

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

蔡昀臻¹、樊國楨^{2*}

¹國立交通大學管理科學研究所、²國立交通大學資訊管理研究所
¹yct1230@gmail.com、²kjf.nctu@gmail.com

摘要

隨著「大數據」、「資料探勘」的興盛發展，保護個人隱私的「去識別化」相關技術與標準化的需求也日益受到重視。相關的隱私洩漏事件已清楚證明了資料庫的公開往往隱藏著當事人的資料被識別的風險；為此，各方提出了各種保護個人隱私資料去識別化的實作方法與訂定相關標準；根基於此，國際標準組織(International Organization for Standardization, 簡稱 ISO)已立項進行「增強隱私之資料去識別化技術」的標準化工作項目。為了平衡個人隱私保護之風險與公開資料分析的效益，本文闡明前述標準化之內涵，並建議在我國「個人資料去識別化過程驗證要求及控制措施」的驗證規範之要求事項與控制措施中宜考慮加入相關規範。

關鍵詞：個人資訊去識別化、標準化、統計公開控制、K-匿名模型、差分隱私模型

Data De-Identification Standardized implementation of Big Data: Based on ISO/IEC 2nd WD 20889：2016-05-30

Yun-Chen Tsai¹, Kwo-Jean Farn²

¹Institute of Information Management Science, National Chiao-Tung University,

²Institute of Information Management, National Chiao-Tung University

¹yct1230@gmail.com、²kjf.nctu@gmail.com

Abstract

With the progress of “big data” and “data mining”, “de-identification” techniques and standards that protect people’s privacy had become more and more important. Privacy leak cases had proved the risk that people’s data been identified come hand in hand with database opening. To solve this, different parties have come up with a variety of de-identification methods and standards. In order to achieve the balance between the risk of personal privacy violation and the benefit of open data, we formulate the standardization content of de-identification and make suggestions also the additional requirements the government and take to protect personal privacy in data base opening.

* 通訊作者 (Corresponding author.): 樊國楨

Keywords: Personally Information De-identification, Standardization, Statistical disclosure control (SDC), K-anonymity model, Differential privacy model

壹、前言

隨著電子科技之一日千里，網際網路的發展已澈底改變了日常之生活，網路逐漸成為人們生活的中心，購物、娛樂、社交、看新聞、聯絡事情等，均經由網路。日常生活之網路化，使用者的資訊已成為如何在「大數據分析」時履行適當之「隱私防護」的議題。

2015年7月17日，面對「開放資料」與「大數據」之「去識別化」議題，行政院張善政副院長根基於經濟部標準檢驗局(BSMI)提出如圖 1.1 所示的方案規劃，公布如表 1.1 所示之行政院推動大數據發展的個人資料保護之標準化工作項目。而「CNS 29191 有要求事項，無控制措施；而 CNS 29100 是保護個人可識別資訊的高階框架，可引用作為『去識別化』控制措施」，是 BSMI 對其執行圖 1 與表 1 之思路的說明[12] [13][15][17][19]。

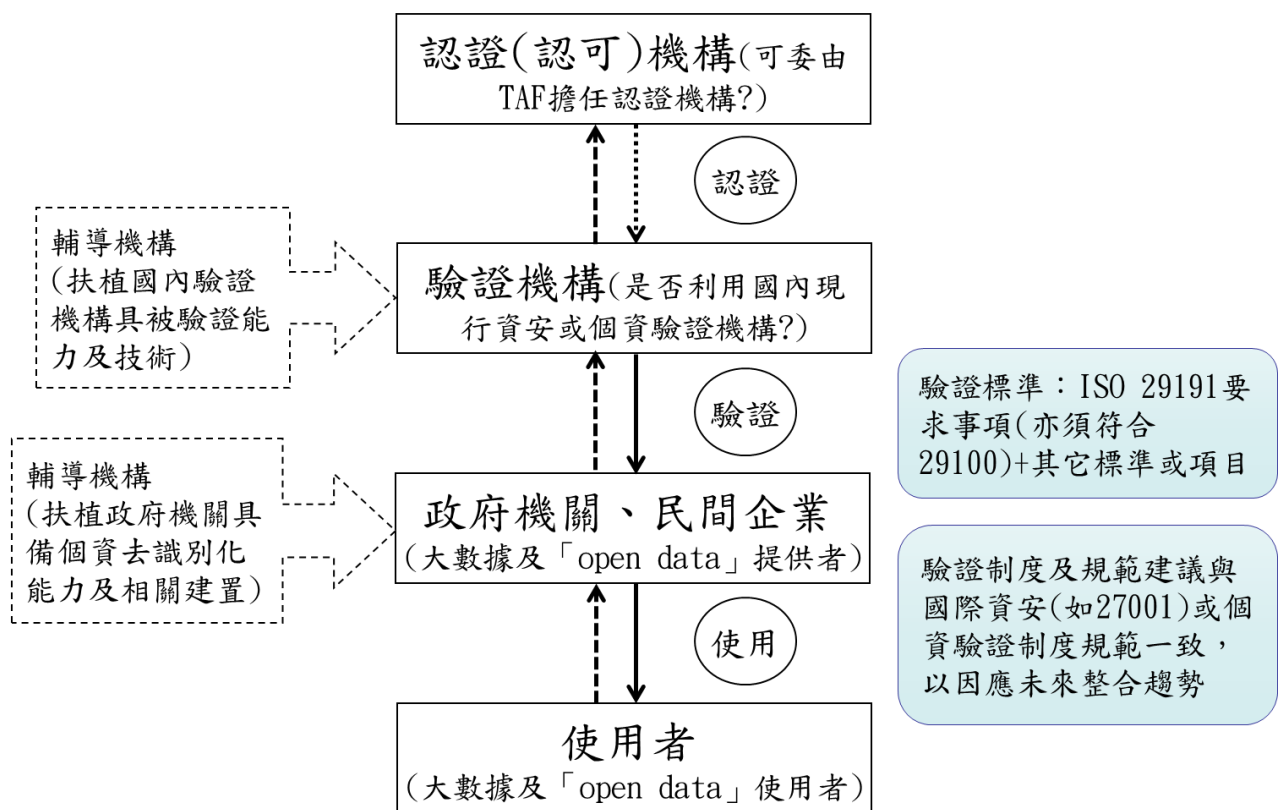


圖 1.1 個人資料去識別化方案規劃一個資去識別化驗證制度體系規劃

資料來源：經濟部(2015)個人資料去識別化之運作機制(簡報資料)，2015-07-14(「研商因應大數據潮流個人資料去識別化可行機制」會議，經濟部標準檢驗局許景行組長簡報)

表 1.1 行政院推動大數據之個人資料保護相關的2項國家標準

標準	CNS 29100：2014-06-04	有關如何管理、確保隱私權之原則框架的國家標準
	CNS 29191：2015-06-10	有關如何去識別化之部分匿名與部分去連結的國家標準
推動作法	<ul style="list-style-type: none"> ● 政院月底將出爐如何取得符合兩標準的標準程序作法 ● 第一步先鼓勵部會取得驗證，下一步鼓勵金融、電信業取得驗證 	
用處	<ul style="list-style-type: none"> ● 去除外界擔心敏感個資外洩疑慮 ● 各部會與業界可以合理應用大數據 	

資料來源：2015年7月17日，大數據發展訂國家標準，經濟日報A1，記者林安妮/台北報導

2015年12月21日，「行政院國家資通安全會報第29次委員會議」，於會議紀錄中將「『個人資料匿名化、去識別化』分列」，同時要求相關機關善加推廣利用前述驗證規範[14]；惟同表 1.2 所示，宜闡明之。

表 1.2 去識別化用語與現有技術之對照

國際標準	ISO/IEC 29191 (2012)	ISO/TS 25237 (2008)	ISO/IEC 29100 (2011)	ICO (2012)	Article 29 (2014)
ISO/IEC 2 nd WD 20889：2016-05-30 中用語					
去識別化 (De-identification)	N/A	De-identification, Anonymisation	Anonymisation	Anonymisation	N/A
遮罩(Masking)	N/A	N/A	N/A	Anonymisation	N/A
可控制的重新識別之擬匿名化 (Pseudonymization with controlled re-identification)	partially anonymous, partially unlinkable	Pseudonymization reversible	Pseudonymization	Anonymisation	Pseudonymization
無可控制的重	N/A	Pseudonymization	Anonymisation	Anonymisation	Pseudonymization

新識別之擬匿名化 (Pseudonymization without controlled re-identification)		irreversible			
隨機 (Randomization)	N/A	N/A	N/A	Anonymisation	Anonymisation
泛化 (Generalization)	N/A	N/A	N/A	Anonymisation	Anonymisation
差分隱私 (Differential Privacy)	N/A	N/A	N/A	N/A	Anonymisation

資料來源：ISO/IEC 2nd WD 20889：2016-05-30, Information technology – Security technology – Privacy enhancing data de-identification techniques, Annex B.

