



圖片來源：unsplash

特稿

行政大數據運用的可能與精進

廖興中¹

國立政治大學公共行政系副教授

〈摘要〉

近來公共行政的實務工作者與學者們紛紛開始擁抱大數據所帶來的機會，同時也必須面對大數據所帶來的諸多挑戰。這些大數據資料集漸漸地被公共管理者用來洞察民眾的意向與掌握社會趨勢。甚至，許多公共行政的研究者開始嘗試使用這些大數據資料集來產生過往無法操作化的概念或指標，並透過理論驗證來產生新的見解。臺灣不論是中央與地方政府也都開始有許多的運用產生。然而，未來唯有針對行政資料管理能力的建立、行政資料產生流程的認識、行政資料品質確保的思考、行政資料開放整合的突破等面向的提升，行政大數據的運用才能更發揮其完全的能量。

關鍵字：行政資料、大數據、公共事務、資料治理

1 電子郵件信箱：liaogis@nccu.edu.tw

壹、大數據在公共事務領域的崛起

公共行政中實務的公務同仁與研究者，在過去時常感嘆缺乏可以用來分析與操作化的資料。在近來二十年間，網際網路的蓬勃已經開始漸漸解決了這樣的問題。現在，公共行政的實務工作者與學者們都開始爭先恐後地擁抱大數據（Big Data）所帶來的機會，同時也必須面對大數據所帶來的諸多挑戰。這些大數據資料集漸漸地被公共管理者用來洞察民眾的意向與掌握社會趨勢的變化。甚至，許多公共行政的研究者則是開始嘗試使用這些大數據資料集來產生過往無法操作化的概念或指標，並透過理論的驗證來產生新的見解（Chen, Roger, Chiang, and Storey, 2012）。本文從一個概述的角度，藉由三個部分的討論來進行分享。首先回顧公共事務與社會政策、政治學、管理學、甚至政策資訊學等研究領域，針對大數據所提出的相關定義。第二部分則是透過臺灣官方與民間的應用個案，觀察臺灣目前大數據運用的情形。最後，則是提出一些未來推動行政大數據運用的挑戰與機會。

貳、不同學科領域對大數據的關注

大數據的定義其實相當多元，並沒有放諸四海皆準的定義。首先從管理學領域的學者來看，正如 George、Haas 與 Pentland（2014）在 *Academy of Management Journal* 所強調的，在主流媒體與實務界對於大數據有一定程度的錯誤理解。這些誤解主要強調資料集的巨大，而較不關注資料的內容或運用（McAfee et al., 2012）。從前述管理學家的觀點來看，大數據事實上主要強調，這些可被收集的資料藉由不同來源資料集的整合，產生各種事件、個人、過程或現象的有意義資訊。

在政治學家的眼中，大數據則被定義為：「藉由創新技術（例如機器學習）使得研究者所收集到的新形態資料（例如社群媒體資料），或者產生成本相對更低的傳統資料」（Clark and Golder, 2015: 65）。同時這兩位學者更強調，當我們越來越有能力產生、收集、存儲與分析大量數據時，將改變眾人對政治世界的理解。這些學者將大數據視為一種對學科領域的挑戰，會大大增加資料收集與分析能力的複雜性。然而，政治學者們也確實認識到更多的資料總比缺乏資料來得好，也因此這些各式各樣來源的大規模資料集，正在挑戰政治科學家的思考方式、研究方法、資料收集方式，甚至是相關理論的發展。大數據更大大地挑戰學者們需要系統性地思考哪些資料是重要的？以及哪些是沒有價值的？

從公共政策學者的觀點來看，像 Pirog（2014）在 *Journal of Policy Analysis and*

Management 的評論便提到這些嶄新與更高品質資料集的可取用性，例如美國聯邦政府 Data.gov 上的開放資料，將大大的轉變公共政策的實驗研究。該評論斷言大數據提供了政府對於每一位民眾更完整的圖像，而更有效地了解民眾將改善公共政策和公共管理的能力。該評論甚至認為大數據的挑戰主要在於資料本身非結構化的特性，因此未來需要透過社會科學家可能不理解的方法進行整理與分析。

