留個資，但此同意仍須符合第 5 條所揭示的誠信與比例原則，不得以籠統或無限期方式規避刪除義務[327]。

由此觀之，雖第 3 條第 5 款形式上禁止限制刪除權，但透過第 11 條第 2 項但書、《施行細則》與行政機關解釋，實務上資料主體仍可透過書面契約約定合理的資料保存期限，達到延緩刪除權行使的效果。因此，我國法並非全然排除資料刪除權的契約限制可能性，反而透過例外條款、施行細則與行政解釋，內建了一定程度

[327] 國家發展委員會發法字第 1100016400 號函。

doi:10.6342/NTU202600641

的制度彈性，使金融機構與資料主體得以約定合理的資料保存期限，兼顧業務需求與個資利用的穩定性。

然而，此種「明文禁止、實質開放」的制度設計，易使人民誤解法律意旨，實應修法調整之。更重要的是，當事人在與大型金融機構互動時，往往面臨資訊不對等、談判能力不足，或受限於行為偏誤（如現狀偏誤、損失怯避等）而傾向接受預設方案，致使其形式上雖有同意，實質保障卻未必充分。憲法法庭於 111 年憲判字第 13 號判決中，即就健保資料的刪除強調，即使資料係經當事人同意或符合法定要件取得，仍應賦予事後刪除與限制利用的權利，從其立場或可反映部分見解對現行例外條款保障不足的疑慮。

此情形正呼應行為經濟學對契約自由的反思──在資訊不對稱與認知偏誤存在的條件下，單純依賴形式同意，未必能確保當事人真正基於理性與充分資訊做出選擇。因此，若要在現行法的彈性空間中兼顧契約自由與權益保障，制度設計上仍有必要透過行為經濟學的「輕推」措施，以矯正偏誤、強化資料主體的實質自主性。

第二款　金融資料分享之對價設計與市場效率

新古典經濟學主張透過契約自由促進資源最適配置，然而，市場失靈現象仍可能造成效率損失。在我國金融資訊市場中，除了前述資訊不對稱的問題外，由大量資料訓練而成的 AI 模型，因具備非競爭性與不可排他性，呈現高度準公共財的特徵，易引發搭便車問題。具體而言，若資料主體在提供訓練資料時需承擔隱私風險，卻未獲得相應補償或對價，則其分享誘因將顯著降低，甚至可能因不願讓他人「免費受益」而拒絕貢獻資料。長期而言，將不利於金融資訊資料庫形成規模經濟，限制資訊流通效率，最終削弱市場功能並損及社會福祉。

觀察我國現行資訊法制，雖現行法在實務上允許資料主體與資料控制者就資料保存期限訂立契約，但對於資料提供行為是否應有對價或補償，並無明確規範。因此，實務上大多無以明確對價為基礎成立的「同意」，而是普遍透過形式化的同

doi:10.6342/NTU202600641

意程序達成契約效果[328]。此種「無對價化」的運作方式，雖能在形式上成立，但欠缺行為經濟學所強調的「突出性（salience）」——亦即讓當事人清楚意識到其行為的價值交換與實質成本。當對價缺乏顯著性時，資料主體往往無法充分評估分享資訊的真實效益與風險，導致誘因不足與選擇偏差[329]。此缺漏不僅可能限制資料主體依個人偏好與利益進行彈性安排，也壓抑了願意以資訊交換回饋者的選擇自由，使制度與市場實際運作產生落差。

依新古典經濟學的預期效用理論，個體在行為決策時會比較行動的預期效益與預期成本，作為採取或拒絕行動的依據[330]。若結合前述經濟學的觀察，在金融資訊場域中，雖提供敏感性個資可能換取更精準的金融服務，但其潛在的隱私風險與準公共財式的搭便車問題，往往降低資料主體的分享意願。因此，若要緩解市場失靈，制度設計應在於提高「提供資料」的淨效益或降低其成本。一方面，可引入明確且具顯著性的對價機制（如現金補償、優惠利率、服務升級），使交換價值對資料主體清晰可見；另一方面，亦可透過將個資提供與進階金融服務資格連結，賦予 AI 模型服務一定排他性（惟須避免不當歧視）。如此，既能在現行法未禁止對價的框架內提升參與誘因，也可兼顧契約自由與市場效率。

第三款　行為偏誤下刪除權契約設計之操作工具

為維護金融服務的穩定與 AI 技術的發展，同時保障資料主體得以依個人意願對其個資作最適安排，本文前文已論述兩項核心方向：其一，在現行法下適度引入契約彈性，使刪除權的行使可依個別情境調整；其二，透過適當的激勵機制，使資

328 觀察我國規模前三大的金融控股公司（富邦金、國泰金與中信金）之個資交互利用或共享同意書，僅國泰金設有類似對價的約定：「本人已知悉可隨時通知貴公司或上述任一子公司停止資料交互運用，且不同意資料交互運用時，將無法優先享有前述公司所提供之整合性資源、優惠措施或服務。」其餘公司在此類條款上則付之闕如。然而，即便國泰金已引入對價概念，其優惠內容仍屬寬泛且缺乏具體化，能否形成具突出性的激勵效果，進而矯正市場失靈與行為偏誤，仍有待實務檢驗。參國泰金控，《資料交互運用同意書》，https://www.cathay-ins.com.tw/insurance/assets/CoMarketing/aa9904.pdf。

329 *See* WILLIAM J. CONGDON ET AL., *supra* note 286, at 22-23.

330 楊岳平，前揭註 252。

doi:10.6342/NTU202600641

料主體更加重視其個資價值並提升分享意願。惟契約自由與市場誘因雖有助於解決部分問題，但在行為經濟學對理性人假設的修正下，仍必須正視資料主體因有限理性與有限意志所引發的行為偏誤，否則制度設計的善意可能在實際決策場景中被折損。

具體而言，資料主體在面對冗長且複雜的個資處理條款時，常因有限注意力而選擇直接略過內容、逕行簽名，忽略其中與隱私風險密切相關的重要細節；又因有限計算能力，在評估長期個資利用可能帶來的風險與效益時，難以全面衡量，僅能作出「尚可接受」而非真正最優化的決策[331]。更為關鍵的是，有限意志會導致時間不一致行為：資料主體雖在計畫階段設定了理性且符合長期利益的隱私策略，但在執行階段卻因短期便利或即時回報而延宕行動，最終錯失行使刪除權或調整隱私設定的最佳時機。換言之，即便提供了契約彈性與誘因，若不處理上述偏誤，制度成效仍可能大打折扣。

觀察我國規模前三大的金融控股公司之個資利用同意書[332]，可見其在版面與語言設計上呈現不同程度的「偏誤暴露度」。以台北富邦銀行[333]與中國信託銀行[334]為例，條款常見用語艱澀、篇幅冗長且要點標示不顯著；加上資訊密度過高，易使

331 有國外文獻指出，「告知與同意」（notice-and-consent）制度中，使用者的同意通常是在高度形式化的程序下取得，例如透過點擊「我同意」或在使用服務時默示同意，而非建立在實質的、具對價意義的協商過程上。批評者認為，這種同意往往是非自願的（使用者幾乎「別無選擇」只能同意，因為避免資料蒐集意謂著必須放棄使用關鍵科技）且資訊不足（隱私政策冗長複雜、數量龐多、可隨時單方更改、資料推論結果難以預測、且第三方流通頻繁），導致實質上並非基於真正協商或交換利益的合意，而只是透過形式化程序達成的契約效果。*See* Kevin Mills, *Consent and the Right to Privacy*, 39 J. APPL. PHILOS 721, 723-24 (2022).

332 我國市值規模前三大的金融控股公司分別為：富邦金、國泰金與中信金。參：聯合新聞網（03/28/2025），〈台灣 3 家金控躋身全球百大銀行〉，https://udn.com/news/story/6839/8638432（最後瀏覽日：2025 年 8 月 12 日）。

333 台北富邦銀行，《個人資料之轉告知聲明暨共同行銷、資料共享同意書》，https://www.fubon.com/banking/document/news/TW/%E5%80%8B%E4%BA%BA%E8%B3%87%E6%96%99%E4%B9%8B%E8%BD%89%E5%91%8A%E7%9F%A5%E8%81%B2%E6%98%8E%E6%9A%A8%E5%85%B1%E5%90%8C%E8%A1%8C%E9%8A%B7%E3%80%81%E8%B3%87%E6%96%99%E5%85%B1%E4%BA%AB%E5%90%8C%E6%84%8F%E6%9B%B8__(112%E5%B9%B49%E6%9C%88%E7%89%88)%20.pdf

334 中國信託商業銀行，《個人資料運用告知同意書》，https://www.ctbcbank.com/content/dam/twcbo/files/eform/sg/SG_PersonalAgreementDoc_PIPA.pdf。

doi:10.6342/NTU202600641

資料主體無心閱讀或缺乏消化能力，進一步加劇資訊不對稱。相較之下，國泰金控則在用詞、版面與呈現方式上較為簡潔明瞭（如採用白話語句、縮短篇幅、強化重點視覺），較能降低行為偏誤並提升理解與自主判斷[335]。此一對照顯示：相同的法律內容，因選擇架構設計不同，可能帶來截然不同的實質保護效果。更可見，即便是市場規模居前的業者，其中兩家在個資同意流程中仍存在隱私保障不足的疑慮，反映我國金融實務在個資與隱私保護的落實上仍有相當大的精進空間。

基於上述脈絡，為在不損及契約自由前提下減輕行為偏誤風險，或可同步導入行為經濟學的輕推設計，作為契約彈性與誘因機制的「操作層工具」。方向上，可包括：提升關鍵資訊揭露的清晰度與顯著性，確保資料主體在決策前能迅速掌握核心風險與利益；設置契約成立後的冷靜期，降低倉促決策或短期誘惑所導致的後悔成本；以及透過預設條款保留刪除權，僅在經過額外確認程序後才同意特定資料使用或保存安排，以減少現狀偏誤與自動選項效應對選擇的干擾。

第五節　小結

綜合前述分析，法律經濟學為個人資料刪除權的制度設計提供了多層次的啟發。從寇斯定理與新古典經濟學的角度觀之，制度應保留一定的契約彈性，容許資料主體與金融機構在資訊對等與自願基礎上，就刪除權限制或資料保存期間作適度約定，甚至引入合理的對價安排，以提升資料流動效率與市場創新動能。然而，行為經濟學的觀察提醒我們，資訊不對稱、有限理性與行為偏誤會削弱契約自由的實質意義，使形式上的同意不必然等同於真正自主的選擇。

因此，在制度設計上，有必要引入輕推工具，例如優化資訊揭露、改善預設規則、設置冷靜期等，以矯正決策偏誤並強化自主性。此外，考量個人資料的準公共財特性與搭便車問題，若欠缺適當誘因，資料主體可能因隱私顧慮而抑制分享意願，限制 AI 模型訓練與金融創新。基此，制度設計上或可探索引入明確且具顯著性的

[335] 參國泰金控，前揭註 328。

doi:10.6342/NTU202600641

對價機制，或將資料提供與特定服務權益連結，在不侵蝕資訊自決權的前提下，兼顧資料供給的穩定性與市場效率。

綜上所述，未來刪除權制度的發展宜採取「有限度的契約自由」與「行為經濟學的輕推設計」雙軌並行的思維：前者著重於確保資料利用的持續性與供給效率，使金融與 AI 應用得以在穩定基礎上發展；後者則關注於矯正資訊不對稱與行為偏誤，確保資料主體在作出決策時具備實質自主性與真實意願。此種雙軌設計既維持市場運作與技術創新的動能，也避免個資保障淪為形式化，可作為制度優化的理論基礎，為後續探討具體規範提供啟發。

doi:10.6342/NTU202600641

第五章 比較法之借鏡

本章將進行比較法分析，並聚焦於歐盟《一般資料保護規則》（General Data Protection Regulation, GDPR）與美國加州《消費者隱私法案》（California Consumer Privacy Act, CCPA）之相關規範。前者為現行全球最具體系性之資料保護法，後者則為美國最具代表性且於實務層面影響廣泛之州級法規。簡言之，兩部法規皆承認資料主體享有刪除其個資之權利，然其針對具體適用條件、例外事由與契約限制安排方面，展現出顯著差異。GDPR 強調刪除權為資訊自決核心，並輔以強制性程序與對資料管理者課與的技術義務加以保障；相較之下，CCPA 雖賦予資料主體刪除請求權，但亦允許企業在履行契約或合理業務需求下保留個資，並對資料交易行為設有明確之選擇權與激勵條款設計空間。本文將透過此二法之比較，分析其對刪除權與資料彈性保存機制之安排，進而為我國制度之調整提供參考依據。

第一節　歐盟法之參考

第一項　歐盟個人資料監管概況

隨著人工智慧與數位技術的快速發展，歐盟近年來積極強化個人資料保護之制度設計，試圖確保 GDPR 在高度動態且跨境運作的科技環境中仍具實效性。

為強化歐洲經濟區（European Economic Area, EEA）各國資料保護機關（Data Protection Authorities, DPAs）之執法能力，歐洲資料保護委員會（European Data Protection Board, EDPB）推動「專家支援計畫」（Support Pool of Experts, SPE），積極建構一個涵蓋法律、資安與 AI 倫理等領域的跨域專家資料庫，協助 DPAs 處理新興科技下所衍生之高度專業與跨境性資料保護案件[336]。該計畫特別聚焦於人

[336] European Data Protection Board, *Support Pool of Experts (SPE) Programme*, https://www.edpb.europa.eu/support-pool-experts-spe-programme_en (last visited May 5, 2025).

doi:10.6342/NTU202600641

工智慧、金融科技、加密演算與雲端運算等領域，並已累積多項針對大型語言模型之隱私風險[337]、演算法偏見評估[338]、與資料主體權利落實機制的研究成果[339]。

雖然歐盟目前尚未針對AI模型中的個人資料處理制定專法或發佈具拘束力的具體技術指引，惟EDPB於2024年所發佈的第28/2024號意見書中，已針對AI模型訓練使用個人資料之合法性提出實質指導[340]。該報告針對數項核心問題提出具體說明，並嘗試建立歐盟成員國間一致的解釋與執行標準，包括：(1) AI 模型是否構成個人資料之判斷需視個案情況而定；(2) 評估模型是否達成「匿名化」的準則；(3) 以「合法利益」作為AI模型個人資料處理之法定依據的適用性；以及(4)模型開發階段中若有違法處理行為，是否影響該模型後續運作的合法性等[341]。此等實務指引，展現出監管機關致力於平衡創新發展與基本權保障之努力。

此外，歐盟亦透過多項新制定的數位法規，如《人工智慧法案》（AI Act）、《資料法案》（Data Act）、《數位服務法》（DSA）與《數位市場法》（DMA），試圖建立涵蓋多層次的監管架構[342]。其中，AI Act為全球首部全面針對人工智慧風險進行分級規範之立法。該法建立風險為本的監理架構，並針對高風險系統設有資料治理、透明性與可追溯等義務，與《一般資料保護規則》（GDPR）保障之個資刪除權密切相關[343]。

---

[337] Isabel Barberá, *AI Privacy Risks & Mitigations – Large Language Models (LLMs)*, European Data Protection Board (Apr. 2025).