說明：

1. 英國資訊專員辦公室 (Information Commissioner's Office, 簡稱 ICO)。
2. Article 29 係歐盟之規範。

圖 1.1 中之 ISO/IEC 29100、ISO/IEC 29191 等標準與規範的用語並不一致，表 1.2 是其對照列表；其中 ISO/IEC 29191 僅為通稱「網路實名制」之「可控制的重新識別之擬匿名化」的管理控制措施之要求事項。

去識別化可能遇到之重新識別攻擊的攻擊者可能來自各方，他們可能是為了展示自己之理論的正確性或是想從資料中獲益。而成功之攻擊並不需要將資料庫完整重現才算成功的攻擊，攻擊者只要用各種方法取得相關去識別化資料，包含向資料庫提出詢問 (Query) 以及直接取得去識別化資料集，能夠在資料中分析出其目標即可以算是成功之攻擊。以目標分類的話攻擊可以分為下列幾種：

1. 重新識別某筆紀錄是否在特定的資料主體中。
2. 重新識別特定資料主體中的特定紀錄。
3. 重新識別越多越好的紀錄與相對應的資料主體。
4. 重新識別特定資料主體是否落在資料集中。

任何的重新識別之攻擊一般而言都會組合多種技術，並搭配可使用的外部信息作為分析資料庫內容之工具。儘管攻擊的目標多變，評估哪些重新識別技術可能會被使用於去識別化之資料庫仍是相當重要的，表 1.3 是常用之重新識別的技術表列。

表 1.3 重新識別(Re-Identificaton)技術舉隅

技術(techniques)	實作方法
單獨挑出 (singling out)	透過觀察特定的特質，將單一資料或少數資料從資料主體(data principal)分離。
連結 (linking)	連接至少兩個以上在相關的資料主體中的紀錄，或是連接在不同資料集中的一組資料主體。
推斷 (Inference)	有不小的機率可以從某一組屬性(attribute)推斷出另一組屬性。
不可分辨之分析 (indistinguishability analysis)	針對特定資料，透過執行計算(computations)或詢問以確定其是否存在於搜尋的資料主體中。

註：資料主體(data principal) 在此指的是單一主體(個人、組織、設備、軟體程序、.....)其需要保護的敏感資料總稱。

資料來源：ISO/IEC 2nd WD 20889：2016-05-30 第 6.3 節。

綜前所述，於第二節經由案例敘述個人資料去識別化的脈絡及其在美國醫療隱私之影響；在第三節，闡明 ISO/IEC 2nd WD 20889：2016-05-30 的思路並探討其與我國個人資料去識別化之法制的關聯；第四節，比較分析我國現行驗證規範試辦單位之實作與以及美國醫療隱私相關法規及其實作；最後，在第五節，提出本文的結論。

貳、資料去識別化

當寬頻網路問市、線上交易啟用、網路社群誕生時，「資訊社會」隱私洩漏之落塵也散落人間，20 世紀著名的隱私洩漏事件之一敘述如下：

1996 年美國麻塞諸塞州(Mass.)之團體保險委員會(Group Insurance Commission, GIC)決定將州政府員工的醫療資料去識別化後發布，供公共醫學研究使用。然而，1997 年麻省理工學院(MIT)博士生 Latanya Sweeney 使用連結攻擊法(Linking attack)，分析使用保留了患者的出生日期、性別及郵遞區號三項屬性；破解前述資料集，找出時任州長(Governor)之 William Weld 的醫療紀錄並寄給州長本人[8]。

娜坦雅·史威尼之研究，深深影響了如圖 2.1 與表 2.2 所示的美國「健康保險可攜式與責任法(Health Insurance Portability and Accountability Act，簡稱 HIPPA)之「安全港法(Safe Harbor)」的隱私規則，其提出之「K-匿名(K-anonymity)法」將圖 2.2 中需要去識別化的「部分識別符(Quasi-Identifiers)」予以在「重新識別(Re-identification)」風險及保留資訊之「使用性(Utility)」間取得妥協。

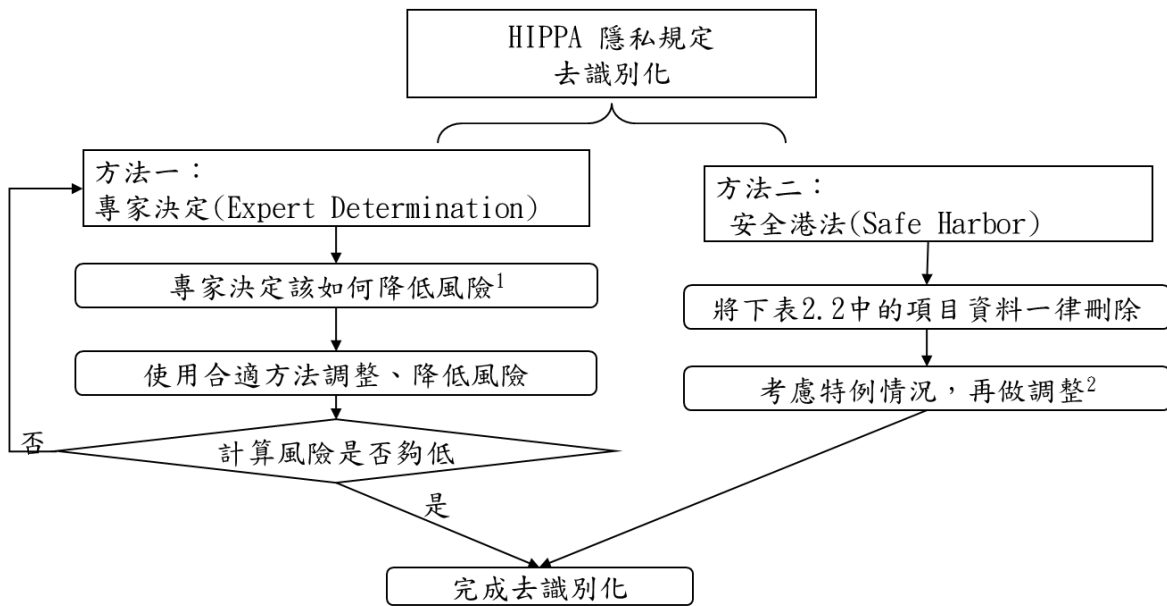


圖 2.1 HIPPA 隱私規定去識別化流程

說明 1：專家一般指在統計、數學、或其他科學領域受過去識別化相關教育或有處理過相關事項的人員；專家會透過評估資料的可複製性(Replicability)、資料來源可用性(Data source Availability)、可辨識性(Distinguishability)與風險評鑑(Assess Risk)來決定資料的安全與否。
說明 2：有些資料因為情況特殊(有該項特徵的人很少)需要手動調整。例如：如果資料裡面有人的職業是總統，則可能需要另外遮蔽該訊息。

表2.2 HIPPA之Safe Harbor 隱私規則規定需要處理的資訊

(A) 姓名	
(B) 所有比「州」小的地理單位，包括街道地址，城市，縣，管轄區，郵遞區號，和它們的相對應的地理編碼，而郵遞區號前三碼則根據人口統計局資料調整： (1) 如果原來的三碼郵遞區號涵蓋的地理單元人口超過20000人;且 (2) 人口為20000以下的地理單元的郵遞區號更改為000	
(C)直接關係到個人資料的日期（年份除外），包括出生日期，入院日期，出院日期，死亡日期；而所有年齡超過89歲及其相關日期資料（含年份）合併為「90歲或以上」之單一類別	
(D) 電話號碼	(L) 車輛識別碼和序號，包括車牌號碼
(E) 傳真號碼	(M) 設備標識符和序列號
(F) 電子郵件地址	(N) 網頁網址(URLs)
(G) 社會安全號碼	(O) 通訊協定編號（IP）地址

(H) 醫療記錄編號	(P) 生物識別技術，包括指紋和聲紋
(I) 健康計劃受益人數	(Q) 正面照片和任何可比較的影像
(J) 帳戶號碼	(R) 任何其他唯一識別號，特徵或代碼等
(K) 證照/執照號碼	

