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

廖建瑜*

壹、前言

一、研究背景

近年來人工智慧（Artificial Intelligence, AI）於醫療領域之應用，已由早期之資料輔助分析與單一任務演算法，快速演進至結合機器學習（machine learning）與深度學習（deep learning）之高度複雜系統，並實際進入臨床診斷、影像判讀、治療建議、慢性病管理及手術支援等核心醫療場域。此一趨勢，不僅改變醫療行為之技術樣態，更深刻動搖傳統醫療法制所預設之從事醫療業務必為人類醫師之結構。

從比較法觀察，AI/ML-based Software as a Medical Device（SaMD），已成為當前醫療監理與責任法制中最具顛覆性的技術類型。其一方面被寄予提升診斷準確性、降低人為錯誤、促進精準醫療之高度期待，另一方面卻同時暴露出既有法制在安全性審查、責任歸屬、醫療專業監督與患者權利保障上的結構性不足[1]。尤有甚者，與傳統醫療器材或藥品不同，醫療AI系統往往具有「上市後持續學習與變異」之特性，使其實際運作狀態，可能隨資料輸入與臨床環境而不斷改變。此一動態性與不確定性，直接挑戰以靜態產品為前提所建構之醫療器材法制與責任歸責模型[2]。

在此背景下，各國逐漸意識到，醫療AI已非僅屬一般科技創新問題，而係涉及患者生命身體安全、醫療專業自主性與國家公共衛生責任之高度風險科技。歐盟透過人工智慧法（EU AI Act）將醫療AI明確列為高風險AI；義大利除繼承歐盟風險分類更具體落實在各種醫療情境使用上；韓國則在專法上亦將醫療運用列為高影響系統；日本則欲以AI醫療作為解決社會問題及少子化結構的基本建設；越南亦不同風險管制注意到AI在醫療領域的特殊性。相較之下，台灣甫通過人工智慧基本法，對醫療領域AI的風險特性與專屬規範模式，尚缺乏系統性梳理與比較法反思。

*本文作者係臺灣高等法院刑事庭審判長兼法官，國立成功大學法律學博士。

註1：Boubker, J. (2021). When medical devices have a mind of their own: The challenges of regulating artificial intelligence. American Journal of Law & Medicine, 47, 427-454.

註2：Grossbard, E. S. (2025). The AI-robotic prescription: Legal liability when an autonomous AI robot is your medical provider. University of Miami Business Law Review, 33, 273-316.

二、問題意識與提出

基於上述背景，本文之核心問題意識在於：醫療領域人工智慧是否適合納入一般性、原則性之人工智慧基本法架構中加以規範，抑或有必要發展具領域特化性質之規範模式？從比較法經驗可見，醫療AI所引發之法律問題，並非僅止於抽象之科技風險或倫理疑慮，而是具體且反覆出現於以下層面：

第一，AI診斷或治療建議錯誤時，究竟應依醫療過失法、產品責任法，或建立混合型責任架構加以處理，學界與實務仍高度分歧[3]。

第二，當AI系統逐步具備高度自主性，甚至實質取代醫師完成醫療決策時，是否仍可維持最終責任必然由醫師承擔之傳統假設，已遭嚴重質疑[4]。

第三，AI系統的黑箱性與資料偏誤，使患者難以理解醫療決策基礎，亦使事後責任追究與司法審查面臨實質障礙[5]。

然而，台灣現行及擬議中的人工智慧基本法，多著眼於一般性風險治理、倫理原則與跨領域通用義務，對醫療AI所具有之高度風險性、專業性與生命法益關聯，是否足以妥適回應，仍有重大疑問。本文即試圖透過比較法素材，反思台灣在選擇醫療AI規範模式時，是否有必要超越單一基本法中心主義，轉而採取分層、分域或混合型之法制設計。

貳、人工智慧於醫療領域之風險特性與規範必要性

人工智慧（AI）進入醫療領域後，固然被期待提升診斷效率、減少人為錯誤並促進精準醫療，但其風險樣態在法益層次、專業結構與技術屬性上，均顯著不同於一般消費性或行政性AI應用。尤有甚者，該等風險往往並非單一、可切割之技術問題，而係同時牽動患者生命身體安全、醫療專業判斷、制度性信任與責任法體系之多重面向，因而形成高度複合且難以回復之法制挑戰。

一、資料偏誤（bias）與演算法偏見（Algorithmic Bias）：由「技術瑕疵」升格為「健康不平等」之法制風險

醫療AI的預測品質高度依賴訓練資料，而醫療資料本身常受資料蒐集結構、醫療可近性與人口統計分布影響。若訓練資料未能充分代表實際使用族群，AI將可能「再製並常態化」既有不平等，使少數族群或弱勢群體面臨較高誤判率，進而擴大既存健康差距[6]。

亦有研究主張，公平性應被置於醫療AI全生命週期（設計、訓練、驗證、部署與監測）之核心地位，並建議於監理審查導入

註3：Tonti, N. (2024). Who to blame? AI or your physician. Quinnipiac Health Law Journal, 27, 219-256.

註4：Grossbard, E. S. (2025). The AI-robotic prescription: Legal liability when an autonomous AI robot is your medical provider. University of Miami Business Law Review, 33, 273-316.

註5：Hamilton, E. (2022). Transparency is a misplaced regulatory focus for holding adaptive software as medical devices (SaMDs) liable for defective design. Washburn Law Journal, 61, 571-601.

註6：Merin, S. F. (2024). Understanding the impact of artificial intelligence on the future practice of law and medicine. University of Detroit Mercy Law Review, 101, 267-286.

「公平性檢核點」以評估訓練資料與臨床證據的代表性與品質，否則安全與有效性之評估將可能在弱勢族群上失真[7]。此一風險之所以構成規範必要性，在於醫療AI的偏誤並非僅造成使用者體驗下降，而可能造成延誤診斷、錯誤治療與資源錯置等不可逆損害，且該等損害常具有制度性與群體性，單靠個案救濟或市場競爭不足以矯正。

二、黑箱性與可解釋性不足：侵蝕醫師專業判斷與患者知情同意之制度基礎

深度學習等模型往往具有黑箱性（black-box），即其輸出結果難以由人類理解或追溯其推論路徑。此一現象在醫療領域的問題，不止於透明度，而是直接衝擊兩個醫療法核心制度：(1)醫師注意義務與最終判斷責任，(2)患者告知同意。就告知同意而言，若醫師在診斷、治療建議或資源配置上實質依賴AI，但患者未被告知AI的參與與其限制，是否構成告知義務違反，學界已有系統性討論。有學者以美國法告知同意doctrine為素材，指出現行法下多數情境未必會因「未告知AI參與」即成立責任，但當AI更不透明、在決策中扮演「過度關鍵角色」、或被用於成本控制等情形，告知義務的射程可能擴張，至少在規範論上有其正當性基礎。

就監理策略而言，亦有文獻警示：將「透明度／可解釋性」設定為主要監理焦點，可能是「錯置的管制重心」。該文以適應型SaMD為例，指出FDA偏重事前透明度要求，但真正關鍵或更有效的風險控制可能在於建構高品質資料輸入與共享基礎設施，以及強化事後監測與資料回饋機制，以確保模型在真實世界的穩健性與安全性[8]。

三、責任歸屬結構的系統性模糊：醫療過失與產品責任的中介地帶

醫療AI傷害事件的責任判斷，常同時牽涉：開發者（資料、模型、更新設計）、製造商（硬體與整合）、醫療機構（採購、部署、治理）、醫師（臨床使用與依賴程度）、維護者（更新與資安管理）等多重主體。當傷害發生時，往往難以辨識究竟是醫師使用不當、或「模型／資料瑕疵」、或更新後漂移所致，導致傳統以單一行為人為中心的過失歸責模型遭遇結構性困難。就責任架構選擇而言，有文獻主張應建立使「醫療過失與產品責任並存」的制度設計，並將AI自主性程度作為責任分配的重要變項，以反映人為失誤與機器失靈的不同成因；同時強調應避免完全交由事後訴訟反應式地形塑責任邊界，而應透過立法與監理先行建立可預期的責任配置[9]。

另有研究則從現行侵權法框架出發，主張

註7：Shabani, M., & Rahimzadeh, V. (2025). Introducing a fairness checkpoint for data quality and evidence during regulatory review of AI/ML-enabled medical devices. Houston Journal of Health Law & Policy, 24(1), 191-218.

註8：Hamilton, E. (2022). Transparency is a misplaced regulatory focus for holding adaptive software as medical devices (SaMDs) liable for defective design. Washburn Law Journal, 61, 571-601.

註9：Grossbard, E. S. (2025). The AI-robotic prescription: Legal liability when an autonomous AI robot is your medical provider. University of Miami Business Law Review, 33, 273-316.

醫療過失與產品責任（設計缺陷、製造缺陷、警示義務）原則上仍可調適AI情境，並提醒「標準照護」本質具有演進性，未來甚至可能出現「不使用AI反而構成過失」的爭議[10]。

四、持續學習與動態變異：傳統「一次性上市審查」的失靈與事後監測必要

與鎖定式（locked）模型不同，持續學習模型可能在上市後持續自我調整，導致同一輸入在不同時間點產生不同輸出。文獻指出，這種特性使得以往偏重事前性能證明的審查邏輯，對於適應型SaMD形成明顯不適配，因為每一次模型變動都可能改變風險輪廓，而頻繁重新審查又將抑制技術真正價值。亦有研究直指，連續學習演算法若在上市後根本性改變設備性質，可能造成重大公共衛生風險，故監理應在允許適度更新的同時，對「變更界限」與「上市後監測」投入更高密度之制度設計，並補強個資保護、資安與安全有效性監督等面向[11]。

五、醫療場域高度信任結構下之規範必要性

醫療行為本質上建立於高度制度性信任之上。患者將生命與身體交付予醫療體系，正是基於國家透過法律制度確保醫療安全、專業能力與責任可追究性。醫療AI若未置於適當規範架構之下，即可能削弱此一信任基礎，進而影響整體公共衛生體系之正當性。

正因如此，多數比較法文獻皆指出，醫療AI不宜僅被視為一般科技創新之延伸，而應被定位為高風險、強公共性之特殊應用領域，需透過加重監理密度、明確責任配置與制度化風險治理加以因應。

參、各國AI法制之一般治理原則與醫療領域規範

目前世界各國有將人工智慧以專法規範除我國外，計有歐盟、義大利、日本、韓國、越南（通過及生效時間如附表1），以下分二層次解構分述

一、人工智慧之一般治理原則與監理架構之比較

分析各國AI治理在司法領域的母法基礎，從比較法觀察，監理架構已呈現出從自律指引轉向實體規制的趨勢，並形成五種不同的治理模型。

（一）歐盟──風險分級為核心之強制規定模型

歐盟人工智慧法（EU AI Act[12]，以下簡稱AIA）的核心結構在於以風險分級（Risk-based approach）作為規範強度配置之依據。該法將AI系統依其對基本權利、民主秩序與公共利益之潛在影響，區分為不可接受風險

註10：Tonti, N. (2024). Who to blame? AI or your physician. Quinnipiac Health Law Journal, 27, 219-256.