[338] Dr. Kris Shrishak, *AI: Complex Algorithms and Effective Data Protection Supervision – Bias Evaluation*, EUROPEAN DATA PROTECTION BOARD (Mar. 2024).

[339] Dr. Kris Shrishak, *AI: Complex Algorithms and Effective Data Protection Supervision – Effective Implementation of Data Subjects' Rights*, EUROPEAN DATA PROTECTION BOARD (Mar. 2024).

[340] European Data Protection Board (EDPB), Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models 7–8 (Dec. 17, 2024), https://edpb.europa.eu.

[341] *Id.*

[342] European Commission, *Shaping Europe's Digital Future*, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/shaping-europes-digital-future_en (last visited May 6, 2025); European Commission, *Regulatory Framework Proposal on Artificial Intelligence*, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (last visited May 6, 2025); European Commission, *The Digital Services Act package*, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package (last visited May 6, 2025).

[343] *Id.*

doi:10.6342/NTU202600641

綜上，歐盟面對科技變遷與 AI 應用挑戰所展現之制度調適能力，不僅體現在法規的持續更新與補充，更透過跨機關協作、專業支援機制與實務指引的累積，逐步建立一套具有預警性、可操作性與基本權導向的資料治理體系，為全球數位監管提供具啟發性的治理典範。本文擬就歐盟《一般資料保護規則》（GDPR）與《人工智慧法案》（AI Act）針對金融應用情境下人工智慧系統所涉及之資料刪除權相關規範加以說明與分析。

第二項　與人工智慧資料刪除權相關之軟法與硬法規範分析

第一款　2016 年 GDPR

歐盟 GDPR 為歐洲資料保護制度的核心法規，自 2018 年 5 月 25 日正式生效，取代舊有的《資料保護指令》（Directive 95/46/EC）。GDPR 適用於在歐盟境內提供商品或服務、或監控歐盟資料主體行為之所有控管者與處理者，具域外適用性[344]。該規則確立了包括存取權、更正權、刪除權（被遺忘權）、限制處理權與資料可攜權等多項資料主體權利，並強化控管者之責任義務，如資料保護影響評估（DPIA）、預設與設計資料保護、資料外洩通報等機制[345]。GDPR 在制度設計上強調「以風險為本」（risk-based approach）與「問責原則」（accountability）[346]，成為全球個資保護立法的重要參考典範[347]。

---

[344] Regulation (EU) 2016/679, art. 3(1)-(2), 2016 O.J. (L 119) 1.

[345] European Commission, *Data Protection in the EU*, https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en (last visited May 6, 2025).

[346] GDPR 不僅強調依風險程度調整合規行為，更透過「問責原則」（Article 5(2)）要求資料控制者能證明其合規性，這是風險導向制度的核心基礎。*See* Regulation (EU) 2016/679, art. 5(2), 24, 25, 2016 O.J. (L 119) 1.

[347] 實務層面，歐洲議會認為風險預防與問責原則有助於達成資料保護的「正和效益」，在必要的緩解機制下提升整體效用；英國 ICO 亦認為，問責制度是展示尊重個人隱私並建立信任的重要機會。*See* EUROPEAN PARLIAMENTARY RESEARCH SERVICE, *supra* note167, at 3; Information Commissioner's Office, *Guide to accountability and governance*, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/accountability-and-governance/guide-to-accountability-and-governance/?utm_source=chatgpt.com (last visited May 6, 2025).

doi:10.6342/NTU202600641

第一目　與刪除權相關之法規內容

歐盟 GDPR 第 17 條明文保障個人資料之「刪除權」，規定在多種情況下，資料主體得請求資料控制者刪除其個人資料。此等情形包括：個人資料已不再符合原始蒐集或處理目的之必要性；資料主體撤回其同意，且不存在其他合法依據；資料主體依第 21 條提出反對處理，且無優先於其權益之合法理由；資料係以違法方式處理；該資料係為實現公共利益、科學或歷史研究目的或統計目的[348]；或依歐盟或成員國法須刪除等[349]。

表面觀之，第 17 條之刪除權似僅能於特定事由下方得行使，然而若與第 7 條第 3 項「撤回同意」之規範相互結合，其實質適用範圍即大幅擴張。第 7 條第 3 項明定，資料主體得「隨時」撤回其同意，且撤回不得影響撤回前基於同意所為資料處理之合法性[350]。此一設計意味著，只要資料主體選擇撤回同意，即將直接觸發第 17 條第 1 項第(b)款所規定之刪除事由。除非符合第 17 條第 3 項所列之極少數例外（如保障言論與資訊自由、履行法律義務、維護公共利益、保障公共衛生、保存檔案以供公共利益、學術或歷史研究，以及法律上權利之主張），否則資料主體在幾乎任何時點皆可透過撤回同意要求刪除其個人資料[351]。換言之，刪除權雖名義上受限於事由，但在實務操作上，因撤回同意的高度便利性，幾乎等同於賦予資料主體一項近乎無限制的刪除權。

此種設計與我國憲法法庭於 111 年憲判字第 13 號判決所揭示之理念，可謂異曲同工。該判決明言：「當事人就獲其同意或符合特定要件而允許未經同意即得蒐

348「科學研究豁免」之所以屬 GDPR 嚴格個資規範下的例外條款，係因 GDPR 的立法初衷在於規範如 Facebook、Amazon 等大型網路平台的資料濫用行為，若一律套用高標準的個資限制，恐對科學研究造成過度衝擊。基於此，立法者考量科學研究所追求的公共利益，以及隱私保護之間的衡平，特別於前言與條文中設置「科學研究豁免」規範，使在特定條件下得以限縮 GDPR 的適用範圍，藉此減緩過度嚴格的規範對學術與科研發展的阻礙，並確保隱私與研究自由之間能取得適度平衡。參何之行、廖貞（2020），〈AI 個資爭議在英國與歐盟之經驗──以 Google DeepMind 一案為例〉，《月旦法學雜誌》，302 期，頁 146、147。

349 Regulation (EU) 2016/679, art. 17(1), 2016 O.J. (L 119) 1.

350 Regulation (EU) 2016/679, art. 7(3), 2016 O.J. (L 119) 1.

351 Regulation (EU) 2016/679, art. 17(3), 2016 O.J. (L 119) 1.

doi:10.6342/NTU202600641

集、處理及利用之個資，仍應保有事後控制權，不因其曾表示同意，或因符合強制蒐用要件，即喪失請求刪除、停止利用或限制利用個資之權利[352]。」兩者雖分別立基於歐盟與臺灣不同的法制脈絡，但均共同彰顯「資料主體事後控制權」之核心精神，凸顯同意並非一經作出即永久不可撤回，而應受到事後自主控制之保障。

此外，為確保資料主體的「同意」確實基於明確告知或積極行為而作出，並具備自主性、具體性、充分知情性、明確性及自由意志等要件，即達到真正的知情同意（informed consent）[353]，GDPR 第 7 條第 1 至第 4 項進一步對基於「同意」的個人資料處理設下嚴格規範，明文界定其有效性條件與運作方式。其核心內涵如下：其一，第 1 項將舉證責任分配給資料控制者，要求其證明資料主體已就其個人資料之處理明確表示同意，此舉不僅提升透明性，也降低資料主體日後行使刪除權的舉證負擔。其二，第 2 項要求，若同意係於一份同時涉及其他事項之書面聲明中作出，則必須以清楚、易於理解、並可與其他條款明確區辨的方式呈現，且應採用淺白、通俗易懂的語言。其三，第 3 項再次確認資料主體享有「隨時撤回同意」的權利，且撤回不得影響撤回前處理之合法性；同時規定資料控制者在蒐集同意之前，必須明確告知資料主體此項撤回權，並確保撤回的方式與提供同意一樣容易。其四，第 4 項進一步強調，在判斷同意是否屬自由給予時，應特別考量契約履行（包括提供服務）是否被不當附加處理非履約所必要個資之條件，以防止同意淪為形式化或被迫之結果[354]。

綜上可見，GDPR 所建構之刪除權與同意規範，並非僅是形式上的權利宣示，而是透過舉證責任、資訊揭露、撤回便利性與契約自由檢視等多層次設計，實質確

[352] 111 年憲判字第 13 號判決理由書第 69 段。

[353] 所謂知情同意（informed consent），又稱告知後同意原則，意指資料控管者負有充分告知之義務，必須以透明方式揭露個資蒐集與利用的相關資訊，使資料主體在獲得完整且清楚的說明後，得基於告知內容或明確肯定的行動，作出具自主性、具體性、充分知情性、明確性與自由意志的「同意」。如此方能確保資料主體之合意真正成立，並據以授權資料控管者蒐集與使用其個人資料。參張陳弘（2019），〈新興科技下的資訊隱私保護—「告知後同意原則」的侷限性與修正方法之提出〉，《臺灣大學法學論叢》，第 47 卷第 1 期，頁 218-220；曾憲立、朱斌妤、陳恭、戴豪君（2024），〈個人資料授權在知情同意機制的優化研究〉，《行政暨政策學報》，第 78 期，頁 69-72。

[354] Regulation (EU) 2016/679, art. 7(1)-(4), 2016 O.J. (L 119) 1.

doi:10.6342/NTU202600641

保資料主體對個人資料得以保持持續性、有效的控制權。此一精神與臺灣憲法法庭所闡釋之個資自主理念相互呼應，共同揭示現代個人資料保護體系中「事後控制權」的核心價值。

第二目　經濟分析觀點下的 GDPR 刪除權

若從本文前一章法律經濟分析所揭示之制度設計方向觀察 GDPR，可從數個面向切入。其一，GDPR 是否允許資料主體在一定範圍內透過契約自由，排除刪除權或延長個人資料保存期間？其二，GDPR 是否容許資料主體以對價方式訂立契約，透過有償交換調整刪除權之適用？其三，GDPR 是否設計足以避免資料主體在締約過程中因行為偏誤而作出不利選擇之機制？

就前兩個問題而論，GDPR 並未賦予當事人以契約自由任意處分或排除刪除權。換言之，刪除權屬於強行規範（mandatory rules），資料主體不得事前合意放棄刪除個資之權利，亦不得將該權利以契約方式讓渡或出售予他人[355]。此一設計，與 GDPR 強調資料保護作為「基本權」而非單純交易標的之立場相互呼應：個人資料之保護緊密結合於人格權與隱私權，不得視為僅具財產價值之資源[356]。

然而，從法律經濟學角度觀察，此種嚴格設計雖能避免刪除權因資訊落差或契約不對稱而遭濫用，卻同時剝奪了當事人進行成本效益權衡的空間。倘若資料主體與控制者間存在高度信任，或雙方確有基於市場交換之偏好（如允許平台留存資料以換取優惠），則現行 GDPR 一概禁止排除刪除權之規範，可能反而導致資源配置效率下降。特別是在數位經濟脈絡下，個人資料常被視為「交換籌碼」或「數據資本」，而全面禁止其契約流通，勢必在某種程度上限制市場之創新可能性。

---

[355] *See* Paul De Hert & Vagelis Papakonstantinou, *The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals?*, 32 COMPUTER LAW & SECURITY REVIEW 179, 192 (2016).
[356] ORLA LYNSKEY, THE FOUNDATIONS OF EU DATA PROTECTION LAW, 35-40 (2015).

doi:10.6342/NTU202600641

至於第三個問題，雖然 GDPR 並未允許資料主體就刪除權透過契約作出特別安排，但在規範資料主體同意個資處理之過程中，其制度設計已融入行為經濟學對人類決策偏誤的洞見。以第 7 條第 2 項為例，若同意條款被納入涉及多事項之契約聲明，則必須以「清楚、易懂且可區辨」之方式呈現。此舉並未剝奪資料主體之選擇權，而是透過精心設計之「選項架構」降低資訊過載與有限注意力所致之決策錯誤，正可視為 Thaler 與 Sunstein 所謂「輕推」（nudge）之典型實踐。此外，第 7 條第 3 項更進一步要求，資料主體應享有隨時撤回同意之權利，且其撤回方式應與給予同意同樣容易。此種設計猶如在同意後保留一段「冷靜期」，使資料主體即便在最初因衝動、過度折現或低估長期風險而同意資料處理，仍可在日後經過反思後予以撤回。換言之，GDPR 並非僅透過形式上的資訊揭露來維護自主，而是實質上藉由制度性輕推來矯正認知偏誤，確保資料主體能持續掌握對個資處理的控制權。

