擬匿名化(Pseudonymization)的大數據(Big Data)之安全標準初探：根基於支付卡(Payment Card)的安全事故與公開金鑰基礎建設(Public Key Infrastructure，簡稱 PKI)之技術脆弱性的議題

樊國楨[1*]、蔡昀臻[2]
[1]異術科技股份有限公司、[2]國立交通大學管理科學研究所
[1]kjf.nctu@gmail.com、[2]yct1230@gmail.com

摘要

隨著「大數據」與「資料探勘」之興盛發展，「開放資料」的去識別化之議題已成為保護個人資料應面對的工作項目，法務部於「我國個人資料保護法有關去識別化之標準」中提出：「應進行整體風險評估，針對不同資料類型或資料提供方式，依比例原則分級控管去識別化程度」，闡明「開放資料」宜達「匿名化(Anonymised)資料(data)」或「不可逆(Non-retraceable)之擬匿名化(Pseudonymised)資料」亦即一般稱為「無可控制的重新識別之擬匿名化(Pseudonymization without controlled re-identification)」的程度較為妥適。根基於此，本文探討已發生之擬匿名化的資訊安全事故，提出其依比例原則宜求索之資料開放的控制措施框架以供探討使用之。

關鍵詞：匿名化、去識別化、開放資料、擬匿名化、脆弱性

Security Standards of the Big Data Pseudonymization: Based on Payment Card Security Incidents and Public Key Infrastructure (PKI) technical vulnerability issues

Kwo-Jean Farn[1*], Yun-Chen Tsai[2]
[1]Exsior Data & Information Technology, Inc,
[2]Institute of Information Management Science, National Chiao-Tung University
[1]kjf.nctu@gmail.com, [2]yct1230@gmail.com

Abstract

With the development of “Big Data” and “Data Mining”, the issue of “open data” has become an issue in protecting personal information. The Ministry of Justice states in “Standards of Personal Information Protection Act related to De-identification” that "The overall risk assessment should be carried out based on the degree of identification in

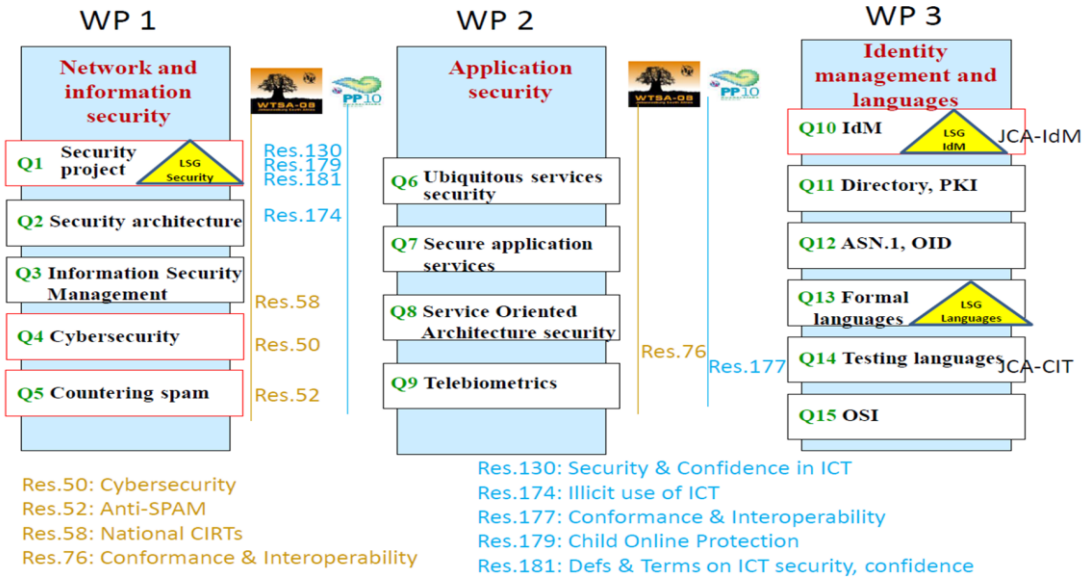
[*] 通訊作者 (Corresponding author.)

accordance with different data types or data provided ways.”. Clarify that it is more appropriate if the “open data” is at the level of "anonymized" or “non-retraceable Pseudonymised”, which is also known as “Pseudonymed without controlled re-identification”. Based on this, this paper discusses on the pseudonymation data security incidents that have occurred, and puts forward the would-be framework of data open controls based on the principle of proportionality.