註11：Gottfried, C., & Rosenstein, S. S. (2025). Who is practicing medicine? Man or machine: Telehealth AI and the corporate practice of medicine. Nova Law Review, 49, 313-340.

註12：EU AI Act https://artificialintelligenceact.eu/ai-act-explorer/

（Unacceptable Risk[13]）、高風險[14]、有限風險（透明度風險）[15]與最低風險四類。在此架構下，AIA透過附件（Annex Ⅲ）列舉高風險情境，使司法、醫療等領域在同一法制邏輯下被一致性處理。

1.不可接受風險

根據AIA第五條所謂不可接受風險意指對人類安全、生計或基本權利構成明顯威脅的人工智慧系統。例如，利用潛意識技術操縱人類行為、針對弱勢群體的剝削、政府進行的社會評分系統（Social Scoring）、在公共場所進行即時遠程生物特徵識別[16]（real-time remote biometric identification, RBI，執法用途除外，且需嚴格條件[17]）。採取完全禁止的規範手段。

2.高風險

⑴定義

根據AIA第6條第3項反面解釋，高風險人工智慧系統被定義為對自然人之健康、安全或基本權利構成重大危害者，透過雙重路徑認定高風險：

①產品安全法路徑（第6條第1項）：若AI系統本身即為歐盟產品安全法規（如玩具、機器、電梯）所涵蓋之產品或其安全組件，且須經第三方合規評估者。

②特定領域路徑（第6條第2項及附件III）：列舉了八大類對基本權利有重大衝擊的領域，包括：

Ⅰ.生物識別、教育培訓、就業人力管理。

Ⅱ.基本公共服務（如社會福利、金融信用評分）。

Ⅲ.執法、移民與邊境管制。

Ⅳ.司法行政與民主程序（附件Ⅲ第8點）。

⑵規範模式

高風險人工智慧系統受到的規範最為嚴格。其監督採取「合規性先行」與「持續性監督」並重的模式，貫穿了從設計開發（Pre-market）、合規評估到部署使用及上市後監測的全生命週期。以下分述之：

①設計與開發階段（提供者《Provider》[18]之義務）

在系統上市前，必須內建下列

註13：AIA Article 5.

註14：AIA Article 6.

註15：AIA Article 50.

註16：AIA Article 5(1).

註17：AIA Article 5(1)(d)、5(2)（一）搜尋特定受害者：包括但不限於：失蹤兒童、人口販運、性剝削被害人、綁架案件被害人（二）預防迫切且重大威脅（三）偵辦特定重大犯罪嫌疑人：附件II、III所列的重大犯罪類型。例如：恐怖主義、人口販運、嚴重暴力犯罪，且須為已發生或高度具體化之犯罪、鎖定特定嫌疑人。

註18：AIA Article 3(3)從文義及體系解釋，所謂提供者指任何自然人、法人或公私機關、機構或其他組織體，不論其是否親自或透過他人開發人工智慧系統或通用型人工智慧模型，只要其以自身名義或商標，將該系統或模型首次置於市場中，或實際投入具體使用情境，即使係無償提供，亦屬之。

技術標準，以確保設計即合規（compliant by design）：

Ⅰ.風險管理系統（第9條）：必須建立、實施及紀錄貫穿整個生命週期的持續、迭代風險管理流程。

Ⅱ.資料治理（第10條）：訓練、驗證與測試數據集應具備代表性、無偏誤，且符合隱私標準。

Ⅲ.技術文件（第11條）：必須編製詳細文件，說明系統預期目的、運作邏輯及合規證明。

Ⅳ.自動記錄（第12條）：系統必須能自動生成日誌（Logs），以便事後可追溯。

Ⅴ.透明度與資訊提供（第13條）：設計上須讓部署者能解讀系統輸出，並附有詳細的使用說明書。

Ⅵ.人類監督（第14條）：設計時應確保人類能有效進行監督，包括能夠理解系統限制、發現異常，並擁有緊急停止按鈕（stop button）或類似中斷機制。

Ⅶ.精確性、穩健性與資安（第15條）

確保人工智慧系統具備抵抗攻擊與減少幻覺的技術能力。

②上市前合規評估（Conformity Assessment）

Ⅰ.合規評估（第43條）及標誌（第48條）：高風險AI必須通過合規性評估程序。多數情況下由提供者進行「內部控制評估」，但特定類型（如生物識別）則須經第三方「指定機構（Notified Bodies）」審查。通過評估後，需貼上CE合規標誌。

Ⅱ.註冊義務（第49條）：提供者必須在歐盟資料庫中註冊該高風險AI系統，確保公眾監督。

③部署與使用階段（部署者《deployer》[19]之義務）

當AI系統進入實務運用（如法院及醫院），部署者需承擔監控義務：

Ⅰ.依指令使用及監控與記錄保持（第26條）：部署者必須確保系統在預期條件按照使用說明操作、確保輸入數據的相關性下，指派具備能力的人員進行監督，當發現風險或嚴重事故時，必須立即停止使用並通知提供者或主管機關。

Ⅱ.基本權利影響評估（Fundamental Rights Impact Assessment, 第27

註19：AIA Article 3(4) 搭配序言13所謂部署者指任何自然人、法人，或公私機關、機構或其他組織體，在其職務、組織或管理權限支配下，使用人工智慧系統者；但自然人僅於其私人生活中，基於非專業、非職業、非商業目的使用人工智慧系統者，不在此限。

條）：在部署高風險系統之前，除擬用於附件Ⅲ第2點所列區域的高風險AI系統[20]外，依公法管轄的機關或提供公共服務的私人實體部署者，以及附件Ⅲ第5(b)及(c)點所指高風險AI系統部署者[21]，應評估使用該系統可能對基本權利產生的影響。為此，部署人員應進行包括：(i)部署者將如何依其預期目的使用高風險AI系統的流程描述；(ii)描述每個高風險AI系統預計在何時及頻率內使用；(iii)在特定情境下，可能受到其使用影響的自然人及群體類別；(iv)根據本款第3點所識別的自然人或群體類別，考慮提供者依第13條所提供資訊，可能影響自然人或群體類別的具體損害風險；(Ⅴ)依照使用說明，說明實施人為監督措施的過程；(Ⅵ)在風險發生時應採取的措施，包括內部治理及申訴機制的安排。在系統首次運作時，也要踐行上開義務[22]。

④上市後監督與問責（後端監督）

Ⅰ.上市後監測（第72條）：提供者必須建立系統性收集、記錄和分析系統性能數據的監測系統。

Ⅱ.嚴重事件通報（第73條）：若發生死亡、嚴重健康損害或侵犯基本權利之事故，必須向監管機關報告。

Ⅲ.市場監督（第75-84條）：各成員國的市場監督機關有權進入機房檢查、要求移除不合規系統，或處以巨額罰款。

3.非高風險之AI管制

AIA法條本身並未明文定義有限風險或最低風險，這些用語主要來自歐盟執委會最初的立法提案說明文件[23]，正式通過的AIA條文中，立法者實際採取的是二分法：高風險AIvs.「其他非高風險AI，嚴格規範僅限前者，後者不設強制性實體義務，僅透過第50條透明資訊原則，使用者應知其正在與AI系統互動，或內容為AI生成。另依AIA第95條

註20：即人工智慧系統旨在作為關鍵數位基礎設施、道路交通管理與運作的安全元件，或供水、天然氣、供暖或電力供應。

註21：5(b)用於評估自然人的信用狀況或建立其信用分數的人工智慧系統，但用於偵測金融詐欺的人工智慧系統除外；5(c)用於自然人風險評估與定價的人工智慧系統，適用於壽險與健康保險。

註22：AIA Article 27(2).

註23：European Commission. (2021). Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021) 206 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 (last visited on 2025/12/27).

（配合序言165）對於非高風險AI業者鼓勵自願採納部分高風險AI的要求；此外，非高風險AI並非處於法律真空，仍可能同時適用GDPR（個資、自動化決策）、消費者保護法、不公平商業行為法及數位服務法（DSA）等。

（二）韓國：基本法形式下之高影響AI治理模型

1.立法簡介

韓國於2025年1月21日頒布人工智慧發展與建立信任基礎基本法（下稱AI基本法），準備於2026年1月22日生效，係繼歐盟第二個以專法形式對於人工智慧加以規範，並且創下法律尚未生效上路前，即於2025年12月30日再加以修正[24]。韓國AI基本法不直接以刑罰或強制行政處分為核心，前半部高度著墨於AI產業、研發、人才、資料、公共部門導入，後半部才逐步引入倫理、透明、安全、影響評估，管制採取先發展、後規範與產業振興的基調。重點在於支援與推廣，此從法律名稱以及此次修正重點將國家人工智慧策略委員會的重組法典化[25]、設立並運營人工智慧研究機構[26]、創造公共部門人工智慧需求[27]、支持人工智慧新創啟動、提供人工智慧專業人才支援、建立提供公共資料作為訓練資料的基礎[28]、支持人工智慧技術教育，以及確保弱勢群體的人工智慧可及性[29]。韓國的管制重點並非全面的限制，而是建立國家級的AI發展支援體系與基礎建設，並僅針對特定高影響（High-Impact）領域進行管理。

2.管制特色：區分高影響人工智慧

所謂高影響人工智慧，依AI基本法第2條第4項規定，係指對人的生命、身體安全或基本權利，具有重大影響或可能造成風險之人工智慧系統，並且係運用於下列特定領域之一者：

⑴能源供應

⑵飲用水生產流程

⑶醫療體系之建構與運作

註24：인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 일부개정법률안（대안）韓國國會議案編號2215126，https://opinion.lawmaking.go.kr/gcom/nsmLmSts/out/2215126/detailRP?utm_source=chatgpt.com,（last visited on 2025/12/27）。

註25：AI基本法第7條第1項「國家人工智慧委員會」變更為「國家人工智慧戰略委員會」。

註26：AI基本法第22條之2人工智慧研究所建立設立與營運AI研究所的法律根據，政府可資助經費。

註27：AI基本法第16條第1項國家及地方政府應制定促進企業與公眾導入AI的政策、第3項國家機關採購時應優先考慮由總統令所定之人工智慧產品或人工智慧服務、第4項對於優先採購造成的損失，除非有故意或重大過失，否則免除相關人員責任。

註28：AI基本法第6條第2項第4款之2人工智慧基本計畫之制定應包含關於依《公有資料提供及利用活性化法》之公有資料，利用其生成學習用資料（訓練資料）、公有資料提供等之範圍與標準，以及其活性化（促進）之事項。

註29：AI基本法第6條第2項第7款人工智慧基本計劃應包括關於保障人工智慧產品或人工智慧服務之「人工智慧脆弱階層」可接近與使用之事項。

(4)醫療器材與數位醫療器材
(5)核能設施與核物質安全
(6)犯罪偵查或逮捕用之生物辨識資料分析（臉部、指紋、虹膜等）
(7)招募、貸款審查等，對個人權利義務有重大影響之判斷或評價
(8)交通工具、設施、系統之核心運作
(9)公共服務中涉及資格認定、決策或費用徵收之國家或公共機關決策
(10)初等與中等教育中之學生評量
(11)其他經總統令指定，對生命、身體或基本權有重大影響之領域韓國針對高影響人工智慧採取管制措施計有：

