「人工智慧基本法」簡介

Introduction to the Artificial Intelligence Basic Act

文 / 臺大醫院倫理中心 周采潔

人工智慧（Artificial Intelligence, AI）成為全世界競逐焦點，廣泛應用，卻也引發隱私侵害、偏見歧視、不公平競爭、安全性疑慮、工作變遷等負面效應及風險，為利技術深耕及產業發展，2026 年 1 月 14 日經總統公布施行之「人工智慧基本法」[1]，將道德倫理、法制整備、資料處理及社會變遷等議題納入政策考量[2]，揭示我國 AI 技術發展及運用以人為本（human-centered）之基本價值、治理原則及政策方針，並責成數位發展部（下稱數發部）制定 AI 風險分類框架與指引，由各目的事業主管機關依權管事項訂定具體規範。

壹、「人工智慧基本法」介紹

一、 人工智慧之定義

本法參考美國國家 AI 創新法案（National AI Initiative Act of 2020）美國法典（U.S. Code）第 9401 章等國際組織規定，所稱人工智慧係指具自主運行能力之系統，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（第 1 條）。

二、 業務執掌及法規調適

本法中央主管機關為國家科學及技術委員會（下稱國科會）（第 2 條第 1 項），行政院應成立國家人工智慧戰略特別委員會，協調、推動及督導全國人工智慧事務，並訂定國家人工智慧發展綱領（第 6 條），由國科會辦理該委員會之幕僚作業。又本

法規範事項包含資源挹注、政府資料公開、個人資料運用、智慧財產權保障、產業發展、跨領域應用等相關法規[3]，按第 2 條第 2 項規定，涉及各目的事業主管機關職掌者，由各目的事業主管機關辦理，且政府應自本法施行後 2 年內，完成法規之制（訂）定、修正或廢止，及行政措施之改進（第 18 條）。

三、 確立研發與應用之基本原則（第 4 條）

(一) 永續發展與福祉：參考 G7 廣島 AI 國際行動規範，兼顧社會公平與環境、經濟之協調發展。

(二) 人類自主：參考經濟合作暨發展組織（OECD）2019 年公布之人工智慧建議書，應以支持人類自主權（Human Autonomy），並尊重人格權（含姓名、肖像、聲音）等個人基本權利與文化價值，確保以人為本之基本價值。

(三) 隱私保護與資料治理：應妥善保護個資隱私，並遵循資料最小化原則（data minimization），意即個人資料之蒐集須適當且具相關性，並僅止於符合資料處理目的所需之程度。另於符合憲法隱私權保障之前提下，促進非個人或非機敏資料之開放及再利用。

(四) 資安與安全：參考新加坡 2024 年生成式 AI 治理架構，應建立資安防護措施，確保系統穩健性（robustness）與安全性。

(五) 透明與可解釋：參考歐盟 2019 年可信賴 AI 倫理準則，應致力權衡決策生成之準確性，並提升使利害關係人理解其影響及決策過程之可解釋性，兼顧使用者及受影響者權益。

(六) 公平與不歧視：演算法應避免產生偏差或歧視之結果，重視社會多元包容。

(七) 問責：參考新加坡 2024 年生成式 AI 治理架構，於 AI 開發運用之生命週期中，應確保開發者、部署者、最終使用者等不同角色，均能承擔相應責任。

四、 風險分類及管理規範

(一) 風險分類層級化管理：數發部應參考國際標準或規範，推動與國際介接之 AI 風險分類框架；各目的事業主管機關應依循前開風險分類框架，識別潛在風險類別，訂定以風險為基礎之管理規範（第 16 條第 1 項）。

(二) 明確高風險應用之責任歸屬：高風險人工智慧之應用，係依據潛在風險及影響程度判斷之，為避免於特定關鍵領域應用時對人民基本權利、生命安全、財產保障或社會秩序造成嚴重危害，政府應針對其可能產生之損害風險，明確責任歸屬及歸責條件，並建立救濟、補償或保險機制（第 17 條）。另基於「聯合國兒童權利公約」之兒少最佳利益原則（Best Interests of the Child），人工智慧產品或系統經認定屬高風險應用者，應明確標示注意事項或警語（第 5 條第 2 項）。

(三) 提供驗證工具：數發部及其他相關機關應提供或建議評估驗證之工具或方法，並應徵詢相關利益團體、產業、學者、社會團體及法律專家之意見（第 5 條第 3 項及第 4 項）。

(四) 風險評估及內控管理機制：政府使用人工智慧執行業務或提供服務時，應進行風險評估及規劃風險因應措施，依業務性質，訂定使用規範或內控管理機制（第 19 條）。

五、 權益保障及資料治理

(一) 權益保障

1. 避免侵害權益：政府應避免 AI 應用造成人民生命、身體（含身心健康）、自由或財產、社會秩序、國家安全、生態環境損害，或出現偏差、歧視、廣告不實、資訊誤導或造假等違反相關法規之情事（第 5 條第 1 項）。前開情形如依現行技術手段仍無法有效管理或降低該應用風險，各目的事業主管機關應依法予以限制或禁止（第 16 條第 2 項）。
2. 個人資料保護：AI 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相關措施或機制（第 14 條）。
3. 保障勞動權益：政府應積極弭平 AI 發展所造成之技能落差，針對 AI 利用所致之失業者，依其工作能力予以輔導就業（第 15 條）。