在資訊管理與相關技術領域的觀點，大數據的討論與運用強調在將不同來源的資料串連起來的技術。而 Janssen 與 Van den Hoven (2015) 更進一步強調大數據主要是透過大量人、事、物之間互動所產生的資料，分析出大量的資訊。對於資訊相關領域的學者來說，一面需要探索出更多收集、儲存資料的方法，一面更需要不斷地設計與測試分析資料模式的演算法。但是這些學者也提到大數據的挑戰主要在於大量資料分析固然有助於我們挖掘社會行為與現象，然而在大數據創造公共價值之前，可能需要解決隱私權的倫理問題 (Boyd and Crawford, 2012)。同時對於大數據分析後特殊的發現，其背後可能的解釋，或者可以採取的相關行動，甚至如何真正能洞察這些發現背後的故事，也會是主要的挑戰 (Frankel and Reid, 2008)。

綜合各領域學者所關注的特點，吾人可以大數據之所以受到關注，並非僅僅因為其資料量龐大的原因，更重要的是這些大數據的出現，很有可能透過有意義的資料整合，透過合適方法的分析，探索出有意義的結果並挖掘出其背後的故事。

參、大數據的可能定義

大數據實際上並沒有一個明確的定義。像 de Goes (2013) 便曾經提示過大數據這個詞過於模糊且範圍過於廣泛，以至於很難明確定義出其意義。本文試著總結過去一些學者們針對大數據所提出定義，試圖更清楚地將大數據的意義初步建構。

Taylor 等學者 (2014) 曾針對許多知名的經濟學家進行訪談，試圖更深入地了解大數據的意義以及相關的運用。這些受訪者主要將資料集的大小與複雜性，視為是大數據關鍵的特徵。這些受訪專家認為，這些可以被利用觀察值與變項，因著資料來源大量開始來自於網路與社群媒體，數量與過去有著截然不同的改變。換言之，資料集的大小與涵蓋的範圍是定義大數據的關鍵特點。就如同 Einav 與 Levin (2013) 所強調的，現在資料可取得的速度更快，涵蓋的範圍也遠遠大過於之前社會研究可用的資料資源了。

許多討論大數據的文獻都集中在網上活動與社群媒體使用所產生的資料 (Tinati et al., 2014)。這類資料的產生可能是透過網上搜索、網路瀏覽歷史記錄、部落格，或是社群媒體貼文、視頻與圖片的分享等。網際網路的發展與數位社群網絡帶來了前所未有的大量資料可以被收集。網際網路的使用與社群媒體資料，創造出大量各式各樣主題的研究，例如個人的情緒 (Dodds et al., 2011)、政治人物的形象管理 (Jackson and Lilleker, 2011) 與集體政治行動 (Segeberg and Bennett, 2011)。

然而大數據不應僅僅侷限在那些透過網際網路收集的資料。這是因為大數據也可以來自商業交易，例如超市或銀行內的購買或交易資料 (Felgate and Fearn, 2015)。大數據也可以來自感測器，例如手機 GPS 追蹤資料 (Eagle et al., 2009)。行政資料，例如教育記錄、醫療記錄，以及稅務記錄也可以被視為大數據的來源 (Chetty et al., 2011)。

一般最常被提到的大數據定義是由 Laney (2001) 所提出的，其將大數據的主要特徵分成資料量 (volume)、可用資料的格式與來源，例如文字、圖片、影片、金融或社群交易或互動等 (variety)、以及資料產生的速度 (velocity)。就如同 Tinati 等學者 (2014) 所強調大數據包羅萬有與即時性是等關鍵特徵一樣。

Schroeder 與 Cowls (2014) 強調，大數據的概念與可以提供研究者使用資料來源的大改變有密切相關。在大數據時代，研究人員花費更多的時間在挖掘有價值或有意義的潛在資料來源。因為社會科學中的傳統資料來源，主要是由研究人員根據其研究問題，再選擇合適的調查工具所產生。相比之下，大數據則是收集各式各樣來源的資料，而且這些資料的產生與收集多半與研究的目無關，甚至研究人員對這些資料或其內容的設計沒有任何介入。一個重要的特徵也可以被納入大數據的定義中，是這些資料不是為了研究目的而收集的，而是可以被社會科學研究者適當地重新利用。