①事前辨識與確認義務

依據韓國AI基本法第33條規定，AI事業者有義務自行事前判斷其AI是否屬高影響AI。若有疑義，得向科學技術情報通信部長請求確認。主管機關得設專家委員會，並制定「判斷基準與案例指引」。

②安全性與信賴性之強化義務

依據韓國AI基本法第34條對於提供高影響AI的事業者，法律明文要求必須採取特定措施，包括：風險管理制度、可解釋性（對結果、主要判斷基準、訓練資料概要之說明）、使用者保護機制、人類之管理與監督、文件化與保存義務、這是具體化的行為義務群，而非單純倫理宣示。

③影響評估制度

依據韓國AI基本法第35條對於事業者提供高影響AI產品或服務前，應努力進行對基本權影響之評估。國家機關使用AI時，應優先選擇已完成影響評估之產品或服務。

④透明性義務（揭露與標示）

依據韓國AI基本法第31條凡屬高影響AI或生成式AI，即須：事前告知使用者此產品或服務係基於AI運作；若為生成內容，須標示其為AI生成；對於高度逼真的虛擬聲音、影像，須以可清楚辨識方式揭露。

⑤安全性檢驗與公共部門使用之偏好規則

依據韓國AI基本法第30條第3、4項規定，AI事業者提供高影響AI時，應努力事前接受安全性、信賴性之檢驗或認證。國家機關使用高影響AI時，應優先考量已完成檢驗認證之產品或服務。

3.非高影響之人工智慧之管制

(1)說明義務

依韓國AI基本法第3條第2項規定受影響之人，對於人工智慧產生最終結果時所採用之主要判斷基準及其原理等事項，於技術上及理性上可行之範圍內，應得以獲提供明確且具實質意義之說明。不論是否屬於高影響人工智慧，影響人得在技術上合理可行範圍內，獲得有意義的說明，法條雖並未直接指定說明

義務之承擔主體，而是以影響人得獲說明之權利型條款作為基本原則；其具體義務歸屬，應依AI系統之實際運用關係，由對外作成決定之部署者負第一線說明責任，並由掌握模型結構與推論邏輯之提供者負協力說明義務。

(2)透明性確保義務

依韓國AI基本法第31條第2項規定人工智慧事業者於提供生成式人工智慧，或提供利用生成式人工智慧之產品或服務時，應標示該結果物係由生成式人工智慧所生成之事實。同條第3項針對利用人工智慧系統提供難以與實際（真實）情形加以區分之虛擬聲音、影像或影片等結果物時，要求人工智慧事業者應以使利用者得以明確識別之方式，告知或標示該結果物係由人工智慧系統所生成。但該結果物屬於藝術性或創作性表現，或構成其一部分者，得以不妨礙展示、欣賞或享用等方式之方式為告知或標示。

(3)民間自律機制

依韓國AI基本法第28條第1項規定人工智慧事業主得自設民間自律人工智慧倫理委員會，而執行同條第2項所示確認人工智慧技術於研究、開發及運用過程中，是否遵循倫理原則及運用之安全性、結果及人權侵害等事項，進行調查與研究、監督。

（三）日本：政策推進導向與軟法整合模型

1.立法簡介

日本AI專法的立法起點可以追溯到2023年5月在日本廣島舉行的G7峰會。作為當年的G7主席國，日本首相岸田文雄發起了「廣島AI進程」（Hiroshima AI Process），旨在針對生成式AI的快速發展建立國際性的治理框架，該年年底促成《開發先進AI系統組織國際指導原則（Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System）》及《開發先進AI系統國際行為準則（Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI System）》上路[30]。於此同時日本政府成立了由學界、業界及法律專家組成的AI戰略會議。該會議成為日本AI政策的最高智囊團，負責審議從技術研發到法律監管的所有重大議題[31]。翌年2月14日由經濟產業省主導設於情報處理推進機構（Information-

註30：劉汶渝，開發先進AI系統國際指導原則暨行為準則之新進程—2024年OECD巴黎部長理事會議，臺灣人工智慧行動網，2024年5月5日，https://ai.iias.sinica.edu.tw/hiroshima-ai-process-comprehensive-policy-framework/（last visited on 2025/12/29）。

註31：内閣府，AI に関する暫定的な論点整理，2023年5月26日AI戦略会議，https://www8.cao.go.jp/cstp/ai/ai_senryaku/2kai/ronten.pdf（last visited on 2025/12/29）。

technology Promotion Agency, IPA）下人工智慧安全研究所（AI Safety Institute, AISI）正式成立，設立目的是為了因應生成式AI的普及，針對AI安全性進行技術評估、制定標準以及加強與國際的合作目的，為立法提供了技術支撐。。

2024年間，日本政府內部針對是否應效仿歐盟對於AI進行嚴格立法進行了深入辯論[32]，最終日本採取的是政策推進導向與軟法整合模式，旨在成為全球最易於開發與應用AI的國家，制定一部基本法，既確立監管框架，又保留促進創新的靈活性，而於2025年3月內閣通過送請國會審議於同年5月28日通過並於6月4日公布，同年9月1日生效之人工智慧相關技術研究開發及活用推進法（人工知能関連技術の研究開発及び活用の推進に関する法律，簡稱AI推進法[33]）。日本法制未採取剛性的風險分級表，而是透過內閣設立人工智慧戰略本部（由首相領導），統籌制定「人工智慧基本計畫」。其監管是動態的，依據AI推進法第13條[34]由國家制定活用事業者應遵守之指引來落實妥適性，而非直接在法律中列舉禁止事項。

2.管制特色──軟法及促進發展

日本AI推進法第1條明確指出法律目的是「鑑於AI是經濟社會發展的基礎技術……謀求AI相關技術的研究開發及活用推進措施之綜合性與計畫性推進」[35]，搭配同法第11條[36]規定國家應採取必要措施以推進AI的基礎研究及成果實用化，顯示日本對AI態度係促進者的角色。而非以處罰或限制為首要目標。體現在第4條國家的職責與任務只有二個：國家有義務依循第三條所定之基本理念，綜合性且計畫性地擬定並實施人工智慧相關技術之研究開發及活用推進相關措施及國家為謀求行政事務之效率化及高度化，應積極推進國家行政

註32：内城喜貴，政府、AIのリスク対策で法規制を検討 国の戦略として利用促進と両立目指す，Science Portal, 2024.6.7,
https://reurl.cc/6bY44M（last visited on 2025/12/29）。

註33：令和七年法律第五十三号人工知能関連技術の研究開発及び活用の推進に関する法律，
https://laws.e-gov.go.jp/law/507AC0000000053（last visited on 2025/12/29）。

註34：AI推進法第13條國家為謀求人工智慧相關技術之研究開發及活用之適當實施，應採取擬定活用事業者應遵守之指引、促進開發人工智慧相關技術利用所伴隨風險之評估及管理手法等必要措施。

註35：AI推進法第1條：本法鑑於人工智慧相關技術為我國經濟社會發展之基礎技術，爰就人工智慧相關技術之研究開發及活用推進相關措施，訂定基本理念、人工智慧相關技術之研究開發及活用推進之基本計畫擬定等措施之基本事項，並設置人工智慧戰略本部，以期與《科學技術・創新基本法》（平成七年法律第一百三十號）及《數位社會形成基本法》（令和三年法律第三十五號）等相關法律之施策相輔相成，謀求人工智慧相關技術之研究開發及活用推進施策之綜合性與計畫性推進，進而貢獻於國民生活之提升及國民經濟之健全發展。

註36：AI推進法第11條：國家為推進人工智慧相關技術之研究開發，應採取推進人工智慧相關技術之基礎研究開發、促進研究開發成果之實用化等必要措施。

機關對人工智慧相關技術之活用。然而第三條所謂基本理念除第1項「委外」（應依循《科學技術・創新基本法，下稱科學技術基本法》第三條所定之科學技術・創新創出振興方針[37]及《數位社會形成基本法》第二章所定之基本理念[38]）補充外，第2、3、5項均係以提升人工智慧相關技術產業之國際競爭力、推進AI活及取得國際主導為宗旨，只有在第4項有論及「人工智慧相關技術之研究開發及活用，鑑於若以不正當目的或不適當方法為之，恐助長用於犯罪、個人資訊洩漏、著作權侵害等危害國民生活平穩及國民權益之情事，為謀求其適當實施，必須採取確保人工智慧相關技術研究開發及活用過程之透明性等必要措施」等語，有概括談到AI管制。

依據日本AI推進法第13條規定國家為確保AI活用之適當實施，應「擬定活用事業者應遵守之指引」以及「促進開發AI利用所伴隨風險之評估及管理方法」，日本政府透過法律授權制定指引，具體的監管細節將透過指引（軟法）來落實，並要求企業遵守指引進行風險評估，保留了企業自主管理的彈性，而非在法律中直接條列具體的禁止事項。而對AI提供之業者責任及義務，僅在第7條規定企業應「依循基本理念……致力於事業活動之效率化、高度化及新產業之創出」，並配合國家措施，屬於宣示性義務而非懲罰性義務。最後

註37：科學技術基本法第三條：鑑於科學、技術和創新創造是日本和人類社會未來發展的源泉，科學技術知識的累積是人類的智力財富，必須積極推進科學、技術和創新，同時尋求與人類生活、社會和自然的和諧，以充分發揮研究人員和其他利用研發成果創建新企業的人員的創造力。

1.在推動科學、技術和創新創造的過程中，必須考慮在各個領域培養均衡的研發能力，兼顧各領域的特點，促進跨學科或綜合性研發，協調基礎研究、應用研究和開發研究，均衡推進學術研究和非學術研究，並促進國家測試研究機構、研發公司、大學等機構、私營企業和其他相關方在國內和國際上的合作。此外，鑑於自然科學與人文學科的相互互動對科技進步與創新至關重要，必須重視二者的和諧發展。

2.推動科技發展必須牢記科技的多重意義，它不僅有助於創新，還有助於創造學術價值，並且必須確保研發過程的公平性。

3.促進創新必須與科技發展有機協調，使科技發展帶來的研發成果與創新成果最大程度地結合。

4.促進科技發展和創新必須以建構一個全體公民都能廣泛享受科技和創新成果的社會為目標。

5.在推動科學技術和創新創造的過程中，必須注意全面利用各領域的科學技術知識，並妥善應對以下及其他社會挑戰：

(1)日本面臨的問題，包括出生率下降和人口老化、人口減少、跨國社會經濟活動的發展。

(2)人類面臨的糧食問題、能源限制、全球暖化等問題。

(3)因科學技術應用所帶來的社會經濟結構變化而產生的就業及其他領域的新挑戰。

註38：數位社會形成基本法第二章基本理念第3條至12條，規範內容為實現全民分享資訊通信科技惠益的社會、促進經濟結構改革，增強國際產業競爭力、實現人民安居樂業、建構充滿活力的社區等、實現公民安全生活的社會、減少資訊獲取等方面的差距、國家和地方政府與私部門的角色分工、保護個人和企業等的權利和利益、因應資訊通信科技的進步及因應社會經濟結構變化帶來的新挑戰。