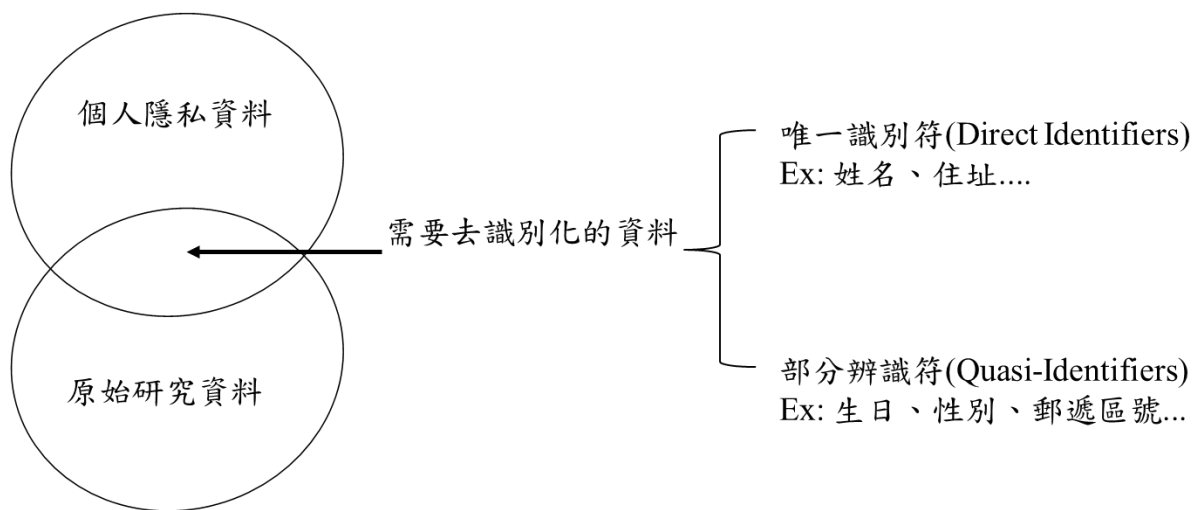


圖 2.2 個人資料去識別化之圖格

於醫療資訊，前述之 HIPPA 的框架，提供資訊之的資料庫宜經由「專家決定」法處理並闡明其「重新識別」的風險估計值，理論上其值應遠低於「安全港法」之「重新識別」的風險估計值；換言之，「安全港法」提供隱私防護與資訊共享間實作的參考準繩。

參、ISO/IEC 2nd WD 20889：2016-05-30之徵求意見稿

2015-12-18，ISO/IEC JTC 1/SC 27/WG 5 發出 ISO/IEC 20889 的計畫編輯(Project editor)，英國倫敦大學之 Chris Mitchell 教授提出的 ISO/IEC 1st WD 20889 草案(Text)之徵求意見稿；在此，探討其思路於後：

1. 去識別化技術

在闡明資料去識別化之實作使用的統計學工具(Statistical tool)與密碼學工具(Cryptographic tool)後(第 9~10 節)，先於第 11~15 節綜整抑制(Suppression)、擬匿名化(Pseudonymization)、泛化(Generalization)、隨機(Randomization)與彙集(Aggregation)之技術敘述如表 3.1 [1][2][3][4][6][9][10][11]：

表3.1 去識別化(De-identification)技術

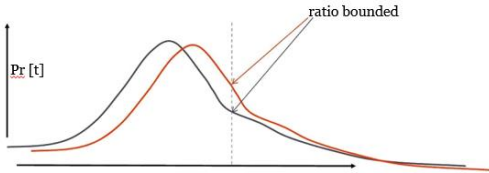
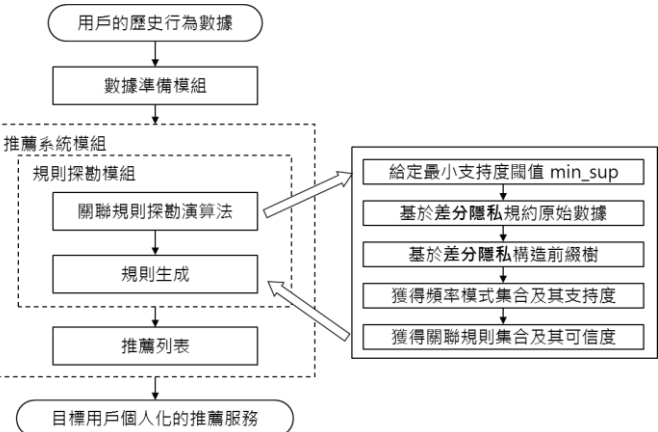
去識別化(De-identification)技術	實作方法
抑制(Suppression)	移除部分資料以降低資料集的準確性，可分成下列幾種類型： <ol style="list-style-type: none"> 1. 紀錄抑制(Record suppression)：刪除單一資料(如偏離值)。 2. 區域抑制(Field suppression)：刪除敏感資料類別。 3. 遮罩(Masking)：將所有唯一識別符移除。
擬匿名化(Pseudonymization)	將原始資料中的唯一識別符(unique identifiers)以擬匿名資料取代，保留不同資料庫之間的連結，但去除個人資料被識別的可能性。 碎映(Hashing)、匿名演算法等密碼學技術是常用之方法。
泛化(Generalization)	將原始資料修改至較概略的資料，但保留資料的準確性，包含下列做法： <ol style="list-style-type: none"> 1. 四捨五入(Rounding)：決定基準後，將原始資料數值調整至接近的泛化數值。 2. 上/下端編碼(Top/bottom coding)：針對資料的極值調整為大於或小於某個數值。
隨機(Randomization)	將原始資料內容修改為接近的資訊，以降低準確性，包含下列做法： <ol style="list-style-type: none"> 1. 增加雜訊(Noise addition)：將資料集中資料加入隨機數值，以減低資料被識別可能性，但保持資料集的統計資訊正確。 2. 置換(Permutation)：保持原始資料，但將資料內容彼此交換。
彙集(Aggregation)	將原始資料用其統計數據(如：平均、分布圖(例：直方圖).....)替代之。

2. 去識別化模型與合成資料

於第 16~18 節綜整 K-匿名模型(K-anonymity model)、差分隱私模型(Differential privacy model)與合成資料(Synthetic data)如表 3.2[3]：

表3.2 去識別化施行模型與合成資料

去識別化(De-identification)施行模型	說明
K-匿名模型(K-anonymity model)	透過隱藏或是概略化部分類識別資訊，使得符合相同條件之個體數量有 k 個以上，並擴增至相關聯的資料屬性等(例：L-diversity 與 T-closeness)，以達到避免識別之目的之系統化技術。

<p>差分隱私模型 (Differential privacy model)</p>	<p>透過數學方式調整資料內容，使得產生的資料在分析使用上和原資料效果相同，同時得以保護當事人之隱私。</p> <p>假定 K 為一給定之 ϵ-差分隱私(differential privacy)若原始資料庫(DB)中所有值，與調整後資料庫(DB')與其僅有單一元素(single element)不同，且於 K 的值域中所有之 S</p> $\frac{\Pr[K(DB) \text{ in } S]}{\Pr[K(DB') \text{ in } S]} \leq e^\epsilon \sim (1 + \epsilon)$ <p>當 K 的輸出是實數時，拉普勒斯(Laplace)機制(Mechanism)為其添加雜訊常用之方法；指數(Exponential)機制採用滿足特定分佈的隨機抽樣，取代添加雜訊之方法來實現差分隱私，擴增其應用範疇。</p> 
<p>合成資料 (Synthetic data)</p>	<p>經由事先定義之統計資料模型，人工創建能於分析面向代表母體，且不能映射到任一資料主體的資料集，實作時可以視為是「隱私防護資料探勘(Privacy Preserving Data Mining, 簡稱 PPDM)與「隱私防護資料發布(Privacy Preserving Data Publishing, 簡稱 PPDP)之融合。</p> <p>資料去識別化之合成資料框架例[11]：</p> 

資訊理論中的事前熵(priori entropy)與事後熵(posteriori entropy)是使用 K -匿名模型時，量測「資料去識別化後之效用」常用的方法，表 3.3 是 SDC 之技術與框架內含及效用損失舉隅，表 3.4 是去識別化的工具與技術，針對表 1.3 之重新識別的技術之屬性列表。表 3.5 是隱私保證層級考量的示意說明，可以作為對應資安防護基準之資料去識別化的比例原則之參考[1][4]。