前面其實針對大數據的特性，有許多學者提出相當多的見解。有許多屬性都被提出來表徵大數據，但大數據是否需要滿足這諸多的特徵才能稱為大數據？Kitchin (2014) 便曾經強調有一些資料可能具有許多被認為符合大數據定義的特徵 (例如大小、來源與速度)，但是其他類型的大數據可能具有幾個不同的或者只有一個特徵，仍然可以被認為是大數據的一種類型。不是所有的大數據的數量都一樣大，也不是所有的大數據都具有快速與即時可用性。也因此，行政資料也可以是某種意義

上的大數據。

肆、臺灣行政大數據運用情形

行政資料被定義為源自行政系統運作所產生的資料，通常由政府各部門或單位在執行相關業務時產生（Elias, 2014）。同樣，Woollard（2014）亦認為行政資料主要是因為政府為了處理註冊、交易與記錄保存等的不同目的的業務而取得的，而且這些行政資料往往與服務的遞送很有關係。這些資料可以來自廣泛的行政系統，例如教育、醫療保健、稅收、房屋或車輛許可等。甚至有一些行政資料是登記而來的，例如出生、死亡、婚姻、選舉登記、及全國人口普查。儘管行政資料在過去大數據的討論核心中並未太受關注，然而從性質上來看，這些行政資料完全符合大數據的定義。

而在考慮在政府部門應用大數據的可能性時，可以先從以下不同的行政功能來觀察：

1. 公共監督—察覺違規行為（例如法律違規）並採取應對措施。
2. 公共管理—透過計畫、許可、禁止、命令等工具與方式，來確保社會行為的合宜、社會關係的良好、社會環境的正常。
3. 公共服務提供—提供社會大眾某些服務或產品（包括基礎設施，例如道路）。

在上述每一類功能當中，大數據的運用會有所不同。在公共監督的業務中，主要會透過利用大數據發現違規行為。在管制的情況中，大數據則可用於觀察管制區域的狀態。因大數據所產生的管理知識則可以使管制政策可以被改進得更好。大數據還可以幫助決策者更了解其所制定出決策的社會影響，並從社會回應中得出更好的結論。對於公共服務提供來說，大數據可用於改善公共服務與公共服務的創新。另一種運用的可能，則是可以用在驗證民眾取得特定社會服務的資格，如此也可以預防或避免社會服務被濫用及誤用。

一、更好的公共監督：察覺違規行為

大數據運用可以為政府部門帶來價值的第一個領域就是察覺違規行為，這是公共監督功能的一部分。政府機關監督特定區域並檢測違規行為，如果發現違規行為，便可以採取相對應的管理措施。這部分的運用主要是利用資訊系統或人工智慧根據

違規行為確認模型，根據大數據偵測到可能的違規行為，並將這樣的個案轉知相關單位。其次，進一步透過大數據分析，則可以提供政府機關違規行為發生的可能區域，並進一步研擬可能對應的監管計畫。不過，監管的行政行動較可能是透過傳統實體的行政方式來執行，例如現場確認或檢查；主要是在減少全自動化下的可能風險，例如：機器誤判的問題（Citron, 2008）。最後，大數據一面可以協助政府快速的偵測到這些違規行為，更在違規行為預測方面提供了更巨大的益處，亦即違規行為的提前預測（Pretty, 2013）。

自從 ETC 的機制開始運行之後，交通部陸續都有運用大數據針對交通管理的相關業務進行分析，包含進入服務區的車輛來自哪些縣市、服務區的離尖峰時間、甚至於事故分析。另外還針對超速項目進行分析，從容易超速的路段、時段等，到 2018 年 4 月更進一步對嚴重超速做資料收集²。甚至透過 ETC 的車牌辨識，可以收集一些違規車輛的動態資料，例如未懸掛車牌的車輛、或是懸掛非制式車牌的車輛，其在高速公路上行駛的資料都也被有意義的分析，例如根據交通部針對 2015 年 3 月至 6 月 ETC 資料的分析，便發現未懸掛車牌的車輛，無明顯常出現的時段或路段，大多發生於通行量較高的路段；但是懸掛非制式車牌的車輛，則較常出現於南部路段³。而藉由 ETC 資料，分析特定車輛行駛路線或常出現路段，便可以移送給警察機關建立攔查策略規劃時參考。