(二) 資料治理：政府應針對 AI 訓練資料之合理利用、研發資金補助、產業扶持、研發責任之歸責範圍等，提供合理使用、扶持及補助措施（第 11 條第 1 項）。另參考歐盟人工智慧法有關支援高品質資料近用之規定，政府應進一步建立資料開放、共享及再利用機制，並提升 AI 使用資料之品質與

數量，確保訓練及產出結果足以展現國家多元文化價值與維護智慧財產權（第 13 條）。

六、 人才培育及促進創新

(一) 人才培育：厚植國民人工智慧與倫理教育知能（第 7 條），鼓勵產官學界人才及技術之跨域合作交流（第 8 條）。

(二) 基礎建設：鼓勵有利於 AI 發展之基礎設施，如建置資料中心、能源設備等（第 8 條），及推動人工智慧研發、應用與基礎建設（第 10 條）

(三) 獎補助及優惠措施：政府應辦理人工智慧相關產業之補助、委託、出資、投資、獎勵、輔導，或提供租稅、金融等財政優惠措施（第 10 條）。

(四) 實驗沙盒：各目的事業主管機關得針對 AI 創新產品或服務，建立或完備其研發及應用服務之創新實驗環境（第 11 條第 2 項），實際應用前，不適用責任歸屬相關規範（第 17 條第 2 項）。

貳、AI 風險分類框架

為兼顧 AI 創新與治理，臺灣採分層管理模式[4]，第一層「人工智慧基本法」，已建立政府推動 AI 發展之基本原則，第二層則由數發部研訂之 AI 風險分類框架，為各部會風險辨識與管理提供一致性決策邏輯，第三層為各目的事業主管機關分別就應用場域及產業性質制定規範。

數發部已依據本法第 16 條第 1 項訂定 AI 風險分類框架（草案），以檢核表協助各目的事業主管機關盤點權管業務之 AI 應用情境、風險辨識、界定高風險樣態、應對措施，預計於 2026 年 3 月底前提報行政院[5]，重點說明如下：

一、 應用情境：包括 AI 應用名稱、使用場景、主要利害關係人、系統特性等項目。

二、 風險辨識：第 1 類為 AI 系統本身的技術設計缺陷，是否涉及歧視、不公平、安全漏洞等；第 2 類為系統部署、操作及人機互動問題，包括使用者是否產生過度依賴、侵害隱私等；第 3 類為廣泛社會結構與環境衝擊，例如是否影響民眾工作權。

三、 高風險樣態界定：針對已識別的風險，進一步評估其影響對象、範圍及損害嚴重性，判斷是否為高風險樣態。

四、 應對措施：管制介入或干預程度由低到高，包括引導業者自律、透明度義務或自我揭露、事前審查或第三方驗證、禁止使用等措施，或建立高風險 AI 應用責任歸屬與救濟機制。

參、結語

「人工智慧基本法」奠定我國 AI 發展與風險治理的法制基礎，以風險分類為治理核心，所揭示永續發展與福祉、人類自主、隱私保護與資料治理、資安與安全、透明與可解釋、公平與不歧視、問責等 7 項原則，不僅提供跨部門政策整合基準，亦成為醫療 AI 應用的價值指引。

數發部訂定之 AI 風險分類框架係為 AI 治理落地的第一步，提供醫療領域原則框架與風險監管方向。鑑於醫療領域 AI 應用，涉及個人數據、網路安全、全民健康保險資料及醫療行為，尚須仰賴各目的事業主管機關如個人資料保護委員會、數發部、衛生福利部制定相關子法與技術指引，包括明確醫療 AI 風險分級與審查要求、建立健保醫療資料合法共享機制，以及完善醫師—開發者—機構間的責任分擔與保險補償制度[6]，具體落實人工智慧基本法精神。

參考文獻

1. 人工智慧基本法。根植法律網。查閱日期：2026 年 3 月 9 日。
2. 張麗卿：臺灣人工智慧基本法制定之必要與倡議。月旦法學雜誌 2023；340：79-101。
3. 章忠信：人工智慧基本法與智慧財產權之保護。金融科技/法律。2026 年 2 月。查閱日期：2026 年 3 月 9 日。
4. 溫怡玲、楊育青：數位發展部部長林宜敬 人工智慧基本法與主權 AI。循證決策集刊 2026 年 3 月 9 日，第 4 期。查閱日期：2026 年 3 月 23 日。
5. 中央社：數發部 AI 風險分類框架 勞動部金管會等 4 部會擬先檢視。工商時報。2026 年 3 月 8 日。查閱日期：2026 年 3 月 10 日。
6. 廖建瑜：醫療領域人工智慧之規範模式選擇 - 從比較法反思台灣人工智慧基本法。全國律師 2026；30(1)：16-47。