人工智慧簡介、監管及國際法制發展趨勢

中原大學財經法律學系助理教授 蔡鐘慶

楔子

什麼叫做人工智慧（Artificial Intelligence, AI）？其實AI跟一般我們對電腦認知不太一樣。一般認為電腦只是幫我們進行計算、存取資料，或是播放影片。人工智慧的概念是希望電腦能夠表現出類似人的智慧，例如感知、推理跟處理事務的能力。早期學者把人工智慧分成幾個重要的研究項目，包括智慧的遊戲、機器人、電腦視覺、專家系統等，另外類神經系統是一種讓電腦由個案中學習能力的技術，後來發展為深度學習技術[1]。

依據國家科學及技術委員會（下稱國科會）之前身科技部2019年9月所發布之人工智慧科研發展指引（AI Technology R&D Guidelines）中指出AI作為引領未來的關鍵浪潮，對人類社會、經濟、政治生活造成重大改變，並於醫療、教育、自駕車等各領域均引起破壞式創新。為持續創造AI正面效益，應連結政府、產業、研發人員、民眾等利害關係人，共同打造值得信賴之AI環境，創立「以人為本」、「永續發展」及「多元與包容」三大核心價值之AI社會[2]。

從1956年艾倫・圖靈（Alan Turing）發表了文章《Computing Machinery and Intelligence》概念後，科學家便不斷追求更為先進的AI技術發展。過去 AI集中在弱人工智慧（Weak AI），即僅具備單一專長並應用於特定領域的AI，能力範圍相對有限。然而，隨著自然語言處理技術的進步，使大型語言模型能夠處理多領域任務，打破AI只能在單一領域運作的限制，標誌著AI技術正朝向強人工智慧（Strong AI）或通用人工智慧（Artificial General Intelligence, AGI）邁進。特別是2022年11月OpenAI將ChatGPT開放給大眾使用後，帶出了生成式AI（Generative AI）的新序幕，在各種文字、圖片、程式碼、聲音／音樂等生成下，

1 黃國禎，人工智慧的發展與教育應用，人文與社會科學簡訊第23卷第1期，2021年12月，頁1-2。

2 科技部，人工智慧科研發展指引（AI Technology R&D Guidelines），2019年9月版，頁1。

生成式對整個社會帶來重大的衝擊。根據國際機構研究Gartner指出，2024年科技趨勢之重點在於生成式AI及其影響，預測在2026年有超過八成企業會在營運環節中部署生成式AI，以此提高生產力及提升服務效能[3]。

自從歐盟執委會於2021年4月提出人工智慧法草案以來，其後續發展備受全球矚目，也吸引歐洲的人權組織、學術團體以及大型科技公司的關注。在多方利益關係者的遊說與介入下，該法案一度陷入僵局，其中生成式人工智慧（Generative AI）亦為爭議焦點。歐洲議會和理事會的人工智慧法草案修正版本中，曾經納入生成式AI的定義與監管條款，然而最後拍板定案以AI系統與基礎模型為監管對象，並未針對生成式AI。理事會、執委會和歐洲議會經過多次三方會談，終於在2023年12月8日就內容達成協議，草案在2024年3月13日交由歐洲議會大會表決，最終以壓倒性的票數通過該法[4]。歐盟理事會於2024年5月22日正式批准《人工智慧法》（Artificial Intelligence Act，下稱AIA），該法於2024年7月12日公告於歐盟的官方公報上，自同年8月1日起生效，成為全球首部全面性監管AI的法律框架。歐盟在2024年通過了AI法案（AI Act）。如果說2024年是AI立法元年，那麼2025年就是AI相關規範落地實施元年。隨著歐盟AI法案於2025年2月1日正式實施，並推動AI責任指令（AILD），各國將在歐盟的引領下，開始思索如何規範AI[5]。