另外像臺北市與新北市政府，由於全國通報的家庭暴力案件約 100 萬件，而雙北市家暴案件就占 27%。因此，臺北市自 2016 年結合民間資料英雄計畫，招集多位統計、數位及科技專長的資料英雄，利用科技整理過去大量的家暴案件資料，初步繪製成易懂的「家暴案件地圖」，研發出預測家暴再發的風險模型，幫助社工人員在社區防治工作及案件處理上能夠更加精確，讓社區的家暴相關危機與風險都能提早被預防⁴。這些個案都是中央與地方政府，為了監測或預測違規或不當行為，透過大數據分析的導入，所衍生出來的運用。

二、更好的管理：現況的掌握與回應

大數據的收集與分析協助政府機關對於許多不同區域的人口、社會、經濟、環境、其他許多的狀況得以掌握與了解，進而可以提出相關的方法來進行管理與回應。

2 請參考洪巧藍，2018，「超速習慣全都錄！ETC 大數據科技執法盯上 10 玩命駕駛」，ETtoday 新聞雲，<https://www.ettoday.net/news/20180624/1198005.htm#ixzz7kL6tP9x7>。2022/11/28。

3 請參考交通部運輸研究所網站，<http://www.iotbd.creatidea.org/website/index.php>。2021/12/01。

4 請參考社會安全網網站，<https://topics.mohw.gov.tw/SS/cp-4534-48106-204.html>。2021/12/01。

這種模式主要強調大量不同來源資料的收集與即時分析，並且呈現資料分析的結果，根據結果行政機關可以提出適當的介入手段或決策。

以國家災害防救中心的個案為例，該中心在原有工作的成果中已經持續累積相當多元的資料，包含監測、預警與衝擊的相關資料。加上更多社會經濟與民生公共物聯網相關資量的建置，更增加許多過去所沒有的資料。透過有效利用這些資料進行災害預警技術的再研發，勢必可以繼續提升防災的能力。甚至由於現今資訊圖像化技術的進步，更可以將這些多元資料轉化成防災的資訊，透過圖像化的方式成為提醒民眾及預警的重要工具⁵。

另外像是環保署所建立的環境物聯網數據分析平臺，主要目的是希望透過人工智慧技術來找出污染排放潛勢區位。傳統的空污管理，主要是透過宣導或是制定排放量的方式進行，但因著數位技術的興起，該行政機關開創人工智慧加上空氣污染大數據的技術，一面能夠更精準鎖定排放來源，一面則能建立空污排放的大數據分析模型，透過全自動的方式針對空氣污染進行預測、推估、告警。這樣大數據分析的導入，使得相關空污稽查人員從過去被動式進行管理、單方面的接受社會檢舉報告的模式；轉換到主動規劃出擊，精準查處可能污染的智慧稽查模式⁶。

三、更好的公共服務提供：更好的服務

對於提供公共服務的政府機關來說，大數據所帶來的明顯益處在於產生了更好的服務。這主要是因為能夠輸入大量有關民眾在需求、或是行為資料的結果。如果一個國家是「智慧的」，那麼許多的公共服務不論在設置、提供、甚至是方式，應該都可以緊密地與民眾端的需求配合。這類的大數據運用，很適合許多公共設施的規畫與服務選址參考。

就像 2019 年由內政部與衛生福利部合作形成的「銀髮安居計畫」，便是以行動健康、照護人力、經濟狀況、環境安全、環境便利與住宅狀況等六個面向的指標，透過加權計算得出「銀髮安居需求指數」，界定出有服務需求的長者，並思考如何導入政策資源給予協助。這個計畫更曾參加「總統盃黑客松」競賽獲得卓越團隊獎。當時該計畫篩選出全臺 20 縣市共計 1,170 名高風險的個案，並交由地方政府進行家訪，並發現 66.7% 有潛在長照需求，同時再將這些結果回饋給內政部，再精進評