綜合以上觀察，GDPR 對刪除權採取強行規範立場，確保其不被任意契約化處分，並藉由制度化的輕推設計矯正行為偏誤，以維護資料主體的持續控制權。此一設計雖強化了刪除權作為基本權的定位，但同時也壓縮了當事人進行成本效益衡量的彈性，對於數位經濟中以個資作為交換籌碼的可能模式，難免存在一定程度的效率疑慮。換言之，GDPR 在保障自主與防範濫用之間，採取了高度保護的取徑，其規範選擇既展現了基本權思維，也暴露出與市場機制間的緊張。

第二款　人工智慧法

歐盟於 2024 年正式通過的 AI Act，是全球第一部針對 AI 系統採用風險導向監理架構的綜合性立法，其制度設計不僅回應了 AI 技術廣泛滲透社會的重要性，也試圖建立一套可預期、分層級的法律責任與義務架構。該法核心理念為「風險為本」（risk-based approach），依據 AI 系統對基本權利與社會利益可能造成的衝擊

doi:10.6342/NTU202600641

程度，將系統劃分為「不可接受風險」、「高風險」、「有限風險」與「最小風險」四類，並分別施以不同程度的法律限制與監管義務[357]。

其中，被列為高風險（high-risk）之 AI 系統，需遵守一系列嚴格規定，包含資料治理（data governance）、透明性（transparency）、可追溯性（traceability）、人為監督（human oversight）與安全穩健性（robustness and accuracy）等[358]。這些義務與 GDPR 中對個人資料的保護要求尤其是刪除權的實踐，有著高度關聯。例如，為確保高風險 AI 系統可回應使用者行使刪除權的需求，開發者與使用者須建立明確的資料刪除流程、紀錄機制與影響評估，以避免資料持續留存對個人權益造成侵害[359]。

在金融領域中，AI 法第 6 條與附錄 III 明確將「信用評等」納入高風險 AI 系統的範疇，故銀行、貸款機構若使用 AI 來評估個人或企業的信用風險，必須全面遵守高風險義務，如前述的資料可追溯與透明揭露，並接受主管機關的事前登記與後續監理。此舉反映歐盟對金融決策過程中 AI 的潛在歧視、資料偏誤與不當自動化決策所帶來風險的重視[360]。

然而，同樣位於金融場域的「金融詐欺偵測」（financial fraud detection）系統，則被 AI 法排除於高風險範疇之外。此一差異顯示歐盟在立法時已考量到風險的社會敏感性與制度成本，對於以公益與系統安全為目的之 AI 應用（如打擊詐欺），採取較寬容的監理態度，以避免造成實務阻礙，並保留技術靈活運用的空間[361]。

再者，針對高風險人工智慧系統，AI 法第 10 條進一步規定，高風險人工智慧系統若使用涉及以資料訓練人工智慧模型的技術，在使用這些資料庫時，開發者應

[357] European Commission, Regulatory Framework Proposal on Artificial Intelligence, Digital Strategy, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (last visited June 19, 2025).
[358] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on Artificial Intelligence (Artificial Intelligence Act), ch. III, § 2, arts. 9–15, 2024 O.J. (L 1689) 12 July 2024.
[359] *Id.*
[360] *Id.* Annex III, § 5(b), 2024 O.J. (L 2024/1689) 1 ("AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems used for the purpose of detecting financial fraud.").
[361] *Id.*

doi:10.6342/NTU202600641

落實資料治理（data governance）與品質控管措施，確保訓練、驗證與測試資料「具代表性、無偏誤、完整且具關聯性」[362]。此一條文雖未明示刪除權之實作機制，然其對資料來源、內容及更新義務之要求，實質上為未來因應資料主體之刪除請求所需之資料可追溯性（traceability）與可控性奠定基礎。

最後，AI Act 第 2 條第 7 項明確指出：該法「不影響」GDPR 的適用[363]。換言之，凡屬處理個人資料之 AI 系統，即使於技術面已採取風險控管，仍應全面遵循 GDPR 所保障之個資刪除權與其他資料主體權利。此一見解亦獲得 EDPB 近期於《第 28/2024 號意見書》之強調：監管機關認為，即便 AI 模型經過匿名化或技術處理，若仍存在可識別風險，仍須受到 GDPR 框架拘束[364]。因此，AI Act 與 GDPR 構成一互補關係，前者以風險治理為核心架構，後者則保障資料主體之個人權益，雙者共同形塑人工智慧於金融等高風險領域中合法與可責治理（accountable AI governance）之基礎。

第三款　小結

綜上所述，GDPR 對刪除權的制度設計展現出強烈的基本權保障取向。透過第 17 條與第 7 條相互搭配，資料主體不僅在特定法定事由下得以行使刪除權，更因「隨時撤回同意」的便利性，而實質上擁有近乎無限制的事後控制權。此外，第 7 條第 1 至第 4 項針對「同意」之效力與運作方式，透過舉證責任分配、資訊揭露、撤回便利性與契約自由審查等多層次規範，進一步確保資料主體的自主性與選擇

[362] *Id.* art. 10(3) (“Training, validation and testing data sets shall be relevant, sufficiently representative, and to the best extent possible, free of errors and complete in view of the intended purpose. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons in relation to whom the high-risk AI system is intended to be used. Those characteristics of the data sets may be met at the level of individual data sets or at the level of a combination thereof.”).

[363] *Id.* art. 2(7)(“Union law on the protection of personal data, privacy and the confidentiality of communications applies to personal data processed in connection with the rights and obligations laid down in this Regulation. This Regulation shall not affect Regulation (EU) 2016/679, ...”).

[364] 歐洲資料保護委員會（EDPB）在《第 28/2024 號意見書》中指出，即便 AI 模型並非設計用來輸出個資，只要其內部參數中仍可能保留從訓練資料中學習到的個人資訊，進而有可能被直接或間接地從模型中提取出來，就可能構成個資處理 。因此，若某 AI 模型中，使用合理可預期手段仍可能取得與訓練資料中個體有關的資訊，即可認定該模型非屬匿名，從而須符合 GDPR 的相關規定。European Data Protection Board (EDPB), Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models 12-14 (Dec. 17, 2024), https://edpb.europa.eu.

doi:10.6342/NTU202600641

權不被形式化的同意程序所侵蝕。這些設計體現出 GDPR 對個資保護的高度重視，並與我國憲法法庭強調的個資事後控制權理念相互呼應。

然而，GDPR 在確保刪除權不得由契約自由任意排除的同時，也壓縮了當事人進行成本效益衡量的彈性，使其在經濟效率與市場創新上面臨挑戰。此種以「強行規範」確保基本權的取徑，不僅背離新古典經濟學所強調的契約自由與最適資源配置邏輯，亦與數位經濟將個人資料視為「數據資本」加以運用的市場邏輯形成結構性張力。換言之，GDPR 在個人基本權保障與經濟效率之間，展現出一種高度保護傾向，其制度選擇在增強資料自主性的同時，也可能相對降低市場靈活性與創新誘因。這樣的規範模式，亦對後續的數位監理發展產生深遠影響。尤其在人工智慧領域，歐盟 AI Act 嘗試在 GDPR 既有的高標準框架下，建立兼顧創新推動與基本權保障的新平衡，凸顯出兩部規範在制度理念上的延續與調整。

第二節　美國法之參考

為進行制度比較並分析人工智慧於金融應用下個人資料刪除權的可行設計，本文擬以加州之 CCPA 及其修正法 CPRA 為美國法之主要觀察對象。此係因 CCPA 自 2018 年生效以來，即被認為已為美國州級隱私法律設定了標準，其對刪除權之設計、資訊揭露義務及處理例外，對金融業與資料密集型企業亦具高度參考價值[365]。

至於加州另於 2023 年通過之《刪除法案》（Delete Act）[366]，其立法重點係針對資料經紀人（data brokers）建立集中刪除機制，雖有助於刪除權之可及性提升，惟其適用對象與本文所聚焦之資料控制者（如金融機構或金融科技業者）不同[367]，故本文不另行深入討論。其他如維吉尼亞州（VCDPA）[368]、科羅拉多州（CPA）

[365] *See* e.g., DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 871 (7th ed. 2023) (describing CCPA as having set the standard for U.S. state privacy laws).
[366] California Delete Act (SB-362, 2023) Cal. Civ. Code § 1798.99.80 et seq. (West 2023).
[367] Secure Privacy, *California's Data Deletion Law: Understanding the California Delete Act for Regulating Data Brokers* (Feb. 7, 2024), https://secureprivacy.ai/blog/california-delete-act-guide. "The Delete Act grants Californians the right to demand that data brokers erase their personal information from their records. This empowers individuals to manage their digital footprint and limit the data used for profiling, targeting, and potentially harmful purposes."
[368] Virginia Consumer Data Protection Act (VCDPA) Va. Code Ann. §§ 59.1-575 to -585 (2023).

doi:10.6342/NTU202600641

[369]與康乃狄克州（CTDPA）[370]等州法，亦有賦予刪除權之規定，但大多仿效 CCPA 框架，尚未展現顯著制度差異[371]，故本文亦不重複分析。

第一項　CCPA 與 CPRA

美國加州於 2018 年通過 CCPA，為全美首部保障消費者個人資料之綜合性法案，賦予消費者查閱、刪除與拒絕販售個人資料之基本權利。然而，原始 CCPA 頒布後即引發眾多爭議與修法訴求，最終於 2020 年經由公投通過 CPRA 進行重大修正[372]。修正後的法案不僅新增如更正權、限制敏感資料使用之權利，亦創設加州隱私保護局（California Privacy Protection Agency, CPPA）作為專責監管機關，並強化刪除權之適用與執行規範。

CPRA 所修正之 CCPA 條文中，明確規範消費者有權請求業者刪除其個人資料，並要求業者在合理範圍內通知相關服務提供者、承包商與第三人共同刪除資料。即便部分情況下可援引例外（如履行契約、保障安全、符合法定義務等），但刪除權原則上仍為不可剝奪之核心權利，尤其在人工智慧應用持續依賴資料訓練與分析的背景下，成為個資自決權落實的制度支柱。由於 CPRA 為 CCPA 的修正案，故以下就該部法案的討論以 CPRA 稱之。

第一款　與刪除權相關之法規內容

第一目　刪除權的基本內容

CPRA 於第 1798.105 條中，明文規範了消費者對於其個人資訊的刪除權。此權利的核心在於，消費者得要求企業將其自消費者蒐集的個人資訊予以刪除，為確保此權利能夠落實，法律同時課予企業一系列義務與限制。詳言之，企業在接獲消費者提出的「可驗證刪除請求」後，必須自其紀錄中刪除相關資訊，並進一步通知

[369] Colorado Privacy Act (CPA) Colo. Rev. Stat. §§ 6-1-1301 to 6-1-1313 (2023).
[370] Connecticut Data Privacy Act (CTDPA) Conn. Gen. Stat. Ann. §§ 42-515 to 42-529 (2023).
[371] DANIEL J. SOLOVE & PAUL M. SCHWARTZ, *supra* note 365, at 871–73.
[372] *Id.*

doi:10.6342/NTU202600641

其合作之服務提供者或承包商，要求其一併刪除[373]。立法者藉此避免刪除權流於形式，確保消費者得以真正掌握其個人資訊流通軌跡，並切斷資訊於不同商業主體間的連結鏈條。顯示 CPRA 對刪除權之理解不僅是單純的資料移除，更是一種「控制力」的落實。

然而，刪除權並非絕對，CPRA 設計出若干例外情形，使企業得以保留特定資訊。例如，若刪除會妨礙完成消費者當初所要求的交易、履行法律義務、或提供消費者合理期待之服務，企業得繼續保留資料。同樣地，資訊也可因確保系統安全、進行除錯、行使言論自由或遵守法律規範而被保留。此外，在公共利益研究（如科學、歷史或統計研究）之情境下，若刪除會導致研究無法進行，並且已取得消費者知情同意，企業亦得繼續保存相關資訊[374]。這些例外條款顯示，CPRA 下的刪除權並非絕對，而是受到平衡考量的限制。

進一步而言，CPRA 特別排除了契約排除刪除權的可能性。依第 1798.192 條規定，消費者基於 CPRA 所享有的權利屬於不可拋棄之權利（inalienable rights），任何契約若試圖透過格式條款或放棄聲明剝奪或限制刪除權，均屬無效[375]。換言之，企業雖可依據法定例外提出抗辯，但不得透過「契約自由」規避法律強制規範[376]。此一設計與 GDPR 的立場高度一致，均將刪除權視為具有公共秩序與基本權性質的強行規範（mandatory rule），藉此避免企業以不對等的契約力量削弱消費者的資料控制權。

第二目　肯認「資料價值化」的市場機制

然而，CPRA 與 GDPR 最大的差異之一，在於其同時引入「資料價值化」的市場機制。依第 1798.125 條，原則上企業不得因消費者行使刪除權或其他隱私權利

373 Cal. Civ. Code § 1798.105 (West 2024).
374 *Id.*
375 *See* Cal. Civ. Code § 1798.192 (West 2024).
376 *Id.*

doi:10.6342/NTU202600641

而對其進行差別待遇，如拒絕提供商品或服務、或收取不同價格[377]。但在(a)(2)及(b)(1)項的例外設計中，企業可以在「合理反映資料價值」的範圍內，向同意資料的蒐集、銷售、共享或保存的消費者提供金錢誘因、優惠折扣，或對拒絕分享資料的消費者收取差額費用[378]。此規範承認個人資料在數位經濟中的經濟價值，並將其明確納入交易架構之中。舉例而言，企業可以向同意分享資料的消費者提供免費服務，或以金錢補償方式換取其資料使用之同意。此種制度設計兼顧了消費者的「控制權」與企業的「商業模式彈性」，在保障隱私的同時亦容許市場機制運作。

