國立臺灣大學法律學院法律學研究所

碩士論文

Department of Law
College of Law
National Taiwan University
Master's Thesis

個人資料去識別化規範制度之建構研析
—以歐盟 GDPR 與美國 HIPAA 隱私規則為借鑑對象

Reconstructing the Legal Framework on De-identified Personal Data: From A Comparative Study of EU GDPR and US HIPAA Privacy Rule

許雅琳

Ya-Lin Hsu

指導教授：李建良 博士

Advisor: Chien-Liang Lee, Ph.D.

中華民國 113 年 7 月

July 2024

doi:10.6342/NTU202401988

誌謝

這篇論文的完成，首要感謝李建良老師的指。論文撰寫的初期，思緒都還很混亂的階段，李老師在聽完我的論述想法後，總是為我理出清楚明確的架構，對我的論文後續的開展延伸，有莫大的啟發和幫助。此外，也非常感謝張陳弘老師和何之行老師擔任此篇論文的口試委員，給予我在架構上和內容上非常詳盡的建議及資訊的補充，使我的論文跟上我國的最新立法發展，討論的內容和面向得以更加完整。

在臺大的八年歲月中，非常感激也非常慶幸能夠結交這些陪伴我走過、將來也將一起邁進的朋友。感謝從大一大二就一起修課、考試、去畢旅，至今也持續相互扶持的芝宜、李芯、其叡和棣為；陪我度過國考時光、每天一起在圖書館念書的彥妏和姿儀；在 WTO 中心共同辛苦完成許多期編輯出刊的在善、宗霖和婉婷，也由衷地感謝在善擔任我論發和口試的紀錄，陪我順利完成畢業的最後一哩路；十分感謝宜盈、佳珺和昀瑋來參加我的論發，以及雖然認識的時間不長，但卻是我碩士最後一年中最常聊天分享、陪我找租屋處和搬家的 Mesh。

最後要感謝我的父母，總是給我無盡的關心、照顧和支援，讓獨自離家北上漂泊的我，始終有歸宿和依靠。

雅琳 謹誌於

臺北

2024/07/18

doi:10.6342/NTU202401988

摘要

去識別化在個人資料保護體制下，被期望能作為開放資料使用與隱私保障兩難下的折衝解套措施。我國個人資料保護法就有規定，公務機關和非公務機關基於公益、學術或統計研究之目的，若有將資料處理達到無從識別特定當事人者，便可在必要範圍內合法對該資料為蒐集、處理和利用，無須取得當事人之同意；我國其他的資料相關法規中，亦有允許以資料去識別化取代當事人同意之類似機制規定。然而，隨著健保資料庫爭議的發生和討論，我國現行去識別化制度規範被顯見存在許多問題和缺失。

在概念釐清上，匿名化與假名化同屬去識別化的概念範圍，前者以永久完全消除資料識別性為目的，後者則指以別名遮蔽識別資訊，達到隱藏資料與個人連結關係之效果，又加密經常與假名化一起提及，但其主要目的是防止資料外洩，與以保障隱私為目的的去識別化本質不同；在政策制度的引進和建構上，須留意到去識別化會減損資料的可利用性，且對非結構化資料能發揮效能十分有限，重新識別風險也難以完全防範，故須強化風險評估和管控措施的實施。

比較法分析發現，歐盟與美國對於去識別化的定位和規範功能的設置有顯著差異。歐盟將去識別化作為資料保護措施，是資料控管者和處理者應履行的義務，實際的放寬管制效果有限；美國法則是在規範中直接明訂去識別化方法，鼓勵規範主體依據法定方法將健康資料進行去識別化後，即可獲得獲得豁免管制，但此制度模式存在對隱私之保障過於不足的問題。

我國法規範模式在限制人民資訊自主權之下，卻幾乎僅以資料去識別化作為唯一、最終的保護措施，且規範標準不明確，實務解釋混亂，對基本權干預明顯欠缺正當性。本研究對我國法規範制度提出的改革建議是：一、以資料揭露模式作為去識別化標準的設置基準框架，使規範逐步明確；二、強化資料使用者的告知義務及

doi:10.6342/NTU202401988

對去識別化資料接收者的行為控管，以防止去識別化資料的恣意使用；三、透過法規遵循稽核和事前救濟途徑完善監督機制，確保政策規範的落實。

關鍵詞：去識別化、資訊隱私、個人資料、重新識別、當事人同意、一般資料保護規則、美國健康保險可攜與責任法隱私規則

doi:10.6342/NTU202401988

ABSTRACT

De-identification, within personal data protection frameworks, aims to balance open data usage and privacy protection. The Personal Data Protection Act allows public and private agencies to use data for public interest, academic, or statistical research without consent if the data is processed to prevent identification of individuals. Other regulations similarly permit de-identification in place of consent. However, controversies, particularly around the health insurance database, have exposed issues and deficiencies in Taiwan's de-identification regulation and system framework.

Conceptually, "anonymization" and "pseudonymization" are both forms of de-identification. Anonymization seeks to permanently eliminate data identifiability, whereas pseudonymization involves using aliases to mask identifying information, thus obscuring the link between the data and individuals. "Encryption" is often mentioned alongside pseudonymization but differs fundamentally in its aim to prevent data leakage rather than protect privacy. In practice, de-identification reduces data usability, is less effective on unstructured data, and cannot completely prevent re-identification risks, necessitating stronger risk assessment and control measures.

Comparative legal analysis shows significant differences between the EU and US approaches. The EU sees de-identification as a protective measure with limited regulatory relief. In contrast, US law explicitly outlines de-identification methods, encouraging entities to de-identify health data accordingly. Once de-identified, the data is exempt from certain regulatory controls. However, this model presents concerns regarding insufficient privacy protection.

doi:10.6342/NTU202401988

Taiwan’s regulatory model limits citizens’ information autonomy by relying almost exclusively on data de-identification as the primary protection measure. The standards are unclear, leading to inconsistent practical interpretations and lacking legitimacy in fundamental rights intervention. This study concludingly proposes the following reforms for Taiwan's regulatory system: 1. Use a data disclosure model as the basis for setting de-identification standards to gradually clarify the regulations. 2. Strengthen the notification obligations of data users and control over the behavior of recipients of de-identified data to prevent arbitrary use. 3. Enhance supervision through regulatory compliance audits and preemptive relief channels to ensure effective implementation of policy regulations.

KEYWORDS: *de-identification, data privacy, personal data, re-identification, informed consent, GDPR, HIPAA Privacy Rule*

doi:10.6342/NTU202401988

目次

doi:10.6342/NTU202401988

doi:10.6342/NTU202401988

doi:10.6342/NTU202401988

doi:10.6342/NTU202401988

圖次

doi:10.6342/NTU202401988

表次

doi:10.6342/NTU202401988

第一章 緒論

第一節 研究動機與目的

如何兼顧個人隱私保護與資料合理使用，一直以來都是資料隱私法制面對的底層課題。我國近年所發生的健康保險資料使用爭議（即「健保資料庫案」），即是公務機關在未經民眾同意下，將民眾去識別化後的健康保險資料提供給第三人建置成資料庫，開放供學術研究申請使用，以促進我國健康保險和醫療研究之發展；本案的公務機關主張，其既然已對民眾的健保資料進行去識別化處理，即應已符合個人資料保護法（以下簡稱「個資法」）合法利用個資之規定要件，無須再經當事人的同意。

確實，我國個資法以及我國其他資料相關法規，多有允許資料去識別化可取代當事人同意，作為蒐集、處理、利用和目的外使用一般和敏感性個資的獨立合法或除外事由，亦具有免除告知義務和資料銷毀義務之規範效果。從資訊隱私的憲法保障意旨來看，任何人欲對他人的資料為蒐集或使用時，原則上均應得到該他人之同意，始得為之，此為憲法對於個人資訊自決權的保障；雖然憲法賦予的保障並非絕對，但若要對資訊資決權進行限制，除須有明確之法律依據外，該法律規範在內容的設置上，仍須符合目的明確正當、手段符合比例原則，且應有法定目的外使用之明文禁止[1]。

然而，許多學說文獻指出，我國法下的去識別化概念過於空泛，且有規範標準模糊、矛盾的不明確性問題；在手段上，又幾乎僅以資料去識別化作為唯一的保護措施，並不足以正當化對於個人資訊自決權的干預；再者，目前的監管機制使資料當事人僅得透過事後的裁判救濟制度，請求法院停止或制裁資料使用者之違法行

[1] 李建良，資料流向與管制環節—個資保護 ABC，月旦法學雜誌，272 期，頁 26（2017 年）。

doi:10.6342/NTU202401988

為，欠缺事前救濟途徑之提供，此應有違反憲法正當法律程序之原則。由此足見，我國個人資料去識別化制度亟需進行通盤的檢討與重新建構。

目前相關的學說文獻，許多主要著重在去識別化認定標準的釐清，就個資法去識別化要件應如何合理適用提出解釋見解；另有許多研究文獻，是以健保資料庫案的情節脈絡為背景，討論在制度設計上，應如何對此種建置大型資料庫的資料使用行為進行合理規範；針對我國去識別化規範機制本身進行研究討論之文獻，則相對有限。

因此，本研究之目的是希望能對我國現行的個人資料去識別化制度，提出較全面的改革建議，以回應改善現行制度在規範面和監管面上，所存在的種種缺失和失衡之問題，在確保合乎憲法保障資訊隱私權的意旨之同時，亦使我國去識別化制度規範模式下的政策目的能被具體實現。

第二節 研究方法與架構

本研究以法制研究、學說整理與裁判案例研析為主要的研究方法，並輔以去識別化的技術相關文獻，了解去識別化的方法性質和運作原理，以作為法規制度研析的參考基礎。在比較法研究上，本研究係以歐盟和美國的去識別化制度，作為參考借鑑的對象，將主要引用規範主管機關所發佈的指引說明和研究報告文獻，對於兩國的規範架構和制度實施模式進行觀察比較；就我國法部分，本研究從我國去識別化的條文內容出發，透過健保資料庫案的歷審裁判、憲法訴訟案中的專家學者意見和大法官理由，以及相關的學說討論文獻，對我國去識別化制度的現況問題予以釐清，並參照歐盟和美國法的規範模式和討論，提出改革建議。

本研究的架構如下：第一章為緒論。第二章將先從技術層面介紹去識別化的概念、原理和性質，就幾個與去識別化經常混淆的名詞用語進行釐清，並將去識別化實施可能帶來的負面影響和局限性一併納入討論。第三章進到比較法研究，分析討

doi:10.6342/NTU202401988

論歐盟與美國法的規範制度建構，主要將聚焦於去識別化的認定標準、去識別化在規範架構下的功能效果、與其他資料使用合法性要件的適用關係以及監督機制，最後提出本研究對於兩制度模式的比較觀察。第四章則回歸我國法，首先對我國現行的去識別化規定進行盤點，以了解我國當前對於去識別化的規範模式，以及規定立法上本身存在的問題；再從健保資料庫案中有關去識別化爭議的討論，檢視我國裁判實務是如何就現行規範為解釋、認定和適用；後將有系統性地梳理我國去識別化規範的制度問題，並提出改革的建議方向和具體措施建構。最後，以第五章作結。

doi:10.6342/NTU202401988

第二章 去識別化規範制度建置的基礎與考量議題

第一節 去識別化在個資保護法制下的目的與功能

第一項 巨量資料的興起與效益

資通訊科技日新月異，使資料產生、蒐集和儲存的成本大幅下降，巨量資料（Big Data）[2]由此躍升為各領域產業的熱門關鍵字詞，隨著各式新興資訊應用的發展與普及、人類社會逐漸邁入數位化時代之下，資料儼然成為 21 世紀的新石油，是打造創新商業模式、產品和服務的驅動力，得提升產業界的生產力和競爭力，又在公部門領域方面，資料的蒐集和使用亦有助於提升政府機關公務運作、服務提供的品質和效率，且可供作公益、學術之研究使用等。

透過演算法（algorithm）[3]或其他資料分析技術進行資料探勘（data mining）[4]處理巨量的數據資料，從中歸納、連結和找出資料間的關聯性和規律性，可從看似雜亂無序的資料堆中發掘出具有價值的知識，用以分析評估、趨勢預測、精準決策和人工智慧（Artificial Intelligence, AI）及自動化之建置，常見的應用情境則包含個人化廣告推送、用戶喜好推薦、醫療診斷、信用評分、聊天機器人等。由

[2] 亦有翻譯為「大數據」。巨量資料的概念可回溯自 2001 年由 META Group 發布的一篇研究報告，內容指出電子商務的興起，將使可被取得處理之資料的資料量規模更大（high-volume）、來源和格式更多樣化（high-variety）且講求即時處理的高速率（high-velocity），要如何因應此三大面向的發展變化，即為資料管理（data management）將要面臨的挑戰。由此，所謂的巨量資料科技，指的就是有效且快速處理此等大量而複雜資料集的新興資料技術。Doug Laney, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, APPLICATION DELIVERY STRATEGIES (Feb. 6, 2001), https://studylib.net/doc/8647594/3d-data-management--controlling-data-volume--velocity--an....

[3] 演算法係指將輸入（input）轉化為輸出結果（output）的一連串電腦指令，可簡單理解為分析資料的一套方法流程，多數的演算法是以統計學為基礎，最常見的如線性迴歸（Linear Regression）等。然而，沒有任何一個演算法得以滿足於所有的分析運算需求，每個演算法的適用性則是取決於資料呈現的格式型態以及應用之目的。EUR. UNION AGENCY FOR FUNDAMENTAL RTS., #BIGDATA: DISCRIMINATION IN DATA-SUPPORTED DECISION MAKING 7 (May 30, 2018), https://fra.europa.eu/sites/default/files/fra_uploads/fra-201 8-focus-big-data_en.pdf.

[4] 資料探勘的過程主要可分為四個步驟—資料蒐集、資料處理、資料分析與採取行動，於資料分析完成後，資料間的規律性和特徵關聯則是以建立模型（model）的方式被描述和記錄，當需要為應用時，即可將輸入的新資料套入先前所建立的模型進行處理，最終以得出推論的結果。陳宗和、蔣彥亭、宋瑞豐、陳姿蓉、陳姿佑，資訊科技全一冊，頁 198-199，臺北：科有（2021 年）。

doi:10.6342/NTU202401988

此，資料可發揮的效用無可限量，各產業、機構單位和政府部門當今無不積極投入資料的蒐集和取得，透過資料的運用創造出最大的價值和商機。

第二項 巨量資料的運作與個資保護之衝突

雖然巨量資料的發展和應用可滿足人們眾多的需求，正在為人類社會帶來許多的變革，但卻同時也引發許多隱私侵害與妨害個人自主之隱憂。隨著隱私保護意識的抬頭，各國紛紛制訂個人資料或隱私保護法規，要求資料的使用應遵循相關個資保護原則、賦予當事人對資料的自主決定和控制權利，以及課予資料使用者應盡相關的告知和保護義務。然而，當今個人資料隱私保護規範的制度設計，卻與巨量資料的資料運用原理有很大的衝突，使得巨量資料的推行和發展遭遇很大的阻礙。

第一款 資料最小化蒐集原則

現行個人資料保護法規要求資料使用者應特定其蒐集資料之目的[5]，且應盡可能最小化資料之蒐集，僅得限於原始目的所必要且適當之範圍內，例如：歐盟一般資料保護規則（General Data Protection Regulation, GDPR）第 25 條更強調，資料控管者（controller）[6]應於產品或服務的設計階段即採取必要措施，以使後續的資料使用行為皆在遵循資料最少化等資料保護原則下進行，且亦須確保產品或服務在系統的預設模式下，僅會就特定目的適當必要範圍內的個人資料進行使用。

[5] Regulation 2016/679 of Apr. 27, 2016, Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC [hereinafter GDPR], art. 5(1)(b), 2016 O.J. (L119) 1, 35 (EU). 我國個人資料保護法（以下簡稱「個資法」）第 15 條和第 19 條第 1 項亦有規定，公務機關和非公務機關蒐集或處理個人資料應有特定之目的。

[6] 控管者是指主要決定個人資料處理（processing）之目的和方法之自然人、法人或其他公私單位機構；又 GDPR 並未如我國個資法，將資料之使用區分為蒐集、處理和利用三種行為，而是統稱為資料之處理（processing），故此處之控管者將包含資料蒐集者、處理者或利用者。控管者之完整定義，詳見 GDPR art. 5(1)(b).

doi:10.6342/NTU202401988

然而，這種先確立目的再開始使用資料的制度模式，實與巨量資料的運作原理大相違背。資料探勘的原理，本即是為大量的資料中發掘出有用的資訊，是一種探索知識的過程與方法，故實務上的情況是往往是盡可能先蒐集愈多的資料，以利未來尚未可知的應用目的[7]，例如商業上將消費紀錄或用戶足跡資料，用於往後新產品和服務的開發等。因此，資料最小化蒐集原則在根本上即與巨量資料和新興資訊科技之原理和目的有所牴觸。

第二款 目的限制原則

規範亦要求，對個資的處理和利用行為，除非符合少數例外事由，否則不得脫離原始資料蒐集目的之必要範圍[8]。然而，巨量資料科技卻十分著重資料的二次利用（secondary use）價值，透過將原始資料與其它或後續取得之資料，不斷進行比對分析，以發掘出資料的隱藏效用，用於其它計劃目的之推行和支援。由於資料未來得再發揮的用途，往往也並非最初蒐集時就得預見，且受限於資料使用目的之設定須特定（specific）、明確（explicit）之規範要求，資料使用者亦難以先劃定一個廣泛、模糊、尚不必要的目的範圍，以避開目的限制原則的約束[9]。更遑論當今的資料處理、分析程序均已逐漸邁向高度自動化，這使得要確保整個資料處理、分析過程中，對於個人資料的使用都維持在最初蒐集目的範圍內，會變得更加困難、所費成本更高、甚至實際上可能無法達成[10]。

[7] 張陳弘，新興科技下的資訊隱私保護：「告知後同意原則」的侷限性與修正方法之提出，臺大法學論叢，47 卷 1 期，頁 206（2018 年）。
[8] GDPR art. 5(1)(b). 個資法第 5 條；個資法第 16 條和第 20 條第 1 項亦規定，公務機關和非公務機關對於個人資料之利用，須與蒐集目的相符合。
[9] Tal Zarsky, *Incompatible: The GDPR in the Age of Big Data*, 47 SETON HALL L. REV. 995, 1006 (2017).
[10] *Id.*

doi:10.6342/NTU202401988

第三款 事前告知後同意原則

資料使用者在對個人資料為蒐集前，須以符合規定要求之方式，向當事人充分告知使用資料之對象、蒐集的目的、欲蒐集的資料範圍內容、使用資料之方式，以及當事人得行使之相關權益[11]；又除非符合少數例外事由，否則須在蒐集時取得當事人個別之明示同意[12]，該資料使用行為始為合法。此外，若欲對所蒐集或自第三人處得取得之個人資料進行目的外之利用，原則上仍須再行逐一向當事人為告知，並取得當事人之個別同意[13]。

且為避免當事人同意之規定流於形式，現行個人資料保護法規多對於同意的有效性，設有嚴格的要件和認定標準，以確保當事人是在充分理解資訊的情況下，為真實且自由的意思表示，例如 GDPR 第 4 條第 11 款即規定，有效之當事人同意應具備四個要素，即須為當事人自主給予（freely given）、針對特定目的（specific）、在受告知（informed）之情況下為之，且須以聲明或其他清楚之行動（a clear affirmative action）表達其同意，不得是模糊的意思表示[14]，並要求資料使用者應就當事人有為同意之事實，負舉證責任[15]。

然而，這種由保障當事人自主決定權，所衍生出的各種資訊告知和取得同意的規範措施，都與資訊科技的演進和功能性有所抵觸。首先，現今愈來愈多的資料蒐集方法，是使用新興資訊採集科技，像是透過應用軟體、感測器等各式系統裝置，

---

[11] GDPR arts. 12-14. 我國個資法下之告知義務，又區分為直接蒐集（第 8 條）與間接蒐集（第 9 條）兩種類型，前者為向當事人蒐集其本人之個人資料時，須向該當事人為告知之情形，後者則是指向當事人以外之第三人蒐集個人資料時，於處理和利用該資料前，須向該當事人另為告知之情形。

[12] GDPR art. 6(1). 我國個資法第 6 條規定，除非取得當事人之書面同意或具備其他例外事由，否則原則上不得就病歷、醫療、基因、性生活、健康檢查或犯罪前科之敏感性個人資料，為蒐集、處理或利用之行為；第 15 條、第 16 條、第 19 條和第 20 條則規定使用一般性個人資料之合法要件。

[13] GDPR art. 6(4). 我國個資法第 16 條和第 20 條第 1 項規定參照。

[14] WP29 亦有針對有效同意發布相關指引，詳細說明此四項要素的具體內涵，並以實際案例示範個別要素的判斷與適用，詳見 Article 29 Data Protection Working Party [WP29], *Guidelines on Consent Under Regulation 2016/679*, at 7 (¶¶ 11-12) (May 4, 2020), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf。

[15] GDPR art. 7(1). 我國個資法第 7 條第 4 項規定參照。

doi:10.6342/NTU202401988

快速且大量地獲得用戶、使用者或其他不特定個人之特徵、行為、通訊和交易紀錄等。

這當中會面臨到的困境是，就大規模的資料蒐集逐一告知並取得同意，在執行面上近乎不可能達成，可能導致許多新興科技應用無從或難以實現，像是臉部辨識、自動駕駛等當今許多的科技應用情境，從取得（輸入）資料開始，一直到資料處理分析最後系統產出反應結果（輸出）等整個流程，都是或甚至必須在非常短的時間內完成，若加上為告知後同意的等待時間，將大幅減損這些科技帶來的便利性，甚至可能引發安全性的問題。

又或者是，為了維持資料使用行為的合法性，只能後續花費大量的成本對資料進行加工處理，例如：為定期更新地圖服務上的街景與當地實際情景一致，Google 須頻繁地在世界各處執行街景紀錄任務，過程中捕捉到的街景圖像多常會包含私人住宅門牌、汽車車盤號碼，而涉及個人資料之使用，殊難想像 Google 在執行街景紀錄任務時，有辦法逐一向被記錄到當事人為告知並取得同意，故須要事後再對所記錄到的影像中涉及個人隱私資訊的部分，進行大量的模糊化處理才能合法使用。

此外，當每次欲對於先前蒐集或取得之個人資料的進行二次利用，都必須再聯繫找回所有當初提供資料的當事人，逐一對其重為告知並取得同意，如此龐大而繁瑣的遵循作業，往往對於科技發展研究的推行造成很大的阻礙。

第三項 去識別化作為隱私保護與資料開放使用的折衝手段

第一款 個人資料保護法制的基礎概念

識別性的概念是界定個人資料，進而決定個人資料保護法規適用範圍的關鍵因素。我國個資法第 2 條第 1 款將個人資料定義為「自然人之姓名、出生年月日、國民身分證統一編號、護照號碼、特徵、指紋、婚姻、家庭、教育、職業、病歷、醫療、基因、性生活、健康檢查、犯罪前科、聯絡方式、財務情況、社會活動及其他

doi:10.6342/NTU202401988

得以直接或間接方式識別該個人之資料」；又個資法施行細則第 3 條進一步說明，「以間接方式識別」是指保有資料之公務機關或非公務機關，單憑資料本身無法直接識別，須與其他資料對照、組合、連結等，始能識別出特定個人。

歐盟 1995 年資料保護指令（Directive 95/46/EC）第 2(a)條與後續發布的 GDPR 第 4(1)條則將個人資料定義為「有關於已被識別（identified）或可被識別（identifiable）之自然人的任何資訊」。對此，歐盟個人資料保護工作小組（Article 29 Data Protection Working Party，以下簡稱「WP29」）[16]於 2007 年針對 Directive 95/46/EC 所發佈的意見書（以下簡稱「2007 年意見書」）中進一步說明，「已被識別」是指某一個人已可被與群體中的其他個人區隔、辨別開來，而「可被識別」則是著重於辨別可能性的判斷，也就是得否透過合理、可行的方式，直接或間接將資料內容所關聯的該個人辨別出來之情形[17]；又英國資訊委員辦公室（Information Commissioner's Office, ICO）於 2017 年發佈的 GDPR 指引（UK GDPR Guidance and Resources）中，則將識別（identifiability）概念的解釋為，當某一個人得被與其他人辨別（distinguished）開來，亦或者是被標定（singled out）出來時，該個人即是處於已被識別或可被識別之狀態[18]。

有國外學者便指出，現行之個人資料保護法規乃是奠基在「可識別個人之資訊（Personally Identifiable Information, PII）」與「不可識別個人之資訊（non-PII）」之區分上[19]，這當中隱含的前提假設是，無涉及 PII 的蒐集、處理和利用行為，並不會

[16] 歐盟個人資料保護工作小組（Article 29 Data Protection Working Party, WP 29）是歐盟 1995 年資料保護指令下建立的諮詢機關，於 2018 年 5 月 25 日 GDPR 正式實施後，即被歐盟個資保護委員會（European Data Protection Board, EDPB）所取代。*Legacy: Art. 29 Working Party*, EDPB, https://www.edpb.europa.eu/about-edpb/who-we-are/legacy-art-29-working-party_en (last visited July 15, 2024).
[17] WP29, *Opinion 4/2007 on the Concept of Personal Data*, at 12 (June 20, 2007), http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.
[18] *What Are Identifiers and Related Factors?*, INFO. COMM'R OFF. [ICO], https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/personal-information-what-is-it/what-is-personal-data/what-are-identifiers-and-related-factors/ (last visited July 15, 2024).
[19] *See e.g.*, Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and A New Concept of Personally Identifiable Information*, 86 N.Y. UNIV. L. REV. 1814, 1814 (2011).

doi:10.6342/NTU202401988

造成對隱私的侵害結果，又或者是說對於隱私的侵害風險低至可忽略之程度，因此，個人資料保護法規僅就 PII 的運用行為進行規範，將 non-PII 排除於法規的適用範圍外[20]。

對此，我國學者亦有認為，我國個資法亦是採取二分法的規範架構，也就是只保障具識別性個資，不保障不具識別性之資料[21]；實務上，法務部及國家發展委會（以下簡稱「國發會」）的解釋函令見解，也多次呼應了此一觀察，例如法務部就曾表示[22]汽車原廠將車輛維修紀錄表進行去識別化處理，使其已無從識別出特定個人，則汽車原廠之後再將該車輛維修紀錄表提供給新車主的對外揭露行為，自無個資法之適用；但倘若該車輛維修紀錄表仍可直接或間接識別特定個人，因其仍屬於個人資料之範疇，故汽車原廠後續利用該資料時，便仍須遵照個資法的規範要求；法務部及國發會對於其他有關公務機關運用資料之法規適用問題，亦多次說明系爭資料得否直接或間接識別特定當事人，將決定各機關單位運用資料之行為是否受個資法之規範[23]。

是故，現行規範對於個人資料範圍的認定，即等同是在對資料的識別性風險設定規範所容許的限度門檻。GDPR 立法理由第 26 點即有表示，在認定個人是否可被識別時，是以資料控管者或其他之人是否可能透過「所有可能合理使用之方法（all the means reasonably likely to be used）」將該個人識別出來，也就是在個人資料的範

[20] *Id.* at 1816.
[21] 張陳弘，個人資料之認定－個人資料保護法適用之啟動閥，法令月刊，67 卷 5 期，頁 95（2016 年）。
[22] 法務部(105)法律決字第 10503500370 號函。
[23] 例如國家發展委員會(111)發法字第 1110000748 號函（節錄）：「關於某特定門牌地址內之戶數及設籍人數，縱使自該等資料本身不能直接識別特定個人，惟仍應視個案情況觀察，若某特定門牌地址仍能透過與其他資料對照、組合、連結等方式，得以間接識別特定現生存之自然人者，則屬於上開個資法第 2 條第 1 款所定之個人資料範圍，有該法之適用」；法務部(103)法律字第 10303507480 號函（節錄）：「故旨揭悠遊卡刷卡交易資料提供後，如無法以直接或間接方式識別特定個人者，尚非本法所稱之個人資料，即無適用本法餘地。惟蒐集者如將前開資料與其他資料對照、組合、連結而得識別特定個人，則屬本法所稱之個人資料，有本法適用」。

doi:10.6342/NTU202401988

圍界定上，並非是採取絕對主義標準[24]，只要有假設上的可能性（hypothetical possibility）[25]可得識別出特定個人，該資料就會被劃歸為個人資料成為規範標的，而是係將識別可能性限縮於「所有可能合理使用之方法」之具體判斷，劃定資料識別風險的門檻以區分否屬於個人資料[26]。

就我國法而言，從個資法第2條第1款和個資法施行細則第3條之規範文字中，亦得看出我國法是以資料的識別性是否已達到「得以間接方式識別特定個人」之程度，來界定該資料是否屬於個人資料。我國法規在法條文字上並未明確說明「間接識別」究竟應如何判斷，法務部於民國100年就記名悠遊卡號是否為個人資料所作成之函釋中則有表示，資料是否得以直接或間接方式識別該個人，應視對於持卡人以外之第三人而言，查詢特定持卡人之卡號「是否有所困難或需耗費過鉅」[27]；在裁判實務上，曾有法院見解表示，間接識別之情形不得毫無限制，解釋上應以該資料得透過「合理、可能且容易」的方式，與其他資料組合、對照、連結，即得識別出特定個人之情形為限[28]。

---

[24] 就識別性的判斷，我國有學者提出絕對主義或相對主義之寬嚴標準區分，絕對主義要求識別性的判斷須窮盡理論上所有的可能性和機會，只要世界上有任何一個人得以任何方式從資料中辨識出特定個人，識別性的要件即滿足；相對之下，相對主義就資料識別性的考量範圍，則是限縮於有實際可能性的情形，也就是說，只是有理論上的高度可能性，尚不足以構成識別性之要件，而是須達到現實中有合理機會（reasonable chances）識別出特定個人之程度，詳見江耀國、黃子宴，個人資料的概念與匿名化：一個認識論的觀點，東海大學法學研究，58期，頁8-9（2019年）。

[25] *Opinion 4/2007 on the Concept of Personal Data*, *supra* note 17, at 15.

[26] 關於GDPR下資料識別性的判斷標準，詳見後述第三章、第一節、第二項。

[27] 法務部(100)法律字第 0999051927 號函（節錄）：「就持卡人以外之第三人而言，倘該個人資料係屬查詢有困難或需耗費過鉅始能足以識別特定個人者，客觀上即屬無法識別之個人資料；惟就貴公司而言，因技術上仍得透過比對記名卡卡號與貴公司內部資料庫系統而得知特定持卡人之個人資料，非屬查詢上有困難或耗費過鉅，該記名卡卡號即屬得以間接方式識別之個人資料」。

[28] 臺北地方法院103年度訴字第255號（節錄）：「足以該當個資法第2條第1款間接識別之個人資料，並非毫無限制，倘未予適當限制，則舉目所見之資料盡屬個資法所保護之個資，則因所有資料均可能經由反覆多層遞次對照、組合、連結而耗費鉅大時間、費用、人力後，或可識別某一特定人，然此洵非個資法為促進個人資料之合理利用之立法目的，且不當壓縮言論自由、資訊自由及公共利益，是解釋個資法第2條第1款所定間接識別之個人資料，應以「合理、可能、容易之方法」為限。」臺北地方法院103年度訴字第212號亦有類似之理由說明。

doi:10.6342/NTU202401988

由此可見，我國實務見解大抵上也是將規範容許的資料的識別性風險範門檻，界定在否達到得合理、容易且在無須耗費過多成本下識別特定個人之程度，以作為是否屬於個人資料之區分[29]。

第二款 以去識別化作為個人資料保護法規適用範圍之調控

科技社會發展與個資保護規範間的衝突，一直以來都是各國主管機關執行監管上的痛點，在二分法之資料保護法制下，去識別化措施便被期待可作為兼顧隱私保護與促進資料有效利用的解套措施[30]。此立論在於，透過將個人資料進行去識別化處理，將資料的識別性風險減低至規範要求的門檻標準以下，在識別性大幅降低後，該資料即已非屬規範定義範圍內之個人資料。

由此，對政策制定者而言，透過鼓勵去識別化的實施確保個人隱私受到一定程度的保障，即可適度鬆綁對於資料運用行為的管制，以貫徹平衡兼顧基本權保障與促進個資的自由流通之立法意旨[31]；另一方面，這對資料運用者而言亦表示，只要先將個人資料進行去識別化處理，之後再對該資料為其他目的之使用或傳輸給第三人等後續處理、利用行為時，則可受到較少、或甚至可不再受到法規的限制。

下圖一係出自英國 ICO 於 2021 年所發布之「匿名化、假名化和隱私強化技術指引草稿之第二章：如何確保匿名化之有效性」文件，藉由該圖的示意內容，得以更清楚地了解去識別化、資料識別性風險與個人資料保護法規二分法體系之關連。圖一中的上方部分標示資料的識別性程度（由右至左遞增）；中間區塊表示資料保護法，以得否「透過可合理使用之方法（means reasonably likely to be used）識別出特定個人」作為判斷標準，認定該資料的識別風險是否達到被界定為個人資料之門檻，

[29] 此處須強調的是，由於我國個資法及個資法施行細則對於個人資料識別性要件的規定過於簡疏，導致實務上對於「得以直接或間接方式識別」究竟應如何解釋適用之見解十分混亂，尚未有形成一致之認定標準和方式，相關函釋與法院裁判之整理分析，詳見張陳弘（註 21），頁 78-88。

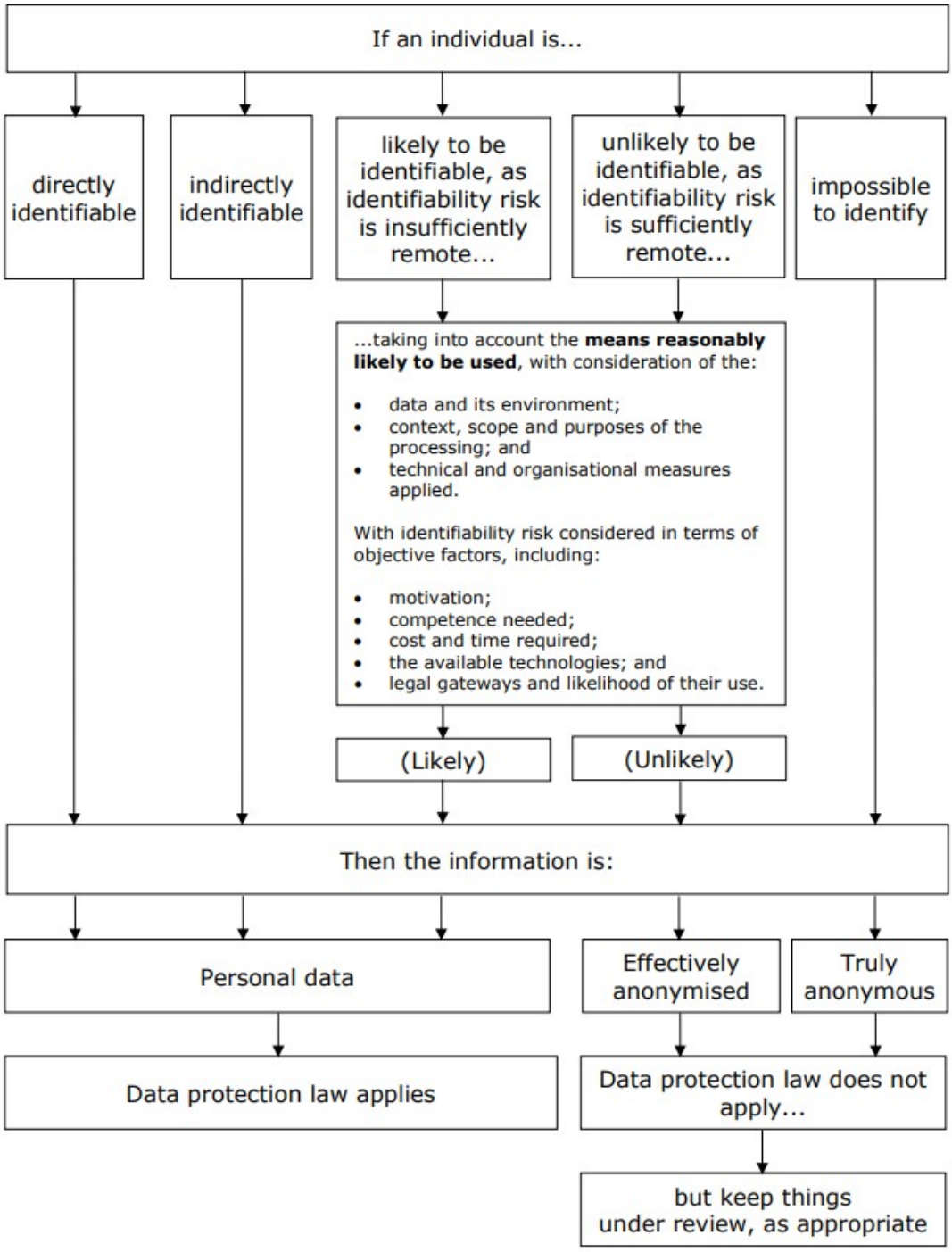
[30] Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1738 (2010).

[31] GDPR art. 1; GDPR r. 3-4. 我國個資法第 1 條規定參照。

doi:10.6342/NTU202401988

此為資料識別性風險在二分法規範架構下之概念關係；去識別化於此的功能則是，將個人資料的識別程度降低至「無法以得合理使用之方法進行識別」之程度，如此即表示該資料的識別性風險已低至足夠微小（sufficiently remote）、可為規範所容忍之程度，而可脫離資料保護法規之適用範圍。

doi:10.6342/NTU202401988



圖一：個人資料範圍之界定與資料識別性程度之關係

資料來源：ICO（2021）[32]

[32] ICO, *Draft Anonymisation, Pseudonymisation and Privacy Enhancing Technologies Guidance, Chapter 2: How Do We Ensure Anonymisation Is Effective?* [hereinafter *Chapter 2 of Draft Guidance*] 9 (Oct., 2021), https://ico.org.uk/media/about-the-ico/documents/4018606/chapter-2-anonymisa tion-draft.pdf.

doi:10.6342/NTU202401988

這種將去識別化作為折衝解套手段的概念，已普遍可見於現行許多的個人資料保護法規的制度設計中，像是 GDPR 立法理由第 26 點即有說明，GDPR 不適用於匿名資訊（anonymous information），故只要將個人資料進行有效之匿名化（effective anonymisation）[33]處理，使資料主體（data subject）不再可得識別時，該資料即轉變成為匿名資訊，不再屬於 GDPR 之適用範圍[34]。

美國健康保險可攜與責任法（Health Insurance Portability and Accountability Act of 1996, HIPAA）亦訂有健康資料之去識別化標準和實施措施之具體規範[35]，規定受保護之健康資料（protected health information）已依專家認定法（Expert Determination）[36]或安全港模式（Safe Harbor）[37]之法定方式進行去識別化後，即非屬於該規範下個人識別性的健康資料（individually identifiable health information）[38]，可被 HIPAA 隱私規則所豁免。

在我國法下，法務部亦表示公務機關或非公務機關如將所保有之個人資料進行去識別化，使之無從直接或間接識別該特定個人者，該資料即非屬個人資料，自無我國個資法之適用[39]；此外，我國最高行政法院裁判見解亦明確表明，若已經去識別

---

[33] WP29, *Opinion 05/2014 on Anonymisation Techniques*, at 9 (Apr. 10, 2014), https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

[34] GDPR r.26 "The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable."

