



圖片來源: pixabay

特稿

全球競爭下政府主導的「資料戰爭」¹

樊家忠²

國立臺灣大學經濟學系教授

〈摘要〉

隨著大數據的發展，如何運用資料來提高競爭力已經是商業常態。然而，資料競爭不只限於商業領域，各國政府也掌握了大量行政資料可以支援各類研究，幫助公共政策的設計與改良，並協助產業發展。本文介紹當前各國政府在行政資料管理與運用上的制度差異，說明北歐制度相對於美國、德國制度的優越性，以及北歐如何利用這樣的優越性吸引全球人才（brain drain）。本文並且以丹麥為例說明該國如何在法規與制度設計上同時達到資料開放與個資保護的雙重目標。本文最後建議臺灣政府應該發展出類似北歐行政資料的開放體系，讓臺灣豐富的行政資料可以創造出研究價值，並幫助臺灣政府建立循證治理的基礎。

關鍵字：政府行政資料，大數據，GDPR，資料科學，個資法

1 本文改寫自作者在「看雜誌」上的兩篇文章，分別為「政府主導的『資料戰爭』（一）」（<https://www.watchinese.com/article/2016/22306>），以及「政府主導的『資料戰爭』（二）」（<https://www.watchinese.com/article/2016/22403>）。

2 電子郵件信箱：elliottfan@ntu.edu.tw

壹、引言

隨著全球化不斷地發展，加深了不同貿易體之間的競爭程度，國家間的競爭型態也不斷地在演變。近年來，「大數據」（Big Data）的發展受到各界的關切，許多人已經注意到數據的應用可以增加企業的營運效率，提高競爭的優勢。

企業間的競爭如此，國家間的競爭也是如此。各國政府都有各種不同的「行政資料」，這些資料涵蓋範圍廣且詳盡，包含了小至個人、家庭層級，大至學校、企業、各類團體的資料。大多數行政資料涵蓋的人口數量很廣，遠超過任何單一廠商或者研究者可以蒐集的範圍。如果能夠滿足個資保護的需要，這些資料潛藏著不可限量的價值。因此，許多國家開始在思考如何運用這些資料在各類公共政策效果的評估上，以及改善政策的設計，甚至幫助企業的發展。

目前國際間的資料戰爭早已開打，只是先前被中美貿易戰、近來被俄烏戰爭給掩蓋了，沒有得到人們普遍的認識。但雖然隱形，資料戰爭卻對國家長期的發展有重大影響。本文簡介臺灣與各國行政資料的使用現況，以及近年來各國政府之間的競爭策略，並提出對臺灣政府的建言。

貳、行政資料的優點

以臺灣為例，政府的行政資料由各部會根據其業務分別管理。主要有衛福部的「健保資料庫」、勞動部的「勞保資料庫」、內政部的「戶籍資料」（包含戶籍登記、出生、死亡、結婚離婚等紀錄）、財政部的「財稅資料庫」、教育部的學籍資料，以及大考中心與招聯會的大學申請、考試、與入學登記、經濟部的「海關資料」、行政院主計總處的「人口普查」與「廠商普查」等。

跟調查資料比起來，行政資料具備以下幾個優點，使其具有特殊的價值。第一，行政資料鮮少有資料遺漏或者受訪者因故不能逐年連續受訪導致「樣本流失」（sample attrition）的問題。這使得行政資料的分析結果有全國的代表性，減少小樣本抽樣時出現偏誤的疑慮。第二，行政資料的樣本數大，這使得針對少數族群（例如原住民、弱勢族群、小區域居民、或者罕見案例）的研究可以進行，調查資料往往受到樣本數量限制，無法照顧這些少數。第三，行政資料往往多年連續，而且對同一個人、家庭、或者廠商進行追蹤，因此能夠提供研究者進行長期的追蹤研究。第四，行政資料準確性較高，一般透過調查取得的資料往往有受訪者不願透漏實情

而謊報的情形（例如低報所得），而行政資料相對準確。

參、各國行政資料使用狀況

一、北歐各國

過去由於個資保護以及技術限制，各國政府習慣的作法通常是不開放行政資料給學術機構或者民間企業使用。但如今除了少數例外，為了發揮行政資料的價值，近年的趨勢是各國開始爭相走向開放的道路。其中北歐尤為全球之冠，成為開放系統的典範。北歐能取得優勢主要是它們跟西歐、美國不同的是北歐國家有身分證號碼以及戶籍登記系統，使得各項政府資料可以用身分證號碼串聯起來。另外，北歐的各國的社會福利政策、其他各項公共政策，以及社會保險都是全國統一開辦的，由中央政府統籌。這使得行政資料都統合在中央政府手裡，且可透過身分證號碼以及戶籍資料串橫向串聯各部會資料，也可以縱向追蹤每個人的時間動態。

