

美國普查局研究資料中心簡介

個別資料提供政府制定決策及研究單位或企業研究分析使用，有助提升統計資料應用效能，且藉由知識淵博的研究人員之嚴格分析，可收檢驗統計資料品質之效。然而，個別資料之提供，須兼顧保密原則，為此，各國統計機構莫不尋求最適釋放方式，而「研究資料中心」或稱「資料實驗室」乃目前歐美採行方式中較成熟完備者，本文以美國商務部普查局為例，簡述其發展供參。

◎曾雪雅、劉麗瓊、王珮倫（行政院主計處第三局科長、專員、科員）

壹、前言

隨資訊技術進步，統計調查個別資料（Microdata）相較於總體資料（Macrodata）有時更能符合研究者細緻化分析的需求，若能在兼顧個人資料保密原則下，適度釋放個別資料供政府制定決策及研究單位或企業研究分析使用，將有助提升統計資料的應用效能。為提供研究者更多資訊，使個

別資料發揮最大效益，並兼顧個別資料保密，各國統計機構莫不尋求最適釋放方式，而「研究資料中心」（Research Data Centres）或稱「資料實驗室」（Data Laboratories）乃目前歐美採行方式之一，其運作方式係在高度安全作業管控下，研究人員於研究資料中心進行讀取個別資料之研究。目前美國、澳洲、紐西蘭、荷蘭、加拿大、義大利等國均有

相關的設置，本文僅以運作已逾20年之美國商務部普查局之研究資料中心（以下簡稱RDCs）為例，簡述其發展供參。

貳、RDCs 的沿革

美國普查局將個別資料區分為公用之個別資料（Public Use Microdata, PUM）和非公用之個別資料二種，前者包括

人口普查、社區調查及消費支出等15種個別資料，在去除可辨識個人身分之欄位資料後，置於普查局網頁（<http://dataferrett.census.gov>）免費供大眾使用。

至於非公用之個別資料，則早在1950年代起，即有零星提供之案例，惟至1980年代初期始建立系統化提供機制。鑑於公用之個別檔案並未包括企業資料，加以其家庭調查之個別資料，僅限於地區、所得和職業等粗略變數，侷限了統計資料的效用與附加價值，因此普查局於1983年成立了經濟研究中心（The Center for Economic Studies；現為美國普查局RDCs的總部），專責執行、協助及支持政府、學術界、企業界進行非公用之個別資料研究。其後，為肆應各界對於個別資料之需求日殷，以及衡酌全美研究者均須赴經濟研究中心（華盛頓D. C.）的

成本考量，普查局分別於1994年及1997年新設2處RDCs，且由於這2處具試驗性及指標性RDCs之運作成功，促使普查局於2000年以後陸續成立RDCs，目前全美總計9處，大多設於知名大學校區，詳如下表。

參、RDCs的運作

一、服務內容與對象

美國普查局提供非公用個別資料給申請合格的研究者使用。目前所提供之資料集總計43種，包括普查局主辦普(調)

查24種、其他機關委辦調查9種、其他政府單位調查5種及其他5種；而研究者主要包括學術界、獨立的研究組織（如國家經濟研究局，NBER）以及各級政府機構之人員。

二、計畫之申請、審核及執行

為保護個別資料隱私，凡欲申請使用非公用個別資料之研究，均經嚴格審核，若資料係由其他機關提供者，並須同時得到該機關允許，通過者應在RDCs執行個別資料之存取，而進入RDCs之研究人

普查局 RDCs 設置地點

1	總部	經濟研究中心
2	分部	波士頓RDCs
3	分部	加州柏克萊大學
4	分部	加州洛杉磯大學
5	分部	密西根大學
6	分部	芝加哥RDCs
7	分部	紐約RDCs
8	分部	康乃爾大學
9	分部	杜克大學

員，在法律上等同普查局員工，須受普查法之限制與約束。茲就研究計畫之申請、審核及執行分述如下：

(一) 研究計畫之申請：

1. 政府、研究單位或企業機構欲使用非公用之個別資料以執行研究計畫者，須先至普查局之經濟研究中心（以下簡稱CES）網頁登記，申請成功者會收到由系統產生的電子郵件，包含密碼及帳號，並利用此帳號向CES提交計畫。
2. 計畫方案之內容應包括計畫概述、研究目的、經費來源、研究人員、執行場所、需用資料及軟體，以及該計畫對普查局業務之助益說明等。研究人員透過線上計畫管理系統提出申請計畫，經RDCs初審及提供修改建議後，研究人員再向CES提交最終計畫以進行最終的審核。

3. 最終計畫必備包括三項pdf格式的文件：①摘要；②計畫說明書，內容包括需用之普查局或非普查局資料；③有利於普查局業務之說明。

(二) 研究計畫之審核：

1. 審核新研究案之研究主題、資料集及揭露風險約歷時2個月，如有需要，亦可邀集外部研究機關（如國家科學基金會，NSF）或委辦機關進行會審。
2. 為審慎保密個別資料，建議研究者使用迴歸分析等統計方法，而儘量避免表列資料，主因後者很難評估揭露之風險；另普查局政策辦公室會就研究計畫與該局政策是否一致進行查核。對於存在揭露風險或與該局政策不明確相符之計畫，再進一步交由資料保管執行政策委員會複審。