[35] 45 CFR §164.514(a)(b).

[36] 45 CFR §164.514(b)(1).

[37] 45 CFR §164.514(b)(2)(i).

[38] 45 CFR §164.502(d)(2) "Health information that meets the standard and implementation specifications for de-identification under § 164.514(a) and (b) is considered not to be individually identifiable health information, i.e., de-identified."

[39] 例如法務部(107)法律字第 10703513050 號函（節錄）：「是以，非公務機關如將保有之個人資料，運用各種技術予以去識別化，而依其呈現方式已無從直接或間接識別該特定個人者，即非屬個人資料，自非個資法之適用範圍」；又如法務部(103)法律字第 10303510410 號函（節錄）：「如交通管理機關將 ETC 資料以代碼、匿名、隱藏部分資料或其他使他人無從辨識該特定個人方式，進行去識別化處理後，已無從識別特定當事人，即無個人資料保護法適用」；內政部亦表示過相同見解，參見內政部(106)內授營建管字第 1060809377 號函（節錄）：「關於公寓大廈區分所有權人會議召集人之書面推選文件，倘將其中推選人之個人資料，運用各種技術予以去識別化（例如遮蔽推選人之姓名、住址及簽名等所有個人資料），而依其呈現方式已無從直接或間接識別該特定個人者，該去識別化之部分即非屬個人資料，自非個資法之適用範圍」。

doi:10.6342/NTU202401988

化作業去除資料內容的個人屬性，該資料則不再屬於受個資法規範之個人資料，後續就該資料的處理和利用也就無須再受到個資法之規制[40]。

第二節 去識別化的原理概念與技術方法

第一項 去識別化的定義與內涵

資料與個人之間的連結關係，主要來自於資料內容中對於個人的各種描述與紀錄，像是：網路購物平台的會員資料通常包含姓名、年齡、出生年月日、地址、手機電話等欄位；又或者是商家發放的消費者意見調查表中，可能會詢問到客戶的職業、收入、已婚或未婚、居住地區、消費目的、平均消費次數或購買時優先的考量因素等。

這些資訊紀錄都和個人具有不同程度上的密切、連結關係，有助於連結、標定、辨別出特定個人的資訊，因而又被稱作識別符號（identifiers）或識別資訊（identifying data）[41]。有些識別符號具有個人專屬性，例如：姓名、身分證字號或私人電子郵件等，單憑此一項資訊內容本身，即有辦法辨別出該資訊指涉的具體對象是何人，此類得直接識別出特定個人的內容要素，即稱為直接識別符號（direct identifiers）或直接識別資訊（directly identifying data）[42]；又即便欠缺這些直接識別

[40] 最高行政法院106年度判字第54號（節錄）：「至於對資料之收受者而言，首應探究，其收受之資料是否還屬「個人資料」。而其判準則為資料內容之「去識別化」作業是否已經完成。如果該資料內容已完成「去識別化」作業，「個人」屬性即已消失，不能再視之為新個資法所規範之「個人資料」，而該資料收受者對資料之後續處理及利用，亦不受新個資法之規範。但若未進行「去識別化」作業，或作業不嚴謹，未達成「去識別化」作業應有之實證效用（即徹底切斷資料內容與特定主體間之連結），該收受之資料仍具「個人資料」屬性時，則應依其收受目的是為「處理」或「利用」而受新個資法對應法規範之規制」。

[41] 例如在 GDPR 在個人資料定義中，便將這些與個人具有密切連接關係的資訊要素，稱作識別符號（identifiers），其中包含（但不限於）姓名、身分證字號、位置資料、網路識別碼或有關個人之身體、生理、基因、心理、經濟、文化或社會認同因素等。GDPR art. 4(1) s.2. 原文為：“an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.”

[42] ISO 25237:2017 cl. 3.21 (directly identifying data) “*data* (3.14) that directly identifies a single individual”; ISO 25237:2017 cl. 3.21 (n.1) “Direct identifiers are those data that can be used to identify a person without additional information or with cross-linking through other information that is in the public domain.” *ISO*

doi:10.6342/NTU202401988

符號，只要集結夠多其他次要的識別符號資訊內容，透過逐漸縮小特定個人所屬的年齡段、職業種類、居住地等屬性範圍，最終仍有辦法成功將特定個人個人從資料集的群體中辨認出來[43]，這類不具有直接識別性、但仍對於識別出特定個人有一定幫助的資訊，則稱為間接識別符號（indirect identifiers）、間接識別資訊（indirectly identifying data）或類識別符號（quasi-identifiers）[44]。

由此，去識別化（de-identification）之功能目的，簡單而言就是在於消除、減低資料與特定個人（資料主體）之間的連接關係。國際標準化組織（International Organization for Standardization, ISO）首於 2008 年制定、後於 2017 年經修正提出之 ISO 25237「健康資訊學—假名化（Health informatics — Pseudonymization）」，係將去識別化定義為「減少識別資料與資料主體間之關聯的任何過程之一般術語[45]」；2018 年之 ISO/IEC 20889「隱私增強資料去識別化術語與技術分類（Privacy enhancing data de-identification terminology and classification of techniques）」，亦強調去識別化技術之實施，是確保 PII 運用行為得持續符合 ISO/IEC 29100 等隱私標準的重要措施，根據 ISO/IEC 20889 之定義，去識別化技術是指「一種將資料集（dataset）進行轉化的方法，目的是為減低資訊可得聯繫至個別資料主體的連結程度[46]」。

美國國家標準暨技術研究院（National Institute of Standards and Technology, NIST）於 2015 年作成有關 PII 去識別化技術之研究報告，報告內容包含對於去識別化技術的功能原理說明、去識別化技術的侷限性與挑戰，以及重新識別（re-

25237:2017(en) Health informatics — Pseudonymization [hereinafter *ISO 25237:2017*], ISO, https://www.iso.org/obp/ui/#iso:std:iso:25237:ed-1:v1:en (last visited July 15, 2024).

[43] *Opinion 4/2007 on the Concept of Personal Data*, *supra* note 17, at 12-13.

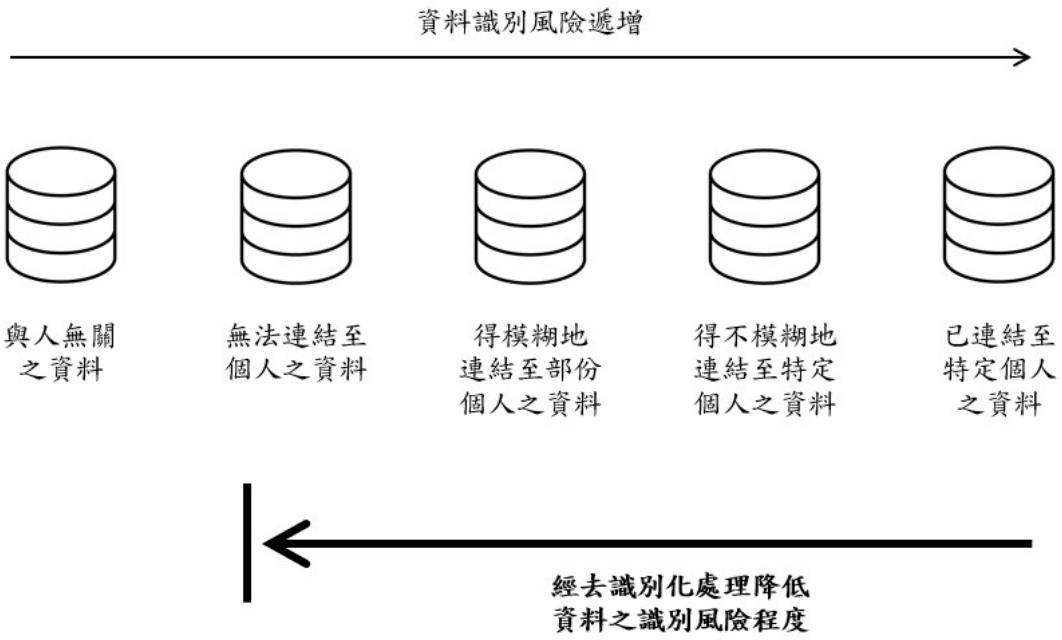
[44] ISO 25237:2017 cl. 3.28 "*data* (3.14) that can identify a single person only when used together with other indirectly identifying data", cl 3.28 (n.1) "Indirect identifiers can reduce the population to which the person belongs, possibly down to one if used in combination." *ISO 25237:2017*, *supra* note 42.

[45] ISO 25237:2017 cl.3.20 (de-identification): "general term for any process of reducing the association between a set of *identifying data* (3.14) and the *data subject* (3.18)." *Id.*

[46] ISO/IEC 20889:2018 cl. 3.7 (de-identification technique): "method for transforming a dataset (3.5) with the objective of reducing the extent to which information is able to be associated with individual data principals (3.4)." *ISO/IEC 20889:2018(en)* Privacy *Enhancing Data De-Identification Terminology and Classification of Techniques*, ISO, https://www.iso.org/obp/ui/#iso:std:iso-iec:20889:ed-1:v1:en (last visited July 15, 2024).

doi:10.6342/NTU202401988

identification）問題之討論等。該研究報告之內容即說明，去識別化是泛指得用以在不同程度上，去除資料中的個人識別資訊之任何工具、方法或演算法[47]；另從識別風險程度的角度切入來看，資料的識別性並非全有或全無，而是有如光譜一般有不同程度的分別（如下方圖二所示）。圖中最左側是與人無關的資料類型，例如：地形資料和降雨數據等，這類資料並不具有任何的識別風險；而圖一愈往右側，代表資料類型的識別風險便逐漸遞增，最右側的資料則屬於得直接連結至特定個人的種類，去識別化技術的作用，即是降低個人資料的識別程度，也就是圖一中的箭頭逐漸向左挪移的概念[48]。



圖二：資料識別性光譜

資料來源：Simson L. Garfinkel（2015）[49]，本研究翻譯繪製

[47] Simson L. Garfinkel, *NISTIR 8053: De-Identification of Personal Information*, NIST 1 (Oct., 2015), https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf. 原文為：“De-identification is a tool that organizations can use to remove personal information from data that they collect, use, archive, and share with other organizations. . . De-identification is not a single technique, but a collection of approaches, algorithms, and tools that can be applied to different kinds of data with differing levels of effectiveness.”

[48] *Id.* at 5-6.

[49] *Id.* at 6.

doi:10.6342/NTU202401988

第一款 去識別化、匿名化與假名化之關聯

去識別化、匿名化和假名化之用語，無論是在技術層面或是規範層面上都經常被並列提及，而有被混淆使用之問題。本研究以下將先從技術面的角度，對此三個術語名詞進行界定和區分，以作為在後續章節中，與我國和比較法上去識別化標準的設置和認定方式相互釐清參照之基礎。

首先，從ISO標準的術語定義來看，ISO 25237將去識別化、匿名化、及假名化分別定義為：匿名化（anonymization）指「將個人資料進行不可逆轉化的過程，使資料主體不再能被資料控管者單獨或與其他方合作下，直接或間接地被識別出來[50]」；去識別化（de-identification）指「減少識別資料與資料主體間之關聯的任何過程之一般術語[51]」；假名化（pseudonymization）是「一種特定的去識別化類型，指在消除與資料主體間之關聯的同時，又在資料主體有關的特定特徵與一個或多個假名之間建立關聯[52]」；ISO/IEC 29100則更具體地說明，假名化是指「將PII中的識別資訊以假名或別名等（alias）取代的一種處理方式[53]」，並於附註中說明強調，假名化後之資料仍無法排除有被資料處理者以外之第三人，得以透過假名或其他資訊識別出資料主體身份之可能性[54]。

由此初步觀察，去識別化的概念範圍相較於匿名化和假名化應最為寬廣，泛指透過消除資料內容與特定個人之關聯性，以降低資料識別風險程度之資料處理技

[50] ISO 25237:2017 cl.3.2 (anonymization): "process by which *personal data* (3.37) is irreversibly altered in such a way that a data subject can no longer be identified directly or indirectly, either by the data controller alone or in collaboration with any other party." *ISO 25237:2017*, *supra* note 42.
[51] ISO 25237:2017 cl.3.20 (de-identification): "general term for any process of reducing the association between a set of *identifying data* (3.14) and the *data subject* (3.18)." *Id.*
[52] ISO 25237:2017 cl.3.42 (pseudonymization): "particular type of *de-identification* (3.20) that both removes the association with a *data subject* (3.18) and adds an association between a particular set of characteristics relating to the data subject and one or more *pseudonyms* (3.43)" *Id.*
[53] ISO/IEC 29100:2011 cl. 3.22 (pseudonymization): "process applied to *personally identifiable information (PII)* (3.7) which replaces identifying information with an alias." *ISO/IEC 29100:2024(en) Information Technology — Security Techniques — Privacy Framework*, ISO, https://www.iso.org/obp/ui/en/#iso:std:iso-iec:29100:ed-2:v1:en (last visited July 15, 2024).
[54] *Id.* at cl. 3.22 (n. 2).

doi:10.6342/NTU202401988

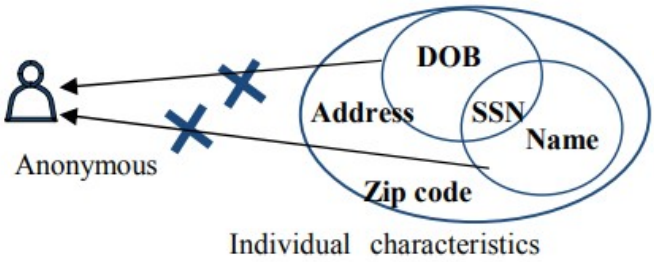
術。匿名化與去識別化的區分乃是在於識別性的降低程度，匿名化是特指資料經去識別化後，資料的識別性已經被完全消除且無可再回復，無法再重新識別出特定個人之情形，這也就是 ISO 標準對於匿名化之定義中所稱的「不可逆轉化」；相較於匿名化是一種絕對性、無法再被重新識別的概念標準，去識別化的效用強度則涵括可被重新識別和無法被重新識別之所有程度範圍[55]。假名化則是指資料去識別化的其中一種作法，專指用代碼或別名取代資料內容中的識別性資訊，達成降低資料和特定個人間之關聯性的效果，但經這種方法處理後的資料仍舊有重新連結或推認出該特定個人之可能性，故無法使資料達到匿名化的程度。

關於匿名化和假名化的差異，可進一步透過資訊、電腦科學等技術領域之文獻進行理解。匿名化技術是「去除」特定個人和資料內容的連結關係，最直接的作法就是將姓名、出生日期、地址、住址郵遞區號等識別資訊自資料中刪除[56]，以達到無法與資料中的其他主體相互辨別的狀態[57]。假名化技術則是以假名遮蔽資料主體的識別資訊，使資料得以在不知悉所規屬之資料主體的狀態下進行運用[58]，這與匿名化的差別在於，假名化並未將資料中的識別資訊永久刪除，而是將識別資訊與原始資料予以分離（separated），兩者之間得透過所替換的假名相互對應（referenced）[59]；也就是說，假名化資料並未能達到完全去識別化（即匿名）的程度，仍保有可追蹤性（traceable anonymity）[60]。匿名化與假名化技術在概念和原理上之不同，可參照下方圖三和圖四的圖解內容。

---

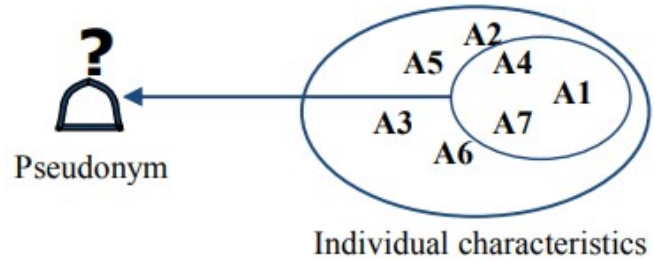
[55] Garfinkel, *supra* note 47, at 3.
[56] Rose Tinabo et al., *Anonymisation Vs. Pseudonymisation: Which One Is Most Useful for Both Privacy Protection and Usefulness of E-Healthcare Data*, INST. OF ELEC. AND ELEC. ENG'R [IEEE] 2 (2010), https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5402501.
[57] Andreas Pfitzmann & Marit Hansen, *Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology (v0.28),* TU DRESDEN 6 (May 29, 2006), https://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.28.pdf.
[58] Tinabo et al., *supra* note 56, at 3.
[59] Johannes Heurix et al., *Pseudonymization with Metadata Encryption for Privacy-Preserving Searchable Documents*, IEEE 3011-12 (2012), https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6149189.
[60] *Id.* at 3013.

doi:10.6342/NTU202401988



圖三：匿名化技術的原理概念

資料來源：Rose Tinabo et al.（2009）[61]



圖四：假名化技術的原理概念

資料來源：Rose Tinabo et al.（2009）[62]

第二款 去識別化技術與加密技術之區辨

另外有必要特別說明的是去識別化和加密之間的關係，加密此一名詞也經常在個資保護法規的去識別化條文規定中，與假名化、匿名或去識別化的其他方法名詞一齊出現。對此，我國健保資料庫憲法訴訟案有大法官在意見書中提及，認為加密應不得充當取代匿名化或假名化，作為有效之去識別化措施[63]。因此，本研究也將先

[61] Tinabo et al., *supra* note 56, at 2.
[62] *Id.* at 3.
[63] 黃昭元大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書，頁 6；謝銘洋大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書，頁 6-7。

doi:10.6342/NTU202401988

從技術的目的和本質進行梳理，嘗試提供去識別化和加密在技術面上有何區別之說明解釋。

加密根據 ISO/IEC 18033-1 提供之定義，是一種可回復（reversible）的資料轉換過程，指運用演算法產生出密碼文字（ciphertext）將原始資料內容（plaintext）進行轉換，以隱藏資訊的內容[64]；加密的可回復性具體展現於，加密演算法在設計上是可透過私鑰將密碼文字進行解密（decryption），使加密內容回復成原先可讀（readable）的形式[65]。因與假名化技術的運作原理有些類似，同為一種遮蔽資料內容的方式，故在規範上，加密也多被認為同樣具有保障資料隱私的效果[66]。

然而，加密的主要功能實則應為防護機敏資訊於未經授權之接近使用[67]，以確保資料內容在存儲和傳輸過程中的機密性和完整性，防範資料洩露和網路攻擊之事件發生[68]。因此，加密和去識別化在功能目的上即有所不同，去識別化技術的重點在於「去個人化（depersonalization）」，透過匿名化或假名化之方式，消除或隱藏資料內容和個人間的連結性，故主要針對處理的會是資料中可識別個人的資訊部分；加密則是一種資料安全維護措施，旨在將所欲保護的資料內容轉化為不可讀的形式，讓未經授權的第三人無法窺視、取得受保護的資料內容，故加密實施的對象不一定是或僅限於個人識別資訊，亦可能包含商業機密和其他重要、敏感的資訊內容。

---

[64] ISO/IEC 18033-1:2021 cl.3.11 (encryption): "(reversible) transformation of data by an *encryption algorithm* (3.12) to produce *ciphertext* (3.7), i.e. to hide the information content of the data." *ISO/IEC 18033-1:2021(En) Information Security — Encryption Algorithms — Part 1: General*, ISO, https://www.iso.org/obp/ui/#iso:std:iso-iec:18033:-1:ed-3:v1:en:term:3.20 (last visited July 15, 2024).

[65] Agencia Española de Protección de Datos [AEPD] & EDPB, *10 Misunderstandings Related to Anonymisation* (misunderstanding 2) (Apr. 27, 2021), https://www.edps.europa.eu/system/files/2021-04/21-04-27_aepd-edps_anonymisation_en_5.pdf.

[66] *Id.* (misunderstanding 2). 歐盟 GDPR 規範架構下即肯認加密雖非匿名技術，但可作為一種假名化的工具，詳見後述第三章、第一節、第三項、第一款。

[67] Heurix et al., *supra* note 59, at 3012.

[68] Amitai Richman, *Pseudonymization Vs Encryption: Understanding the Differences*, K2VIEW (Aug. 25, 2023), https://www.k2view.com/blog/pseudonymization-vs-encryption/#How-Pseudonymization-Works-Preserving-Privacy-Through-Data-Anonymization.

doi:10.6342/NTU202401988

加密的內涵和功能，亦可從我國電子簽章法之規範內容予以理解。現行電子簽章法第 2 條第 3 款將數位簽章定義為「將電子文件以數學演算法或其他方式運算為一定長度之數位資料，以簽署人之私密金鑰對其加密，形成電子簽章，得以公開金鑰加以驗證」，且鑒於數位簽章的安全性和公信力，113 年 6 月 27 日公告之修正草案表明數位簽章有「推定」為本人簽署之效力[69]，符合電子簽章所要求「能辨識及確認電子文件真偽」之要件，屬於電子簽章的一種[70]。

第二項 去識別化的工具原理與實施程序

第一款 常見的去識別化工具

去識別化的方法眾多，運作原理亦有所不同。實務上有許多的方法流程，是將姓名等直接識別資料先移除，後再將身分證字號和其他重要明顯的識別資訊進行多道加密或代碼編譯程序為處理[71]，達到隱藏遮蔽識別資訊的效果，使窺探者不易從中為資訊的串連和推測，此為較簡易的去識別化實施方式。

歐盟 WP29 於 2014 年針對匿名化技術所發布的意見書中（以下簡稱「2014 年意見書」），則將資料匿名化工具依其原理性質區分為兩大類。第一大類的作法主要是在原始資料集中加入雜訊、亂數（即假資料），將資料在資料集中的次序弄亂，使窺探者不易從當中進行連結推導，此類別下的工具有隨機亂數法（Randomization）及差分隱私法（Differential Privacy）[72]；另一大類則是將資料內容概略化，如將確切數字改成區間範圍，或是將文字敘述以更上位的概念取代，如將職業欄位下的法官和律師統一改為法律從業者，抑或是使資料集中的資料達到一定

[69] 電子簽章法修正草案修正總說明，頁 2。
[70] 同上註，頁 4。
[71] 我國全民健康保險研究資料庫的建置流程，即是由衛生福利部中央健康保險署，先刪除健康保險與就醫資料中的姓名資訊，並以加密程式將病患、醫師和藥師的 ID 和醫事機構代碼進行兩道加密後，在提供給國家衛生研究院創建加值資料檔案供產學研究申請使用。有關我國健保資料庫去識別化的程序流程，詳見後述第四章、第二節、第一項、第二款。
[72] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 12-16.

doi:10.6342/NTU202401988

程度的同質性，此類常見的去識別化工具包含彙集法（Aggregation）和 K 匿名框架（K-anonymity）等[73]。另外還有一種思維作法是，以深度學習方法模擬出虛擬的合成資料直接取代原始資料作為使用。

1. 隨機亂數法（Randomization）

隨機亂數法的原理是透過調整資料屬性的數值內容（value），以降低識別資訊的準確性（veracity），當資訊的內容較不精確時，就表示會有更多個人的特徵性質同時符合該資料的紀錄描述 ，如此一來，便可提高窺視者從一項資料內容去推導另一項資料內容的困難程度，達到降低推論風險（inference risks）的效果[74]。

在具體操作上，主要是透過添加干擾（Noise Addition）與置換（Permutation）的方式進行。添加干擾是指在保留原始資料的分配、平均值、變異數等統計特性之前提下，於選定的資料屬性數值中加入隨機值[75]；置換則是添加干擾法下的一種特殊類型，是指不修改原始的數值內容，而是將同一資料集中相同資料屬性紀錄的數值相互交換（swap），例如將資料集中所有當事人的身高數值紀錄交換弄亂，達到保留數值的整體分布和範圍之下，使數值內容和資料主體間無法相互對應連結[76]。

2. 彙集法（Aggregation）

彙集法，又稱「聚合法」，是指將原始資料的內容以統計整合後的形式釋出，如加總值、平均值、最大值、最小值，或是各式的分配或分布圖等[77]。此法以統計數

[73] *Id.* at 16-19.
[74] 祝亞琪、魏銷志，行動支付之個人資料去識別化方法，電腦稽核，34 卷，頁 24（2016 年）。*Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 12.
[75] 祝亞琪、魏銷志（註 74），頁 24。
[76] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 13-14.
[77] 祝亞琪、魏銷志（註 74），頁 24。

doi:10.6342/NTU202401988

值（aggregate statistic）[78]取代直接釋出逐比原始資料之作法，不會揭露任何單一的資料內容，由此達到隱蔽個人身分資訊的效果。

3. K 匿名框架（K-anonymity）

K 匿名框架最早是由 Latanya Sweeney 博士在其 1998 年的研究論文中所提出的資料隱私保護方法[79]，Sweeney 博士發現，僅僅刪除或遮蓋資料中當事人的姓名、電話或地址等直接識別資訊，並不足以使資料達到匿名化的效果[80]，這是因為窺探者仍得運用所掌握的其他資訊，經過組合比對後重新辨認出該筆資料背後所連結代表的特定個人，故 K 匿名框架的提出即是為應對，這種間接識別資訊所產生的再識別風險。

K 匿名框架的原理，可透過我國學者所舉的案例情境進行理解。即假設某一患病檢查資料表中，有一筆性別為女性、身高為 150 公分且體重 64 公斤的病患病況紀錄，且該資料表當中正好沒有其他女性病患資料的身高和體重數值也落在 150 公分和 64 公斤左右，這時若有窺探者剛好從其他資訊來源中確知，具有這些屬性的女性其資料必定有在此資料表中，那麼該窺探者便得以透過資訊的比對中，得知該名女性病患有罹患哪些疾病[81]。對此，Sweeney 博士想出的解法是將資料集中的資料維持一程度的同質性，具體而言，就是讓資料集中各種欄位交集的組合下都有一定數量筆的資料，這個數量標準的設定就是 K 值。K 值愈大，代表資料集中符合相同欄位內容條件的資料筆數愈多，去識別化的程度也就愈大，窺探者便會愈難猜測出代表特定個人的資料究竟是哪一筆。

[78] Ohm, *supra* note 30, at 1715.
[79] *See generally* Latanya Sweeney, *K-anonymity: A Model for Protecting Privacy,* 10(5) INT'L J. UNCERTAINTY, FUZZINESS & KNOWLEDGE-BASED SYS. 557 (2002).
[80] *Id.* at 558-59.
[81] 李思壯、黃彥男，數位時代之數位隱私保護，國土及公共治理季刊，7 卷 4 期，頁 34（2019 年）。

doi:10.6342/NTU202401988

在具體操作上是使用泛化法（Generalization）與抑制法（Suppression）將資料進行處理，使資料集達到 K 值設定的目標。泛化法是指將資料的屬性內容概略化，例如將具體的數字修改成區間範圍，在降低資料粒度之下仍保有數據的真實性；在方法上，可透過捨去法（Rounding）將原始資料的每個數值上調或下調至決定的基準值，或是以頂部/底部編碼（Top- / Bottom- coding）的方式，將具體的資料數值調整為大於或小於某數值[82]。抑制法則是包含紀錄抑制（Record Suppression）刪除單一資料（偏離值）、區域抑制（Field Suppression），即刪除敏感性資料的類別，以及遮罩（Masking）移除資料集中存在的唯一識別符號等方法[83]。

4. 差分隱私法（Differential Privacy）

差分隱私法是由微軟的資安專家 Cynthis Dwork 於 2006 年所提出的資料隱私保護措施，以統計資料庫（statistical database）為適用標的，主要是為防止窺探者可能從釋出的統計數值中，比對得出特定個人隱私資訊的風險[84]。其原理是讓資料庫在刪除任何特定個人之資料紀錄的情況下，經過演算法運算後的輸出結果都不會有顯著變化[85]，使窺探者無法透過觀察資料庫的輸出結果，得知有關特定個人的準確資訊。

差分隱私法是透過加入運算後的雜訊，使資料集的集合分布符合統計學上的定義。下方公式為差分隱私的基礎定義公式[86]， D_1 和 D_2 為僅相差一個元素的資料集合，帶入一個隨機演算法（randomized function）K，參數ϵ則表示在所有可能產生的運算結果範圍 S 中，演算法 K 在鄰近資料集 D_1 和 D_2 中產生相同輸出結果的概率值，也就反映了隱私保護的程度。參數ϵ值愈小，表示在統計資料庫中查詢時，多或少一

[82] 祝亞琪、魏銷志（註 74），頁 24。
[83] 蔡昀臻、樊國楨，大數據之資料去識別的標準化實作初探：根基於 ISO/IEC 2nd WD 20889：2016-05-30，資訊安全通訊，22 卷 4 期，頁 8（2016 年）。
[84] *See generally* Cynthia Dwork, *Differential Privacy: A Survey of Results*, *in* THEORY AND APPLICATIONS OF MODELS OF COMPUTATION 1 (Manindra Agrawal ed al. eds, 2008).
[85] *Id.* at 1.
[86] *Id.* at 2.

doi:10.6342/NTU202401988

個資料元素項目所得出的結果機率分布愈相似，窺探者就愈難分辨此輸出結果究竟是來自 D_1 或 D_2，如此一來，所能達到的隱私保護效果也就愈佳。

$$\Pr[\mathcal{K}(D_1) \in S] \leq \exp(\epsilon) \times \Pr[\mathcal{K}(D_2) \in S]$$

差分隱私法在應用上，可分為交互式（interactive）與非交互式（non-interactive）兩種使用情境模式。交互式模式是發生在統計資料查詢介面，使用者（user）向資料庫（database）發出查詢統計值的指令（query）後，由負責系統（curator）將使用者發出的查詢指令和資料庫的回覆（response）加入運算後的雜訊後再進行回傳；非交互式模式則是負責系統將原始資料集進行雜訊處理後，便直接將資料集整個釋出[87]。

5. GAN 生成資料法

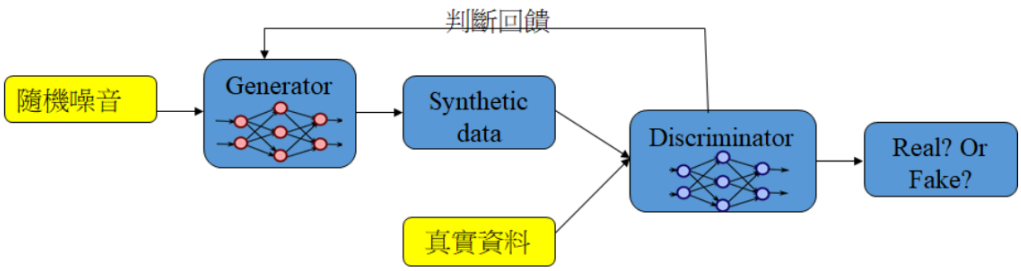
GAN 生成資料法是透過對抗生成網路（Generative Adversarial Network, GAN）學習模擬原始資料，以產生出的虛擬合成資料（synthetic data）取代原始資料用於後續的分析研究。只要合成資料能夠愈接近原始資料，使用合成資料進行分析的準確度，與完全使用原始資料相比的誤差即會愈小。

GAN 技術產生合成資料的原理是透過「生成網路」與「鑑別網路」兩組類神經網路的反覆交錯訓練，不斷提高合成資料的擬真性[88]。具體的操作流程如下圖五，先將隨機變數輸入 GAN 生成器（generator）產生一組合成資料，再由 GAN 鑑別器（discriminator）就每次產出的合成資料與原始資料進行辨識，從擬真度評估的回饋中反覆修正調整 GAN 生成器，直到所產出的合成資料成功通過 GAN 鑑別器的辨識為止。

[87] *Id.* at 1. 王紹睿，淺談人工智慧系統的隱私資訊安全保護機制，科儀新知，215 期 6 卷，頁 78（2018 年）。

[88] 資策會科技法律研究所，合成資料（synthetic data），2020 年 10 月，https://stli.iii.org.tw/article-detail.aspx?no=6 4&tp=1&d=8532（最後瀏覽日：2024 年 7 月 15 日）。

doi:10.6342/NTU202401988



圖五：GAN 技術產出合成資料之操作流程

資料來源：黃維中（2021）[89]

相較於上述的工具方法，GAN 生成資料法的優勢在於，提供隱私保障之餘仍可維持資料的可分析性，且較不會有大幅增加資料運算的成本之問題[90]。我國工業技術研究院自 2017 年起，也開始採用 GAN 生成資料法作為資料去識別化的措施，[91]並發展出支援跨機構間資料整合的延伸應用，即各機構先將手中的原始資料透過 GAN 技術形成合成資料完成去識別化後，再運用橫向整合、資料配對等統計方法將各機構的合成資料進行隨機配對後，匯出整合後的資料集供後續使用，對於資料傳遞、共享過程中的隱私風險提供解套[92]。

最後，關於去識別化方法的採擇須強調的是，去識別化工具雖有實際效用強度的強弱分別，但並非實際效用強的工具方法就必定能夠單獨提供足夠、適當的隱私安全保障；在工具方法的採擇上，往往是依據處理資料性質的不同，考量資料的具體使用情境和資料經去識別化後呈現的屬性和可利用性，選擇搭配不同的工具方法

[89] 黃維中，人工智慧應用下的隱私保護與個資去識別化，科技部全球事務與科學發展中心，2021 年 10 月 18 日，https://trh.gase.most.ntnu.edu.tw/tw/article/content/247（最後瀏覽日：2024 年 7 月 15 日）。

[90] 王若樸，【從 K 匿名法、GAN 和統計整合練兵，再攻聯合學習】工研院揭 3 種去識別化方法，iThome，2021 年 5 月 21 日，https://www.ithome.com.tw/news/144539（最後瀏覽日：2024 年 7 月 15 日）。

[91] 同上註。

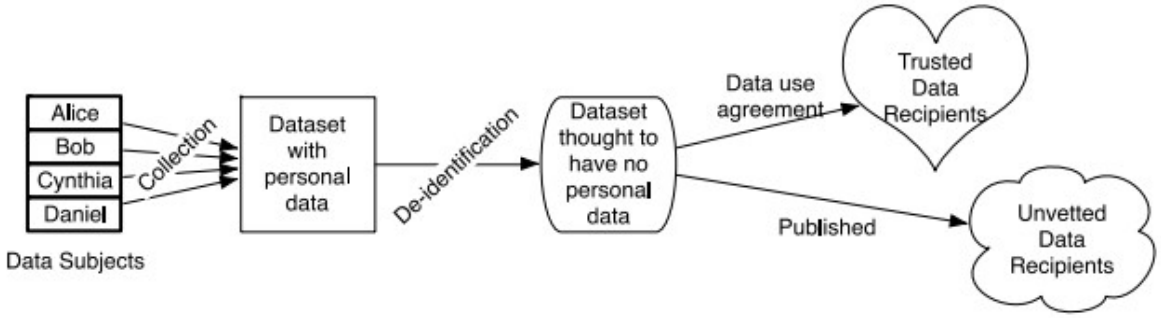
[92] 同上註。

doi:10.6342/NTU202401988

以達到預期的隱私保護和利用效果。以下將詳述業界實務慣常使用的去識別化工具的背景、原理和實施作法。

第二款 去識別化的實施流程

去識別化技術實施的一般機制流程，可參考下方圖六。資料蒐集者向資料主體（data subjects）蒐集其資料後，會將取得的資料整合彙集成資料集（dataset），當中包含許多個人的識別資訊；將該資料集進行去識別化處理完畢後，會產生出一個不包含識別資訊的新資料集，此時總共便有兩個資料集，一個是原始資料（即保有原先所有的識別資訊），另一個則是識別資訊經調整、模糊化後的新資料集，至於該經去識別化後的新資料集，則會取代原資料集被資料蒐集者所使用，由此達到降低隱私風險的效果；資料蒐集者後續亦可能透過資料使用協議（Data Use Agreement, DUA）的簽訂，將新資料集提供給受信任的第三方（Trusted Data Recipients），或是直接將經去識別化處理的資料集公開予大眾使用。



圖六：去識別化處理的程序流程

資料來源：Simson L. Garfinkel（2016）[93]

[93] Garfinkel, *supra* note 47, at 9.

doi:10.6342/NTU202401988

第三項 去識別化的考量因素與侷限

第一款 去識別化對資料效用的減損影響

資料的價值在於其背後所記錄的具體內容，最徹底得以達成去識別化效果的作法固然是直接刪除或永久遮蔽資料中所有的識別資訊，以全然消除資料和個人的連結關係，充分保障隱私，但這麼做會使資料內容變得十分空洞，而無法從中分析萃取出有用的知識和發現，使資料的效用（utility）目的盡然喪失。

隱私安全與資料效用的發揮本質上是兩個相互拉鋸的概念目的，基於去識別化工具的技術原理，去識別化往往會犧牲到資料的可分析性和可利用性。像是彙集法僅呈現資料的統計值，抑或是 K 匿名法使資料集中符合相同欄位內容交集的資料保持有一定數量，這些方法簡單而言就是將資料內容變得概略化、同質化，以防止特定個人在資料集中被輕易標定[94]，然而，當資料的數值內容愈概略、同質化的同時，資料進行統計分析得出的結果也就會愈不精準、愈難進行準確的預測，從而使得資料的效用價值大打折扣。

隨機亂數法、差分隱私法等方法，主要是以加入亂數雜訊的原理方式，達到資料去識別化的目的，資料的隱私安全程度即取決於對於原始資料的擾亂程度。然而，加入假資料或對原始資料進行擾亂，也都會隨之造成資料運算處理難度和成本的增加；甚至在許多時候，資料的項目內容間彼此間是具有邏輯關係或統計上的相關性，故資料內容互換打亂的作法，則可能會破壞資料的邏輯性[95]，從而導致運算結果發生錯誤誤差。

去識別化的目的終究還是要讓資料可以被進行分析和使用，因此，如何在維持資料可分析性之前提下得以最大程度地提供隱私保障的安全性，則是資料使用者在選擇採用去識別化工具時的重要考量因素。

---

[94] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 16.
[95] *Id.* at 14.

doi:10.6342/NTU202401988

第二款 對非結構化資料實施去識別化的技術困境

依據資訊呈現的格式和型態，資料在資料科學領域中可大致區分為結構化資料（structured data）和非結構化資料（unstructured data）。結構化資料是指有經過高度統整（organized）、預先定義化（predefined）、較容易被機器演算法所讀取辨認（decipherable）的資料類型[96]，例如：表格資料內的所有數據紀錄內容都可對應至所屬欄位類別；非結構化資料則是泛指資訊內容並非依特定編排或可預測的格式、規則存儲呈現的資料類型，例如：影像、語音或電子郵件、網路貼文等任何以自然語言組成的文字訊息等[97]。

