



統計調查檢誤資訊作業之探討

行政院主計總處專責掌理全國國勢普、抽查，及統計行政管理與對外提供統計服務等資訊化作業，為整合統計資料處理資源，及提供便捷的檢誤作業服務，爰規劃建置「共通性檢誤作業資訊系統」，以滿足多元資料檢核需求，增進統計資訊推廣及應用。

李哲綸（行政院主計總處主計資訊處研究員）

壹、前言

依據統計法定義：「統計調查」係指各機關基於統計目的及業務需要，向個人、住戶、事業單位、機關或團體舉辦之調查，包括基本國勢調查、指定統計調查及一般統計調查。

行政院主計總處（以下簡稱主計總處）統籌我國基本國勢調查與大型抽樣調查業務如國富統計、受僱員工薪資調查及生產力統計，近期將舉辦三大國勢普查包括 109 年人口及住宅普查（以下簡稱人口普

查）、109 年農林漁牧業普查（以下簡稱農業普查）、110 年工業及服務業普查（以下簡稱工商普查），涵蓋全國範圍，所需人力、物力及經費龐大，資料浩繁，程序相當複雜。

各機關舉辦之指定或一般統計調查繁簡不一，或以委外或建置系統調查，為確保資料正確性，各類調查資料均須進行檢誤作業，若有一套系統能適用各種統計調查檢誤作業，將能減少重覆投資、擲節政府經費，主計總處遂規劃「共通性檢誤作業資訊系統」，先以具有龐大資料量之三大國勢普

查為使用對象，制定標準化資料檢核程序，並運用新興資訊技術強化系統功能，提升資料處理效率。未來將再評估適用其他統計調查之可行性。

貳、共通性檢誤作業資訊系統簡介

資料檢誤工作為接收紙本問卷、OCR 辨識、網路填報等來源資料，依統計調查問項之檢核條件，反覆地進行正確性校驗及修正，直至符合資料處理要求。將檢誤完成資料，透過問項計算或格式整併，產製統計結果編製前資料。

目前係運用三大國勢普查資料，進行大量普查資料檢核，可容納百萬筆普查資料，適用各業、表別檢誤條件、調查表、更正審核等機制，藉以完成三大國勢普查檢誤作業，實證系統符合檢誤資料處理需求，具大量資料檢核時效優勢（圖 1）。

參、共通性檢誤作業 資訊系統特點

因應統計調查各自檢誤

作業程序之差異性，本系統採取共通性模組設計，制定資料標準規則、親和友善檢誤設定、更正審核程序、彈性資料整併等，以利程序異中求同；又考量作業繁雜及資料量龐大，亦須強化系統效能與充足資源。本系統以三大國勢普查資料運行實例，具有下列特色：

一、建立共通資料格式， 制定資料標準規則

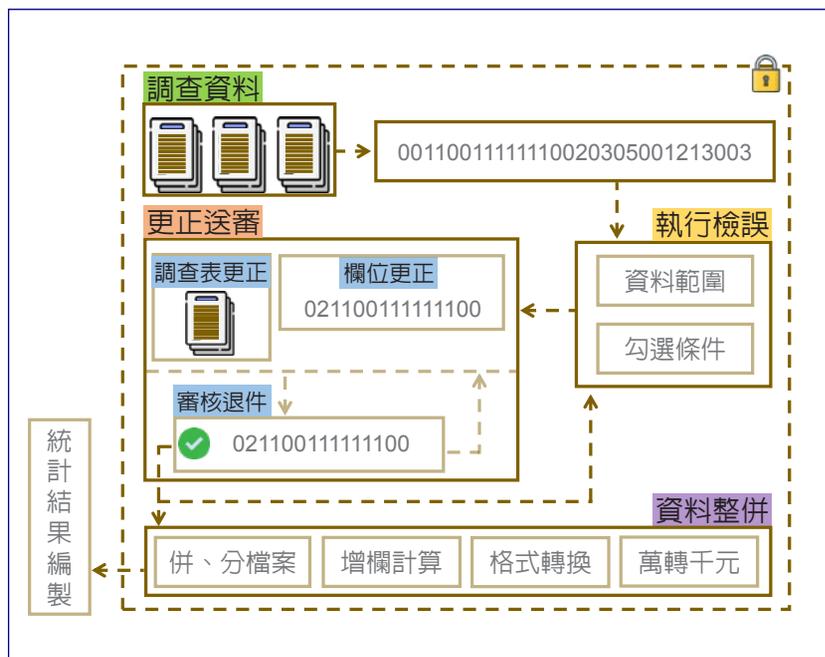
三大國勢普查之調查對象各有不同，人口普查具單宅多戶多人之住戶情形；農業普查具諸多農作物及畜牧與漁類品項；工商普查具單一企業單位之經濟產銷結構，以致問項資料差異性極大，須訂定資料標準規則，建立共通資料格式介面，達到異質性資料通用需求。