Keywords: Anonymised, De-identification, Open data, Pseudonymised, Vulnerability

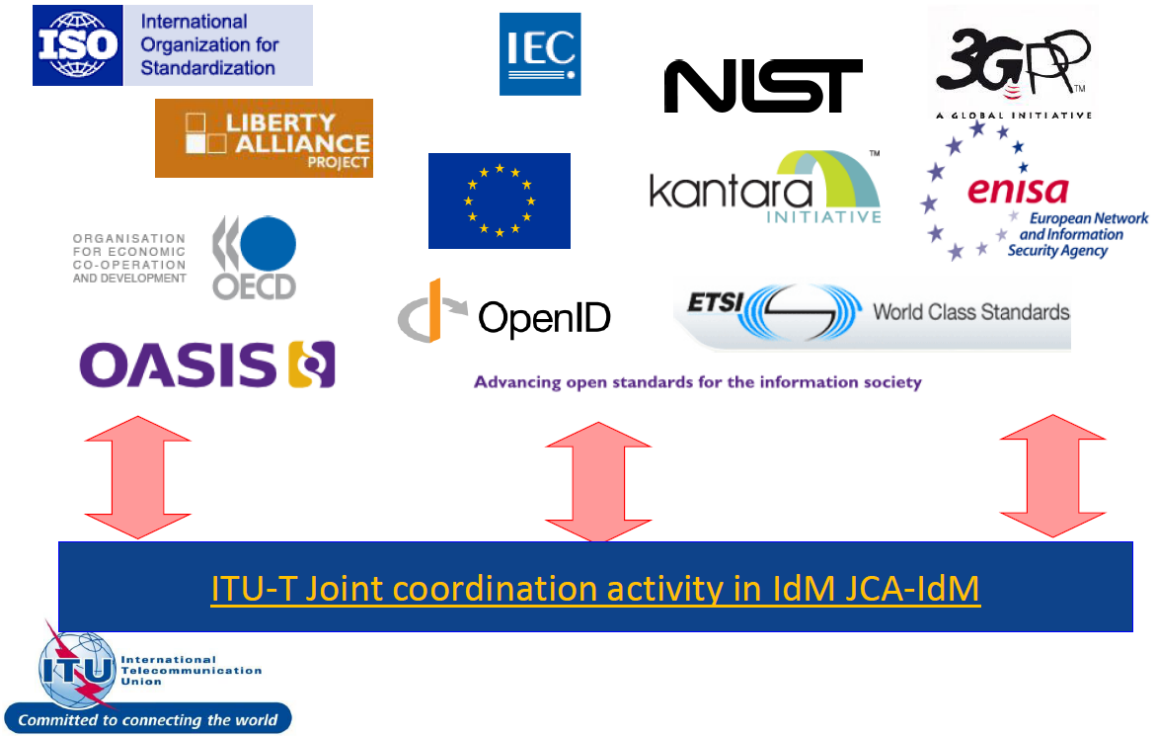
壹、前言

隨著網際網路、無線通訊與數位電視等科技之日新月異及電子技術的一日千里，電信網路、網際網路以及廣電網路(三網)融合已成為資訊社會生活之情境。當寬頻網路問市、線上交易啟用、網路社群誕生時，數位社會的落塵也散落人間，許多伴隨著新世代網路之資訊安全問題紛紛湧現，如何建立相關之生態體系並確保其可信賴性，已成為推動下一世代網路(Next Generation Network，簡稱 NGN)社會的全世界之共識；NGN 的安全(Security)標準系列(Y.2700~Y.2779)由國際電信聯盟(International Telecommunication Union，簡稱 ITU)之第 17 技術研究小組主責(ITU-T SG 17)其工作項目與公開金鑰基礎建設(Public Key Infrastructure，簡稱 PKI)相關技術標準化與合作團體分如圖一及圖二所示。



資料來源：http://www.itu.int/ITU-T/studygroups/com17/index.asp (2013-03-29)

圖一：ITU-T SG 17工作項目框架(2013-03-28)



圖二：ITU-T SG 17於身分管理(Identity Management)之合作團體

隨著 PKI 技術使用之普及，PKI 已成為三網融合身分管理安全的基石，其公鑰(Public Key)存取之脆弱性引申出的資訊安全議題已成為其標準化之實然與應然的具攸關性之問題；在另一方面，NGN 以封包傳輸為核心，提供無所不在的服務之特性，使得網際網路平台提供者的安全成為三網融合具攸關性之安全議題。

2012 年的上半年，臺灣之資訊安全一點也不平靜，先是驚傳海巡署情報處被植入惡意程式竊取情資；接著，法務部資訊處亦遭攻陷，洩露公務機密。前述 2 機構均已通過資訊安全管理系統(Information Security Management System，簡稱 ISMS)的驗證多年，後者曾獲頒資訊安全成就獎，於事故發生後，回覆記者詢問時，答以：「技不如人」；至 2012 年 8 月 30 日止，行政院主責資訊安全之官員，仍於會議中表示，張政務委員認為我國 ISMS 已有績效，惟若於 2012 年 10 月 1 日之後方發生前述的資訊安全事件，海巡署及法務部均將面對個人資料保護法第 28 條之 N.T.$ 200,000,000.的求償？本文作者之一於 2011 年 1 月 4 日曾致函前法務部資訊處陳處長，闡明此地對 ISMS 驗證的要求事項可能過低之觀點；2012 年的臺灣駭客年會(HITCON 2012)之 2012 年 7 月 21 日，下午 13 時至 13 時 35 分的「現實生活中之密碼分析(Cryptoanalysis in Real Life)」場次中，說明於臺灣地區的 PKI 存在可離線破解憑證(Certificate)對應之公鑰(Public Key)的密鑰(Private Key)之後檯作業脆弱性(Vulnerability)經確認屬實的歷程，即

為例證；換言之，資訊安全管理的要求事項不如人亦為前述資訊安全事故發生的根本原因之一[1~3,16~18,21~23]，值得反思。根基於此，本文就我國已發生過的資訊安全事故，於第 2 節闡明擬匿名化(Pseudonymization)之脆弱性；在第 3 節探討擬匿名化宜求索的技術性與組織性之防護措施框架；最後，在第 4 節提出本文之結論。

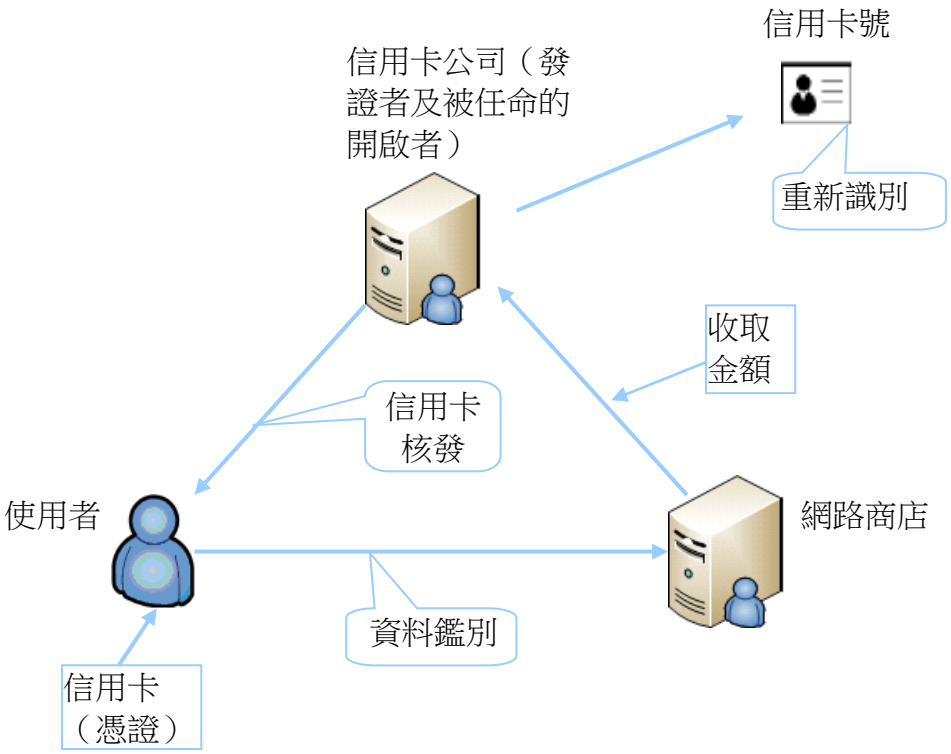
貳、信用卡之擬匿名化控制措施缺失的大數據資訊安全事故

使用者有付費帳號，諸如銀行帳號及信用卡帳號等。在網路商店購物時，使用者有一合法付費帳號是重要的。然而，輸送確實的帳號到網路商店是有風險的，因為可能被網路商店濫用。另一方面，網路商店在接收顧客確實帳號上也有風險，並且網路商店可能需要付出額外成本以保存該等資料安全不被危害。為資料鑑別及資料保護之目的，CNS 29191 之部分匿名、部份去連結鑑別的要求事項已應用於此情境。