承前所述，常見的去識別化的作業方式，就是將資料中的識別資訊進行移除或以其他代碼符號取代，這過程中會須要先「偵測找出（detection）」資料內容中的識別資訊，然後再將識別資訊予以「遮蔽（masking）」。結構化資料由於有固定的格式和定義方法，故較容易可透過演算法或程式進行大量、系統性的查找、標註和處理[98]，得以有效率且精準地執行去識別化，但是結構化資料其實反而僅占當今資料量的少數部分，據估計，超過85%以上的商業資訊是以非結構化的資料形式呈現[99]，非結構化資料的增長速度也約為結構化資料的三倍[100]。

然而，相對於結構化資料，要對非結構化資料進行去識別化則困難許多。在結構化資料中，可清楚呈現各項識別資訊和所歸屬資料主體的對應性，且每項識別資訊的內容描述或數值的屬性都有被清楚定義（例如：表格第二列數值為受測者的體重，單位為公斤，四捨五入到小數點後第一位），但在非結構化資料中，這些對應

---

[96] *Structured Versus Unstructured Data: What's the Difference?*, IBM (June 29, 2021), https://www.ibm.com/think/topics/structured-vs-unstructured-data.
[97] *See* Mary Shacklett, *Structured vs Unstructured Data: Key Differences*, DATAMATION (Nov. 23, 2023), https://www.datamation.com/big-data/structured-vs-unstructured-data/.
[98] *Id.*
[99] *Glossary: Unstructured Data,* RESOURCES.DATA.GOV, https://resources.data.gov/glossary/unstructured-data/ (last visited July 15, 2024).
[100] Robert Heeg, *Possibilities and Limitations, of Unstructured Data*, RESEARCHWORLD (Feb. 20, 2023), https://researchworld.com/articles/possibilities-and-limitations-of-unstructured-data.

doi:10.6342/NTU202401988

關係、屬性定義和紀錄格式等規則皆不存在，故在偵測應被遮蔽的識別資訊過程中，經常會發生標註錯誤的問題（例如將「艾迪森氏症（Addison's disease）」等疾病名稱誤認為是個人姓名予以刪除），往往需要實際檢查資料內容中的前後脈絡才能正確進行判斷[101]，這使得非結構化資料的去識別化，目前仍須大量仰賴人工方式完成實施[102]。

再者，在非結構化資料類型中，除了文字檔案外，影像、視訊和音訊等多媒體資料（mutimedia data）來源中，亦可能含有可識別個人的相關資訊，例如：Google Map 的街景影像可能捕捉到私人的住址外觀、監視錄影器畫面拍下個人的臉部外觀特徵和出沒地點的確切時間、錄音檔案含有特定個人的聲紋資訊等。有鑑於這類資料多屬於高維度的資料（high dimensional data），資料本身的變數多，包含雜訊和紀錄不精確的可能性便會提高，導致資料處理過程上的不易，誤差的機會較高，且會須要耗費較多的運算能力和存儲空間[103]；此外，要從不同存儲格式的聲音或影像欄位中萃取出特定識別資訊[104]，進而再對當中含有識別資訊的音訊、影像部分進行修改[105]，也都是技術上的一大挑戰。由此，對於非結構化資料而言，所能選擇有效的去識別化工具技術，尤其是進行資料遮蔽的工具，相對於結構化資料即是非常有限[106]。

[101] Emily M. Weitzenboeck et al., *The GDPR and Unstructured Data: Is Anonymization Possible?*, 12(3) INT'L DATA PRIV. L. 184, 188 (2022).
[102] Fadi Hassan et al., *Anonymization of Unstructured Data Via Named-Entity Recognition*, *in* MODELING DECISIONS FOR ARTIFICIAL INTELLIGENCE 296, 297 (Vicenç Torra et al. eds, 2018)
[103] 翁慈宗，資料探勘的發展與挑戰，科學發展，442 期，頁 39（2009 年）。
[104] 同上註；Garfinkel, *supra* note 47, at 32.
[105] Garfinkel, *supra* note 47, at 33.
[106] Weitzenboeck et al., *supra* note 101, at 189.

doi:10.6342/NTU202401988

第三節 去識別化後的重新識別問題

第一項 何謂重新識別

第一款 重新識別的定義內涵

重新識別（re-identification），又稱為「再識別」或「去匿名化（deanonymization）」，是指意圖辨認出已在去識別化過程中被去除、隱藏的個人身分和資訊[107]，以將個人與其識別資訊再度串連起來為目的之程序過程[108]。英國ICO在其所發布的「匿名化：管理資料保護風險實務守則（Anonymisation: Managing Data Protection Risk Code of Conduct）」文件中，則將重新識別定義為，運用資料比對（data matching）或其他相類技術將匿名化資料轉變回個人資料的過程[109]。

重新識別風險的存在，主要是來自於去識別化措施本身的侷限和缺失，畢竟無任何一項去識別化技術工具有辦法完全去除資料的識別性風險，特別是尚須兼顧考量資料可利用性之情況下。由於即便資料已經去識別化處理，仍難以完全阻絕之後再遭到他人刺探、洩露的可能性，這種資料主體的身分資訊嗣後再被重新識別出來的可能性，又被稱為資料去識別化固有的「殘餘風險（residual risk）[110]」。

第二款 實際發生的重新識別案例情形

以下三個在美國的隱私事件和研究，皆為資料管理者在將個人資料進行去識別化處理，確信識別資訊已無從識別後將資料釋出，嗣後卻仍成功被重新識別，最終導致隱私洩露之結果。

---

[107] Garfinkel, *supra* note 47, at 9.
[108] Info. and Priv. Comm'r of Ont. [IPC], *De-identification Guidelines for Structured Data*, at 2 (June, 2016), https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf.
[109] ICO, Anonymisation: Managing Data Protection Risk Code of Conduct, at 6 (Nov., 2012), https://ico.org.uk/media/1061/anonymisation-code.pdf.
[110] Michèle Finck & Frank Pallas, *They Who Must Not Be Identified—Distinguishing Personal from Non-Personal Data Under the GDPR*, 10(1) INT'L DATA PRIV. L. 11, 35 (2020).

doi:10.6342/NTU202401988

1. 麻州醫療資料洩露案

第一個事件是來自 Latanya Sweeney 博士的研究，其以 1990 年的美國人口查資料為研究材料，結論證明僅須要和郵遞區號五碼之三項資訊，即可識別出 87%的美國人民；又 53%的美國人民可透過其居住地區、性別和出生日期被成功識別[111]。在其後續的研究中，更有示範其是如何運用公開資料進行重新識別。美國麻薩諸塞州的團體保險機關（Group Insurance Commission, GIC）將該州政府機關人員及其眷屬共約 135,000 名人的醫療資料紀錄，經去識別化處理後提供和販售給學術研究者及產業人士使用[112]。

GIC 稱在釋出醫療資料前，已有將姓名、地址和社會安全碼等明顯識別資訊移除，但出生日期、居住地郵遞區號、性別和敏感性健康醫療紀錄內容則仍有予以保留[113]。但 Latanya Sweeney 博士卻仍藉由麻州政府所公開的選民登記名冊資料，透過當中所登載之姓名、地址、郵遞區號、出生日期和性別等選民資訊，與 GIC 釋出的資料集相互比對後，最終成功在醫療資料中標定出麻州州長的醫療紀錄。[114]

2. AOL 隱私洩露案

第二個事件是發生在 2006 年 8 月，美國網路搜尋服務提供者 America Online LLC（AOL）將約 65 萬名用戶在當時近三個月內總計超過 2000 萬筆的網頁搜尋和瀏覽紀錄資料，公開在其所建置的研究資源網站供大眾下載使用[115]。公開的資料內容

[111] Latanya Sweeney, *Simple Demographics Often Identify People Uniquely* 2 (Carnegie Mellon Univ., Data Priv. Working Paper No. 3, 2000).
[112] *Id.* at 2.
[113] Ira S. Rubinstein & Woodrow Hartzog, *Anonymization and Risk*, 91(2) WASH. L. REV. 703, 711 (2016).
[114] Sweeney, *supra* note 111, at 2.
[115] *Public Comments: August 2006 – FTC Complaint About Search AOL Data Releases*, WORLD PRIV. FORUM (Aug. 16, 2006), https://www.worldprivacyforum.org/2006/08/public-comments-ftc-complaint-about-search-aol-data-releases/.

doi:10.6342/NTU202401988

包含姓名、IP 位置、搜尋和網頁點擊的時間紀錄、地址、社會安全碼，以及其他用戶可能在搜尋欄框輸入的任何資訊[116]。

AOL 在釋出資料前有將資料中之姓名、IP 位址等顯著（直接）識別資訊改以隨機代碼的方式呈現，但基於資料可供研究使用之考量，在處理上仍保留資料的對應關係，使資料使用者有辦法對應出每筆搜尋資料內容是來自哪一個代碼編號的用戶。即便 AOL 已將所釋出的資料進行去識別化，但不幸地的是，由於資料中所包含的關鍵字搜尋和網頁瀏覽紀錄內容過於大量且鉅細靡遺，仍已足以讓資料背後用戶的真實身份被嗣後揭露。

紐約時報的記者便實際嘗試，並在報導中說到他們並未花費太多氣力，就能夠從「手指麻木」、「60 歲單身男人」、「狗到處撒尿」、「喬治亞州利伯恩縣的園木設計師」、「喬治亞州格林納特縣 Shadow Lake 區的出售房屋」等搜尋關鍵字中，找出編號 4417749 用戶的真實姓名，並得知其為 62 歲的寡婦、家住在喬治亞州利伯恩縣、養著三隻小狗，且喜歡經常上網搜尋其朋友患有的疾病問題[117]。

3. Netflix 電影評分紀錄識別案

第三個事件同樣發生在 2006 年，Netflix 為提升其電影推薦系統的效能，便舉辦演算法開發競賽，提供近 50 萬名訂閱用戶在 1999 年至 2005 年期間超過 1 億筆的電影評分紀錄，作為比賽中使用的訓練資料。Netflix 聲明所釋出的資料紀錄是從資料集中隨機挑選而來，僅占所有紀錄資料的百分之十；訓練資料內容包含用戶名、電影別、評分分數及評分時間四個欄位，其中用戶名和電影別是以隨機編碼的方式取代呈現，且資料都有再以置換法進行去識別化處理[118]。

---

[116] Michael Arrington, *AOL Proudly Releases Massive Amounts of Private Data*, TECHCRUNCH (Aug. 7, 2006; 9:17 AM), https://techcrunch.com/2006/08/06/aol-proudly-releases-massive-amounts-of-user-search-data/.
[117] Michael Barbaro & Tom Zeller Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, THE N.Y. TIMES (Aug. 9, 2006), https://www.nytimes.com/2006/08/09/technology/09aol.html.
[118] Arvind Narayanan & Vitaly Shmatikov, *Robust De-Anonymization of Large Sparse Datasets*, *in* 2008 IEEE SYMPOSIUM ON SECURITY AND PRIVACY 111, 118 (2008).

doi:10.6342/NTU202401988

然而，事後 Arvind Narayanan 和 Vitaly Shmatikov 兩位德克薩斯大學奧斯汀分校學生，嘗試將 Netflix 釋出的資料與電影評分入口網站（Internet Movie Database, IMDb）上公開的電影評分紀錄進行參照，在 50 名所追蹤在 IMDb 留下評分的用戶中，即有 2 名用戶可在 Netflix 的去識別化資料中被比對識別出來[119]。Arvind Narayanan 和 Vitaly Shmatikov 的研究結果證明，其實只要須要部分、不多的額外資訊，且即便該用以結合比對的資訊有些許程度的誤差，仍已足以重新識別 Netflix 釋出的訓練資料[120]。

上述三個案例顯現出，要從去識別化資料中拼湊探得有關特定個人資訊，其實比想像中容易；此也呼應到 Paul Ohm 所觀察到的「釋出後即遺忘（Release-and-Forget）」現象之問題隱憂，其指出許多資料管理者將已經去識別化處理的資料紀錄釋出（如對大眾公開、提供予第三人或供自己組織內部使用等）後，便不會再去關關注、追蹤該資料紀錄後續發生的影響。這種將資料進行去識別化後隨即釋出的模式在實務上十分常見，且通常含有許多隱私漏洞，往往最會淪為重新識別的攻擊目標[121]。

不過，也並非所有的重新識別的實施，都是出於刺探他人隱私、使己方事業組織得利，或是欲使去識別化實施者遭受責難等負面意圖等。在某些情形中，個人或組織是在取得授權允許且具備正當目的之前提下，才對已去識別化的資料進行重新識別攻擊，例如：受委託對去識別化實施品質進行測試評估，事前已簽署保密協議且過程中皆有依循安全保護程序，或是基於學術研究或新聞價值目的，在未違反 DUA 的範圍內的合法實施等[122]。

[119] *Id.* at 122-23.
[120] *Id.* at 120-23.
[121] Ohm, *supra* note 30, at 1711-12.
[122] Garfinkel, *supra* note 47, at 10.

doi:10.6342/NTU202401988

第二項 重新識別攻擊

在資訊傳遞和儲存十分迅速且便利之下，窺探者往往得以透過公開資訊、資料販售或盜取等各種管道，取得大量有關個人的片段性資訊，並透過先進的技術在大量的資料中進行快速的組合比對。

第一款 串連攻擊

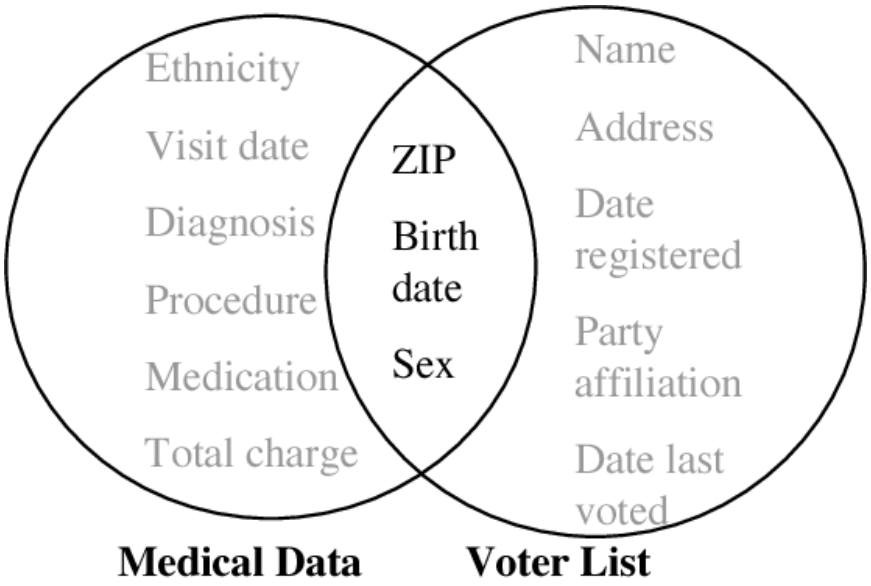
串連攻擊（linkage attack）是指以利用背景資料、公開資訊等其他輔助資訊（auxiliary information）進行比對，將特定個人連結至去識別化資料集中的某筆紀錄資料[123]，由此探得有關該個人的身分和其他個人資訊。上述的 AOL、Netflix 和 GIC 隱私事件即是屬於此類，在此三個實例中，第三人都是經由掌握網路資訊或與去識別化資料，在所含內容上具有一定相關程度的其他來源資料，透過對比兩資料中所含有的重疊資訊別，使資訊成功串連。

例如在 Netflix 案中，Netflix 所釋出的用戶評分資料和同屬提供電影評分資訊的 iMDb 網站上公開的評分資訊，由於產業和平台性質的類似性，故在無論是在所資料主體（用戶群）和所蒐集的資料內容（電影別和評分結果等）都具有一定程度的相關性，因此，即便 Netflix 已將用戶名和電影別進行去識別化處理，仍得經由 iMDb 網站上的公開資訊對照揭露出 Netflix 訓練資料中特定評分紀錄背後的個人身分，進而對應出有關該個人在 iMDb 和 Netflix 上的其他的更多評分紀錄，從而得知該個人的對電影類別的偏好。

[123] Rubinstein & Hartzog, *supra* note 113, at 711.

doi:10.6342/NTU202401988

又如在 GIC 案中，雖然 GIC 已將醫療資料中的姓名、地址和社會安全碼等識別資訊予以刪除，但 Latanya Sweeney 博士仍透過選民登記名冊，與 GIC 醫療資料所交集的資料欄位（如下圖七），成功對應出醫療資料紀錄歸屬的個人，由此得知有關該個人的健康及醫療狀況。



圖七：麻州醫療資料洩露案的資料對應關係
資料來源：Latanya Sweeney（2002）[124]

串連攻擊的預防與最大化資料效用間的兩難在於，單單僅是將直接識別資訊刪除或遮蔽處理，往往遠不足以抵擋串連攻擊，從上述案例中已明顯可證，經由輔助資訊中對於有關個人間接識別資訊的累積，亦能夠將經去識別化處理所隱藏的個人身分予以揭穿。然而，若欲降低、甚至幾乎完全減削串連攻擊的成功率，便會須要將原始資料中更多，甚至是幾乎是大部分的間接識別資訊予以刪除或遮蓋，抑或是同時將原始資料中同筆資料不同資料項目間的對應性打亂或消除，但隨著愈多的資

[124] Sweeney, *supra* note 79, at 559.

doi:10.6342/NTU202401988

訊被刪除遮蔽，資料集中資料被弄得愈雜亂，資料的內容也就會欲空泛、可分析性便會愈低，使資料變得愈來愈不好利用。

第二款 假名還原

另一種導致資料被重新識別的可能原因是假名還原（pseudonym reversal），意即所實施的去識別化技術方法遭窺探者破解，使得原先隱藏的識別資訊被復原，導致資料主體身分被揭露。[125]此原因類型發生通常被歸屬於假名化類型的去識別化工具。假名化技術的一大特徵在於可還原性，其可還原性來自於所生成的代碼和所取代的原始資料值之間的對應關係，因此，掌握代碼生成的演算法程式、加密鑰匙或編碼簿之人，便有辦法破解該對應關係，將去識別化資料的資訊進行還原。

正常而言，這些演算法程式、加密鑰匙或編碼簿等轉換工具，會由原始資料的蒐集者組織內部或受其委託執行去識別化的單位所保管，在委託執行去識別化的情形下，通常會以簽訂 DUA 或其他契約方式約定、限制對於去識別化資料的還原行為。然而，倘若轉換工具不慎外流，或對應關係的機制被未經授權的第三人所破解，即會發生隱私洩露的結果。

第三人破解還原假名化資料的難易程度，取決於代碼、字串等假名的產生方式（隨機或演算法生成）、鑰匙的可取得性，以及用以取代原始資料的假名是特別專屬或重複使用的。此外，隨著假名使用的時間愈長，所對應代表的原始資料量愈多，對於隱私所造成的風險也就愈大[126]。

[125] Garfinkel, *supra* note 47, at 10.
[126] Garfinkel, *supra* note 47, at 17.

doi:10.6342/NTU202401988

第三項 重新識別風險的控制

第一款 去識別化無法提供絕對隱私保障

重新識別的發生基本上已經無可全然避免，有見解就指出在 Microsoft、Facebook、Amazon 等大型跨國科技公司對個人資料的廣泛蒐集之下，這些資料巨人早已囤積大量、詳細的資料足供進行比對、參照，故恐怕經過任何處理、遮蔽的資料對這些公司而言，都算是可識別個人的資料[127]。

各界亦逐漸湧現對於去識別化有效性的質疑聲浪，甚至以「匿名化已死」等標語來表達重新識別對於去識別化信任造成的打擊[128]；亦有學者認為去識別化相較於其他隱私保障措施中被過度重視、強調[129]；另有意見指出對於去識別化的過分倚賴，放任仍殘有識別風險的資料自由流通，反而將導致對個人隱私更嚴重的危害結果，此外，應建立的正確思維認知應是，我們往往只能「努力嘗試」將資料進行匿名化或去識別化，而不是一概以為資料一定都能達到匿名化或完全去識別化之狀態[130]；抱持類似看法的學者亦有呼籲，完美的匿名化應是一種迷思，應將匿名化理解為將隱私風險最小化的程序過程，而不是一再地追求匿名化能作為隱私安全性的保證，故認為去識別化在制度規範設計上應著重於風險導向[131]。

第二款 重新識別風險的評估

為因應重新識別的發生與其所衍生對隱私及其他權利的負面影響，在選擇去識別化工具時，即須同時考量各去識別化工具是否得以有效抵禦或降低重新識別攻擊的發生，這當中最重要的就是對於重新識別風險的評估和控制程序。

---

[127] Patricia S. Calhoun & Patricia M. Carreiro, *De-Identified Data Exception in HIPAA Poses A Litigation Risk*, STAT (Apr. 16, 2020), https://www.statnews.com/2020/04/16/de-identified-data-exception-hipaa-litigation-risk/.
[128] Rubinstein & Hartzog, *supra* note 113, n. 18.
[129] *See id.* at 709.
[130] *See* Ohm, *supra* note 30, at 1744, 46-48.
[131] Woodrow Hartzog & Ira Rubinstein, *The Anonymization Debate Should Be About Risk, Not Perfection*, 60(5) COMMC'N ACM 22, 24 (2017).

doi:10.6342/NTU202401988

在重新識別風險的影響因素上，重新識別攻擊的成功與否，與窺探者的技能和其掌握的資訊息息相關。一種常見重新識別風險的分類方式，即是依據窺探者的目的及其掌握資訊的具體程度，分為以下三個風險情境類型：

（1）檢察官風險（prosecutor risk）：係指在窺探者已知去識別化資料中含有某一特定個人的資料之情境下，窺探者能夠透過其他輔助資訊，正確地將該個人從去識別化資料中重新識別的可能性[132]。

（2）新聞記者風險（journalist risk）：係指窺探者先前不知道去識別化資料中究竟含有哪些個人的資料，卻仍得以從中成功重新識別出任何一個人的資訊的可能性。[133]在此情境下，窺探者進行重新識別攻擊的目的，通常不是要針對某一特定個人，而是欲使釋出資料的組織機構遭受打擊[134]。

（3）行銷業者風險（marketer risk）：在此情境下，窺探者欠缺對於去識別化資料中內容的背景資訊，其基於行銷推廣等類似目的，希望能夠從去識別化資料中盡可能地重新識別出愈多的資料[135]。因此，行銷業者風險關注的是，有多少比例的資料可能被窺探者揭露，可能被重新識別的資訊比例愈大，行銷業者風險也就愈高。

另一方面，有鑑於無法達到完全的去識別化，在風險評估程序上則會須要先設定可接受的重新識別風險，作為評估的門檻。舉例來說，加拿大安大略省資訊及隱私機關（Information and Privacy Commissioner of Ontario, IPC）基於政府機關所經手和公開的資料量漸增，為提升政府機關處理資料的安全性，便於 2016 年所提出「結構化資料去識別化指引文件（De-identification Guidelines for Structured Data）」，提

[132] Fabian Prasser et al., *The Importance of Context: Risk-Based De-Identification of Biomedical Data*, 55(4) METHODS INFO. MED. 347, 349 (2016).
[133] *Id*. at 349.
[134] Khaled El Emam et al., *Protecting Privacy Using k-Anonymity*, 15(5) J. AM. MED. INFORMATICS ASS'N 627, 628 (2008).
[135] Prasser et al., *supra* note 132, at 349.

doi:10.6342/NTU202401988

供機關在對資料進行去識別化時應遵循的程序之參考建議，當中，IPC 即說明去識別化的實施強度應與資料釋出後所致生的重新識別風險相稱[136]，並以最低可接受的重新識別風險門檻，作為可否被認定為去識別化資料之判斷[137]。

在具體程序步驟上[138]，由於資料的釋出模式將影響資料的可及性（availability）和可受控制保護之程度，故 IPC 建議的第一步驟是先決定資料究竟是以公開、半公開或不公開的形式釋出[139]；進而，在可接受重新識別風險門檻的設定上，應考量資訊的敏感程度、資訊的範圍和具體內容、資料主體的多寡、資訊外洩和不當使用對資料主體可能致生的傷害程度、資訊的揭露是否有經資料主體同意、資訊是否為資料主體主動提供（隱私期待程度低），以及資料主體是否有明示同意個人資訊以去識別化方式揭露供進階使用等因素[140]。

又如歐洲藥品管理局（European Medicines Agency, EMA）針對人體用醫藥產品臨床報告資料之發布政策，所制定的外部指引文件中亦表示，為在資料效用和降低重新識別風險間求取最適衡平，在確保隱私已受妥適保護下，盡可能最大化地保留資料的實用效能[141]，須透過風險評估程序透過可接受重新識別風險門檻之設定，以決定去識別化實施具體應達到的程度[142]。此外，EMA 也強調重新識別的風險程度與資料揭露的情境方式息息相關，在資料是對大眾公開的情況下，因難以控管資料的

---

[136] *De-identification Guidelines for Structured Data*, *supra* note 108, at 9.
[137] *Id.* at 10.
[138] IPC 建議的去識別化程序共包含九個步驟：一、決定資料的釋出模式（Determine the Release Model）；二、對資料內容的識別性進行分類（Classify Variables）；三、設定可接受的重新識別風險門檻（Determine an Acceptable Re-identification Risk Threshold）；四、衡量資料風險（Measure the Data Risk）；五、衡量情境風險（Measure the Context Risk）；六、計算整體風險（Calculate the Overall Risk）；七、進行資料去識別化（De-identify the Data）；八、評估資料的效用（Assess Data Utility）；九、程序紀錄歸檔（Document the Process）。*Id.* at 6-20.
[139] *Id.* at 7-8.
[140] *Id.* at 10.
[141] Eur. Med. Agency [EMA], *External Guidance on the Implementation of the European Medicines Agency Policy on the Publication of Clinical Data for Medicinal Products for Human Use*, at 43 EMA/90915/2016 (v.1.4) (Oct. 15, 2018), https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/external-guidance-implementation-european-medicines-agency-policy-publication-clinical-data-medicinal-products-human-use-versio n-14_en.pdf.
[142] *Id.* at 70.

doi:10.6342/NTU202401988

後續流通和使用，故應設定較低的可容許重新識別風險程度，又若資料僅限於在內部非公開地進行交換揭露，因尚得透過契約、內部相關安全措施控管、限制資料的使用，故可接受的重新識別風險程度則得提高放寬[143]。

第二款 重新識別風險的難以控制性

即便對於重新識別攻擊的發生有所體認，也意識到重新識別風險評估的重要性，然而，基於對公開資訊的不可掌握性，導致重新識別風險在本質上具有不斷變動和難以預測的特性。

申言之，想要完整掌握當前究竟已經有哪些資訊存在基本上非常困難，此外，也難以預料和控制未來又會有哪些資訊以何種形式被釋出公開。[144]在這當中，網路和社群平台也是一個重要影響因素，隨著社交文化行為模式的改變，人們每日在社群帳號中所上傳有關自己生活細節的貼文、照片和影音，都已成為窺探者可利用進行串連、比對的資訊來源。[145]

此外，縱使已採取當下所得預估重新識別風險最小的去識別化措施，但隨著資料的產生、資訊的釋出以及資料分析比對技術的演進，窺探者能夠獲取的資訊將愈來愈多、比對分析的能力和成本將愈來愈低，導致重新識別風險將隨時間經過逐漸增加。由此，在無法控制未來資訊的釋出和技術發展程度之情況下，也就非常難以對於重新識別風險進行完善的評估預測。

---

[143] *Id.* at 40.
[144] *Anonymisation: Managing Data Protection Risk Code of Conduct*, *supra* note 109, at 18.
[145] Ohm, *supra* note 30, at 1725.

doi:10.6342/NTU202401988

第四節 小結：去識別化規範制度面臨的挑戰

去識別化是降低或消滅資料與個人間之關聯程度的資料處理程序，是一種隱私強化、保護措施。長久以來，在二分法的資料保護規範框架下，人們認為只要將個資經過去識別化處理，達到永久無法識別特定個人之匿名狀態，因資料的識別風險已消滅或大幅下降，即可脫離資料保護規範的適用範圍，得自由對去識別化資料進行運用和交換。

引進去識別化作為衡平隱私保護與社會進步發展的制度措施，看似是一個兩全其美的方法，但去識別化的實施並非百利無一弊害，蓋去識別化所帶來的隱私保障效果，其實是需要在不同程度上犧牲資料的可利用性和可分析性而來，因此，在規範制度上，將會面臨的兩難是，為保障個人隱私，去識別化的標準要求須較嚴格，但規範設置較嚴格標準，卻可能造成資料無法進行有效利用，使兼顧促進資料使用之立法意旨落空；另一方面，去識別化在技術層面上其實仍有其侷限性，尤其是未經定義、型態格式複雜的非結構化資料，較難透過機器演算法精準地辨認、截取當中的識別資訊，並進行移除或遮蔽。

在期望去識別化技術能長久且周延地保障資料使用的隱私安全，為制度開放資料自由流通和運用提供背書的同時，隨著資訊量大爆炸和資料技術工具的進步，重新識別風險的發生所導致層出不窮的隱私外洩侵害事件，也徹底動搖著人們對於去識別化有效性的信心。對此，許多學說便有呼籲，政策制度的制訂不應再過度倚賴、以去識別化作為唯一的隱私保護或放寬管制的條件機制，蓋去識別化實則僅能降低資料的識別程度，但無法提供永久、絕對之隱私保障。

總結而言，本研究認為去識別化得減低個人資料的識別程度，有助於緩解個人資料運用可能產生的隱私侵害風險，應可作為資料保護規範合理放寬的基礎。去識

doi:10.6342/NTU202401988

別化對於資料效用的減損影響，以及如何有效防控重新識別風險，則是去識別化規定制度妥善建置應納入評估的考量重點。

doi:10.6342/NTU202401988

第三章 借鑑比較法：兩種去識別化制度的規範模式

第一節 歐盟 GDPR 下的去識別化規範

GDPR 是於 2016 年 4 月 27 日通過，自 2018 年 5 月 25 日生效時起正式取代 1995 年的資料保護指令，作為歐盟地區具通則性之個資保護規範。在與各會員國內國法之適用關係上，蓋 GDPR 在法律位階上提升為「規則（regulation）」之層級，因此，無須再經各會員國內部立法程序通過實施，在生效起即直接在各會員國內產生規範拘束效力[146]；GDPR 規範中亦有設置部分之空白條款（opening clauses），授權各會員國得制定國內法詳細定義、補充之[147]。因此，歐盟之個資保護規定是以 GDPR 作為主要的規範框架，至於 GDPR 中相關要件之具體定義和實施措施，則須另參照各會員國內國法之規定。

第一項「匿名」與「假名」概念之區分

GDPR 中並未直皆使用去識別化一詞，而是以「匿名」與「假名」兩個代表不同去識別化程度的規範名詞，來表示去識別化在 GDPR 中的不同規範定位和功能類型。GDPR 第 4(5) 條將假名化（pseudonymisation）定義為是「一種資料處理方式，使個人資料在未對照其他額外資訊的情形下，無法從中辨認出該資料所歸屬的資料主體；額外資訊必須被分開保存，且須實施技術上和組織上之措施，以確保該個人資料無法被歸屬至已被識別或可得識別之特定個人」[148]。至於匿名，匿名化並未被訂入 GDPR 本文的規範內容，但立法理由第 26 點中則有對何謂匿名資訊進行說明闡

[146] THE EU GENERAL DATA PROTECTION REGULATION (GDPR) A COMMENTARY 10 (Christopher Kuner et al. eds., 2020).
[147] LUKAS FEILER ET AL., THE EU GENERAL DATA PROTECTION REGULATION (GDPR): A COMMENTARY 16-17 (2018).
[148] 原文為：“‘pseudonymisation’ means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.”

doi:10.6342/NTU202401988

釋，其中表示匿名資訊（anonymous information）是指「非有關於已被識別或可得識別之個人之資訊，抑或是經一定處理後，已轉變成為無法再識別出資料主體之匿名狀態之資訊」[149]。

在規範功能的區別上，GDPR 的立法理由有強調，匿名資訊非個人資料，不在GDPR的適用範圍內，對於匿名資訊的處理活動也不會受到GDPR之規制[150]，也就是說，資料的匿名與否關係到的是 GDPR 規範效力是否發動的問題；相較之下，假名化則與 GDPR 的適用與否無關，蓋將個資進行假名化處理的目的，是在於降低資料使用對於資料主體可能致生的風險，以及作為得幫助資料控管者和處理者履行其資料保護義務之手段，本規則明文引進假名化也並非有意排除任何其他資料保護措施之採行實施[151]。

再者，WP29 於其所發布的意見書中亦有明確指出，假名化並非屬於 GDPR 規範下的匿名化措施範圍，這是因為假名化的作用僅在於「減低」資料主體與資料集中識別資訊之間的連結性（linkability），而非完全除去[152]。由此，經假名化之個人資料，因仍得結合其他額外資訊識別出特定個人之身分，性質上應屬可得識別或得間接識別（indirectly identifiable）出特定個人之資訊[153]，仍屬於個人資料之範疇，而有本規則之適用。

關於匿名資料和假名資料在 GDPR 的概念定義下究竟應如何區分之疑問，可從歐洲議會（European Parliament）針對 GDPR 草案（GDPR Interservice Draft[154]，下稱

---

[149] 原文為：“anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.”
[150] GDPR r. 26.
[151] GDPR r. 28.
[152] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 20.
[153] *Opinion 4/2007 on the Concept of Personal Data*, *supra* note 17 at 18
[154] GDPR 草案（GDPR Interservice Draft）係由歐盟執委會下之司法與消費者保護總處（Directorate-General for Justice and Consumers, DG JUST）於 2011 年 11 月 29 日所公布，歷經多次調整後，最終的草案版本於 2012 年 1 月 25 日定案。Kuner et al. eds, *supra* note 146, at 4.

doi:10.6342/NTU202401988

「GDPR Proposal」）的討論文件中，嘗試尋找可能的解釋論點。

歐洲議會當時指派議員（Member of European Parliament, MEP）Jan Philipp Albrecht作為GDPR草案的主要報告員（rapporteur），Jan Philipp Albrecht後於2013年1月8日針對GDPR Proposal，向歐洲議會下的公民、司法與內政委員會（Committee on Civil Liberties, Justice and Home Affairs，下稱「LIBE委員會」）提出了一份草案修正報告書（下稱「Albrecht Report」），當中創設假名資料（pseudonymous data）此一新的資料類別，稱假名（pseudonym）是指對應於個別內容的特殊符號，認為其功用在於「阻止對特定個人之直接識別，但允許可能對特定個人進行標定（single out）」[155]；匿名資料（anonymous data）之概念，則是強調須無論是單獨來看或與其他相關資料相結合，都無法或需要花費不成比例之成本始得直接或間接連結至特定個人[156]，始足當之。

最終LIBE委員會所提出的協議草案版本（compromise text）有參照Albrecht Report的定義內容，進一步定義假名資料為「只要有將額外資訊分開存放，且要求應實施技術上和組織上措施以確保資料的無法歸屬性，使在不參照結合其他額外資訊之情況下，即無法歸屬至特定個人之個人資料」[157]，內容上已幾乎等同於現行GDPR第4(5)條對於假名化的規定說明文字。

綜合上述來看，GDPR中假名化的功能在於消除資料的直接識別性，透過以假名將資料中的明顯識別資訊適當遮蔽，達到單從資料的內容本身無法辨別出資料歸屬主體的去識別效果；然而，經假名化後的資料若有其他相關連的額外資訊可供參

[155] Eur. Parl. Doc. (COM 2012)0011 (2013), at 65 (Amendment 85) ("'pseudonym' means a unique identifier which is specific to one given context and which does not permit the direct identification of a natural person, but allows the singling out of a data subject").
[156] *Id.* at 16 (Amendment 14) ("This Regulation should not apply to anonymous data, meaning any data that can not be related, directly or indirectly, alone or in combination with associated data, to a natural person or where establishing such a relation would require a disproportionate amount of time, expense, and effort . . . .")
[157] Eur. Parl. Doc. EP-PE_TC1-COD(2012)0011, at 99 (art. 4(2a)) ("'pseudonymous data' means personal data that cannot be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution").

doi:10.6342/NTU202401988

照、比對，便很容易可藉由資料間內容的結合、連結，成功得出假名化資料所遮蔽的資訊內容，進而間接辨認出資料主體的身分，這也是 GDPR 第 4(5) 條才會要求額外資訊須分開保存，以降低重新識別的風險，由此，假名化資料仍具有間接識別性，故仍屬於個人資料之範疇。

相對之下，在 GDPR 之制度規範中，匿名資料所指的是不具有識別性的資訊，其所要求得不僅是從單一資料之內容本身無法識別特定個人外，也須要達到透過串連其他資料的資訊內容，也無法或至少難以在合理成本的投入下，推知出特定個人的身分資訊之程度，始足當之。

以前章所提到的重新識別事件（美國麻薩諸塞州案、AOL 案及 Netflix 案）為例，就本研究的觀點而言，三個案件中原保有資料的機關、事業單位所採取的去識別化方法（移除、以隨機代碼取代姓名、地址、社會安全碼等識別資訊），在GDPR的規範脈絡下，都僅屬於假名化措施的概念範圍，蓋事件後續的發展足以證明，案例中所實施的資料處理措施，並未能防免資料主體的身分，仍舊在與其他公開資訊比對過程中被推敲揭露之隱私洩露結果。因此，案件中經處理後的資料（州政府機關人員之醫療資料紀錄、用戶網路搜尋和瀏覽紀錄及平台電影評分紀錄）之去識別化程度，應皆尚未達到 GDPR 下的匿名化要求標準，不符合立法理由第 26 點定義下之匿名資訊，故仍應屬於 GDPR 下的個人資料。

第二項 匿名資訊的界定

匿名資訊的判斷，變相是在界定個人資料的範圍，功能上等同於 GDPR 規範適用與否的啟動閥。在此脈絡下，匿名資訊概念和標準的採擇設置，即是作為二分法資料保護法制下，用以兼顧隱私保護與促進資料有效利用的衡平措施[158]，立意在於透過將個人資料進行匿名化處理，減低資料的識別性至 GDPR 規範所要求的識別性

[158] Ohm, *supra* note 30, at 1738; *Chapter 2 of Draft Guidance*, *supra* note 32, at 12.

doi:10.6342/NTU202401988

門檻標準以下，在確認識別風險大幅降低後，該資料即已非個人資料，是為匿名資訊，該資料的處理活動則不受到 GDPR 的規制，可供自由運用。

第一款 匿名資訊的範圍

GDPR 的規則諮詢機關歐盟個資保護委員會（European Data Protection Board, EDPB）肯認要全然消除資料與個人的連結關係，使原具有識別性之資料，轉變為永久不可回復、識別風險為零的完全匿名狀態，在資通訊科技和重新識別技術進步發展下，在現今的技術層面上已經非常困難[159]；此外，GDPR 對於個人資料也並非是賦予絕對性之保障[160]，因其之立法意旨本在於保障自然人基本權的同時，亦須兼顧促進資料的流通使用[161]。由此，GDPR 並不要求資料須達到，不管在任何情況或假設前提下，皆絕無識別出特定個人之可能性，才符合匿名資料的範圍標準。

GDPR 立法理由第 26 點即有明確說明，在認定個人是否可被識別時，應考量資料控管者或其他之人所有可能合理使用之方法（all the means reasonably likely to be used）；又 WP29 的 2007 年意見書中也進一步強調，若僅是有假設上的可能性（hypothetical possibility）得以辨別出特定個人，尚不符合識別性之要件，關聯於該個人的資料也尚不足以被認定為是個人資料[162]。

詳言之，在識別能力的判斷基準上[163]，雖然識別主體並未限於使用資料之人或保有資料之事業機關內部，而是包含其他任何人，但其他之人仍然必須是實際上存

---

[159] EDPB, *EDPB Document on Response to the Request from the European Commission for Clarifications on the Consistent Application of the GDPR, Focusing on Health Research* [hereinafter *EDPB Document on Health Research*] ¶ 47 (Feb. 2, 2021), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_replyec_questionnaireresearch_final.pdf.

