

大數據產業的崛起

近年來，大數據（big data，或譯為海量資料、巨量資料）為學術界與實務界熱烈討論的新興領域。產業界與學術界將大數據分析（big data analytics）視為第五波科技革命後的明星產業，未來人類社會將出現以資料為驅動力的經濟（data-driven economy）。根據國際數據公司¹（International Data Corporation, IDC）的調查報告指出，大數據科技與服務市場是一個快速成長、產值高達數百億美元的全球新興產業，每年成長率約為 27%，至 2017 年預估產值將高達 324 億美元，約等於過去資通訊科技（ICTs）總產值的六倍之譜。McAfee 與 Brynjolfsson (2012) 指出，大數據正在掀起一場管理革命，企業首先必須設法改變傳統的決策制訂文化，從「我們想些什麼」（What do we think?）轉變為「我們知道些什麼」（What do we know?），業主必須充分運用大量的客戶資訊，客觀地瞭解大量數據所呈現的意義與價值，然後做出妥適的企業決策，快速地提升企業的營運效益。

聯合國潘基文秘書長推動一項名為聯合國全球脈動計畫（United Nations Global Pulse, 2014），該計畫最重要的任務是發起大數據應對氣候挑戰（The Big Data Climate Challenge），期盼號召全球一流的科學家與技術專家能夠充分運用大數據資料，為全球氣候變遷的因應措施，研議出以資料為

驅動力的證據（data-driven evidence），以建立大眾對於使用大數據資料能夠改善地球環境產生十足的信心。Robertson (2013) 指出，以健康產業為例，美國每年大約創造 30 億元產值，可見大數據已成為「大商機」。事實上，大數據存在於各個領域，包括醫療健康、食品安全、氣候、能源與環保、智慧交通等，在不違反資料隱私保護的原則下，如何妥善利用這些巨量資料，挖掘有價值的知識與決策資訊，成為當前政府必須面對的一項「大挑戰」。

雖然大數據分析成為一門跨學科的顯學，但其運用涉及非常敏感的資料隱私問題。早在 2006 年，當臉書剛開始問世時，哈佛大學研究團隊就利用 1700 位臉書用戶，探討大學生在興趣與友誼之間的長期發展關係，這個研究係假設該團隊使用臉書社群之資料是匿名的，可以讓研究者自由地進行探究（Lewis, Kaufman, Gonzalez, Wimmer, & Christakis, 2008）。不過，這個論點被學者反駁，即便是依據隱私的要求進行個人資料的匿名化，以現代科技水準而言，亦可順利透過多種資料庫的組合與連結，輕易地去匿名化（de-anonymize），多數的個資對象並未知悉其資料已被使用過。2012 年，臉書聘用一群學者在未明確告知 689000 位使用者的情況下，對他們進行有關情緒相互感染的研究，引起全球用戶對於臉書侵犯隱私的疑慮。儘管用戶在加入臉書之前，都必須「同意」臉書的要求，

1 國際數據公司是全球著名的資訊技術、電信行業與消費科技諮詢、顧問與活動服務專業提供商，在全球擁有超過 1000 名分析師，對 110 多個國家的技術發展趨勢與業務行銷機會進行深入分析（International Data Corporation, 2014）。

但同意並不代表未來可以使用，此意味著現行徵求客戶的同意方式已經流於形式，毫無意義。

本研究旨在探討，為因應大數據時代的來臨，台灣產業界在適用資料隱私保護法制上遭遇何種問題？應如何整合公私部門的力量，以公私夥伴(public-private partnership)途徑，在保障資料隱私的前提下，推展台灣的大數據分析產業？本研究首先檢視大數據的崛起及其概念，其次探討歐美各國對於大數據隱私管制的法令規定，最後再針對現行台灣資料隱私所可能產生的問題進行分析，並提出可行的政策建議，以供決策當局參考。

大數據的意義

大數據主要是 Mayer-Schünberger 與 Cukier 於 2013 年出版的《大數據》而出名，該書是亞馬遜網路書店的排行榜暢銷書之一，執教於英國牛津大學，專研網路治理與管制(internet governance and regulation)的 Mayer-Schünberger 教授及《經濟學人》(The Economist)雜誌的資深編輯 Cukier 指出，今天人類社會從科學到醫療保健、從銀行到網路，涵蓋的行業各式各樣，講的是同一件事——世界上的資料正在迅速增長，不僅超過電腦機器能處理的數量，甚至遠超過我們的想像。Google 每天就得處理超過 24PB 的資料（1 petabyte 等於 1000 terabyte），是美國國會圖書館所有紙本資料量的數千倍。Youtube 服務用戶達 8 億，每秒上傳影片長度超過 1 小時。臉書每小時上傳 1000 萬張新照片，按讚或留言至少超過 30 億次，資料量之大難以想像(Mayer-

Schünberger & Cukier, 2013, p.16)。

前科技部長張善政引用 2013 年 KPCB 投資基金合夥人 M. Meeker 估計，2012 年全球網際網路用戶達 24 億，年成長率為 8%，這些網路透過各種不同的網路應用工具，資料產生量不僅非常龐大，且資料類型相當多元，不是傳統格式化資料所能儲存與分析的，因而開啓了資料蒐集、儲存、處理、分析的新頁，乃促成大數據分析產業的興起(Thomas, McAfee, & Brynjolfsson, 2014, p. 5-7)。因而，2013 年被數位科技媒體記者譽為大數據元年(Savvas, 2013)。

根據美國白宮總統辦公室(Executive Office of the President, the White House)公布的大數據報告(Executive Office of the President, 2014, pp. 2-3)指出，大數據的概念需視你是何種專家而定，多數定義必然呈現出，大數據是由於當前雲端科技的發達，足以快速地擷取、累積、利用與處理個人資料，以致於呈現出數量大、速度快與種類雜的三種特徵之資料。換言之，當前資料獲取速度非常之快，且具有極大的涵蓋面，甚至包含過去所無法觀察或測量的資料型態。由許多科技工具、偵測感應器、網際網路轉換、電子郵件、影片、按讚或其他類似工具所產生的資料類型是大量的、多元的、繁雜的、長期的或分散化的，必須發展出整合性的分析工具挖掘其價值。

Boyd 與 Crawford (2012, p. 663)指出，大數據時代已經開始了，電腦專家、物理學家、經濟學家、數學家、政治與社會學家、生物科技專家等投入大量有關人們、