會有銀行或信用卡公司，將核發帳號給使用者。會有想要從網路商店購買某些商品的使用者。網路商店會有一軟體查核使用者是否有合法之付費帳號，同時若使用者有，則執行商品銷售過程。銀行或信用卡公司將扮演被任命的開啟者之角色。

開啟其銀行帳號或信用卡帳號之使用者參與從事銀行或信用卡公司之核發過程。使用者收到一憑證。當使用者想要購買某些網路商店的商品時，使用者參與從事網路商店的使用者鑑別並證明該使用者具有申請銀行或信用卡公司的合法付費帳號。經由部分匿名，部份去連結鑑別之使用，網路商店不需要得知該使用者的確實帳號，即可驗證使用者確實在申請組織有一合法帳號。日誌項目將為商品名稱；購買日期及使用者宣稱有帳號的銀行或信用卡公司名稱。交易腳本為去連結的，亦即不可能從日誌項目中發現曾經使用該同一帳號購買的其他商品。網路商店轉送該交易腳本給申請組織。組織、銀行或信用卡公司，執行交易腳本上的重新識別過程，並成功地獲得確實的付費帳號。

使用者對網路商店未得知其帳號是自在的，網路商店對其不需要得知隱私資訊是自在的。前述之交易腳本中的授權碼等資訊係由信用卡核發時僅於信用卡紀錄之卡號外加的 3 或 4 位數經由密碼演算法產生，作為交易腳本之「身分」證明，圖三是其擬匿名化框架的示意說明。



圖三： 信用卡之隱藏付費帳號框架

前述之「敏感驗證資料」於「支付卡產業安全標準(Payment Card Industrial Data Security Standard, 簡稱 PCIDSS)」的要求事項，如圖四所示，圖四中之「PCIDSS 要求 3.4 係指 PCIDSS 第 3.4 項的要求事項」；此要求事項，PCIDSS 至今仍同。在資訊系統中僅用於列印交易腳本並不允許儲存；2011 年 7 月以前，我國一金融機構因未遵循前述之要求事項，儲存前述的「敏感驗證資料」，因於大數據之前述的「敏感驗證資料」，可以計算出(例：字典攻擊法)據以產生前述交易腳本所需之信用卡紀錄之卡號外加的 3 或 4 位數字，用以製造偽卡，致發生如表一、表二與表三所示之資訊安全事故。

	資料元素	允許儲存	需要保護	PCI DSS 要求 3.4
持卡人資料	主帳戶 (PAN)	是	是	是
	持卡人姓名[1]	是	是[1]	否
	業務代碼[1]	是	是[1]	否
	失效日[1]	是	是[1]	否
敏感驗證資料[2]	完整磁條資料[3]	否	不存在	不存在
	CAV2/CVC2/CVV2/CID	否	不存在	不存在
	PIN/PIN 資料塊	否	不存在	不存在

說明：PAN：Primary Account Number。

備考：

1. 這些資料元素如果連同 PAN 一起儲存，則必須對其進行保護。該保護措施應當符合 PCI DSS 對持卡人資料環境一般保護的相關要求。此外，如果在業務過程中收集與消費者相關的個人資料，其他立法（例如，與消費者個人資料保護、隱私、身分盜竊或資料安全有關的法律）可能要求對此類資料進行特別保護，或要求對公司操作進行適當披露。然而，如果不對 PAN 進行儲存、處理或傳輸，則 PCI DSS 不適用。
2. 敏感驗證資料不應在驗證後儲存（即便是經過加密的）。
3. 來自磁條、晶片上的磁條圖形或其他地方的全磁軌資料。

資料來源：PCI安全標準委員會(Security Standards Council)：PCI DSS要求與安全評估程序，v1.2，2008-10。

圖四：PCI DSS 適用性資訊

表一：擬匿名化的ISO/IEC 29191：2012(E)附錄B之控制措施缺失的大數據資訊安全事故案例

事例	信用卡交易時隱藏付款帳戶之機制，遵循 ISO/IEC 29191：2012(E)的要求事項；其應用，交易授權後，不得儲存驗證敏感資料為其資訊安全之控制措施；臺灣一金融機構因未落實前述控制措施，2000-01~2011-05 間被盜賣，造成 52 家發卡銀行數十億元的財產損失。
參考資料	支付卡產業資料安全標準(Payment Card Industry Data Security Standard，PCIDSS)(1.0 版)與臺灣高雄地方法院檢察署檢察官起訴書(2003-01-10)。
備考	擬匿名化於歐盟屬個人資料保護之範疇。

表二：財金公司資訊外洩案例

資料來源	1. 2002 年 9 月 17 日，自由時報 1 版，記者詹遜鴻/台北報導。 2. 2002 年 9 月 18 日，聯合報 1 版，記者曹敏吉、欒丕智/連線報導。 3. 2002 年 9 月 18 日，中國時報 3 版， 吳江泉、林憲祥/高雄報導。
案例	高雄地檢署查緝黑金中心今(2002)年 9 月 17 日召開記者會，宣佈和調查局高雄縣調查站等單位共同查獲國內黑幫分子黃 OO 等偽卡集團，勾結財政