[160] GDPR r.4.

[161] GDPR art. 1.

[162] *Opinion 4/2007 on the Concept of Personal Data*, *supra* note 17, at 15.

[163] 此處涉及識別能力相對性之概念，即有鑑於對資料技術之熟稔程度和所掌握有關當事人其他資料的質與量之差異，將使得連結、組合、比對相關資料而識別出個人之能力，在不同個人、組織、單位間有巨大的落差；比如實際使用資料之人或單位，往往具備較完善之資料處理技術條件和資源以及資料取得權限，故往往較其他一般人更容易可從資料中進行串連、比對，進而成功識別出特定個人。因此，在認定資料是否具有識別性，必須先確認究竟是以具備誰的識別能力作為判斷依據，此即為識別性判斷的主觀標準。有關識別能力相對性之概念和各種標準區分，詳見范姜真媺，個人資料保護法關

doi:10.6342/NTU202401988

在且可能，不能只是假設上存在而已[164]；「可能合理使用之方法」則是指限於有實際可能性的識別方式，對此，前言第 26 點中亦有說明，「可能合理使用之方法」在認定上應考量識別所需的時間和成本，並且參酌資料使用當下現有的科技技術水準，以及技術在資料使用期間中演進的可能性等客觀因素。若是受法律禁止，或事實上需耗費不合理（unreasonable amount）的人力、時間和成本始得識別出特定個人，則非屬於可能合理使用的識別方法[165]。

總結而言，在 GDPR 的規範架構下，匿名資訊與個人資料互為反義詞，非個人資料即屬匿名資訊，反之亦然，故是否為匿名資訊的認定，即同樣適用上述立法理由第 26 點有關個人資料識別性要件的判斷標準[166]。由此，在 GDPR 的標準下，匿名資訊具體而言須是對於任何人而言，都無法透過可能合理使用之方法，將資料內容連結至特定個人，實屬當之。

第二款 匿名化技術有效性的檢驗

歐盟 GDPR 雖肯認匿名化得使個人和社會在享受開放資料（open data）的效益，並確保隱私侵害風險的降低和控制，但也同時提醒資料控管者須留意有重新識別的殘餘風險問題，強調在選擇實施資料匿名化措施方法前，應對於各種去識別化技術的優勢和劣勢有足夠的了解，以確保匿名化作業程序的設計之於個案資料使用活動的合適性[167]。

有鑑於此，WP29 之 2014 年意見書的內容目的，便是針對實務上常見去識別化技術的效用和限制進行分析，提醒資料控管者未來設計實施匿名化程序時，應注意

於「個人資料」保護範圍之檢討，東海大學法學研究，41 期，頁 98（2013 年）；范姜真媺，大數據時代下個人資料範圍之再檢討—以日本為借鏡，東吳法律學報，29 卷 2 期，頁 8（2017 年）；邱忠義，談個人資料保護法之間接識別，月旦裁判時報，30 期，頁 101（2014 年）。

[164] 江耀國、黃子宴（註 24），頁 13。

[165] Case C-582/1, Breyer v. Bundesrepublik Deutschland, ECLI:EU:C:2016:779, ¶ 46 (Oct. 19, 2016).

[166] *EDPB Document on Health Research*, *supra* note 159, at 11 (¶ 45).

[167] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 3.

doi:10.6342/NTU202401988

選用合適有效之措施方法；此外，對於資料控管者而言，其主要在乎的會是應為採取怎樣的匿名化處理措施，才能夠使資料達到 GDPR 對匿名資訊所要求的狀態標準，GDPR 的監管機關也等於是在此份意見書中，示範如何認定一個去識別化技術是否足夠先進，能夠在當前重新識別技術的發展程度下，提供有效的隱私保障，故當中所揭示的匿名化有效性檢驗標準，則也是提供資料控管者作為一項重要的參考依據。

2014 年意見書首先說明，「識別」不單單僅是從資料內容中找出特定個人的姓名或地址等資訊以得知該個人之身份而已，更包含透過標定（single out）、連結（linkability）和推論（inference）之潛在識別方法[168]。由此，WP29 便提出去識別化技術必須要通過這三個潛在識別項目的檢驗，始為歐盟個資保護規範下所認可的有效匿名化（effective anonymisation）：

（1）標定：指將有關特定個人之所有紀錄或資訊自資料集中單獨辨識出來，能夠從資料集中特定出該個人之身分，與其他資料主體辨別開來[169]。舉例來說，即便有將履歷資料集所有檔案中的姓名資訊以編碼取代，但若仍能夠依據其他有關特定個人的敘述特徵中推知出該個人的身分，則該履歷資料集就不是匿名資訊[170]。

（2）連結：指結合、串聯有關於特定個人或特定群體之多個資訊，不問資訊是否出於同一資料來源，藉以拼湊得出有關該個人之身分，將該個人成功識別出來[171]。

（3）推論：係指以關於特定個人之各種資訊作為基礎，得從中推斷、猜測、描繪或預測出有關該個人的細節特徵或特質[172]。舉例而言，有一個關於問卷調查受

[168] *Id.* at 10.
[169] *Id.* at 11.
[170] *Data Protection Guide for Small Business: Secure Personal Data*, EDPB, https://www.edpb.europa.eu/sme-data-protection-guide/secure-personal-data_en (last visited July 15, 2024) [Hereinafter *Data Protection Guide for Small Business*].
[171] *Id.*
[172] *Id.* at 12.

doi:10.6342/NTU202401988

訪者納稅義務資訊的資料集，假設所有 20 歲到 25 歲的男性受訪者均免繳稅，那麼在年齡和性別資訊已知的情況下，便可以推論出特定個人是否為應納稅的特徵狀態，此即為推論的具體示範[173]。在判斷是否有辦法對個人的特徵進行推論時，應考量的資料範圍應包含：從識別資訊已經刪除或泛化（generalised）而不完整的資料集、來自同一資料集中看似無明顯直接關聯性的片段資訊，以及窺探者本身擁有或有合理期待可取得的其他資訊（如公開資訊等）中[174]。

準此，匿名化處理作業措施是否足夠有效之判斷方式，便是去一一檢視資料經處理後，是否仍得從資料集中標定出特定個人、是否仍可能將其他資料中的紀錄內容連結至該個人，以及得否從經處理後的資料內容中推論出有關該個人的特徵等任何資訊。

2014 年意見書對於各該去識別化技術的檢驗分析結果是，常見的去識別化技術中，並沒有任何一項技術方法通過上述有效性標準的檢驗，得以完全有效防免不被標定、連結和推論的潛在識別風險[175]。由此，WP29 最終給出的結論和建議是，匿名化措施的採取必須依具體個案的情況而定，不同資料使用情境下的最適處理方法將有所不同，且通常會須要搭配不同類型的技術方法，以達到最佳有效的匿名化效果；此外，也強調資料匿名化並非為一次性的任務，資料控管者應避免落入「釋出後即遺忘」的錯誤慣習，而是應定期進行風險評估和控制，以確保維持匿名化處理作業程序的有效性[176]。

[173] *Data Protection Guide for Small Business*, *supra* note 170.
[174] *Chapter 2 of Draft Guidance*, *supra* note 32, at 7.
[175] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 23.
[176] *Id.* at 24.

doi:10.6342/NTU202401988

	是否仍存在標定的風險?	是否仍存在連結的風險?	是否仍存在推論的風險?
假名化（Pseudonymisation）	是	是	是
添加干擾（Noise Addition）	是	不太會存在（May not）	不太會存在
代換（Substitution）	是	是	不太會存在
彙集法（Aggregation）/ K 匿名框架（K-anonymity）	否	是	是
L 多樣性（L-diversity）	否	是	不太會存在
差分隱私法（Differential Privacy）	不太會存在	不太會存在	不太會存在
雜湊函數（Hashing）/ 標記化（Tokenization）	是	是	不太會存在

表一：匿名化有效性測試的檢驗結果

資料來源：WP29（2014）[177]，本研究翻譯繪製

第三項 假名化的規範功能

承前所述，假名化的作用目的乃是在於「降低」，而非「去除」個人資料之使用對資料主體產生的風險，故假名化並非能作為 GDPR 規範下獨立之資料處理合法性要件事由，應僅是「得幫助義務主體（即資料控管者和處理者）履行」其個人資料保護義務之手段措施之一。

[177] *Id.*

doi:10.6342/NTU202401988

以下將整理介紹歐盟法所歸納認定的假名化方法種類，並詳細討論假名化在GDPR下的規範功能、措施方法選用上的制度配套措施，以及與資料處理合法事由的適用關係。

第一款 歐盟 GDPR 對於假名化方法的認定

WP29 的 2007 年意見書中提到假名化最常見的作法，就是將資料中如姓名、出生日期和地址等直接顯識別資訊以對應的編碼（code）取代，此種經過密鑰編碼轉換後的資料（key-coded data）[178]因有關個人的明顯識別資訊已被遮蔽，即可避免他人從資料本身的內容便得直接獲知特定個人相關資訊之情形。

嗣後的 2014 年意見書則有對於假名化技術的類型和運作原理，為更全面的描述介紹，其中另將私鑰加密（encryption with secret key）等各式加密機制、雜湊函數（hash function）和標記化（tokenization）劃歸為假名化措施的範圍內[179]。加密是將資訊內容轉換為代碼的技術過程，必須要以正確的私鑰才能夠將資料內容解密，由此確保資料內容不會被未經授權、許可之人接近取用[180]；雜湊函數則是可將任何隨機的內容數值轉換為固定長度字符序列，例如將「Hello」此一字詞帶入函數 SHA256 中可轉換為「10412375520728980495627208923307217150057402816369478709346449899807316016551」此一 256 位元字串的雜湊值[181]；標記化技術常見於金融行業，用來保護客戶信用卡卡號等交易資料不被他人竊取，主要是透過單向加密機制（one-way encryption）或是運用索引函數（index function）分配序列號或隨機生成

[178] *Opinion 4/2007 on the Concept of Personal Data*, *supra* note 17, at 18
[179] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 21-22.
[180] *Data Protection Guide for Small Business*, *supra* note 170.
[181] SHA256(Hello) = 10412375520728980495627208923307217150057402816369478709346449899807316016519。AEPD & EDPB, *Introduction to the Hash Function As A Personal Data Pseudonymisation Technique* at 5 (Oct., 2019), https://www.edps.europa.eu/sites/ default/files/publication/19-10-30_aepd-edps_paper_hash_final_en.pdf.

doi:10.6342/NTU202401988

碼[182]，得出對應的標記值（token value）以取代原始資料值[183]。

WP29 認為這些假名化技術方法的原理，基本上就是透過演算法或函數等轉換程式，將資料中的一個或多個資料屬性值（attributes），如：姓名、出生日期或地址等欄位內容以編碼數字取代呈現，來達到遮蔽資料的效果，但這些類型的處理方法都並不足以單獨使個資轉變為無識別性的匿名資訊狀態[184]。這是因為「加密」與「匿名化」，在技術的功能目的上，即有根本上的不同，「加密」作為一種保全資料安全的措施，其作用在於為人員、設備或硬體 / 軟體組件之間的通信、接取渠道提供保密性，以防範竊聽或意外的資訊洩露事件的發生；相對之下，「匿名化」的功能才是透過防止資料屬性與資料主體間潛在關係的連結，以達到避免對個人識別之效果目的[185]。

第二款 假名化作為一種適當保護措施

假名化技術於控制、降低資料處理風險方面的重要性，具體可見於 GDPR 第 24 條、第 25 條和第 32 條中，所要求資料控管者和處理者應履行資料保護義務之制度規範中。

有鑑於個人資料之運用，對於當事人自由權利課可能造成廣泛而嚴重之侵害風險[186]，GDPR 第 24(1) 條便規定資料控管者應考量資料處理的性質、範圍、情境和目的，以及其他諸多可能和嚴重的風險因素，實施技術上和組織上之適當措施（appropriate technical and organisational measures），以確保其資料處理活動有遵循 GDPR 之規範要求。

---

[182] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 21.
[183] Yash Mehta 著，曾祥信譯，資料標記化：遮蔽資料的新方法，CIO Taiwan，2022 年 10 月 17 日，https://www.cio.com.tw/data-tagging-a-new-way-to-mask-data/（最後瀏覽日：2024 年 7 月 15 日）。
[184] *Opinion 05/2014 on Anonymisation Techniques*, *supra* note 33, at 29.
[185] *Id.*
[186] GDPR r.75.

doi:10.6342/NTU202401988

相較於第24條為一般性規定之性質，第25條則是更具體化地要求資料控管者，在決定資料處理方式之時和進行資料處理的過程中，都必須實施假名化等技術上和組織上措施，以確保資料最小化等處理原則之落實，保障資料主體之權利（第 25(1)條）；此外，資料控管者亦應實施適當的技術上和組織上措施，確保資料處理的作業程序在預設模式下（by default）[187]，僅會就特定目的和必要範圍內的個資為處理（第 25(2) 條），資料控管者若未履行第 25 條之規定將有行政罰鍰之制裁[188]。

第 32(1) 條則是規定資料控管者和處理者，皆應在考量現有技術、執行成本、資料處理之性質目的和風險後，實施包含加密與假名化在內，與資料安全風險相稱之適當技術上和組織上措施。

此處應注意的是，GDPR 逐條釋義文獻有特別強調，同樣是課予應實施技術上和組織上適當措施之義務，但第 24 條和第 32 條兩者所欲著重實現之目的並不相同。第 24 條是歸責性原則（accountability）的具體落實，除了要求資料控管者須遵循GDPR的原則規範的外，且規定資料控管者有義務須揭示其遵循的程序過程[189]，故技術上和組織上措施於此的作用目的，主要在於協助、確保資料處理活動的合法性（lawfulness）[190]；相較之下，第 32 條針對的是資料安全風險，要求實施適當措施以維護個人資料的機密性（confidentiality）、完整性（integrity）和可用性（availability），避免資料遭意外毀損、遺失、更換或是在未經授權的情況下被揭露或近用[191]。

---

[187] 有關第 25(2)條中「預設（by default）」具體意涵，詳見 EDPB, *Guidelines 4/2019 on Article 25 Data Protection By Design and By Default*, at 11 (¶¶ 40-41) (Oct. 20, 2020), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf.

[188] 最高可被處以 10,000,000 歐元或前一會計年度全球年營業額百分之二之行政罰鍰。GDPR art. 83(4)(a).

[189] Kuner et al. eds., *supra* note 146, at 25 ("The principle of accountability enshrined in Article 24 GDPR . . . the principle of accountability places responsibility firmly on the controller to take proactive action to ensure compliance and to be ready to demonstrate how it has done so"); FEILER ET AL., *supra* note 145, at 142-43.

[190] LUKAS FEILER ET AL., *supra* note 147, at 160.

[191] GDPR r. 83; *id.* at 160-61.

doi:10.6342/NTU202401988

「技術上和組織上的適當保護措施」一詞在 GDPR 的條文規定中頻繁出現，EDPB 在針對第 25 條所公布之指引文件中有說明，「技術上和組織上的適當保護措施」所指涉的範圍很廣，可以是任何資料控管者可能在資料處理活動中部署、實施的措施和方法[192]，對個人資料為假名化處理即是其中一項可選用的方法[193]。至於資料控管者具體而言究竟應實施哪些措施，EDPB 表示在措施的選用上，應針對資料處理程序中各該階段環節所著重的資料保護原則，採取對應合適的技術上組織上措施方法[194]。

對此，第 25 條指引文件中有列舉出落實各該資料保護原則[195]，所應採行的措施方法和實行重點，其中建議資料控管者得採取雜湊函數和加密技術，以限制個資遭他人二次利用之可能性[196]；再者，第 25 條指引文件亦強調資料最小化原則與資料的可識別程度有所關聯，倘若後期階段僅需要使用資料內容的統計數值，即可達成資料處理的目的，無須再用到得識別特定個人之逐筆資料，資料控管者應立即將刪除資料或將資料匿名化，若仍須使用具識別性之資料，則應將資料為假名化之處理，以降低對資料主體的侵害風險[197]；資料控管者亦應設置資料刪除和匿名化的內部程序和設施，且確保經匿名化和刪除後的資料不會被重新識別或復原，並應執行相關的測試[198]，此為存儲限制原則之落實；至於完整性和保密性原則，則是要求個人資

[192] *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, *supra* note 187, at 6 (¶ 8).
[193] 除此之外，尚包含將個人資料存儲為機器可讀取的結構化（structured）格式、使資料主體得以介入干預資料的處理過程、提供有關個資存儲的資訊、設置惡意軟體偵測系統、對員工進行網路衛生（cyber hygiene）的基礎培訓、建立隱私和資訊安全管理系統、透過合約約定要求資料處理者應實施特定措施以落實資料最小化原則等。*Id.* at 6 (¶ 9).
[194] *Id.* at 7(¶ 14) ("Article 25 does not require the implementation of any specific technical and organizational measures, rather that the chosen measures and safeguards should be specific to the implementation of data protection principles into the particular processing in question").
[195] 第 25 條設計及預設之資料保護原則具體包含：透明性原則（Transparency）、合法性原則（Lawfulness）、公平性原則（Fairness）、目的限制原則（Purpose Limitation）、資料最小化原則（Data Minimisation）、準確性原則（Accuracy）、存儲限制原則（Storage Limitation）、完整性和保密性原則（Integrity and Confidentiality），以及歸責性原則（Accountability）。*Id.* at 3.
[196] *Id.* at 20 (¶ 72).
[197] *Id.* at 21 (¶¶ 75-76).
[198] *Id.* at 25 (¶ 82).

doi:10.6342/NTU202401988

料和軌跡檔案（logs）應經過雜湊函數、加密等假名化處理，以防範資料外洩等資安風險之發生[199]。

第三款 假名化與資料處理合法性基礎的配合關係

前面已有提到假名化並無法取代第 6 條所規定之個資處理合法性要件，作為獨立的合法基礎，亦無法排除在為處理資料時，應遵循第 5 條各項資料保護原則之義務。假名化在 GDPR 規範制度下的功能定位和規範效果，可從 GDPR 第 89 條有關特殊資料處理活動的規範模式，即該條與第 5 條、第 6 條、第 9 條有關特種資料處理行為及當事人權利行使規定之適用關係，進行具體觀察。

第 89(1) 條規定基於公共利益、科學或歷史研究或統計之目的，而為個人資料之處理者，應實施保障資料主體自由、權利之適當保護措施（appropriate safeguards），以確保資料最小化等原則之實現。又在實現上開目的可行之範圍內，對於資料的處理或進階處理，應以將資料假名化或甚至是匿名化[200]等之方式為之。逐條釋義文獻便有說明，第 89 條規定並不能單獨作為個資處理行為的合法性基礎[201]；EDPB 於 2019 年針對基於健康研究目的個資運用行為適用 GDPR 規定之提問與爭議，所發布的官方回應文件中（以下簡稱「健康研究目的回應文件」），也明確

---

[199] *Id.* at 26 (¶ 85).

[200] 從第 89(1)條條文最後一句規定，若公共利益、科學或歷史研究或統計之目的，可透過對資料為進階處理且不允許或不再允許對資料主體進行識別之方式來實現，那麼則應須以此方式進行資料處理；前言第 156 點也補充說明，若資料控管者已評估過，以不對資料主體進行識別之情況下實現上開目的的可行性，且亦已有採取適當安全措施，便得基於以公共利益、科學或歷史研究或統計之目的為進階處理，從兩段規定和說明內容來看，GDPR 認為在仍得實現公共利益、科學或歷史研究或統計目的之情況下，應將資料進行匿名化處理，由此，雖然本條文僅有列舉假名化為應採取之適當措施，並未有提及匿名化，但應仍得解釋推論出，匿名化和假名化同為本條所要求應實施的適當保護措施。Kuner et al. eds., *supra* note 146, at 1247.

EDPB 於 2020 年所發布新冠疫情期間為科學研究目的運用健康資料之指引中，亦有強調若得使用匿名資料進行科學研究，則應須將資料進行匿名化處理，以合於資料最小化原則和存儲限制原則。EDPB, *Guidelines 03/2020 on the Processing of Data Concerning Health for the Purpose of Scientific Research in the Context of the COVID-19 Outbreak* [hereinafter *Guidelines Scientific Research in COVID-19 Outbreak*], at 10 (¶ 46) (Apr. 21, 2020), https://www.edpb.europ a.eu/sites/default/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_en.pdf.

[201] Kuner et al. eds., *supra* note 146, at 1245.

doi:10.6342/NTU202401988

表示資料控管者若欲蒐集利用個資進行健康研究，除了須依照第 89(1) 條的規定採取保護措施外，仍須滿足第 6 條一般個資處理行為之合法事由（經資料主體同意、為履行契約義務所必要、遵守法律義務所必要、為保護資料主體或他人之重大利益所必要、基於公益執行職務或受委託行使公權力所必要），且須再具備第 9(2) 條所規定處理特種個資之豁免要件[202]。

此外，EDPB 也強調假名化和匿名化的實施，並無法完全排除 GDPR 規範所課予資料控管者和處理者在為資料處理時，應遵循之資料處理原則和保護資料主體之義務[203]。首先，第 89 條和告知義務的關係上，EDPB 在健康研究目的回應文件，以及針對新冠疫情期間為科學研究目的運用健康資料之指引中，均有表示 GDPR 第 13 條和第 14 條[204]課予資料控管者之資訊告知義務，是落實透明性原則之重要制度規範，相關例外事由之條款規定必須嚴格解釋[205]。

原則上實施假名化等安全措施，無法當然免除資料控管者應依照第 13 條或第 14 條之規定，於取得個資和對個資為進階處理前，告知資料主體有關個資處理之相關資訊，確保資料主體可適時主張其權利，但 EDPB 在上述兩份文件中均有特別討論到[206]，若資料控管者是以間接蒐集者之地位，欲將個人之健康資料用於科學研究目的之目的外使用，倘若得證明事實上無可能逐一向資料主體提供資訊，或告知義務之履行須投入不成比例之成本時，始得援引第 14(4)(b) 條[207]之除外規定，在有實施第

[202] *EDPB Document on Health Research*, *supra* note 159, at 12 (¶ 54).
[203] *Id.* at 12 (¶ 49).
[204] GDPR 的告知義務分為直接向資料主體蒐集個人資料（第 13 條），以及自他處（間接）蒐集個人資料（第 14 條）兩種類型。李世德，GDPR 與我國個人資料保護法之比較分析，台灣經濟論衡，16 卷 3 期，頁 87（2018 年）。
[205] *EDPB Document on Health Research*, *supra* note 159, at 9 (¶ 32); *Guidelines Scientific Research in COVID-19 Outbreak*, *supra* note 200, at 8 (¶ 29).
[206] *EDPB Document on Health Research*, *supra* note 159, at 9 (¶ 34); *Guidelines Scientific Research in COVID-19 Outbreak*, *supra* note 200, at 8-9 (¶¶ 33-40).
[207] 原文為：“the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to the conditions and safeguards referred to in Article 89(1) or in so far as the obligation referred to in paragraph 1 of this Article is likely to render impossible or seriously impair the achievement of the objectives of that processing. In such cases the controller shall take appropriate measures to

doi:10.6342/NTU202401988

89(1) 條適當保護措施之前提下，以公告（making the information publicly available）等其他得保障資料主體權利和利益之程序、措施取代之。然而，蓋第 13 條直接蒐集個資之告知義務規定並未設有除外條款，因此，若資料主體以直接蒐集者之地位，即便是基於第 89 條之特殊目的，且有實施第 89(1) 條所要求之假名化等保護措施，仍無例外得減輕或免除應向資料主體為資訊提供之義務要求[208]。

再者，關於目的限制原則第 5(1)(b) 條有規定，個人資料之蒐集目的應特定、明確且合法，不得對個資為原始蒐集目的外之進階處理；而第 89(1) 條為實現公共利益、科學或歷史研究及統計目的所為之進階處理，不應被視為與原始蒐集目的不相容（incompatible）。關於相容性原則，GDPR 立法理由有說明，若處理個人資料不是基於原始蒐集目的者，則資料處理之目的必須在可相容於原始蒐集目的之情況下，始允許為之[209]；故進階處理若未取得資料主體之同意，原則上須依照第 6(4) 條規定之各項要素進行相容性之判斷[210]，其中即包含是否有實施加密或假名化等適當保護措施；而第 5(1)(b) 條後半段則是相容性推定（compatibility presumption）之規定，將基於第 89 條之目的而為個資之進階處理推定與原始蒐集目的相容。

對於第 5(1)(b) 條相容性推定之適用，EDPB 則有強調必須在有實施第 89(1) 條之適當安全措施之前提下，始得援引第 5(1)(b) 條推定將個資用於科學研究目的為進階處理，與原始蒐集目的具有相容性，而不違反目的限制原則[211]；至於具體應實施哪些技術上或組織上之措施，以及合法性基礎的要求，EDPB 則表示尚有賴透過之後

protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available."

[208] *EDPB Document on Health Research*, *supra* note 159, at 9 (¶ 35).

[209] GDPR r.50.

[210] GDPR 第 6(4)條規定應考量的因素包含但不限於： (a) 原始蒐集目的與欲為進階處理之目的間的連結性（link）、 (b) 當初蒐集個人資料之背景脈絡，尤其是資料主體與資料控管者之關係； (c) 個人資料之性質，是否屬於第 9 條或第 10 條所規定之特種個資； (d) 進階處理可能對資料主體產生的後果影響； (e) 是否有實施加密或假名化等適當保護措施。

[211] *EDPB Document on Health Research*, *supra* note 159, at 6 (¶ 20)

doi:10.6342/NTU202401988

的指引文件更清楚地說明[212]。不過，EDPB 在健康研究目的回應文件中已有明確回覆，當援引第 5(1)(b) 條將個人健康資料用於科學研究目的之進階處理時，原先蒐集個人健康資料時所該當之第 9(2) 條豁免事由，將無法自動延續作為進階處理時使用健康資料之合法基礎[213]，也就是說，對於特種個資為第 89 條目的之進階處理時，仍須再額外獨立具備第 9(2) 條之任一豁免事由，始足當之，故第 5(1)(b) 條推定相容性之規定並無法排除或取代第 9 條，作為處理特種個資之合法性基礎，應特別留意。

然而，去識別化措施在 GDPR 下的規範功能，也並非全然僅是作為加諸於資料控管者和處理者之義務負擔，第 89 條有授權歐盟法或由會員國各自訂定內國法，限縮資料主體在 GDPR 第三章中部分權利之行使。第 89(2) 條規定為實現科學、歷史研究或統計之目的而處理個人資料者，歐盟或各會員國得以立法程序限制資料主體行使第 15 條（查閱權）、第 16 條（更正權）、第 18 條（限制處理權）、第 21 條（拒絕處理權）之權利；倘若處理個人資料是基於公共利益之目的者，第 89(3) 條授權得限制之權利，除第 15 條、第 16 條、第 18 條、第 21 條外，則尚包含第 19 條（更正、刪除或限制處理之通知義務）和第 20 條（資料可攜權）。

這裡須強調的是，第 89(2) 條和第 89(3) 條均要求必須在已滿足第 89(1) 條所要求應實施假名化或匿名化等適當保護措施之條件下，才能夠在為實現科學、歷史研究、統計和公共利益之資料處理目的必要範圍內，以立法之方式限縮資料主體上述之權利。至於具體應實施的技術上組織上的措施內容，以及對資料主體的權利限縮程度，則會因各會員國所制定的法律規範內容而在歐盟國家之間有所落差。

第四項 監管機關的糾正權限

從上述的討論可知，在 GDPR 的規範制度下，假名化的實施是保障資料安全和遵循第 5 條之資料處理原則的重要義務設置；而除了採取有效的匿名化方法將個人資

[212] *Id.* at 6 (¶ 21).
[213] *Id.* at 7 (¶ 22).

doi:10.6342/NTU202401988

料轉變為匿名資訊，即可脫離 GDPR 的適用範圍以外，資料控管者亦有義務將資料進行匿名化處理，以符合資料最小化和存儲限制原則之要求。總結而言，匿名化和假名化在 GDPR 下更大部分的目的作用，乃是課予資料控管者和處理者之資料保護義務內容，以確保資料處理活動的合法性，並保障資料主體的自由和權利。

GDPR 第 58(2) 條賦予監管機關（supervisory authority）[214]糾正權限（corrective powers），得採取法所規定之事前和事後救濟手段，確保GDPR的有效遵循。若未確實實施假名化或匿名化等適當保護措施，導致所欲進行的資料處理活動有違反GDPR規定之可能時，監管機關即得向資料控管者過處理者發出警告（warnings）[215]，作為事前監督措施，目的在於阻止違法侵害的實際發生[216]；亦得直接指示資料處理者或處理者，命令其應在特定期間內以特定方式使資料處理活動符合 GDPR 之規定[217]；或甚至得暫時或永久限制對資料之處理[218]。若確實已有違反規則之情事發生，監管機關則得對資料控管者或處理者發出告誡（reprimands）[219]；或是得依個案違反之嚴重情形，單獨或併同上述其他糾正措施[220]，在第 83 條所定之數額範圍內處以行政罰鍰[221]。

第 83(4)(a) 條即規定，資料控管者若未依第 25 條之規定，在設計和預設階段採取包含假名化和匿名化之適當安全措施，或資料控管者或處理者若未依第 32 條之規定實施加密等假名化措施保護個人資料之資料安全，最高可處 10,000,000 歐元或前一

[214] 依據 GDPR 第 51(1) 條之規定，各會員國應設置至少一個個資保護之獨立監管機關，負責監控各會員國境內 GDPR 制度的有效實施和執行，以保障資料處理活動下個人的基本權和自由，以及促進個人資料在歐盟境內的自由流通。GDPR art. 51(1); GDPR r. 117; EDPB, *Opinion 39/2021 on Whether Article 58(2)(g) GDPR Could Serve As A Legal Basis for A Supervisory Authority to Order Ex Officio the Erasure of Personal Data, in A Situation Where Such Request Was Not Submitted By the Data Subject*, at 3-4 (¶ 1) (Dec. 21, 2021), https://www.edpb.europa.eu/system/files/2022-01/edpb_opinion_202139_article_582g_gdpr_en.pdf.
[215] GDPR art. 58(1)(a).
[216] Kuner et al. eds, *supra* note 146, at 946.
[217] GDPR art. 58(1)(d).
[218] GDPR art. 58(1)(f).
[219] GDPR art. 58(1)(i).
[220] GDPR art. 83(2).
[221] GDPR art. 58(1)(b).

doi:10.6342/NTU202401988

會計年度全球年營業額百分之二之行政罰鍰；又對於違反監管機關依上述第 58(2) 條所發出之暫時性或終局性資料處理限制或命令者，監管機關最高可處以 20,000,000 歐元或前一會計年度全球年營業額之百分之四之行政罰鍰（第 83(5)(e) 條、第 83(6) 條）。

在糾正措施的採擇上，行政罰鍰乃是作為確保 GDPR 規制效力的最終執行手段；然而，倘若個案違規情事實屬輕微，處以行政罰鍰將會對於規範主體，施加不成比例的制裁負擔之情況下，監管機關則可能選擇僅發出告誡即為足以[222]，以符合適當性、必要性和比例性[223]。

第二節 美國法：HIPAA 隱私規則

美國目前並沒有如同歐盟 GDPR 般具有通則性效力的隱私和資料保護規範，在聯邦立法層級上，是由各聯邦部門（sectoral）就其所職掌的事業類別、針對不同產業所運用之個人資料，各別制定規範隱私和資料保護之規定[224]，例如：1974 年隱私權法（the Privacy Act of 1974）規範聯邦政府機構系統中個人資料的收集、維護、使用和傳播[225]；2000 年由美國聯邦貿易委員會（Federal Trade Commission, FTC）制定生效，保障兒童網路隱私及個人資料之兒童線上隱私權保護法（the Children's Online Privacy Protection Act, COPPA）[226]；賦予學生本人和家長對於子女教育學習紀錄之存取、請求修改和其他相關控制權限之家庭教育與隱私保護法（the Family Educational

[222] Paweł Hajduk, *The Powers of the Supervisory Body in the GDPR As A Basis for Shaping the Practices of Personal Data Processing*, 45(2) REV. EUR. & COMPAR. L. 57, 69 (2021).
[223] GDPR r. 129 ("each measure should be appropriate, necessary and proportionate in view of ensuring compliance with this Regulation").
[224] *Data Protection Laws and Regulations USA 2023-2024*, ICLG.COM, https://iclg.com/practice-areas/data-protection-laws-and-regulations/usa (last visited July 15, 2024).
[225] 5 U.S.C. § 552a. *Privacy Act of 1974*, OFF. OF PRI. AND CIV. LIBERTIES, U.S. DEP. OF JUSTICe, https://www.justice.gov/opcl/privacy-act-1974 (last visited July 15, 2024).
[226] 15 U.S.C. §§ 6501. *Children's Online Privacy Protection Act*, FED. TRADE COMM'N, https://www.ftc.gov/legal-library/browse/statutes/childrens-online-privacy-protection-act (last visited July 15, 2024).

doi:10.6342/NTU202401988

Rights and Privacy Act, FERPA）[227]等。

此外，美國亦有部分州有針對隱私和資料保護另訂州法[228]，像是加州於 2018 年制訂通過，2020 年 1 月 1 日正式施行之加州消費者隱私保護法（the California Consumer Privacy Act, CCPA），即為美國各州首部一般性之個人資料資料保護法[229]。總結而言，美國的隱私和資料保護規範，即是由聯邦的部門式立法和各州州法組結而成[230]。

在本研究主要討論的去識別化議題方面，美國各聯邦部門立法和州法因所規範資料類型和制度保護目的之差異，對於所保護的個人資料的種類和定義範圍、去識別化資料排除適用規定，以及對於去識別化制度的建置和著墨程度均有所不同，而其中則以 1996 年制定之健康保險可攜性與責任法（Health Insurance Portability and Accountability Act of 1996, HIPAA），就去識別化的措施方法、對應的法律效果和相關配套措施，有較完整的規範設置和主管機關之實施指引，故也是學界進行美國去識別化制度比較法研究之主要參考對象[231]。

由此，就美國法制度的部分，本研究亦將以 HIPAA 為主軸，詳細說明 HIPAA 下的去識別化規定、去識別化措施應如何操作實施，以及制度在規範面及實施面上的缺失。此外，為解決 CCPA 和 HIPAA 在健康資料去識別化方面，所存在的規範適用衝突問題，加州參議院於 2020 年 8 月 31 日正式通過，加州州長嗣於同年 9 月 25

[227] 20 U.S.C. § 1232g. *Family Educational Rights and Privacy Act (FERPA)*, U.S. DEP. OF EDU., https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html (last visited July 15, 2024).

[228] 美國各州隱私和資料保護規範之制定現況，可詳參此網站：Andrew Folks, *US State Privacy Legislation Tracker*, IAPP, https://iapp.org/resources/article/us-state-privacy-legislation-tracker/ (last visited July 15, 2024).

[229] 國家發展委員會，加州消費者隱私保護法(CCPA)規範重點說明，頁 1。

[230] 賈忻蓉，簡析美國 HIPAA 下個人健康資料用於研究之合法要件，科技法律透析，34 卷 4 期，頁 41（2022 年）。

[231] 例如翁清坤，個人資料之去識別化與再識別化風險：法律之觀點，臺大法學論叢，52 期 3 卷，頁 654-663（2023 年）；吳全峰、許慧瑩，健保資料目的外利用之法律爭議－從去識別化作業工具談起，月旦法學雜誌，272 期，頁 56-57（2018 年）；張陳弘（註 7），頁 234-238。

doi:10.6342/NTU202401988

日正式簽署 CCPA 修正案（AB-713 號法案）[232]，限縮 CCPA 的適用範圍與 HIPAA、聯邦人體研究保護通則（Federal Common Rule for Human Research Subjects）和加州醫療資訊保密法（California's Confidentiality of Medical Information Act, CMIA）之規定調和銜接，並增訂數項針對去識別化資料利用和販售行為之義務措施條款[233]，為 HIPAA 去識別化規定中所無。因此，本研究亦將 AB-713 號法案納入研究討論之對象，以更充分、完整地了解美國的去識別化制度的發展趨勢。

第一項 HIPAA 隱私規則的適用範圍

HIPAA[234]是 1996 年制定管理個人健康資訊的聯邦法規，美國衛生及公共服務部（U.S. Department of Health and Human Services, HHS）於 2000 年 12 月發布 HIPAA 的隱私規則（Privacy Rule），由 HHS 下之民權辦公室（Office for Civil Rights, OCR）負責實施和執行[235]，是規範個人可識別之健康資訊之使用、揭露，以及保障當事人控制權利的隱私標準，確保當事人之隱私權，在個人健康資訊自由流通之情況下，能受到充分之保障，兼顧促進資料有效運用、實現高品質醫療照護之福祉。

HIPAA 隱私規則的規範對象是，受規範機構（covered entities）及其商業夥伴（business associates）對於受保護健康資訊（protected health information, PHI）之使用和揭露行為。在規範主體上，受規範機構包含[236]提供健康計畫（health plans）的醫療照護院所、處理醫療健康資訊的中間機構單位（health care clearinghouses）以及所

---

[232] *AB-713 California Consumer Privacy Act of 2018*, C.A. LEGIS. INFO. (Sept. 29, 2020; 2:00 PM), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201920200AB713.

