從負責任AI執行論醫療AI開發及應用

吳英志*

壹、前言

政府為大力發展智慧醫療產業及將智慧醫療軟體導入於臨床治療、健康照護、公共衛生及醫學研究等領域中，除了從2020年1月15日公布醫療器材管理法，且於該法第三條[1]明定如用於診斷、治療、緩解或直接預防人類疾病之醫療器材軟體（包含醫療器材人工智慧軟體），則應同受醫療器材管理法規定所管制，以確保醫療器材使用安全、效能及品質。另於2021年9月9日修正生技醫藥產業發展條例，將「再生醫療」、「精準醫療」及「數位醫療」[2]納入生技醫藥產業發展條例適用範圍，大力鼓勵智慧醫療產業發展。又於2023年11月22日數位醫療發展條例草案公聽會[3]中，對於在數位醫療與智慧醫療發展架構下，通訊診察治療辦法應如何修正、數位創新醫療價值評估、創新（智慧）醫療器材之健保給付沙盒等有關數位醫療創新議題進行廣泛討論；末由總統府於2024年8月成立健康台灣推動委員會，推出11項台灣健康政策項目，其中為優化醫療工作環境與鼓勵醫療創新發展，更推出兩大健康政策方向[4]，分別為「智慧醫療結合健康照護」及「健康台灣深耕計畫」，希望透過智慧科技醫療導入臨床診斷、臨床照護、疾病預防、公共衛生防疫監測、醫學研究應用、藥物開發應用及醫療行政管理等方式，優化醫療照護流程和效率及提高臨床醫療精準性與安全性。

然而，人工智慧運用於臨床治療、健康照護、公共衛生及醫學研究等各項領域固然深受上述各項政府政策獎勵及法令鬆綁所支

＊本文作者係眾勤法律事務所律師，全國律師聯合會醫藥及健保法制委員會副主任委員。

註1：醫療器材管理法第3條：「本法所稱醫療器材，指儀器、器械、用具、物質、軟體、體外診斷試劑及其相關物品，其設計及使用係以藥理、免疫、代謝或化學以外之方法作用於人體，而達成下列主要功能之一者：一、診斷、治療、緩解或直接預防人類疾病。二、調節或改善人體結構及機能。三、調節生育。」

註2：生技醫藥產業發展條例第4條第7款：「數位醫療：指以巨量資料、雲端運算、物聯網、人工智慧、機器學習技術應用於健康醫療照護領域，且用於提升疾病之預防、診斷及治療，並經主管機關會同中央目的事業主管機關審定之創新性產品或技術。但屬人工智慧或機器學習技術之醫療器材軟體，由中央目的事業主管機關審定。」

註3：立法院社會福利及衛生環境委員會（2023），「數位醫療發展條例草案」公聽會報告。

註4：衛生福利部（2025），《2025年健康台灣全國論壇，健康台灣整體推動成果》。

持，惟人工智慧應用卻也存在著醫療隱私侵犯、AI系統安全性不足、AI算法偏差、歧視與黑箱，以及AI系統錯誤與責任歸屬等不可預測風險，是以衛生福利部資訊處於2025年間成立負責任AI執行中心，並參考世界衛生組織（WHO）提出使用人工智慧於醫療領域的六大原則，以及國科會公布「人工智慧基本法草案」揭示七大原則，提出三大AI治理策略即AI落地三大實施策略，以提高AI於臨床應用的透明度與效率，及增進公眾對於AI技術在醫療領域應用的信任與接受度。

惟有疑問的是，負責任AI執行中心所提出三大實踐AI治理策略是否真的完全符合我國「人工智慧基本法草案」七大倫理原則要求？其具體實踐策略內涵為何？倘若醫療機構未遵守三大AI治理實施落地策略，造成醫療機構AI治理缺陷或發生錯誤，該醫療機構是否應對於違反三大AI治理落地實踐策略，負起相應法律責任？均屬人工智慧導入醫療領域過程中最具有爭議治理問題與法律責任歸屬問題。