同時，CPRA 對於企業在經濟誘因安排下的行為亦設有嚴格規範。依據第 1798.125 條第(b)(2)–(4)項，企業若提供與個人資料處理有關的優惠或經濟誘因，必須依第 1798.130 條明確告知消費者該誘因的內容，並取得消費者的明確同意。該同意必須建立在已清楚揭露計畫之重要條款的基礎上，且消費者得隨時撤回之[379]。此外，條文明確禁止企業透過不公、不合理、具脅迫性或高利剝削性的方式，誘使消費者同意參與此類計畫[380]。由此可見，CPRA 不僅在刪除權上排除了契約限制的可能性，更在經濟誘因領域嚴格規範企業行為，以防止透過「變相交易」或「利益綁架」削弱消費者對自身資料的控制權。

第三目　刪除權請求之程序規範

除了前述針對刪除權與激勵性誘因的實體規範之外，CPRA 第 1798.130(a)條更進一步建構了一套企業回應消費者請求的核心程序規範。此條款不僅提供消費者具體且明確的權利行使管道，同時對企業在接收與處理相關請求時的義務、期限與資訊交付方式提出具體要求，目的在於確保消費者的實體性權利能夠在程序層面獲得真實而有效的落實。

---

377 *See* Cal. Civ. Code § 1798.125(a)(1) (West 2024).
378 *See* Cal. Civ. Code § 1798.125(a)(2)-(b)(1)(West 2024).
379 Cal. Civ. Code § 1798.125(b)(2)-(3)(West 2024).
380 Cal. Civ. Code § 1798.125(b)(4)(West 2024).

doi:10.6342/NTU202600641

首先，第 1798.130(a)(1)條要求企業須提供「多元且便利」的管道，使消費者能夠提出資訊揭露、更正或刪除之請求[381]。此舉避免企業僅流於形式化的設置，防止以程序設障礙削弱權利行使，並直接體現保障權利可及性與公平性的核心精神。其次，第 1798.130(a)(2)條則針對「企業回應與處理消費者請求的時效」作出規範，原則上企業必須於 45 日內完成處理，必要時得申請合理展延，但不得無故拖延[382]。此一時間限制的設置，實質上避免了企業以程序拖延削弱刪除權的效力。

最後，第 1798.130(a)(2)與(a)(3)條則細緻化規範了「資訊交付方式與揭露內容」。其要點包括：企業須提供多種行使刪除權的途徑，且資訊揭露必須以書面形式交付，並應使用消費者能夠理解與操作的易用格式。法律特別強調資訊應以便於傳輸的方式提供，使消費者得以不受阻礙地將資訊移轉至其他實體。至於揭露的內容，則涵蓋企業直接蒐集以及透過服務提供商、承包商間接取得的所有個人資訊，並要求以清晰且具體的方式呈現，避免因技術或專業門檻而造成理解困難[383]。

綜觀第 1798.130(a)條，其核心意旨在於建立一套兼顧便利性、時效性與透明性的「程序保障」制度。立法者透過要求企業提供便利管道、設定明確處理時限、規範交付方式與揭露內容，確保消費者不僅在實體權利上享有刪除、揭露與更正等選項，更能在程序層面獲得有效保障。換言之，此條款構成了 CPRA 落實資料自決權的制度基石，使刪除權不再停留於抽象宣示，而能在實務操作中展現具體可行性與法律可預測性。

第二款　經濟分析觀點下的 CPRA 刪除權

從法律經濟分析的制度設計方向觀察 CPRA，可從以下兩個面向切入。其一，CPRA 是否允許資料主體在一定範圍內透過契約自由，排除刪除權或延長個人資料

[381] Cal. Civ. Code § 1798.130(a)(1) (West 2024).
[382] Cal. Civ. Code § 1798.130(a)(2) (West 2024).
[383] Cal. Civ. Code § 1798.130(a)(2)-(3) (West 2024).

doi:10.6342/NTU202600641

保存期間？且此契約是否得以有償方式訂立？其二，CPRA 是否設計足以避免資料主體在締約過程中因行為偏誤而作出不利選擇之機制？

在契約安排方面，CPRA 已透過第 1798.192 條明文禁止資料控制者要求消費者「放棄權利」或「限制其權利之行使」，因此刪除權屬於不能契約排除的強行規範。企業不得透過格式條款、會員合約、服務條款或其他交易安排，剝奪消費者依法刪除其個人資料的權利。此設計反映出立法者已意識到「契約自由」在數位市場中的侷限性：由於個資交易關係通常存在高度資訊不對稱，消費者普遍缺乏議價能力，若任由企業透過合約排除刪除權，將使消費者在形式上「自由同意」，卻在實質上喪失核心隱私保障。因此，CPRA 將刪除權提升為不可放棄的最低標準，確保其不會被契約架構掏空。

然而，儘管 CPRA 在立法設計上並未允許消費者以「出賣刪除權」或「換取經濟誘因」的方式放棄刪除請求，其仍一定程度承認了「資料價值化」的市場機制，並透過第 1798.125 條的「經濟誘因」制度加以規範。依據該條，企業得在「合理反映資料價值」的範圍內，向同意分享資料的消費者提供金錢獎勵或優惠折扣，亦可對拒絕分享資料的消費者收取差額費用。換言之，消費者可透過與企業締結一定年限的資料保存契約，以換取相應的補償。

此種設計在一定程度上呼應了新古典經濟學強調的「契約自由」理念，使資料流通與利用更趨有效率。特別是在人工智慧與刪除權的情境下，雖然「可隨時撤回同意」的制度對保障消費者權益至關重要，但也可能對模型訓練的穩定性造成干擾。若能結合合理的經濟誘因與嚴謹的程序保障，則有助於在維護消費者選擇權的同時，緩解資料供給不足所引發的公共財疑慮，為協調個資刪除權與 AI 發展之間的張力提供可能的解方。

此外，CPRA 在制度設計上亦回應了行為經濟學所揭示的有限理性與行為偏誤問題。具體而言，企業若欲透過折扣、獎勵、積分等經濟誘因鼓勵消費者提供個資

doi:10.6342/NTU202600641

或延長保存，必須符合數項要件：首先，企業須依第 1798.130 條規定，明確揭露誘因內容及重要條款；其次，消費者必須在充分知情後提供明確同意，且有隨時撤回之自由；最後，企業不得使用不公、不合理、具脅迫性或剝削性的誘因方式。這些規範有效降低了消費者因短視、過度折現或受到行銷操縱而低估隱私價值的風險。

從法律經濟分析角度來看，CPRA 在此展現出對「行為偏誤」的深層回應。透過「資訊揭露義務」「隨時撤回權」與「禁止不公平誘因」等制度設計，立法者試圖矯正消費者可能出現的不理性選擇，避免因小額誘因而犧牲長期隱私保障。此一制度安排兼顧了市場運作與消費者保護：一方面保留了部分契約自由與市場誘因的空間，另一方面則以強行規範與程序保障確保刪除權的核心價值不致流於空洞。綜合而言，CPRA 透過禁止契約放棄刪除權、嚴格規範經濟誘因、以及導入行為經濟學思維的保障設計，體現了「有限契約自由」與「反行為偏誤」的制度理念。這樣的安排不僅避免刪除權被市場力量架空，也在一定程度上平衡了數位隱私保護與資料經濟的需求，形成與 GDPR 可相互對照的規範路徑。

第三節　小結——歐美個人資料刪除權之法制比較

第一項　契約自由與刪除權限制之差異

透過上述分析可知，在契約自由與刪除權限制的制度設計上，歐美採取了截然不同的路徑，或多或少反映出其背後法律文化與價值取向的差異。歐盟 GDPR 對於透過契約方式限制刪除權採取嚴格禁止態度，強調刪除權的絕對性與資訊自主性。具體而言，GDPR 明文規定，資料主體得隨時撤回資料利用的同意，且禁止將非必要之個資處理作為契約履行條件[384]。換言之，即便雙方於契約中明確約定限制刪除權，該約款亦不具拘束力。此一立場體現了歐陸法系將個人資料保護視為基本

[384] Regulation (EU) 2016/679, art. 7(3)-(4), 2016 O.J. (L 119) 1.

doi:10.6342/NTU202600641

權之價值定位，認為刪除權不可讓渡或稀釋，而應確保資料主體於契約關係中仍享有完整的資訊自主權。

相較之下，美國加州 CPRA 採取相對溫和且具彈性的立場，允許在一定範圍內透過契約安排來處理個人資料。具體而言，雖然第 1798.192 條明文禁止以契約排除刪除權，但同時承認企業得以提供合理折扣、優惠或其他對價，引導消費者在明示同意下交換個人資料。此種設計並非全面否定契約效力，而是要求企業在提供誘因時充分揭露相關內容，並確保消費者得以隨時撤回同意。換言之，加州模式企圖在市場交易自由與資訊自主權之間取得平衡，藉由透明揭露與退出機制降低資訊不對稱的風險。其結果並未完全排除契約對刪除權的影響，反而在制度上承認個人資料的經濟價值，允許企業透過合理誘因維持資料的穩定供給，並進一步正當化資料的持續利用。

綜合而言，歐盟模式強調刪除權的不可讓渡性，透過嚴格禁止條款維護其實效，避免因契約安排而削弱資料主體的控制力；相對之下，美國加州則側重於市場機制下的自主選擇，允許在有限範圍內透過財務誘因延長資料保存，以緩解刪除權過度絕對化可能帶來的制度困境。此一對照不僅彰顯歐陸與美國隱私保障理念上的差異，也為其他法域在制度設計上提供了兩種典型的參考模式。本文將此「契約自由與刪除權限制之差異」整理如下表，以資比較。

doi:10.6342/NTU202600641

國家/地區	對以契約限制刪除權的態度	核心規範內容與重點	制度意旨
歐盟（GDPR）	嚴格禁止	● §7III：資料主體可隨時撤回同意 ● §7IV：不得將不必要資料當作契約履行條件 ● 否定以契約對刪除權施加約束力，即便有對價亦然	強調資訊自主不可讓渡，刪除權具絕對性，不容透過契約削弱
美國加州（CPRA）	溫和彈性，有限允許	● 第 1798.125 條：企業可提供「財務性誘因」換取資料 ● 須明示同意且可隨時退出 ● 不否定契約效力，但需設退出機制與揭露誘因內容	設計誘因制度兼顧使用穩定與資訊自主，允許透過契約合理延長資料保存期限

【表 五】GDPR 與 CPRA 就契約自由與刪除權限制之差異

本研究自行整理製作

第二項　行為經濟學設計下的同意制度比較

就行為經濟學所強調的「輕推」設計理念，歐美法制在同意制度上皆有所回應。首先，兩者皆要求資訊揭露必須清楚、具體、易懂，以避免條款包裹與資訊超載。GDPR Art. 7(2)明定，同意不得被夾帶於其他聲明之中，必須以清楚、易辨識的方式呈現[385]；而 CPRA 則要求企業在事前以清楚易懂的方式揭露經濟誘因的內容與重要條款[386]，並於網站首頁設置「Do Not Sell/Share」等明顯入口，以集中化方式管理消費者的選擇[387]。

其次，在同意的動態性上，兩者均強調「可隨時撤回」，且撤回不得比給予更為困難。GDPR 規定，資料主體得隨時撤回同意，且不得因此受到不利對待[388]；

[385] Regulation (EU) 2016/679, art. 7(2), 2016 O.J. (L 119) 1.
[386] Cal. Civ. Code § 1798.130(a) (West 2024).
[387] Cal. Civ. Code § 1798.135(a) (West 2024).
[388] Regulation (EU) 2016/679, art. 7(3), 2016 O.J. (L 119) 1.

doi:10.6342/NTU202600641

CPRA 同樣保障消費者可隨時撤回同意，並要求退出經濟誘因計畫須具備明確且便利的流程，確保消費者在整個互動過程中持續保有選擇權，且企業不得因消費者行使此類權利而對其進行差別待遇或歧視[389]。

最後，在誘因規範方面，CPRA 更進一步引入「禁止不公平誘因」的設計，明文規定企業不得採取不公、不合理、具脅迫性或剝削性的經濟誘因，以避免消費者因短視、過度折現或資訊不足而作出不利於長期隱私利益的決策[390]。相較之下，GDPR 雖較少著墨於具體的誘因設計，但其對「同意真實性」與「隨時可撤回」的高度要求，亦能在實質上減少決策偏誤所帶來的風險。

因此，GDPR 偏向以嚴格的程序要求來確保同意的真實性與有效性，而 CPRA 則透過限制誘因與退出機制來降低行為偏誤的風險。本文進一步整理兩者同意制度對行為經濟學問題之回應，並以表格呈現如下。

面向	歐盟 GDPR	美國加州 CPRA
資訊揭露	Art. 7(2)：同意不得夾帶於其他聲明中，必須清楚、具體、易辨識，避免條款包裹與資訊超載。	§1798.130(a)：企業須事前已清楚易懂的方式揭露經濟誘因內容與重要條款；§1798.135：網站首頁需設置「Do Not Sell/Share」等顯著入口，集中化管理選項。
撤回機制	Art. 7(3)：同意得隨時撤回，且不得因此受到不利對待；撤回方式不得比給予更困難。	§1798.125(b)(3)：退出經濟誘因計畫須設有明確且便利的流程；消費者可隨時撤回同意，企業不得因行使權利而歧視。
防止不公平誘因	GDPR：透過「真實自願性」要求避免脅迫或不當綑綁。	GDPR§1798.125(b)(4)：禁止不公、不合理、具脅迫性或剝削性的誘因設計。