[233] 資策會科技法律研究所，加州消費者隱私保護法修正法案重點說明，2021 年 3 月，https://stli.iii.org.tw/article-detail.aspx?no=64&tp=1&d=8630（最後瀏覽日：2024 年 7 月 15 日）。

[234] Health Insurance Portability and Accountability Act of 1996, Public Law 104-191, 110 Stat. 1936 (Aug. 21, 1996).

[235] U.S. Dep't of Health and Human Serv. [HHS], *Summary of the HIPAA Privacy Rule*, at 1, https://www.hhs.gov/sites/default/files/privacysummary.pdf.

[236] 45 CFR § 164.104(a).

doi:10.6342/NTU202401988

有以電子方式傳輸個人健康資訊的醫療健康服務提供者（health care providers）[237]；而商業夥伴則是指協助或經受規範機構委託，為受規範機構執行有關 PHI 使用和揭露活動的個人、健康資訊傳輸服務機構或其次承包商[238]。

所保護的資料範圍則是PHI，PHI專指由受規範機構或其商業夥伴，以電子、紙本或口頭等任何媒介形式，進行儲存或傳輸的個人可識別性健康資訊[239]。HIPAA 隱私規則中尚有區分另外兩種資料類型，分別是「經去識別化後的健康資訊（de-identified health information）」與「有限資料集（limited data set）」，這兩種資料的使用和揭露，在 HIPAA 隱私規則下分別獲得不同程度上的管制鬆綁[240]。

第二項 HIPAA 隱私規則所規範的去識別化方法

HIPAA 隱私規則對去識別化（de-identification）標準的定義是，健康資訊經處理後，達到「無合理基礎可相信該資訊得用以識別特定個人（no reasonable basis to believe that the information can be used to identify an individual）」之狀態[241]，此時該健康資訊因已不具備個人可識別性，非屬 PHI，可完全脫離 HIPAA 隱私規則的適用範圍。在制度設計上，HIPAA 隱私規則有具體訂明兩種去識別化方法，即「專家認定法（Expert Determination）」和「安全港模式（Safe Harbor）」，允許受規範機構[242]

---

[237] 有關健康計畫（health plans）、處理醫療健康資訊的中間機構單位（health care clearinghouses）和醫療健康服務提供者（health care providers）的詳細定義說明，請詳見 *Summary of the HIPAA Privacy Rule*, *supra* note 235, at 2-3。

[238] 45 CFR § 164.103.

[239] 45 CFR § 164.103; *Summary of the HIPAA Privacy Rule*, *supra* note 235, at 3 ("the Privacy Rule protects all 'individually identifiable health information' held or transmitted by a covered entity or its business associate, in any form or media, whether electronic, paper, or oral").

[240] 但仍可能落入其他州法或聯邦層級隱私和資料保護法規之適用範圍。

[241] 45 CFR § 164.154(a). 原文為："*Standard: De-identification of protected health information*. Health information that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual is not individually identifiable health information."

[242] 受規範機關除了自己執行去識別化程序外，以得授權委由其商業夥伴代其完成法定去識別化方法的實施。HHS, *Guidance Regarding Methods for De-Identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule* [*hereinafter Guidance on De-identification of Protected Health Information*], at 5 (Nov. 26, 2012), https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhsdeidguidance.pdf.

doi:10.6342/NTU202401988

只要完成上述實施任一法定之去識別化方法，即可自由運用該經去識別化之個人健康資訊，無須再遵循 HIPAA 隱私規則有關使用和揭露之限制規定[243]。

第一款 專家認定法

專家認定法[244]是指，需要由一個具備適當知識和經驗之專家（expert），透過一般認可之統計和科學原則方法進行評估後，認定健康資訊經去識別化後，可能被預期之資訊接收者（anticipated recipient）單獨或結合其他合理可取得之資訊，識別出特定個人之風險已降低至非常微小（very small）的程度[245]。此法並沒有特定具體應實施的去識別化方法或作業程序，受規範機構得自由採取任何的去識別化方法，只要有依上述規定之方式通過專家的風險評估，即符合 HIPAA 隱私規則的去識別化要求；此外，HIPAA 隱私規則有要求，須將專家評估使用的方法和分析結果妥善記錄[246]，於 HHS 請求時提供[247]。

OCR 於 2012 年發布了一份 HIPAA 隱私規則去識別化方法的指導文件（以下簡稱「OCR 指導文件」），當中整理有關專家認定法和安全港模式在實施面上的重要問題，並提供詳細的解釋說明，有助於受規範機關對於方法實行的理解和遵循。在專家的資格要求上，OCR 表示，可為統計、數學或其他任何科學領域背景之專業人士，OCR 會主要針對受規範機構所選任之專家，在相關專業領域的學經歷以及健康資訊去識別化技能的實作經驗，進行審查[248]。

再者，OCR 指導文件中亦有提供專家認定法的參考實施流程，供受規範機關作為具體操作上的依循參照。OCR 所建議的實施程序可分為以下幾個步驟[249]：

---

[243] 45 CFR § 164.502(d)(2).
[244] 45 CFR § 164.514(b)(1).
[245] 45 CFR § 164.514(b)(1)(i). 原文為：“Applying such principles and methods, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information.”
[246] 45 CFR § 164.514(b)(1)(ii).
[247] *Guidance on De-identification of Protected Health Information*, *supra* note 242, at 12.
[248] *Id.* at 10.
[249] *Id.* at 12.

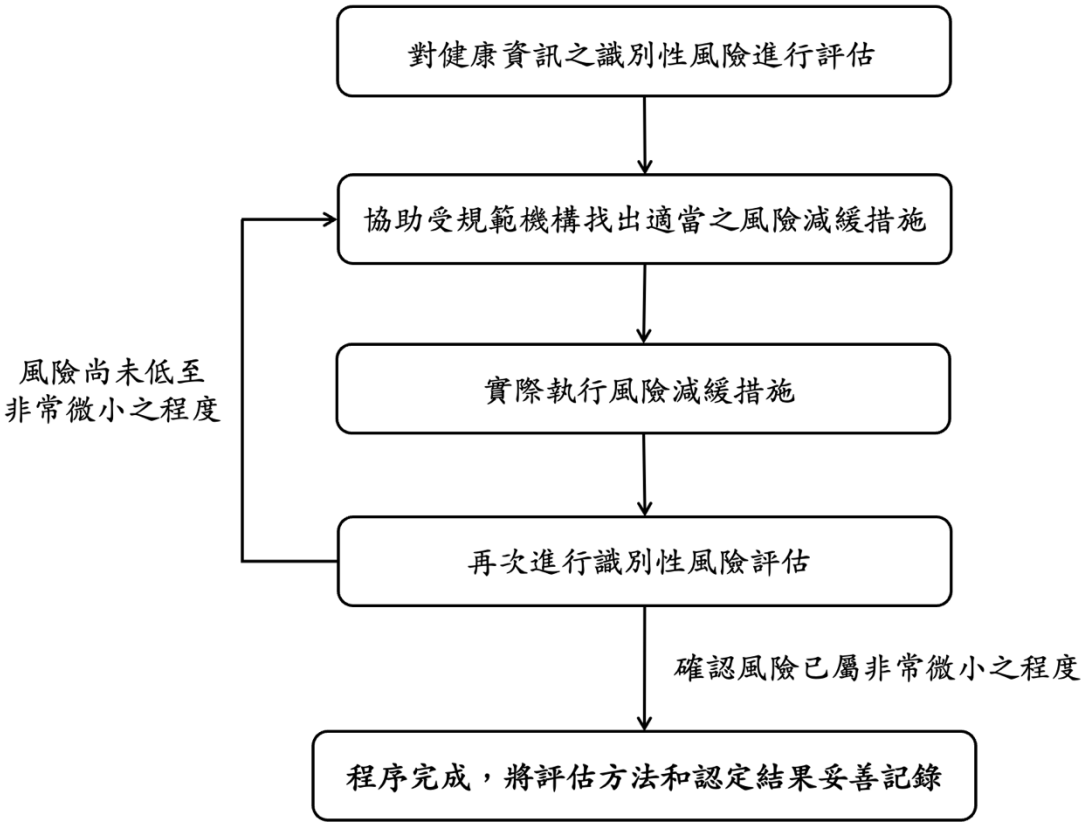
doi:10.6342/NTU202401988

步驟一：專家對健康資訊被預期接收者成功識別之可能性風險程度進行評估。

步驟二：專家向受規範機關或其商業夥伴提供指導，建議可採取哪些統計或科學上方法措施，以有效減緩個案中所使用之健康資訊的識別性風險。

步驟三：經受規範機構或其商業夥伴接受同意後，專家實際執行其所建議得有效減緩識別性風險的措施方法。

步驟四：專家再次評估經處理後健康資訊的識別性風險，確認該風險已有低於非常微小之程度。



圖八：專家認定法操作流程

資料來源：OCR（2012）[250]，本研究翻譯繪製

[250] *Id.* at 13.

doi:10.6342/NTU202401988

從 OCR 給出的參考流程看來，若專家起初認定的識別性風險仍高於「非常微小」的程度，專家得與受規範機構討論後，採取適當處理方法調整個案中的健康資訊，以協助使受規範機構保有的健康資訊符合法規規定之去識別化標準。至於專家所協助採取的作法，通常會是透過調整資料的屬性或數值，以消除資料中的識別資訊和要素[251]，OCR 指導文件中亦有列舉示範幾項常見的技術方法，包含；抑制法（suppression）、泛化法（generalization）、置換法（perturbation）及 K 匿名框架（k-anonymity）[252]。

專家認定法的優點在於，此法是針對每個個案中的健康資料，以專門的措施方式進行個別的識別性風險評估[253]，方法上較為細緻。然而，相較於安全港模式，專家認定法在實務上則並不常被使用，原因在於專家認定法的執行成本較高，且專家的人選不多，要找到符合資格條件的專家不容易[254]；此外，即便已有 OCR 指導文件的說明指示，但專家認定法的實施對於適用者而言，仍有許多模糊之處，像是專家的具體資格條件究竟為何、重新識別的風險究竟要多低，才符合規範所要求的「非常微小」之門檻標準、重新識別風險又應該如何計算和量化，以及統計上和科學上方法又應該如何選用等等[255]。這些適用上的疑義都使專家認定法的執行面上充滿障礙。

第二款 安全港模式

相對於專家認定法，安全港模式的操作就顯得明確、直觀許多。在此法下，只要滿足以下兩項要件，即符合去識別化之要求[256]：一是須先將規則所羅列，有關個

[251] *Id.* at 18.
[252] *Id.* at 19-21. 關於此處所提及技術的概念和運作原理，亦可詳見前述第二章、第二節、第二項、第三款。
[253] Nat'l Comm. on Vital and Health Stat. [NCVHS], *Recommendations on De-identification of Protected Health Information Under HIPAA*, at 8 (Feb. 23, 2017), https://www.ncvhs.hhs.gov/wp-content/uploads/2013/12/2017-Ltr-Privacy-DeIdentification-Feb-23-Final-w-sig.pdf.
[254] *Id.*
[255] Garfinkel, *supra* note 47, at 23; *id.*at 5-6
[256] 45 CFR § 164.514(b)(2).

doi:10.6342/NTU202401988

人之 18 項識別資訊於健康資訊中移除；再者，受規範機構須無「確實知悉（have actual knowledge）」該已移除 18 項識別資訊的健康資訊，仍可被單獨或結合其他資訊識別出資料主體[257]。

HIPAA 隱私規則所規定應移除的 18 項識別資訊包含：1. 姓名；2.「州」以下的地址位置資訊，含街道名、市、郡、區及其相對應之郵遞區號（若是可從人口普查局公開查得的郵遞區號前三碼，則不在此限）；3. 出生日期、入出院日期、死亡日期等所有與個人直接關連的日期資訊（不包含「年度」），以及超過 89 歲之年齡資訊；4. 電話號碼；5. 傳真號碼；6. 電子郵件地址；7. 社會安全碼；8. 醫療紀錄編號；9. 醫療保險受益人編號；10. 帳戶號碼；11. 證照號碼；12. 車輛識別特徵和序列編號，含車牌號碼；13. 裝置識別特徵和序列編號；14. 網址； 15. 通訊協定位址碼；16. 指紋、聲紋等生物識別特徵；17. 全臉部影像照片及其他任何可比擬之影像照片；及 18. 其他任何特殊可識別之號碼、特徵或代碼[258]。以上的識別資訊，乃是大量參考自 Sweeney 博士的研究所設定，尤其是類識別資訊（即間接識別資訊）的辨識[259]。根據美國范德堡健康資訊隱私實驗室 2010 的統計研究顯示，經安全港模式進行去識別化後的健康資料，在美國各州地區間的重新識別風險約介於 0.01%-0.25%之間[260]。

即便已經移除上述 18 項之識別資訊，倘若受規範機構「實際知悉」健康資訊仍可被重新識別，該資訊則仍為受 HIPAA 隱私規則保護之健康資訊。對此 OCR 指導文件有進一步解釋，「實際知悉」是指，清楚且直接地（clear and direct）知道該健康資訊仍可被用來識別特定個人，即受規範機構已心知肚明該健康資訊實際上並非去識別化的資訊[261]；此外，OCR 亦有明確表示，「實際知悉」並不包含受規範機構僅

[257] 45 CFR § 164.514(b)(2)(ii).
[258] 45 CFR § 164.514(b)(2)(i).
[259] Garfinkel, *supra* note 47, at 25.
[260] Kathleen Benitez & Bradley Malin, *Evaluating Re-identification Risks with Respect to the HIPAA Privacy Rule*, 17(2) J. AM. MED. INFORMATICS ASS'N 169, 169 (2010).
[261] *Guidance on De-identification of Protected Health Information*, *supra* note 242, at 27.

doi:10.6342/NTU202401988

是單純知道有某個特定的重新識別方法或研究之情形[262]，而須是對於特定某個資料的「具體知悉（specific knowledge）」[263]。

總結而言，安全港模式的優點是，對於所有的類型、性質的健康資料，都是採取一致、相同的去識別化方法，也就是將上述 18 項識別資訊進行移除，這對於受規範機構來說明顯較簡便且容易遵循，故為受規範機構普遍選擇的作法；然而，基於去識別化的實施往往伴隨資料品質、效用降低的代價結果，本須根據資料的性質、種類和預定用途，選擇適當的去識別化方法，故將安全港模式一體適用於每個個案中，是否皆確實能夠達到去識別化的效果和資料效用的維持間之衡平，恐有疑義。

第三項 去識別化標準的放寬：有限資料集

有限資料集[264]是 HIPAA 隱私規則下，介於 PHI 和去識別化資訊間的第三種資料類型，因其去識別化程度並非完全，故仍屬於 PHI 的概念範圍，但被例外允許得在限定目的範圍內，且滿足特定要件之條件下，得由受規範機構在未經當事人同意授權之下，自由進行使用和揭露。附帶一提，在 HIPAA 隱私規則的規範架構下，除有限資料集外，另外設有三種允許受規範機構，得在未經資料主體授權的情況下對 PHI 為使用、揭露的例外事由，分別是：僅供擬訂研究框架或相類之研究準備目的所須之用（reviews preparatory to research）[265]、基於研究目的使用或揭露已逝者之資訊（research on decedent's information）[266]，或是取得研究倫理審查委員會或隱私審查委員會之審查核准（board approval of a waiver of authorization）[267]。

---

[262] *Id.* at 28.
[263] *Recommendations on De-identification of Protected Health Information Under HIPAA*, *supra* note 253, at 6 ("this requirement does not refer to having general knowledge of studies and methods about re-identification and risks, but, rather, it refers to specific knowledge relating to the particular dataset in question.")
[264] 45 CFR § 164.514(e).
[265] 45 CFR § 164.512(i)(1)(ii).
[266] 45 CFR § 164.512(i)(1)(iii).
[267] 45 CFR § 164.512(i)(1)(i).

doi:10.6342/NTU202401988

有限資料集的設置的目的，是因為有許多研究者反應，安全港模式要求刪除過多的識別資訊，導致資料難以提供流行病學、健康照護服務和其他重人口、地域性等需要資料主體的地址、年齡和期日資訊之研究使用[268]。有鑑於此，HIPAA 隱私規則便在 2002 年最終版本修正文件中，增訂有限資料集作為安全港模式的去識別化放寬版本，但限定受規範機構只能在研究、公共健康和醫療照護營運之目的範圍內，且須有與資料接收者簽訂 DUA 之前提下，才能自由使用和揭露[269]。

在具體的要件上，有限資料集僅要求移除 16 項之個人識別資訊，與安全港模式之 18 項識別資訊相比，允許保留得辨別資料主體之重要期日的完整日期資訊（如出生年月日和入出院日期等），以及其他任何特殊可識別之號碼、特徵或代碼；此外，在地址方面則允許多保留鎮或市、州和郵遞區號之資訊內容[270]。

由於有限資料集的去識別化並非完全，具有更高程度的重新識別風險[271]，故 HIPAA 隱私規則規定，受規範機構必須與有限資料集的接收者簽定 DUA，對控管資料接收者的行為進行約束和控管。DUA 的約定內容具體而言須包含[272]：1. 禁止對資料為 DUA 約定許可範圍或法律規定外的使用或揭露；2. 須實施適當安全措施，預防資料被用於 DUA 約定範圍以外之使用或揭露活動；3. 若有發現資料被用於 DUA 約定範圍外之使用或揭露活動時，資料接收者應即向受規範機構進行通報；4. 須確保其所提供資料的第三方亦同意所有之資料使用行為限制和義務內容；以及 5. 禁止對資料進行識別或聯繫資料的當事人。

此外，2002 年最終版本修正文件中有表示，對於 DUA 的作成形式沒有特定之要求，可以締結正式契約的方式為之，抑或是在合作意向書中明訂上述要求內容，

---

[268] Stacey A. Tovino, *The Use and Disclosure of Protected Health Information for Research Under the HIPAA Privacy Rule: Unrealized Patient Autonomy and Burdensome Government Regulation*, 49 S.D. L. REV. 447, 457 (2004).
[269] 45 CFR § 164.514(e)(1).
[270] 45 CFR § 164.514(e)(2).
[271] 美國范德堡健康資訊隱私實驗室的研究顯示，有限資料集的重新識別風險則約在 10%-60%之間。Benitez & Malin, *supra* note 260, at 169.
[272] 45 CFR § 164.514(e)(4)(ii)(C).

doi:10.6342/NTU202401988

均無不可；若受規範機構欲建立有限資料集供自己內部之研究使用，亦得比照要求掌管敏感資料之員工簽署保密協定的方式，與相關人員簽定相類似的協議即可[273]。

惟此處須強調的是，若資料接收者違反DUA約定的限制，除非該資料接收者亦為 HIPAA 隱私規則下的規範主體，可認定其未履行義務故不符合有限資料集的例外事由規定[274]，否則 HHS 並無法直接對該資料接受者進行懲罰制裁，因其畢竟並非HIPAA 隱私規則的規範對象。在資料接收者不具備受規範機構地位之情況下，亦不表示提供其資料的受規範機構，須對資料接收者違反 DUA 的行為負責，此時，受規範機構須做的是，在知悉資料接收者違約情事後，應採取適當措施停止資料接收者的違約行為、治癒所造成的損害，若情況未見好轉，則應依規定[275]停止繼續為資料之提供，並向 HHS 進行通報[276]。

第四項 HIPAA 去識別化制度的檢討和建議

依照現行的規範，只要受規範機構完成實施專家認定法或安全港模式之方法，其所保有的健康資訊轉變為去識別化資料後，HIPAA 隱私規則就不會再對受規範機構就該資料之後續使用行為和揭露方式，進行任何的監督和管制。詳言之，現行制度幾乎沒有任何防範不當揭露或無正當理由對去識別化資料進行重新識別之措施規定，對於去識別化的健康資料，HIPAA 隱私規則並未有要求，受規範機構須與資料接收者簽定 DUA，以控管資料接收者的資料運用行為[277]；又若受規範機構或其商業夥伴有不當揭露或重新識別之情事，造成對資料當事人之隱私侵害，亦無任何制裁措施或救濟途徑之提供，至多僅有因重新識別使資料回復為 PHI 後，後續的使用和

[273] Standard for Privacy of Individual Identifiable Health Information 67 no. 157 Fed. Reg. 53128, 53236 (Aug. 14, 2002).
[274] 45 CFR § 164.514(e)(4)(iii)(B).
[275] 45 CFR § 164.514(e)(4)(iii)(A).
[276] Standard for Privacy of Individual Identifiable Health Information 67 no. 157 Fed. Reg. 53128, 53236 (Aug. 14, 2002).
[277] *Guidance on De-identification of Protected Health Information*, *supra* note 242, at 22, 29.

doi:10.6342/NTU202401988

揭露行為，須再度受 HIPAA 隱私規則規範之法律效果而已[278]。另一方面，HIPAA 隱私規則亦未提供健康資料的當事人，對於其已經去識別化的健康資訊，有任何知悉權利上的保障，受規範機構後續在對去識別化資料進行使用或揭露前，都不需要對資料當事人為告知。

然而，如同一再強調地，沒有任何去識別化措施得提供絕對性的保障，以專家認定法和安全港模式處理後的資料，仍會有可得識別個人資訊的殘存，且最初所評估實施去識別化後的識別風險，亦會隨時間經過而變動，無法作為永久性的保障[279]。由此，HIPAA 隱私規則完全脫手管制去識別化資料的作法，將會招致的結果就是，已去識別化的健康資料可能被受規範機構、資料接收者或其他第三人任意揭露、與其他資訊進行連結和推論，導致資訊的洩露和個人隱私的侵害；更嚴重的問題在於，這一連串從健康資訊被受規範機構為去識別化處理、後續被受規範機構任意使用或對他人提供，一直到被不當比對發生隱私洩露的整個過程，都可能是在資料當事人不知情的情況下進行和發生，使資料當事人無從及時主張權利或請求救濟。

針對上述 HIPAA 隱私規則的制度隱憂和缺失，HHS 的諮詢機構國家重要健康統計委員會（National Committee on Vital and Health Statistics, NCVHS）於 2017 年提出的建議文件中，有提供相關建議整理如下：

1. 建立後續對去識別化資料的監管措施

有鑑於近年資料技術進步，以及去識別化措施本身在效果和時效上的侷限性，NCVHS 認為現行允許去識別化後即遺忘的制度，對於個人隱私的保障已有所不足[280]。由此，NCVHS 在建議文件中非常強調對於去識別化資料的管理，包含妥善運用

[278] 45 CFR § 164.502(d)(2)(ii); *id.* at 9.
[279] *Recommendations on De-identification of Protected Health Information Under HIPAA*, *supra* note 253, at 8.
[280] *Id.* at 10.

doi:10.6342/NTU202401988

DUA 等契約方式、授權機制等，對去識別化資料的揭露和不當使用進行限制，以有效控制重新識別的風險[281]；以及應採取加密和其他安全措施，對去識別化資料提供安全防護[282]。

再者，NCVHS 也建議，HHS 應評估課予受規範機構和其商業夥伴，應持續追蹤紀錄去識別化資料後續揭露情形之義務，且須依資料當事人之請求，告知有關其去識別化資料後續揭露、使用情形之規範可行性[283]，以及建立對於去識別化資料的安全、重新識別防範和管理措施實施的責任歸屬[284]。

2. 強化告知義務和去識別化程序的透明性

NCVHS 也指出，HIPAA 隱私規則的去識別化規定，在透明性方面的落實上尚有不足之處。目前的規定僅要求受規範機構實施專家認定法時，須將專家的認定結果和評估方法詳實記錄，NCVHS 認為安全港模式亦須設有記錄義務之規定[285]，且兩者方法的實施過程和結果的紀錄內容，都應向資料當事人提供[286]。

此外，NCVHS 也建議即便去識別化資料不在 HIPAA 隱私規則的適用範圍，至少也應該要讓資料主體得以知悉其健康資訊即將被進行去識別化，以及後續將被使用於何種目的[287]。

3. 完善監督管道和提供指導示範

至於 HHS 的職責部分，NCVHS 認為有必要強化 HHS 對於重新識別風險發生的監控，作法上應著重於資訊的即時蒐集，NCVHS 的建議是建立公眾通報系統[288]，任

[281] *Id.* at 13-14 (recommendation 2).
[282] *Id.*
[283] *Id.* at 16 (recommendation 11).
[284] *Id.* at 16 (recommendation 9).
[285] *Id.* at 15 (recommendation 7).
[286] *Id.* at 6.
[287] *Id.* at 16 (recommendation 8).
[288] *Id.* at 16 (recommendation 10).

doi:10.6342/NTU202401988

何人或單位若發現對去識別化資訊的不當使用，可能致生重新識別隱私侵害的疑似情事，可直接通報給 HHS，使 HHS 得以掌握充分資訊、及時介入。

另一方面，NCVHS 也建議 HHS 應設置去識別化執行的資訊平台，用以公布專家認定法和安全港模式的優良實施範例，供受規範機構和商業夥伴相互參照，如此應有助於解決找不到適任的統計和科學領域專家之問題[289]；又應針對去識別化執行人員應具備的能力和訓練程度，提供相關的書面指引和資源管道[290]；此外，HHS 亦應就如何進行重新識別風險評估、如何管控去識別化資料的揭露和去識別化資料的安全防護等面向，提供相關之政策指導和資源工具，以增進受規範機構和商業夥伴有效應對重新識別問題之能力[291]。

第五項 CCPA 的修正法案對 HIPAA 去識別化制度之調和

在州法和聯邦法相互羅織的立法現況下，便經常會衍生出法規適用的疑義，蓋 CCPA 和 HIPAA 隱私規則各有不同對於去識別化資料的定義範圍和認定方式，使得在過去受規範機構即便已符合 HIPAA 隱私規則的去識別化要求而得豁免，卻可能還是須受到 CCPA 的適用管制，造成法規遵循的困境。對此，加州便於 2020 年 9 月 25 日通過 AB-713 號法案，將 CCPA 的適用範圍與 HIPAA 等其他聯邦法和加州州法間進行調和，以解決去識別化健康資料的規範適用問題。

第一款 擴張 CCPA 的豁免範圍

AB-713 號法案對於 CCPA 的修正包含兩個部分：一是擴張 CCPA 的豁免範圍；二是增訂使用或販售去識別化資訊的義務規範。首先，CCPA 修正前的豁免範圍，在資料種類的部分，原先僅及於 HIPAA 隱私規則的規範標的，即受規範機構或商業夥

[289] *Id.* at 14 (recommendation 3).
[290] *Id.* at 14 (recommendation 4).
[291] *Id.* at 14 (recommendation 5).

doi:10.6342/NTU202401988

伴所蒐集之 PHI[292]；在豁免規範對象的部分，原先僅免予規範受規範機構在符合 HIPAA 隱私規則下所為的資料運用行為[293]，商業夥伴則並不在修正前的豁免對象範圍內。

對此，AB-713 號法案新增條款，將 CCPA 的豁免範圍擴張納入原為 HIPAA 隱私規則定義下之受規範機構或商業夥伴所蒐集、製作、傳輸和保有，且已依專家認定法或安全港模式進行去識別化處理之健康資訊[294]；然而，倘若經去識別化資訊後續又被重新識別，便會不再符合豁免的資格條件，須回復受到 CCPA 的規管[295]。另一方面，除了受規範機構外，AB-713 號法案亦將 CCPA 豁免範圍，擴張及於其商業夥伴，其資料運用行為在符合 HIPAA 隱私規則的限度範圍內，亦可排除於 CCPA 的適用範圍外[296]。

第二款 增訂使用或販售去識別化資訊的義務規範

對於 CCPA 此次的修正，本研究認為更值得關注的是所增訂的義務和限制部分。相較於 HIPAA 隱私規則不再介入去識別化健康資訊的任何使用和揭露行為，AB-713 號法案則原則禁止 CCPA 的規範對象，即企業（business），對於依上述規定豁免的去識別化健康資訊，進行或企圖（attempt to）進行重新識別，除非是符合以下之例外事由和目的：1. 經受規範機構委託或依其指示，實施治療、執行付款或醫療照護營運活動；2. 從事公共健康活動；3. 從事健康和人體研究；4. 依與合法保有去識別化資訊者之契約，協助其執行去識別化技術有效性的測試、分析或驗證，且契約內容設有禁止對重新識別資訊為其他使用或揭露，以及執行完畢應返還或銷毀之限制義務條款；或是 5. 應法規之要求[297]。

[292] CAL. Civ. Code. § 1798.145(c)(1)(A).
[293] CAL. Civ. Code. § 1798.145(c)(1)(B).
[294] CAL. Civ. Code. § 1798.146(a)(4)(A).
[295] CAL. Civ. Code. § 1798.146(a)(4)(B).
[296] CAL. Civ. Code. § 1798.146(a)(3).
[297] CAL. Civ. Code. § 1798.148(a).

doi:10.6342/NTU202401988

再者，AB-713 號法案亦有規定任何販售，或將上述去識別化資訊提供授權之行為活動，都必須在契約中載明以下內容和限制條款：1. 所販售或授權的去識別化資訊中，有包含病患的去識別化資料之聲明；2. 本契約禁止買受者或被授權者對去識別化資訊進行或試圖進行重新識別；以及 3. 除法規另有要求外，否則買受者或被授權者不得將去識別化資訊揭露、提供給其他第三人，除非與該第三人之間亦有締結相同或更嚴格的限制條款[298]。此外，修正規定亦要求企業須在其隱私政策聲明內容中，說明是否有將病患的去識別化資訊進行販售或授權之情事，以及所採用的去識別化方法，是依據 HIPAA 隱私規則下的專家認定法還是安全港模式[299]。

附帶一提，加州在 2020 年 11 月 3 日以公投方式表決通過第 24 號提案（Proposition 24），訂定加州隱私權法（California Privacy Rights Act, CPRA）作為加州之新消費者隱私保障規範，主要是在 CCPA 的制度架構和規範內容的基礎上進行調整，針對原先保障不足的部分加強規定，並創設獨立專責機關加州隱私保護署（California Privacy Protection Agency, CPPA），執掌 CPRA 的調查、執行和行政命令的制定[300]，新法於 2023 年 1 月 1 日起生效實施。CPRA 與 CCPA 之規範適用關係起初曾有疑義[301]，直到 CPPA 公告制定加州消費者隱私保護法規則（the California Consumer Privacy Act Regulations, CCPA Regulations），調和現行 CCPA 規範內容與 CPRA 一致，CCPA Regulations 已完成立法程序，於 2023 年 3 月 29 日起正式生效[302]。

[298] CAL. Civ. Code. § 1798.148(c).
[299] CAL. Civ. Code. § 1798.130(a)(5)(D).
[300] 有關 CPRA 與 CCPA 的比較差異，可詳見：張陳弘，美國加州消費者隱私保護法制之最新發展與比較法啟示，當代法律，6 期，頁 26-32（2022 年）。
[301] Cathy Cosgrove, Top-10 Operational Impacts of the CPRA: Part 2 – Defining "Business" Under the Law, IAPP (Dec. 22, 2020), https://iapp.org/news/a/cpras-top-operational-impacts-part-2-defining-business/
[302] *California Consumer Privacy Act* Regulations, CA. PRIV. PROTECTION AGENCY, https://cppa.ca.gov/regulations/consumer_privacy_act.html (last visited July 15, 2024).

doi:10.6342/NTU202401988

第三節 小結：歐盟法與美國法去識別化制度的比較

由上述對於兩制度的整理分析，可顯見歐盟 GDPR 和美國 HIPAA 隱私規則對於去識別化的規範方式有很大的差異。歐盟 GDPR 對於去識別化是採綜合判斷法，斟酌考量識別所須的時間和成本、資料處理技術的水準和科技發展等所有客觀因素，認定資料是否達到資料控管者或其他人，皆無法透過所有「合理可能」使用的方法，識別出特定個人之程度。此等同於是對個案中的資料識別風險進行評估，故學說上又有稱此法為「風險評估法（risk-based approach[303]）」，在不同的資料處理時空脈絡下，所要求達到的去識別化程度會有所不同。

美國 HIPAA 隱私規則則是直接在條文中，明定規範所認可的去識別化方法，只要依照法定方法對資料進行處理，該資料便成為該法下的去識別化資料，對該資料的使用行為便可豁免管制。尤其，相較於歐盟法的個案判斷模式，美國法的安全港模式甚至不區分資料的類型、範圍和使用目的，一律適用同一標準，只要刪除特定的 18 項資訊，即符合去識別化的標準。

兩者之利弊在於，歐盟法個案認定之方式，使所適用的去識別化標準，符合對應個案資料處理的識別風險程度，較得確保對隱私的充分保障，但缺點則是欠缺明確性，導致各監管機構的認定結果常有不一致的問題[304]，資料控管者和處理者也難以遵循應對；美國的安全港模式採取單一的作法標準，雖然對規範適用者而言直觀簡明易操作，亦容易為監督機關檢查判斷，但對於某些較高風險的個案，可能會有隱私保障不足的問題，且無法應對重新識別風險的變動性。

此外，兩規範制度下對於去識別化規範功能的設置，亦呈現明顯的差異。美國 HIPAA 隱私規則允許，即便資料在未去識別化完全之狀態下，只要規範主體滿足額外的義務要件，仍得以在限定目的範圍內，例外允許對半去識別化的資料自

[303] Finck & Pallas, *supra* note 110, at 14.
[304] *Id.*

doi:10.6342/NTU202401988

由運用。同樣是僅達到部分去識別化之情形，歐盟則未在合法性基礎的要求上，對於假名化資料的處理行為有任何的放寬，至多僅能在必要範圍內，限制資料主體少部分權利之行使，且必須是由歐盟或其會員國以獨立作用法有明文規定之方式為之[305]；如同一再強調地，假名化在 GDPR 下的功能目的，實為是資料控管者和處理者履行其資料保護義務所採取的措施，並無法單獨作為或取代 GDPR 第 6 條和第 9 條下，本應具備的合法性要件和豁免事由，也無法免除第 5 條所列各項資料處理原則之遵循義務；此外，資料控管者或處理者在 GDPR 第 25 條和第 32 條下，甚至是有義務必須採取假名化和匿名化等適當安全措施，否則將會面臨行政罰鍰等制裁。

美國 HIPAA 隱私規則與歐盟 GDPR 規範模式之差異，應與兩部法規範之立法時空背景有很大的關聯。美國 HIPAA 隱私規則制定於 2000 年，對去識別化資料所要求之安全港模式和專家認定法方法標準，亦為起初之版本延續至今並未修正調整；歐盟 GDPR 則是將近二十年後，為因應數位化時代下個人隱私侵害風險不斷變動升高之產物。美國 HIPAA 隱私規則為較早期之規範，對於資料保護之規範思維較為單純，認為資料去識別化措施已足以提供足夠之隱私保障，而高度倚重去識別化作為放鬆管制以最大化資料效能價值發揮的政策技術；到了歐盟 GDPR 訂立之時代，在重新識別事件層出不窮下，去識別化早已被普遍認知無法提供完足之隱私保障，使得資料保護則轉向風險評估導向之規範模式，去識別化之規範功能便亦逐漸趨於保守。

[305] 劉靜怡健保資料庫釋憲案專家諮詢意見，頁 18。

doi:10.6342/NTU202401988

第四章 我國個資去識別化規範的現況缺失與重新建構

第一節 我國法規範中之去識別化規定樣態

第一項 我國的資料保護規範概覽

個人資料保護法為我國個人資料保護之通則規範，任何對於個人資料的蒐集、處理和利用行為，皆須符合個人資料保護法之規定要求。除一般性之個人資料保護法外，我國亦就人體試驗、人體研究之實施和檢體[306]資料和衍生物之保存等事宜，訂有人體試驗管理辦法和人體研究法等之規定；此外，為加速對疾病成因和新藥開發等[307]醫療研究進展，提供政府機關、醫事機構、研究單位和法人蒐集運用生物資料[308]、生物檢體之平台，亦有制定人體生物資料庫管理條例，就人體生物資料庫之建置、管理和運用進行規範。

另一方面，我國另有特別針對政治檔案之徵集、整理、保存和應用開放訂定政治檔案條例，使政治檔案得以在兼顧當事人隱私和轉型正義精神下開放利用；又配合國民法官制度上路實施，為保障國民法官、備位國民法官之隱私使其得公正、客觀執行職務，國民法官法第 40 條第 2 項亦授權司法院，制定國民法官個人資料保護辦法，針對國民法官、備位國民法官或候選國民法官之個人資料，具體規範相關的保護方式與程序。

---

[306] 檢體資料，依據人體研究法第 4 條第 2 款之規定，係指「人體（包括胎兒及屍體）之器官、組織、細胞、體液或經實驗操作產生之衍生物質」；人體生物資料庫管理條例則是使用生物檢體一詞，按該法第 3 條第 1 款之規定，係將生物檢體定義為「自人體採集之細胞、組織、器官、體液或經實驗操作所產生，足以辨識參與者生物特徵之衍生物質」，意涵上應同於人體研究法中所稱之人體檢體。

[307] 例如由國衛院所建置之國家衛生研究院人體生物資料庫，即是以肝細胞癌和肺癌之成因研究和治療新藥之開發為目的所建置。國家衛生研究院人體生物資料庫，人體生物資料庫簡介，https://biobank.nhri.edu.tw/info/（最後瀏覽日：2024 年 7 月 15 日）。

[308] 現行之人體生物資料庫管理條例似尚未將生物資料個別納入規範，而為因應精準醫療之發展需求，主管機關衛福部於 2023 年 2 月 23 日所發布之預告修正草案中，修訂生物資料庫之定義，將生物資料庫所得蒐集、處理的資料項目包含生物資料和生物檢體，以擴大其運用，並於修正草案第 3 條第 3 款中增訂生物資料的定義為「人體基因、生物標記、醫學影像、健康紀錄及源自生物檢體之數據、資料及資訊」。

doi:10.6342/NTU202401988

另一個重要的部分是對於金融資料的管制保護規範，金融機構所經手掌握的客戶資料中必定有涉及個人資料的部分，自然適用個資法之相關規定，至於金融法規本身，則是設有諸多對客戶個人資料和交易資料的保密義務規定，要求不得與外部單位、金控子公司彼此間[309]或甚至是單一機構不同部門間[310]任意分享。但這些金融保密義務規定幾乎是零星四散於各金融法規中，規範內容的規劃上則不若個資法全面而完整[311]。

除上述之法規外，個人資料檔案安全維護計畫和處理實施方法亦是建構我國資料保護法規重要的拼圖板塊。有鑑於某些行業，如：銀行、醫院和保險等，因保有大量個資且重要的個資檔案，應負擔較重之資料安全保管義務，由此個資法第 27 條則授權各中央目的事業主管機關得針對特定行業，制定該行業應遵循之個資安全維護計畫和處理方法標準（第 3 項），並得指定特定非公務機關應依據該標準訂定具體之適當安全措施計畫（第 2 項），以防止個資遭竊取、竄改、滅失、洩漏。目前各該目的事業主管機關已有針對諸多行業，制定相應之個資檔案安全維護辦法，例如：金融監督管理委員會指定非公務機關個人資料檔案安全維護辦法、數位經濟相關產業個人資料檔案安全維護管理辦法、醫院個人資料檔案安全維護計畫實施辦法等，行業別範圍尚有大眾運輸業、製造及技術服務業、照護機構和短期補習班等。