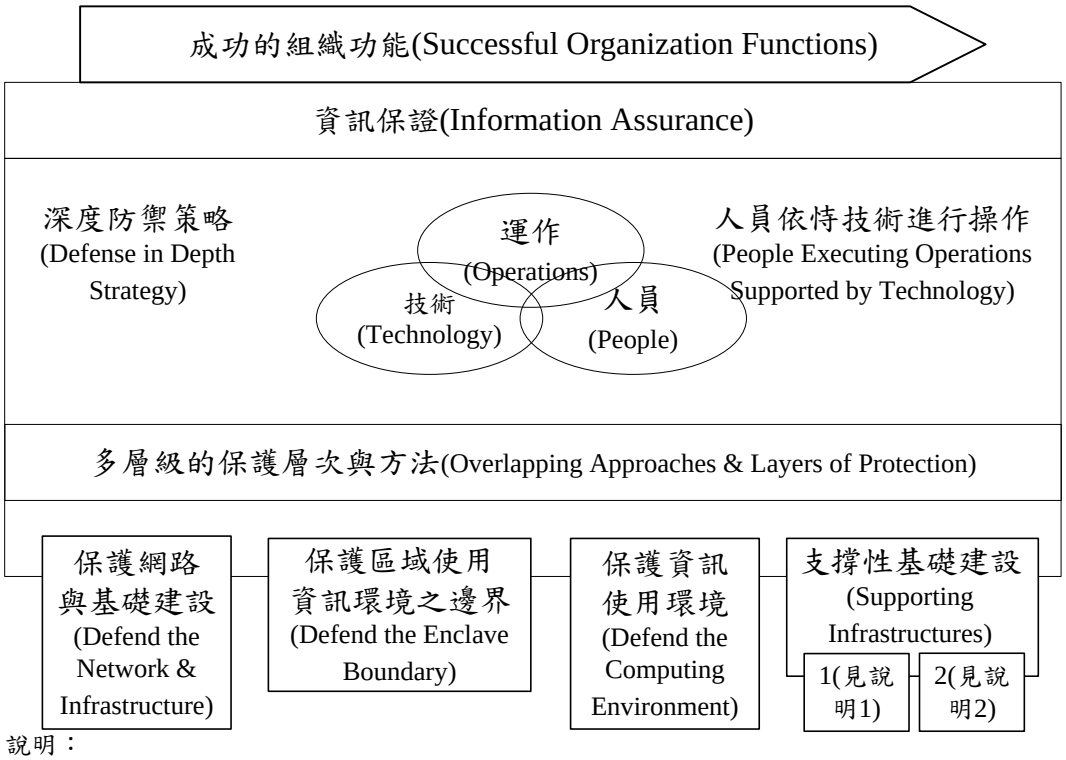
	部所屬的財金資訊股份有限公司(簡稱財金公司)職員張 OO、張 OO、賴 OO 等三名工程師，盜取客戶信用卡內外碼資料高達 100 萬筆以上及金融卡資料，再由北、中、南 3 個大盤商銷售給偽卡集團，製成偽卡盜刷；據銀行初步統計，遭盜刷金額至少有 N.T.$ 3,500,000,000.-。
說明	財金公司發言人今年 9 月 17 日說明自 2001 年下半年起，財金公司將信用卡資料改為亂碼處理，所以外洩資料應以 2001 年下半年以前為主；至於檢調單位所說金融卡資料外洩 10,000 筆部分，應是簽帳金融卡(Debit Card)並非一般之金融卡。
教訓	資訊安全管理依恃技術控制措施已是我國必須正式的問題。

表三：財金公司人員疑似涉及信用卡案具體改善成果

資料來源	林真真(2002) 善盡企業責任化解危機總動員，財金資訊，Vol. 24, 頁 10~13。
系統面	● 90 年 6 月：信用卡授權交易資料庫敏感性資料以亂碼儲存。 ● 90 年 9 月：建置自動化風管監控系統，加強權限人員取得資訊之控管。 ● 91 年 3 月：建置 PRISM 風險偵測系統，提供發送消費簡訊警示服務。 ● 91 年 10 月：取消信用卡系統傳輸層紀錄檔(Log)之敏感性資料儲存。 ● 91 年 11 月：為根本杜絕資料外洩可能，將信用卡授權交易資料庫敏感性資料由原亂碼儲存方式，改為完全不留存。
業務面	● 90 年 7 月：全面清查備份光碟並銷毀相關之敏感性資料。 ● 90 年 12 月：提供銀行針對高風險特店進行檢核功能。 ● 91 年 1 月：進行信用卡作業小組之輪調。 ● 91 年 9 月：建置安控訊息主控台，監控使用者登入使用狀況。 ● 91 年 10 月：改善作業人員之端末設備只能查詢資料，無法複製資料。取消光碟備份作業，改由 File System Server(檔案管理伺服器)主機管理。建置符合 CNS 17800/BS7799-2 標準規範之資訊安全管理系統，訂定資訊安全政策，建立有效管理制度，並藉由該系統循環運作及持續改善之模式特性以強化各項安控機制。 ● 91 年 11 月：建置作業環境實體安全控制系統。