於附則第2條規定政府應考量國際動向及技術變化，檢討本法之施行狀況，並在必要時採取措施。這體現了日本法制設計上保留了隨技術演進而調整的空間，強調敏捷治理，法律應保持靈活性，並強調在發展中解決風險，而非預先設立過多障礙；而依第19條於內閣設置人工智慧戰略本部，並於第22條由總理大臣充任部長，負責第18條AI基本計畫之草擬，並由內閣會議決定，確保政策的長期性與綜合性。這種以舉國體制推動AI，這與歐盟側重於市場監管機構的角色不同。

2025年（令和7年）12月19日高市總理召開第三次人工知能戰略本部會議[39]，推出人工知能基本計劃草案即以「用可信賴的AI重振日本」為副標，該計畫旨在平衡「創新促進」與「風險應對」，並將核心措施分為四個面向，以實現AI創新與社會變革的良性循環。

(1)AI活用的加速的推進

該計畫強調AI的社會應用，特別是從政府部門開始帶頭使用。推動「政府AI」（Government AI），要求中央省廳職員普遍使用生成式AI，以提高業務效率與品質。並用以解決社會課題，積極支援醫療、防災、農林水產業、製造業等領域的AI導入，特別是針對日本面臨的人口減少與人手不足問題，並且進行制度改革：以AI活用為前提，檢視並修改現有法規與制度。

(2)AI開發力的戰略與強化

目標是利用日本的優勢數據，強化自主開發能力，實現「反轉攻勢」。重點在基礎設施強化，加速整備AI基礎設施，包括確保足夠的計算資源、數據中心、以及高性能AI半導體的研發與供應。

利用日本在產業、醫療、研究等領域積累的高品質數據，開發具備日本文化與習慣的「可信賴AI」基礎模型以及積極引進國內外頂尖AI人才，並透過產學官合作強化研究開發。

(3)AI治理的主導

日本將延續「廣島AI進程」的國際成果，主導AI治理框架的建立。強化「AI安全研究所」（AISI）的功能，將人員規模擴充至現行的兩倍，以提升對AI模型安全性、公平性及風險的技術評估能力繼續主導「廣島AI進程」等國際框架，確保AI治理的國際互操作性。在風險應對針對深度偽造（Deepfake）、網路攻擊等AI惡用風險，制定指引並支援AI生成內容的判別技術（如電子浮水印）開發。

(4)向人工智慧社會持續轉型

著重於社會結構的持續變革，以

註39：人工知能戰略本部，總理的一日，2025.12.19，
https://www.kantei.go.jp/jp/104/actions/202512/19jinkoutchinou.html（last visited on 2025/12/29）。

適應AI時代的來臨。從人才育成必須初等教育階段開始提升全民AI素養。推動針對各行各業的AI轉職訓練，並強調在AI時代，人應提升創造力、思考力、判斷力，以實現人與AI的最佳協作。最後持續探討AI相關的民事責任歸屬、智慧財產權保護與利用、以及AI對就業市場的影響等議題[40]。

（四）越南：國家管理與行政管制導向之模型

越南國會於2025年12月10日通過《人工智慧法》（Luật Trí tuệ nhân tạo），法律編號為134/2025/QH15[41]，並定於2026年3月1日生效。這項立法使越南成為全球第5個擁有全面AI法律框架的國家。越南與歐盟AI管制的共同核心皆採取「基於風險的方法」（Risk-based approach），但在風險分級架構、管制細節與特定議題的處理上存在顯著差異。

1.風險分級採三層級

越南人工智慧法第9條第1項規定AI系統分為三類：(1)高風險：是指可能對組織、個人的生命、健康、權益及合法利益，國家利益、公共利益、國家安全造成重大損害的系統。(2)中風險：是指因使用者無法識別互動對象是人工智慧系統或內容由系統生成，而可能導致混淆、影響或操縱使用者的系統。(3)低風險：不屬於前兩類的系統。同條第2項規定人工智慧系統的風險分類基於以下標準確定：對人權、安全、安防的影響程度；系統的使用領域，特別是必要領域或直接涉及公共利益的領域；使用者範圍及系統影響規模。

關於分類要求供應商（Provider）投入使用前自行為之，若尚未確定風險等級，供應商可依據技術檔案請求科學技術部指導分類，部署方（Deployer）可繼承供應商的分類結果，並有責任確保系統在使用過程中的安全和完整；若修改、整合或變更功能導致產生新風險或更高風險，應配合供應商重新進行分類[42]。檢查、監督依據系統風險等級進行：(1)高風險人工智慧系統進行定期檢查或在有違規跡象時檢查；(2)中風險人工智慧系統透過報告、抽樣檢查或獨立組織評估進行監督；(3)低風險人工智慧系統在發生事故、反映或需要保障安全時進行追蹤、檢查，不給組織、個人產生不必要的義務對高風險AI系統管制要求與歐盟相近：

(1)全生命週期管理：針對高風險系統，兩者均要求從設計、訓練到部署後的持續監控。越南人工智慧法第14條明確要求建立風險管理制

註40：人工知能基本計画（案），https://www8.cao.go.jp/cstp/ai/ai_plan/aiplan2025_draft5.pdf （last visited on 2025/12/29）。

註41：Luật Trí tuệ nhân tạo https://vanban.chinhphu.vn/?pageid=27160&docid=216334&classid=1&typegroupid=3 last (visited on 2025/12/29).

註42：越南人工智慧法第10條第1、2、4項。

度、數據治理、技術檔案保存、人類監督與事故通報，這與歐盟法案第9-15條的要求幾乎一致。

(2)事前合規評估：越南人工智慧法第13條規定高風險系統在「投入使用前」必須進行合規評估（Conformity Assessment）。

(3)透明度與標示：越南人工智慧法第11條規要求供應商與部署方使用者必須知道其正在與AI互動，並對生成的內容（如Deepfake）進行標記

對高風險AI系統管制要求越南與歐盟相異主要在高風險定義範圍，越南針對國家利益、公共利益及國家安全造成重大損害，具體清單授權行政權頒布[43]；在國家政策與原則中，明確提及AI發展必須遵守道德規範與越南文化價值，這顯示其立法具有強烈的主權與在地化色彩[44]；關於特定高風險系統必須經過第三方依法登記或承認的組織進行合規認證[45]，與歐盟多數高風險系統由廠商自評，僅生物辨識等少數需第三方審查有所不同；AI受控試驗機制（沙盒）機制，可在特定條件下豁免部分法律責任，用以測試高風險或新興AI解決方案[46]。

此外，越南風險分類外，亦有絕對禁止行為在第7條規定：

①利用、侵占人工智慧系統以實施違法行為，侵害組織、個人的權益及合法利益。

②開發、提供、部署或使用人工智慧系統於以下目的：a）實施法律規定的禁止行為；b）使用偽造要素或模擬真實人物、事件，有目的地、系統性地欺騙或操縱人類的認知、行為，對人類的權益及合法利益造成嚴重損害；c）利用弱勢群體（包括兒童、老年人、身心障礙者、少數民族或民事行為能力喪失、受限者、認知及行為控制困難者）的弱點，對其自身或他人造成損害；d）製作或傳播可能對國家安全、社會秩序及安全造成嚴重危害的偽造內容。

③違反關於數據、個人數據保護、智慧財產權及網路安全的法律規定，收集、處理或使用數據來開發、訓練、測試或運作人工智慧系統。

④阻礙、無效化或歪曲人類對人工智慧系統依本法規定進行的監督、干預和控制機制。

⑤隱瞞必須公開、透明或解釋的資訊；塗改、歪曲人工智慧活動中強制性的資訊、標籤、警示。

⑥利用人工智慧系統的研究、測

註43：越南人工智慧法第9條第3項、第13條第4項。
註44：越南人工智慧法第4條。
註45：越南人工智慧法第13條第2項。
註46：越南人工智慧法第21條。

試、評估或檢定活動實施違法行為。

2.越南對「中度風險」AI的特殊管制

越南法對於中度風險（Rủi ro trung bình）的管制比一般僅要求透明度的模式更為嚴格，增加了行政通報的義務。

(1)強制通報（第10條第3款）：

供應商在投入使用前，必須透過「人工智慧單一窗口入口網站」向科學技術部通報分類結果。這點比歐盟更嚴格（歐盟主要針對高風險註冊）。

(2)透明度與標籤義務（第11條）：

必須讓使用者識別正在與人工智慧系統互動；供應商應確保對生成的音訊、圖像、影片必須有機器可讀的標記；部署方有責任確保為模擬、仿真真人外貌、聲音或再現真實事件而由人工智慧系統生成或編輯的音訊、圖像、影片必須貼上易於識別的標籤，以與真實內容區分

(3)解釋責任（第15條第1款）：

①當國家機關在稽查、檢查中要求，或有風險、事故跡象時，供應商有責任就使用目的、功能描述層面的運作原理、主要輸入數據及風險管理、系統安全措施進行解釋；解釋不要求揭露原始碼、詳細演算法、參數集或商業機密、技術機密。

②當國家主管機關在稽查、檢查或處理事故過程中要求時，部署方有責任就運作、風險控制、事故處理及保護組織、個人權益進行解釋。

③使用者有責任遵守關於人工智慧系統通報、標籤的規定。

(4)監督機制（第10條第5項b項）

透過報告、抽樣檢查或獨立組織評估進行監督。

（五）義大利──歐盟AI專法之落地實踐規範

義大利於2025年9月23日通過人工智慧法（第132號法律[47]）並於2026年1月1日生效上路，成為歐盟國家中第一個另成立專法建構國家層級的AI管理架構。該法指定義大利數位局（AgID）與國家網路安全局（ACN）為國家人工智慧主管機關[48]。AgID主要負責促進AI創新發展，並管理合格評定機構的通知與評估程序；ACN則負責AI系統的監管執法（包括檢查和制裁），以及在網路安全領域推動AI的應用。兩機關共同建立並運作AI實驗空間（沙盒）以支援監管創新。此外，ACN被指定為市場監管機構及歐盟機構單一聯絡點，AgID則擔任通知機關，確保義大利與歐盟在AI監管上的銜接。為了整合政策，

註47：Disposizioni e deleghe al Governo in materia di intelligenza artificiale, https://www.appaltiecontratti.it/disposizioni-e-deleghe-al-governo-in-materia-di-intelligenza-artificiale-2/ (last visited on 2025/12/31).