貳、負責任AI執行準則與實踐

衛生福利部為落實上述總統府健康台灣推動委員會所訂定「智慧醫療結合健康照護」及「健康台灣深耕計畫」兩大政策計畫，乃於「智慧醫療結合健康照護」政策計畫部分，啟動三大AI中心計畫，分別為負責任AI執行中心、臨床AI取證驗證中心及AI影響性評估中心，以解決智慧醫療落地、取證及給付三大問題。另於「健康台灣深耕計畫」範疇三政策計畫部分[5]，導入智慧科技醫療於臨床醫療與醫療機構，並強化智慧醫療科技之資安治理、資料治理與AI治理。

又為落實上述「健康台灣深耕計畫」AI治理，衛生福利部負責任AI執行中心除了參考世界衛生組織提出的保護人類自主性（Protect Autonomy）、促進人類福祉、安全及公共利益益（Promote Human Well-being and Safety and the Public Interest）、確保透明性、可解釋性及可理解性（Ensure Transparency, Explainability, and Understandability）、促進責任感及當責性（Foster Responsibility and Accountability）、確保包容性和公平性（Ensure Inclusiveness and Equity）及促進可持續性AI（Promote Sustainable AI）的六大原則外，更以「人工智慧基本法草案」第三條揭示七大倫理治理原則為準則，訂定三大實踐AI治理策略，並鼓勵國內各家醫療院所參考衛福部負責任AI執行中心成立模式，以及衛福部三大實踐AI治理策略訂定方式，成立各家醫院專屬負責任AI執行中心及各家醫院專屬三大實踐AI治理策略，以期各家醫院導入人工智慧軟體於臨床治療中及醫療器材上之使用風險可被監控與預防。

一、負責任AI執行準則

我國為因應人工智慧技術創新之速度和可能面臨的挑戰，經行政院院會審查，於114年8月28日通過「人工智慧基本法草案」，

註5：衛生福利部（2025），「健康台灣深耕計畫範疇3導入智慧科技醫療懶人包」。

以期建立我國AI發展的基本方針。另外，我國人工智慧基本法草案以立足於「鼓勵創新、兼顧人權」的核心理念，揭示永續性、人類自主性、隱私保護及資料治理、安全性、透明性及可解釋性、公平性、可問責性等七大基本原則[6]，明確定錨國家發展方向，同時謹慎平衡產業創新與風險管理的關係，作為引導我國各政府機關研發與應用人工智慧之原則[7]。

（一）永續性

人工智慧研發與應用除追求科技與經濟發展外，更應兼顧社會公平、環境與經濟間之協調發展，並追求有益於人類及地球之發展，進而促進人類與地球之永續發展（sustainable development），故於基本法草案明定人工智慧之研發與應用應兼顧社會公平及環境永續，降低可能之數位落差，使國民適應人工智慧帶來之變革。

（二）人類自主性

人工智慧研發與應用目的乃是追求人類社會之輔助發展，而不是讓人工智慧取代人類，故人工智慧研發與應用應尊重人類自主權及個人人格權（含姓名、肖像、聲音）等基本權利與文化價值，並朝向以人為本的科技發展模式及人類可監督發展模式，方符合人類創造人工智慧輔助人類社會發展之目的。因此，基本法草案參考2019年經濟合作暨發展組織公布「人工智慧建議書」，明定人工智慧研發與應用目應支持人類自主權，尊重人格權等個人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。

（三）隱私保護及資料治理

人工智慧研發與應用過程中需要大量資料輸入，故關於資料之蒐集、處理與應用過程中，是否能確保資料安全與個人資料隱私最小侵害，乃成為目前人工智慧科技發展的最大疑慮。故而，基本法草案參考2022年美國「AI權利法案藍圖」，明定人工智慧之研發與應用：應妥善保護個人資料隱私，避免資料外洩風險，並採用資料最小化原則；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。