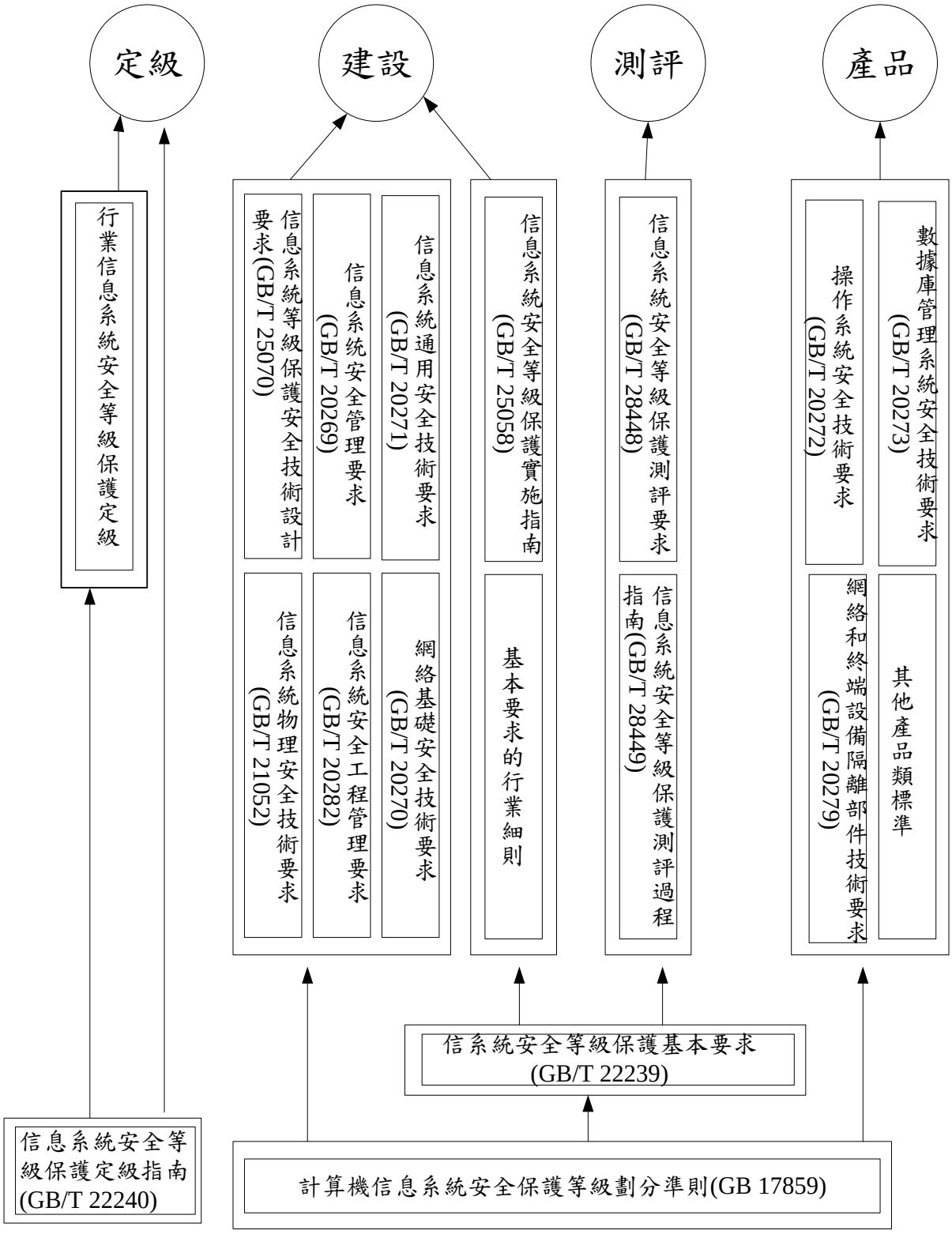
參、資訊系統安全防護比例原則初探

中國大陸自 2002 年起，師法美國如圖五所示之資訊保證技術框架(Information Assurance Technical Framework，簡稱 IATF)，逐步建立如圖六之 ISMS 標準化框架；至 2007 年，正式以法規要求 ISMS 的控制措施事項，表四是其技術面向之簡述，其中海巡署情報處及法務部資訊處被攻陷的資訊系統，於我國資訊與資訊系統分類分級之行動方案中，應均屬「高級」亦即表四中的第四級資訊系統，惟諸如「雙因子鑑別」、「強制性存取控制」等均未要求，尚未達到表四中第三級系統之技術及測評的準繩；表四中第三級系統採用之「程式可信執行保護」除「惡意代碼防範」外，尚包含回復等功能。有效之防禦是 ISMS 的核心工作，「人員依恃技術進行作業」，表五、表六與圖六是其例證，合規之 IATF 能提供事半功倍的防禦有效性，遵循表五、表六及圖七之資訊系統，媒體報導之攻擊方法的成功率應極低。



說明：
1. 公開金鑰基礎建設(Public Key Infrastructure，簡稱PKI)/
   金鑰管理基礎建設(Key Management Infrastructure，簡稱KMI)。
2. 偵測(Detect)與回應(Respond)。
3. 資料來源：IATF (Information Assurance Technical Framework)Release3.1,2002-09, Figure1- 8,Page1-14

圖五：深度防禦示意說明



圖六：中華人民共和國信息系統安全等級保護相關標準體系框架

表四：中國大陸資訊系統分級之伺服機技術要求舉隅

	第二級	第三級	第四級
鑑別	通行碼	雙因子(例：通行碼與符記)	同第三級，惟要求其中 1 因子不可偽造(例：生物識別)
安全標記	無	無	建議安裝
存取控制	自主性存取控制(DAC)	強制性存取控制(MAC)	同第三級
可信路徑	無	無	必須具備
稽核存底	稽核檔	必須具備稽核分析工具	必須安裝集中處理之稽核分析工具
殘留資訊保護	無	必須具備刪除工具	同第三級
完整性檢查	無	必須具備完整性檢查工具	同第三級
電磁防護	電磁干擾之防護	關鍵設施之屏蔽防護	關鍵區域之屏蔽防護(例：必須位於屏蔽室內)

說明：
1. DAC: Discretionary Access Control。
2. MAC: Mandatory Access Control。
3. 參考資料：GB/T 22239:2008，信息安全技術－信息系統安全等級保護基本要求。

表五：中華人民共和國資訊系統分級安全保護環境表列

使用範圍	安全功能	安全等級				
		一	二	三	四	五
安全計算環境	用戶身分鑒別	*	*	*	*	第五級資訊系統之安全保護環境個案規範
	自主訪問控制	*	*	*	*	
	標記和強制訪問控制			*	*	
	系統安全審計		*	*	*	
	用戶資料完整性保護	*	*	*	*	
	用戶資料保密性保護		*	*	*	
	客體安全重用		*	*	*	
	惡意代碼防範	*	*			
	程式可信執行保護			*	*	
安全區域邊界	區域邊界訪問控制			*	*	第五級資訊系統之安全保護環境個案規範
	區域邊界包過濾	*	*	*	*	
	區域邊界安全審計		*	*	*	
	區域邊界惡意代碼防範	*	*			
	區域邊界完整性保護		*	*	*	
安全	通信網路安全審計		*	*	*	

通信網路	通信網路資料傳輸完整性保護	*	*	*	*	
	通信網路資料傳輸保密性保護		*	*	*	
	通信網路可信接入保護			*	*	
安全管理中心	系統管理		*	*	*	第五級資訊系統之安全保護環境個案規範
	安全管理			*	*	
	審計管理		*	*	*	
資料來源：中華人民共和國國家質量監督檢驗檢疫總局 (2009) 信息安全技術－信息系統等級保護安全設計技術要求(GB/T 25070：2010 報批稿)，2009-03-12。						

表六：網路安全等級劃分之組織性與技術性控制

1. 資料來源：中華人民共和國財政部(2005)涉密辦公網白皮書 v1.1。
2. 2004 年 10 月 19 日之前，分成預算專網、內網與外網，3 網之間物理隔離，自成體系。
3. 2004 年 10 月 19 日之後，分成涉密辦公網(藍線)、預算專網(紅線)、內網(灰線)與外網(綠線)；藍線網之基礎建設均通過防電磁泄漏、數據傳輸安全及可追責性等之資訊安全性要求，已於 2005 年 2 月 9 日完成前期建設。
4. 使用者登入涉密辦公網時，須將用戶鎖(機密級)插入 USB 接口，一旦用戶鎖被拔出，網路將立即斷開。用戶鎖不可隨身攜帶、移交他人代管或帶離辦公區，不使用時，須將用戶鎖拔下並鎖存。
5. 涉密辦公網之終端設備均安裝與用戶鎖相對應之數據密碼卡(機密級)。