註48：第20條第1項AgID：負責促進AI的創新與發展，以及驗證機構的通知、評估與監測程序。ACN：負責AI系統的監管（包括檢查與制裁），以及涉及網路安全方面的AI推廣與發展。

義大利法並在總理府下設立國家AI戰略與協調委員會（由總理或其授權人主持）[49]，協調公私部門在數位創新與AI領域的活動。

1.風險分級管制

義大利AI專法基本上承襲歐盟的風險導向監管思路。法案開宗明義要求對AI可能帶來的經濟社會風險及對基本權利的影響進行監管[50]，並強調在AI全生命週期中落實網路安全和相稱的安全控制措施[51]，以確保系統穩健可靠。然而，義大利法並未另立一套本國風險分類體系，而是直接引用歐盟AI法規的定義[52]：例如「人工智慧系統」和「人工智慧模型」均依照歐盟2024/1689號條例第3條的定義。義大利明文規定，本法對通用目的AI系統與模型（即廣泛用途的基礎模型等）不產生超出歐盟AI法規的新增義務[53]。因此，實質上義大利將高風險系統的要求、風險等級劃分直接與EU法對接。例如，一旦EU AI Act生效，義大利將透過授權立法把高風險AI的評估認證、監管執法權等納入國內法。

2.管制特色──特定領域落地規範

義大利專法側重AI在特定領域規範，在第二章第7-15條分別在醫療衛生與身心障礙領域、醫療領域AI實現的科學研究與實驗、關於電子健康檔案衛生領域監測系統及數位衛生治理（前四者詳後述醫療領域說明）外，有下列具體使用情境規定：

(1)個人資料處理

義大利AI專法第9條對於AI訓練所使用所涉個資問題，直接規定衛生部長將在120天內頒布法令，規範利用AI和機器學習進行研究實驗的個人資料（包括特殊類別數據）處理，採用最簡化模式，包括建立研究用的特殊實驗空間（沙盒）

(2)勞動領域

義大利AI專法第11條對於AI使用於勞動關係情境下可能產生問題規定如下：

①AI應用於改善工作條件、保護勞工身心完整性、提升績效品質與生產力。

②勞動領域的AI使用必須安全、可靠、透明，不得違反人類尊嚴或侵犯隱私。雇主必須告知勞工AI的使用。

③AI在勞動關係管理中必須保證遵守勞工不可侵犯的權利，不得有性別、年齡、種族、宗教等歧視。

(3)專業性職業領域

義大利AI專法第13條對於AI於專業性職業領域情境規定如下：

①專業性職業領域中使用AI僅限於

註49：第19條第4項。
註50：第1條第3項。
註51：第3條第6項。
註52：第2條。
註53：第3條第5項。

作為專業活動的工具和輔助，且A不得取代成為實質上提供專業判斷的主體。

②為確保信任關係，專業人士必須以清晰簡單的語言告知客戶其使用的AI系統相關資訊。

(4)公共行政

義大利AI專法第14條對於AI於公部門領域使用情境規定如下：

①公共行政部門使用AI旨在提高效率、減少程序時間、提升服務質量，並確保運作的可知性與可追溯性。

②AI僅作為行政處分活動的輔助工具，人類仍是決策與程序的唯一負責人。

③採取措施確保AI的負責任使用及使用者的培訓

(5)司法活動中使用

義大利AI專法第15條對於AI於司法領域使用情境規定如下：

①在司法活動中，法律解釋、事實與證據評估及採取措施的決定權始終保留給法官。

②司法部規範AI系統在司法服務組織、工作簡化及附屬行政活動中的使用。

③在（UE）2024/1689條例完全實施前，普通司法機關的AI實驗與使用需經司法部授權。

④司法部長在法官培訓計畫中推動關於AI的教學活動。

3.增定使用AI犯罪或違規處罰

針對不當使用AI處罰部分，義大利AI專法第26條（刑法及其他刑事規定之修正）規定如下：

(1)刑法之修正如下

①於第61條第11款之9[54]後，新增下列一款：

第11款之10：以人工智慧系統實施犯罪行為者，於該等系統因其性質或使用方式，構成詭詐手段，或其使用已妨礙公共或私人防禦，或加重犯罪結果者。

②於第294條末尾，新增下列一項：

如詐欺行為係以人工智慧系統為手段實施者，處二年以上六年以下有期徒刑。

③於第612條之3後，增訂第612條之4如下：

第612條之4（以人工智慧系統生成或變造內容之非法散布罪）

凡未經本人同意，轉讓、刊登或以其他方式散布經由人工智慧系統偽造或變造，且足以使人誤認其真實性之影像、影片或聲音，而致他人受不法損害者，處一年以上五年以下有期徒刑。

本罪須告訴乃論。

但如該行為與其他應依職權追訴之犯罪相牽連，或係對於因年齡或身心障礙而無行為能力之人

註54：相當於我國刑法第57條。

所為，或係因其職務行使而針對公權力機關或公務人員所犯者，則不以告訴為限，逕依職權追訴。

(2)民法第2637條[55]末尾，新增下列規定

如該行為係以人工智慧系統實施者，處二年以上七年以下有期徒刑。

(3)1941年4月22日第633號《著作權法》第171條第1項[56]，於a-bis款後，增列如下）

違反第70-ter條及第70-quater條之規定，自網際網路或資料庫中重製或擷取文字或資料，不論是否透過人工智慧系統為之。

(4)1998年2月24日第58號立法命令所定《金融中介統一法》第185條第1項[57]末尾，新增下列規定：

如該行為係以人工智慧系統實施者，處二年以上七年以下有期徒刑，並科歐元二萬五千元以上六百萬元以下罰金。

4.限制未成年人使用AI

此外，義大利AI專法第4條第4項14歲以下未成年人使用AI技術及相關資料處理需經行使親權者同意。14歲以上未滿18歲的未成年人，若相關資訊易於理解，可自行表達同意。

（六）小結

各國AI治理正呈現從「自律指引」轉向「實體規制」的趨勢，但立法模式各有側重：

1.**歐盟**：確立「風險分級」之強制規定模型，將AI區分為不可接受、高風險等四類，針對高風險AI實施從設計到上市後監測的全生命週期嚴格合規監管。

2.**韓國**：採「高影響AI」治理模型，基調為「先發展、後規範」與產業振興。僅針對能源、醫療等特定領域之「高影響AI」課予事前確認、安全強化及透明性義務，避免全面限制。

3.**日本**：採「軟法整合」與「政策推進」導向，旨在成為最易開發AI之國家。透過《AI推進法》授權制定指引（軟法）落實監管，保留企業自主彈性，不直接於法律列舉禁止事項。

4.**越南**：採「國家管理」與「行政管制」導向，除高風險外，對「中風險」AI亦課予嚴格的行政通報義務，並明文列舉絕對禁止行為，展現強烈主權與在地化色彩。

5.**義大利**：作為歐盟法之落地實踐，除直接引用歐盟風險定義外，特重醫療、司法及勞動等特定領域之規範，並修法增訂利用AI犯罪之加重刑責。

這顯示全球雖有風險管控共識，但在管制強度（硬法vs.軟法）與監理重心（安全vs.發展）上仍有顯著差異。

註55：本條為資本市場犯罪（市場操縱）。

註56：未經授權的利用行為。

註57：利用未公開之重大資訊，從事金融商品交易或洩漏該資訊的行為。

二、各國在醫療領域AI運用之管制

（一）緒論：醫療人工智慧監管的全球轉向

人工智慧技術在醫療保健系統中的整合，代表了現代醫學史上最深刻的變革之一。從基因序列分析、影像診斷輔助到手術機器人的自動化操作，AI承諾為人類帶來前所未有的診療效率與精準度。然而，這場技術躍進同時也引發了一系列複雜的法律、倫理與安全挑戰，迫使各國政府必須在「促進創新」與「保障病患安全」之間尋求微妙的平衡。2024年至2025年間標誌著全球AI立法監管的關鍵轉折點，主要司法管轄區紛紛從軟性的指導原則（Soft Law）轉向具有法律約束力的成文法典（Hard Law）。

以下針對歐盟、義大利、日本、韓國與越南這五個具有代表性的司法管轄區，剖析各國法規文本，探討其在醫療領域的具體管制規範、倫理要求及實施機制。

（二）歐洲聯盟：風險導向的監管範式與醫療器材的雙重管制

1.監管邏輯

歐盟AIA對於醫療領域的監管核心在於其分類機制。該法案並非單純列舉各類醫療器材，而是透過結構性的定義將醫療AI納入高風險範疇。根據法案第六條第一款，若AI系統本身是一個產品，或者作為產品的安全組件，且該產品受歐盟附錄I所列的調和立法規範，並因此需要進行第三方合格評定，則該AI系統自動被歸類為高風險系統。附錄I明確列入了醫療器材法規（MDR, 2017/745）與體外診斷醫療器材法規（IVDR, 2017/746）。這意味著，凡是屬於MDR分類下的Ⅱa級、Ⅱb級及Ⅲ級醫療器材，以及大部分IVDR管制的體外診斷設備，若其功能依賴於AI技術，將直接受到AIA的高風險監管。

除了上述基於產品安全法規的分類外，法案附錄Ⅲ進一步擴大了高風險的定義範疇，明確將「用於緊急醫療保健病患分流（Triage）」的AI系統列為高風險。這類系統雖然可能不直接接觸人體，但其演算法決定了急救資源的分配與優先順序，一旦發生錯誤判斷，將直接導致病患死亡或重傷。歐盟立法者透過此一條款，補強了傳統醫療器材法規可能忽略的「流程管理」風險，確保所有涉及生命攸關決策的演算法皆受到嚴格審視。

這種分類邏輯的深層意涵在於，歐盟不信任單純的市場機制能解決醫療AI的安全問題。透過將醫療AI鎖定在高風險類別，歐盟強制要求這些系統在進入市場前（Ex-ante）必須通過嚴格的合規性測試，而非僅在發生事故後進行補救。這建立了一個極高的市場准入門檻，要求開發者在設計階段就必須將歐盟的基本權利價值觀寫入程式碼之中。

2.提供者的義務：數據治理與技術穩健性

對於被歸類為高風險的醫療AI系統，歐盟《人工智慧法案》第三章規定了詳盡的提供者義務。其中，數據治理（Data Governance）的要求尤為關鍵。在醫療領域，演算法的偏差可能導致對特定種族、性別或年齡層病患的誤診。因此，法案第十條強制規定，用於訓

練、驗證和測試的高風險AI系統的數據集，必須符合特定的品質標準。這些數據集必須具有相關性、代表性，且在可能的範圍內無錯誤且完整。供應商必須審查數據集是否存在可能影響健康安全的偏差，並採取適當措施進行檢測與修正。這意味著，如果一個皮膚癌診斷AI僅使用白種人的皮膚影像進行訓練，該系統將無法通過歐盟的合規審查，因為它對有色人種病患構成了潛在的歧視與健康風險。

此外，法案對「技術穩健性、準確性與網路安全」（Accuracy, Robustness and Cybersecurity）提出了具體要求。醫療AI系統必須具備對抗「對抗性攻擊」（Adversarial Attacks）的能力。在醫療影像診斷中，惡意攻擊者可能透過微幅修改影像像素來欺騙AI做出錯誤診斷。歐盟法規要求提供者必須在系統設計中內建防禦機制，防止此類攻擊導致的誤診風險。同時，系統必須具備「日誌記錄」（Record-keeping）功能，自動記錄系統運作期間的事件，以確保在發生醫療事故時，能夠追溯是數據輸入錯誤、演算法邏輯錯誤還是人為操作失誤，這為未來的醫療責任歸屬提供了關鍵的數位證據基礎。