二、以丹麥為例

丹麥政府所蒐集的行政資料涵蓋面非常廣泛，他們利用個人的身分證字號（personal identification numbers, CPR），廠商 ID（companies' identification numbers, CVR），還有住宅 ID（data about buildings and accommodations, BBR）將所有關於個人出生、家庭、教育、就業、所得、消費、財富、居住、遷徙、醫療、福利、死亡等超過 250 種主題的資料都串連起來，集中在丹麥統計局（Statistics Denmark, SDK）管理³。

同時，資料不但開放給本國大學以及研究機構使用，外國研究者也可以透過跟當地學術單位合作取得資料（參見圖 1）。尤其，資料也開放民間企業或組織可以將自己的資料與行政資料合併進行分析，以幫助企業進行市場分析與各項研究。

為了方便使用資料，不在當地的人還可以透過個人電腦遠端連線作業，這大幅增加了使用資料的方便程度。特別是在 Covid-19 肆虐下，各國因防疫需求而面臨不同程度的管制措施，部分行政機關內存放資料之單位亦因此須暫時關閉。然而，北歐透過遠端系統來使用資料的管道並不會因此而中斷。

3 請參見丹麥統計局網頁：<https://www.dst.dk/en>。2022/09/26。

The screenshot shows the 'STATISTICS DENMARK' website with a navigation bar at the top containing 'Statbank', 'Contact Statistics Denmark', 'Information services', 'Press', and 'Dansk'. Below the navigation bar are links for 'Find statistics', 'Products & services', 'Consulting abroad', 'About us', and a search icon. The main content area is titled 'SUBJECTS' and features a grid of nine categories, each with an icon and a list of sub-topics:

People	Labour and income	Economy
Names, population, health, elections, housing, church, gender equality...	Employment, unemployment, earnings, income, wealth...	GDP, price index, national accounts, imports and exports, consumption, property trading...
Social conditions	Education and research	Business
Criminal offences, social benefits for senior citizens, cash benefits, placements...	Number of students, education programmes, innovation...	Enterprise statistics, sectors, accounts, entrepreneurship, bankruptcies...
Transport	Culture and leisure	Environment and energy
Cars, goods transport, passenger transport, infrastructure, traffic accidents...	Film, media, museums, music, digital behaviour, sports...	Emissions, environmental-economic accounts, organic production and trade, area...

資料來源：丹麥統計局 (<https://www.dst.dk/en>)。2022/ 09/26。

圖 1：丹麥統計局關於開放資料給研究使用的網頁

北歐各國行政資料的完備程度與開放程度，不但吸引了這些國家的國內學者大量地利用這些資料進行研究，同時也吸引了國際各界學者的參與，其中包括為數眾多的國際頂尖研究團隊。由於許多實證研究計畫必須有大量的詳細個體資料才能進行，北歐因此成為研究者的樂園。北歐的系統等於是讓國際的研究人才免費替北歐國家效力，因為他們的研究成果都可以供作這些政府做政策參考，使各項公共政策規劃更加完善。換句話說，北歐透過一個資料開放系統吸引全世界的人才（brain drain）。

三、美國學界憂心忡忡

北歐的開放制度在其他國家並不多見，美國現行的分治制度（decentralized system）就與北歐的集中制度（centralized system）大不相同。首先，美國聯邦政府並沒有一個像丹麥統計局那樣的全國性系統，將各個不同中央行政單位的資料集中起來管理，而是各單位、各資料中心、各州政府與地方政府有著不同的規定與管理方式。因此，研究者如果要做全國性、各州比較時，需花費相當多的時間成本來蒐集資料，當然也未必能順利取得使用資料的管道。其次，跟北歐國家比起來，美國無論在資料整合的程度以及開放的程度都低很多，大多數的研究者都只能使用一些局部的或者單一的行政資料進行研究，跟丹麥統計局將所有 250 項行政資料串連起來的情況有天壤之別。