表七：RSA 密碼系統例

公鑰：

n 為兩質數 p、q 的乘積，(p、q 需保持隱密)

e 和 $(p-1)\times(q-1)$ 互質

密鑰：

$d = e^{-1} (\text{mod } (p-1)\times(q-1))$

加密：

$c = m^e (\text{mod } n)$

解密：

$m = c^d (\text{mod } n)$

以下是一個簡單的例子：假設 $p = 5$、$q = 11$，則

$n = p \times q = 55$

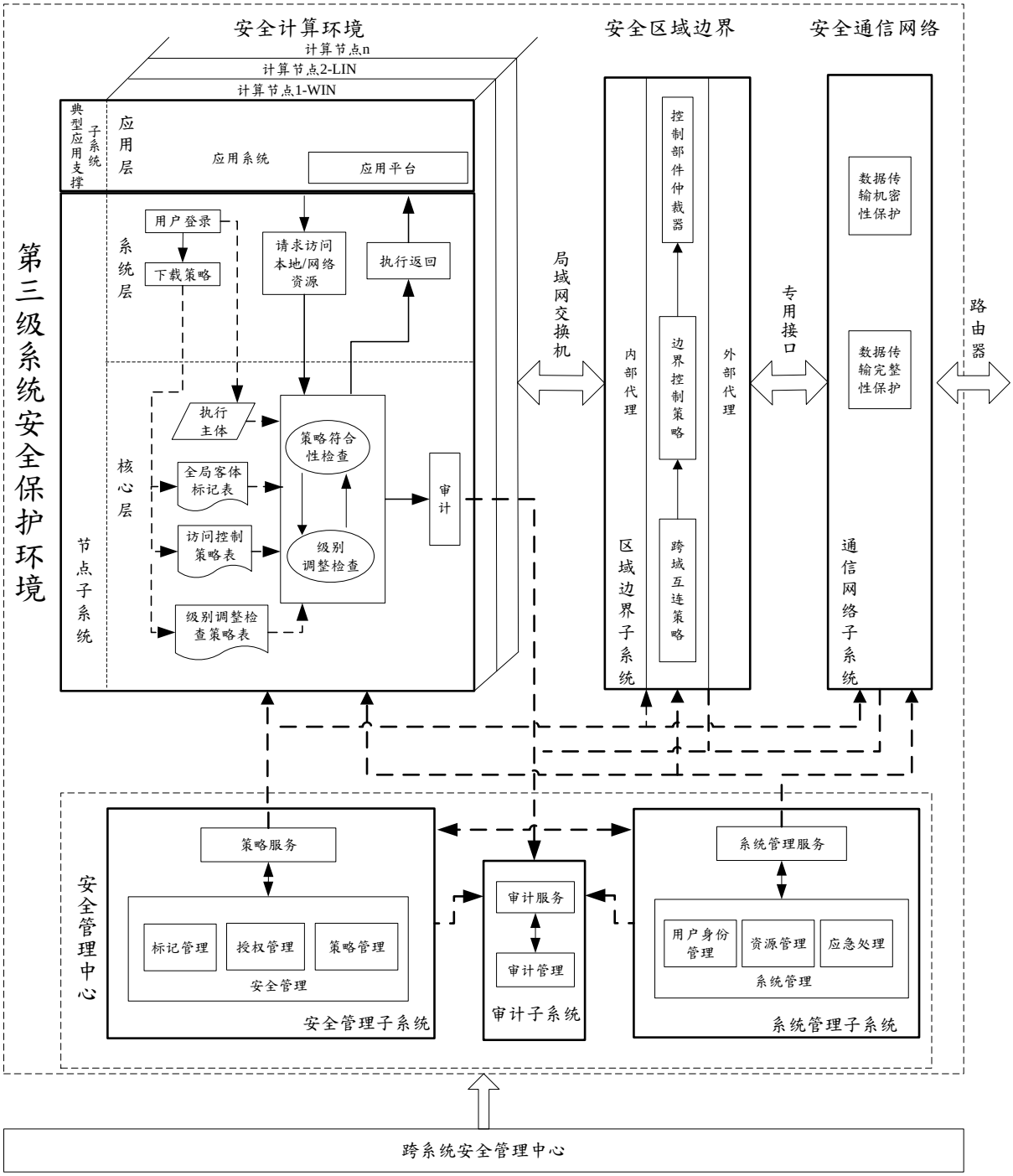
公鑰 e 必須和以下的數沒有相同的因數：

$(p-1)\times(q-1) = 4\times10 = 40$

隨機選擇 e 為 7，在這情況下：

$d = 7^{-1} (\text{mod } 40) = 23$

接著公開 e 和 n (公鑰)，d 保密(密鑰)，p、q 銷毀



資料來源：中華人民共和國國家質量監督檢驗檢疫總局 (2009) 信息安全技術－信息系統等級保護安全設計技術要求(GB/T 25070：2010 報批稿)，2009-03-12。

圖七：中華人民共和國第三級系統安全保護環境結構與流程示例

「居安思危，思則有備，有備無患」，以前述 PKI 之資訊安全要求事項為例；任何一個植基於密碼學之資訊安全機制，其在應用上安全與否，均仰賴金鑰管理之良莠，有如一個非常安全的銀行金庫，一旦鑰匙遭竊或是被人得知開鎖密碼，那麼這座金庫很難對抗那位拿著鑰匙或是已知開鎖密碼之入侵者一般；因此，如何防護使用者的密鑰被破解是 PKI 應要求之控制措施；PKI 之「公鑰資料庫」宜於要求事項中規範如表五中的「標(符)記(Token)與強制性訪問(存取)控制及屬性基(Attributed based)存取控制」以及表四中之「可信路徑」，其控制措施通常使用密碼學的技術實作之。

2012 年 7 月 21 日，周立平先生提出我國公開金鑰基礎建設(Public Key Infrastructure，簡稱 PKI)後門作業之脆弱性[18] ，至 2013 年 3 月 17 日，PKI 前門作業的明文與密文相同之脆弱性的控制措施仍待開展[1]；簡述如後[21] ，表七之例中，於表八可知 10, 11, 21, 34, 44, 45 之密文相同於明文且由取最大公因數之計算等即可解出密鑰，前述在 1979 年發表的 PKI 後門及前門之脆弱性的控制措施並不困難，本文作者之一應要求提供 2 大 PKI 業者參考的方法如後：