389 Cal. Civ. Code § 1798.125(b) (West 2024).
390 Cal. Civ. Code § 1798.125(b)(4) (West 2024).

doi:10.6342/NTU202600641

制度意旨	偏向以嚴格程序要求確保同意真實性與有效性，避免資訊過載與不當條款。	透過限制誘因、退出機制與揭露義務，降低行為偏誤風險，並承認資料價值化的市場機制。

【表 六】GDPR 與 CPRA 就行為經濟學設計下的同意制度差異

本研究自行整理製作

doi:10.6342/NTU202600641

第六章 AI 時代下刪除權之法制展望：我國制度的建議

如前所述，我國《個資法》雖在文義上維持刪除權的不可拋棄性，卻透過例外條款與行政函釋釋放了一定契約自由空間，形成「初次歸屬於資料主體，但實務上允許有限契約彈性」的混合模式。此一「形式禁止、實務容許」的結構，一方面提供了制度上的彈性，另一方面卻導致適用上的不確定性，加上缺乏確保同意真實性與撤回可行性的配套，實質保障因此被削弱。

更進一步地，現行法並未明確規範資料提供是否應有補償，致使實務多以形式化同意取代對價機制。資料主體未能清楚意識其個資的經濟價值，不僅壓抑了分享誘因，也降低了市場效率。同時，行為偏誤的存在（如資訊過載、現狀偏誤），使得刪除權縱然存在，實際行使卻常流於形式。

整體而言，我國法制在刪除權契約制度上既存在「形式禁止、實質容許」的矛盾，也欠缺對價設計與行為經濟學式的保障措施，導致刪除權在實務中難以發揮完整效果。未來若要在契約自由與隱私保障之間取得更佳平衡，應考慮明確開放有限度的契約彈性，並引入對價機制與「輕推」式的制度設計，以確保資料主體的知情同意並提升分享誘因。

在此背景下，為了更清楚揭示不同制度選擇背後的價值取向與調和邏輯，並為我國未來修法或制度調整提供具體借鏡，本文將於本章透過比較分析檢視歐美法制的設計經驗，藉此勾勒可行的規範路徑，並在此基礎上提出對我國個資刪除法制的制度建議。

第一節　AI 應用與個資保護的前提

在進行法制的比較之前，必須先確認一個核心前提，即 AI 模型是否構成個人資料，從而有個人資料保護相關法制的適用？依據 GDPR 第 4 條第 1 項與第 26 號前言，個人資料的定義極為廣泛，凡是「任何關於已識別或可識別之自然人之資訊」

doi:10.6342/NTU202600641

均屬之[391]。換言之，唯有在資訊已完全無法再識別特定個人的情況下，始得視為匿名資料。

歐洲資料保護委員會（EDPB）在《第 28/2024 號意見書》中指出，即便 AI 模型並非設計用來直接輸出個資，只要其內部參數仍可能保留從訓練資料中學習到的個人資訊，並可被合理手段直接或間接提取，即屬於個資處理[392]。現有研究亦已揭示，部分模型存在洩漏個資的風險，例如透過訓練資料提取或模型重現等攻擊方式[393]。因此，僅因模型已完成訓練，即推定其屬於匿名資料處理，顯然並不妥當，而應依個案具體判斷。

由此可知，AI 模型若要被認定為匿名資料，其條件極為嚴格。如本文於第三章所述，目前最具潛力的解方之一是以合成資料作為訓練資料。然而，技術文獻已指出，若合成資料本身係由真實個資訓練生成，並進一步作為模型輸入或比對依據，即使合成資料並非真實個資，仍可能被鏈結或推論回特定個體[394]。若依歐盟標準檢視，此方案恐難以徹底迴避刪除請求。

另一種更能符合法律要求的方式是完整重訓模型（精確機器遺忘），即直接移除個資並重新進行模型訓練。然而，對於大型模型而言，此方式成本極高。例如 GPT-4 的訓練過程需動用數千部高效能伺服器連續運算數十日，並伴隨龐大的能源與財務消耗[395]。若僅為回應單一資料主體之刪除請求而頻繁重訓模型，於實務上難以負擔，亦將對技術發展造成嚴重阻礙。

綜合以上觀察，依據 GDPR 與 EDPB 的標準，AI 模型難以透過匿名化主張迴避資料刪除請求。在現行技術水準下，刪除權與 AI 模型的運作確實存在結構性扞

[391] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), art. 4(1), recital 26, 2016 O.J. (L 119) 1.
[392] EDPB, *supra* note 364, at 12-14.
[393] *Id*.
[394] Khaled El Emam et al., *supra* note 233.
[395] Meem A.M., *supra* note 65, at 3.

doi:10.6342/NTU202600641

格。此一張力不僅對 AI 的進一步發展構成潛在隱憂，在金融應用等高度仰賴資料穩定性的領域，更可能影響市場穩定性。是以，未來制度設計必須積極思考如何在隱私保障與技術創新之間尋求可行的調和方案。

第二節　引入有償契約自由機制的可能

如第四章所述，法律經濟分析除強調契約自由外，更重視「初次權利歸屬」的劃定。依循寇斯理論，唯有先確認刪除權究竟歸屬於資料主體抑或資料控制者，市場參與者才能在權利界線清晰的基礎上，透過契約自由進行調整與交易，以達致最適資源配置。在資訊對稱與自願條件下，契約確實有助於雙方偏好的體現與風險分擔的合理化；但若缺乏明確的初始界定，契約安排便可能淪為權利真空下的弱者讓渡，導致交易失靈。套用至個人資料處理領域，刪除權的初次劃定即為制度的核心基礎；在此基礎上，契約才能作為平衡資料自主與資料利用的中介架構，若能妥善設計，則有助於同時兼顧隱私保障與技術創新。

回到我國現行《個資法》觀察，立法者雖透過第 3 條將刪除權明文賦予資料主體，確立了初次權利歸屬，但同時又在第 11 條第 2 項但書與《施行細則》下，開放「職務或業務必要」或「經當事人書面同意」的例外，使企業得以在目的達成後繼續保存個資。此種設計表面上維持刪除權的不可拋棄性，卻透過例外條款與行政解釋釋放了一定契約自由空間，形塑出「初次歸屬予資料主體、但實務上允許有限契約彈性」的混合模式。從新古典經濟學角度觀察，這確實為市場留下調整空間，有利於滿足金融業務或 AI 訓練的穩定需求。

與此相比，美國加州 CPRA 在第 1798.105 條同樣明文保障消費者刪除權，並在第 1798.192 條明文規定刪除權屬於不可拋棄的基本權，任何契約若試圖排除或限制此權利，均屬無效。換言之，立法者選擇以強行規範確保刪除權地位，避免企業利用契約不對等削弱消費者隱私保障。歐盟 GDPR 則採取類似立場，第 17 條同樣將刪除權視為資料主體的基本權利，並僅允許在公共利益、法律義務或學術研究

doi:10.6342/NTU202600641

等有限情境下加以限制。企業無法透過契約自由排除刪除權，立法者以強行規範方式避免刪除權淪為私法自治下的弱勢讓渡。換言之，無論是 GDPR 或 CPRA，均將刪除權定位於不可由當事人拋棄的「絕對權利」，其價值取向明確地傾向資料主體保障。

但與 GDPR 的「純粹強行規範」相比，CPRA 另於第 1798.125 條引入「資料價值化」制度，允許企業在合理反映資料價值的前提下，透過優惠、折扣或金錢誘因吸引消費者同意共享資料。此種對價化的運作方式，能在交易中強化行為經濟學所強調的「突出性」——亦即讓當事人清楚意識到其行為所帶來的價值交換與實質成本。當對價具備顯著性時，資料主體不僅更能充分評估分享資訊的效益與風險，也能在意識到交換條件後作出更符合自身偏好的判斷。這種設計使資料主體得以依個人利益與偏好進行靈活安排，同時也拓展了願意以資訊換取對價者的選擇自由，促進制度規範與市場機制之間的契合，減少因資訊不對稱或心理偏誤導致的落差。

換言之，CPRA 雖否定契約排除刪除權的可能性，但卻透過「資料價值化」在刪除權之外創造了市場誘因的調和機制，較 GDPR 更貼近經濟學強調的「效率與公平並行」。相較之下，我國雖實務上容許資料控制者與資料主體透過契約約定保存期間，然《個資法》本身並未明文規範此類安排，更未設計與對價化相連動的制度。此種欠缺，容易導致契約僅流於形式，既無法凸顯交易的實質價值，也難以降低因認知偏誤而產生的決策落差。

本文以為，CPRA 的制度設計顯示刪除權與契約自由並非全然衝突，而是可以在「不剝奪基本權利」的前提下，透過市場化的對價誘因機制達成互補。我國未來修法，或可將資料保存契約及其對價安排明文規範於具法律位階的《個資法》中，既確保刪除權不被架空，又能兼顧金融業務與 AI 發展的穩定需求。

doi:10.6342/NTU202600641

第三節　行為經濟學視角下的「輕推」制度

在肯認契約自由與對價機制能為市場帶來正面效果的同時，仍須注意，若欠缺確保同意真實性的配套，消費者可能因資訊不對稱或談判力不足，而僅能作出流於形式的同意，進而削弱刪除權的保障。此一風險與行為經濟學揭示的偏誤現象相呼應，例如現狀偏誤與有限注意力，常使資料主體即便在法律上享有形式權利，卻在實務上失去實際自主空間。是以，雖然理論上立法者應允許金融機構與消費者透過契約自由約定是否以及如何限制個資刪除權，以提升資料利用之效率，但仍必須設計相應的制度配套，以確保能兼顧刪除權保障與資料利用穩定性的雙重需求。而行為經濟學「輕推」理論的運用，除了上節所提之「對價機制」，尚有「隨時撤回機制」與「資訊揭露機制」，以下即就此二配套措施加以說明。

第一項　隨時撤回機制

歐美法制為避免資料主體因有限理性與有限意志而做出短視決策，普遍賦予其「冷靜期」與「隨時撤回」的權利。此設計正是對行為經濟學所揭示的「現在偏誤」與「損失厭惡」的制度化回應：消費者往往低估長期約束的成本，卻過度重視即時好處，從而增加未來後悔與被剝削的風險。撤回權則提供了一種動態保障，讓資料主體在偏好或情境改變時，得以修正原先的決定，避免因一次性承諾而被鎖定在非最適狀態。

更重要的是，隨時撤回機制同時改變了契約雙方的誘因結構。金融機構在設計合約或優惠方案時，必須考慮消費者可能撤回的風險，從而避免設計過度偏頗或具強制性的條款。這不僅降低了資訊不對稱下的剝削空間，也迫使制度更貼近「公平契約」的理想狀態。換言之，撤回權既是保障資料主體控制權的動態工具，也是矯正行為偏誤的制度裝置，使市場結果更接近理性模型下的長期均衡。

以 GDPR 為例，其第 7 條第 3 項明文規定資料主體得「隨時撤回」同意，且不得以契約條款加以剝奪。加州 CPRA 則於第 1798.125 條允許企業提供財務性誘

doi:10.6342/NTU202600641

因以換取消費者明示同意分享個資，但同樣要求保留隨時退出的權利。此一設計雖會增加資料利用的不確定性，但其立法目的正在於回應行為經濟學揭示的現實困境——防止資料主體因短視與資訊落差陷入長期不利。

相較之下，我國《個資法施行細則》第 21 條第 1 款雖允許契約約定個資保存期間，作為刪除權行使之例外，但並未建立普遍性的「撤回同意」制度，而僅要求保存期限須具體明確，不得以籠統或無限期方式逾越誠信與比例原則。換言之，我國現行制度主要仰賴「定期契約」加上比例原則，來避免企業利用不公平條款侵蝕消費者刪除權，雖能發揮一定制衡功能，但缺乏隨時撤回所提供的動態保障與行為矯正效果。

因此，本文認為，若能引入「有償契約」並結合「隨時撤回同意」的雙重機制，將更能兼顧市場效率與個人保障。其理由在於：現行僅允許無償的定期保存契約，實際上仍可能侵蝕人民的資訊自決權，因為資料主體往往因「現在偏誤」或短視選擇而低估長期承諾所帶來的風險。反之，有償契約一方面透過實質對價使資料的經濟價值具體化，避免企業僅以象徵性誘因即長期占有使用權；另一方面也迫使資料控制者在契約設計與誘因安排上更為審慎，以免藉由資訊優勢規避公平義務。

固然，隨時撤回機制可能增加資料供給的不穩定性，對企業或 AI 模型的長期運作構成挑戰。然而，從激勵設計的角度觀之，正因消費者得以隨時撤回，企業反而更有動機提供能真實反映資料價值的對價，以降低撤回風險並維持資料供給的穩定。如此一來，不僅能避免價值被低估而導致的不均衡交易，也能透過行為矯正機制減少消費者因有限理性而受困於不利承諾。換言之，此種制度安排實際上在「個人控制」與「資料利用」之間建立起相對均衡的配置基礎，並進一步提升市場資源分配的效率與公平性。