表3.3 SDC 常用方法舉隅

方式名稱	實作方法例	SDC 評估
單元抑制 (Cell-suppression)	<p>針對特別敏感的資料優先抑制其數據的讀取(Primary suppression)，其次如果數據可能造成優先抑制數據被識別，也需抑制其數據讀取(Secundary suppression)。判別資料是否為優先抑制讀取之敏感資料常見方法有：</p> <ol style="list-style-type: none"> 1. (n, k)-dominance：如果資料中小於等於 n 個個體，卻對整體數值影響大於 k 比例，視為敏感資料 2. pq-rule：在未公開資訊前，個體對整體數值的影響估計在 p% 內；公開後影響估計在 q% 內，視為敏感資料。 3. p%-rule：當 q=100% 時，為 pq-rule 的特例。 	<ul style="list-style-type: none"> ● 效用損失(Utility loss)：可以用次識別數據(Secundary suppression)被抑制的數量代表，並可加入重要性權重估計。 ● 洩漏風險 (Disclosure risk)：計算敏感資料的破解可能性，如果在預設的安全值範圍內，視為安全。
差微分隱私法模型 (Differential privacy model)	<p>在一定範圍內可以加入固定的雜訊 (noise) ϵ，但仍然保持其原始的統計資訊。例如：可以利用動差生成函數 (Moment Generation Function, MGF) 控制，使得加入雜訊的動差生成函數與原來之要求(例：1~n 階的動差相差近於 0) 相同，於實際上兩者存在誤差 ϵ 之系統化技術。</p>	<ul style="list-style-type: none"> ● 效用損失：可用修改後資料與實際資料的差距總合表示。 ● 洩漏風險：ϵ 可視為其被識別風險。
K 匿名模型 (k-Anonymity model)	<p>透過隱藏或是泛化唯一識別符與部分識別符資訊，使得符合相同條件之個體數量有 k 個以上，並擴增至相關聯的敏感(個人隱私)資料屬性等(例：L-diversity 與 T-closeness)，以達到避免識別之目的之系統化的技術。</p>	<ul style="list-style-type: none"> ● 效用損失：可以利用統計數值結果估算。 ● 洩漏風險：若僅使用 k-匿名法，則 k 可視為其風險值。

表3.4 去識別化工具與技術之屬性(Properties)列表

Technique Name	Output Data	Data truthfulness	Applicable to types of values	Applicable to types of attributes	Reduces the risk of			
					單獨挑出 (singling out)	連結性 (linking)	推斷 (Inference)	不可分辨之分析 (indistinguishability analysis)
統計工具 (Statistical tools)								
取樣(Sampling)	Micro.							
彙集 (Aggregation)	Stat.	N.A.	Continuou s	Sensitive attributes	Yes	Yes	Yes	Partially
加密工具 (Cryptographic tools)								
確定性加密 (Deterministic encryption)	Micro.		All	Identifiers, quasi-identifiers, and sensitive attributes	No	Partially	No	No
保序加密 (Order-preserving encryption)	Micro.		All	Identifiers, quasi-identifiers, and sensitive attributes	No	Partially	No	No
同態加密 (Homomorphic encryption)	Micro.		All	Identifiers, quasi-identifiers, and sensitive attributes	No	No	No	No
同態秘密共享 (Homomorphic secret sharing)	Micro.		All	Identifiers, quasi-identifiers, and sensitive attributes	No	No	No	No

抑制 (Suppression)		Yes						
遮罩(Masking)	Micro.	Yes	Categorical	Identifiers	Yes	Partially	No	No
區域抑制(Local suppression)	Micro.	Yes	Categorical	Identifiers and quasi-identifiers	Partially	Partially	Partially	Partially
紀錄抑制(Record suppression)	Micro.	Yes	N.A.	N.A.	Partially	Partially	Partially	Partially
取樣(Sampling)	Micro.	Yes	N.A.	N.A.	Partially	Partially	Partially	Partially
擬匿名化 (Pseudonymization)		Yes	Categorical	Identifiers	Yes	Partially	No	No
泛化 (Generalization)		Yes	All, subject to meaning	Identifiers, quasi-identifiers, and sensitive attributes				
四捨五入(Rounding)	Micro.	Yes	Continuous	Ditto	No	Partially	Partially	Partially
上/下端編碼(Top/bottom coding)	Micro.	Yes	Ordinal	Ditto	No	Partially	Partially	Partially
隨機 (Randomization)		No		Identifiers, quasi-identifiers, and sensitive attributes				
增加雜訊(Noise addition)	Micro.	No	Continuous	Ditto	Partially	Partially	Partially	Partially
置換(Permutation)	Micro.	No	All	Ditto	Partially	Partially	Partially	Partially
微資料集彙集(Micro aggregation)	Micro.	No	Continuous	Quasi-identifiers, and sensitive attributes	No	Partially	Partially	Partially
彙集 (Aggregation)	Stat.	N.A.	Continuous	Sensitive attributes	Yes	Yes	Yes	Partially

說明：

1. 微資料集(Mircodata, 縮寫 Mirco.)定義為：由與個人資料主體(individual data principals)相關之記錄組合而成之資料集(dataset)。
2. 編者註：於 2015 年 10 月坦帕會議(Tampa meeting)，已同意再新增相關成本計算一欄於表中，唯其適當內容尚未決定。

資料來源：ISO/IEC WD 20889.2：2016-05-27, Information technology – Security technology – Privacy enhancing data de-identification techniques, pp. 21.

表3.5 ISO/TS 25237：2008-12-01第5.1.5.2節之隱私防護保證層級(Levels of assurance of privacy protection)的考量

	層級 1：識別個人資料元件關聯之風險(the risks associated with the person identifying data elements)。	層級 2：彙集資料變數關聯之風險(the risks associated with aggregating data variables)。	層級 3：母體資料庫中離群值關聯之風險(the risks associated with outliers in the populated database)。
保護措施	1. 移除明確可識別資訊以及可以輕易獲得的間接識別資訊	1. 滿足層級一要求 2. 考慮攻擊者會利用外部資料	1. 滿足層級二要求 2. 離群值納入考量
實作方法	應用「經驗法則(rule of thumb)」刪除個人資料。	將外部之各種資料庫納入考量，再刪除配合外部資料比較，可能識別的相關訊息	實作上有難度，目前尚無系統方法(通常為依個案設計)。
對應資安防護基準之安全等級	普	中	高
適用資料類型	一般性資料：資料外洩不致影響機關權益或僅導致機關權益輕微受損。	敏感性資料：外洩將導致機關權益嚴重受損。	機密性資料：外洩將危及公共安全、導致機關權益非常嚴重受損。

「個人資料保護法施行細則」第十七條，明文規定之「.....所稱無從識別特定當事人，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。」已被經濟部標準檢驗局於 2015 年 8 月 18 日提交行政院研商「個人資料去識別化過程驗證要求及控制措施」之驗證規範草案，採用為「驗證標準」；惟其無具攸關性之資料去識別化過程的「效用(utility)」亦即前述「資料去識別化後之效用」的諸如「資訊損失之閾值」的要求事項，亦無如表 1.3 所示之估計「重新識別風險」的規範；已成為資料去識別化核心技術之差分隱私與合成資料亦未述及，宜參照 ISO/IEC 2nd WD 20889 徵求意見稿的內蘊精進之。

肆、實作案例討論

「九層之塔，起於累土；千里之行，始於足下。」，前面所述之技術與模型雖然完備卻只是紙上談兵，其實際施行時仍需視個別資料庫之資料性質差異、其公開的目的、可以承受的風險.....等因素，做相對應的修正與調整。而風險評鑑也是如此，針對不同的實作方法與規劃，需要施行的風險評估也不相同。以下將藉由國內外實際案例，探討

如何實作資料去識別化與相關風險評鑑，並透過比較國內外規定之差異，以期能討論出如何制定更完善的去識別化要求事項。

4.1 國外案例 – 以HIPAA為例

4.1.1 實作方法

美國於1996年制定的HIPAA是國際間去識別化施行時間數一數二長，規範也較完整的例子。因此，如何實作「去識別化」可以作為參考，敘述如下。

HIPAA的相關隱私政策主要是針對PHI(受保護之醫療資訊, Protected Health Information)所設計的去識別化方法。其方法如圖2.1所示。HIPAA做法可以分為兩種：專家判斷(Expert Determination)(以後簡稱專家法)與安全港法(Safe Harbor)。專家法需要由相關領域的專家判斷該資料是否符合去識別化要求。其優點是可以保留的訊息相較之下較多，有利於研究分析，但是缺點是實作上可能需要花比較長的時間和人力。安全港法相較之下執行方法較為單純，僅需處理所有如表2.2所列之訊息後再做少部分調整，但是缺點是少數特殊情況很有可能會被忽略，而使個人資料洩漏。其假設與實作的說明如後：

1. 假設資料存取者並未具有實際相關知識，能將資訊以單獨或與其他資訊組合用以識別資訊之當事人。
2. 實作時需移除直接識別符(如：姓名)、間接識別符(如：居住地)、唯一識別碼(如：醫療紀錄編號)。