發文作業：於發文時，先計算文本加上時戳之文件彙紀值，再使用密鑰簽章後，檢查其數位簽章值是否與文件彙紀值相同？若相同，則更換時戳，直到數位簽章值及文件彙紀值不相同時方停止；然後，執行發文作業。

簽章確認作業：

1. 說明：
   1.1 m：文件彙紀值，
   1.2 c：數位簽章值，
   1.3 r：收文方，
   1.4 s：發文方，
   1.5 d_r：收文方密鑰，
   1.6 d_s：發文方密鑰，
   1.7 e_r：收文方公鑰，
   1.8 e_s：發文方公鑰，
   1.9 ttp：後門。
2. 收文方作業之 1：將$\{s, c^{d_r} \pmod{n}\}$送交後門。
3. 後門作業：
   3.1 計算$[c^{d_r} \pmod{n}]^{e_r} \pmod{n} = c_{ttp}$，
   3.2 計算$[c_{ttp}]^{e_s} \bmod (n) = m_{ttp}$，
   3.3 將$m_{ttp}{}^{e_r} \pmod{n}$送交收文方。
4. 收文方作業之 2：
   4.1 使用 d_r 解出 m_{ttp}

4.2 檢查 m_{ttp} 是否等於 m

5. 備考：發文方應使用過濾軟體確保 $m \neq c$，當 $m = c$ 發生時，更改文件之時戳直到 $m \neq c$ 時方發文。

表八：RSA 中明文與密文相同

1. $\gcd(10 \pm 0, 55) = 5,\ \gcd(10+1, 55) = 11$
2. $\gcd(11-1, 55) = 5,\ \gcd(11 \pm 0, 55) = 11$
3. $\gcd(21-1, 55) = 5,\ \gcd(21+1, 55) = 11$
4. $\gcd(34-1, 55) = 11,\ \gcd(34+1, 55) = 5$
5. $\gcd(44 \pm 0, 55) = 11,\ \gcd(44+1, 55) = 5$
6. $\gcd(45-1, 55) = 11,\ \gcd(45 \pm 0, 55) = 5$

5 和 11 即為 p 和 q，當 RSA 機制中密文與明文相同，而非 0、1 或 $n-1$ 的值被人知道的時候，我們可以解出 p、q 及 d 的機率大於或等於

$$1 - \frac{(\gcd(e-1, p-1) - 2) \times (\gcd(e-1, p-1) - 2)}{(1 + \gcd(e-1, p-1)) \times (1 + \gcd(e-1, p-1))}$$

e、d 均不為 2 且明文與密文相同的數目是 9 的時候，上述的機率就是 1[1,6,21]。

綜前所述，RSA 密碼系統之公鑰資料庫因其前/後檯的弱點，雖符合如表九所示之「無可控制的重新識別之擬匿名化」的要求，仍不宜公開，宜使用「授權研究者模型(The Qualified Investigator model)」之有限存取資料(limited access data)的方式，供研究使用。

表九：去識別化用語與現有技術之對照

國際標準 ISO/IEC 2nd WD 20889：2016-05-30 中用語	ISO/IEC 29191 (2012)	ISO/TS 25237 (2008)	ISO/IEC 29100 (2011)	ICO (2012)	Article 29 (2014)
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	Pseudonymisation	Anonymisation	Pseudonymisation
無可控制的重新	N/A	Pseudonymi	Anonymisat	Anonymisat	Pseudonymi

識別之擬匿名化 (Pseudonymization without controlled re-identification)		zation irreversible	ion	ion	sation
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。
3. 包含統計揭露控制(Statistical Disclosure Control，SDC)之「匿名化(anonymization)」優於(is stronger than that of)「去識別化(De-identification)」係歐盟的觀點，其法規使用前述之內蘊SDC的匿名化定義。

使用密碼學之擬匿名化技術，於公開金鑰幾均存在前述巧門的後門與前門之弱點[2,6]，我國「個人資料去識別化」的標準應修訂之[19]。

肆、結論

2012 年 1 月歐盟開啟整合「個人資料保護指令(Directive 95/46/EC)」、「電子通訊隱私指令(Directive 2002/58/EC)」與「電信網路改革指令(Directive 2009/136/EC)」三大個人資料以及隱私防護指令之法制，其以單一規則(Regulation)簡化機關/構與企業的法規遵循義務並促進單一數位市場；2016 年 4 月 14 日，歐洲議會通過，2016 年 4 月 27 日公布之「一般資料保護規則(General Data Protection Regulation，簡稱 GDPR)」，於其前言(26)除沿用表 9 的「匿名化」：「匿名資訊，(anonymous information,)意指釋出之資訊與已識別或可識別的自然人無關，或是對匿名化遞交之該個人資料而言，資料主體已不是或不再是可識別的，即資訊」；同時於條款(article)4 之第 5 項定義內蘊 SDC 的「擬匿名化(Pseudonymisation)」：「 意指個人資料經過處理後，在沒有提供其他額外資訊的情況下已無法將個人資料歸類於特定資料主體，且前述之額外資訊需與個人資料分離，並接受技術及組織的控制措施以及量測(measure)以確保該個人資料無法歸類於已識別或可識別之自然人。」使用密碼技術實作表 9 中，「可控制的重新識

別之擬匿名化」與「無可控制的重新識別之擬匿名化」，在大數據的環境中均存在理論上之弱點，GDPR 的「擬匿名化」定義方符合諸如「支付卡產業資料安全(Payment Card Industrial Data Security Standard，PCIDSS)」等事實標準之規範，ISO/IEC JTC 1/SC 27/WG 5 已立項進行 ISO/IEC 27551 之 GDPR 的「擬匿名化」相關之標準制定的工作項目中。GDPR 於條款 11 內蘊「去識別化(de-identification)」之規範，將個人資料區分為「已識別(identified)」、「能識別(identifiable)」、「條款 11 之已去識別(de-identified)」與「匿名化(anonymous)」4 個層級，於條款 15、16、17、18、20 及 21 規範前述 4 個層級的前 3 項資料分於存取控制、資料儲存、適當的資料安全等之要求事項，於條款 25 將 PbD(Privacy by Design(從設計著手控制隱私))納入規範並於其前言(78)闡明：「在與個人資料處理有關之自然人的權利及自由之保護，須採取適切的技術性以及組織性措施，方能確保此(GDPR)規責之要求能被遵循。」，於條款 32 闡明對不同層級之個人資料的包含擬匿名化、加密(encryption)等技術之資訊安全的處理、於條款 35 規範「資料防護衝擊評鑑(data protection impact assessment)」，於條款 37、38 及 39 規範「資料防護專員(Data Protection Officer，簡稱 DPO)」的權責。條款 83 明定對於一般性的違法，罰款上限是 10,000,000 歐元或上一年度全球營業收入之 2%的數額高者；對於嚴重之違法，罰款上限是 20,000,000 歐元或全球營業收入的 4%之數額高者。在 GDPR 的整備(readiness)工作項目，由 DPO 主責；歐盟於 2010 年即規範前述 GDPR 條款 37~39 之 DPO 的權責，預作準備。GDPR 於條款 40、41、42 與 43 規範行為準則以及驗證，相關單位幾均公布採用 ISO/IEC 27001 作為其包含個人資料去識別化之「個人資料管理系統」要求事項合規的驗証規範。