我們可以將資料集比喻為共存容器，此容器包括資料集名稱、數值或文字內容物、資料數量等，各普查依此資料標準訂定，完成普查檢誤作業標準資料集。統計調查亦依本系統之資料標準規則，自行選擇線上介面、離線資料規則檔，設立各類型檢誤資料集，及提供設定續填多張綁定服務，以利統計調查編號歸戶（下頁圖 2）。

二、簡易檢誤設定介面， 節省人力與時程

三大國勢普查之業、表別執行檢誤，依各自調查問項訂定檢誤條件，使驗證普查資料正確性，故檢誤條件質量非常

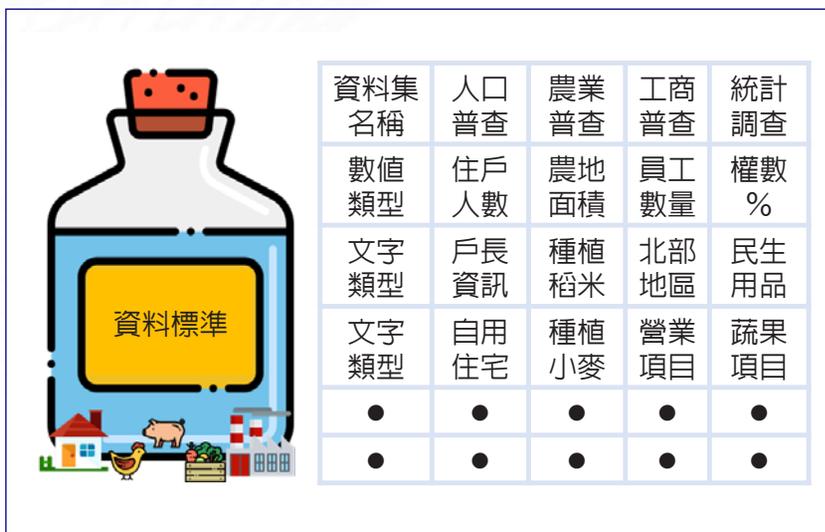
圖 1 共通性檢誤作業資訊系統示意圖



資料來源：作者自行繪製。

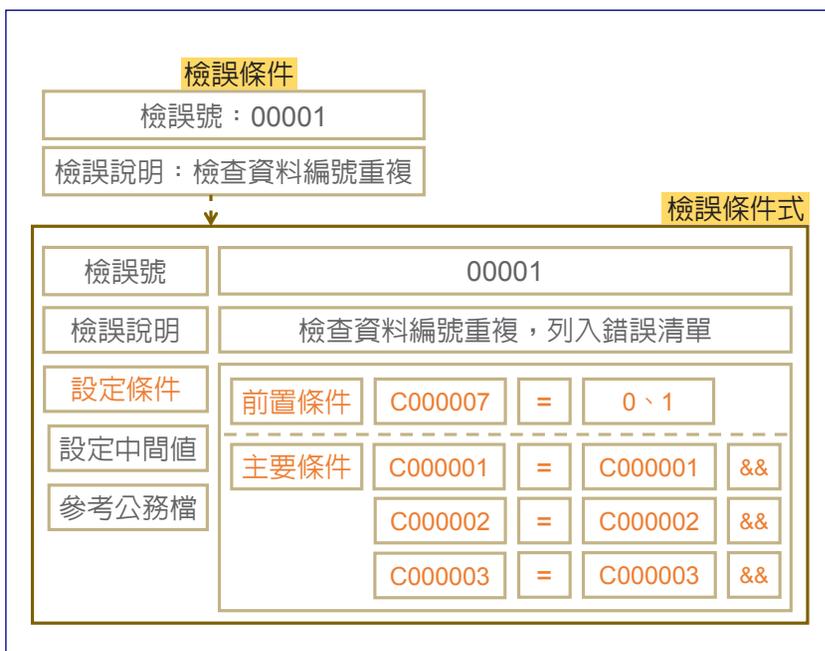
論述 » 管理 · 資訊

圖 2 資料格式標準示意圖



資料來源：作者自行繪製。

圖 3 設定檢誤條件式示意圖



資料來源：作者自行繪製。

重要，每一業、表別約有百條檢誤條件，均需人工逐條撰寫程式相當繁瑣及費時，面對各普查尖峰期或週期重疊時，須在期限內及時完成數千條條件，極花費人力、時間等資源。

為提升條件設定作業效率與自動化程度，爰採客製化開發程式模組，設計檢誤條件設定介面，以簡易點選組合呈現條件語法，提供匯入或複製單筆、整批規則，及資料格式欄位連動檢誤條件，減少重複性調整操作，具親和友善、便利再用性之檢誤設定，降低檢誤條件設定花費成本。