美國加州議會則於去（2024）年8月28日通過《尖端人工智慧模型安全與創新法》（Safe and Secure Innovation for Frontier Artificial Intelligence Models Act，編號SB-1047）草案（下稱加州人工智慧模型法草案），以因應尖端人工智慧技術，例如生成式人工智慧（Generative artificial intelligence）對於社會造成之潛在風險、威脅或負面影響，同時促進人工智慧技術發展於促進創新與防範風險間取得平衡。由於加州聚集美國眾多的新創產業，故加州的立法對於美國先進人工智慧技術之監管將產生重要影響[6]，儘管最後被加州州長否決，但也凸顯了生成式人工智慧的重要性。我國國科會也於2024年7月15日公告：預告制定「人工智慧基本法」草案[7]，均顯示人工智慧的重要性及監管的迫切性。

---

3 國家科學及技術委員會，「人工智慧（AI）推動現況與未來方向」專題報告，立法院第11屆第1會期教育及文化委員會，2024年6月5日，頁1。

4 沈娟娟，歐盟公布人工智慧法，建立全球首部AI全面監管框架，資策會科法所，2024年7月12日，available at https：//stli.iii.org.tw/article-detail.aspx?no=64&tp=1&d=9213（last visited 2025.2.16）

5 馮震宇，2025 AI 法規實施元年：歐盟AI法案領軍國際治理趨勢，能力雜誌第827期，2025年1月，頁95。

6 張允亭，美國加州通過《尖端人工智慧模型安全與創新法》草案，為尖端AI模型建立嚴格的安全標準，財團法人電信技術中心，2024年12月11日，available at https：//www.ttc.org.tw/News/more?id=43b9586b9ef8424e8e6df79c1aa48009（last visited 2025.2.16）

7 中華民國113年7月15日科會前字第1130048999號。

AI（人工智慧）的監管

AI的快速發展對社會、經濟及倫理層面帶來深遠影響。然而，由於AI技術的複雜性及其潛在風險，監管問題成為全球關注的議題：AI技術的應用涵蓋醫療、金融、交通、國防等多個領域，帶來許多便利，但同時也伴隨風險。例如，深度學習技術可能產生偏見（bias），導致歧視性決策；自動化決策系統的黑箱問題（black box）使結果難以解釋。此外，對個人隱私的影響也不容忽視。因此，制定合理的監管框架以降低風險至關重要。目前，各國及國際組織已針對AI監管提出不同框架與法規例如OECD《人工智慧原則》、美國《人工智慧權利法案》、歐盟《人工智慧法》：

1. **OECD《人工智慧建議書》**：經濟合作與發展組織（OECD）提出AI發展應符合人權、透明度與可問責性（OECD, 2019）[8]。
2. **美國《人工智慧權利法案》**（Blueprint for an AI Bill of Rights）：美國白宮於2022年發布該法案，旨在保護個人權益，確保AI技術透明、公正與問責（White House, 2022）[9]。
3. **歐盟《人工智慧法》**（AI Act）：此法案為全球首部全面監管AI的法規，採用風險分級管理，對高風險AI應用設立嚴格規範[10]，該法已於2024年7月12日公告於歐盟的官方公報上，自同年8月1日起生效。

儘管AI監管已取得初步進展，但仍面臨諸多挑戰，例如技術發展速度快於監管立法，AI技術快速演進，使監管法規難以跟上創新步伐；全球標準不一致：不同國家對AI監管的立場各異，導致國際間協調困難；此外過度監管可能抑制創新，而監管不足則可能帶來倫理與安全問題。特別是許多AI系統難以解釋決策過程，如何確保責任歸屬成為最大挑戰。為因應上述挑戰，AI監管或許可從以下幾個方向思考：

1. **促進國際合作**：建構全球統一的人工智慧（AI）監管框架，確立技術發展的法規基礎與公平競爭機制，以降低跨國監管差異對技術發展與市場公平性的影響。透過多邊國際協調機制，促進各國在監管標準協同合作，以提高全球治理的整體效能。
2. **強化透明度與可解釋性**：要求人工智慧開發與應用機構揭露其決策機制，確保模型運行的可解釋性與可追溯性，以提升社會對技術的信任度，特別是AI所生成之決策對於利害關係人有重大影響，為保障決策過程之公正性，在AI系統、軟體及演算法等技術