3.透明度與人類監督：反對黑箱醫療

歐盟法規強烈反對黑箱技術在醫療中的應用。法案第13條規定了嚴格的透明度義務，要求高風險AI系統必須設計成能讓使用者（即醫師與護理人員）理解其產出結果。提供者必須提供詳細的使用說明書，解釋系統的性能特徵、限制以及預期的準確率。這不僅是為了讓醫生知道如何操作，更是為了確保醫生知道何時不該信任AI。與此相輔相成的是第14條的「人類監督」（Human Oversight）要求，高風險AI系統必須設計成能被人類有效地監督。在醫療場景中，這意味著AI不能完全取代醫生的判斷，醫生必須保留干預、中斷或駁回AI建議的權力。法規要求建立「人機介面」（Human-Machine Interface）工具，確保人類監督者能夠理解系統的功能限制，並具備足夠的權限來防止自動化偏誤（Automation Bias）意即人類過度依賴系統建議而忽略自身專業判斷的傾向。這項規定在法律層面上確立了「人為最終決策者」的地位，防止了醫療責任的真空。

4.禁止的AI運用與醫療例外

雖然法案第二章列出了一系列被禁止AI的運用，但在醫療領域，法案留下了必要的倫理例外空間。例如，法案原則上禁止使用AI進行情緒識別（Emotion Recognition）系統，特別是在工作場所或教育機構。然而，法案明確排除了醫療或安全目的的使用。這意味著，如果情緒識別AI是用於診斷憂鬱症患者的病情變化，或是監測中風患者的臉部表情以預警，則屬於合法範疇。這種細緻的區分顯示了歐盟立法者試圖在保護公民免受侵入性監控與促進醫療創新之間取得平衡，確保禁止性規定不會阻礙具有正當治療目的的技術發展。

5.雙重法規的整合挑戰

對於醫療器材製造商而言，歐盟AI法案的最大挑戰在於如何與現有的MDR/IVDR法規整合。法案的設計初衷是讓AI的合規性評估成為現有醫療器材合格評定程序的一部分，即「單一審核」機制。製造商不需要分別申請兩張證書，而是應在同一套技術文件中同時證明符合MDR的安全要求與AI法案的數據治理及演算法透明度要求。然而，這也意味著現有的認證機構必須具備審核AI演算法的專業能力，這對歐盟的監管基礎設施提出了巨大的能力建設挑戰。

（三）義大利：機構化治理與醫師主體的法律保障

1.醫師決策權的絕對性與反歧視原則

義大利AI專法第7條是針對醫療領域AI使用的核心條款。該條文以極為明確的語言界定了AI在醫療中的角色：「輔助」。法律規定，醫療領域的AI系統僅構成預防、診斷、治療及處方選擇流程的輔助工具，決策權始終保留給醫療專業人員。這一條款具有深遠的法律意涵，它不僅排除了全自動化醫療診斷的合法性，更在法律責任上鎖定了醫師的主體地位。無論AI的算法多麼精準，最終的診療責任由人類醫師承擔。

此外，第7條還引入了嚴格的非歧視原則，法律明文禁止在醫療系統引入AI時，根據歧視性標準來選擇或設定醫療服務的獲取條件。這是在回應演算法可能因數據偏差而導致醫療資源分配不公的擔憂（例如，AI可能因為保險支付能力或居住地數據而降低對特定病患的治療優先級）。義大利法律要求，AI的引入不得減損病患平等獲取醫療服務的權利，並賦予病患知情權，即病患有權被告知其診療過程中使用了AI技術。這將「告知同意」（Informed Consent）的概念從單純的醫療處置延伸到了使用的技術工具層面。

2.國家區域衛生服務局（AGENAS）的核心角色與國家AI平台

與歐盟AIA側重於市場監管不同，義大利AI專法展現了強烈的國家治理色彩。第10條授權建立一個專門支援醫療治療及社區援助的AI平台，並指定「國家區域衛生服務局」（AGENAS）負責該平台的設計、實施與管理。更關鍵的是，法律指定AGENAS為該平台內收集和生成數據的「處理控制者」（Data Controller）。這一舉措將醫療AI的核心資源即數據，置於國家公共衛生機構的直接控制之下，而非任由私人科技巨頭壟斷。

AGENAS透過該平台向衛生專業人員提供非約束性建議（non-binding advice），這再次呼應了AI的輔助性質。透過由國家機構統一管理AI平台，義大利試圖解決醫療資源碎片化的問題，確保全國各地的醫療機構都能獲得標準化、高品質的AI輔助工具，而不是讓富裕地區的醫院獨享先進技術。AGENAS還被賦予了制定數據匿名化及合成數據程序指導方針的權力（第8條），使其成為義大利醫療AI技術標準

的實際制定者與守門人。

3.數據沙盒與科學研究的公共利益

為了促進醫療AI的研發，義大利法律在隱私保護與數據利用之間開闢了一條綠色通道。第八條宣告，為了預防、診斷、治療疾病及藥物開發等目的進行的AI研究與實驗數據處理，具有重大公共利益（Major Public Interest）。這一法律定性至關重要，因為它允許研究機構在履行告知義務後，對去識別化的個人資料進行「二級使用」（Secondary Use），而無需重新取得每位病患的單獨同意。這解決了長期以來困擾醫療AI研發的數據授權瓶頸。第9條進一步要求衛生部長在120天內頒布法令，建立專門用於研究的「監理沙盒」（Regulatory Sandbox）。這個沙盒機制允許研究者在受控的環境下，利用真實世界的健康數據測試AI演算法，同時享有一定程度的監管豁免或簡化程序。這顯示義大利政府試圖透過制度創新，將該國打造為醫療AI研發的樞紐，同時利用AGENAS的監管能力確保這些實驗不會侵害病患隱私。這種「國家平台+監理沙盒」的雙軌制，構成了義大利推動醫療AI的獨特模式。

（四）日本：國民生活平穩與跨學科的倫理防線

1.替代人類判斷的技術定義與風險認知

日本AI推進法第2條將AI相關技術定義為「以人工方法替代人類認知、推論及判斷等智力能力」的技術。在醫療語境下，這一定義直指核心：當AI用於診斷癌症或規劃手術路徑時，它實際上是在「替代」醫師的認知與判斷過程。正因為承認這種替代性，法律隱含了對技術可靠性的極高要求。如果一個系統被設計來替代人類進行攸關生死的判斷，那麼它必須具備與人類相當甚至更高的穩定性。第3條的基本理念進一步闡述了這種風險意識。法律明確指出，AI是「安全保障觀點上之重要技術」。這裡的安全保障不僅指國家防衛，更涵蓋了社會基礎設施與公共衛生的安全。第3條第4項發出了嚴厲的警告：若AI以不正當目的或不適當方法使用，恐將助長危害國民生活平穩及國民權益的事態。對於醫療AI而言，所謂「不適當方法」可能包括使用了偏差數據、演算法缺乏解釋性，或是未經充分驗證即投入臨床使用。為了防範這些危害，法律強制要求必須採取措施確保研發與活用過程的「透明性」。在日本的監管邏輯中，透明性是信任的前提，也是保護國民生命安全的第一道防線。

2.跨學科研發義務：人文與科學的融合

日本法第6條對研究開發機關（如大學、國立研究所）提出了一項獨特且具前瞻性的要求：為了有效推進研發，必須「綜合活用人文科學及自然科學等多元領域之知識」。這一條款在醫療AI領域具有關鍵意義。它意味著，開發醫療AI不能僅僅是資訊工程師（自然科學）的工作，還必須納入倫理學家、法學家、社會學家（人文科學）以及臨床醫師的觀點。這種跨學科的要求是為了確

保AI系統在設計之初就考慮到了生命倫理（Bioethics）與社會接受度。例如，在開發長照機器人或臨終關懷AI時，單純的技術效率（如省電、反應速度）不能凌駕於對病患尊嚴的考量之上。法律強制要求這種知識融合，是為了防止技術發展脫離人性的軌道，確保AI系統在介入人類身體與生命時，能夠符合社會的倫理期待。

3.事業者的責任與社會契約

日本AI推進法第7條規定了「活用事業者」（即AI產品與服務的提供者，包括醫療器材商）的責任。業者不僅被要求利用AI提升業務效率，更負有「依循基本理念」進行事業活動的法律義務。這意味著，企業在追求利潤的同時，必須主動維護「國民生活的平穩」。如果一家醫療AI公司推出了存在嚴重缺陷的產品，導致社會大眾對醫療系統產生恐慌，這不僅是產品責任問題，更是違反了本法的基本理念。此外，法律要求業者必須「配合國家與地方公共團體實施的施策」。這為政府未來頒布具體的醫療AI指導方針（Guidelines）預留了法律效力，構建了一種政府與企業間共同維護社會安全的社會契約。

（五）韓國：高影響AI的精準定義

1.醫療領域的高影響AI定義

根據韓國AI基本法第2條第4項，高影響AI」是指對人類生命、身體安全及基本權利有重大影響或可能造成風險的AI系統。法律明確列舉了屬於此類別的醫療領域應用：

(1)保健醫療提供與利用體系（第3款）

依據《保健醫療基本法》所建立的醫療服務提供與利用系統。這涵蓋了醫院的資訊管理系統、遠距醫療平台以及國家級的健康數據網絡。

(2)醫療器材與數位醫療產品（第4款）

依據《醫療機器法》開發的醫療設備，以及依據《數位醫療製品法》定義的數位醫療產品（如AI軟體醫療材SaMD）。

(3)生物特徵識別（第6款）

用於犯罪偵查的生物特徵分析（如指紋、虹膜、臉部識別）。雖然主要針對執法，但涉及人體生理數據的處理，與醫療隱私高度相關。

這種列舉式的定義方式消除了模糊地帶，明確告知產業界哪些醫療AI應用將受到最嚴格的監管。這與韓國政府試圖在確保安全的前提下，快速推動醫療AI商業化的意圖相符。

2.人工智慧事業者經營者之解釋義務

根據韓國AI專法第3條第2項受影響者有權獲得關於AI決策標準與原則的明確且有意義的說明，針對高影響AI事業經營者必須證明其系統的「可靠性」與「安全性」。這不僅是技術標準，更包含了「可解釋性」（Explainability）的要求。在醫療場景中，這意味著當AI建議進行某項高風險手術或拒絕某項保險給付時，病患有權要求經營者解釋該建

議背後的邏輯。

此外，法律要求國家與地方政府在制定政策時，必須保障身心障礙者與高齡者等「AI弱勢群體」的參與權與使用便利性（第3條第5項）。這體現了韓國法規中的包容性原則，確保醫療AI的進步不會造成數位落差，排除了社會中最脆弱的群體。

（六）越南：數位主權與必要領域的國家管控

1.醫療作為受限制的必要領域

越南AI專法第6條第1、2項明確規定，AI在各行業的應用必須遵守風險管理原則。對於直接影響人類生命、健康、合法權益或社會秩序的「必要領域」，必須實施更嚴格的風險管理。醫療領域被明確指定為此類領域之一。法律對醫療AI提出了三項具體要求：