5 請參考行政法人國家災害防救科技中心網站，<https://www.ncdr.nat.gov.tw/Page?itemid=25andmid=22>。2022/11/28。

6 行政院環境保護署環境物聯網，<https://iot.epa.gov.tw/iot/guide/analysis>。2022/11/28。

估指標⁷。許多城市都面對到人口老化與住宅老舊的問題，透過跨部會的資料整合、指數加權與大數據分析，快速地界定潛在需求者。更有助於長照計畫與老舊社區更新或增設電梯計畫的有效執行。

另外，臺北市交通局與叫車業者合作取得一週的資料，包括電招、App 叫車、路招等，藉以了解民眾上車及下車需求，透過地圖套疊圖資，掌握民眾在臺北市各個路段上車、下車的空間分布情形。特別是在甚麼路段、時間、以及使用甚麼樣的方式上車等資料，藉以觀察民眾上下車的熱點或熱區。這樣的運用發現計程車電招的資料分布較為隨機零散，而路招則會以大馬路為主。另外，大安區與信義區都是計程車招呼的熱區，但從時間軸來看，信義區夜生活較多，所以白天與晚上的搭車的密度明顯不同。羅列出臺北市計程車招呼的百大熱點路段後，交通局則會派員現場勘察，依當地路況、路寬是否允許，決定是否設置新的計程車招呼站，來滿足的可能的需求⁸。

肆、未來的挑戰

從前面的論述看來，政府部門擁有許多重要的行政資料，而這些行政資料往往也某種程度吻合大數據的特性。甚至在臺灣近來已經有許多不同的行政機關開始大量的運用這些行政資料進行相關大數據的分析運用。然而，未來仍然有許多面向是政府部門可以努力精進突破的部分。

一、行政資料管理能力的建立

本文注意到，大數據型態的行政資料，與過去傳統類型的社會科學資料相比，可能更不具系統性，並且需要行政或研究同仁針對資料進行清理與整理來促進這些資料被分析的可能性 (Einav and Levin, 2013)。資料的清理與整理，主要是準備並整理資料來滿足這些大數據分析的任務目標。這些其中的工作包括重組資料、重新編碼、構建變項 (Mitchell, 2010)。例如：有許多地方政府開放 1999 民眾陳情派工的資料，然而這其中有許多資料的座標值可能有缺漏甚至誤植，這些都必須進行資料欄位的檢查與清理。同時，若需要結合人口、社會經濟指標進一步分析，那麼便需要先將所有點位歸納到每個村里或行政區，如此才能跟村里或行政區的相關資料進行接合並分析。未來行政同仁在資料管理的能力，確實有建立的必要性。一旦

7 銀髮安居計畫—大數據助攻找出 1.8 萬長照需求長者，<https://www.ankecare.com/article/1435-25836>。2022/11/28。

8 蘇文彬，2021，「臺北市大數據應用實例大公開」，iThome，<https://www.ithome.com.tw/news/143186>。2022/11/28。

行政機關同仁對於資料管理的能力漸漸具備，將會有更多行政資料可以被整理成為有助於大數據分析的材料。

二、行政資料產生流程的認識

如同本文前面所提到的，行政資料的產生主要是由於各個行政機關包羅萬象的行政業務執行所產生。這些資料的收集與產生並沒有與研究的目的相契合。因此研究者必須透過對於這些行政資料產生與蒐集的充分了解，來掌握與了解行政資料。如此研究者才能對於這些資料的正確性與有效性，有一定程度的認識。在運用行政資料進行大數據分析時，才知道可能的研究限制。以至今許多社會科學研究常運用的戶籍人口統計資料為例，研究者就必須認識到戶籍人口資料建立的流程，也才會了解這中間可能出現「籍在人不在」的可能限制。例如許多民眾因著各樣理由會登記在某一個戶籍中，但卻在其他地區生活、工作，然而所有的戶籍資料就是透過登記的方式來建立，並沒有事後的追蹤與校正，這樣的落差便會對人口的評估產生一定的落差。這也就是為何目前內政部開始透過與手機公司資料的合作，進行重新的加權推估，希望藉此來矯正前面所提到的問題。簡單地說，這許多資料正確性與有效性問題的察覺，都有賴於行政同仁與資料的使用者對於這些行政資料產生的流程有充分的認識，方能察覺，進而思考如何結合其他資料與方法來進行矯正。