「借箸代籌」，我國宜先制定「個人資料保護細則」規範實作其組織與技術的控制措施及其量測之框架並要求一定規模的機關/構必須設置 DPO 之職務，主責資料去識別化、重新識別風險評鑑，PbD、資料防護衝擊評鑑等工作項目；先期，政府以身作則，於 2017~2020 年間分成「DPO 養成的參與以及訓練」及「從符規(compliance)到問責(accountability)的 PIMS 之實作」2 階段，要求表十中的「資安防護基準之安全等級」的高級與中級之機關/構完成 PIMS 建制，再推廣至民間。

表十: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)」刪除個人資訊。	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(6310 機關/構，資料來源：2016-11-06，自由時報，A2 版)	中(572 機關/構)	高(123 機關/構)
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。

「技不如人」還是「要求事項不如人」？「橘逾淮為枳」，前法務部資訊處陳泉錫處長於 2012 年 7 月 27 日上午 10 時 30 分至 12 時接受訪談[16]時表示：「個人認為 ISO/IEC 27001 認(驗)證，僅針對文件表單統一標準，但實際執行部分幫助有限。」，前述 PKI 之弱點未見有效要求其應有的控制措施並將其公鑰資料庫公開，成為不設防之脆弱性的資訊安全事件即為例證。當政府推動開放資料時，不規範汽車應配備安全帶、安全汽囊、安全(幼兒)座椅等技術要求，防止交通安全事件/事故之有效性自然比不上列入要求事項的國家，當圖三之深度防禦已臻成熟並擴增至包含個人資料防護與資料去識別化的廣度防禦之 ISMS 控制措施實作的此時，我國 ISMS 之政策，在 2017 年~2021 年的全國資訊與通信安全會報之工作計畫中，是具攸關性的議題，宜進行深入之分析及探討[1~15,19,20] ，並制定適當的行動方案。

[誌謝]

本文作者謹在此對審稿者增進本文水平之貢獻，與林樹國博士與黃健誠博士協助整理資料的辛勞，致衷心之謝忱。

參考文獻

[1] G.R. Blakly, and I. Borosh “RSA Public Key Cryptosystems do not always conceal messages”, *Computers and Mathematics with Applications*, Vol. 5, No. 3, pp. 169~178.
[2] http://eprint.iacr.org/2012/064.pdf/ (2017/02/27)。
[3] http://iac.dtic.mil/csiac/iapolicychart.html (2016/11/21)。
[4] http://www.fidis.net/ (2016/11/17)。
[5] http://www.nist/nstic/ (2016/10/31)。
[6] W.C. Kuo, C.S. Laih, M.J. Gau, and C.C. Chang “On the Number of Messages which cannot be Concealed in LUC”, *IEICE Trans. Fundamentals*, Vol. E80-A, No. 11, pp.2218~2224.
[7] NIST, “Request for Information (RFI): Framework for Reducing Cyber Risks to Critical Infrastructure”, 2013.
[8] NIST, “Security and Privacy Controls for Federal Information Systems and Organizations”, 2013, retrieved from http://dx.doi.org/10.6028/NIST.SP.800-53r4 (2015/07/17).
[9] NIST, “Assessing Security and Privacy Controls in Federal Information Systems and Organizations”, 2014, retrieved from http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53Ar4.pdf (2015/07/17).
[10] Office Journal of the European Union, “General Data Protection Regulation”, 2016-04-27.
[11] OMB, “Annual Report to Congress：Federal Information Security Modernization Act”, 2016-03-18.
[12] The White House, “National Strategy for Trusted Identities in Cyberspace: Enhancing Online Choice, Efficiency, Security, and Privacy”, 2011-04-15.
[13] The White House, “Critical Infrastructure Security and Resilience”, *Presidential Policy Directive/PDD-21*, 2013-02-12.
[14] The White House, “Improving Critical Infrastructure Cybersecurity”, *Executive Order/EO 13636*, 2013-02-12.
[15] U.S.C. § 2521, “Federal Information Security Modernization Act of 2014”, 2014, Retrieved from http://www.congress.gov/113/publ283/PLAW-113publ283.pdf (2015/03/15).
[16] 中華民國資訊軟體協會，“行政院「完備我國資訊安全管理法規之分析」委託研究計畫期中報告 (初稿) ”，2012-08-17。
[17] 李詩慧等，“中共網軍駭進法務部”，*壹週刊*，第 591 期，頁 46~50，2012-09-20。
[18] 周立平， “Cryptoanalysis in Real Life” (Presentation), 2012-07-21 P.M. 13:00~13:45, HITCON 2012。備考：周教授於 2012-07-21 簡報中提出之脆弱性－「若取得一定數量的公鑰資料，則 $n_1 = p_1 \cdot q \wedge n_2 = p_2 q \rightarrow q = \gcd(n_1, n_2)$ 」，謹此敘明。
[19] 法務部，“法律決字第 10603500720 號(書函)”，2017-02-13。
[20] 許瀞文，“花錢就能拿證書 台灣資安玩假的？”，*今周刊*，第 849 期，頁 54~56，2013-04-01/07。

[21] 樊國楨，“電子商務高階安全防護－公開金鑰密碼資訊系統安全原理”，*財團法人資訊工業策進會資訊與電腦出版社*，頁 145~151、頁 154~156、頁 296~301，1997。
[22] 樊國楨，“探討資訊安全管理之信函”(個人淺見，僅供參考)，(2011 年第 2 季)資訊安全管理系統標準化系列討論會(會議資料)，頁 79~97，2011-03-09。
[23] 戴志楊等，“海巡署 3,000 機密外洩”，*時報周刊*，第 1794 期，頁 34~37，2012-07-06/12。