統計調查所需檢誤條件，可轉化成本系統之檢誤條件式，由介面設定條件程序，點選前置、主條件欄位，或參考公務檔、計算中間值等，為避免條件語法有誤，提供檢視語法及標準化語法資訊供驗證，達到便利性地檢誤條件式操作（圖 3）。

三、快速有效更正程序，資料異動審核機制

三大國勢普查之業、表別執行更正，由檢誤後之待調整資料，進行欄位數值修正，須不斷持續重複檢誤、更正，直至符合資料除錯要求。本系統如同紙本調查表般，套用於線上調查表直觀地呈現填表資訊，連結檢誤條件自動查找表內需修正問項，及提供參考公務資訊，以利加速除錯程序。

考量檢誤資料異動上達百萬問項，評估有效資料更正流程，本系統具分派修正、異動審核機制，將大量異質資料依錯誤條件數，彈性分派予多位修正人員，由修正人員於資料更正後，將修正結果提交予業、表別承辦人員檢視、確認，始完成資料更正作業。

統計調查之更正程序，就資料量多寡選擇線上單筆、批次之更正送審與審核確認，亦提供離線批次更正檔服務，本系統設計共通性更正處理模式，以適用各類型統計調查更正需求，全部異動、審核內容皆留存，確保資料更正管控稽核。

四、彈性資料整併，快速轉換統計資料

三大國勢普查資料經檢誤處理後，須產出檢核完畢乾淨資料，提供業務單位進行資料整併、格式轉換及統計前處理，如調查問項外之參考性資訊、劃分行政區層級如縣市鄉鎮村里，以符合各普查業、表別所需匯出格式，作為下階段結果表系統、統計軟體、視覺化應用之資料來源。