這樣的現象讓許多美國的學者開始憂心，他們擔憂美國自從二次大戰以來在

研究資料上以及實證研究上的優勢，會隨著行政資料的重要性日增而漸漸喪失。四位重量級美國經濟學家 David Card（2021 年諾貝爾獎得主）、Raj Chetty、Martin Feldstein，以及 Emmanuel Saez 就曾經撰寫過專文分析美國在行政資料使用上的各項限制，導致許多美國的研究者轉向去使用北歐的行政資料。另外，由於資料的限制，許多重要的議題不能利用美國的資料進行研究，而國外的研究成果也未必能夠拿來了解美國本土的問題。他們對這樣的現象表示擔憂，認為這將使美國慢慢喪失在各項研究上的優勢（Card et al., 2010）。

有鑒於此，美國多位學者近年來開始呼籲美國政府應該改善行政資料的管理系統，並有效地開放給學術研究使用。美國著名的經濟學家 David Grusky 等人於 2019 年曾撰文描述美國在實證研究上的困境，於是他們呼籲美國政府成立 American Opportunity Study，一個在功能上類似丹麥統計局的資料中心，可以將政府各項行政資料、調查資料以及各個研究蒐集到的資料進行全面的彙整，並開放給學術研究使用。他們認為這樣的資料中心將有助於各個領域的研究，特別對於個體的長期追蹤研究有所助益（Grusky et al., 2019）。

更具前瞻性的是，前段所提四位經濟學家向美國政府提出建言，希望未來能建立一個跟北歐各國中央整合型系統不同的制度。他們提出的構想是由民間來經營各項資料中心，因為「中央統一」向來不是美國制度的傳統。他們認為由民間設立的資料中心互相具備競爭性，而競爭會提高服務品質。這些經濟學家的信念與盤算是，一個開放民營的競爭系統將遠比一個政府公務員主導的資料中心來得有效率，所以這樣的系統將會超越北歐，使得美國重新奪回全球的資料優勢地位。

四、德國出奇兵以致勝

北歐的優勢不但讓美國憂慮，也讓西歐那些自詡「進步」的國家感到壓力。德國為了迎頭趕上，近年來設置了兩項「武器」。第一，政府不但在德國各地建設符合資安標準的資料中心，並且在美國幾所頂尖大學內設立資料中心，讓這些大學的使用者不需到德國總部，即可透過鄰近的資料中心進行遠端連線取得大量行政資料。第二，資料申請容易，而且免費。這樣的設計其實也是想吸引美國的研究人才投入對德國的實證研究，也就是想搶北歐的生意。但是德國的資料內容遠比北歐貧乏。前面提到過丹麥統計局將每個國民所有從出生到死亡等超過 250 種主題的行政資料都串連起來，而德國的行政資料目前主要是個人就業資料，其豐富性完全無法跟丹麥相提並論。針對這個限制，德國當局也在未來發展計畫裡提出要將各個政府部門

的行政資料串連起來。

五、以教育研究為例來跨國比較

教育應該是與醫療並列為政府施政的兩個最重要的領域，可是跟臺灣在醫療領域的大量研究相比，利用臺灣的資料來進行教育相關議題的研究卻遠遠不及。美國西北大學的 David Figlio 教授、Emory 大學的 Krzysztof Karbownik 教授、與挪威經濟研究所的 Kjell Salvanes 教授在 2016 年的一篇論文中，整理了近幾十年來發表在嚴肅學術期刊中利用政府行政資料來研究教育議題的論文，其結果非常怵目驚心。如表 1 所示，在兩位學者整理的數百篇論文當中，92% 以上都是使用北歐國家或者美國國內的行政資料，其中有 23% 是關於北歐的研究，這顯然是北歐開放資料鼓勵研究的成果（Figlio et al., 2016）。

令人驚訝的是，用亞洲資料的只有 4 篇，僅占全數論文的 1.1%，比南美洲國家的論文數量比例 4.9% 都少得多，這應該是亞洲國家普遍將行政資料封閉起來的結果，而使用臺灣行政資料的研究僅僅只有 1 篇。臺灣在教育議題上研究的匱乏，主要的原因正在於行政資料的限制。想要對於教育議題進行深入的研究，除了需要教育部的各項行政資料之外，還需要合併戶籍資料以便了解個別人的家庭背景，以及合併勞動市場的資料（例如勞保與公保資料）才能了解個人在受教育之後的參與勞動市場的結果。可惜的是，這些關於教育、家庭、勞動的行政資料通通不在政府開放之列，更遑論不同資料的相互串聯。