[309] 金融控股公司法第 42 條第 1 項有規定：「金融控股公司及其子公司對於客戶個人資料、往來交易資料及其他相關資料，除其他法律或主管機關另有規定者外，應保守秘密。」
[310] 銀行法第 28 條第 4 項有規定：「銀行經營信託及證券業務之人員，關於客戶之往來、交易資料，除其他法律或主管機關另有規定外，應保守秘密；對銀行其他部門之人員，亦同。」
[311] 楊岳平，金融科技時代下金融資料共享法制之發展與限制－評「金融機構間資料共享指引」，台灣法律人，17 期，頁 101（2022 年）。

doi:10.6342/NTU202401988

第二項 我國法下之去識別化規定

第一款 個人資料保護法之去識別化規定

法務部的解釋函令，將去識別化定義為「指透過一定程序的加工處理，使個人資料不再具有直接或間接識別性而言[312]」；個資法和個資法施行細則之條文文字，則並未直接使用去識別化一詞，而是以個資法第 6 條第 1 項第 4 款、第 9 條第 2 項第 2 款、第 16 條第 5 款、第 19 條第 1 項第 4 款和第 20 條第 1 項第 5 款規定中之「資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人」，為去識別化的制度條文規定。

個資法第 6 條第 1 項為蒐集、處理和利用敏感性個資之合法性規定，其原則上禁止對於病歷、醫療、基因、性生活、健康檢查及犯罪前科之敏感性個資之蒐集、處理和利用行為，但公務機關或學術研究機關倘若是基於醫療、衛生和犯罪預防之目的，為進行相關統計研究且已將資料進行處理使其無從識別特定個人者，則得例外允許對病歷、醫療、基因、性生活、健康檢查及犯罪前科之敏感性個資為蒐集、處理和利用；至於一般性個資，第 19 條第 1 項規定，將資料進行處理使之達到無從識別之程度，亦可作為學術研究機構基於公益目的進行統計或學術研究，對一般性個資為蒐集、處理之合法事由，無須再取得當事人之同意。

第 9 條則是間接蒐集個資告知義務之規範，公務機關或非公務機關蒐集非由當事人提供之一般性或敏感性個資[313]時，在對資料為處理和利用前，應對該當事人履行告知義務，但倘若是基於公益之統計或學術研究目的，且已有將資料進行處理使其無從識別特定個人者，則可無須再對當事人為告知。

此外，使資料無從識別特定當事人，亦是公務機關和非公務機關欲對個資進行目的外利用之途徑，第 16 條和第 20 條第 1 項即規定，公務機關或學術研究機關倘若

[312] 法務部(107)法律字第 10703513050 號函。
[313] 敏感性個資準用第 8 條和第 9 條直接和間接蒐集個資的告知規定，個資法第 6 條第 2 項規定參照。

doi:10.6342/NTU202401988

是基於公益目的進行統計或學術研究，則可在有將資料經過處理達到已無從識別特定當事人之條件下，對個資為原始蒐集目的外之利用，無須再逐一向當事人取得同意。

從上述條文的規範內容來看，我國和歐盟 GDPR 在資料去識別化規範功能的制度設計上，似乎有比較大的不同。同樣是為公共利益、學術研究和統計之特殊資料使用目的，資料去識別化在 GDPR 之規範下，僅是必須實施的適當保護措施之一，兒不得排除個資處理的一般性合法性要件以及敏感性個資處理之豁免事由條件；然而，在我國法下，無論是對於一般性或敏感性個資，資料去識別化均可取代當事人同意等其他事由要件，「單獨作為」個資蒐集、處理和利用的合法性基礎。

也就是說，資料去識別化在我國個資法的制度規範下，確實有實質上放寬、減輕對規範主體（即公務機關和非公務機關）資料使用行為之管制和義務的功能效果，或許並未達到如同美國 HIPAA 隱私規則，有明文豁免去識別化資料之程度，但整體而言，本研究認為我國個資法在去識別化制度的選擇上，應較近似於美國 HIPAA 隱私規則的政策立場態度。

第二款 人體生物資料庫管理條例等規範之去識別化規定

我國法對於生物資料、檢體資料和衍生物等生物特徵資料之去識別化規定，即與上述個資法之條文，在用語使用有明顯之不同。相較於個資法中之「無從識別特定之當事人」，人體研究法、人體生物資料庫管理條例和人體試驗管理辦法中的去識別化規定則是使用「去連結」一詞，規定資料經去連結後，得以限制資料當事人退出權和同意權之行使。

「去連結」依據人體研究法第 4 條第 3 款規定，係指「將研究對象之人體檢體、自然人資料及其他有關之資料、資訊（以下簡稱「研究材料」）編碼或以其他方式處理後，使其與可供辨識研究對象之個人資料、資訊，永久不能以任何方式連結、

doi:10.6342/NTU202401988

比對之作業」；人體生物資料庫管理條例對於「去連結」一詞之定義與人體研究法類似，該法第 3 條第 7 款規定去識別化係指「使資料永久無法和可供辨識之個資或其他資訊進行連結、比對之作業程序」。

人體研究法要求在實施研究前，研究主持人須先擬定研究計畫，經倫理審查委員會審查通過（第 5 條第 1 項）；研究主持人原則上，應就通過審查後的研究計畫內容，取得研究對象之同意，除非該研究案符合本法授權主管機關所公告，免取得同意之研究案件範圍（第 12 條第 2 項），其中便包含，是使用自合法之生物資料庫中取得的去連結或無法辨識特定個人之資料進行研究之研究案件[314]。此外，第 19 條亦規定，研究材料（包含人體檢體、自然人資料及其他有關之資料和資訊）在研究結束或保存期限屆至後，原則上必須銷毀，但倘若已取得當事人同意或已去連結，則不在此限（第 1 項）。再者，若已將研究材料去連結，使用該材料超過原先書面同意之範圍時或提供國外研究使用時，則無須再經過報請審查、告知或取得同意之程序（第 2 項和第 3 項反面解釋）。

人體研究法並未對於研究材料的蒐集和處理行為有特別之規定，故此部分應仍回歸適用個資法之規範內容，從上述之規定內容中可知，人體研究法在制度設計上，亦肯認去連結與當事人同意互為擇一、取代關係，只要有將研究材料進行去連結化，則視為已對個人權益提供足夠之保障，即可直接免除資料銷毀義務，且在對研究材料為目的外使用時，更無須再重新進行審查、告知或取得同意之程序。

人體試驗管理辦法和人體生物資料庫管理條例，在立法上也同樣肯認，資料去連結有取代當事人同意之規範功能。人體試驗管理辦法第 14 條即規定「受試者之生物檢體、個人資料或其衍生物，於人體試驗結束後，應即銷毀。受試者同意提供再

[314] 衛福部依據人體研究法第12條第2項之授權，制定得「免取得研究對象同意之人體研究案件範圍」規定：「研究案件符合下列情形之一者，得免取得研究對象之同意：……二、自合法之生物資料庫取得之去連結或無法辨識特定個人之資料、檔案、文件、資訊或檢體進行研究。但不包括涉及族群或群體利益者」，行政院衛生署(101)衛署醫字第 1010265083 號函。

doi:10.6342/NTU202401988

利用者，應經審查會審查通過，未去連結者應再次取得受試者書面同意」。至於人體生物資料庫管理條例，其中第 8 條有關參與者退出權和資料庫設置者資料銷毀義務之規範中，亦規定當參與者要求退出參與生物資料庫時，設置者和自資料庫取得資料之第三人，原則上即應銷毀該參與者已提供之檢體和其他相關資料，除非該資料部分已經去連結或有取得參與者之書面同意繼續使用。

除了義務免除和取代當事人同意之功能作用外，去識別化在人體生物資料庫管理條例中，亦作為一種資料保護措施。在生物資料庫的運用上，第 18 條第 1 項即規定，設置者對生物檢體和相關資料為儲存、運用和揭露時，應以編碼、加密、去連結和其他無法識別參與者身分之方式為之。這種課予規範主體，在個資的使用、揭露過程中，應實施去識別化作業程序之資料保護義務規定，則呼應到歐盟 GDPR 將假名化作為適當保護措施，主要目的在於確保資料處理行為合乎資料最小化、存儲限制等資料處理原則，而非減輕規範主體對當事人之義務，或放寬資料使用合法性要求之功能制度設計。

第三款 政治檔案條例中之去識別化規定

政治檔案條例係於民國 108 年 7 月 4 日制定，並曾於民國 112 年 12 月 8 日進行過一次修正，主管機關為國發會，旨在規範政治檔案之清查、徵集、編纂、整理保存、開放使用等事項。

本法之去識別化規定，主要是設置在政治檔案對外開放申請提供之環節中，政治檔案之申請，根據申請者的身分，分為檔案當事人（第 8 條）和非檔案當事人（第 9 條）申請兩種[315]。以申請人為檔案當事人時，第 8 條第 3 項規定所申請之政治檔案倘若有涉及個人隱私者，檔案局應在指定地點供其閱覽、抄錄，或將涉及個人隱私

[315] 檔案當事人係指「政治檔案中遭逮捕、調查、偵查、起訴、通緝、審判、執行或其他受公權力侵害之人」，政治檔案條例第 3 條第 2 款規定參照。

doi:10.6342/NTU202401988

之部分經分離處理使其無從識別特定之個人後，供其複製，除非該個人已死亡或已同意複製[316]。

至於非檔案當事人之申請，則是依照所申請政治檔案之存在年限和隱私性，分別適用不同管制程度之措施辦理。屆滿 30 年之政治檔案，原則上比照檔案當事人申請之規定，涉及個人隱私之檔案應在指定地點閱覽、抄錄，若須複製，除非該個人已死亡，否則須經過分離處理之去識別化作業「或是」取得該個人之同意（第 9 條第 1 項第 1 款本文）；已屆滿 30 年但未滿 70 年且內容涉及檔案當事人高度個人隱私者，申請閱覽、抄錄和複製皆須檔案當事人已死亡或取得其書面同意，始得提供，「且」檔案局在提供該等檔案前，應將當中涉及檔案當事人高度隱私之部分進行分離處理（第 9 條第 1 項第 1 款但書、第 2 項）；至於未屆滿 30 年但有涉及個人隱私之政治檔案，則必須經過該個人之同意始得在檔案局指定之地點提供閱覽、抄錄或複製，未取得同意則一律不允許提供（第 9 條第 1 項第 2 款）。

從上述之規定內容，可看出政治檔案條例乃依據申請人之身份、檔案的存在年限和內容的隱私程度，設置不同之檔案揭露條件。意即，在檔案當事人為申請人，以及非檔案當事人申請已屆滿 30 年內容不涉及高度隱私之政治檔案，在此二種情況下，政府機關（構）只須滿足將檔案進行去識別化「或」取得當事人同意兩者其中一個條件，即可將該政治檔案供申請人進行複製；在非檔案當事人為申請人之後面兩種情況中，也就是所申請政治檔案之存在年限在未屆滿 70 年，且內容涉及檔案當事人或其他個人之隱私者，政府機關（構）則必須取得檔案當事人或該個人之同意，或法所規定之其他例外事由要件，始得提供申請人進行閱覽、抄錄或複製，故不得僅以有實施去識別化作業，取代作為單獨之合法要件事由。

[316] 從檔案當事人之定義和第 8 條第 3 項之規範內容綜合來看，即可明瞭檔案當事人在概念上並不等同於個資法中之資料當事人，換言之，檔案當事人本人所涉之政治檔案中應亦可能含有關於其他個人之個人資訊，因此，此處才會規定檔案當事人對於其本人所涉之政治檔案，若有涉及其他個人之隱私部分，仍須取得該個人之同意或實施分離保護之去識別化措施，檔案局始得提供檔案當事人對此部分之內容進行複製。

doi:10.6342/NTU202401988

第四款 國民法官個人資料保護辦法中之去識別化規定

國民法官個人資料保護辦法係由司法院會同行政院，依據國民法官法第 40 條第 2 項授權於民國 111 年 3 月 29 日所訂定，是一部我國當前較新的個資保護法規，故在去識別化制度措施的設置和監控機制的權責部署上，有相較於個資法更為完善之處，以促進本辦法去識別化要求之落實。

本辦法要求，司法院應開發代碼、遮隱部分資料等方式自動去識別化之程式提供地方法院使用（第 56 條第 1 項），並應建置個人資料去識別化管理作業程序，內容應包含個人資料去識別化之操作管理辦法和相關風險評估機制（第 56 條第 2 項）；在具體實施上，係規定司法院應下設國民法官個人資料保護委員會，由該委員會負責統籌執行個人資料去識別化作業程序之相關事項，包含去識別化作業程序之研擬，以及評估和驗證去識別化措施之適當性（第 57 條第 1 項），事項辦理之經費概由司法院編列預算支應（第 57 條第 2 項）。

關於去識別化之規定，本辦法針對審理本案之法院及辦理國民法官行政業務之地方法院使用個人資料檔案之行為，在第 17 條第 1 項有規定，各法院在進行內部之行政作業流程中，原則上應將涉及國民法官、備位國民法官及候選國民法官個人資料之公文書，以密件方式處理，除非已將個資部分以代號或其他適當方式遮隱，且遮隱之資料有經隔離保護；此外，在個資處理的注意事項上，第 20 條第 1 項則有要求「法院於筆錄記載提及特定國民法官、備位國民法官或候選國民法官時，應以代號或其他適當方式遮隱其個人資料，且以適當方式隔離保護對照識別資訊」。

另一方面，在個資提供和目的外使用的規範上，則有區分內部（對司法院）提供和對外（對學術、研究機構）提供，兩種資料揭露情形分別規定。在資料內部提供之情形中，依據第 25 條之規定，司法院為推動國民法官制度及進行成效評估時，得向地方法院調取國民法官、備位國民法官及候選國民法官之個人資料檔案（第 1 項）；地方法院在提供該等個人檔案予司法院前，必須先以代號或其他適當方式遮

doi:10.6342/NTU202401988

隱部分個人資料，使資料達到無從直接辨識及聯繫特定個人之狀態，並應將對照識別之資訊進行隔離保護（第 2 項）。

就對外提供之部分，第 24 條第 1 項規定審理本案之法院或地方法院，在提供國民法官案件之新聞予各大媒體時，除非有經過當事人之書面同意，否則在所提供的新聞內容中，不得揭露國民法官、備位國民法官及候選國民法官之個人資料。此外，本辦法亦有類似個資法基於學術目的使用個資之條款，規定司法院得將其所保有之國民法官、備位國民法官及候選國民法官個人資料檔案，提供給符合一定條件之國立研究院、公私立學校和依法設立之學術研究機構，進行國民法官制度研究目的使用（第 26 條第 1 項）；對此，申請者須向司法院提出書面研究計畫，內容除說明研究之概要、目的、方法及欲申請使用之個資種類範圍外，尚須載明保護個資的安全維護計畫、終止使用後之資料處理辦法及相關安全責任聲明（第 26 條第 2 項）；司法院若核准，其所提供之個資，亦須是有經過遮隱，使其無從直接識別及聯繫特定個人，且對照識別資料亦須有經隔離保護（第 26 條第 1 項）；倘若司法院認定申請人有第 27 條第 1 項所列之不合格情事，不予提供個人資料時，仍得斟酌改以提供無從直接或間接識別特定個人之數據（第 27 條第 2 項）。

綜合上述之規定可觀察到，無論於內部處理個資的行政作業流程中，還是對內部或對外提供、揭露個資時，本案法院、地方法院和司法院均應「隱蔽部分個人資訊並對識別資訊進行隔離保護」，以維護當事人之資料隱私。於此，去識別化在國民法官個人資料保護辦法下，乃是作為一種的適當保護措施，近似於歐盟 GDPR 要求資料控管者應實施加密、假名化等措施，以確保資料處理活動符合資料保護原則和合法性之目的旨趣。

又根據不同的資訊揭露模式，除了以媒體為揭露對象時，必須取得當事人書面同意，始得在所提供的資訊中透露個資內容以外，當提供資料的對象是司法院或經申請核准通過之學術、研究機構單位時，本辦法則僅有規定必須先將涉及個資之內

doi:10.6342/NTU202401988

容部分適當遮隱和分離保護，未有提及是否應取得當事人之同意。對此，本辦法在制度設計的態度上似乎是認為，既然已經課予資料去識別化之義務，且另有設置其他審查要件，那麼在提供資料給司法院和學術研究為研究、政策評估之利用時，則無須再取得當事人之同意。

第三項 我國去識別化規定的混亂問題

從上述的法規盤點整理中可明顯發現，去識別化在我國各該資料保護法規中的條文用語均不相同，此便導致我國法制度下，去識別化標準和措施要求的混亂和矛盾。對此，學者吳全峰和許慧瑩便有表示，我國的個資法相關規範，未將假名、匿名、編碼、代碼、隱藏部分資料和去連結等例示之去識別化工具名詞適當定義，亦欠缺對此等例示工具之實施方式、資料狀態和與去識別化概念關聯性之明確說明，使得資料在何種條件下得被視為符合規範要求之去識別化標準出現混淆，進而導致個資保護之混亂[317]。

此外，亦有許多學者指出，我國各該個資保護法規之去識別化規定，本身亦存在混淆匿名和假名概念的規範矛盾問題。個資法施行細則第 17 條將「無從識別特定之當事人」解釋為，指「個人資料以代碼、匿名、隱藏部分資料或其他方式，無從辨識該特定個人者」。對此，學者劉定基和劉靜怡則指出，「代碼」和「隱藏部分資料」從技術面上，仍使資料在經過處理後保有與特定個人連結之可能性，故應屬於「假名化」技術的範疇；「匿名」則是永久、不可回復地切斷資料與特定個人間的連結關係，使資料真正不再得以直接或間接識別特定個人。從而，個資法施行細則第 17 條將兩種在特定、效用強度及有無個資法適用之結果不同的去識別化方法並列，已存在明顯邏輯上的矛盾，使人難以理解個資法所規定之去識別化要求標準究

[317] 吳全峰、許慧瑩（註 231），頁 49。

doi:10.6342/NTU202401988

竟為何[318]。此一個資法去識別化客觀標準的規範模糊矛盾問題，也導致法院在個案中面臨適用法律適用、審理認定上之困境[319]。

「無從識別特定當事人」究竟是要求應將個資去識別化至「匿名化」還是「假名化」的程度，在我國行政機關和憲法法庭間之見解亦有抵觸。法務部之解釋函令是說明，當資料以代碼、匿名、隱藏部分資料或其他使他人無從辨識該特定個人方式，進行去識別化處理後，該資料即已非屬個人資料，無個資法之適用[320]。在法務部此一解釋脈絡下，無從識別特定當事人之標準，所關係的是個資法之適用範圍，即相當於歐盟 GDPR 下對於「匿名資訊」之認定標準。

然而，我國憲法法院的判決結果則是認定，無從識別特定當事人應是指要求，將個資之識別程度降低至「非可直接識別」，使資料不含可直接識別特定當事人之資訊，但仍「可能得間接識別」特定個人之狀態。其表示，「無從識別特定之當事人」之要求程度，從文義上來看，可先明確排除「可直接識別該個人」之資料型態，又再與個資法第 2 條第 1 項有關個人資料之定義內容合併觀察，亦可推知「無從識別特定之當事人」，並非是要求資料之去識別化程度須達到「完全匿名且無還原可能性」之狀態，畢竟倘若資料已永久匿名，即已脫離個人資料之範疇，不再是個資法之規範對象[321]。從憲法法庭上述之理由看來，其所謂「非可直接識別」但仍「可能得間接識別」之狀態，則應較近似於「假名化」之概念程度。

除個資法外，人體生物資料庫管理條例中的去識別化條款，亦被認為存在類似的規範矛盾和模糊問題。本法第 18 條第 1 項條文內容將「去連結」、「編碼」和「加密」之不同資料處理方式予以並列。根據本法的定義，「編碼」係指為以代碼取代可供辨識之個人資訊，使資料達到難以辨別個人身分之作業方式（第 3 條第 5

[318] 劉定基健保資料庫釋憲案專家諮詢意見，頁 9；劉靜怡（註 308），頁 10。
[319] 健保資料庫案裁判見解之分歧，詳見後述第四章、第二節。
[320] 法務部(103)法律字第 10303510410 號函、法務部(106)法律字第 10603513040 號函、法務部(107)法律字第 10703513050 號函。
[321] 憲法法庭 111 年憲判字第 13 號第 40-41 段。

doi:10.6342/NTU202401988

款）；「加密」則是將足以識別參與者個人身分之資料、訊息，轉化為無可識別之過程（第 3 條第 6 款）；對於「去連結」，則是要求應達到永久無法以任何方式連結、比對 5 之程度（第 3 條第 7 款）。

對此，學者張陳弘即有認為，「編碼」和「加密」之方法技術與「去連結」之效用強度相比，應有相當程度之落差，「編碼」和「加密」應僅為去識別化中的「假名化」概念，尚不及於「去連結」所要求，永久切斷資料與當事人連結關係之「匿名化」效果[322]；學者李寧修也同樣認為，此處之「去連結」、「編碼」和「加密」係屬不同程度之去識別化技術，人體生物資料庫管理條例第 18 條在此應是要求，資料庫建置者應依據不同之資料運用行為，採取相應效用程度的去識別化措施[323]；又根據謝銘洋大法官和黃昭元大法官，在健保資料庫憲法訴訟案所提出之部分不同意見書中的內容，兩位大法官更是認為，「加密」技術在性質上應為資料安全維護措施，其功能目的是在於防止資料外洩，此與將資料匿名化或假名化之去識別化措施，是屬於不同的概念範疇[324]，兩者無法相互充當替代[325]。

本研究的理解則是，如同本文第二章中對於去識別化和加密技術的區分說明[326]，在技術層面上，「加密」確實與「去識別化」是屬於不同的概念範疇，「加密」技術的功能目的在於維護資料安全，「去識別化」則是隱私保障；但也因為兩

[322] 張陳弘，健保資料二次使用之個人資料保護立法芻議－111 年憲判字第 13 號【健保資料庫案】判決之回應，輔仁法學，66 期，頁 348-349（2023 年）。

[323] 李寧修，個人資料合理利用模式之探析： 以健康資料之學術研究為例，臺大法學論叢，49 卷 1 期，頁 16（2020 年）。

[324] 對此，黃昭元大法官的說明是，去識別化是要斷開資料與當事人之間的連結關係，將資料轉化為難以辨識、甚至永久無法再連結、比對之狀態；加密的原理則是類似於「加鎖」，也就是透過改寫資料內容，將資料上鎖（如放入保險箱），使他人無法在未經准許下獲知資料紀錄的內容，前者才是針對「資料隱私」之保障機制，後者則是著重於「資料安全」之防護，詳見：黃昭元大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書，頁 6。

[325] 謝銘洋大法官表示，資料經加密後固然在形式上無法辨別其內容，但若沒有先經過匿名或假名之去識別化處理，一旦加密機制被解密或破解，資料的內容即會完整呈現，此時，在無須結合或串連其他資料之情況下，就可以從資料內容中輕易識別特定個人，由此，加密技術可作為資料經去識別化後，下一步實施的資料安全防護措施，但無法取代匿名化或假名化方法，作為一個獨立有效的去識別化措施，詳見：謝銘洋大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書，頁 6-7。

[326] 詳見前述第二章、第二節、第一項、第一款。

doi:10.6342/NTU202401988

者在操作原理上有近似之處，尤其是「加密」和「假名化」均是將資料內容轉化成編碼符號以取代原始值，故加密也經常被認為具有去識別化的效果，像是 GDPR 即將私鑰加密等加密技術列為假名化的一種工具方法[327]，同為資料保護的適當安全措施。

回到人體生物資料庫管理條例第 18 條第 1 項之問題，本研究則認為，本條是在規範生物資料庫設置者，在儲存、運用和對資訊為揭露時，應注意實施的保護措施義務，此即相當於 GDPR 所要求資料控管者，有義務採取匿名化、假名化或加密等適當保護措施，以確保資料最小化、處理合法性等資料保護原則之落實[328]。因此，本條並列此三種在技術性質和效用強度不同之方法措施，參照 GDPR 的制度規定，應可解釋為係要求設置者，在建置、運用生物資料庫而使用、揭露個人資料時，應考量資料的內容種類、資料使用之目的以及對資料當事人可造成的影響結果，採取合適之資料保護措施，以保障資個人隱私和安全。

第二節 從健保資料庫案看個資法去識別化的規範問題

第一項 本案事實背景與訴訟歷程

原告蔡季勳、邱伊翎、施逸翔等八人曾於民國 101 年間，向被告中央健康保險署（以下簡稱「健保署」）告知，拒絕被告將其所蒐集之原告個人健保資料，提供給第三人國家衛生研究院（以下簡稱「國衛院」）和行政院衛生福利部（以下簡稱「衛福部」），用於全民健康保險業務以外之目的。被告健保署則回函表示，其將人民健保資料對外提供使用，係為促進全民健康保險之研究，以提升國內醫療衛生領域之發展，相關行為活動均係依法辦理且有輔以嚴格之資料管理安全措施，以確保健保資料之合理使用。原告不服提起訴願遭駁回後，遂即向臺北

[327] WP29 和 EDPB 僅有明確強調加密和匿名化技術的區分，加密無法使資料達到匿名化的效果，詳見前述第三章、第一節、第三項、第一款。
[328] 例如 GDPR 第 25 條、第 89 條第 1 項之規定，詳見前述第二章、第一節、第三項、第二款。

doi:10.6342/NTU202401988

高等行政法院提起本案訴訟，請求法院撤銷訴願決定和原處分，判命被告健保署作成決定，准予停止將原告健保資料提供給第三人國衛院和衛福部為學術研究之用。第三人國衛院和衛福部則以參加人之地位，輔助被告健保署參與本案訴訟。

臺北高等行政法院於民國103年作成判決（102訴36）駁回原告之主張請求；原告不服原判決提起上訴，最高行政法院判決（103 判 600）以本案審理應適用新個資法之規定而非舊法，原審法院於此有判決適用法規不當之違誤，將原判決廢棄發回；臺北高等行政法院重新調查審理後（103 訴更一 120），仍認定被告健保署提供原告個人健保資料予第三人之目的外利用行為具合法性，判決原告敗訴；經原告再次提起上訴，遭最高行政法院駁回後（106 判 54），本案終局定讞。

第一款 健保資料的流動過程

我國自民國 84 年起實施全民健康保險制度，以衛福部為主管機關[329]，並特設健保署掌理全民健康保險業務[330]，健保署便因協助衛福部辦理全民健康保險業務，取得人民之納保及就醫等資料[331]（以下合稱「健保資料」）。

健保署蒐集健保資料的方式，依據全民健康保險法第 80 條及全民健康保險醫療費用申報與核付及醫療服務審查辦法第 3 條和第 4 條之規定，主要是透過保險醫事服務機構，主動申報和上傳民眾之醫療費用和健保 IC 卡就醫紀錄而來；此外，基於醫事審查需要，健保署亦得請求保險醫事服務機構提供病歷或診療相關證明文件。全民健康保險為強制投保，只要符合資格條件之所有國人都必須參加，根據衛福部之統計，全民健康保險的納保率已達 99.9%之國人[332]。由此可見，健保署在近 30 年來

[329] 全民健康保險法第 4 條規定參照。
[330] 衛生福利部中央健康保險署組織法第 1 條規定參照。
[331] 健保署所蒐集的健保資料具體包含：健保醫療服務申報類總表、明細、醫令檔案、健保承保類檔案、健保卡上傳類檔案、檢驗檢查上傳類檔案（包含醫學影像資料）、虛擬私人網路上傳類檔案、特約醫事機構類檔案、健保給付項目及支付標準類檔案等，憲法法庭 111 年度憲判字第 13 號第 4 段。
[332] 衛生福利部，111 年全民健康保險醫療統計，民國 112 年 12 月 21 日發佈，可取得自：https://dep.mohw.gov.tw/dos/lp-5103-113-xCat-y111.html。

doi:10.6342/NTU202401988

已蒐集累積具相當規模的健保資料，為醫藥衛生領域研究中極具代表性的實證資料。

由此，為促進全民健康保險相關研究、提升國內醫療發展之目的，健保署便將所擁有的健保資料，分別提供予國衛院及衛福部建立、推動資料庫加值利用服務。健保署自民國 87 年起委託[333]國衛院建置全民健康保險研究資料庫，並於民國 89 年起開始對外提供學術研究目的申請使用，在委託期間[334]，健保署會於每年 6、7 月間，於前一年所蒐集的健保資料中，選取可供研究使用的資料檔案部分，交由國衛院建置資料庫和製作各式加值資料檔案。

另一方面，就衛福部與其內部所屬機關健保署之間，健保署亦將健保資料中之健保醫療服務申報類總表、明細、醫令檔案及健保承保類檔案[335]，提供給衛福部統計處，衛福部統計處下之衛生福利資料科學中心（前身為健康資料加值應用協作中心，以下簡稱「資科中心」或「協作中心」）[336]則再從統計處，取得該等健保資料後自行加工建置成資料庫後[337]，開放由符合資格要求限制的政府機關、醫事機構和學術單位申請適用。

第二款 所實施的去識別化作業方法

就國衛院之全民健康保險研究資料庫建置流程（即健保資料由納保民眾、健保署到國衛院，最後對外揭露之流動過程）而言，健保署從各保險醫事服務機構處取

[333] 按健保署與國衛院所簽訂之全民健康保險研究資料庫建置及管理契約第 2 條規定，國衛院受委託之標的為「進行資料加值服務，負責健保研究資料庫之建置、對外提供、開放使用等相關事務與勞務」。

[334] 此委託建置關係已於民國 105 年 6 月 28 日終止，自該日起已不再由國衛院提供資料庫加值服務，憲法法庭 111 年度憲判字第 13 號第 5 段、最高行政法院 106 年度判字第 54 號，旁碼 1292-1301。

[335] 憲法法庭 111 年度憲判字第 13 號第 5 段。

[336] 是衛福部於民國 98 年 1 月 31 日依據國民健康資訊建設計畫所成立，並於民國 100 年 2 月 1 日正式對外營運，旨在作為健康資料提供平台，堆廣資料加值利用以產生具應用價值的集體資訊，增進公共衛生決策品質及相關學術研究統計應用，進而達到增進公共利益和全民福祉之目的，臺北高等行政法院 103 年度訴更一字第 120 號，旁碼 649-661。

[337] 中央健保署 1110411 法規範憲法審查言詞辯論書，頁 8。

doi:10.6342/NTU202401988

得健保資料後，會先去除資料中之姓名內容，並針對病患、醫師和藥師的 ID 以及醫事機構代碼等敏感欄位，以演算法程式進行第一次加密轉換編碼處理後，存入健保署之倉儲資料庫；健保署將健保資料從其倉儲資料庫中提供給國衛院前，則會再將資料進行第二次加密轉換編碼處理，前後共經過兩次加密程序後，始成為國衛院所取得用以建置資料庫之原始資料。上開加密程序皆是由健保署資訊組所進行，其中負責實施加密程序的人員，是以已編寫好的加密執行檔進行操作，與加密程式的開發者為不同人，僅程式開發者知悉加密邏輯，實際操作人員因不知加密邏輯，故無從將資料進行還原[338]。

國衛院在對外提供資料庫加值服務前，亦會另外針對每一申請案別或申請人，再進行個別加密處理，使最終每一申請案別或申請人所拿到資料的 ID 欄位所呈現的均會是不同的加密數值。此外，為提升隱私安全防護，自民國 102 年起亦不再提供病患之完整出生年月日，改為僅提供出生年月[339]。

另一方面，就衛福部之部分，其下之協作中心取得健保資料後，會先將資料中的身分證字號、ID 及醫事機構代碼等敏感欄位進行代碼加密，且不會留存健保署所提供之原始資料。協作中心在將資料庫之資料對外提供前，亦會再進行兩層代碼加密程序，使加密欄位僅呈現流水號代碼；每層加密程序分別由不同的作業人員執行，且每件申請案依其使用目的之不同，有不同之加密方式，以確保無法進行相互比對[340]。

另外，在更審程序中，就健保資料從健保署流動至衛福部過程中的去識別化程序部分，衛福部有補充，現已改為由衛福部派專人至健保署執行加密作業，再將加

[338] 臺北高等行政法院 102 年度訴字第 36 號，旁碼 367-370；中央健保署 1110411 法規範憲法審查言詞辯論書，頁 11-13。
[339] 臺北高等行政法院 102 年度訴字第 36 號，旁碼 314-318。
[340] 臺北高等行政法院 102 年度訴字第 36 號，旁碼 367-370；衛生福利部 1110411 法規範憲法審查言詞辯論意旨書，頁 7。

doi:10.6342/NTU202401988

密後的資料攜回資科中心建置使用，由此使得原始資料在整個過程中均不會離開健保署[341]。

第二項 個資法去識別化條款在本案中的適用爭議與裁判見解

本案涉及健保署將健保資料，提供給國衛院建置全民健康保險研究資料庫，以及提供給衛福部建置資科中心，是否符合個資法第 6 條、第 15 條和第 16 條[342]有關個資蒐集、處理和利用行為之合法性要件。本案之健保資料係屬於個資法第 6 條「有關病例、醫療和健康檢查」之特種敏感資料，除非滿足第 6 條第 1 項下之事由，否則原則上禁止蒐集、處理和利用；又公務機關就個資為蒐集和處理應符合個資法第 15 條之規定，且若欲對個資為原始蒐集目的外之利用，僅得在滿足個資法第 16 條但書事由之下，始為合法。

由此，本案的爭點即是，健保署將健保資料提供給國衛院和衛福部建置資料庫和資科中心，是否該當第 15 條第 1 款所規定之執行法定職務必要範圍內之行為；以及該提供資料之行為，是否是基於公益、學術或統計研究目的所必要，健保資料的去識別化程度是否達到規範所要求之程度，故符合個資法第 6 條第 1 項第 4 款和第 16 條但書第 5 款規定之「基於公共利益為統計或學術研究而有必要，且資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人」要件。

本研究以下將透過本案中所涉及有關去識別化之爭點，觀察我國行政法院究竟是如何在具體個案中，解釋適用個資法「資料經過提供者處理後或經蒐集者依其揭

[341] 臺北高等行政法院 103 年度訴更一字第 120 號，旁碼 956-959。

[342] 在本案新舊法適用之爭議上，原一審法院（北高行 102 訴 36）有認為，因當時新（現行）個資法第 6 條尚未施行，考量本案被告所利用之健保資料係為新個資法第 6 條所規範之敏感特種資料，既然該條尚未施行，自亦不應該適用新法下所劃分規範一般資料之規定（即新個資法第 15 條和第 16 條），而是應回歸適用未予區分一般和敏感特種資料之舊個資法第 7 條和第 8 條之規定，作為判斷資料使用行為合法性之規範依據。
此裁判見解遭到上級審法院（最高行 103 判 600）廢棄發回，最高行法院認為本案之健保資料雖屬新個資法第 6 條「有關醫療或健康檢查」之個人資料，但其蒐集、處理和利用行為之合法性，仍應以新（現行）個資法第 15 條和第 16 條之規定進行判斷。

doi:10.6342/NTU202401988

露方式無從識別特定之當事人」之去識別化規定，認定本案所實施的去識別化作業程序，是否達到個資法所要求「無從識別特定之當事人」之標準程度，據以探知我國司法實務是如何操作個資法之去識別化規定。

第一款 一審法院（北高行 102 訴 36）

一審法院雖認定，本案應適用舊個資法第 7 條和第 8 條有關公務機關蒐集、處理和目的外利用行為[343]之規定，但仍有表示見解認為，因健保資料經加密處理後，已無個人姓名或身分證字號等資訊，相關敏感欄位資訊均以代碼形式呈現，已無從依個人身分連結至特定健保資料，由此肯定健保署提供利用之健保資料，應有符合新（現行）個資法第 16 條但書第 5 款「資料經過提供者處理後或蒐集者依其揭露方式無從識別特定之當事人」之情形。

第二款 更審法院（北高行 103 訴更一 120）

更審法院審理結果肯定，健保署提供健保資料予國衛院和衛福部，經處理後依其揭露方式，已符合「無從識別特定當事人」之去識別化要求程度。

更審法院參酌個資法第 2 條第 1 款個人資料定義之反面解釋，認定去識別化的意涵應係指「透過一定程序的加工處理，使個人資料不再具有直接或間接識別性」。關於去識別化程度之標準，更審法院表示，資料揭露者於資料釋出過程所採取之各項安全維護及風險控管措施，均會影響資料於揭露過程是否已達去識別化程度之判斷；在風險之控管上，公務機關應實施整體風險影響評估，並得以契約方式

[343] 即是對應現行法第 15 條和第 16 條，修正前的舊法條文中尚無「資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人」之規定內容。詳見：舊個資法第 7 條：「公務機關對個人資料之蒐集或電腦處理，非有特定目的，並符合左列情形之一者，不得為之：一、於法令規定職掌必要範圍內者。二、經當事人書面同意者。三、對當事人權益無侵害之虞者。」第 8 條：「公務機關對個人資料之利用，應於法令職掌必要範圍內為之，並與蒐集之特定目的相符。但有左列情形之一者，得為特定目的外之利用：一、法令明文規定者。二、有正當理由而僅供內部使用者。三、為維護國家安全者。四、為增進公共利益者。五、為免除當事人之生命、身體、自由或財產上之急迫危險者。六、為防止他人權益之重大危害而有必要者。七、為學術研究而有必要且無害於當事人之重大利益者。八、有利於當事人權益者。九、當事人書面同意者。」

doi:10.6342/NTU202401988

禁止資料使用者對去識別化資料進行重新識別，和設置其他有關資料使用之限制；更審法院認為，倘若個案的風險程度較低，資料去識別化的程度則可相對放寬。

再者，關於去識別化的主觀標準，更審法院認為，個資經去識別化後是否已達到無從識別的程度，應以「資料接收者」的角度進行判斷。因此，資料經編碼加密處理後，即便資料保有者（即提供者）仍保有解密的工具方法，但只要資料保有者未將解密方法等工具提供給資料使用者（即接收者），使資料使用者無從以其自提供者所取得的資料，再與其他公眾可得之資料進行對照、組合、連結而識別出特定個人，該釋出的資料即可認定為無直接或間接識別性，達到去識別化規範的要求程度。

又健保署和衛福部雖有自承，衛福部是從健保署取回尚未加密的原始資料，再自行進行加密並保留解密金鑰，原告便據以主張，健保署將尚未加密的資料提供給衛福部，並未履行規範所要求健保署應負擔之去識別化義務。對此，更審法院則認為，所謂「資料經過提供者處理後或蒐集者依其揭露方式無從識別特定之當事人」，應是指「提供者處理後」或「蒐集者依其揭露方式」擇一情況達到無從識別特定個人之程度即可，並非要求須「併同」達到。由此，考量既然現行作業方式已改為由衛福部派人至健保署進行加密，再將加密後資料攜回，且衛福部在對外釋出加值資料之前後，均有實施資料使用控管和安全維護措施，故應已足以滿足上開規定之要件。