Project Data Sphere為一全世界共享癌症腫瘤資料的資料庫，其個人資料安全保護也採用HIPAA之設計，關於HIPAA專家法之實作將採用該資料庫做為實作說明。

專家法在去識別化設計上除了將直接識別符(如：姓名、社會安全碼)移除外，將資料依類別分為有部分識別可能性與沒有部分識別可能性，列舉如表4.1所示。除了具有獨特可識別號碼的病歷號碼與就醫場地編號屬於可部分識別符外，絕大部分的人口統計資訊也因為可以在個公開資源中找到也被歸類為可以部分識別。臨床資訊部分，大部分的醫療訊息並不會洩露當事人身分，但就醫日期與不良事件因為通常可以從中推斷出如死亡日期等訊息，所以被列為有潛在可能的部分識別。

表4.1 資料樣本中各類資料之可用性

資料屬性	具體資料類別	可部分識別
獨特可識別之號碼 Unique Identifying Numbers	病歷號碼	是
	就醫場地	是
一般人口統計資訊 General Demographics	出生日期	是
	死亡日期	是
	種族	是

	性別	是
	就醫地點	是
臨床資訊 Clinical Information	就醫日期	有潛在可能
	不良事件發生(例如：死亡)	有潛在可能
	治療方式	否
	診斷報告	否
	實驗室測試數據	否
	藥物治療	否
	身體質量指數(BMI)	否

專家法判斷是否需要處理之部分識別資訊與安全港法(表2.2所列)類似，惟差異在於專家法對於地理位置與年齡的限制較多，包含對於地理位置要求提升到以「區域」為單位，對於年齡能公開的範圍限制則降到84歲以下；而對於時間相關的資訊要求則較為寬鬆，包含死亡日期可以揭露到該年的相對周，而就醫日期則在不會造成直接識別出生或死亡日期的前提下被保留。同時，表4.1中的獨特可識別之號碼將用亂數替代，但仍持續代表該患者。

4.1.2 風險評鑑

HIPAA之實作方法，如先前所述，分為專家法及安全港法。但是因為安全港法是制式化的處理，其風險值不用再評估。下面將針對專家法的風險評鑑做法，進一步敘述。

在HIPAA中由於專家法能夠保留的資訊較多，也是較為偏好的方法。在這個情況下，如同Bradley Malin 在Project Data Sphere報告中所述，針對特定資料集清楚表達風險評鑑如何施行與其是否足夠安全是相當重要的[6]。

專家法中的專家一般指在統計、數學、醫療臨床或其他科學領域受過去識別化相關教育或有處理過相關事項的人員。專家會透過評估資料的可複製性(Replicability)、資源可用性(Resource Availability)、可辨識性(Distinguishability)與風險評鑑(Assess Risk) 並將資料依原則分成低、高風險區如表4.2，來決定資料的安全與否。專家們也依據表4.2來判斷申請存取資料庫的使用者可能可以完成資料重新識別到什麼程度。

表4.2 專家用於判斷醫療資料可識別性(identifiability)之決策原則

原則(Principle)	描述	舉例
可重複性 (Replicability)	將醫療資料依照其是否會重複發生於個體的機率排序。	低(Low): 病人口腔疾病風險與嚴重性。
		高(High): 病人的「人口統計數據」(例如：生日)則相對穩定。
資源可用性	決定哪些外部資源含有患者的	低：實驗室報告通常不對相關領

(Resource Availability)	識別符及可複製的醫療訊息，以及哪些人具有存取權限。	域外的人公開。 高：患者的身分及其「人口統計數據」通常可在公開資源中取得(例如生日、婚姻...)
可識別性 (Distinguish)	決定醫療資料中個體資訊可以被區分的程度。	低：依據估算，生日、性別與3位地區碼(Digit ZIP Code)在美國大約0.04%的人口具有唯一性[10]，也就是說只有少數的可由這組合被識別。 高：依據估算，生日、性別與5位地區碼(Digit ZIP Code)在美國大約50%的人口具有唯一性[9]，也就是說大多數的人可由這組合被識別。
評鑑風險 (Assess Risk)	上述的可重複性、資源可用性、可識別性越高，其風險也越大。	低：評鑑結果雖然可能具有可識別性，可是通常無法獨立複製結果也不會有許多人可以取得該資訊。 高：「人口統計數據」大多有高度可識別性、容易取得與複製。

資料來源：Bradley Malin, A De-identification Strategy Used for Sharing One Data Provider's Oncology Trials Data through the Project Data Sphere® Repository, pp.11,2013.

雖然風險實際上應該是更連續的，但專家們認為這個將資料依特性、種類分高、低的做法更容易說明各個原則是如何影響風險評鑑的，也較易執行評估。

在安全港法方面，其風險是透過計算機率估計而得，估算方式在此不作詳細討論。簡化來看，可將年齡、郵遞區號等特徵分成不同的組別，若總共有n個人，其中期待有i個人具有相同的特徵(落在同組中)，則其組別的機率可以表示為[5]：

$$f_i(n) = \binom{n}{i} b^{n-1} (b-1)^{n-i}$$

將k個組別計算之機率加總後可得到，安全港法之識別機率，表示為：

$$r_k(n) = \sum_{i=1}^k f_i(n)$$

若以Project Data Sphere方式計算，安全港法中約0.48%的美國人口具有唯一性。若進一步依專家法設定調整出生年份(85歲以上歸為同類)與居住地區(全部歸類於同一區)，估算出約0.000001%美國人口具有唯一性。

由於不確定資料庫使用者能夠取得那些資訊，Bradley Malin提出在保持統計數據(如：人口數目)一致的情況下，將資料採用隨機(Randomized)修改資料內容，以降低識別風險[5]。然而，若單純只保留統計數據的一致性，難保會有些統計結果因而改變。因此，若改採用如表3.3之合成資料模型，透過保留其統計結果再生成之數據，應可使其風險值再下降。

4.2 國內案例 – 以財政部財政資訊中心為例

4.2.1 實作方法

財政資訊中心率先運用「個人資料去識別化」驗證並於104年11月底前完成驗證程序；其依據104年9月17日會議紀錄函送、由經濟部標準檢驗局研訂之「個人資料去識別化過程驗證要求及控制措施」，實施實作。成為國內實作資料去識別化的首例，亦是目前(105年5月31日止)國內唯一實例。

該實作案例係以K匿名法(K-Anonymity)進行去識別化處理，設定目標為95%以上記錄均可釋出。若考慮戶籍地址及所得總額進行分群，戶籍地址概化處理需擴大至二級發布區或村里界，滿足K 門檻值22 之群組所佔紀錄方大於95%。該案例去識別化實作程序如圖4.1所示。

由於K-匿名法之間接識別欄位愈多，則分群愈細，每群之紀錄數量愈少，能達一定K門檻值之資料顆粒度將愈加粗糙。此案例最後考量資料顆粒度不宜過粗，故採用表4.3之方式處理，但不提供所得淨額、應納稅額、扶養人數及扣除額人數等欄位。