---

8 OECD Principles on AI, available at https：//www.oecd.org/going-digital/ai/principles/（last visited 2025.2.16）

9 Blueprint for an AI Bill of Rights, available at https：//www.whitehouse.gov/ostp/ai-bill-of-rights/

10 European AI Act, available at https：//digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai（last visited 2025.2.16）

發展與應用上，包括但不限於對於模組、機制、組成、參數及計算等進行最低限度的資訊提供與揭露，以確保一般人得以知悉人工智慧系統生成決策之要素，此外AI發展與應用階段，應致力權衡決策生成之準確性與可解釋性[11]。

3. **建立動態監管機制**：採取適應性監管框架（adaptive regulatory framework），確保監管政策能夠隨著人工智慧技術的快速演進而調整，以兼顧創新與風險管控。例如金融業運用人工智慧（AI）指引即指出持續監控與精進：金融機構宜對已部署之 AI 系統進行維護、監控、記錄及審查，及依據風險評估考量因素提供適當資源，並使管理階層瞭解已部署之 AI 系統的表現及其他相關問題。在適當的情況下，監控可以包括自主監控，例如AI系統可以被設計為自動報告其預測的信賴水準，並定期審查風險管理機制，以促進其有效性，如建立內部審查與監測機制，依風險基礎定期評估AI系統是否符合原先運用目的及風險程度，以使AI系統符合政策與指導方針，並及時解決可能存在之問題[12]。

4. **推動公私協力合作**：促進政府、企業及學術機構間的多方治理（multi-stakeholder governance）機制，共同參與人工智慧監管政策的制定與執行。透過公私協力（public-private partnership），平衡監管要求與技術創新需求，確保政策能夠反映不同利害關係人的利益，並提升AI治理框架的適應性與包容性，例如金融機構必要時可邀請不同領域人員參與 AI 評估過程，包括人力資源、行為科學、法律、倫理、永續發展等領域，以協助為AI之發展提供正確的方向[13]。

AI國際法制發展趨勢

我國國科會於2024年7月15日預告制定「人工智慧基本法」草案，該草案總說明中提到人工智慧技術雖帶來社會及經濟效益，同時也可能對個人或社會帶來新的風險或影響。鑑於人工智慧技術創新之速度及可能面臨之挑戰，全球主要國家皆致力在不妨礙技術發展下，尋求建立人工智慧之治理方針與原則。經濟合作暨發展組織（OECD）於2019年5月通過「人工智慧建議書」，提出基本價值原則，並給予各國政策制訂者相關建議；同年歐盟發布「可信賴人工智慧倫理準則」（Ethics Guidelines for Trustworthy AI），確保人工智慧發展所需之共同倫理原則。於此之後，如歐盟於2021年提出「人工智慧法」（Artificial Intelligence Act），2024年通過審議、美國於2022年發布「AI權利法案藍圖」（Blueprint for an AI Bill of Rights）、加拿大亦於2022年提出「人工智慧資料法草案」（Artificial Intelligence and Data Act），皆著重於建立人工智慧技術發展之原則並建立大眾信任；美國白宮於2023年發布「發

[11] 科技部，人工智慧科研發展指引（AI Technology R&D Guidelines），2019年9月版，頁3。

[12] 金管會，金融業運用人工智慧（AI）指引，2024年6月，頁8。

[13] 同前註。

展與使用安全且可信任的AI行政命令」（Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence）訂立聯邦各部門人工智慧發展之推動任務[14]，美國加州議會更是於2024年8月28日通過《尖端人工智慧模型安全與創新法》草案，對於生成式人工智慧加以規範（法案被州長否決）。