(1)確保病患絕對安全：這是法律上的最高指導原則，不容許任何妥協。

(2)在實際使用條件下的可靠性：要求AI系統不僅在實驗室環境下有效，在越南現實的醫療環境（可能面臨設備老舊、網路不穩等挑戰）中也必須保持穩定。

(3)依據法律保護健康數據：強調數據的主權與隱私保護。

此外，依同條第4項授權政府各部會（如衛生部）針對其管轄領域制定詳細的安全要求與部署條件，賦予了行政機關極大的裁量權來規範醫療AI的准入。

2.事前通報制度

對於高風險與中風險系統，越南實施了嚴格的「事前通報」制度。第10條第1項規定，供應商在投入使用前，必須先進行自我分類，並建立分類檔案。更關鍵的是，供應商（Provider）必須透過「人工智慧電子單一窗口門戶」（AI One-Stop Portal）向科技部通報分類結果。這意味著所有在越南運行的醫療AI系統都必須在政府資料庫中註冊備案。這種集中化的管理模式讓越南政府能夠即時掌握全國高風險AI的部署情況，並在必要時進行干預。如果供應商無法確定風險等級，法律允許其向科技部申請指導，這進一步強化了政府在風險認定上的最終解釋權。

3.部署者的責任與系統完整性

越南法律特別強調「部署者」（Deployer，即醫院或醫療機構）的責任。第10條第2項規定，部署者雖然可以繼承供應商的分類結果，但必須負責維護系統的安全性與完整性。如果部署者對系統進行了修改或整合（例如將AI模型整合到醫院的電子病歷系統中），導致產生了新的或更高的風險，部署者必須與供應商協調，重新進行分類。這條規定防止了AI系統在在地化部署過程中因修改而產生的安全漏洞，確保了責任鏈條的完整性。

三、小結：殊途同歸的監管光譜

綜合分析五國的監管架構，我們可以發現全球醫療AI監管正在形成一種「殊途同歸」的趨勢：儘管各國的出發點與手段不同，但最終都指向了「人類監督」與「風險分級」

這兩個核心原則。

（一）風險定義的差異

1.**歐盟**：採取「產品責任」視角，將醫療AI視為需要通過合格評定的產品，其風險定義與既有的醫療器材法規高度整合，強調市場准入的技術門檻。

2.**日本**：採取「社會平穩」視角，將醫療AI視為人類判斷的替代品，風險在於其可能破壞國民生活的安寧，因此強調跨學科的倫理檢視。

3.**韓國**：採取「後果導向」視角，聚焦於「高影響」，將監管視為建立社會信任、推動產業發展的必要基礎設施。

4.**越南**：採取「國家主權」視角，將醫療AI視為必須納入國家登記體系的必要領域，強調政府的集中管控。

5.**義大利**：採取「機構治理」視角，透過AGENAS這樣的專責機構來管理數據與平台，並以法律強行保留醫師的最終決策權，體現了對專業主義的堅持。

（二）控制機制的比較

1.**事前審查**

歐盟與越南都強調事前的合規性審查或通報。歐盟依賴第三方認證機構（Notified Bodies），越南則依賴國家單一窗口。

2.**事中監督**

義大利透過AGENAS的平台進行即時的數據管理與建議；日本則要求業者持續保持透明度並配合國家施策。

3.**人類介入**

義大利的法律最為強硬，明文規定決策權「始終」屬於醫師；歐盟要求設計上必須允許人類監督；越南則要求保持人類的控制權。

（三）數據治理的創新

各國都意識到數據是AI的血液。歐盟要求數據集的無偏差性；義大利則開創性地設立了「數據沙盒」與「重大公共利益」條款，試圖在GDPR的框架下為醫療研發解套；韓國則強調生物特徵數據的嚴格管控。

總結而言，未來的醫療AI監管將不再是單純的技術標準之爭，而是關於「人機關係」的法律定性。無論是義大利的「醫師主權」、日本的「生活平穩」、還是歐盟的「基本權利」，各國法律都在試圖回答同一個問題：在演算法日益強大的時代，我們如何確保醫療決策依然具有人性、當責性與安全性。這些立法實踐顯示，全球正從放任發展的「軟法時代」邁向嚴格管控的「硬法時代」，而醫療領域正是這場監管革命的最前線。

肆、台灣醫療領域人工智慧之規範模式分析與建議解決方案——代結論

一、立法架構分析

（一）基本法性質與分散式治理

人工智慧基本法於2025年12月23日經通過立法三讀，為台灣首部全面性的人工智慧（AI）治理框架法律。依據本法第一條及立法理由，本法定位為宣示國家人工智慧治理基本價值與政策方向的基本法，而非直接創設具體處罰規定的作用法，其目的在於提供跨部門政策整合的共同價值基準，避免技術過早僵化，故本法採「分散式治理」架構，

由行政院成立「國家人工智慧戰略特別委員會」統籌（第6條）；又設有中央主管機關為國家科學及技術委員會[58]（第2條第1項）；另由數位發展部負責推動「風險分類框架」（第16條）；而本法所定事項，涉及各目的事業主管機關職掌者，又交由各目的事業主管機關辦理（第2條第2項），舉例來說，醫療領域的具體應用與規範，則由目的事業主管機關即衛生福利部依據本法原則及風險框架辦理（第16條第2項）。

（二）抽象的核心原則與風險分類

1.抽象的核心原則

本法第4條1至7款揭示之「永續發展與福祉」[59]、「人類自主[60]」、「隱私保護與資料治理[61]」、「資安與安全[62]」、「透明與可解釋[63]」、「公平與不歧視[64]」及「問責[65]」等抽象原則，涵蓋歐盟及各國目前AI管制上位法理基礎。此等原則提供醫療AI應用的價值指引，例如要求醫療AI系統應重視患者隱私與資料安全、避免演算法偏見歧視病患族群、提高模型決策透明度與可說明性，以及強調人類（醫師）對AI決策的最終監督權責。

2.抽象風險分級管制

本法第16條是管制架構的核心，規定數位發展部（MODA）應參考國際標準推動「風險分類框架」，並協助各目的事業主管機關（如衛福部）訂定「以風險為基礎之管理規範」，而各目的事業主管機關（如醫療領域的衛生福利部）應依此框架，針對不同風險等級的AI應用制定相應管理措施，並協助產業制定自律指引。若AI應用出現侵害人民生命、身體、自由或財產，破壞社會秩序、國家安全或生態環境，或偏差、歧視、廣告不實、資訊誤導或造假等第5條列舉的情形，則屬不可容忍的風險，應依法限制或禁止之。此機制類似歐盟「風險分級管理」理念，但又未明示風

註58：本法有出現中央主管機關僅在第18條第2項在法規2年訂定或修正之真空期，既有法規未有規定者，由中央目的事業主管機關會商中央主管機關，依本法規定解釋、適用之。

註59：第四條第1款永續發展與福祉：應兼顧社會公平及環境永續。提供適當之教育及培訓，降低可能之數位落差，使國民適應人工智慧帶來之變革。

註60：第四條第2款人類自主：應以支持人類自主權、尊重人格權等人類基本權利與文化價值，並允許人類監督，落實以人為本並尊重法治及民主價值觀。

註61：第四條第3款隱私保護與資料治理：應妥善保護個人資料隱私，尊重企業營業秘密，避免資料外洩風險，並採用資料最小化原則；同時在符合憲法隱私權保障之前提下，促進非敏感資料之開放及再利用。

註62：第四條第4款資安與安全：人工智慧研發與應用過程，應建立資安防護措施，防範安全威脅及攻擊，確保其系統之穩健性與安全性。

註63：第四條第5款透明與可解釋：人工智慧之產出應做適當資訊揭露或標記，以利評估可能風險，並瞭解對相關權益之影響，進而提升人工智慧可信任度。

註64：第四條第6款公平與不歧視：人工智慧研發與應用過程中，應盡可能避免演算法產生偏差及歧視等風險，不應對特定群體造成歧視之結果。

註65：第四條第7款問責：應確保承擔相應之責任，包含內部治理責任及外部社會責任。

險分類分幾級及分級標準，為醫療AI等高風險應用預留更嚴格監管空間。

3.高風險人工智慧責任與救濟

(1)高風險人工智慧之標示

依本法第5條第2項規定政府應以兒少最佳利益為原則，人工智慧產品或系統經中央目的事業主管機關會商數位發展部認定為高風險應用者，應明確標示注意事項或警語。至於如何認定評估為高風險，由數位發展部及其他相關機關提供或建議評估驗證之工具或方法。

(2)高風險人工智慧之責任歸屬與救濟

依本法第17條明定政府應就高風險人工智慧之應用，明確其責任歸屬，並建立救濟、補償或保險機制。

此條文為未來的「AI醫療事故補償制度」開啟了法源依據。鑑於AI判斷錯誤的因果關係難以證明，傳統侵權行為法（過失責任）可能對病患保護不足。本法暗示了可能採行類似「藥害救濟」的無過失補償或強制責任保險模式。

二、立法架構之優缺點分析──從醫療領域

（一）優點（Pros）

1.**彈性與適應性**：醫療AI技術迭代極快（如從CNN到Transformer架構）。基本法僅確立原則（如人類自主），細節留給衛福部透過指引調整，可避免法律修訂不及導致的技術僵化。

2.**尊重專業自主**：醫療是高度專業化領域。基本法模式允許衛福部依據臨床實務定義風險，而非由不熟悉醫療的數位部或國科會主管機關強行介入。

3.**產業衝擊緩衝**：相較於歐盟AIA直接施加重罰與繁重的合規義務，台灣模式給予產業較長的適應期（Buffer Period），有助於本土中小企業與新創發展。

4.**促進AI發展的法律解釋原則及環境**：本法第11條第1項後段人工智慧研發及應用之法規解釋與適用，如與其他法規扞格，在符合本法第4條基本原則之前提下，以促進新技術與服務之提供為優先原則。本法第13條第1項政府應建立資料開放、共享及再利用機制，以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。

（二）缺點（Cons）與結構性缺陷

1.**規範真空期**：基本法通過後，至各部會完成作用法修訂前（依第18條規定為兩年內），存在法律空窗期。此期間若發生AI事故，責任歸屬仍不明確。特別是現在已在市場上運作之AI醫療系統或器材，是否應從頭來過或者就地合法，並不明確。

2.**跨部會協調成本**：以醫療情境而言，醫療AI涉及個人數據（個人資料保護委員會[66]，下稱個資會）、網路安全（數位部）全民健康保險資料[67]及醫療行為

註66：個人資料保護法第1-1條主管機關為個人資料保護會。

註67：依114年12月19日公布全民健康保險資料管理條例第2條主管機關為衛福部。

（衛福部）。分散式立法可能導致「多頭馬車」、「兩人三腳」，例如屬於個資的利用權限，若個資會與衛福部見解不同，將阻礙研發訓練。

3.**缺乏強制力**：基本法多為宣示性條文。若無具體罰則配合，第四條的「可解釋性」與「公平性」恐流於道德勸說，廠商可能為追求效能而犧牲安全性。

4.**欠缺具體定義與課責主體**：本法除第3條對人工智慧有加以定義外，其他對於人工智慧開發者、提供者、部署者及使用者均欠缺定義，因此，本法所要求之義務應該要求何人履行以及未履行應由誰來負責，付之闕如。而欠缺高風險定義與分類，導致何種醫療AI屬「高風險」本法未細列。雖第5條授權主管機關認定高風險AI並要求標示警語，但醫療AI的高風險情境需要具體界定，例如診斷類AI是否一律高風險？輔助決策類呢？。缺乏細分類標準可能導致實務認定模糊。