本系統採用共通格式轉換、計算模組，適用統計調查檢誤過程，可彈性之資料轉化服務，便利資料格式變動，以

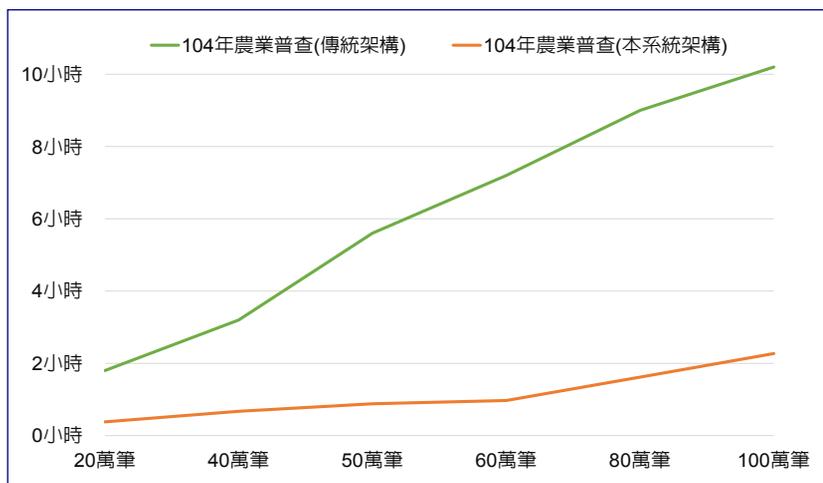
快速轉換統計資料應用，

五、運用新興技術，具有大量資料處理效用

資訊科技日新月異，適逢大數據、雲端運算技術與環境日趨成熟，帶動大量計算技術發展，本系統從基礎建設環境、系統架構至服務運行，皆突破固有作業框架，以分散式運算（Distributed Computing）、上載記憶體（In Memory）技術之高效能多工、分散處理機制，提升檢誤作業執行效率，充分運用環境資源達到優化系統效能。

本系統係運用 104 年農業

圖 4 系統實作服務驗證情況



資料來源：作者自行繪製。



普查資料實作驗證，確實大幅提升檢誤作業效率，如上頁圖 4 所示本系統架構相較於傳統架構提升約 5.5 倍平均執行工時。目前已完成 109 年人口普查、農業普查之試驗調查的資料處理階段，確認本系統具資料檢核工作正確性，將持續地提供正式 109 年人口普查、農業普查與 110 年工商普查之檢誤作業。

肆、系統整合應用成效

一、整合檢誤作業異質程序、達成系統共用效益

本系統推動異質調查之檢誤作業標準化，整合資訊應用促進資源共享成效，減少不同調查須重複建置與維運系統，並囊括多樣態統計調查資料，具符合三大國勢普查差異特性之複雜性需求，並具友善簡易資料檢核、自動化處理程序、操作設定再用性，便捷資料處理應用，達成行政效用與節省人力。

二、善用資訊科技、有效提升作業效能

因應統計調查作業繁雜及資料量龐大，本系統結合新興技術，就大量資料處理效能、關聯公務檔案統計、彈性格式轉換等需求，開發共通性模組、運用分散式運算及上載記憶體等技術，並設計直觀式便利操作介面，改善舊有資料流程，簡化業務程序，以提升檢誤正確性與時效。

伍、未來展望

鑒於統計調查環境愈趨多元，需求性愈來愈廣，本總處積極推展統計資訊服務，適當地結合業務、新資訊技術，創新統計資訊處理應用，並善用多年建置檢誤系統、共通性資訊系統累積經驗，將資訊化能力提升作業效率，以更具親和性與彈性設計介面，提供便利、簡化資料檢核流程，可於有限人力、時間成本，有效地提升檢誤作業績效。

未來，賡續精進共通性檢誤作業，持續整合各統計調查檢誤工作，積極推廣系統使用範圍，朝向跨機關系統橫向發展，讓統計調查使用者有標

準化之系統作業方式，期許全國統計人員，進行統計調查事務處理時，可透過資源共享、系統介接等協同作業，降低重複投入資訊系統建置與維運成本，進而節省公帑發揮共用資源效益，以切合統計業務發展趨勢。

參考文獻

1. 統計法（107年6月20日修正），行政院主計總處。
2. 公務統計方案實施要點（107年12月19日修正），行政院主計總處。
3. 各機關辦理指定統計調查總覽（檢索於109年7月3日），行政院主計總處。❖