表 1：1990 - 2015 以來利用行政資料研究教育議題的論文發表統計

地區	數量（篇）	比例
北美洲 (North America)	253 (美國：240)	68.9%
北歐五國 (Nordic countries)	85	23.2%
南美洲 (South America)	18	4.9%
亞洲 (Asia)	4 (臺灣：1)	1.1%
大洋洲 (Oceania)	1	0.9%

資料來源：統計資料翻譯自 Figlio et al. (2016)。

肆、個資保護與 GDPR

想當然爾，北歐的開放資料系統引發了個資保護的疑慮。同時，歐盟為提升個人資料保護的規範，並建立起一個歐盟各國統一的規範原則，於 2016 年通過了 GDPR (General Data Protection Regulation)，也引發了北歐的資料開放系統是否跟

GDPR 相容的疑慮。

針對這個問題，筆者於 2017 年親自採訪了丹麥統計局的資料主管 Ivan Thaulow，得到關於上述問題的答案。自西元 1970 年代起，丹麥行政資料的主管機關——丹麥統計局（Statistics Denmark）就開始提供個體資料給官方及民間機構，作為各類研究使用。丹麥政府處理資料的原則是：

一、政府蒐集的行政資料屬於整體公眾，個人沒有權利要求將自己的資料自資料庫中移除；

二、基於促進公共利益的研究用途下，資料使用不需要尋求個人的同意。

丹麥統計局是根據《丹麥統計局組織法》（Act on Statistics Denmark）組成，該法中匯集了聯合國的《官方資料基本準則》（Fundamental Principles of Official Statistics），以及歐盟的《歐洲統計應用規範》（European Statistics Code of Practice）。無論是《丹麥統計局組織法》，或是《歐洲統計應用規範》，都明確規範個體資料可以應用在研究用途、為了公共利益、科學或歷史研究的用途、或是統計用途⁴。因此丹麥開放資料的做法並沒有違背 GDPR 的精神。

另外，為了防止個資洩漏的弊端，除了個人資訊去識別化之外，丹麥統計局也實施一種連坐法。要申請使用統計局的資料，必須申請人（例如某大學教授）的所屬研究機構（某大學某系所）跟統計局簽約，由該機構來負責監督資料的使用。一旦出現違規現象，整個研究機構都會被處分以及停權，處分強度視違規程度而定。當在考慮核發授權時，丹麥統計局會非常小心地評估申請機構，特別是當申請機構是屬於私部門時，丹麥統計局會將申請機構的可信度納入考慮，包括評估機構所有者、職員們的教育水準、研究經驗，以及過去的信用。一旦某個機構得到授權，所有跟該機構合作的研究者都可以申請使用個體資料，丹麥統計局並不會單獨授權給任何個人。

Thaulow 對筆者表示，丹麥這套制度已經成功運行多年，違規的事件非常少，而且幾乎都是技術違規，也就是由於對資料使用規則不熟悉所導致的失誤。例如某丹麥知名大學的經濟系就曾經因為某個年輕教授列印了 13 張螢幕上的畫面（上限

4 有關丹麥行政資料使用的立法、授權，以及相關規定的更多資訊，請詳見：<https://www.dst.dk/en/OmDS/lovgivning>。2022/09/26。

是 10 張），而導致全系所有教授被停權三個月，可想而知該年輕教授的下場。

伍、臺灣行政資料使用現況

臺灣中央政府各部會的行政資料（個體資料），除了政府為了政策研究需求而徵求學界參與的計畫案之外，普遍來說並未開放給學界進行研究。少數的例外是衛生福利部的「衛生福利資料科學中心」有提供制式的管道與申請流程讓研究者取得健保資料（衛生福利部頁面如圖 2）。該中心不僅提供了「全民健康保險研究資料庫」，亦有其他相關的加值資料檔，主要來自衛福部與其他部會進行的各類調查。該中心在把所需資料去識別化之後，提供給國內的研究者在中心內的電腦上使用。

