

# State of the Art in Data Science Research Ethics and Law (上)

## ～臺大醫院倫理中心研究倫理研習會紀要

整理 / 臺大醫院倫理中心 王劭慈、戴君芳

### 壹、前言

2024年11月1日，臺大醫院於兒醫大樓 B1 講堂舉辦「智慧醫療與資料科學研究倫理與法律」研究倫理研習會，這是繼2023年12月8日倫理中心舉辦此主題研習會（演講紀要請見本中心第二至四期季刊）後，再次邀請多位資料及法律領域之專家學者，分別對於「資料法制的回顧與前瞻」、「健康資料研究國際趨勢」、「資料研究倫理與公民參與義務」、「健康臺灣引領大數據應用之契機與挑戰」、「生成式 AI 在醫療機構之隱私權保護及倫理議題」等議題進行研討。當天特別邀請臺大醫院高淑芬副院長開場致詞，高副院長督導研究倫理業務，長期支持倫理中心推展研究倫理業務，她在致詞中指出，資料是推動研究與創新的關鍵。無論是早期利用健保資料，還是現在的大數據與 AI，研究進步都離不開病人和受試者的信任支持。她分享了自身研究的經驗，特別強調承諾資料保護的重要性，但也坦言，隨著資料規模越來越大，研究者個人的承諾無法完全保障資料安全，必須依靠完善的法規及機制來補足。高副院長也提到，近期健保資料釋憲後引發的退出權議題，對研究者和政策制定者都提出了挑戰。如何在推動研究的同時平衡關於大數據及隱私權等倫理規範，成為未來需要解決的核心問題。感謝倫理中心的用心安排，這場活動邀請多位權威的專家分享，期盼能帶來更多啟發，為臺灣的資料使用規範上開創新局。

負責籌備規劃本次活動之倫理中心蔡甫昌主任隨後說明了本次活動的背景，他回顧臺灣健保資料庫的相關訴訟與釋憲過程，指出近期憲法法庭要求三年內完善法規，讓資料研究的規範化成為重點。特別提到，病歷資料與健保資料之間的差異，對倫理審查和研究者提出了新挑戰。他也特別感謝研究倫理委員會委員們的努力，在平時自身的工作負荷外協助審查計畫的合規性與倫理性，並幫助研究者計畫內容

更完善，並感謝研究者送案至研究倫理委員會審查，使研究倫理委員會更具價值。蔡主任表示，這場活動的主題非常貼近當前研究的需要，相信能從講者分享之內容更為精進研究、倫理及法規之發展。

## 貳、資料法制之回顧與前瞻：演講者臺北醫學大學人文暨社會科學院李崇僖院長

在過去幾年，大數據與人工智慧成為國際熱點，無論是國家政策還是產業界，都將資源、資金集中於這些前瞻技術。然而，幾位國內科技巨頭參與討論 AI 發展策略之重要場合，抱怨著目前在數據使用及著作權等相關議題，因此隨著技術進步，我們面臨一系列法制上的挑戰：從個資保護到數據使用的專利權問題，要領先國際發展，必須制定更具創新性且前瞻性的資料法制，這是一個關鍵並需突破的瓶頸。本演講將從法律史的角度，分享如何面對當前的法制瓶頸，並探索未來的突破方向。

### 資料是未來最重要的資源

2017 年，《經濟學人》以一篇專題文章揭示了資料已取代石油，成為全球最有價值的資源，這是最早提出且經常被引用之觀點。隨著資料經濟發展，不僅改變了產業發展的焦點，也讓壟斷議題更加凸顯。早在 2016 年，世界經濟論壇首次提出第四次工業革命的概念，從蒸汽機、電力、資訊，再到當前的智能革命，每一波工業革命都帶來科技、經濟與社會的深刻變化。而這些變化必然牽動法制的革新，法制若能率先調整，則可在科技發展中成為國際之領導者，此攸關未來的競爭力。

### 個人資料保護原則背景

臺灣曾在第三波產業革命的後段跟上國際核心圈，如今我們面臨第四波革命的機遇與挑戰。1969 年是一個產業發展的座標，當時因電腦出現且因其易於記錄、複製及流通等特性而對個人隱私影響有所疑慮而使著重隱私保護之強化。1980 年代，OECD（經濟合作暨發展組織）因應電腦科技普及，首次提出個人資料保護原則，以保障個人可控制自己資料可被如何蒐集及利用。1990 年代歐盟即以 OECD 原則制定法律，其內涵包括「最小流通、最小蒐集」原則。第四波工業革命典範是思想的轉換，隨著科技進步，資料的意義已從隱私威脅轉變為創新驅動的關鍵資源。在大數據和 AI 的加持下，資料分析技術得以釋放更多未知的價值，但現行法規仍停留在 80 年代個資保護的思維模式，面對新的科技發展並未以更前瞻的觀點落實法