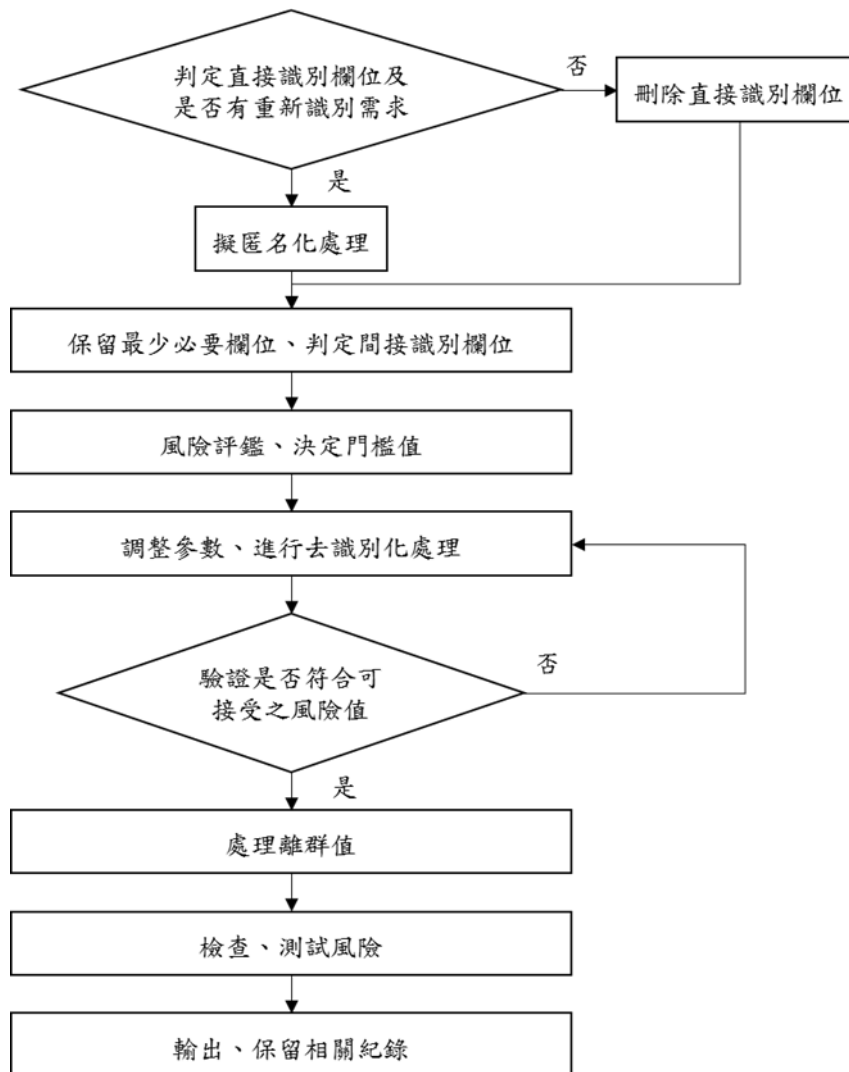


圖 4.1 財政資訊中心個人資料去識別化實作程序

表4.3 財政部財政資訊中心個人資料去識別化實作結論

威脅模型	資料開放
重新識別要求	無
K 門檻值	22
間接識別欄位	戶籍地址、所得總額
低度間接識別欄位	無
泛化處理方式	戶籍地址：村里 所得總額：10 等分位組
資料損失(%)	以群組計：18.8% 以筆數計：2.87%

4.2.2 風險評鑑

在財政部財政資訊中心實作個人資料去識別化時，其中有一個步驟就是進行隱私風險(衝擊)評鑑。風險值的估計方式為：風險值 = 衝擊值 × 重新識別可能性。

其中風險值設定可接受範圍為小於1.2；衝擊值由衝擊構面評分決定(如圖4.2)，在這個例子中，設定造成影響的構面有8個，並依嚴重程度給予1、3、5分，也因此衝擊值會落在8~40分之間。此實作案例之衝擊值為26分。

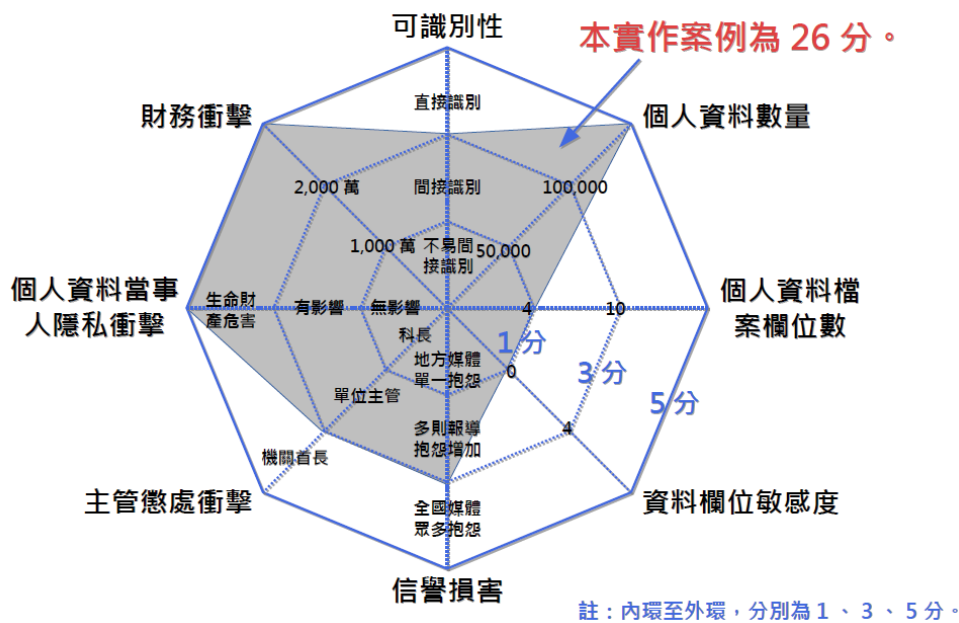


圖 4.2 衝擊構面評分

資料來源：財政部財政資訊中心，2016，個人資料去識別化過程驗證案例報告，2016年第2季資訊安全管理系統標準化系列討論會 會議手冊，頁198，2016-04-19。

重新識別可能性則由重現性、資源可用性、區別性三者中擇其最高者為代表，三者計算方式敘述如下：

$$\text{重現性估算公式為：重現性} = \text{權重} \times \frac{\text{鄧巴數}}{\text{資料集之總個體數}}。$$

$$\text{資源可用性估算公式為：資源可用性} = \frac{\text{可重新識別資料筆數}}{\text{重新識別測試資料筆數}}。$$

$$\text{區別性估算公式為：區別性} = \text{攻擊機率} \times \text{猜中的最高機率}。$$

在這個例子中可能性以區別性代表，其值為 $1/22$ ，因為K值設定為22，代表相同屬性之紀錄至少有22筆。最後再將衝擊值與重新識別可能帶入後可以計算出，此實作案例之風險值 $= 26 \times 1/22 = 1.18$ ，風險確實落在可接受範圍內[16]。然而，直接將兩個評斷標準相乘得到的風險值，是否真的具有代表性，仍需要討論。

經過詢問承辦人(謝明峰簡任分析師)後，其表示現已將實作方法由K-匿名法改成L-多樣性法。在K-匿名法中，雖然規定要透過調整使得符合相同條件之個體數量有K個以上，但是因為很有可能在這K個個體中具有特定敏感資料者只是少數，容易受到如表1.3提的單獨挑出(singling out)的攻擊。

L-多樣性法是以K-匿名法為基礎，增加了確保在K個個體中，至少有L個個體具有不同之敏感資料，以降低被識別的機率。此實作改採用L-多樣性法後，雖不確定L值為多少，但可以確定其小於或等於原先設定之22，重新識別的風險也會大於原先計算的 $1/22$ 。也就是說，此實作之風險值宜再上修，或是需要再調整其去識別化做法以滿足原先設定之風險值。

4.3 案例比較

由於HIPAA之安全港法僅需處理表2.2所列之項目，其風險已由政府計算與負責，在此僅討論HIPAA之專家法。財政部財政資訊中心實作是以行政院公布之《個人資料去識別化過程驗證要求及控制措施》為實作之參考依據(行政院，2015)。故在此以HIPAA隱私規定之專家法與《個人資料去識別化過程驗證要求及控制措施》針對其使用之去識別化方法與風險評估方式相互比較。

在實作技術方面，兩者皆是選用泛化(Generalization)、抑制(Suppression)作為主要之去識別化技術。然而，和HIPAA不同，財政部財政資訊中心實作中並沒有考慮到資料公開後的還能否被用於研究目的之效用問題。如果資料效用低於實作去識別化的成本，那麼基於比例原則的考量，或許選擇不公開該資料庫會是較為務實的做法。

在隱私衝擊評鑑方面，HIPAA則並未特別規範之。其原因在於美國已於其電子化政府法(E-Government Act of 2002)第2篇(Title)第208節(Section)中將隱私影響評鑑(Privacy Impact Assessments)納入規範中，以確保施行電子化政府後，對於個人隱私資訊各機構能提供有足夠的保護。為了追蹤電子化政府法所要求之指標效能，主責的管理與預算辦公室(Office of Management and Budget, OMB)依據聯邦資訊安全管理法(Federal Information Security Management Act, FISMA)，亦即電子化政府法之第三章，要求各機關/構繳交「FISMA 實作計畫」報告，要求各機關/構以經濟有效(cost-effectively)的方式降低資訊安全風險至可接受之範圍，而HIPAA也在其規範範圍中；也就是說，HIPAA其完整風險評鑑流程應如圖4.3所示。