三、行政資料品質確保的思考

在進行行政資料大數據分析前，研究人員應該質疑行政資料資源的品質。首先，研究人員需要考慮這些資料被收集的過程中，行政人員是否因為一些機制建立與否，而會在建立行政資料的流程中產生敷衍了事的可能性（Goerge and Lee, 2001）。因此，在行政業務相關資料建立與收集的過程中，一面除了考慮透過數位工具自動記錄，或是自動除錯以避免人工輸入的問題之外，更重要的是與另一個來源的行政資料進行比對。兩個不同系統中或許會同時存在臺灣醫療院所的資料，如此就可以藉由這兩個不同資料來源的資料進行比對，也就可以找出具有差異的資料，進而確認出正確的資料。再者，研究人員應該可以了解是否有一個審核資料的機制（例如資料是否在任何時候都會進行交叉檢查）。最後要留意某些行政資料會與行政績效評估有關，這類的資料往往會因為行政機關為了滿足績效評估的要求，可能會在行政作為或是行政資料收集方面產生影響資料品質的可能行為。

四、行政資料開放整合的突破

首先許多的行政資料並不直接屬於研究社群，取得資料的過程往往有許多的限制，並且要花費相當的時間來索取。即便現今有許多開放資料不斷地被釋放出來，但是仍然有許多資料無法被開放。其次，許多不同來源資料的串聯，往往會遇到串聯代碼欄位沒有標準化的問題，導致無法串接不同資料來源的各個欄位。例如各鄉鎮市區代碼在內政部的資料庫裡有其編號，但在其他部會未必採用與內政部相同的代碼，甚至可能使用郵遞區號或其他代碼。這樣會造成資料界接上面的困難或是不便。因此許多行政資料中的個體或單位應該再未來思考如何建立跨機關的單一代碼，例如個人的身分證字號，便是一個不同行政機關個人資料中最好串聯資料的標準化欄位。最後，跨域資料的整合還是有賴各機關開放的文化，以及對於大數據多樣資料整合所帶來巨大效益的認同與想像，如此才能避免行政機關各自握著大量的資料，卻是缺乏整合資料的可能性。

政府是一個透過行政業務執行過程收集到許多大量資料的組織。這些行政大數據的運用，對於政府在公共監督、公共管理、與公共服務提供等各方面，預期都會產生許多正面的影響。而且臺灣政府在這些面向中，有許許多的應用已經不斷地被產生。然而，未來的精進方向，還是有賴政府機關與資料使用者不斷的再強化。唯有針對行政資料管理能力的建立、行政資料產生流程的認識、行政資料品質確保的思考、行政資料開放整合的突破等面向的提升，行政大數據的運用才能更發會其完全的能量。

參考文獻

- Boyd, D., and Crawford, K. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon," *Information, Communication and Society*, vol. 15, no. 5, pp. 662~679.
- Chen, Hsinchun, Roger H. L. Chiang, and Veda C. Storey. 2012. "Business Intelligence and Analytics: From Big Data to Big Impact." *MIS Quarterly*, vol. 36, no. 4: pp. 1165~88.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *The Quarterly Journal of Economics*, vol.126, no. 4, pp. 1593~1660.
- Citron, D. K. 2008. "Technological Due Process," *Washington University Law Review*, no. 85, pp. 1249~1313.

