

創新變革精進獎勵項目

農業大數據整合建置與應用

現今農業發展瓶頸多屬混合型問題，涉及層面包含生產面、貿易面、需求面、水資源運用、農地管理、農業補貼及永續發展。為適時提供跨領域農業政策擬定，增進資料聯結應用效能，行政院農業委員會統計室歷時 3 年進行異質性資訊系統整合，完成農業大數據建置。本文簡介農業大數據的建置流程，及跨界運用的實例。

行政院農業委員會統計室（王科長翠華、邱專員郁筑）

壹、農業資料種類

農業統計資料來源管道有農業管理或補助登記的相關公務類資料，還有傳統統計調查、感測器蒐集之航照稻作、產地批發及零售市場的農產品價量資訊、疫病防治的 GPS 追蹤資料、網路輿情及意向調查等（圖 1）。無論資料蒐集方式為何，理論上背後大多會建置一套資訊系統，經清查行政院農業委員會（以下簡稱本會）暨所屬機關至少約有 70 個農、漁、牧資料

圖 1 農業統計資料來源



資料來源：作者自行繪製。

系統，其中屬結構性資料或涉及政府統計者約有 38 個，為目前農業大數據整合之主要對象。

貳、農業大數據建置流程

38 個資訊系統代表背後有著 38 個不同團隊、不同邏輯及不同的程式設計，若欲進行跨域整合，必須先了解各單位原始資料建置格式與定義之歧異，重新規劃各資料庫相關定義、對應分類、匯入週期、匯入檢核條件等作業規範，並進行格式化、轉錄檢核條件設定及代碼標準化等過濾清洗程序：

一、背景資料整理

總體而言資料庫的數值資料比文字資料品質優、補貼資料比管理登記資料確度佳、新資料庫又比舊資料庫好。許多資訊系統的背景資料欄位內容常見亂碼、填寫不完整、空值或寫法不一致等情況（圖 2），加深不同系統間資料串連難度，因此本會統計室投入相當多的時間及心力逐一進行資料除錯、轉換及修正等背景資料碰檔、整理工作，以提升資料庫使用價值。

二、建立檢誤條件

本會統計室第一次轉錄灌溉農地之總面積為 92 萬公頃，明顯高於本會對外公布的 37 萬公頃，經業務單位再次檢核全

國 17 個農田水利會的原始資料後，發現其中一個水利會原始資料的統計單位與其他水利會不同，故形成如此結果。因此，即使為品質較佳的數值資料，在進行資料串接之前，仍須設立簡單的檢誤條件，避免造成後續相關政策分析之偏誤。

三、代碼標準化

由於各資訊系統開發廠商不同、開發時間不一，建立串連變數統計標準化或進行代碼對照，方能將資料予以串連整併。例如內政部的地段地號代碼有 6 碼及 7 碼兩種選擇，致本會各單位資訊系統亦出現兩者其一的呈現方式，因此必需建立轉換對照標準才能處理不同單位的土地資訊（下頁表 1）。另農糧署的資料庫也因補貼方式或用途不同，而有不同的作物代碼或中文名稱，也需先將代碼轉為相同後，才能進一步分析應用（下頁表 2）。

基礎資料經過過濾清洗後，

圖 2 背景資料概況

縣市代碼	農戶姓名	農戶身分證號	性別	出生年	教育程度代碼	農戶住址	住家電話
09007	陳超正	ax54195868	F	58	3	連江縣南竿鄉	083621111
09020	推婷	U215897653	F	40	6	金門縣金城鎮珠浦東路10巷1號	(08)2327999
10003	黃永遠		M	50		桃園縣桃園市中德里中正路100巷1號	03-3111987
10007	紀事本	C221254796	F	55		臺北市南海路37號	2312-6369
10008	倪一?	A111226998	M	49	2	台南市中正路7號	0953-677-111
10009	曾開心	E2223456+6	F	59	4	雲林縣斗六市八德里文化路633巷14號	0523-57900
10010	梅涼馨	F155666999	M	44	1	嘉義縣太保市太保里%%*@太保小段377、378號	05-3712222