第三款 最高行法院定讞判決（最高行 106 判 54）

本審法院則認為[344]，健保署所實施的去識別化作業，並未符合個資法「無從識別特定當事人」之去識別化檢驗標準，健保署提供給衛福部的健保資料，應仍屬尚未完全去識別化之個人資料。

[344] 就健保署提供健保資料予國衛院建置全民健康保險研究資料庫之部分，本審法院認為既然健保署和國衛院間的委託建置關係已於 105 年 6 月 28 日終止，國衛院已不再進行健保資料加值服務之提供，

doi:10.6342/NTU202401988

其理由在於，健保署轉交資料前，既然是（改為）由衛福部派專員至健保署執行加密作業，完成加密後再由衛福部將資料攜回，此即顯示衛福部之內部人員也具有還原資料與資料主體間連結之能力，而非僅由健保署「單方」掌握還原能力，並未達到「完全切斷資料內容與特定主體間之連結線索」的去識別化標準，故對於衛福部而言，該等資料應仍尚未去識別化，仍為受個資法保護之個人資料。

此外，法院亦附帶說明，去識別化的功能和作用只在於確保社會大眾看到資料內容時，無法從該資料內容中輕易推知該資料所屬的主體，不及於刻意所鎖定特定主體，主動搜尋與該主體身分相關的其他資訊，再和個人資料進行連結之情形。本審法院認為，後者情況下的隱私權侵犯問題，應歸因於一連串不當之私人資訊探究行為，不在去識別化的設置目的範圍內。

雖然未通過去識別化的檢驗標準，但本審法院仍肯定，健保署所實施的去識別化作業，已使健保資料的可識別性大幅降低，且考量衛福部與健保署具上下隸屬關係、僅有兩機關內部負責相關業務之少數成員具有解密還原之工具，保密流程只要稍加改進，就能符合去識別化標準。經衡量審酌後，本審法院最終仍認為，健保署將尚未完全去識別化的健保資料交予衛福部建置資料庫，並未違反比例原則，因此，在此結論上認定原判決的判斷結論並無違誤，維持原判決、駁回上訴。

第四款 憲法法庭判決（111 憲判字 13）

本案原告於窮盡所有救濟途徑後，認為本案最高行法院定讞判決所適用之個資法規和全民健康保險法等相關規範有違憲疑義，便於 106 年 12 月 5 日向憲法法庭提出釋憲聲請，請求宣告該等法規範違憲。憲法法庭嗣於民國 111 年 8 月 12 日作成 111

並已將自健保署取得的原始資料和光碟資檔等交還健保署，且銷毀所有加值資料檔，不保留備份，自此上訴人（即原告）之請求應業已獲得滿足，無訴訟權能之存在，故本審法院最終以欠缺訴之利益為由，駁回上訴人此部分之請求主張。

doi:10.6342/NTU202401988

年憲判字第 13 號判決，針對個資法第 6 條第 1 項但書第 4 款以及全民健康保險法第 79 條和第 80 條，是否牴觸憲法第 22 條保障資訊隱私權之意旨，為裁判宣告。

判決主文一宣告個資法第 6 條第 1 項但書第 4 款之規定合憲。憲法法庭表示，本條「無從識別特定之當事人」之要求程度，從條文文義和由個資法規範整體觀察，即可知此處之「無從識別特定之當事人」，係指將個資之識別程度降低至「非可直接識別」，使資料不含可直接識別特定當事人之資訊，但仍「可能得間接識別」特定個人之狀態，故認定個資法第 6 條第 1 項但書第 4 款並未違反法律明確性原則。

再者，憲法法庭認為個資法第 6 條第 1 項但書第 4 款，將強制人民容忍其高敏感特種資料供蒐集、處理和利用，限制於僅供公務機關或學術機構為公益、學術研究目的使用之必要範圍，並要求應「至遲於揭露時」實施去識別化措施，將個資轉變為「非可直接識別」特定個人之資料，使一般人採取當時存在的技術和合理成本，在不使用額外資訊下不能識別特定當事人，應已足以大幅降低蒐集、處理和利用個人健保資料等高敏感特種資料，對個人資訊隱私權所造成的侵害，故認定個資法第 6 條第 1 項但書第 4 款之規定內容，已屬對隱私權之最小侵害手段且與其目的相稱，未牴觸憲法第 22 條保障資訊隱私權之意旨。

雖說如此，關於個資蒐用的程序和現行法律規範概況，憲法法庭則表示，從個資法和其他相關法規整體觀察，我國欠缺個資保護之獨立監督機制；且針對本案就健保資料之保存、利用、所應遵循之法律要件和正當程序，以及避免遭濫用和不當洩露之適當防護機制等事項，皆尚未有明確之規定，有違法律保留原則；又就個資目的外利用之情形，並未賦予當事人得請求停止利用之權利，亦與憲法保障資訊隱私權之意旨未合。

總結而言，憲法法庭判命我國相關機關應於三年內建置個資保護獨立監督機制之相關法制，且應針對健保資料庫之資料儲存、處理、對外傳輸和提供利用之相關事項，補足具體法律依據，以及明訂停止目的外利用權利行使之相關規範。

doi:10.6342/NTU202401988

第三項 本案法院裁判見解的整理和評析

第一款 對於個資法去識別化規定的解釋適用

1. 客觀標準：「無從識別特定當事人」認定標準的分歧

一審法院基本上認為，只要資料經處理後「已無從依個人身分連結至特定資料」，也就是只要有將資料中之姓名、身分證字號和其他敏感欄位等明顯識別資訊刪除或遮蔽，即已符合無從識別特定之當事人之要求，而未予考量加密後之資料和其他公開資料串連比對之可能性。

更審法院則強調，在去識別化標準的設置與認定上，應將風險評估的高低程度、有無透過契約限制資料接收者的利用行為，以及資料釋出過程所採取的安全維護和風險控管措施等因素納入考量。更審法院在判決理由中即有明白表示，倘若所評估的風險較低，則可放寬接受以專屬代碼、雙向加密等可被還原之擬匿名化方式[345]進行去識別化，故即便資料保有者仍保有解密之工具方法，只要其未將解密還原等工具提供給資料使用者，使資料使用者無從將其所釋出的資料，與其他公開可得的資訊進行連結、比對、組合而重新識別出特定個人者，該資料即有達到個資法規定之去識別化標準。

更審法院最終則斟酌考量，健保署在將健保資料移交提供前已經雙重加密、資料申請均係經過審查、僅提供部分欄位的資料檔或抽樣提供以遵循最小資料原則，且衛福部訂有相關安全控管之作業規範，限制申請人僅能於實體隔離之獨立作業區進行資料分析，並限制攜帶其他資料防止進行破解比對之行為發生等，確保個資不易外洩後，肯認本案已達到去識別化程度之標準。

---

[345] 擬匿名化資料，係指識別資訊以編碼或別名取代之資料，即相當於「假名化資料」，又可再細分為「不可逆」與「可逆」兩種樣態，「不可逆之擬匿名化」是指以非專屬代碼、單向加密（one-way cryptography）等處理方式，使任何人（包含原資料保有者）均無法透過資料比對重新直接或間接識別出特定個人者；「可逆之擬匿名化」則是指，透過專屬代碼、雙向加密（two-way cryptography）或其他方式進行處理，使處理後的編碼資料仍得由原資料保有者，以代碼與原始識別資料對照表或解密工具還原為識別資料之情形，詳見：林裕嘉，公務機關利用去識別化資料之風險評估及法律責任（上），司法周刊，1852 期（2017 年）。

doi:10.6342/NTU202401988

最高行法院定讞判決之裁判見解本身即有自相矛盾的問題。其一方面表示，去識別化作業應達到「完全切斷資料內容與特定主體間之連結線索」之程度，始謂符合去識別化之標準，隨後又駁回上訴人之主張，認為去識別化的功能，只是在確保「大眾無法從資料內容中輕易推知該資料所屬的主體」，從而認定，上訴人所主張「必須達到將其他公開資料與資料庫中之一筆資料進行對照、組合、連結亦無法建立特定個人與該筆資料之連結」之程度，已脫離去識別化之法定標準，兩種闡釋應有矛盾[346]。再者，最高行法院定讞判決同更審法院之判決見解，均肯認若僅有資料提供者單方掌握解密還原方法，仍符合去識別化之要求，但既然已知資料內容必定可由部分人士進行還原，又何以能算是已「完全切斷」個人與資料內容間的連結？由此足證，前述所提到個資法去識別化的規範模糊矛盾問題，已導致司法機關在為具體裁判時，就「無從識別特定之當事人」之規定已有難以適用判斷之情形[347]。

2. 主觀標準：以誰的能力作為無從識別的判斷基準

本案原告即上訴人主張，若其將特定個人之資訊收入資料庫進行查詢，最終會僅有一人符合查詢條件，進而即可推知出該特定個人之身分，由此可佐證經加密作業後之健保資料仍有間接識別性，並向法院聲請調查證據，請求至國衛院和資科中心對健保資料檔之資料進行查詢測試。對此，一審法院、更審法院及最高行法院定讞判決均駁回原告即上訴人此一調查證據之聲請，一審法院及更審法院認為，原告所主張在已知特定個人之身分前提下，將所知的其他資訊輸入進行確認之調查方式，與一般社會大眾事前未先知悉特定個人之身分情況下之隱私曝露情形有異，故未予採納原告即上訴人之主張；最高行法院定讞判決則表示，上訴人所述情形之防範並非去識別化效用強度之法定標準範圍內，因去識別化的功能目的僅在於確保社會大眾無法從資料內容輕易推知該資料所屬之主體。

[346] 劉定基（註 321），頁 11。
[347] 同上註；劉靜怡（註 308），頁 11。

doi:10.6342/NTU202401988

原告即上訴人所主張之調查方式所採取的是較嚴格的主觀標準，其認為應以已掌握特定主體之其他相關資訊之人（可能是特定個人之親友、資料使用者組織內部人士或其他具備特殊資料技術能力之人）作為識別主體，來判斷資料去識別化後是否已達到無從識別之程度；至於上述法院判決之回應，則是採取最寬鬆之一般人識別能力標準，也就是只要從未有或未能取得額外資訊之人的角度看來無從識別即可。

3. 去識別化義務的責任主體

本案另一個重要爭點是，究竟是健保署還是衛福部，負有履行將健保資料進行去識別化的義務，此涉及「經過提供者處理後或經蒐集者依其揭露方式」應如何解釋之問題。

原告主張，此處之「提供者」和「蒐集者」均應係指健保署，健保署作為義務責任主體，必須在將健保資料提供給國衛院和衛福部前，即將資料去識別化至無從識別之程度，義務才有履行完成，因此，健保署竟將尚未加密之資料提供給衛福部，即是違反其應負擔之去識別化義務。更審法院見解則是認為，「蒐集者」在本案應是指間接蒐集健保資料之衛福部，故按本條文之規定方式，去識別化義務由健保署時或衛福部擇一履行即可，故即便健保署在轉交健保資料時，尚未執行去識別化作業，只要衛福部將加值資料檔最終對外揭露提供時，有將健保資料去識別化至滿足無從識別之要求程度，去識別化義務即算圓滿履行。

我國行政機關的解釋函令見解，亦與上開更審法院之見解相呼應。法務部認為個資法第 16 條但書第 5 款之意旨，並非是要求「公務機關所保有之個資一律均須達無從以直接或間接方式識別該特定個人者後，方得提供予其他公務機關或學術研究機構，重點在於統計或學術研究成果發表時，依其呈現或揭露方式必須已無從識別

doi:10.6342/NTU202401988

特定之當事人[348]」，換句話說，倘若經過提供者已將直接識別資料加工變成間接識別資料，再提供給研究機構進行統整分析，該機構嗣後是以無從識別特定個人之狀態方式將研究成果公開，該目的外利用行為即為合法[349]；國發會也將此處之「蒐集者」解釋為是自公務機關取得資料之一方，其表示個資法第 16 條但書第 5 款所稱之「依其揭露方式」無從識別特定之當事人，重點在於規範資料接收者揭露、呈現統計研究成果時，須使資料處於無從識別特定個人之狀態，即足以保障當事人之權益[350]。

由此，實務見解基本上均認為，「經過提供者處理後或經蒐集者依其揭露方式」之規定意旨，係要求資料至遲必須在蒐集者（在跨單位資料串接之情境下，即是指資料接收者）對外揭露研究成果時，呈現無從識別特定個人之狀態。因此，在此意義下，只要提供者移交資料時或蒐集者（資料接收者）公開研究成果時，擇一將個資進行去識別化處理達到無從識別之程度，去識別化義務即算履行，目的外利用之行為即為適法。

第二款 健保資料庫案的後續評論意見

總結而言，本案一審法院、更審法院和最高行定讞法院，最終均肯認健保署得在未告知、取得人民同意之情況下，基於醫療研究發展和公益之目的，將人民之個人健保資料提供給國衛院和衛福部建置資料庫使用。

關於本案裁判法院上述對於個資法去識別化規定的解釋適用，學者劉靜怡和張陳弘即有批評，只要至遲於資料接收者對外公開揭露資料前，有將健保資料完成去識別化處理即可之裁判見解，無異是允許資料提供者（即原蒐集、保有個資之機關單位，本案之健保署），可將未經處理、尚可直接識別之資料，提供給「蒐集者」

[348] 法務部(104)法律字第 10403508020 號函。
[349] 法務部(105)法律字第 10503510730 號函。
[350] 國家發展委員會(110)發法字第 1100017815 號函。

doi:10.6342/NTU202401988

（即欲進行統計研究之其他公務機關和學術研究單位，本案之國衛院、衛福部、甚至是申請者），只要最終對外揭露前，有任一方將資料轉變為無從識別之狀態即可，此種混亂的解釋方式，恐將會對人民的資訊隱私權造成嚴重的侵害[351]。

再者，本案所採取的去識別化方法，即刪除姓名資訊以及將病患、醫師和藥師的 ID 以和醫事機構代碼等敏感欄位進行編碼加密處理，充其量也僅相當於達到「假名化」的去識別化程度而已。然而，本案裁判見解卻仍允許健保署可將半完成加密程序的健保資料，提供給國衛院和衛福部建置大型資料庫，開放提供給公眾申請使用，先不論對照歐盟 GDPR 規範下，假名化根本不得作為獨立運用個資的合法性基礎，即便與美國 HIPAA 隱私規則相比，其所要求豁免管制應達到的去識別化程度是必須刪除高達 18 項之識別資訊，亦比本案中僅針對部分幾項識別資訊進行有還原可能性的編碼加密，來得嚴謹許多。

在去識別化工具的分析評論上，學者吳全峰在健保資料庫憲法訴訟案所提出的專家意見中更說明，本案雖然有採取多道加密程序，但均只是針對特定欄位重複進行加密，這只能表示，該特定個欄位資訊的資料安全已達一定程度之保障，但也僅止於使第三人要還原其內容的難度提升，並無法排除第三人可在無須還原該等欄位內容的前提下，可透過其他資料比對方式重新識別出特定個人[352]；此外，其更指出美國受規範機構的資料庫僅存放單一機構個別的資料，各機構的資料庫間無法相互串接互通，然而，台灣單一保險人體制下的健保資料庫則與這種封閉型資料庫不同，其所涵括的是全國人民的健保資料，而非僅限於單一醫事機構下的有限資料範圍，且在利用層面上，亦允許重大傷病檔案進行內部串接，衛福部資科中心所提供

[351] 劉靜怡（註 308），頁 11；對此，亦有學者見解認為，欲對資料進行目的外利用之人，即應負有將資料去識別化的義務，故本案應以健保署作為將健保資料進行去識別化之義務人，詳見：張陳弘，國家建置全民健康保險資料庫之資訊隱私保護爭議－評最高行政法院 106 年度判字第 54 號判決，中原財經法學，40 期，頁 208（2018 年）。

[352] 吳全峰健保資料庫釋憲案專家諮詢意見，頁 3。

doi:10.6342/NTU202401988

之檔案亦得供與其他資料庫進行比對[353]。由此，我國之健保資料庫實具有更高的隱私風險疑慮，但本案法院裁判見解對於健保資料的去識別化要求，卻遠不及美國 HIPAA 隱私規則要求的嚴謹標準。

至於規範面的後續檢討上，本案原告於窮盡所有救濟途徑後，認為本案最高行法院定讞判決所適用之個資法規和全民健康保險法等相關規範有違憲疑義，便於 106 年 12 月 5 日就向憲法法庭提出釋憲聲請，憲法法庭嗣於民國 111 年 8 月 12 日作成 111 年憲判字第 13 號判決，宣判個資法第 6 條第 1 項但書第 4 款之規定合憲。

憲法法庭認為，本條規定既然已將得強制人民容忍其高敏感特種資料供蒐集、處理和利用之情形，限制在僅供公務機關或學術機構為公益、學術研究目的使用之必要範圍，且有要求應「至遲於揭露時」實施去識別化措施，將個資轉變為「非可直接識別」特定個人之資料[354]，如此應已足以大幅降低蒐集、處理和利用個人健保資料等高敏感特種資料，對個人資訊隱私權所造成的侵害，故認定個資法第 6 條第 1 項但書第 4 款之規定內容，已屬對隱私權之最小侵害手段且與其目的相稱，未牴觸憲法第 22 條保障資訊隱私權之意旨[355]。

本案專家學者之意見則多是反對，認為本案行政機關對於民眾健保資料，進行「強制性、大規模、長時間、持續性」之蒐集、處理和利用[356]，對於人民隱私權可能致生的侵害風險，恐是一般個資之使用行為所無可比擬，係屬對隱私權干預和權利限制程度極高之情形[357]。然而，現行個資法之規範架構卻未有區分不同的學術研究類型、資料的性質和資料使用的規模和方式型態，一概僅以對外揭露資料前為去識別化處理之義務，作為唯一之隱私保障機制，未針對重新識別風險之防範，課予公務機關或學術研究單位應遵循之行為義務，如實施事前風險評估程序、禁止將去

[353] 同上註，頁 7。
[354] 憲法法庭 111 年度憲判字第 13 號第 53-54 段。
[355] 憲法法庭 111 年度憲判字第 13 號第 59 段。
[356] 劉定基（註 321），頁 3-4。
[357] 劉靜怡（註 308），頁 6。

doi:10.6342/NTU202401988

識別化資料與其他資料進行比對還原等，更未明確賦予當事人拒絕、請求退出之權利，故難謂現行個資法係對人民之資訊隱私權，在如本案中大規模強制使用高敏感性特種個資之情況下，已提供適足之實體法保障可正當化對人民基本權之干預[358]。

第三節 我國去識別化制度的完善建構

第一項 我國個資去識別化制度的問題梳理

第一款 去識別化與當事人同意要件的關聯

從本章第一節對於我國去識別化規定的盤點整理中，可發現無論是個資法或其他有關資料管理使用的法規範中，皆存在得以資料去識別化於限定條件下，取代或放寬當事人同意，作為義務減輕、放寬管制之規定設計。

像是個資法在個資的蒐集、處理、利用及個資目的外使用之合法性規定中，將「使資料無從識別特定之當事人」與取得當事人同意並列為不同之基礎事由，且允許個資的間接蒐集人，在有將資料去識別化處理之情況下，對該資料後續為處理和利用時，可無須再對當事人為告知；人體研究法亦放寬使用去連結或無法辨識特定個人之資料的研究案件，經倫理審查委員會審查通過後，可免再取得研究對象之同意；又人體研究法、人體試驗管理辦法和人體生物資料庫管理條例中，亦均有規定只要擇一滿足資料去連結化或取得當事人之同意，即可例外免除資料銷毀之義務；政治檔案條例和近年制定之國民法官個人資料保護辦法，於資料揭露之相關規定中，亦有類似的放寬機制。以上的規定內容即顯示出，我國立法者在一定程度上，似乎是肯認透過資料去識別化大幅降低資料的識別風險，已足以對個人隱私提供足夠之保障。

[358] 劉定基（註 321），頁 14-18；劉靜怡（註 308），頁 11-14；李寧修健保資料庫釋憲案專家諮詢意見，頁 17-18；范姜真微健保資料庫釋憲案專家諮詢意見，頁 12-13。

doi:10.6342/NTU202401988

在制度設計上，是否應允許去識別化具有放寬當事人同意之功能，學者張陳弘從多面向資料隱私保護利益的理論觀點切入，認為資料去識別化的目的，主要是保護個人資訊隱私權中的秘密性或親密性利益，同意權之賦予則是保障當事人對其資料的控制利益，既然兩者的保護利益有別，就不應當然認為兩者處於擇一、取代或相互連動之關係，而是只要去識別化後的資料仍屬個人資料，當事人對其個資的控制利益就應該要被尊重和保護[359]。

此外，學者樓一琳和何之行也表示，依照大法官釋字第 509 號、第 535 號、第 585 號和第 603 號有關隱私權的解釋意旨，我國憲法係以「保障個人資訊自主權」為原則，以「符合比例原則之限制」為例外[360]，然而，我國無論是在規範本身的內容或是司法實務運作的適用見解，現況下卻均認為只要有採取一定程度的去識別化處理和管理措施，即可限制當事人事前同意權之行使，又無明文賦予當事人事後退出之權利，此已使當事人之資料自主控制權幾乎喪失，與大法官所要求基本權之限制應符合比例原則之憲法精神，恐有違背[361]。

從資訊隱私的憲法保障意旨來看，任何人欲對他人的資料為蒐集或使用時，原則上均應得到該他人之同意，始得為之，此為憲法對於個人資訊自決權的保障，但大法官於釋字第 603 號解釋理由中也表示，憲法對資訊隱私權的保障並非絕對，得在符合憲法比例原則之條件下，透過法律明確規定適當進行限制；也就是說，我國採行以資料去識別化放寬當事人同意之制度機制，除了須有法律依據外，在條文內容中，亦必須明訂、限制資料使用之目的範圍，且使用資料的手段方法上不得損益失衡[362]。

---

[359] 張陳弘（註 354），頁 239-243。
[360] 樓一琳、何之行，個人資料保護於雲端運算時代之法律爭議初探暨比較法分析：以健保資料為例，臺大法學論叢，46 卷 2 期，頁 412-414（2017 年）。
[361] 同上註，頁 415-416。
[362] 李建良（註 1），頁 26。

doi:10.6342/NTU202401988

我國的去識別化條文，雖然有特定資料使用之目（例如個資法中僅限定於公共利益、學術、統計研究、醫療、衛生或犯罪預防等），惟除了要求規範主體應對資料進行去識別化處理外，似乎就沒有其他要求應實施或遵循的配套管制措施。因此，本研究認為，我國規範制度的問題或不在於「是否」允許限制當事人行使同意權，而是允許限制同意權的手段措施並不合乎比例原則，不足以正當化對當事人資訊自決權的干預限制。

第二款 去識別化的標準設置不明確

我國制度下，資料去識別化實質上有放寬對資料使用行為的管制以及義務減輕的規範效果，那麼在法規中明訂具體、可預見、易被遵循的去識別化門檻標準和判斷依據，即是此一制度得有效實施、運作的基石。像是美國 HIPAA 隱私規則，即透過法定去識別化方法的設置，尤其是安全港模式下明確列出應刪除的識別資訊項目，雖然有見解指出對於不同的資料性質、預定用途目地，不適合一體適用相同的去識別化措施[363]，但本研究認為，從制度可行性的觀點來看，安全港模式作為一個已相對較直觀、易為資料使用者所遵循實施，且使監督單位事後容易進行檢核判斷的標準方法，即便有其缺失和不足，但至少已經可以確保對個人隱私皆達到一定效果程度的保障。

然而，綜觀我國各法規中的去識別化規定條文，對於各種去識別化用語多缺乏明確的定義內涵，甚至在同一條規定中，有混用假名化與匿名化概念的標準矛盾問題；憲法法庭和解釋函令對於資料去識別化後，效果究竟是放寬適用抑或是直接不再適用個資法之規定，也出現上下見解不一致的情況；至於健保資料庫案中，歷審法院對「無從識別特定當事人」的解讀和認定，亦存在分歧。

[363] 翁清坤（註 231），頁 661。

doi:10.6342/NTU202401988

對此，學者范姜真微對健保資料庫憲法訴訟案所提出的專家意見書中，也透過與日本個資法制度的比較觀察，凸顯我國個資法去識別化條款的規範模糊問題。在其意見書中有介紹到，日本個資法有將個資去識別化程度分為「假名化加工資料」與「匿名化加工資料」，並清楚界定前者係指以在未與其他資料結合比對下，即無法識別特定個人之方式加工而成之個資（第 2 條第 5 項）；後者則是限縮在係以法定措施進行匿名加工、加工至依社會一般人之能力無法從資料中識別特定個人之程度，且無法被回復還原成個資者（第 2 條第 6 項）；並針對兩種資料分別制定在處理、使用上應遵循的行為義務規範[364]。

在與美國 HIPAA 隱私規則直接明定規範認可的法定去識別化方法，以及日本個資法對兩種程度標準的清楚定義和明確的判斷要件相比之下，我國卻是連法條規範內容本身就未臻明確，實務上又尚未能形成穩定的解釋說明或裁判見解，此即形同遊戲規則都尚未制定完畢就開放參與者進場，所導致的結果就是，野心勃勃的人總是會先以自己認為符合規則的方式大肆行事，發現自身權益受損的人想搬出規則控訴有違規行為的發生，最終也不一定能討回公道、獲得伸張。

第三款 原資料保有者的義務過於寬鬆

從健保資料庫案中即可看出，在現行個資法的適用結果下，即是允許機關單位只要有先將個資經過去識別化處理，就可以將去識別化後的資料自由提供、分享給第三人進行其他公益、學術研究和統計目的之利用，去識別化在此便好似一種快速通關機制，只要宣稱資料都有經過去識別化處理，後續對於資料的各種移轉、利用和揭露，都可以在未經當事人同意，甚至根本無用對當事人為告知的情況下，快速操作執行完畢。

[364] 范姜真微（註 361），頁 6-10。

doi:10.6342/NTU202401988

對比美國 HIPAA 隱私規則，至少在對於未經去識別化完全的有限資料集，尚有規定須在有以契約方式限制資料接受者資料使用行為的條件下，才可例外放寬無須經當事人授權直接使用；我國個資法，同樣肯認原資料保有者（如本案之健保署）得將尚未達無從識別的資料提供給其他機關單位，卻完全未有課予原資料保有者，應控管其所移轉去識別化資料之接收者，不得為可能侵害當事人隱私之資料使用行為，或任何防止資料被重新識別之行為控管義務。

除個資法外，國民法官個人資料保護辦法中的資料揭露規定，也存在對去識別化資料之後續使用和揭露管制不足的問題。此辦法也同樣僅有要求司法院，在將其所保有的國民法官、備位國民法官及候選國民法官個人資料檔案，提供給獲准審請的學術機構或研究單位進行制度研究目的之運用前，應先將資料內容進行遮隱並以適當方式隔離保護識別資訊，雖然申請者提出申請的計畫內容中應包含所實施的安全維護措施，但此與須在契約內容直接約定，禁止資料接收者對取得的資料進行重新識別，或其他針對防範重新識別風險的措施義務，仍有很大的落差。

由此，我國規範現況就使得原資料保有者，不僅無須負將個資完全去識別化之義務，就可在未告知且未取得當事人同意之情況下，將資料移轉交給其他機關單位，對於其所提供資料的接收者，後續如何對該資料進行利用和揭露，亦不負有任何管理、監督責任，目前的實際情況，幾乎就僅仰賴機關單位自律，內部自行制定資料使用和揭露之安全維護計畫和控制措施[365]，欠缺強制性的義務課予或歸責規範，此恐仍使原資料保有者利用政策促進資料效用之美意，對去識別化資料的使用流於恣意。

[365] 例如衛福部資科中心有訂定「衛生福利部衛生福利資料應用管理要點」，內容規定資料庫的建置作業程序、申請使用的資格限制和申請流程，以及資料使用者應遵守之作業方式和安全措施。

doi:10.6342/NTU202401988

第四款 欠缺對於去識別化資料的監督機制

我國去識別化規定的另一大問題是，在制度設計上，同樣落入美國 HIPAA 隱私規則「去識別化後即遺忘」的缺失謬誤，也就是欠缺應對重新識別風險的監督機制和救濟措施，從而形成無法防止、即時介入後續可能衍生的重新識別和隱私侵害風險之規範保護漏洞。

如同前面章節有一再提到，重新識別風險並非永久不變，而是會隨著資訊技術的進步，以及資料量的增長、可取得資訊的不斷新增和累積，逐漸升高，因此去識別化也不是只要有實施過一次，即可一勞永逸、保障隱私無虞。起初有達到去識別化規範標準的資料，之後也可能會因被重新識別的風險變高，故理論上應不得再允許可在未經當事人同意或未告知當事人之情況下，繼續進行使用或揭露。

然而，當前的實際情況是，當事人往往都是等到發現自己的個資已被輾轉使用後，也才只能透過提起行政訴訟，主張機關單位的資料使用行為違反個資保護法規的規定，請求停止繼續對資料進行使用，幾乎沒有任何事前監督機制，去持續追蹤、控管去識別化資料的識別風險。歐盟 GDPR 所授予獨立監管機構的糾正措施，除了高額的行政罰鍰外，尚有告誡、要求停止資料處理活動，以及得命資料控管者應為一定必要處置等多種方法得動用，相較之下，我國資料保護法規的救濟途徑和監督措施就單薄許多，僅以事後追究刑事責任[366]或處以罰鍰[367]為主，欠缺得立即阻止侵害風險發生的警告或其他暫時性措施。

對此，在健保資料庫案當時即有許多意見指出，我國在監管面上，目前僅仰賴公務機關、學術研究機關的內部資料管理措施和事後的司法救濟途徑，欠缺由外部

[366] 個資法第 41 條規定：「意圖為自己或第三人不法之利益或損害他人之利益，而違反第六條第一項、第十五條、第十六條、第十九條、第二十條第一項規定，或中央目的事業主管機關依第二十一條限制國際傳輸之命令或處分，足生損害於他人者，處五年以下有期徒刑，得併科新臺幣一百萬元以下罰金」；第 44 條則規定公務員犯本章之罪者，加重其刑。

[367] 個資法第 47 條、人體研究法第 22 條、人體生物資料庫管理條例第 25 條與第 24 條第 1 項第 5 款和第 6 款之規定參照。

doi:10.6342/NTU202401988

獨立監管機關，在事前審核、控管資料使用活動之合法性，以提前防範隱私侵害結果發生之他律機制。謝銘洋大法官便有表示，隱私權的侵害具有不可回復性，且往往是在當事人不知情的情況下發生，多會造成當事人難以為權利侵害之主張，故僅有事後司法救濟途徑，實在無從對個人資訊隱私權提供足夠之保障[368]；許宗力大法官亦直指，現行個資法欠缺個資保護之獨立監督機制，以及相應之組織上與程序上保護規定，應未符合限制基本權應實施之憲法上正當法律程序之要求[369]。

第二項 我國個資去識別化的制度發展建議

我國法下普遍允許資料去識別化可取代當事人同意，作為蒐集、處理、利用和目的外使用一般和敏感性個資的獨立合法或除外事由，得免除規範主體對當事人之告知義務或資料銷毀義務。我國法目前使去識別化具有實質上義務限制放寬功能的立法，在規範標準本身不明確，且欠缺足夠配套措施和監督機制之情況下，對於當事人同意權之干預限制，應有違憲之虞。由此，本研究希望透過以下建議面向的提供，改善我國法制度所存在的缺失問題，以平衡、彌補當事人資訊自主控制權益，在當前制度規範模式下受到限制減損的情形，嘗試探尋、建構我國個資去識別化政策目的模式的合理性和正當性。

第一款 以資料揭露模式作為去識別化標準訂定之基本框架

雖然憲法法於 111 年憲判字第 13 號判決中，認定個資法「無從識別特定當事人」之規定文字並未違反法律明確性原則[370]，但多數學者意見以及本研究上述之討論分析，均認為個資法和其他資料相關規範之去識別化標準內涵，亟須待進一步充實。

---

[368] 謝銘洋大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書，頁 14-17。
[369] 許宗力大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書，頁 1-2。
[370] 憲法法庭 111 年憲判字第 13 號第 38-41 段；又本件憲法訴訟案之審查範圍，僅有包含個資法第 6 條第 1 項但書第 4 款與健保法第 79 條和第 80 條之規定，未包含個資法第 16 條但書第 5 款等其他個資法之去識別化條文。

doi:10.6342/NTU202401988

學說多建議，應依個資運用之目的及蒐集、處理和利用規模之不同，分別訂定不同的合法要件和去識別化標準門檻，不可一概而論[371]。本研究亦認同，理想上去識別化的要求標準應與資料運用存在的隱私風險相當，始符合比例原則以正當化對個人資訊自主權的限制。然而，姑且不論立法技術上是否有辦法明確界定不同的資料運用情境，進而設定不同的適用標準，考量重新識別風險的變動性，為避免日後需要頻繁對個資法的內容進行修法，本研究認為，可先著重於在個資法中訂定去識別化標準的基本框架，更具體的標準門檻細分則可授權行政機關以法規命令的方式補充之。

本研究的初步建議是，以資料的揭露模式作為去識別化標準基本框架的制定基礎，也就是依據去識別化資料預定的開放程度，設置所應達到的去識別化標準門檻。倘若資料去識別化後是僅供自己單位內部使用，因所得接觸資料的人相對最能被管控，故對於去識別化的要求程度可相對最低；若是預定會將去識別化資料移轉予特定第三人之半公開揭露模式，資料的去識別化程度則應提高；又若資料經去識別化的目的是未來欲開放供大眾使用，由於此情境下隱私風險通常最高且最難預料和控制，對於重新識別風險容忍程度應最低，故對於資料去識別化的要求應適用最嚴格的標準。

衛福部 112 年所公布之人體生物資料庫管理條例預告修正草案，似乎即是朝向類似的方向進行調整。考量規範所要求去識別化之方法、手段、技術和操作之程序應與當時的科技水準看齊，為保留法制上的彈性空間，該修正草案刪除「編碼」和「加密」之規定文字，並參照歐盟 GDPR 和日本醫療資訊法案關於匿名化之解釋，修正該法中的「去連結」定義為「以匿名化或其他資訊科技方式，使參與者之生物資料，永久無法以合理之方式連結、比對參與者身分」，並另增訂「去識別化」之

[371] 劉靜怡（註 308），頁 13；范姜真微（註 361），頁 18。

doi:10.6342/NTU202401988

概念，將之定義為「參與者個人資料經處理後，無從識別特定參與者」（草案第3條第8款）[372]。進一步看來，修正後之人體生物資料庫管理條例，似有意將「去連結」和「去識別化」在規範功能上作出區分，此可見於修正草案將生物資料庫之運用（現行第 18 條），再細分為資料庫之「內部使用」與「對外提供」兩種行為類型，對於內部管理行為，係要求設置者在儲存、處理其所有之生物資料和生物檢體時，應將資料中參與者之身分「去識別化」（草案第 24 條第 1 項前段[373]）；至於在提供生物資料和生物檢體為研究和技術研發之用時，則要求應以「去連結」的方式為之（草案第24條第1項後段）。

第 24 條修正說明提到，生物資料庫的內部管理，為兼顧實務操作的可能性，應採取措施使資料無從識別特定參與者之身分；對外提供資料時，則應使資料永久無法和個人身分進行連結、比對[374]。從此段說明內容上來看，本研究推測立法者之立場似是認為，資料的內部使用階段，由於資料的接近取得是被限制於特定組織單位內部，隱私風險較可控，故相對於個人隱私權之保障，可更導向維護資料的可利用性，故允許資料得以「非去連結方式」保存在資料庫中[375]，在對外釋出前的內部使用階段中，在措施採取上僅須作到使資料無從識別特定參與者之身分即可；到了對外提供大眾投入運用之階段時，因資料公開揭露後遭串連攻擊之可能性增加，此時則應轉而偏重重新識別風險之預防，故要求去識別化措施應達到永久去連結之強度效果。

---

[372] 人體生物資料庫管理條例修正草案總說明，頁 4-5。

[373] 人體生物資料庫管理條例修正草案第 24 條第 1 項：「生物資料庫設置者，就其所有之生物資料、生物檢體為處理時，應以去識別化為之；提供生物醫學研究及生物科技研發運用時，應以去連結為之；整合平臺設置者處理及提供生物資料時，亦同」。

[374] 草案第 24 條修正說明第二點：「生物資料庫於內部管理生物資料、生物檢體時，基於保障參與者隱私權，並顧及實務操作之可能性，應採取正當程序，使之無從識別特定參與者身份；另對外提供生物醫學研究及生物科技研發運用時，應永久無法連結、比對特定個人身分」。

[375] 現行法與修正草案對於生物資料庫之定義，皆說明基於後續提供研究、研發運用之需求，生物資料和生物檢體等相關資料在資料庫中，係以非去連結方式保存，現行法第 3 條第 4 款、修正草案第 3 條第 1 款規定參照。

doi:10.6342/NTU202401988

不過較可惜的是，修正草案中對於「去識別化」和「去連結」的定義仍過於空泛、不明確，未清楚說明兩者在效用強度或方法措施上究竟應如何區分，但即便如此，本研究仍認為，從此一修正內容觀之，似乎得探見立法者在健保資料庫案後，或有意以內部使用和對外提供之資料使用階段行為之劃分，搭配要求不同去識別化措施之實施。由此，以資料揭露模式作為去識別化、資料保護立法措施的新格局，進而擴及影響個資法和其他資料保護法規中去識別化規範之修法，值得持續關注。

第二款 資料使用者的行為限制與義務責任

目前規範關於資料使用者對個資進行去識別化處理，以及後續所進行之去識別化資料利用活動的責任義務之管制有過於寬鬆之問題。綜觀我國各該資料保護法規也並未明文將去識別化資料予以豁免，111 年憲判字第 13 號的判決理由中亦已清楚表明，經去識別化處理後的資料仍具間接識別性而仍屬個人資料。由此，對於去識別化資料所為之各種利用行為，本就該持續受到各該資料保護法規的規範與限制。

1. 強化對去識別化資料當事人的告知義務

在已放寬取得當事人同意的情況下，倘若又允許資料使用者對去識別化資料的移轉、利用和對外提供等一切使用行為活動，都可以在當事人不知情的情況下秘密進行，此無疑是形同完全剝奪當事人對自己資料自主控制權之行使，更導致當事人根本無從發現其個人隱私即將或已經遭受侵害，無法即時尋求救濟維護自身權益。