律制度，使我們無法充分發揮資料的潛力。

國際上對於個人資料保護之規範，美國是以隱私保護為原則，主要採去識別化及保密機制，歐盟則是以自主權的概念擬定相關規範，但礙於歐盟有許多會員國，要整合建立新制度在溝通成本較高，臺灣則具有靈活性和快速溝通協調的特點，在這場法制改革中原本有望扮演領先角色，使我們能在國際間率先構建更符合時代需求的資料管理的法制規範。

### 個人資料保護法之事前同意模式

現行的個資法採事前同意模式，對資料蒐集與利用設定了特定目的與事前同意的框架。然而，此一模式在大數據的應用中顯得過於僵化。大數據異質及串流之特性，需要先蒐集大量資料，才能從中挖掘未知價值，而現行法制卻要求在蒐集前明確用途，導致許多創新研究及人才因法規風險而裹足不前。

此外，事前同意模式過於強調資料的用途限制，卻缺乏事後監督其實際使用與透明機制。此架構造成大型業者憑藉免費且完整的服務，輕易獲取大規模個資，也進一步壟斷資料資源。而當事人一旦同意，卻缺乏便捷的管道追蹤資料的實際使用情形，宜著重在資料後續如何被使用應該更有意義。

### 大數據之基本法律問題及解決方式

大數據與人工智慧的發展，依賴於資料的匯流與多元應用。然而，現行個資法要求在資料蒐集時需明確告知用途，並獲得當事人同意的機制，常常與大數據的資料驅動研究模式相衝突。資料匯流所需的廣泛應用權限，往往超出當初的同意範疇，這不僅限制了技術發展，也對投資者和人才產生了阻礙效應。

為了解決這些問題，有以下幾種解決方式：

#### 一、概括同意模式 ( general consent )

BioBank 即採用概括同意的方式，在取得當事人同意時不限定特定研究目的，而是廣泛用於生醫研究，並透過法框架規範資料的管理與使用。

#### 二、動態同意模式 ( dynamic consent )

動態同意讓當事人透過資訊平台隨時掌握資料的使用情況，並可根據需要撤回或修改同意範圍。這種方式提高了資料應用的靈活性與透明度，並能有效

回應當事人的知情權需求。

### 三、退出機制 ( opt-out ) 與例外

在回顧性研究中，例如以非為研究建置之資料庫擬使用進行分析時，不反對使用退出機制，但需著重探討退出機制的例外情況，特別是在公共利益與學術研究的背景下，如何在保護隱私與促進研究之間取得平衡。

## 健康資料的回溯性應用與法規創新的挑戰

在健保資料庫的利用爭議中，憲法法庭認為基於學術研究目的提供資料的行為合憲，但同時指出現行個資法在保護當事人權益方面尚有不足，特別是缺乏獨立專責委員會，因此，即將成立個資保護專責委員會，類似消保官處理消費問題概念，這將為資料治理提供更完善的支持。

在健康資料的研究與應用中，前瞻收案模式相對較為成熟，無論是透過人體研究法取得 IRB 批准，還是基於人體生物資料庫管理條例 ( 如 BioBank ) 進行資料蒐集，其法律基礎皆清晰穩固。尤其是 BioBank 模式，通過概括同意為資料未來的彈性應用提供了更大空間。

回溯性資料庫的法制問題仍存有諸多未解議題。例如，對於已蒐集的健康資料之利用，欠缺使用之明確規定，目前衛生福利資料管理條例草案已送交行政院審議，有關當事人得請求停止利用之相關規定及其例外適用範圍，衛生福利資料管理條例草案針對退出權的規範包括：

- 一、資料可用於目的外利用，但必須允許當事人請求全部或部分的停止使用。
- 二、退出權的例外僅限於三種情境，包括維護國家安全等情況。

上述對於公共衛生研究而言，退出權例外的狹窄範圍可能導致研究資料不足，影響醫療科技效益評估的準確性。例如，在評估細胞治療成效時，需要完整的資料庫作為對照基礎，而不完整的資料將使研究結果大打折扣，進而影響健保政策對新療法的支持與給付範圍。因此，在本法案預告期間，有提出建議可補充例外條款，允許受政府委託的公共衛生、醫療、科技效益評估研究不受退出權影響。這種設計既兼顧公益性，也為健康資料研究提供了必要的法規支持。