綜合言之，該制度一方面在「個人控制」與「資料利用」之間建立更合理均衡，另一方面亦能使市場資源配置更貼近真實的社會成本與價值，應值我國借鏡。

doi:10.6342/NTU202600641

第二項　資訊揭露機制

在數位資料治理的脈絡下，資訊揭露機制成為確保「同意」真實性與有效性的核心條件。行為經濟學早已指出，個人在決策過程中往往受限於有限理性與認知偏誤（如現狀偏誤、損失厭惡與資訊過載），容易低估長期風險、過度依賴預設選項，致使同意淪為形式化勾選，而非基於理性與充分資訊的自主選擇。若制度設計未能針對此類偏誤加以矯正，形式上的「同意」即可能淪為企業濫用資訊優勢的正當化工具，難以發揮保障資料主體的應有功能。

因此，制度設計必須確保揭露內容具備清晰性、可理解性與顯著性，讓資料主體能在短時間內辨識可行選項並正確認知個資的經濟價值與潛在風險。若欠缺這樣的揭露安排，資訊過載與模糊選項將使消費者無法有效行使自主權，進而削弱同意作為保護機制的正當性。

歐盟 GDPR 第 7 條第 2 項對「同意」的明確性與可辨性設有高度要求，特別指出若同意係在含其他內容之書面聲明中作出，應以「清楚、易於理解且可明確區辨」的方式呈現，否則其拘束力不成立。第 12 條亦規定控管者應採取適當措施，以「簡潔、透明、易理解且易取得的形式」向資料主體揭露資訊（如處理目的、資料接收者等），此設計回應了行為經濟學對「條款包裹」與「選項模糊」造成之認知負擔的批評，透過形式規則強化資料主體辨識與判斷能力。

無獨有偶，加州消費者隱私法在制度設計上亦已明顯反映行為經濟學之「輕推」理念。根據 CPPA 第 1798.135(a)(1)-(2)條，企業若欲販售、共享個人資訊或使用敏感個資作為法定用途以外之目的，必須在首頁設置清楚可辨的連結，例如「Do Not Sell or Share My Personal Information」及「Limit the Use of My Sensitive Personal Information」，若企業不欲設置上開連結，同條第(a)(3)項亦容許業者採用單一且明確標示的整合連結，以提供消費者對於個資使用限制之全面選項。

doi:10.6342/NTU202600641

相較之下，我國個資法僅在第 6 條敏感性個資蒐集之同意要求書面同意，對於同意條款之可讀性、易理解性、與是否預設選擇等關鍵行為經濟面向，並無實質規範。且在實務運作中，常見以一紙制式條款涵蓋所有資料利用項目，導致資料主體難以辨識權利內容，弱化同意之自主性與有效性。舉例而言，觀察富邦銀行與中信銀行的個資共享同意書，皆存在內容繁雜、用詞艱澀、未突顯重點等問題[396]。

綜上所述，從比較法觀察可見，歐盟與美國加州均或多或少在法規範上已超越形式性「同意」機制的思維，朝向結合心理行為洞察與設計導向的制度模式發展。除 GDPR 透過第 7 條對同意之明確性與可辨性設下高標準外，CCPA/CPRA 亦要求企業於首頁設置清楚標示的連結，並允許集中設計單一整合連結以便人民行使拒絕權，進一步要求隱私政策定期揭露與更新，以降低使用者搜尋與行動成本，對現狀偏誤與認知超載構成制度性矯正[397]。

因此，若我國未來欲強化同意機制的實質保障，或可透過將結合「輕推」措施的資訊揭露規範訂明於法律中，以促使金融機構積極進行制度改革，確保同意不再僅是形式性勾選，而是能真實反映資料主體意願的實質選擇。此舉一方面能降低資訊過載與現狀偏誤對決策的干擾，另一方面亦能在程序上強化透明度與可辨識性，使消費者能即時掌握其所承擔的風險與所獲得的利益。

396 台北富邦銀行，前揭註 333；中國信託商業銀行，前揭註 334。

397 值得注意的是，儘管歐美均就資訊揭露有呈現方式的要求，然實際觀察歐盟市值最大的桑坦德銀行（Santander Private Banking）之隱私聲明，可見其雖透過重點標示與分點列項等方式提升條款的可讀性與理解度，但整體內容仍顯冗長繁複，對於增進消費者自主性與實質同意的效果恐有限。相較之下，美國銀行（Bank of America）之隱私聲明則以簡明易懂的語言、精煉的內容以及表格化的呈現方式，將關鍵資訊清楚展現，更有助於消費者在短時間內掌握要點，作出知情同意的判斷。*See* Santander Private Banking, *Privacy Note*, https://www.pb-santander.com/pdf/pdf/PrivacyNoticeBSISAEnglish.pdf; Bank of America, *Concumer Privacy Note*, https://www.bankofamerica.com/security-center/consumer-privacy-notice/.

doi:10.6342/NTU202600641

第四節　小結

綜上所述，我國《個資法》雖在形式上將刪除權定位為不可拋棄之基本權，但透過例外條款與行政解釋，實質上仍釋放出一定契約自由空間，形成「明文禁止、實務容許」的矛盾結構。此種混合模式固然提供了一定制度彈性，卻也導致適用的不確定性，並因欠缺確保同意真實性與撤回機制等配套，使刪除權保障流於形式。加之現行法並未明文規範資料提供是否須附對價，使實務上大多以形式化同意取代實質補償，弱化了資料主體對其個資價值的認知，不僅抑制分享誘因，也降低市場效率。

比較法的觀察顯示，GDPR 與 CPRA 皆將刪除權定位為不可由當事人放棄的「強行性權利」，但同時採取不同調和機制：GDPR 強調純粹的強制規範，嚴格限制契約自由；而 CPRA 則另建「資料價值化」制度，允許企業在揭露透明、可隨時撤回與公平對價的條件下，以優惠或金錢誘因交換資料利用，兼顧隱私保障與市場效率。相較之下，我國現行制度既未能提供足夠的保障機制，亦缺乏具體的市場化設計，致使刪除權難以兼顧實務需求與個人保障。

因此，本文建議，我國未來修法可考慮三項方向：第一，明文承認有限契約彈性，使金融機構得在明確規範下與資料主體約定資料保存期間；第二，引入「有償契約」制度，以真實對價凸顯資料價值並提升分享誘因；第三，結合行為經濟學式的「輕推」設計，建立資訊揭露、隨時撤回等保障機制，以矯正偏誤並確保同意之真摯性。此種制度設計不僅能在「個人控制」與「資料利用」間建立新的均衡，也能使市場資源配置更貼近真實社會成本與價值，為 AI 時代下我國刪除權法制的前瞻路徑。

doi:10.6342/NTU202600641

第七章 結論

本文透過對人工智慧技術發展與個人資料刪除權之間互動關係的分析，指出兩者在制度運作上存在結構性張力。在以深度學習為核心的 AI 技術架構下，模型效能仰賴長期、穩定且完整的資料庫累積；在金融服務快速普及後，資料穩定性更直接關聯到金融體系的可靠運作。然刪除權的強化則勢必牽動模型再訓練與營運持續性，形成「資料穩定供應」與「資訊自主權」之間難以迴避的衝突，並成為當代資料治理的核心難題。

就台灣現行法制觀之，《個資法》第 11 條第 3、4 項固然明文保障刪除權，並以第 3 條第 5 款禁止預先放棄之；然而，施行細則與行政解釋卻容許企業透過契約約定資料保存期限。此種「文義禁止、實質容許」的混合模式，雖保有一定彈性，卻也造成適用上的不確定性，並弱化了保障效果。更重要的是，我國制度中並未建立「對價機制」與「隨時撤回」等配套，致使同意常流於形式，誘因無法對準資料的真實價值，再加上現狀偏誤、資訊過載與有限注意等行為偏誤，進一步折損刪除權的實效。

本文嘗試透過經濟學視角尋求調和方案。從古典經濟學出發，刪除權若被全面強行化，將過度限制契約自由，不利於資料於市場中有效配置與價值發掘。反之，若允許契約在合理範圍內延緩刪除或約定保存期間，並以明確期間上限與具體對價補償作為條件，則有助於資料主體更真切地衡量分享資訊的風險與效益，在知情下作出偏好一致的選擇。如此設計既能兼顧資訊自主與資料穩定性，亦能提升制度效率。

然而，僅靠契約安排仍不足以克服市場失靈與行為偏誤。行為經濟學已揭示，消費者在面對資訊揭露與選項設計時，往往因有限理性與偏誤（如現狀偏誤、現在偏誤、資訊過載）而無法做出真正理性的選擇。若制度僅停留於形式化同意，便可能淪為企業濫用資訊優勢的正當化工具。因此，除了引入「有償契約」之外，更需

doi:10.6342/NTU202600641

同步導入「輕推」式保障措施，例如隨時撤回機制與資訊揭露機制，作為契約自由的重要補充。唯有如此，方能在「市場效率」與「資訊自主」之間建立更合理的均衡，確保刪除權於 AI 時代不致淪為形式化口號。

比較法上，歐盟 GDPR、美國加州 CPRA 與台灣《個資法》呈現出三種不同取向。GDPR 採取高強度的不可讓渡路徑，刪除權不得透過契約排除或弱化；CPRA 則在同樣保障刪除權為基本權的前提下，引入「資料價值化」機制，允許透過優惠折扣等財務誘因交換資料使用，但必須符合明確揭露、明示同意與隨時撤回的條件，以市場誘因確保資料供給的穩定性。相較之下，我國雖透過契約約定保存期限在實務上容許一定彈性，卻欠缺對價顯著性與可逆保障的明文設計，導致同意流於形式，資訊不對稱風險擴大。

在同意制度的設計上，GDPR 著重程序高標與權利可逆，要求清楚、可辨且易懂的揭露，並確保「撤回不得比同意更困難」，藉此減少資訊過載與現狀偏誤。CPRA 則更側重於操作性「輕推」，透過首頁明顯入口、單一整合連結、經濟誘因之具體揭露、隨時退出及禁止不公平誘因等設計，直接降低搜尋與行動成本，並矯正消費者短視與過度折現的傾向。相較之下，我國在相關規範上付之闕如，難以支撐資料主體的「知情同意」。

綜上，本文主張我國未來應建構一套兼具效率、正義與可操作性的資料治理體系。此體系應以「有償契約自由之明文化」為基礎，並輔以「行為經濟學導向的輕推原則」，特別強化資訊揭露與隨時撤回等措施，引導個人作出理性且自主的資料處分選擇。唯有如此，方能在個人資訊自主與人工智慧發展之間取得可持續性的平衡，避免刪除權流於形式，同時兼顧金融穩定、市場效率與技術創新。

doi:10.6342/NTU202600641

參考文獻

中文文獻

一、 專書

李惠宗（2022），《憲法要義》，第 9 版，元照。

二、 專書論文

王文宇（2019），〈第一章金融法制與金融監理〉。收於：王文宇（等著），《金融法》，頁 1-24，台北：元照。

李震山（2001），〈論個人資料之保護〉，收於：《行政法爭議問題研究（上）》，頁 657-664，五南。

李震山（2020），〈論資訊自決權〉，收於：《人性尊嚴與人權保障》，頁 239-314，元照。

黃昭元（2005），〈無指紋則無身分證？——換發國民身分證與強制全民捺指紋的憲法爭議分析〉，收於：《民主、人權、正義——蘇俊雄教授七秩華誕祝壽論文集》，頁 461-508，元照。

廖福特、翁逸泓（2008），〈兩難？共存？——國家處理個人資料與資訊隱私權保障之糾葛〉，收於：《二十一世紀公法學的新課題－城仲模教授古稀祝壽論文集》，頁 225－317，新學林。

三、 期刊論文

何之行、廖貞（2020），〈AI 個資爭議在英國與歐盟之經驗——以 Google DeepMind 一案為例〉，《月旦法學雜誌》，302 期，頁 127-156。

李沛宸（2019），〈實施歐盟個人資料保護規章對人工智慧發展之影響〉，《財金法學研究》，第 2 卷第 1 期，頁 125-156。

doi:10.6342/NTU202600641

李震山（2005），〈來者猶可追，正視個人資料保護問題──司法院大法官釋字第六○三號解釋評析〉，《台灣本土法學雜誌》，第 76 期，頁 222-234。

周振鋒(2019),〈論機器人投資顧問之興起與投資人之保護──以美國法為中心〉,《東吳法律學報》，30 卷 4 期，頁 69-107。

林勤富（2025），〈歐盟《人工智慧法》之制度設計、規範內涵與治理侷限〉，《中研院法學期刊》，第 36 期，頁 1-119。

邱文聰（2009），〈從資訊自決與資訊隱私的概念區分──評「電腦處理個人資料保護法修正草案」的結構性問題〉，《月旦法學雜誌》，第 168 期，頁 172-189。