3. 五項審核標準：

- ① 對普查局業務有益：普查

局訂定13項評估標準來衡量研究計畫是否有助於該局之業務，其中包括可提高調查資料品質、改善未回應問項之設算、增加調查資料之附加價值等；並特別規定，若須使用聯邦政府稅收資料，該計畫必須符合其中5項標準；若否，僅須符合1項以上即可。

- ② 具科學價值：評估研究計畫對於現有知識是否有貢獻。
- ③ 具使用非公用個別資料之必要性。
- ④ 具可行性：研究計畫必須顯示出該研究可成功執行。
- ⑤ 揭露風險：需評估研究計畫的結果是否有資料外洩的風險。

(三) 研究人員之特殊宣誓身分：

研究計畫一旦通過審核，

接觸保密資料的研究人員必須接受背景調查，並按捺指紋等；之後，普查局便賦予研究者特殊宣誓身分（Special Sworn Status, SSS），研究期間，不論有意或無心揭露個別保密資料，依法與普查局員工負相同刑責，最高可處5年以上有期徒刑或科25萬美元罰金，或二者併罰；且永不再核准其使用個別資料。

（四）RDCs執行長的角色：

RDCs執行長除負責提供個別資料防制揭露手冊、進行揭露分析、灌輸研究人員保密文化外，還扮演CES與研究人員間之溝通橋樑，因此執行長需具備研究背景。

（五）簽約與收費：

1. 凡經核准之研究計畫，研究人員須與CES簽訂正式合約，載明研究範圍、使用的資料集、計畫期程及使用RDCs的時間與頻次。
2. 各RDCs有統一的收費標

準，一般包括直接的實驗成本，如人事、場所及電腦等費用；另如須產生新的資料集，或與其他資料集合併等額外服務，則另支付特別費用。

3. 研究結果若對普查局有顯著貢獻，如協助資料庫發展等，則有折扣優待；此外，對在校學生之研究費用亦另有補助優惠。

肆、RDCs的安全機制

個別資料保密係RDCs之最高指導原則，依照普查局所訂之標準，任一RDCs都有其安全操作場所及程序，綜觀其保密機制，摘述重點如下：

一、嚴格限制進出人員

除了普查局員工外，僅放行具備特殊宣示身分者，包括RDCs工作人員、電腦工程師及獲准可在RDCs內執行研究

計畫之研究人員。SSS人員須簽署保密協定，並備有普查局發給的安全許可證。

二、安全的電腦環境

（一）RDCs之電腦設定為無法複製資料到任何儲存媒體、與外界隔絕、無法上網及使用email，而研究人員亦不准攜帶個別的電腦設備，包括隨身碟、壓縮程式等，嚴禁將保密資料從RDCs辦公室攜出。

（二）其電腦環境由過去PC與Unix Workstations為主的區域網路，改為精簡終端機（Thin Client），存取中央伺服器的作業環境，讓所有實際資料被安全置於普查局中央伺服器上，且其與外部網路及普查局內網均設有防火牆隔絕，而研究人員於RDCs內使用精簡

終端機，透過加密的通信線路存取中央伺服器資料，其安全係數不但增加，更減省資料存於RDCs內之管理成本。

三、其他操作規定

- (一) 所有的分析必須於RDCs內完成，只有在RDCs工作人員監督下，才可列印報告。
- (二) 研究結果，必須經過揭露審核；研究人員離開RDCs前，必須將所有存於設備中的相關結果予以清除。
- (三) 同意放行之研究結果，RDCs的管理人員將email給研究者。

伍、RDCs的研究成果

為保存完整的研究紀錄，並有利於未來之研究，CES希望在RDCs進行的研究論文均

上載CES網頁論文篇數

單位：篇

年	論文篇數	年	論文篇數	年	論文篇數
1988	4	1994	15	2000	13
1989	9	1995	13	2001	19
1990	8	1996	15	2002	24
1991	10	1997	16	2003	19
1992	14	1998	18	2004	20
1993	17	1999	17	2005	29

備註：2006年截至10月23日止計22篇。

上載至該中心網頁之研究報告討論專區。由其網頁資料顯示，上載論文由1988年4篇增加至2005年的29篇，而今（2006）年1—10月亦已達22篇（各年上載論文篇數詳上表），顯示美國利用個別資料進行之研究呈增加趨勢。而就近二年研究內容觀察，應用於經濟面之研究占七至八成，探討主題相當廣泛，內容包括產業生命週期、進口物價壓力、技術成長、工廠生產力動態、生產要素價格等；餘二至三成則為社會面之研究，研究主題多以勞動、社會保險及貧窮等

議題為主。

陸、結語

將非公用個別資料提供給知識淵博的研究人員，藉其研究結果促進決策品質提升，可提高統計資料的附加價值；且透過研究者嚴格分析，亦可藉以檢驗統計資料品質，故若能確實做到個別資料保密，對供需雙方而言，可謂互蒙其利。由美國普查局研究資料中心發展二十餘年之經驗及成果顯示，在保密原則下，個別資料之提供應用已為時勢所趨。❖