三、醫療AI良善治理的解決方案建議

我國在制定人工智慧基本法後，為使其在醫療領域切實發揮作用，尚需配套措施加強高風險醫療AI的治理。以下提出三方面建議：

（一）建立AI醫療系統風險分級架構與監管工具

依本法第16條第1項以風險為基礎之管理規範意旨，醫療AI的風險程度可類比醫療器材，依對患者潛在危害大小進行分級分類。建議衛生福利部參考醫療器材風險矩陣模式，制定AI醫療系統風險分級辦法。例如：

1.**低風險級（如健康管理聊天機器人、醫療知識QA）**：不直接影響臨床決策，可採取自律為主、較寬鬆的監管方式，要求開發者遵守基本的個資與廣告真實原則即可。

2.**中風險級（如輔助判讀影像的AI工具、病歷語音轉錄系統）**：對診斷有一定影響但醫師仍全權掌控，可要求事先功能測試驗證與使用者教育。監管工具包括產品註冊備查、臨床評估報告提交，以及醫院內部的技術評估委員會審核後方可使用。

3.**高風險級（如AI自動判讀並提出診斷建議、AI治療劑量調控系統）**：可能直接影響診療結果，若失誤將對病患生命健康造成嚴重後果。對此類高風險醫療AI，應比照高階醫療器材管理：強制通過主管機關核可或第三方認證，提供足夠的臨床試驗證據證明安全有效；上市前需取得許可或執照，上市後納入持續監管（如定期報告、重大事件通報）。同時可借鑒歐盟做法建立AI系統登錄制度，公開高風險AI的用途、性能和適用範圍，接受社會監督。

在監管工具上，建議發揮現行醫療器材管理體系的作用。將AI醫療軟體納入「軟體作為醫療器材」（SaMD）的監管框架，明確不同風險級別對應的審查程序與技術標準。例如，要求高風險AI軟體提交算法描述與可解釋性報告、訓練資料多樣性分析、模型性能（靈敏度、特異度）評估等文件。為解決持續

學習問題，可引入「變更管制」制度即規定AI模型若經重大版本更新，視同新產品需重新評估批准；較小幅度的更新則須向主管機關備案，並定期提交性能監測報告，以確保模型演進不導致安全隱患。此外，監管機關可建立沙盒機制或試驗性質許可，允許醫療AI在嚴格條件下先行試用，同時收集實證數據來完善後續審查標準。

（二）合法運用健保資料訓練本土AI模型

發展高品質的本土醫療AI，需要大量具代表性的醫療數據作為模型訓練素材。台灣擁有涵蓋99%以上人口的全民健康保險資料庫，以及完整的電子病歷系統，這是本土AI模型的珍貴資源，是訓練醫療AI的寶庫。然而，憲法法庭111年憲判字第13號判決指出，健保資料庫在「缺乏獨立監督機制」及「缺乏當事人退出權（Opt-out）」的情況下，強制提供學術利用有違憲之虞。

此判決對AI產業造成重大影響，但也指引了合憲的解方：

1. **建立獨立監督機制**：政府已修訂《個人資料保護法》，將成立獨立的「個人資料保護委員會」（PDPC），取代過去由衛福部自行球員兼裁判的局面。
2. **退出權與公共利益之平衡**：判決承認「公共利益」可作為限制退出權的理由。因此，新制《全民健康保險資料管理條例》中，應明確將「促進醫療AI發展以增進全民健康」定義為重大公共利益，在高度去識別化及資安防護下，允許例外豁免或限制退出權，以避免數據偏差（Bias），可惜新公布法條仍欠缺上開例外及商用之許可。

政府應以「數位主權」高度，由國科會與衛福部主導，利用健保影像與病歷數據，訓練專屬於台灣的「醫療基礎模型」（Medical Foundation Model）。這不僅是技術問題，更是落實本法第13條「維護國家文化價值」及「展現國家多元文化價值」之目標，建議採取以下措施利用健保數據：

(1)聯邦式學習（Federated Learning）[68]：不將原始健保數據匯出，而是讓AI模型進入健保署的封閉環境（Sandbox）進行訓練，僅將訓練好的參數（Weights）帶出。這既符合本法第四條的「資料最小化」原則，又能徹底解決個資外洩疑慮。

(2)合成資料（Synthetic Data）[69]之開發：利用健保大數據生成不含真實個資但具備相同統計特徵的「合成資料」，供新創業者進行初期模型訓練與驗證。

（三）明確醫師主導原則下的責任分擔機制

醫療AI雖能輔助決策，但基於病患安全與專業倫理，醫師應始終保有最終決策權。為落實「醫師作最終決策主體」原則，同時合

註68：Federated Learning聯邦學習簡介，清華大學，https://biic.ee.nthu.edu.tw/blog-detail.php?id=2（last visited on 2026/01/05）。

註69：范晏儒，合成資料（synthetic data），2020.10.，https://stli.iii.org.tw/article-detail.aspx?no=64&tp=1&d=8532（last visited on 2026/01/05）。

理分配AI引入後的責任，建議從法律和制度上做出以下設計：

1.醫療決策責任歸屬

在醫師法或相關醫療法規中明文化，醫療AI的診斷或治療建議僅具參考性質，執業醫師須對最終診斷與處置承擔法律責任。換言之，AI不能取代醫師的醫療判斷義務，即便醫師採納AI建議亦須確認其合理性後再執行。此條款將使法官在處理醫療糾紛時有明確依據，通常仍以審視醫師是否盡到專業注意義務為判斷標準。對醫師而言，這強調了審慎使用AI的義務，避免過度信賴技術而疏忽臨床經驗和病人個別狀況。

2.AI提供者與醫療機構責任

為平衡責任，加強對AI提供者的約束，可考慮在醫療器材管理辦法或未來的AI子法中，增設AI提供者義務與責任條款。包括：要求廠商對其AI產品進行充分測試與風險告知，保證符合核准時的性能指標；若因演算法明顯缺陷或未盡告知義務造成患者損害，廠商需承擔產品責任或行政罰則。醫療機構在引入AI時，亦應負有審查選擇責任與培訓監管責任。例如醫院應設置AI技術評估委員會，審慎評估AI工具的適用性；在使用過程中監控AI建議品質，及時反饋給廠商改進。若醫院明知AI存在重大問題仍放任使用，對患者損害可能需承擔民事賠償責任（類似醫院用人不當或設備維護不當的責任形式）。

3.責任保險與賠償機制

建立覆蓋醫師、醫院與提供者的多層次保險制度作為風險分擔工具。醫師已普遍參與醫療責任險，建議擴充保險範圍，納入AI決策相關風險。對於AI產品供應商，則可要求投保產品責任保險或專責AI保險，確保一旦其技術缺陷導致事故，有財務資力賠償受害患者。政府亦可考慮設立醫療AI風險補償基金，由政府、醫療機構與產業按比例出資。當發生責任難以歸屬或多方均無過失但患者受害的情形時（如AI罕見失誤案例），基金可對患者進行無過失補償，體現對創新風險的社會承擔。這一機制類似於疫苗傷害補償制度，在鼓勵AI創新的同時保障患者基本權益。

4.制定指引與教育培訓

由衛福部或醫師公會發布醫療AI臨床使用指引，明訂醫師在使用AI輔助時的標準作業流程和注意事項。例如，要求醫師在參考AI診斷時，需結合傳統診斷結果複核，不可單憑AI結論下判斷；遇AI建議與臨床觀察不一致時，以醫師專業判斷為準等。並將此納入醫師繼續教育課程，提升醫療人員對AI的正確認知與風險意識。

透過以上措施，可形成「醫師主導—廠商擔保—機構監督—保險支撐—無過失補償」的責任生態：醫師對患者直接負責，廠商對其產品品質負責，醫院對AI使用環境負責，而保險與補償基金為意外損害提供經濟保障。如此一來，在鼓勵AI輔助診療應用的同時，各相關方的責任邊界清晰，病患權益亦能獲得多層保障。

四、總結

台灣人工智慧基本法為AI治理立下重要里程碑，在醫療領域提供了基本的原則框架與高風險監管方向。然而，面對醫療AI技術的特殊風險與高度專業性，僅有基本法的原則性規定仍嫌不足。透過比較歐盟、南韓、越南、義大利經驗可見，各國皆採行了風險分級、專法配套與責任釐清等措施，以平衡AI創新發展與風險管控。展望未來，我國應加速制定醫療AI相關的子法與技術指引，將基本法的精神落實到具體規範中，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。唯有如此，才能在保障病患安全與權益的同時，釋放人工智慧對醫療品質與效率提升的巨大潛能，實現科技與人權並進的智慧醫療新時代。

附表1

通過順序	國家／區域	法律名稱（中文／原文）	通過日期	生效／施行時點	規範核心與重點（摘要）
1	歐盟	人工智慧法（EU Artificial Intelligence Act, Regulation (EU) 2024/1689）	2024-03-13	2024 年起分階段生效（多數核心義務於 2026 年全面適用）	採風險分級管制（不可接受／高風險／其他）；對高風險AI實施全生命週期硬法監管；醫療AI因連動醫療器材法制，原則列為高風險；強調資料治理、合規評估、人類監督、上市後監測與高額制裁。
2	韓國	人工智慧發展與建立信任基礎基本法（인공지능 발전과 신뢰 기반 조성에 관한 기본법）	2025-01-21	2026-01-22	採「高影響人工智慧」模型；僅對生命、健康或基本權高度影響之領域（含醫療）課予較具體義務；兼顧產業振興、影響評估、透明揭露與人類監督。
3	日本	人工智慧相關技術研究開發及活用推進法（人工知能関連技術の研究開発及び活用の推進に関する法律）	2025-05-28	2025-09-01	政策推進與軟法整合模式；未採剛性風險分級，而由國家制定AI基本計畫與指引；重在促進應用、風險評估與妥適性治理，醫療AI以行政指引為主要治理工具。
4	義大利	人工智慧法（Legge 23 settembre 2025, n.132- Disposizioni in materia di intelligenza artificiale）	2025-09-23	2026-01-01	歐盟AI Act之國內落地實施型專法；指定國家主管機關並設監理沙盒；對醫療、司法、勞動等特定領域明文規範；強調醫師與法官為最終決策主體，並增訂AI犯罪加重處罰。
5	越南	人工智慧法（Luật Trí tuệ nhân tạo）	2025-12-10	2026-03-01	國家管理／行政管制導向；採三層風險分類（高／中／低）；對中風險AI亦課通報與標示義務；並明文列舉禁止行為，醫療AI多屬高風險。
6	中華民國（臺灣）	人工智慧基本法	2025-12-23	另定（配合子法與行政配套分階段施行）	原則性基本法定位；著重倫理、風險治理、透明與人類監督；未設具體風險分級，醫療AI仍仰賴醫療器材法、個資法及後續部門法與指引補強。