上述為已通過實施之法案係歐盟人工智慧法，其他多為草案或行政命令，依據經濟部國際貿易署對於歐盟人工智慧法簡介，該簡介表示：近年AI技術逐漸成熟，並可簡易地落實在當今社會經濟的各個部門，其影響可跨越歐盟會員國間之國境，歐盟各會員國間對於AI監管之不一致性恐破壞單一市場，並對使用、進口、開發AI系統等行為增加法規不確定性，歐盟執委會於2021年4月提出歐盟AI法案，該法案係全球首部針對AI之全面性法規，旨在確保歐盟開發及使用之AI係值得信賴，提供保護人們基本權利之保障措施，於歐盟建立一致的AI市場，並為相關技術之投資與創新創造支持性環境，該法案已於2024年8月1日生效，適用歐盟全體會員國[15]：

（一）以風險為基礎：歐盟AI法案（下稱法案）採用基於風險的方法來監管AI系統，將AI引發之潛在風險區分不同級別。

（二）適用對象：依據法案第2條及第3條規定，自AI供應鏈各階段之提供者、進口商、經銷商、佈署者等，及歐盟境內受影響之自然人，皆為本法案適用對象，另依AI用途目的設有免適用情況。

（三）適用對象之義務：對於高風險AI系統及具系統風險之通用AI模型，法案針對相關人士規定不同之義務如下：

（1）高風險AI系統：包含提供者之義務、歐盟境外第三國之提供者之義務、進口商之義務、經銷商之義務、部署者之義務、AI價值鏈之責任。

（2）具系統風險之通用AI模型：依據法案第 3 條規定，通用AI（GPAI）模型係指一種AI模型，包括使用大規模數據進行自我監督 （self-supervision）訓練之AI模型，該模型顯示出顯著的通用性（generality），且無論以何種態樣進入市場，皆能夠執行各種不同的任務，並可整合到下游各種系統或應用程式中，但排除在進入市場前用於研究、開發、試驗活動（prototyping activities）之AI模型，規範相關提供者之義務。

---

[14] 參閱國科會2024年7月15日發布之人工智慧基本法草案總說明。

[15] 許莉美，歐盟人工智慧法案（AI Act）生效一文之附件，經濟部國際貿易署，2024年8月23日，available at https：//www.trade.gov.tw/Pages/Detail.aspx?nodeID=45&pid=789058（last visited 2025.2.16）

歐盟AI法案採3階段實施模式[16]：

（1）第1階段：AI法案通則及不可接受風險的AI系統禁令，已在2025年2月10日實施。

（2）第2階段：通用AI模型、第三方認證機構、會員國公告合格評估機構，以及違反法案罰則，將於2025年8月1日實施。

（3）第3階段：該法案附件III清單中的高風險AI系統的相關義務，將於2027年8月1日實施，生效前已上市的通用AI模型提供者，應在36個月內符合AI法案規定。

代結論

歐盟推出一套全面的人工智慧規則，以確保用戶安全和責任，而美國在川普總統領導下卻朝著相反的方向發展，放寬限制AI監管，並賦予科技行業更多影響政策的權力，這可從川普總統的就職典禮上，特斯拉與Space X的執行長馬斯克、Open AI執行長阿特曼以及Meta執行長祖克柏等科技界億萬富翁坐在緊挨川普家人的第二排，並在同一天，川普總統撤銷了其前任拜登總統所頒布的一項行政命令，取消了拜登政府期間實施的多項人工智慧管理措施和計劃。在接下來的幾天裡，川普邀請了Open AI、軟銀（Soft Bank）和甲骨文（Oracle）的執行長到白宮，宣布了他所謂的「迄今為止規模最大的人工智慧基礎設施項目」，在未來四年內，這項名為「星際之門」（Stargate）的超級項目將投資多達5000億美元用於人工智慧基礎設施。川普總統還簽署了一項行政命令，要求在180天內制定一份「人工智慧行動計劃」，該計劃旨在「維持和增強美國在人工智慧領域的主導地位」，儘管這一政策的細節尚不明確，但普遍預計科技巨頭將獲得更大的自由以開發新型人工智慧技術。這符合川普總統整體上放鬆AI監管的方向[17]。