衛生福利部統計處
Ministry of Health and Welfare

請輸入關鍵字 進階

熱門關鍵字： 防疫補償 COVID-19 隔離 確診

關於本處 組織架構 業務職掌 焦點新聞 活動訊息 公告訊息 招標資訊 本部各單位及所屬機關

統計處

衛生福利資料科學中心專區
重要公告及下載專區
統計資料發布時間及變更
衛生福利統計分析
死因統計
國民醫療保健支出(NHE)
衛生福利統計指標
衛生福利公務統計
衛生福利統計調查
統計書刊及電子書
性別統計專區
政府開放資料集

衛生福利統計分析

顯示條件查詢

共 26 筆資料，第 1/2 頁， 1 2 ▶▶

標題	發佈日期	更新日期	檔案格式	下載次數
國際身心障礙者日衛生福利統計通報	110-12-02	110-11-30	PDF	4960
醫師節衛生福利統計通報	110-11-10	110-11-10	PDF	3033
世界視覺日衛生福利統計通報	110-10-13	110-10-13	PDF	8838
世界心理健康日衛生福利統計通報	110-10-08	110-10-13	PDF	8541
國際失智症日衛生福利統計通報	110-09-17	110-09-16	PDF	6175
世界過敏性疾病日衛生福利統計通報	110-07-01	110-07-01	PDF	5962
世界衛生日衛生福利統計通報	110-03-23	110-03-23	PDF	2118
世界聽力日衛生福利統計通報	110-02-22	110-03-23	PDF	2453
世界抗癌日統計通報	110-01-28	110-01-28	PDF	5722

資料來源：衛生福利部統計處（<https://dep.mohw.gov.tw/DOS/lp-4445-113.html>）。2022/09/26。

圖 2：衛生福利部統計處－衛生福利統計分析頁面

雖然在使用上有場地、收費、計畫需經過審查等限制，相關的健保資料仍然普遍被政府、學術、與醫療機構的研究者使用，產出了大量的專業研究成果。根據衛生福利部的網站統計顯示，直至 2011 年為止，國內研究人員向衛福部申請的研究

案已經超過 300 件，僅僅在 2008 到 2010 這三年間利用健保資料發表於國際期刊的論文就超過 200 篇，其中發表在高影響力（影響係數大於 3）的期刊就有 78 篇。這顯示了健保資料的開放有助於提高國內醫藥、公共衛生研究成果，這是臺灣其他研究領域無法望其項背的。很不幸的是，不久前憲法法庭作出判決，認定健保資料目前的使用部分違憲。

目前除了衛生福利資料科學中心之外，政府其他各部會舉凡勞動、教育、戶籍、警政、貿易、個別廠商等個體行政資料基本上都沒有開放。其結果就是不論是學界的研究者與各界的專家都無法透過這些豐富的實證資料來了解各項議題，造成臺灣學術發展與政策研究上莫大的損失。

因此，我們呼籲政府部門參照其他國家模式進一步開放行政資料的使用。尤其臺灣擁有身分證字號與完善的戶籍系統，且有集中納保的勞工保險與健康保險制度，使臺灣政府擁有與北歐媲美的行政資料庫。

陸、結語

基於以下三個理由，筆者建議臺灣政府的行政資料應該發展出類似北歐的開放系統，由行政院成立一個獨立於其他部會的資料中心，統一管理並整合各部會的行政資料。第一，臺灣有身分證字號以及完善的戶籍登記系統，不但可以完整記錄一個人從出生到死亡所有的軌跡，各部會的資料也可以透過身分證字號進行串聯，建立起類似丹麥的整合資料型態。第二，由行政院跨部會層級來進行整合，方便各部會間的協調與授權，能最大化資料範圍。第三，臺灣社會逐漸重視個人權益的保護，由中央政府來主辦資料中心，建立出一個通行的規範讓大家遵守能避免爭議。

參考文獻

- Card, D., Chetty, R., Feldstein, M.S., and Saez, E. 2010. "Expanding Access to Administrative Data for Research in the United States," *Political Economy: Structure & Scope of Government eJournal*.
- Figlio, D. N., Karbownik, K., and Salvanes, K. G. 2016. "Education Research and Administrative Data," in *Handbook of the Economics of Education* (Amsterdam: Elsevier), vol. 5: pp. 75~138.

Grusky, D. B., Hout, M., Smeeding, T. M., and Snipp C. M. 2019. “The American Opportunity Study: A New Infrastructure for Monitoring Outcomes, Evaluating Policy, and Advancing Basic Science,” *The Russell Sage Foundation Journal of the Social Sciences*, vol. 5, no. 2: pp. 20~39.