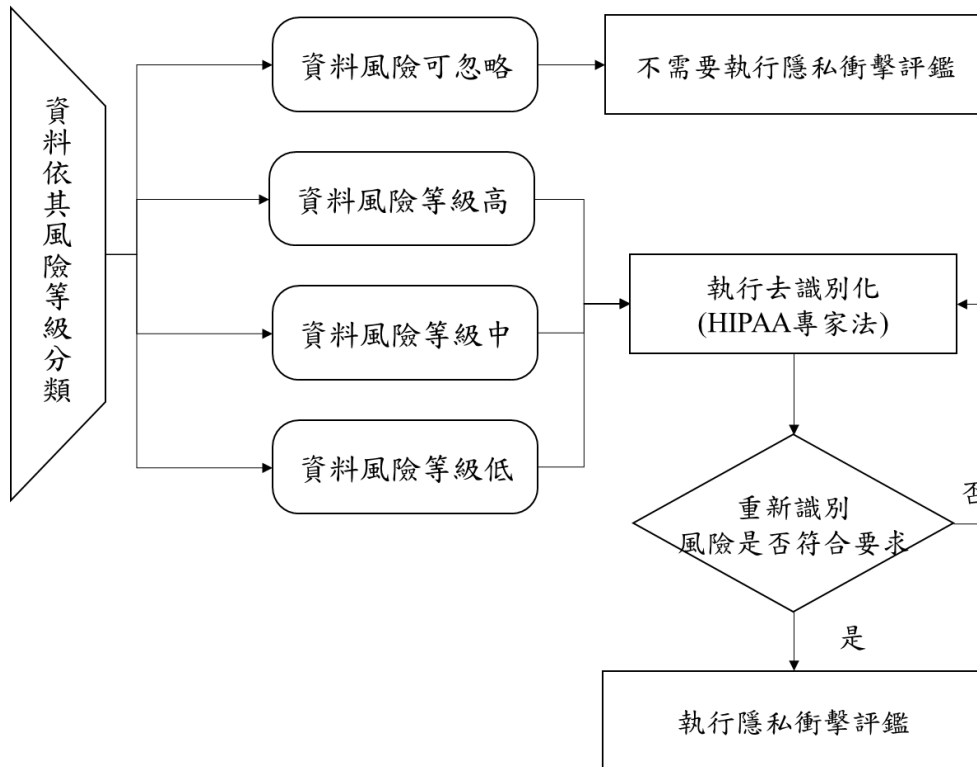


圖 4.3 HIPAA 風險評鑑流程圖

另一方面，財政資訊中心的實作也於衝擊值做了相對應的分析與評估，但是，該案例是特別針對要處理的資料設計對應之風險評鑑的，若是要評估其影響性，財政資訊中心的例子在這個部分評估似乎較為完善，惟未進行「重新識別風險評鑑」是其缺失(備考：2016-07-11經工研院此計畫執行者電話確認)，敘述於後。

在重新識別之風險評鑑上，由於財政資訊中心的案例並沒有考慮資料效用，因此也並未進行表1.3之單獨挑出等重新識別應計算的風險，僅透過計算重新識別可能性，而得到風險數值。但是，被重新識別的機率計算結果為1/22，已遠大於HIPAA任何一種方法的重新識別風險機率之估算。經與財政部財政資訊中心承辦人於電話及電子郵件確認，目前計畫已改採L=2之L-多樣性法，亦即其風險值可能需要上修。另一方面，專家法的風險值理論上應遠低於安全港法，Bradley Malin估計其風險值可以比安全港法小約100倍[5]，也就是重新識別風險值約僅為1/100000。

然而，若再將財政資訊中心案例K匿名法之K值增大，其資料顆粒會更加粗糙、可用性也會降低。因此，本論文建議此實作不妨考慮改採用表3.3所提的差分隱私模型(Differential privacy model)，透過些微調整資料庫資料，可以同時兼顧資料效用與個人隱私保護。

在判別安全性的部分，HIPAA之專家法是以專家分析判斷作為依據；而財政資訊中心的案例則是以驗證合格證書為依據。一般相信專家團的判斷具有一定的公信力，能保證其做法的安全性；但由於驗證合格證書是由稽核人員所發出，稽核人員的知識與技能

是否足夠會成為該案例是否真的安全的關鍵。

在遇到糾紛處理時，HIPAA之安全港法已有政府保證、專家法也有專家團背書，較容易作為舉證或判決之依據。而國內的制度憑藉的是稽核人員的專業程度，遇到糾紛時較容易被質疑。以現今情境，本論文認為，若在不改變國內制度的前提下，需要提升稽核人員的專業度，以降低不被信任的可能性。

綜上所述，彙整HIPAA隱私規定之專家法與《個人資料去識別化過程驗證要求及控制措施》針對其使用之去識別化方法與風險評估方式比較如表4.5所示。

表4.5 HIPAA隱私規定之專家法與《個人資料去識別化過程驗證要求及控制措施》實作之比較

	HIPPA 去識別化隱私規定之專家法	《個人資料去識別化過程驗證要求及控制措施》(財政部財政資訊中心實作為例)
實作隱私衝擊評鑑	無特殊規範，遵循 FISMA 要求	有
重新識別之風險評鑑	有	無
風險評估方式	以高、低等級區分	估算風險值，確認落在可接受範圍。
風險值估計	約 1/100000	大於或等於 1/22
資料效用(data utility)	有保留用於研究目的之資料效用[6]。	無
判別安全性依據	以專家報告為依據[18]。	以驗證合格證書為依據
糾紛處理難易	依據較可被大眾信任，較容易被接受做為參考。	稽核人員之專業度是決定是否被接受的關鍵。

伍、結論

隨著資訊科技之進展，大數據與開放資料等應用需求的發展已勢不可當，如何保護數據發布之個人隱私及敏感資訊洩漏，已成為資訊社會關注的議題。

參照ISO/IEC 2nd WD 20889徵求意見稿的思路，應先將「資料去識別化後之效用」與「差分隱私」納入「個人資料去識別化」的隱私防護之標準化；以及如圖5.1所示，擴增「資訊安全管理系統(Information Security Management System, 簡稱ISMS)的要求事項與控制措施之「個人資料管理系統(Personal Information Security Management, 簡稱PIMS)的標準化之進程僅為一端。我國因個人可識別資料去識別化已是法規的要求事項，宜參照「健康資訊安全管理系統之控制措施(ISO 27799)」在「資料去識別化」的議題要求遵循ISO/TS 25237般，擴增之。

2004年6月14日，行政院院臺規字第0930086121號函頒之「行政院所屬各機關主管法案報院審查應注意事項」的第三點第(四)款規定：「法案衝擊影響層面及其範圍，包括成本、效益及對人權之影響等，應有完整之評估。」，以「個人資料去識別化過程驗證要求及控制措施」行政規則的法制作業之過程與試辦機關的實作結果評估，其「法規衝擊評估(Regulatory Impact Analysis, RIA)」作業宜精進之。

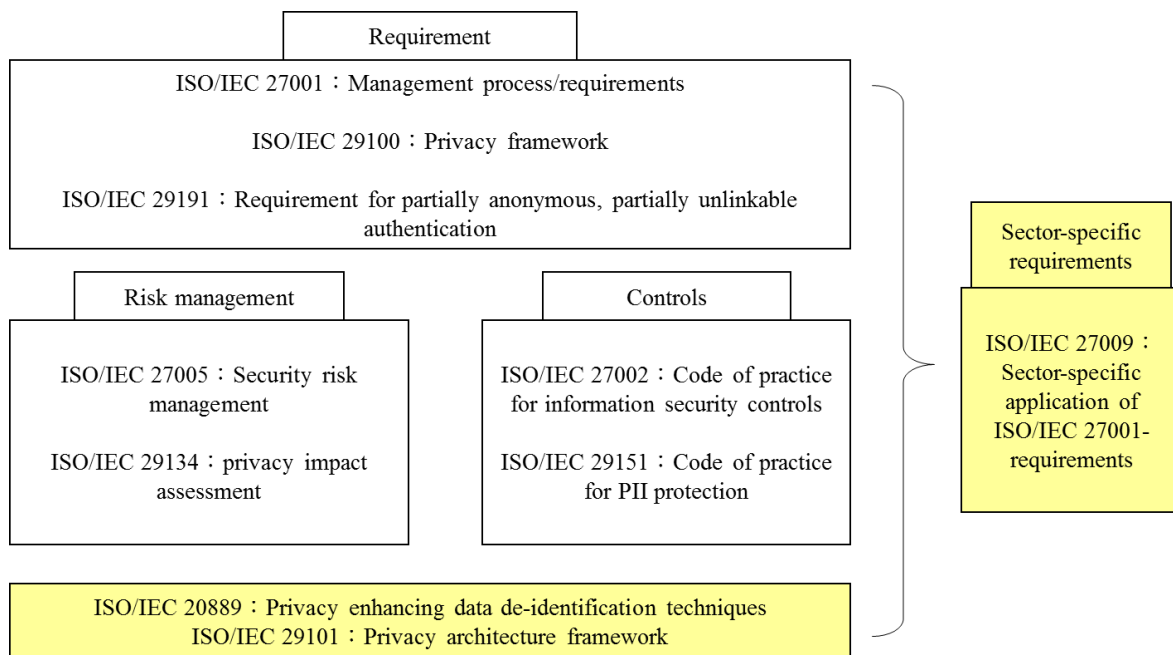


圖 5.1 個人資料匿名化與去識別化之隱私資訊管理系統驗證框架

資料來源：ISO/IEC 2nd CD 29151, Figure 1, page X, 2015-12-16

隨著「個人資料去識別化」等隱私防護議題實作之開展，僅確保資訊系統的機密性 (Confidentially, C)、完整性(Integrity, I)與可用性(Availability, A)並不足以確保民眾數位生活福祉；2014年12月，歐盟已正式發布將CIA擴增：