（四）透明性及可解釋性

透過人工智慧所生之自動化決策可能對於利害關係人產生重大影響，故須確保人工智慧自動化決策過程之公平性與自動化決策結果之準確性，並讓使用者及受影響權益之人均可理解自動化決策可能產生之風險與影響及其決策過程之可解釋性，以達可信賴人工智慧決策之目標。據此，基本法草案爰參考2019年歐盟「可信賴AI倫理準則」，明定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。

（五）安全性

人工智慧之研發與應用過程中是否能確保資料安全與系統穩健性（robustness），乃是目前人工智慧科技發展的最大疑慮之一，故而參考2022年美國「AI權利法案藍圖」及

註6：數位發展部，人工智慧基本法草案第三條立法條文及其理由說明。
註7：數位發展部新聞稿，行政院通過「人工智慧基本法草案」，2025年8月28日。

2024年新加坡「生成式AI治理架構」，於基本法草案中明定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及安全性。

（六）公平性

為降低人工智慧自動化決策對於利害關係人之影響，故應儘量確保人工智慧自動化決策之公平性，避免人工智慧演算法產生偏差或歧視之結果。準此，基本法草案參考2022年美國「AI權利法案藍圖」，明定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。

（七）可問責性

為避免人工智慧開發者、部署者或最終使用者濫用人工智慧演算法與自動化決策結果，造成人工智慧之研發與應用對於人類產生危害、人文社會秩序破壞、利害關係人受到偏差或歧視對待或個人資料與隱私保護受到嚴重之侵害或不可回復損害之侵害等不利後果，故應致力於人工智慧開發者、部署者與使用者之負責任機制建立，以維護社會穩定及人類永續發展。據此，基本法草案參考2024年新加坡「生成式AI治理架構」，並明定人工智慧之研發與應用當確保人工智慧研發與應用過程中不同角色承擔相應之責任，包含內部治理責任及外部社會責任。

二、負責任AI執行實踐

衛生福利部為鼓勵各家醫院機構遵守人工智慧七大倫理原則，推動負責任AI執行的落地管理，乃提出資訊安全維護與隱私保護、九大透明性及可解釋性分析及AI生命週期循環監管的三大AI治理實踐策略，以提升臨床醫療安全與信賴度

（一）資訊安全維護與隱私保護

為遵循上述人工智慧基本法草案第三條揭櫫隱私保護及資料治理原則與安全性原則，衛福部要求醫療機構應提出符合個人資料保護之資訊安全維護與隱私保護管理辦法，以確保個人資料保護受到最安全保護。

關於隱私保護管理部分[8]，係指在AI輸入個人資料時，應先移除可識別個人資料或對於個人資料進行加密處理，以降低個人資料被識別風險。另外，應明確規範個人資料之儲存地點與保存期限、外部管理人員可接觸個人資料權限，及建立定期查核個人資料機制及完善個人資料外洩應變機制。最後，針對需要上傳個人資料至醫院外雲端以執行AI推論運算，必須採取最高等級的保護措施，以防止個人資料於跨機構傳輸或異地傳輸過程中遭到個人資料被竊取或外洩風險。

關於資料隱私安全防護部分[9]，係指個人資料在AI應用上需要上傳至院外雲端情境下，醫療機構為降低系統被入侵或資料遭濫用的風險，應落實資訊安全防護機制。其具體防護措施建議可採取多層次的網路安全措施，包括設置防火牆與入侵偵測／防護系統（IDS/IPS），以防止未授權存取。另外，可透過網路分割，限制對關鍵系統與敏感資料的存取

註8：衛生福利部資訊處（2024），《衛生福利部資訊處三大類型智慧醫療中心技術手冊》，第7頁。
註9：同上註。

範圍；再者，可使用自動化掃描工具定期檢查潛在資訊安全弱點。最後，應明確規範外部管理人員接觸個人資料的權限及其所應負責任，並建立起定期查核資訊安全弱點機制，以確保資訊安全政策能在醫療機構日常營運中能持續被執行。

