

治理防禦對策

數據不落地

優先採用地端部屬的開源大模型，或與簽署BAA的合規醫療雲合作

導入差分隱私

在模型訓練階段加入數學雜訊，確保即使駭客進行成員資格推論攻擊，也無法反推出特定的單一病例數據

嚴格落實人機協作

限制AI產出的文本只能在院內受限制的HIS介面內呈現與編修，禁止任何形式的分享或導出

論壇貴賓

主持人

天主教輔仁大學
數據科學中心執行長 蒲若芳

與談貴賓

天主教輔仁大學附設醫院
智慧醫療中心主任 張孟霖

天主教輔仁大學
研究倫理中心主任 李憶菁

聯新國際醫療集團
法務長 張濱璿



新北市立
聯合醫院



115年

科技與醫療倫理論壇 AI應用與隱私

2026年7月20日

指導單位

新北市政府
法務部廉政署

主辦單位

新北市政府政風處
新北市政府衛生局
新北市立聯合醫院



關於論壇

目的

人工智慧（AI）的興起，已然成為醫療場域不可或缺的輔助角色，深刻影響醫療實務的諸多環節，隨著AI的大幅應用，相關新興問題也逐漸浮現，更衍生多變的倫理挑戰。

議題	主要焦點
AI與醫療文書應用與倫理爭議	- 醫院導入AI輔助撰寫病歷或護理紀錄之實際應用現況 - 醫療現場如何因應現況所遭遇操作上的困境，及相關倫理問題 - AI生成的相關文書發生錯誤，責任歸屬，以及後續衍生損害賠償等法律責任應如何釐清
AI時代的隱私攻防戰	- AI應用醫療現場發生或潛藏相關隱私外洩風險 - 追求AI模型訓練精準度與落實病患隱私保護之間，如何取得平衡 - 傳統知情同意原則的適用問題

生成式AI於醫療文書中的應用與倫理爭議

典型錯誤	態樣
轉錄錯誤	聽錯藥名、劑量與單位
幻覺	無中生有的病史、用藥及檢查紀錄
遺漏	漏掉關鍵的發現
脈絡錯置	例如家族史寫成個人史或左右錯置

- 簽名即負責
 - 應善盡「合理核對」義務
 - 主管機關將「未查核即採納」列為風險
- 合理核對怎麼做？
- 優先核對高風險欄位
 - 特別小心否定敘述
 - 口述時維持結構化習慣（主訴-病史-處置）

AI時代的隱私攻防戰

風險	外洩場景
訓練階段未完全去識別未病患隱私	- 駭客或惡意使用者可透過特定的提示詞不斷試探模型，誘使AI吐出真實個資 - 透過交叉比對重新識別出該名病患
直接使用消費級的公開AI工具	當直接將未去識別化的病例草稿、病理報告複製貼上到對話框，或將門診錄音黨上傳至雲端轉譯，這些敏感數據會脫離進入第三方伺服器，甚至作為對方公開訓練數據，造成不可逆的永久性外洩
提示詞注入攻擊與越權存取	院內非相關科別的人員，甚至被外部駭客入侵帳號，可能透過精心設計的提示詞，繞過傳統HIS系統的權限查核，誘騙AI將資料庫中的敏感個資包裝成摘要輸出