范姜真媺（2013），〈個人資料保護法關於「個人資料」保護範圍之檢討〉，《東海大學法學研究》，第 41 期，頁 91-123。

徐彪豪(2015),〈被遺忘權近期發展──歐盟法院判決週年後回顧與本土觀察〉,《科技法律透析》，第 27 卷第 11 期，頁 50-70。

翁清坤（2008），〈臺灣與美國金融機構分享客戶個人資料之法律界限〉，《輔仁法學》，第 35 期，頁 69-162

張文村、呂宜穎、詹雅慧、林淑萍、陳宇軒、葉育惠（2019），〈人工智慧技術應用於中小企業徵信之初探〉，《電工通訊季刊》，第 2 季，頁 81-88。

張志偉（2017），〈記憶或遺忘，抑或相忘於網路──從歐洲法院被遺忘權判決，檢視資訊時代下的個人資料保護〉,《政大法學評論》,第 148 期,頁 1-68。

張志偉(2017),〈從資訊自決與資訊隱私的概念區分,檢視被遺忘權的證立問題〉,《萬國法律》，第 211 期，頁 2-15。

張陳弘（2019），〈新興科技下的資訊隱私保護—「告知後同意原則」的侷限性與修正方法之提出〉，《臺灣大學法學論叢》，第 47 卷第 1 期，頁 201-297。

許炳華（2015），〈被遺忘的權利：比較法之觀察〉，《東吳法律學報》，第 27 卷第 1 期，頁 125-163。

doi:10.6342/NTU202600641

陳家駿（2024），〈以人為本之 AI 及其透明性與可解釋性──我國人工智慧基本法草之省思與建議〉，《教育暨資訊科技法學評論》，13 期，頁 81-96。

陳銓雄、劉庭妤（2011），〈從「個人資料保護法」看病患資訊自主權與資訊隱私權之保護〉，《月旦民商法雜誌》，第 34 期，頁 23-57。

曾憲立、朱斌妤、陳恭、戴豪君（2024），〈個人資料授權在知情同意機制的優化研究〉，《行政暨政策學報》，第 78 期，頁 67-93。

楊岳平(2020)，〈演算法時代下的投資顧問監理議題──以理財機器人監理為例〉，《月旦民商法雜誌》，67 期，頁 28-50。

楊岳平（2023），〈金融消費者保護法制與監理的法律經濟分析觀點──以行為經濟學的應用為中心〉，《月旦法學雜誌》，339 期，頁 19-41。

楊智傑(2015)，〈個人資料保護法制上「被遺忘權利」與「個人反對權」：從 2014 年西班牙「Google v. AEPD」案判決出發〉，《國會月刊》，第 43 卷第 7 期，頁 19-43。

劉青峰(2023)，〈COVID-19 疫情下資訊自決權之研究－以歐洲人權公約第 8 條作為比較法對象〉，《中原財經法學》，第 50 期，頁 227-315。

劉靜怡（2010），〈不算進步的立法：「個人資料保護法」初步評析〉，《月旦法學雜誌》，第 183 期，頁 147-164。

劉靜怡（2012），〈社群網路時代的隱私困境：以 Facebook 為討論對象〉，《臺大法學論叢》，第 41 卷第 1 期，頁 1-70。

劉靜怡（2022），〈違憲之後－111 年憲判字第 13 號判決簡評〉，《當代法律》，第 11 期，頁 6-11。

蘇柏鳴（2018），〈徵信與 Fintech 發展淺談〉，《金融聯合徵信》，第 32 期，頁 17-20。

四、 學位論文

doi:10.6342/NTU202600641

呂承儒（2022），《論金融控股公司共同行銷之個人資料保護法制》，國立臺灣大學法律學研究所碩士論文。

五、 決議、解釋、函令或研究意見

金融監督管理委員會（2020），《金融科技發展路徑圖報告書》。

金融監督管理委員會（2024），《金融業運用人工智慧（AI）指引》。

國家科學及技術委員會（2024），《人工智慧基本法草案總說明及條文》。

謝碩駿，〈行政機關資料蒐集與個資保護〉，《Human Rights 法務局人權系列》，臺北市政府（2020）。

六、 計畫報告

谷湘儀、洪志麟、賴冠妤、陳國瑞與呂馥伊（2016），《機器人投資顧問（Robo-Advisor）國外實務及相關法令與管理措施之研究》，資產管理產業發展與人才培育基金委託專題研究。

陳安斌、陳莉貞、蘇秀玲、郭怡君（2018），〈我國發展機器人理財顧問之研究〉，《中華民國證券投資信託暨顧問商業同業公會委託報告》，載於：https://webline.sfi.org.tw/download/resh_ftp/AMEDFund/%E6%88%91%E5%9C%8B%E7%99%BC%E5%B1%95%E6%A9%9F%E5%99%A8%E4%BA%BA%E7%90%86%E8%B2%A1%E9%A1%A7%E5%95%8F%E4%B9%8B%E7%A0%94%E7%A9%B6.pdf.。

七、 網路資料與其他

工商時報（2024），〈年關近慎防詐騙！北富銀 AI 洗防系統正式啟用 全國首創 AI 洗錢防制模型〉，載於：https://www.ctee.com.tw/news/20240206701304-430304。

doi:10.6342/NTU202600641

工商時報（2024），〈銀行業 台北富邦銀行秉持 361 度服務 多一度溫暖客戶心〉，載於：https://www.ctee.com.tw/news/20240505700163-439903。

工商時報（2025），〈AI 扮利器 銀行阻詐戰力升〉，載於：https://www.ctee.com.tw/news/20250206700162-439901。

中央通訊社（2018），〈中信銀導入 AI 核貸只要幾分鐘〉，載於：https://tw.news.yahoo.com/%E4%B8%AD%E4%BF%A1%E9%8A%80%E5%B0%8E%E5%85%A5ai-%E6%A0%B8%E8%B2%B8%E5%8F%AA%E8%A6%81%E5%B9%BE%E5%88%86%E9%90%98-123729970.html。

台新銀行（2021），〈台新銀「手 t 貸」AI 智能徵審 數位信貸撥款加速 10 倍〉，載於：https://www.taishinbank.com.tw/TSB/personal/common/news/TSBankNews-002349/。

台新銀行（2025），〈台新三十 智能客服 Rose 再升級 領先改版照顧無障礙需求〉，載於：https://www.taishinbank.com.tw/TSB/personal/common/news/TSBankNews-003658/。

自由時報（2025），〈機器人理財資產規模達 131 億 客戶數破 21.6 萬人〉，載於：https://ec.ltn.com.tw/article/breakingnews/5022828。

吳品樺（2024），〈隱私強化技術：平衡資料保護與資料應用〉，數位發展部，https://moda.gov.tw/press/multimedia/blog/12810。

呂依舫（2021），〈【美股研究報告】金融科技新創 Affirm 的魅力何在？與美國前三大電商合作，更成為亞馬遜獨家 BNPL 供應商！〉，《CM Money 投資網誌》，載於：https://www.forecastock.tw/article/yvonnelu-6b5a5686-d959-11ef-89f2-04e31b60a914。

林宏軒、李肇棠、江滄明（2018），〈深度學習的訓練資料準備與平台之演進發展〉，《電腦與通訊》，載於：https://ictjournal.itri.org.tw/xcdoc/cont?xsmsid=0M236556470056558161&sid=0M265628861865857206。

doi:10.6342/NTU202600641

金融監督管理委員會（2025），〈金管會公布金融業應用人工智慧(AI)調查結果〉，載於：https://www.fsc.gov.tw/ch/home.jsp?id=96&parentpath=0,2&mcustomize=news_view.jsp&dataserno=202505200001&dtable=News。

經濟日報（2024），〈智能理財客戶 小資族最多 多數選擇定期定額〉，載於：https://money.udn.com/money/story/5613/8306733。

遠見雜誌（2025），〈AI 法案之父：法案漏洞多，監管仍待完善〉，載於：https://www.ctee.com.tw/news/20250103700988-431003。

數位時代（2021），〈中國信託以客戶視角發展聊天機器人 機器人客服小 C 會傳 LINE、講電話 多元溝通模式滿足客戶需求〉，載於：https://www.bnext.com.tw/article/63044/bankingmyway-07。

聯合新聞網（2025），〈台灣 3 家金控躋身全球百大銀行〉，https://udn.com/news/story/6839/8638432。

doi:10.6342/NTU202600641

英文文獻

一、 專書

CONGDON, WILLIAM J., et al. (2011), POLICY AND CHOICE: PUBLIC FINANCE THROUGH THE LENS OF BEHAVIORAL ECONOMICS.

COOTER, ROBERT & THOMAS ULEN (2016), LAW AND ECONOMICS, 6th ed.

EL EMAM, KHALED (2020), ACCELERATING AI WITH SYNTHETIC DATA.

HRADEC, J. et al. (2022), MULTIPURPOSE SYNTHETIC POPULATION FOR POLICY APPLICATIONS.

LYNSKEY, O. (2015), THE FOUNDATIONS OF EU DATA PROTECTION LAW.

MCCARTHY, JOHN (1997), WHAT IS ARTIFICIAL INTELLIGENCE?

NEGNEVITSKY, MICHAEL(2011), ARTIFICIAL INTELLIGENCE: A GUIDE TO INTELLIGENT SYSTEMS, 3rd ed.

SHAVELL, STEVEN (2004), FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW.

SOLOVE, DANIEL J. & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 871 (7th ed. 2023).

THALER, RICHARD H. & CASS R. SUNSTEIN (2009), NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS (Revised & Expanded ed.).

TOM BIGHAM ET AL(2024), AI AND RISK MANAGEMENT: INNOVATING WITH CONFIDENCE, DELOITTE 8.

二、 期刊論文

Abioye, S. et al. (2021), Artificial Intelligence in the Construction Industry: A Review of Present Status, Opportunities and Future Challenges, 44 JOURNAL OF BUILDING ENGINEERING 10.

doi:10.6342/NTU202600641

Acquisti, A. et al.(2016), The Economics of Privacy, 54 JOURNAL OF ECONOMIC LITERATURE 442.

Akerlof, G.A. (1970), *The Market for Lemons: Quality Uncertainty and the Market Mechanism*, 84(3) THE QUARTERLY JOURNAL OF ECONOMICS 488.

Akerlof, George A. (1970), The Market for Lemons: Quality Uncertainty and the Market Mechanism, 84(3) QUARTERLY JOURNAL OF ECONOMICS 488.

Beaulieu-Jones, B.K. et al. (2019), *Privacy-Preserving Generative Deep Neural Networks Support Clinical Data Sharing*, 12(7) CIRC CARDIOVASC QUAL OUTCOMES 2.

Bellovin, S.M. et al. (2019), Privacy and Synthetic Datasets, 22 STANFORD TECHNOLOGY LAW REVIEW 1.

Bourtoule, Lucas et al. (2021), Machine Unlearning, IEEE SYMPOSIUM ON SECURITY AND PRIVACY 141.

Brown, T.B. et al. (2020), Language Models Are Few-Shot Learners, 33 ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 1877.

Bubb, Ryan & Richard H.P. (2014), How Behavioral Economics Trims Its Sails and Why, 127 HARVARD LAW REVIEW 1593.

Calo, Ryan (2014), Digital Market Manipulation, 82(4) GEORGE WASHINGTON LAW REVIEW 995.

Cao, Yinzhi et al. (2015), Towards Making Systems Forget with Machine Unlearning, IEEE SYMPOSIUM ON SECURITY AND PRIVACY 463.

Chen, Xue-Wen & Xiaotong Lin (2014), Big Data Deep Learning: Challenges and Perspectives, 2 IEEE ACCESS 514.

Chong Huang et al. (2017), Context-Aware Generative Adversarial Privacy, 19(12) ENTROPY 656.

doi:10.6342/NTU202600641

Chowdhery, A. et al. (2022), PaLM: Scaling Language Modeling with Pathways, 240 THE JOURNAL OF MACHINE LEARNING RESEARCH 11324.

Chundawat, V.S. et al. (2023), Zero-Shot Machine Unlearning, 18 IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY 2345.

COASE, Ronald H. (1960), The Problem of Social Cost, 3 JOURNAL OF LAW & ECONOMICS 837.

Cohen, J.E. (2000), Examined Lives: Informational Privacy and the Subject as Object, 52(5) STANFORD LAW REVIEW 1373.

Craswell, Richard (1993), Property Rules in Unconscionability and Related Doctrines, 60(1) UNIVERSITY OF CHICAGO LAW REVIEW 1.

Elizabeth, B.K. (2002), Privacy and Financial Institutions: Current Developments concerning the Gramm-Leach-Bliley Act of 1999, 21 ANNUAL REVIEW OF BANKING LAW 167.

Emam, K.E. et al. (2020), Evaluating Identity Disclosure Risk in Fully Synthetic Health Data: Model Development and Validation, 22(11) JOURNAL OF MEDICAL INTERNET RESEARCH e23139, 736.

Fernández, Ana (2019), Artificial Intelligence in Financial Services, 2 ECONOMIC BULLETIN 4.

Gal, M.S. & Orla Lynskey (2024), Synthetic Data: Legal Implications of the Data-Generation Revolution, 109 IOWA LAW REVIEW 1087.

Ginart, A.A. et al. (2019), Making AI Forget You: Data Deletion in Machine Learning, 32 ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 3518.

Hert, P.D. & Vagelis P. (2016), The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals?, 32 COMPUTER LAW & SECURITY REVIEW 179.

doi:10.6342/NTU202600641

Hinton, G. E. & Salakhutdinov, R. R. (2006), Reducing the Dimensionality of Data with Neural Networks, 313 SCIENCE 504.

Hinton, G. E., et al. (2006), A Fast Learning Algorithm for Deep Belief Nets, 18 NEURAL COMPUTATION 1527.

Jolls, Christine et al. (1998), A Behavioral Approach to Law and Economics, 50(5) STANFORD LAW REVIEW 1471.

Kahneman, Daniel & Amos Tversky (1974), Judgment under Uncertainty: Heuristics and Biases, 185 SCIENCE 1124.

Kahneman, Daniel & Amos Tversky (1979), Prospect Theory: An Analysis of Decision Under Risk, 47(2) ECONOMETRICA 263.

Kalyan, K.S. (2024), A Survey of GPT-3 Family Large Language Models Including ChatGPT and GPT-4, 6 NATURAL LANGUAGE PROCESSING JOURNAL 100048.

Lambert, Thomas A. (2017), From Gadfly to Nudge: The Genesis of Libertarian Paternalism, 82 MISSOURI LAW REVIEW 623.

LeCun, Y., et al. (2015), Deep Learning, 521 NATURE 436.