- Clark, W. R., and Golder, M. 2015. “Big Data, Causal Inference, and Formal Theory: Contradictory Trends in Political Science?: Introduction,” *PS: Political Science and Politics*, vol. 48, no. 1, pp. 65~70.
- de Goes, J. 2013. “Big Data is Dead. What’s Next.” **Venturebeat.com** guest blog post: <https://venturebeat.com/business/big-data-is-dead-whats-next/> (access: November 28, 2022)
- Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., and Danforth, C. M. 2011. “Temporal Patterns of Happiness and Information in A Global Social Network: Hedonometrics and Twitter,” *PloS One*, vol. 6, no.12, e26752.
- Eagle, N., Pentland, A., and Lazer, D. 2009. “Inferring Friendship Network Structure by Using Mobile Phone Data,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 36, pp. 15274~15278.
- Einav, L., and Levin, J. D. 2013. “The Data Revolution and Economic Analysis,” *National Bureau of Economic Research Working Paper* (no. w19035).
- Elias, P., Lane, J., Stodden, V., Bender, S., and Nissenbaum, H. 2014. “A European Perspective on Research and Big Data Analysis,” in *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (Cambridge: Cambridge University Press), pp. 173~191.
- Felgate, M., and Fearne, A. 2015. “Analyzing the Impact of Supermarket Promotions: A Case Study Using Tesco Clubcard Data in the UK,” in *The Sustainable Global Marketplace* (Switzerland: Springer Cham), pp. 471~475.
- Frankel, F., and Reid, R. 2008. “Big Data: Distilling Meaning from Data,” *Nature*, vol. 455, pp. 30.
- George, G., Haas, M. R., and Pentland, A. 2014. “Big Data and Management.” *Academy of Management Journal*, vol. 57, no.2, pp. 321~326.
- Goerge, R. M., and Lee, B. J. 2002. “Matching and Cleaning Administrative Data,” *Studies of Welfare Populations: Data Collection and Research Issues* (Washington, DC: The National Academies Press), pp. 197~219.
- Jackson, N., and Lilleker, D. 2011. “Microblogging, Constituency Service and Impression

- Management: UK MPs and the Use of Twitter. *The Journal of Legislative Studies*, vol. 17, no. 1, pp. 86~105.
- Jackson, N., and Lilleker, D. 2020. "Microblogging, Constituency Service and Impression Management: UK MPs and the Use of Twitter," in *The Impact of Legislatures* (London: Routledge), pp. 414~433.
- Janssen, M., and van den Hoven, J. 2015. "Big and Open Linked Data (BOLD) in Government: A Challenge to Transparency and Privacy?" *Government Information Quarterly*, vol. 32, no. 4, pp. 363~368.
- Kitchin, R. 2014. "Big Data, New Epistemologies and Paradigm Shifts," *Big Data and Society*, vol. 1, no. 1, pp. 1~12.
- Laney, D. 2001. "3D Data Management: Controlling Data Volume, Velocity and Variety," *META Group Research Note*, vol. 6, no. 70, pp. 1.
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., and Barton, D. 2012. "Big Data: the Management Revolution," *Harvard Business Review*, vol. 90, no. 10, pp. 60~68.
- Mitchell, M. N. 2010. *Data Management Using Stata: A Practical Handbook* (College Station, TX: Stata press).
- Pirog, M. A. 2014. "Data Will Drive Innovation in Public Policy and Management Research in the Next Decade," *Journal of Policy Analysis and Management*, vol. 33, no. 2, pp. 537~543.
- Pretty, I. 2013. "Big Data in the Public Sector," Capgemini: www.in.capgemini.com/media-analysts/resources/video/big-data-in-the-public-sector (accessed: September 10, 2015).
- Schroeder, R., and Cowls, J. 2014. "Big Data, Ethics, and the Social Implications of Knowledge Production," in Data Ethics Workshop, KDD@ Bloomberg, August (vol. 24, pp. 1~4).
- Seegerberg, A., and Bennett, W. L. 2011. "Social Media and the Organization of Collective Action: Using Twitter to Explore the Ecologies of Two Climate Change Protests," *The Communication Review*, vol. 14, no. 3, pp. 197~215.

Taylor, L., Schroeder, R., and Meyer, E. 2014. Emerging Practices and Perspectives on Big Data Analysis in Economics: Bigger and Better or More of the Same?" *Big Data and Society*, vol. 1, no. 2.

Tinati, R., Halford, S., Carr, L., and Pope, C. 2014. "Big Data: Methodological Challenges and Approaches for Sociological Analysis," *Sociology*, vol. 48, no. 4, pp. 663~681.

Woollard, M. 2014. "Administrative Data: Problems and Benefits. A perspective from the United Kingdom," in Duşa, A., Nelle, D., Stock, G. and Wagner, G. G., eds., *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences* (Berlin: SCIVERO Verlag), pp. 49~61.