在告知義務的建置上，具體應告知的內容和義務履行方式，可參照 NCVHS 的建議文件，以及加州 AB-713 號法案對 CCPA 的修訂內容。為使當事人獲得充分的知悉，資料使用者在實際實施去識別化前，即應保留適當期間先向當事人告知其資料即將要被去識別化，內容至少應包含去識別化方法和操作程序的描述，及資料去識別化後的預定用途；後續如有再將去識別化資料公開揭露、移轉予第三人，或供其他目的之使用之情事，亦須事前向當事人為告知，使當事人得持續追蹤、了解其資

doi:10.6342/NTU202401988

料的動態情況、目前掌管其個資者確切為何人。至於告知的方式，除了個別進行告知外，倘若涉及的當事人人數眾多，或是有難以逐一聯繫、履行成本過高或其他顯有困難之情事，則可允許以公告或其他合理方式行之。

在此之下，本研究認為現行個資法第 9 條第 2 項第 4 款，允許間接蒐集者基於公益、統計或學術研究目的，在處理或利用已完成去識別化的資料前，得無須另向當事人履行告知義務之規定，應參考 GDPR 第 14(4)(b) 條之規定，修改為由資料使用者舉證履行告知須投入不成比例之成本，始得例外以公告方式進行告知，而非逕予免除；再者，去識別化實屬個資法定義下的資料「處理」行為[376]，在直接蒐集個資的情形，依照個資法第 8 條規定的文義，公務機關和非公務機關僅有在「蒐集」個人資料的當下，才有義務須對當事人進行告知，然而，考量我國法下去識別化實已可取代當事人同意等其他事由，作為合法處理、利用個資的獨立要件，應修法要求公務機關和非公務機關，仍須就去識別化資料的處理、利用和目的外利用之情事，對當事人為告知，以補償、正當化對於當事人同意權的限制干預。

2. 新增重新識別的一般性限制規定

針對重新識別行為的防範，首先，應在個資法中增訂禁止資料重新識別行為的一般規定，對象應包含原先實施去識別化者本身，也就是不允許其在將資料進行去識別化後，又恣意再將資料還原識別。具體的規範內容，可參考 AB-713 號法案新增條文的規定方式，即原則上禁止公務機關和非公務機關，對已去識別化的資料進行或企圖（attempt to）進行重新識別，只有在限定的特殊目的或活動事由的範圍內，始得例外允許為之，具體的開放事由，則可授權各該事業目的主管機關另行訂定子法補充之，以應對不同產業領域之合理需求。

[376] 法務部(107)法律字第 10703512280 號函（節錄）：「個人資料去識別化之行為應定性為個資法第 2 條第 4 款所稱之『處理』」。

doi:10.6342/NTU202401988

為抑制例外允許重新識別可能衍生的隱私風險，在控管方式上，可另參考 HIPAA 隱私規則對於重新識別所進一步設置的限制規定[377]，例如：要求機關單位應採取適當措施以妥善保管對照資訊，且不得將解碼簿等任何重新識別的工具或機制揭露予其他人等。

3. 課予對資料接收者行為之控管義務

為防止去識別化資料被不當利用或揭露，導致成為重新識別攻擊的對象，或是被用以作為窺探他人隱私的對照參考資訊，應課予資料保有者，須對於其所提供去識別化資料的接收者之後續使用行為，盡到約束和控管的義務責任。具體的作法可參考 HIPAA 隱私規則中，有限資料集所適用的規定內容，即要求資料保有者須在其與去識別化資料接收者的契約中，加入約束資料接收者不當行為之條款內容，一併作為合法利用以及目的外利用去識別化資料的合法要件，例如：禁止將去識別化資料為契約約定範圍外之使用或揭露、資料接收者為履行契約義務將去識別化資料揭露予第三人時，須確保該第三人亦同意遵守本契約有關去識別化資料使用行為之限制內容等。

回歸我國個資法，即是在第 19 條與第 20 條關於非公務機關蒐集、處理、利用之規範中，增訂條款規定非公務機關應以契約或其他具有拘束效力之方式，限制、約束其所提供去識別化資料之接收者的不當資料使用和揭露行為，始得依上述第 19 條第 1 項第 4 款或第 20 條第 1 項第 4 款之規定，移轉、揭露去識別化資料。

[377] 45 CFR § 164.514(c): Implementation specifications: Re-identification. 原文為：“A covered entity may assign a code or other means of record identification to allow information de-identified under this section to be re-identified by the covered entity, provided that: (1) Derivation. The code or other means of record identification is not derived from or related to information about the individual and is not otherwise capable of being translated so as to identify the individual; and (2) Security. The covered entity does not use or disclose the code or other means of record identification for any other purpose, and does not disclose the mechanism for re-identification.”

doi:10.6342/NTU202401988

第三款 監督機制的設置與執行

個資去識別化程度所應達到的標準訂定，以及對於資料使用者所規範的行為義務，都需要有完善的監督機制與有效的執行措施，才得以真正發揮制度規範的規制效力，落實防範重新識別風險、保障個人隱私的功能意旨。然而，我國過去長久以來，皆未存在有負責統籌個資保護相關事項的獨立專責機關，而是分散由各該事業目的主管機關個別執掌。終於，憲法法庭在 111 年憲判字第 13 號中，宣告要求我國應於 3 年內就個資保護建置組織上及程序上之監督防護機制，我國才正式於 112 年 12 月 5 日成立「個人資料保護委員會籌備處」（以下簡稱「籌備處」），開始逐步進行我國個資監督機制的規劃和建置。

目前關於籌備處的準備、建置進度資訊並不多，僅粗略公布此籌備處主要掌理的事項為以下九點：一、個人資料保護委員會籌設事項之總體規劃、協調及推動；二、個人資料保護委員會組織法規之研擬；三、個人資料保護法規之訂修、解釋及協調推動；四、對公務與非公務機關個人資料保護事務監督、查核、通報、陳情等相關機制之規劃；五、個人資料保護相關教育訓練、宣導與人才培育之規劃、推動及執行；六、個人資料保護相關科技與應用模式之研析、評估、發展、交流及推廣；七、國內、外個人資料保護法制、政策與執行之研究、追蹤、調查及統計；八、國際個人資料保護事務合作、參與、交流之規劃及推動；九、其他有關個人資料保護委員會籌設事項[378]。

關於上述第四項有關籌備處的監督、查核、通報、陳情等機制之規劃，本研究認為應首重兩大事前機能面向，即「法規遵循稽查」與「事前救濟管道」的建立。

「法規遵循稽查」的目的是督促、確保政府機關和私人機構單位對於個人資料的使用行為和活動，都是在符合規範要求和標準的情況下進行。在去識別化的監管

[378] 個人資料保護委員會籌備處，https://pipa.pdpc.gov.tw/（最後瀏覽日：2024 年 7 月 15 日）。

doi:10.6342/NTU202401988

議題上，具體採取可以的作法包含：檢驗正在使用的去識別化資料是否有達到規範要求的去識別化標準；透過內部作業流程規則、個資使用處理紀錄和合約文件的調取，檢查確認資料使用者是否有每次確實對當事人進行告知，在契約或其他商業往來文書中，是否有確實納入去識別化資料使用行為之限制內容，以及例外允許進行資料重新識別時，是否是在遵守相關限制和安全措施保護義務的情況下進行等。

「事前救濟管道」則是使監管機關得以即時介入，即將發生或已經正在進行的隱私侵害事件，阻止損害的發生和擴大。具體落實的方法，首先，是建置申訴專線或網站申訴專區等檢舉通報管道，讓一般民眾或機構單位內部吹哨者，在發現去識別化個資疑似被不當利用或移轉、未獲得充分告知，或是有正在進行不當重新識別行為等違法情事時，得報請監管機關介入調查，監管機關亦可藉此管道進行情報資訊的蒐集，以利後續責任的釐清和追究；再者，是可參考歐盟 GDPR 下的糾正權限機制，在個資法中增訂賦予監管機關，除了罰鍰以外，亦得對有違法之虞或已經違法的機關單位，發出警告使其盡快履行去識別化資料之告知義務、補正對去識別化資料接收者行為應盡之控管義務、要求進行重新識別者應採取特定之安全措施，又倘若有重大違規行為或屢次勸告不聽之情事時，並得命暫時或永久停止對去識別化資料的使用行為活動。

第三項 法制進展：衛生福利資料管理條例草案

第一款 草案規範內容簡述

回應 111 憲判字 13 之判決意旨，衛福部於 113 年 3 月 1 日公告預告制定「衛生福利資料管理條例」草案，針對包含健保資料在內之個人衛生健康相關資料，基於醫療、衛生、公益或統計研究之目的，對外釋出予政府機關和學術研究機構單位之目的外利用相關事項，進行規範[379]。

[379] 衛生福利資料管理條例草案總說明，頁 2。

doi:10.6342/NTU202401988

該草案之規範標的為基於辦理健保業務，依衛福部組織法和主管之作用法規所蒐集之衛生福利資料，具體包含個人衛生資料、個人福利資料與其他非個人資料（草案第 4 條）。草案第 5 條第 2 項規定主管機關（即衛福部）得命其所屬機關，將所蒐集之衛生福利資料假名化後，傳輸至衛福部進行串聯並於資料庫中保存；本草案之假名化規定於第 3 條第 4 款，係指「資料經處理或加工後，非透過其他資訊對照，不能識別其身分，且該其他資訊應分開存放，並採取技術上或組織上之保護措施」，此與歐盟 GDPR 假名化之規範定義近乎相同；假名化之具體程序和作業方法授權由衛福部訂定之（草案第 5 條第 3 項）。

關於衛生福利資料之對外釋出，草案第 10 條明確將申請者限定於政府機關（構）、醫療機構、學術研究機構、大學，及受政府機關委託之大學、法人、機構。此外，針對經申請者利用行為之管制，草案第 13 條規定申請者就衛生福利資料之利用，應依衛福部或其所屬機關指定之方式、時間和場所進行（第 1 項）；且要求申請者對衛生健康資料之利用結果，內容不得含有可識別個人個人身分之資訊（第 2 項）；並禁止申請者不得將為申請目的範圍外之利用，或將所取得的資料供他人使用（第 3 項）。申請者如有違反草案第 13 條利用行為之限制禁止規定，依草案第 22 條可由衛福部或其所屬機關處以新臺幣 50 萬元以上 250 萬元以下之罰鍰，限期未改正並得中止或終止該申請者之利用行為，及令其一年內不得再為申請。

此外，應 111 憲判字 13 之要求，該案第三章增訂資料停止使用之相關規範，草案第 16 條明文賦予當事人得申請請求停止對其衛生福利資料之目的外利用，經衛福部和其所屬機關同意註記後，即不得再將該資料對外釋出供目的外之利用（草案第 17 條第 1 項）。

再者，呼應 111 憲判字 13 之意旨，草案第 6 條規定衛福部應設置「衛生福利資料諮詢議會」作為本法之監督機制措施，負責目的外利用申請之爭議處理、衛生福利資料目的外利用和管理政策之擬定、安全維護措施之諮詢、受理當事人請求停止

doi:10.6342/NTU202401988

目的外利用之申請、協助衛福部對受託機構之輔導、評估和監督，以及其他相關事項之建議。

第二款 本研究評析

首先，就去識別化標準而言，於衛福部所屬機關將衛生福利資料回傳給衛福部之階段，該草案僅要求各該所屬機關應將資料進行假名化處理，消除資料的直接識別性，嗣後則以申請人應將資料再去識別化至不含有可識別個人資訊之程度。本研究認為，此標準規範模式對去識別化義務之分配並不合理，衛福部將衛生福利資料建置成資料庫並對外開放申請使用，為主要進行目的外利用之主體，但卻是由申請人負擔最終將資料去識別化之義務；另從風險評估之角度而言，衛福部及其所屬機關對外釋出衛生福利資料，始為整個資料目的外利用過程中隱私風險大幅升高之階段，故資料去識別化於此階段即應大致完備，而不應是申請者將資料研究成果發表之前。

再者，該草案雖有針對申請人不當利用衛生福利資料之行為，訂有罰則規範，但卻遺漏第 5 條衛福部所屬機關回傳資料之假名化義務，設置任何制裁或監督措施，對於重新識別之防範恐尚有不足。

不過，值得肯定的是，草案第 15 條有要求衛福部及其所屬機構，應依個資法第 17 條之規定，公開其持有衛生福利資料之相關資訊，包含提供目的外利用之類別，以及請求停止他人利用之申請方式、受理機關等事項，以利當事人為權利之行使主張。此規範內容即呼應本文所強調，應強化資料使用者對去識別化資料告知義務之建議，期許衛生福利資料諮議會應確實監督此義務之履行情況，使草案之立法意旨得充分落實。

doi:10.6342/NTU202401988

第五章 結論

去識別化作為資料去個人化的處理技術，在個人資料保護法制 下，被期待可作為開放資料運用促進研究發展與確保隱私適足保障兩難下的政策解套措施；透過將資料的識別程度降低至個人資料識別性要件的界定門檻下，使資料脫離規範適用的管制範圍，在確保隱私保障之同時，得供自由流通和使用。

本研究以上述立論作為出發點，第二章先從去識別化的相關概念和技術特性進行探討。在去識別化相關用語的區辨上，匿名化與假名化同為去識別化技術的一種，但在效用強度上，前者須永久、不可逆地消除資料識別性，以完全切斷資料與個人間的關聯性，後者則專指以別名遮蔽識別資訊的方法，使單從資料內容本身無法看出其所指涉的特定個人，但仍若與其他資訊進行比對，則仍可能得以識別出個人身份；又加密雖然在個資保護規範中經常與假名化並列出現，但回歸技術的本質功能來看，加密主要是以維護機密、防止資訊內容被他人讀取為目的的資料安全措施，此與透過遮蔽資料中的識別資訊，以達到隱藏資料與個人間連結關係之假名化技術，並非全然等同。

再者，基於去識別化技術的實施原理和侷限性，在政策制度的引進和建構上，須留意去識別化的實施會同時減損資料的可利用性，對去識別化效用強度所設的標準愈高，雖然能更高程度上保障個人隱私，但卻可能導致資料分析準確性的大幅降低與資料處理成本劇增的後果。另一方面，去識別化實際能對產生多大的隱私保障功效，亦有許多的疑點，其一是，以目前的資料處理技術來看，對於非結構資料所能達到的去識別化效果仍有相當大的限制；其二則是去識別化的殘餘風險問題，數位化時代下資料比對技術的進步和資料量增長的趨勢，導致無論如何都可能再透過與其他資訊的比對，進而重新識別出去識別化資料背後的個人身份，故現今已普遍

doi:10.6342/NTU202401988

肯認無法達到完全的去識別化，因此，如何降低、控制重新識別的風險，是制度設計上必須斟酌考量之課題。

本研究第三章進行比較法制度的研究分析，發現歐盟 GDPR 和美國 HIPAA 隱私規則中，對於去識別化在整體規範上目的定位和效果功能的設計，有很大的差異。GDPR 在去識別化的用語和標準上，有匿名和假名化之分，其制度模式下雖然肯認個資經有效匿名化變為匿名資訊後，就可不再屬於 GDPR 的管制適用範圍，但其對匿名化技術有效性的檢驗標準甚嚴格，故能透過資料匿名化得到豁免管制實際上恐怕是非常困難；假名化在 GDPR 中則是一種資料保護措施，不具有取代或充當作為獨立資料處理合法性基礎的功能地位，反而實為資料控管者或處理者應盡的義務，無管制放寬的功能效果。

相對之下，美國 HIPAA 隱私規則則是鼓勵受規範機構，只要依照法定方法將資料進行去識別化後，便得自由對資料進行利用和揭露，無須再取得當事人之同意授權，惟此種對去識別化資料完全豁免的制度模式，在當今重新識別風險的籠罩下，已浮現出對個人隱私兼顧程度明顯不足的問題。加州 2020 年通過的 AB-713 號法案，在將依 HIPAA 隱私規則規定之方式，進行去識別化處理的健康資料，納入 CCPA 豁免範圍的同時，亦新增規範主體對於去識別化健康資料的透明性和契約管控義務，故美國的去識別化制度規範，未來是否會對完全豁免模式逐漸縮緊調整，值得繼續觀察。

第四章則回歸我國法規範的現況檢討與分析。首先，本研究認為，我國法對於去識別化概念的規範，應較近似於美國 HIPAA 隱私規則的制度模式，此主要展現在，我國現行法規範多將資料去識別化與當事人同意之取得視為擇一、取代關係，例如：生物特徵資料相關之規定，多允許研究者或資料庫建置者倘若已將資料進行去識別化，則無須經當事人同意，即可免除原本應將資料銷毀之義務；又個資法不僅允許間接蒐集者只要已將資料完成去識別化，對該資料後續之處理利用則無須再

doi:10.6342/NTU202401988

對當事人為告知，且更於個資的蒐集、處理、利用和目的外利用的規定中，使資料去識別化得取代當事人同意等其他要件，單獨作為一項合法事由；政治檔案條例及國民法官個人資料保護辦法，在資料揭露使用的相關規範中，亦均有類似的制度安排等等。

然而，我國法規範模式限制人民的資訊自主權，卻幾乎僅以資料去識別化作為對當事人之唯一、最終保護措施，不僅現行條文對於去識別化標準的訂定，有規範標準矛盾且定義模糊的不明確問題，導致行政機關與憲法法庭的闡釋說明相牴觸，以及實務裁判適用認定見解的混亂；且在未要求規範主體應盡去識別化資料的告知義務或行為控管義務之情況下，亦將使得去識別化資料被恣意使用或提供，造成當事人巨大的隱私風險；在程序面上，又欠缺事前救濟措施和相關監督機制之提供和提行，使得當事人之權益無法獲得即時之保障。

最後，本研究嘗試對我國的去識別化制度提出改革建議。第一，以資料去識別化之預定揭露模式，作為重新建置去識別化規範標準的基準框架，使我國法下去識別化的內涵和門檻要求得逐步明確化、合理化。二是，增訂規範主體使用去識別化資料的告知義務和重新識別行為之禁止規定，並且有責任應就其所提供資料之接收者使用行為進行控管，避免規範主體對去識別化資料的處理、使用流於恣意，導致對當事人更大的隱私侵害風險。三則是監督層面的完善，因應我國甫成立個人資料保護委員會籌備處，在監督機制的設置上，建議未來的專責機關應定期實施稽查程序，確保機關單位對告知義務、行為控管義務及重新識別行為禁止等去識別化規範之遵循；再者，除事後的司法救濟外，亦應建立申訴通報機制，作為人民可請求事前救濟的管道，防止隱私侵害的實際發生和擴大，落實正當法律程序憲法原則之精神。

doi:10.6342/NTU202401988

參考文獻

一、中文部分

（一）書籍

陳宗和、蔣彥亭、宋瑞豐、陳姿蓉、陳姿佑（2021），資訊科技全一冊，臺北：科有。

（二）期刊論文

王紹睿（2018），淺談人工智慧系統的隱私資訊安全保護機制，科儀新知，215 期 6 卷。

江耀國、黃子宴（2019），個人資料的概念與匿名化：一個認識論的觀點，東海大學法學研究，58 期。

吳全峰、許慧瑩（2018），健保資料目的外利用之法律爭議－從去識別化作業工具談起，月旦法學雜誌，272 期。

李世德（2018），GDPR 與我國個人資料保護法之比較分析，台灣經濟論衡，16 卷 3 期。

李建良（2017），資料流向與管制環節—個資保護 ABC，月旦法學雜誌，272 期。

李思壯、黃彥男（2019），數位時代之數位隱私保護，國土及公共治理季刊，7 卷 4 期。

李寧修（2020），個人資料合理利用模式之探析： 以健康資料之學術研究為例，臺大法學論叢，49 卷 1 期。

林裕嘉（2017），公務機關利用去識別化資料之風險評估及法律責任（上），司法周刊，1852 期。

邱忠義（2014），談個人資料保護法之間接識別，月旦裁判時報，30 期。

doi:10.6342/NTU202401988

范姜真嫩（2013），個人資料保護法關於「個人資料」保護範圍之檢討，東海大學法學研究，41 期。

范姜真嫩（2017），大數據時代下個人資料範圍之再檢討—以日本為借鏡，東吳法律學報，29 卷 2 期。

祝亞琪、魏銷志（2016），行動支付之個人資料去識別化方法，電腦稽核，34 卷。

翁清坤（2023），個人資料之去識別化與再識別化風險：法律之觀點，臺大法學論叢，52 期 3 卷。

翁慈宗（2009），資料探勘的發展與挑戰，科學發展，442 期。

張陳弘（2016），個人資料之認定－個人資料保護法適用之啟動閥，法令月刊，67 卷 5 期。

張陳弘（2018），國家建置全民健康保險資料庫之資訊隱私保護爭議－評最高行政法院 106 年度判字第 54 號判決，中原財經法學，40 期。

張陳弘（2018），新興科技下的資訊隱私保護：「告知後同意原則」的侷限性與修正方法之提出，臺大法學論叢，47 卷 1 期。

張陳弘（2022），美國加州消費者隱私保護法制之最新發展與比較法啟示，當代法律，6 期。

張陳弘（2023），健保資料二次使用之個人資料保護立法芻議－111 年憲判字第 13 號【健保資料庫案】判決之回應，輔仁法學，66 期。

楊岳平（2022），金融科技時代下金融資料共享法制之發展與限制－評「金融機構間資料共享指引」，台灣法律人，17 期。

賈忻蓉（2022），簡析美國 HIPAA 下個人健康資料用於研究之合法要件，科技法律透析，34 卷 4 期。

樓一琳、何之行（2017），個人資料保護於雲端運算時代之法律爭議初探暨比較法分析：以健保資料為例，臺大法學論叢，46 卷 2 期。

doi:10.6342/NTU202401988

蔡昀臻、樊國楨（2016），大數據之資料去識別的標準化實作初探：根基於ISO/IEC 2nd WD 20889：2016-05-30，資訊安全通訊，22卷4期。

（三）網路資料

Yash Mehta 著，曾祥信譯，資料標記化：遮蔽資料的新方法，CIO Taiwan，2022年10月17日，https://www.cio.com.tw/data-tagging-a-new-way-to-mask-data/。

王若樸，【從K匿名法、GAN和統計整合練兵，再攻聯合學習】工研院揭3種去識別化方法，iThome，2021年5月21日，https://www.ithome.com.tw/news/144539。

個人資料保護委員會籌備處，https://pipa.pdpc.gov.tw/。

國家衛生研究院人體生物資料庫，人體生物資料庫簡介，https://biobank.nhri.edu.tw/info/。

黃維中，人工智慧應用下的隱私保護與個資去識別化，科技部全球事務與科學發展中心，2021年10月18日，https://trh.gase.most.ntnu.edu.tw/tw/article/content/247。

資策會科技法律研究所，加州消費者隱私保護法修正法案重點說明，2021年3月，https://stli.iii.org.tw/article-detail.aspx?no=64&tp=1&d=8630。

資策會科技法律研究所，合成資料（synthetic data），2020年10月，https://stli.iii.org.tw/article-detail.aspx?no=6 4&tp=1&d=8532。

（四）其他

人體生物資料庫管理條例修正草案總說明。

中央健保署1110411法規範憲法審查言詞辯論書。

內政部(106)內授營建管字第1060809377號函。

doi:10.6342/NTU202401988

行政院衛生署(101)衛署醫字第 1010265083 號函。

吳全峰健保資料庫釋憲案專家諮詢意見。

李寧修健保資料庫釋憲案專家諮詢意見。

法務部(100)法律字第 0999051927 號函。

法務部(103)法律字第 10303507480 號函。

法務部(103)法律字第 10303510410 號函。

法務部(104)法律字第 10403508020 號函。

法務部(105)法律字第 10503510730 號函。

法務部(105)法律決字第 10503500370 號函。

法務部(106)法律字第 10603513040 號函。

法務部(107)法律字第 10703512280 號函。

法務部(107)法律字第 10703513050 號函。

范姜真微健保資料庫釋憲案專家諮詢意見。

國家發展委員會(110)發法字第 1100017815 號函。

國家發展委員會(111)發法字第 1110000748 號函。

國家發展委員會，加州消費者隱私保護法(CCPA)規範重點說明。

許宗力大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書。

最高行政法院 103 年度判字第 600 號判決。

最高行政法院 106 年度判字第 54 號判決。

黃昭元大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書。

電子簽章法修正草案修正總說明。

臺北地方法院 103 年度訴字第 212 號民事判決。

臺北地方法院 103 年度訴字第 255 號民事判決。

臺北高等行政法院 102 年度訴字第 36 號判決。

doi:10.6342/NTU202401988

臺北高等行政法院 103 年度訴更一字第 120 號判決。

劉定基健保資料庫釋憲案專家諮詢意見。

劉靜怡健保資料庫釋憲案專家諮詢意見。

衛生福利部，111 年全民健康保險醫療統計。

衛生福利部 1110411 法規範憲法審查言詞辯論意旨書。

衛生福利資料管理條例草案總說明。

憲法法庭 111 年憲判字第 13 號判決。

謝銘洋大法官憲法法庭 111 年憲判字第 13 號判決部分不同意見書。

二、英文部分

（一）書籍

Dwork, Cynthia (2008), *Differential Privacy: A Survey of Results*, *in* THEORY AND APPLICATIONS OF MODELS OF COMPUTATION 1, Manindra Agrawal ed al. eds.

FEILER, LUKAS ET AL. (2018), THE EU GENERAL DATA PROTECTION REGULATION (GDPR): A COMMENTARY.

Hassan, Fadi et al. (2018), *Anonymization of Unstructured Data Via Named-Entity Recognition*, *in* MODELING DECISIONS FOR ARTIFICIAL INTELLIGENCE 296, Vicenç Torra et al. eds.

Narayanan, Arvind & Vitaly Shmatikov (2008), *Robust De-Anonymization of Large Sparse Datasets*, *in* 2008 IEEE SYMPOSIUM ON SECURITY AND PRIVACY 111.

THE EU GENERAL DATA PROTECTION REGULATION (GDPR) A COMMENTARY, Christopher Kuner et al. eds., 2020.

doi:10.6342/NTU202401988

（二）期刊論文

Benitez, Kathleen & Bradley Malin (2010), *Evaluating Re-identification Risks with Respect to the HIPAA Privacy Rule*, 17(2) JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION 169.

Emam, Khaled El et al. (2008), *Protecting Privacy Using k-Anonymity*, 15(5) JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION 627.

Finck, Michèle & Frank Pallas (2020), *They Who Must Not Be Identified—Distinguishing Personal from Non-Personal Data Under the GDPR*, 10(1) INTERNATIONAL DATA PRIVACY LAW 11.

Hajduk, Paweł (2021), *The Powers of the Supervisory Body in the GDPR As A Basis for Shaping the Practices of Personal Data Processing*, 45(2) REVIEW OF EUROPEAN AND COMPARATIVE LAW 57.

Hartzog, Woodrow & Ira Rubinstein (2017), *The Anonymization Debate Should Be About Risk, Not Perfection*, 60(5) COMMUNICATIONS OF THE ACM 22.

Ohm, Paul (2010), *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA LAW REVIEW 1701.

Prasser, Fabian et al. (2016), *The Importance of Context: Risk-Based De-Identification of Biomedical Data*, 55(4) METHODS OF INFORMATION IN MEDICINE 347.

Rubinstein, Ira S. & Woodrow Hartzog (2016), *Anonymization and Risk*, 91(2) WASHINGTON LAW REVIEW 703.

Schwartz, Paul M. & Daniel J. Solove (2011), *The PII Problem: Privacy and A New Concept of Personally Identifiable Information*, 86 NEW YORK UNIVERSITY LAW REVIEW 1814.

doi:10.6342/NTU202401988

Sweeney, Latanya (2002), *K-anonymity: A Model for Protecting Privacy,* 10(5) INTERNATIONAL JOURNAL ON UNCERTAINTY, FUZZINESS AND KNOWLEDGE-BASED SYSTEMS 557.

Tovino, Stacey A. (2004), *The Use and Disclosure of Protected Health Information for Research Under the HIPAA Privacy Rule: Unrealized Patient Autonomy and Burdensome Government Regulation*, 49 SOUTH DAKOTA LAW REVIEW 447.

Weitzenboeck, Emily M. et al. (2022), *The GDPR and Unstructured Data: Is Anonymization Possible?*, 12(3) INTERNATIONAL DATA PRIVACY LAW 184.

Zarsky, Tal (2017), *Incompatible: The GDPR in the Age of Big Data*, 47 SETON HALL LAW REVIEW 995.

（三）官方文獻

Article 29 Data Protection Working Party, *Guidelines on Consent Under Regulation 2016/679* (May 4, 2020), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf.

Information Commissioner's Office, *Draft Anonymisation, Pseudonymisation and Privacy Enhancing Technologies Guidance, Chapter 2: How Do We Ensure Anonymisation Is Effective?* (October, 2021), https://ico.org.uk/media/about-the-ico/documents/4018606/chapter-2-anonymisa tion-draft.pdf.

Information Commissioner's Office, Anonymisation: Managing Data Protection Risk Code of Conduct (November 2012), https://ico.org.u k/media/1061/anonymisation-code.pdf.

Agencia Española de Protección de Datos & European Data Protection Board, *10 Misunderstandings Related to Anonymisation* (April 27, 2021), https://www.edps.europa.eu/system/files/2021-04/21-04-27_aepd-edps_anonymisation_en_5.pdf.

doi:10.6342/NTU202401988

Information and Privacy Commissioner of Ontario, *De-identification Guidelines for Structured Data* (June, 2016), https://www.ipc.on.ca/wp-content/uploads/2016/08/Deeidentification-Guidelines-for-Structured-Data.pdf.

Agencia Española de Protección de Datos & European Data Protection Board, *Introduction to the Hash Function As A Personal Data Pseudonymisation Technique* (October, 2019), https://www.edps.europa.eu/sites/default/files/publication/19-10-30_aepd-edps_paper_hash_final_en.pdf.

European Data Protection Board, *EDPB Document on Response to the Request from the European Commission for Clarifications on the Consistent Application of the GDPR, Focusing on Health Research* (February 2, 2021), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_replyec_questionnaireresearch_final.pdf.

European Data Protection Board, *Opinion 39/2021 on Whether Article 58(2)(g) GDPR Could Serve As A Legal Basis for A Supervisory Authority to Order Ex Officio the Erasure of Personal Data, in A Situation Where Such Request Was Not Submitted By the Data Subject* (December 21, 2021), https://www.edpb.europa.eu/system/files/2022-01/edpb_opinion_202139_article_582g_gdpr_e n.pdf.

European Medicines Agency, *External Guidance on the Implementation of the European Medicines Agency Policy on the Publication of Clinical Data for Medicinal Products for Human Use*, EMA/90915/2016 (v.1.4) (October 15, 2018), https://www.e ma.europa.eu/en/documents/regulatory-procedural-guideline/external-guidance-implementation-european-medicines-agency-policy-publication-clinical-data-medicinal-products z-hu man-use-versio n-14_en.pdf.

European Parliament Document (COM 2012)0011 (2013).

European Parliament Document EP-PE_TC1-COD(2012)0011.

doi:10.6342/NTU202401988

United States Department of Health and Human Services, *Summary of the HIPAA Privacy Rule*, https://www.hhs.gov/sites/default/files/privacysummary.pdf.

United States Department of Health and Human Services, *Guidance Regarding Methods for De-Identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule* (November 26, 2012), https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhs _deid_g uidance.pdf.

Standard for Privacy of Individual Identifiable Health Information 67 no. 157 Federal Register 53128 (August 14, 2002).

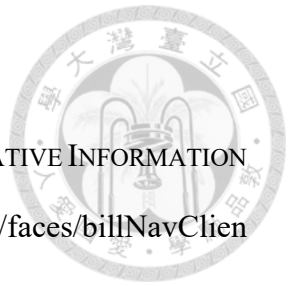
Article 29 Data Protection Working Party, *Opinion 4/2007 on the Concept of Personal Data* (June 20, 2007), http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.

Article 29 Data Protection Working Party, *Opinion 05/2014 on Anonymisation Techniques* (April 10, 2014), https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

European Data Protection Board, *Guidelines 4/2019 on Article 25 Data Protection By Design and By Default* (October 20, 2020), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf.

European Data Protection Board, *Guidelines 03/2020 on the Processing of Data Concerning Health for the Purpose of Scientific Research in the Context of the COVID-19 Outbreak* (April 21, 2020), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_en.pdf.

doi:10.6342/NTU202401988



（四）網路資料

AB-713 California Consumer Privacy Act of 2018, CALIFORNIA LEGISLATIVE INFORMATION (September 29, 2020; 2:00 PM), https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201920200AB713.

Andrew Folks, *US State Privacy Legislation Tracker*, IAPP, https://iapp.org/resources/article/us-state-privacy-legislation-tracker/

Arrington, Michael, *AOL Proudly Releases Massive Amounts of Private Data*, TECHCRUNCH (August 7, 2006; 9:17 AM), https://techcrunch.com/2006/08/06/aol-proudly-releases-massive-amounts-of-user-search-data/

Barbaro, Michael & Tom Zeller Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, THE NEW YORK TIMES (August 9, 2006), https://www.nytimes.com/2006/08/09/technology/09aol.html.

Calhoun, Patricia S. & Patricia M. Carreiro, *De-Identified Data Exception in HIPAA Poses A Litigation Risk*, STAT (April 16, 2020), https://www.statnews.com/2020/04/16/de-identified-data-exception-hipaa-litigation-risk/.

California Consumer Privacy Act Regulations, CALIFORNIA PRIVACY PROTECTION AGENCY, https://cppa.ca.gov/regulations/consumer_privacy_act.

Children's Online Privacy Protection Act, FEDERAL TRADE COMMISSION, https://www.ftc.gov/legal-library/browse/statutes/childrens-online-privacy-protection-act.

Cosgrove, Cathy, *Top-10 Operational Impacts of the CPRA: Part 2 – Defining "Business" Under the Law*, IAPP (December 22, 2020), https://iapp.org/news/a/cpras-top-operational-impacts-part-2-defining-business/.

doi:10.6342/NTU202401988

Data Protection Guide for Small Business: Secure Personal Data, EUROPEAN DATA PROTECTION BOARD, https://www.edpb.europa.eu/sme-data-protection-guide/secure-personal-data_en.

Data Protection Laws and Regulations USA 2023-2024, ICLG.COM, https://iclg.com/practice-areas/data-protection-laws-and-regulations/usa.

Family Educational Rights and Privacy Act (FERPA), U.S. DEPARTMENT OF EDUCATION, https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html.

Glossary: Unstructured Data, RESOURCES.DATA.GOV, https://resources.data.gov/glossary/unstructured-data/

Heeg, Robert, *Possibilities and Limitations, of Unstructured Data*, RESEARCHWORLD (February 20, 2023), https://researchworld.com/articles/possibilities-and-limitations-of-unstructured-data.

ISO 25237:2017(en) Health informatics — Pseudonymization, INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, https://www.iso.org/obp/ui/#iso:std:iso:25237:ed-1:v1:en.

ISO/IEC 18033-1:2021(En) Information Security — Encryption Algorithms — Part 1: General, INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, https://www.iso.org/obp/ui/#is o:std:iso-iec:18033:-1:ed-3:v1:en:term:3.20.

ISO/IEC 20889:2018(en) Privacy *Enhancing Data De-Identification Terminology and Classification of Techniques*, INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, https://www.iso.org/obp/ui/#iso:std:iso-iec:20889:ed-1:v1:en.

ISO/IEC 29100:2024(en) Information Technology — Security Techniques — Privacy Framework, INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, https://www.iso.org/obp/ui/en/#iso:std:iso-iec:29100:ed-2:v1:en.

doi:10.6342/NTU202401988

Legacy: Art. 29 Working Party, EUROPEAN DATA PROTECTION BOARD, https://www.edpb.europa.eu/about-edpb/who-we-are/legacy-art-29-working-party_en.

Privacy Act of 1974, OFFICE OF PRIVACY AND CIVIL LIBERTIES, U.S. DEPARTMENT OF JUSTICe, https://www.j ustice.gov/opcl/privacy-act-1974.

Public Comments: August 2006 – FTC Complaint About Search AOL Data Releases, WORLD PRIVACY FORUM (August 16, 2006), https://www.worldprivacyforum.org/2006/08/public-comments-ftc-complaint-about-search-aol-data-releases/.

Richman, Amitai, *Pseudonymization Vs Encryption: Understanding the Differences*, K2VIEW (August 25, 2023), https://www.k2view.com/blog/pseudonymization-vs-encryption/#How-Pseudonymization-Works-Preserving-Privacy-Through-Data-Anonzymization.

Shacklett, Mary, *Structured vs Unstructured Data: Key Differences*, DATAMATION (November 23, 2023), https://www.datamation.com/big-data/structured-vs-unstructured-data/

Structured Versus Unstructured Data: What's the Difference?, IBM (June 29, 2021), https://www.ibm.com/think/topics/structured-vs-unstructured-data.

What Are Identifiers and Related Factors?, INFORMATION COMMISSIONER'S OFFICE, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/personal-information-what-is-it/what-is-personal-data/what-are-identifiers-and-related-factors/.

（五）其他

Case C-582/1, Breyer v. Bundesrepublik Deutschland, ECLI:EU:C:2016:779 (October 19, 2016).

doi:10.6342/NTU202401988

EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS, #BIGDATA: DISCRIMINATION IN DATA-SUPPORTED DECISION MAKING (May 30, 2018), https://fra.europa.eu/sites/default/fil es/fra_uploads/fra-2018-focus-big-data_en.pdf.

Garfinkel, Simson L., *NISTIR 8053: De-Identification of Personal Information*, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (October, 2015), https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf.

Heurix, Johannes et al. (2012), *Pseudonymization with Metadata Encryption for Privacy-Preserving Searchable Documents*, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6149189.

Laney, Doug, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, APPLICATION DELIVERY STRATEGIES (February 6, 2001), https://studylib.net/doc/8%20647594/3d-data-management--controlling-data-volume--velocity--an....

National Committee on Vital and Health Statistics, *Recommendations on De-identification of Protected Health Information Under HIPAA* (February 23, 2017), https://www.ncvhs.hhs.gov/wp-content/uploads/2013/12/2017-Ltr-Privacy-DeIdentification-Feb-23-Final-w-sig.pdf.

Pfitzmann, Andreas & Marit Hansen, *Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology (v0.28),* TU DRESDEN (May 29, 2006), https://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.28.pdf.

Sweeney, Latanya (2000), *Simple Demographics Often Identify People Uniquely* (Carnegie Mellon University, Data Privacy Working Paper No. 3).

Tinabo, Rose et al. (2010), *Anonymisation Vs. Pseudonymisation: Which One Is Most Useful for Both Privacy Protection and Usefulness of E-Healthcare Data*, INSTITUTE

doi:10.6342/NTU202401988

OF ELECTRICAL AND ELECTRONICS ENGINEERS, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5402501.

doi:10.6342/NTU202401988