以新北仁康醫院負責任AI中心採用智慧病房AI設備為例[10]，新北仁康醫院向緯創醫學科技股份有限公司採購慧病房AI設備，該等AI設備包含BestShape VS（生理感測器），BestShape Care（動作感測器）。其中BestShape VS（生理感測器）設備獲得CE、NCC、FCC、日本telec認證，TFDA認證（衛部醫器製字第007955號），日本福祉器材認證TAIS コード：02037-000004。BestShape Care（動作感測器）則獲得CE、NCC、FCC、日本telec認證，TFDA認證（衛部醫器製字第007955號），日本福祉器材認證TAIS コード：02037-000004。前述兩項AI設備均獲得臺灣衛生福利部驗證及國際認證，足以證明AI設備安全性與有效性已可確保資料隱私安全防護。

（二）九大透明性原則及可解釋性分析

為遵守上述人工智慧基本法草案第三條揭示透明性及可解釋性原則，衛福部要求醫療機構應落實AI使用的九大透明性原則及可解釋性分析，以確保資料運算過程中公開透明及運算結果可被解釋。

所謂九大透明性原則[11]是指醫院在使用AI工具時，必須在醫院專門網站上公開使用AI重要資訊，包括AI資料來源、AI訓練模型、AI驗證資料、AI適用的臨床情境、FDA驗證結果等，其具體落地九大步驟為AI介入臨床診斷的基本資料及輸出結果（Details and output of the intervention）、AI介入臨床診斷目的（Purpose of the intervention）、AI介入臨床診斷限制與警告（Cautioned Out-of-Scope Use of the intervention）、AI資料輸入及算法模型（Intervention development details and input features）、確保AI資料輸入與運算的公平性（Process used to ensure fairness in development of the intervention）、確保AI輸入與運算正確性的外部驗證（External validation process）、AI模型表現結果的量化指標（Quantitative measures of performance）、AI介入臨床診斷的持續監控（Ongoing maintenance of intervention implementation and use）及持續驗證AI輸入及運算正確性與公平性（Update and continued validation or fairness assessment schedule）。

所謂可解釋性分析[12]係指人工智慧運用於臨床診斷過程中，令使用人工智慧作為輔助診斷的醫師，能夠理解和解釋人工智慧模型如何做出預測或決策，以強化化人工智慧使用透明性與可信任，其在醫療領域常見的可解釋性分析方法為SHAP值解釋方法（SHapley Additive exPlanations）和顯著性圖解釋方法（Saliency Maps）。

以臺北榮民總醫院腦轉移瘤腦轉移瘤人工

註10：新北仁康醫院負責任AI中心官網，智慧病房AI設備，https://njkh.com.tw/ai-benefit/。

註11：衛生福利部資訊處，前揭註8，第8-12頁。

註12：衛生福利部資訊處官網，https://aicenter.mohw.gov.tw/AC/cp-7202-82645-208.html。

智慧輔助診斷系統DeepMets運用九大透明性原則為例[13]，首先，該DeepMets輔助判斷系統係以腦部磁振造影為分析標的，並透過偵測及自動標註全部腦中疑似腦轉移瘤病灶的輪廓及腫瘤面積之方法，輔助放射診斷科醫師辨識腦部磁振造影中疑似的腦轉移腫瘤，增加醫師診斷準確率。其次，該人工智慧輔助診斷系統DeepMets不能取代醫師作為診斷病人腦轉移瘤病灶的唯一依據，仍應由醫師依據專業醫療判斷，決定病人應如何接受放射手術治療。再者，該DeepMets輔助診斷系統係以臺北榮民總醫院1000多例腦部轉移腫瘤個案的磁振造影影像及腫瘤圈註資訊為基礎，結合健保署磁振造影影像資料，進行DeepMets人工智慧模式建置與訓練，以提高磁振造影影像系統偵測及辨識之精準率。此外，該DeepMets為確保人工智慧系統開發過程公平性，於系統開發初期即納入不同性別、族群與年齡病患資料，進行人工智慧資料的輸入與訓練。另外，臺北榮總醫院醫療人工智慧發展中心團隊透過來自國內19家醫院不同廠牌磁振造影系統掃描參數所生磁振影像，驗證DeepMets系統模式及效能，完成醫療器材查驗登記前的臨床驗證，並取得衛福部核發醫療器材許可證。其後，DeepMets系統更透過擴大於臺北榮民總醫院臨床使用數及接受歐美多個醫療中心數百例以上臨床測試等方法，持續維護DeepMets系統高精準率。同時，也透過院內醫師們回饋使用DeepMets系統經驗予人工智慧系統研發團隊，令研發團隊可以繼續優化DeepMets系統。最後，因DeepMets使用過程中會於原始偵測影像中圈註病灶，使醫師能從影像中得到AI圈註之病灶及範圍，與已知臨床腫瘤大小、形狀及紋理等特徵，了解DeepMets輔助判斷是否精準與正確，從而得知病人是否患有腦轉移瘤，令DeepMets系統使用具有可解釋性。