## AI 時代是法規競爭的時代

在大數據與人工智慧的時代，法規創新是國家競爭力的關鍵所在。臺灣若能率先在健康資料治理上實現法律突破，不僅能解決資料應用的瓶頸，也將在國際舞台上取得領導地位，這需要各界共同努力。

## 參、健康資料研究國際趨勢：演講者臺北醫學大學數據處許明暉數據長從 MIT 的 MIMIC 資料庫說起

MIMIC ( Medical Information Mart for Intensive Care ) 資料庫由美國麻省理工學院 ( MIT ) 與哈佛醫學院的 Beth Israel Deaconess Medical Center ( BIDMC ) 合作開發，收錄了超過 5 萬名重症病人的臨床數據，經去識別化處理後，開放給全球研究者申請免費使用，目前的版本已經是第四版 ( MIMIC-IV )。MIT 開發團隊不僅提供資料庫，還編寫了教科書《Secondary Analysis of Electronic Health Records》教導使用，並在全球舉辦 Datathon，推廣資料分析的應用。今年，MIMIC 的資料 Datathon 在臺灣舉辦時，數據松之父 Leo Celi 教授 ( MIT / Harvard ) 鼓勵大家使用本地之數據，因為他們在全球多個國家都已經這樣進行。我們花費許多努力才通過本國研究倫理審查，建立臺灣 MIMIC 的資料集提供活動使用，活動中並對於「為何去識別資料在臺灣應用仍有諸多限制」有些討論及反思。

## 臺灣與國際資料法規的差異

在美國根據 Common Rule 規範，符合 HIPAA 規定去識別後的健康資料不被視為人體研究 ( human subject research )，因此不需要經過 IRB ( 人體研究審查委員會 ) 審查。但在臺灣，即使是去識別的資料，仍需要符合種種法規限制，使研究者在使用資料時面臨更多挑戰。這一法規差異不僅限制了臺灣資料庫的利用率，也導致國際研究者更傾向於選擇使用其他國家的資料庫，影響了臺灣在健康資料研究中的競爭力。

## 技術進步與研究環境的反差

現代健康資料研究蓬勃發展，主要得益於三個因素：強大的硬體運算能力、成熟的資料分析工具以及更加友善的國際資料分享機制。但臺灣的法規趨勢卻朝著限制性方向發展，讓研究者的使用機會逐漸減少。這對於本地資料庫的推廣與應用形成了阻力，也可能迫使研究者轉向使用國外資料庫，對臺灣整體研究環境造成不利影響。

## 全球資料庫應用的研究優勢

國際間多種健康資料庫的開放與應用為研究者帶來了豐富的研究機會。例如美國的「All of Us」資料庫，除了包含電子病歷和基因資料，還整合了穿戴設備數據，為研究者提供全面而多元的資料支持。這些資料經過去識別處理，根據美國法規不再視為人體研究，研究者申請使用時不需經過複雜的審查程序。另外，「All of Us」要求使用機構需具備數據治理能力，包括購置保險以應對資料洩露風險等，這樣的管理方式不僅保障了資料安全，也促進了資料使用的便利性。「All of Us」並著重透明化，定期公布使用之機構，機構簽約窗口亦會收到定期報告，了解機構內研究者使用情形。

另一個全球性的資料庫為 TriNetX，是以龐大的數據量和便捷的分析平台吸引了越來越多的研究者參與。其提供訂閱制的商業模式，讓資料使用更加高效，研究者可以在極短時間內完成數據分析，顯著提升研究效率。值得注意的是，臺灣在健保資料的國際學術產出量上，去年起已被韓國超越，顯示出臺灣健康資料的利用在國際競爭中已不再具有優勢。

## 資料民主化與研究效率的平衡

除了上述資料庫，Columbia University 提出的 OHDSI 平台則致力於資料民主化，提倡使用開源工具將各地的醫院資料庫轉換為統一的資料模型，各地將個別之資料以下載之開源工具進行各自之分析，不須將資料輸出，只要將結果提供，研究者再將個別之結果彙整即可，實現分散式的數據分析。這種方式符合資料共享與隱私保護的理念，但缺點是過程繁瑣且效率較低，常需耗時一年才能完成一項跨國研究。相比之下，TriNetX 平台在效率上有明顯優勢，不同資料治理模式各有取捨。

## 臺灣的健康資料研究未來展望

去識別的健康資料若繼續被視為人體研究，將成為臺灣研究環境的一大關卡。這不僅增加研究者的操作成本，也限制了國際合作的拓展。期盼專家學者們共同推動法規改革，使臺灣的研究能與國際接軌，讓下一代研究者能夠更自由地使用本土資料，實現更多的學術突破。