1. 去連結性(Unlinkability)：隱私相關之資料不能跨資料庫彼此連結。
2. 透明性(Transparency)：可以在任何時間理解與重建，包含法規、技術以及組織設置之所有隱私相關的資料處理。
3. 調解性(Intervenability)：對計畫與正在進行之隱私相關的資料處理，能進行合理的干預。

前述CIA定義之擴增，已納入於2016年4月成案的通稱為「從設計著手保護資料(Data protection by design)」之「隱私工程(Privacy engineering)」的標準化計畫之先期研究的內容中；2013年10月15日公布通稱為「以預設機制進行資料保護(Data protection by design)」的「隱私架構框架(Privacy architecture framework)」之ISO/IEC 29101與「隱私工程」的標準，通稱為「從設計著手保護隱私(Privacy-by-design, PbD)」，PbD及「資料最小化

(Data minimization)」是個人資料防護實作之原則[1]。

美國聯邦政府自2008年起，開展整合個人資料管理系統與資訊安全管理系統之整合性安全管理系統(Integrated Security Management System, IISMS)的工作項目；2013年4月，公布第4版之NIST SP 800-53作為IISMS的控制措施之作業規範，並於2014年12月公布第4版的NIST SP 800-53A作為其稽核作業評鑑之標準；2014年12月18日，美國公佈「聯邦資訊安全現代化法(Federal Information Security Modernization Act of 2014, FISMA 2014)」的3552(b)(3)(B)條款將「個人隱私(Personal Privacy)」納入並實作之，於送交美國國會的2015財年之FISMA 2014的年度報告中已闡明「資料去識別化」具有效性；2015年5月，公布同前述歐盟擴增CIA之「分離性(Disassociability)」、「可預測性(Predictability)」及「可管理性(Manageability)」的3項定義之徵求意見草案[7]；換言之，前述美國聯邦政府的整合性安全管理系統(IISMS)已建立法源。

「他山之石，可以攻玉」，我國正進行中的「資訊安全管理法」宜納入個人資料之防護，並由主管機關公布表3.5中普、中、高級之「個人資料重新識別風險」的「閾值」之參考值(例：HIPPA的安全港法(普級)、HIPPA之專家法(中或高級))，俾供法務部法律事務司基於「個人資料去識別化」的RIA，遵循「行政程序法」第7條，提出之「比例原則」的參考，以利實作。

致謝詞：本文作者謹在此對ISO/IEC JTC 1/SC 27/WG 1之友人提供ISO/IEC 2nd WD 20889：2016-05-30、ISO/IEC 2nd CD 29134：2015-12-16以及ISO/IEC 2nd CD 29151的盛情，與審稿者提昇內容水平之意見，致衷心的謝忱！

備考：本文完整內容將於2016-11-25/26之「第20屆全國科技法律研討會」中發表，參考文獻中的公文函件均為「資訊公開法」應公開資料，惟幾均無供直接檢索之網址，應審稿者要求，囿於文長，於參考文獻中加註說明。

參考文獻

- [1] European Union Agency for Network and Information Security (ENISA), “Privacy and Data Protection by Design – from policy to engineering”, 2014-12.
- [2] S. L. Garfinkel, “De-Identification of Personal Information”, *NIST IR 8053*, October 2015.
- [3] ISO/IEC WD 20889：2016-05-30, Information technology – Security technology – Privacy enhancing data de-identification techniques.

- [4] ISO/TS 25237:2008-12-01, Health informatics – Pseudonymization. (備考：2015-06-17，此份標準於 ISO/TC 215/WG 4 自 2012-08 起之審核，已完成行政程序，ISO 公佈結論：不修訂，繼續使用)
- [5] B. Malin, “Sharing Pre-Competitive Clinical Trials Data to Facilitate Cancer Research: The Data Sphere Experience”. Retrieved from http://ottawagroup.ohri.ca/docs/Bradley_Malin_2013.pdf, 2013.
- [6] B. Malin, “A De-identification Strategy Used for Sharing One Data Provider’s Oncology Trials Data through the Project Data Sphere® Repository”, 2013.
- [7] Office of Management and Budget, “Annual Report to Congress : Federal Information Security Modernization Act”, March 18, 2016.
- [8] L. Sweeney, “Uniqueness of simple demographics in the US population.” *Technical report*, Carnegie Mellon University, 2000.
- [9] L. Sweeney, “k-anonymity: A model for protecting privacy.” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05), 557-570, 2002.
- [10] L. Sweeney, “Testimony before the National Center for Vital and Health Statistics Workgroup for Secondary Uses of Health information.” August 23, 2007.
- [11] 丁麗萍、盧國慶，發明專利申請公布號，CN：104050267；發明名稱：基於關聯規則滿足用戶隱私保護的個性化推薦方法及系統；申請日：2014-06-23；申請人：中國科學院軟件研究所；中華人民共和國知識財產局，申請公布日期：2014-09-17。
- [12] 行政院，行政院院臺法字第 1010056845 號令，2012。(個人資料保護法除第 6 條及第 54 條條文外，其餘條文自 2012 年 10 月 1 日施行；2016 年 2 月 25 日，院臺法字第 1050154280B 號函，自 2016 年 3 月 15 日施行)
- [13] 行政院，〈我國個人資料保護法有關去識別化之標準〉，院臺科字第 1040144764 號函(附件 1)，2015。(闡明「法務部」對「去識別化」應遵循「比例原則」等法律意見)；〈個人資料去識別化過程驗證要求及控制措施〉，院臺科字第 1040144764 號函(附件 2)，2015。(ISO/IEC 29191 之要求事項已敘明為「選項」)
- [14] 行政院資通安全辦公室，院臺護字第 1050150057 號函，2016-01-05。(2015 年 12 月 21 日，「行政院國家資通安全會報第 29 次委員會議紀錄」決定：(四)經濟部標準檢驗局研訂之個人資料去識別化驗證標準及驗證要求與控制措施，可供推動大數據分析或開放資料等業務，作為個人資料匿名化、去識別化之準則，請相關機關善加推廣運用)
- [15] 法務部，法律字第 10303513040 號函，2014。(2014 年 11 月 17 日，「法務部」函釋闡明：「去識別化之個人資料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」)
- [16] 財政部財政資訊中心，個人資料去識別化過程驗證案例報告，2016 年第 2 季資訊安全管理系統標準化系列討論會 會議手冊，頁 192~208，2016-04-19。

- [17] 最高行政法院，判字第 600 號，2014。(2012 年 5 月 9 日起，上訴人主張「衛生福利部中央健康保險署」釋出給第三者之「個人全民健康保險資料」並不符合「個人資料保護法」，要求行使「刪除權」，被拒絕後，提起訴願/訴訟，均遭駁回，上訴人猶不服，再提起上訴；2014 年 11 月 13 日，「最高行政法院」判決主文：「原判決廢棄，發回臺北高等行政法院」，開啟我國「個人資訊去識別化」法規實作之議題)
- [18] 楊智傑，“美國醫療資訊保護法規之初探：以 HIPAA/HITECH 之隱私規則與資安規則為中心”，*軍法專刊*，第 60 卷第 5 期，頁 79-116，2014 年 10 月。
- [19] 經濟部，經標授字第 10420050540 號函，2015。(2015 年 7 月 27 日，經濟部召開「研議政府機關個人資料去識別化之適用標準」，會中作者之一應行政院此議題承辦人朱俊銘檢察官力邀，提出：「僅採用 ISO/IEC 29191 之不足，宜納入 ISO/TS 25237 與 PbD 的第一份標準：ISO/IEC 29101 等觀點」；行政院指定之試辦單位承辦人謝明峰簡任分析師亦提出其綜覽之文獻僅提及 ISO/TS 25237，未見 ISO/IEC 29191 的發言；法務部李世德科長發言提出 ISO/IEC 29191 之名稱即不符合「法務部」函釋的要求，ISO/TS 25237 看起來比較像是相關標準。2015 年 8 月 3 日，經濟部函送前述會議之會議紀錄決議納入 ISO/TS 25237 等，作為中長期目標。)