（三）AI生命週期循環有效性監測

為符合上述人工智慧基本法草案第三條規定公平性原則，衛福部要求醫療機構導入AI系統至臨床醫學應用過程中，應對於AI系統在整個生命週期中的有效性進行持續監測和評估，以確保AI系統準確性、可靠性及公平性。

所謂AI生命週期循環監測有效性[14]，係指為確保AI系統導入臨床醫療環境中表現能夠持續符合預期準確性、可靠性及公平性，故對於AI準備導入臨床醫療之開發與部署階段進行有效性監測，AI系統導入臨床後運行、維護與改進亦持續進行定期有效性監測，以確保AI系統導入臨床醫療環境中的全生命週期均能夠獲得完整監測。

AI生命週期有效性監測的循環步驟[15]為1.AI產品設計與臨床醫療契合度；2.臨床資料篩選與標註；3.AI模型優化；4.AI產品開發；5.AI產品准入許可及監測；6.AI產品部署與臨

註13：臺北榮民總醫院負責AI執行中心官網，腦轉移瘤腦轉移瘤人工智慧輔助診斷系統DeepMets，https://wd.vghtpe.gov.tw/maic/Fpage.action?muid=20196&fid=18228。

註14：衛生福利部資訊處，前揭註8，第16頁。

註15：衛生福利部資訊處官網，https://aicenter.mohw.gov.tw/AC/cp-7202-82645-208.html。

床整合；7.實際臨床治療監測管理；8.AI產品通過持續有效性監測後上市販賣。例如：以實際檢測乳癌的AI影像診斷系統為例，首先，AI產品開發階段須先經過廣泛影像資料訓練；其後，進行臨床醫療環境部署後，需要持續監測AI影像訓練有效性；再者，隨著新的乳房攝影影像資料加入，必須對AI算法模型進行再訓練，並對算法模型偏差部分進行檢查和修正，及檢查AI系統的診斷結果是否與專家放射科醫師的診斷結果一致，以確保AI算法模式準確率及AI系統可靠性；最後，蒐集臨床醫師和放射科醫師的回饋意見，對於AI實際使用過程中問題進行改善。

參、結論

綜上所述可知，衛生福利部為配合總統府健康台灣推動委員會所提出「智慧醫療結合健康照護」及「健康台灣深耕計畫」政策方向，啟動負責任AI執行中心、臨床AI取證驗證中心及AI影響性評估中心，以解決智慧醫療落地、取證及給付三大問題。同時，為謹慎平衡產業創新與風險管理的關係，及協助國內醫療機構能將智慧醫療科技導入臨床診斷、臨床照護、疾病預防、公共衛生防疫監測、醫學研究應用、藥物開發應用及醫療行政管理等醫療領域，乃由負責任AI執行中心依循人工智慧基本法草案所揭示七大倫理原則即永續性、人類自主性、隱私保護及資料治理、透明性及可解釋性、安全性、公平性及可問責性，訂定「負責任AI實施指南」，透過「負責任AI實施指南」輔導國內各大醫療機構從資訊安全維護與隱私保護、九大透明性原則及可解釋性分析與AI生命週期循環有效性監測之三大執行面向建立起醫療體系及醫療機構AI治理制度。