加拿大全球事務部公布在2025年2月在舉行巴黎的全球AI峰會上簽署歐洲委員會的「人工智慧與人權、民主和法治架構公約」。若國會批准該條約，將加強對AI發展和使用的監管，包括要求AI公司透明化並揭露資訊，違反者將面臨嚴厲刑事責任，表彰加拿大致力以法律確保AI發展在可靠範圍內進行。反觀美國副總統JD Vance拒絕簽署聲明，強調過度監管可能會阻礙AI行業發展。儘管態度不同，巴黎峰會顯示出許多國家思考研擬可行監管措施以因應AI安全和環境問題，並逐步達成國際共識。包括加拿大、歐盟、印度和中國在內的60個國家簽署了「包容與永續人工智慧聲明」，呼籲縮小AI應用能力的差異，避免市場過度集中，強調開放和透明。此外巴黎峰會同時促成了「永續人工智慧聯盟」的成立，目標是制定AI環境影

[16] 馮震宇，2025 AI法規實施元年：歐盟AI法案領軍國際治理趨勢，能力雜誌第827期，2025年1月，頁96。

[17] 德國之聲，「川普2.0」大幅放鬆人工智慧監管，優先考慮國安與利益：歐洲還能堅持理念多久？available at https：//www.storm.mg/article/5316469（last visited 2025.2.16）

響的衡量標準，並優化演算法以減少計算複雜性，雖然不具約束力，但為未來合作提供了合作平台[18]。

我國金管會2024年6月20日發布之「金融業運用人工智慧（AI）指引」，該指引特別針對第三方業者的監督管理，金融機構重視責任的分工，明訂「金融機構運用AI系統時宜辨識可自行監控風險的程度，並對自身較無控制權的部分或事項，透過契約等方式與合作廠商明訂風險監控的責任分工」，並訂定金融機構委託第三方業者導入AI系統相關作業時宜採行的監督管理措施，包括就停止委託的情形訂定適當的資料或系統遷移機制等，以及金融機構運用AI系統時，係以風險為基礎落實核心原則，並要求金融機構應確保AI系統運作的可解釋性，但考量金融機構如委託其他業者研發或購入AI系統，可能因業者商業機密無法完全得知AI系統運作細節，因此將「可解釋性」限縮在金融機構可清楚說明「自行或委託研發並使用之AI系統」如何運作及其預測或決策過程背後之邏輯，而在透明性部分，為提升市場對金融機構AI系統的信任度，明訂金融機構如有需要，亦可規劃透過發布報告、技術文件或於網站上揭露相關資訊等方式，主動讓利害關係人（stakeholder）知悉其運用AI系統之做法[19]，是金融業都必須要瞭解與參考，而未來在歐盟人工智慧法全方面監管與美國川普政府放寬AI監理間，會激起何種火花，國際間AI監管方向亦值得我們持續注意。

蔡鐘慶

現任中原大學財經法律學系助理教授、台灣財產法暨經濟法研究協會理事

專長公司法、證券交易法、保險法、金融科技法、金融消費者保護法、永續治理，並曾於臺灣證券交易所、金融消費評議中心服務，擁有家族信託規劃顧問師、公司治理、永續發展、證券商高級業務員、投信投顧業務員等十餘張金融證照。

[18] 陳重江，AI智慧峰會彰顯加拿大對AI安全與永續性的重視，2025年2月15日經濟部國際貿易署全球商機資訊，available at https：//www.trade.gov.tw/Pages/Detail.aspx?nodeID=45&pid=797660（last visited 2025.2.16）

[19] 金管會發布「金融業運用人工智慧（AI）指引」新聞稿，2024年6月20日，available at https：//www.fsc.gov.tw/ch/home.jsp?id=96&parentpath=0,2&mcustomize=news_view.jsp&dataserno=202406200001&dtable=News（last visited 2025.2.16）