Leitner, Georg, et al. (2024), The Rise of Artificial Intelligence: Benefits and Risks for Financial Stability, FINACIAL STABILITY REVIEW 104.

McCarthy, John (1978), History of Lisp, 13 SIGPLAN NOTICES 217.

McCulloch, Warren S. & Walter Pitts (1943), A Logical Calculus of the Ideas Immanent in Nervous Activity, 5 BULLETIN OF MATHEMATICAL BIOPHYSICS 115.

Mills, Kevin (2022), Consent and the Right to Privacy, 39 JOURNAL OF APPLIED PHILOSOPHY 721.

Narechania, T. N. & Sitaraman, G. (2023), An Antimonopoly Approach to Governing Artificial Intelligence, 24(9) VANDERBILT LAW RESEARCH PAPER.

Nguyen, Q.P. et al. (2020), Variational Bayesian Unlearning, 33 ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 16025.

doi:10.6342/NTU202600641

Patel, S. et al. (2021), Colour Image Encryption Based on Customized Neural Network and DNA Encoding, 33 NEURAL COMPUTING AND APPLICATIONS 14533.

Qu, Youyang et al. (2024), Learn to Unlearn: Insights into Machine Unlearning, 57(3) COMPUTER 79.

Schnitzler, T. et al. (2021), SoK: Managing Longitudinal Privacy of Publicly Shared Personal Online Data, 2021 PROCEEDINGS ON PRIVACY ENHANCING TECHNOLOGIES 229.

Schwartz, P.M. (1997), Privacy and the Economics of Personal Health Care Information, 76(1) TEXAS LAW REVIEW 1.

Schwartz, P.M. (2004), Property, Privacy, and Personal Data, 117(7) HARVARD LAW REVIEW 2056.

Shokri, R. et al. (2017), Membership Inference Attacks Against Machine Learning Models, IEEE SYMPOSIUM ON SECURITY & PRIVACY 3.

Simon, Herbert A. (1955), A Behavioral Model of Rational Choice, 69 QUARTERLY JOURNAL OF ECONOMICS 99.

Sommer, Daniel M. et al. (2022), Athena: Probabilistic Verification of Machine Unlearning, 3 PROCEEDINGS ON PRIVACY ENHANCING TECHNOLOGIES 268.

Sovern, Jeff (1999), Opting in, Opting out, or No Options at All: The Fight for Control of Personal Information, 74(4) WASHINGTON LAW REVIEW 1033.

Stiglitz, Joseph E. (1989), Markets, Market Failures, and Development, 79 AMERICAN ECONOMIC REVIEW 197.

Sunstein, Cass R. (2013), The Storrs Lectures: Behavioral Economics and Paternalism, 122(7) YALE LAW JOURNAL 1826.

Sunstein, Cass R. et al. (1998), A Behavioral Approach to Law and Economics, 50 STANFORD LAW JOURNAL 1471.

doi:10.6342/NTU202600641

Thaler, Richard H. (2016), Behavioral Economics: Past, Present, and Future, 106(7) AMERICAN ECONOMIC REVIEW 1577.

Tversky, Amos & Daniel Kahneman (1973), Availability: A Heuristic for Judging Frequency and Probability, 5(2) COGNITIVE PSYCHOLOGY 207.

Tversky, Amos & Daniel Kahneman (1991), Loss Aversion in Riskless Choice: A Reference-Dependent Model, 106 QUARTERLY JOURNAL OF ECONOMICS 1039.

Wu, Yinjun et al. (2020), Deltagrad: Rapid Retraining of Machine Learning Models, 119 PROCEEDINGS OF THE 37TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING 10355.

Xu, Jie et al. (2021), Machine Unlearning: Solutions and Challenges, 14(8) JOURNAL OF LATEXX CLASS FILES 1.

Xu, Yongjun et al.(2021), Artificial Intelligence: A Powerful Paradigm for Scientific Research, 2(4) THE INNOVATION 100179.

Zeller, Bruno et al. (2019), The Right to Be Forgotten—The EU and Asia Pacific Experience (Australia, Indonesia, Japan and Singapore), 2019(1) EUROPEAN HUMAN RIGHTS LAW REVIEW 23.

Zouari, G. & Marwa A. (2021), Customer Satisfaction in the Digital Era: Evidence from Islamic Banking, 10(9) JOURNAL OF INNOVATION AND ENTREPRENEURSHIP 2.

三、 研討會論文

Brophy, Jonathan & Daniel Lowd (2021), Machine Unlearning for Random Forests, 139 PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MACHINE LEARNING 1092.

Chen, M. et al. (2021), When Machine Unlearning Jeopardizes Privacy, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MACHINE LEARNING 896.

Guo, Changsheng et al. (2020), Certified Data Removal from Machine Learning Models, 359 INTERNATIONAL CONFERENCE ON MACHINE LEARNING 3832.

doi:10.6342/NTU202600641

Marchant, N.G. et al. (2022), Hard to Forget: Poisoning Attacks on Certified Machine Unlearning, 36(7) PROCEEDINGS OF THE AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE 7691.

Mitchell, Eric et al. (2022), Memory-Based Model Editing at Scale, 162 PROCEEDINGS OF THE 39TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING 15817.

Rebedea T. et al. (2023), *Nemo Guardrails: A Toolkit for Controllable and Safe LLM Applications with Programmable Rails*, in CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING 431.

四、 專書論文

CAMERER, COLIN F. et al. (2004), BEHAVIORAL ECONOMICS: PAST, PRESENT, FUTURE, in ADVANCES IN BEHAVIORAL ECONOMICS 1(Colin F. Camerer et al. eds.).

Kahneman, Daniel et al. (2004), Experimental Tests of the Endowment Effect and the Coase Theorem, in ADVANCES IN BEHAVIORAL ECONOMICS 1325 (Colin F. Camerer et al. eds.).

Thaler, Richard H. et al. (2013), Choice Architecture, in THE BEHAVIORAL FOUNDATIONS OF PUBLIC POLICY 428 (Eldar Shefir ed.).

Viale, Riccardo (2018), Understanding Financial Behaviour for Better Policy Making: An Introduction, in THE BEHAVIOURAL FINANCE REVOLUTION 2 (Riccardo Viale et al. eds.).

五、 案例

Court of Justice of the European Union, Google Spain SL – Agencia Española de Protección de Datos (AEPD), Case C-131/12, ECLI:EU:C:2014:317 (Judgment of May 13, 2014).

doi:10.6342/NTU202600641

French Conseil d'État, Ligue des droits de l'Homme – Région PACA, Decision No. 426311 (Judgment of Feb. 27, 2020).

六、 決議、解釋、函令或研究意見

Article 29 Data Protection Working Party (2014), Opinion 05/2014 on Anonymisation Techniques (WP216).

Barberá Isabel (2025), AI Privacy Risks & Mitigations – Large Language Models (LLMs), EUROPEAN DATA PROTECTION BOARD.

Bernanke, Ben S. (2010), Causes of the Recent Financial and Economic Crisis, TESTIMONY, FEDERAL RESERVE.

European Data Protection Board (2024), Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models.

European Data Protection Board(2024), Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, Version 1.0.

Organisation for Economic Cooperation and Development (2019), Recommendation of the Council on Artificial Intelligence, OECD.

Sartor, Giovanni (2020), The Impact of the General Data Protection Regulation (GDPR) on Artificial Intelligence, EUROPEAN PARLIAMENTARY RESEARCH SERVICE, Study No. 641530.

Shrishak, Kris (2024), AI: Complex Algorithms and Effective Data Protection Supervision – Bias Evaluation, EUROPEAN DATA PROTECTION BOARD.

Shrishak, Kris (2024), AI: Complex Algorithms and Effective Data Protection Supervision – Effective Implementation of Data Subjects' Rights, EUROPEAN DATA PROTECTION BOARD.

U.S. Department of the Treasury (2024), Managing Artificial Intelligence-Specific Cybersecurity Risks in the Financial Services Sector.

doi:10.6342/NTU202600641

HARVARD BUSINESS REVIEW ANALYTIC SERVICES (2021), THE EXECUTIVE'S GUIDE TO ACCELERATING ARTIFICIAL INTELLIGENCE AND DATA INNOVATION WITH SYNTHETIC DATA.

World Intellectual Property Organization (WIPO) (2019), WIPO Technology Trends 2019: Artificial Intelligence.

八、 網路資料與其他

Ahmed, Nabilaa et al., Deepfake Imposter Scams Are Driving a New Wave of Fraud, BLOOMBERG(Aug. 21, 2023), https://www.bloomberg.com/news/articles/2023-08-21/money-scams-deepfakes-ai-will-drive-10-trillion-in-financial-fraud-and-crime.

Bousquette, Isaballe, Deepfakes Are Coming for the Financial Sector, WALL STREET JOURNAL (Apr. 3, 2024 7:00 AM), https://www.wsj.com/articles/deepfakes-are-coming-for-the-financial-sector-0c72d1e5

Brightwood, S. & Henry J. (2024), Data Privacy, Security, and Ethical Considerations in AI-Powered Finance, RESEARCH GATE 1,2, https://www.researchgate.net/publication/379078709_Data_privacy_security_and_ethical_considerations_in_AI-powered_finance.

Copeland, Michael, What's the Difference Between Artificial Intelligence, Machine Learning and Deep Learning?, NVIDIA (July 29, 2016), https://blogs.nvidia.com/blog/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/.

Deloitte, Generative AI is Expected to Magnify the Risk of Deepfakes and Other Fraud in Banking, https://www2.deloitte.com/us/en/insights/industry/ financial-

doi:10.6342/NTU202600641

services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html.

Gartner, Gartner Identifies Top Trends Shaping the Future of Data Science and Machine Learning (Aug. 1, 2023), https://www.gartner.com/en/newsroom/ press-releases/2023-08-01-gartner-identifies-top-trends-shaping-future-of-data-science-and-machine-learning.

Koubaa, A., GPT-4 vs. GPT-3.5: A Concise Showdown, (Mar. 24, 2023), https:// doi.org/10.20944/preprints202303.0422.v1.

Manab, Meem Arafat, Eternal Sunshine of the Mechanical Mind: The Irreconcilability of Machine Learning and the Right to be Forgotten (Mar. 6, 2024), https://arxiv.org/abs/2403.05592.

Mejia, Niccolo, AI for Credit Scoring – An Overview of Startups and Innovation, EMERJ (Dec. 10, 2018), https://emerj.com/ai-sector-overviews/ai-for-credit-scoring-an-overview-of-startups-and-innovation/.

Newton Investment Management, Synthetic Data: Fuel for the AI World (May 22, 2023), https://www.newtonim.com/us-institutional/insights/blog/synthetic-data-fuel-for-the-ai-world/.

Radford, A., et al. (2019), Language Models are Unsupervised Multitask Learners, OPENAI BLOG, https://cdn.openai.com/better-language-models/language_models _are_unsupervised_multitask_learners.pdf.

RISE Research Institutes of Sweden, Generative AI Does Not Run on Thin Air!, RISE (Feb. 21, 2024), https://www.ri.se/en/news/blog/generative-ai-does-not-run-on-thin-air.

Rothmann, C.C., Protecting Data Privacy: A Baseline for Responsible AI, CSIS, (Jul. 18, 2024), https://www.csis.org/analysis/protecting-data-privacy-baseline-responsible-ai.

doi:10.6342/NTU202600641

Savage, Neil, Breaking into the Black Box of Artificial Intelligence, NATURE (Apr. 1, 2022), https://doi.org/10.1038/d41586-022-00858-1.

Slinas, Diego, Intelligent Customer Service in 2025: A Guide on Benefits & Implementation, CLOUDTALK (Dec. 20, 2024), https://www.cloudtalk.io/blog/intelligent-customer-service.

Swire, P.P., Markets, Self-Regulation, and Government Enforcement in the Protection of Personal Information, NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION, https://www.ntia.doc.gov/page/chapter-1-theory-markets-and-privacy.

Vondohlen, Eric, How Does AI-Based Credit Scoring Fare Against Traditional Credit Scoring?, FINEXTRA (Nov. 27, 2018), https://www.finextra.com/blogposting/16343/how-does-ai-based-credit-scoring-fare-against-traditional-credit-scoring.

White House Office of Science and Technology Policy, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People (2022), https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/.

Zhang, Dawen et al. (2024), Right to Be Forgotten in the Era of Large Language Models: Implications, Challenges, and Solutions**,** arXiv:2307.03941.

Zhao, Wayne Xin et al.(2025), A Survey of Large Language Models (version 16), https://arxiv.org/abs/2303.18223.

doi:10.6342/NTU202600641

德文文獻

一、 專書

BÄCKER, M. (2006). INFORMATIONELLE SELBSTBESTIMMUNG UND INFORMATIONSTECHNISCHE SYSTEME, MOHR SIEBECK.

HOFFMANN, C. et al. (2015). DIE DIGITALE DIMENSION DER GRUNDRECHTE – DAS GRUNDGESETZ IM DIGITALEN ZEITALTER, NOMOS.

二、 期刊論文

Bäcker, M. (2012). Grundrechtlicher Informationsschutz gegen Private. 51(1) DER STAAT, 91.

Luch, A. D. et al. (2014). Ein Recht auf Vergessenwerden als Ausprägung einer selbstbestimmten digitalen Persönlichkeit – Anmerkung zum Urteil des EuGH v. 13.5.2014 (Google), 49(6) EUROPARECHT (EUR) 698.

Von Lewinski, K. (2015). Staat als Zensurhelfer – Staatliche Flankierung der Löschpflichten Privater nach dem Google-Urteil des EuGH, ARCHIV FÜR PRESSERECHT (AFP) 1.

doi:10.6342/NTU202600641