說明：以上為擬真的虛構個資。
資料來源：作者自行整理。

怪符號

格式不一

創新變革精進獎勵項目



表 1 各系統地段代碼呈現方式

地段碼別	7 碼 縣市代碼 (1 碼) + 鄉鎮代碼 (2 碼) + 段代碼 (4 碼)	6 碼 地政事務所代碼 (2 碼) + 段代碼 (4 碼)
系統名稱	航照稻作檔 (農糧署)、 休耕轉作檔 (農糧署)、 小地主大佃農檔 (農糧署)、 敏感性作物檔 (農糧署)	全國農地地籍資料清冊 (資訊中心)、 農田水利會灌溉地籍清冊 (農水處)

資料來源：作者自行整理。

表 2 各系統作物代碼呈現方式

系統名稱	農情調查系統	休耕轉作系統
作物代碼、名稱	602 鳳梨 001 硬質玉米	602 鳳梨、016 契作鳳梨 112 飼料玉米

資料來源：作者自行整理。

圖 3 統計調查先期作業

【受查戶】						
農戶編號	調查姓名	電話	地址	生日	原層別	流水號
00000001	徐太宇	08-7777111	屏東縣曾正鄉美麗村8鄰真心路528號	310101	23	4153

【戶籍檔】			【福利相關檔】			
姓名	生日	關係	姓名	農保	老農津貼	子女教育補助
徐太宇	380101	戶長	徐太宇	Y	Y	
林真心	440303	妻	林真心	Y		
徐偌宣	600520	長女	徐偌宣	Y		
徐至模	821010	長孫	徐至模	Y		Y

【作物資訊】			【畜禽資訊】		【農業補助收入】			
作物名稱	面積 (公頃)	收購金額 (萬元)	畜禽細品項	在養數量 (頭、隻)	補助金額 (元)	天然災害	停灌補助	禽流感補償金
稻穀	1.01	14.0	1 種公豬	2	4,320	4,320	39,000	N
食用玉米	0.05		2 種母豬	52				
			3 哺乳小豬	92				

資料來源：作者自行繪製。

我們再依資料庫品質高低作為整合順序，每年匯入約 2 千萬筆、250GB 之農業基礎資料，完成本會農耕類、畜牧類及福利類等農業生產巨量資料庫之建置。

參、分析應用案例

一、支援統計調查先期作業

近年隨隱私權高漲，調查資料蒐集困難，透過農業大數據整合資訊，有效運用於農家所得調查、畜情調查及畜產品生產成本調查，減少調查問項，並列出受查戶近 2 年公務資料總歸戶，作為調查員調查前參考文件，不僅提升本會農業調查之資料確度，還可縮短實地訪查時間，減少受查者負擔，並降低受查戶懷疑詐騙集團疑慮，提升統計調查效能（圖 3）。此外，我們也建立調查結果回饋機制，針對農民關心的價量、出口及百萬農業所得收入的作物資訊，利用本會最新

調查結果及價量預測，編製簡易統計圖表，回饋農戶即時資訊，提升政府與農戶良性互動關係。

二、支援政策論述

103 年下半年以來水情不佳，致 104 年一期稻作部分區域辦理停灌休耕作業，並進行相關補償措施。為評估停灌區域依休耕或轉作提供之各項補償方案之合理性，經運用停灌補償區域清冊資料與農耕類資料庫相互碰檔，觀察請領停灌補助者之 103 年生產行為，提供本會農水處及農糧署擬定種植綠肥、轉作契作作物等六大停灌補償標準，減少農民受停灌影響之經濟損失。

三、協助政策規劃

農業施政主要圍繞人、地、物、水資源 4 大面向，透過人、地、作物及水資源資料串聯，可配合機關施政需求，適時提供整合碰檔結果，協助農業政策規劃。如農業灌溉區

域內、外產值專題研究，透過地段地號串連 9 個公務資料庫，進行交叉比對，了解灌區內農地之種植作物，並推算灌區內、外農作物產值，協助未來農業用水運用效能的提升。另配合綠能農業，結合地理資訊系統，進行沼氣發電示範之優先地理位置及推估成效之研究，審視目前 2 千頭以上養豬場群聚分布，選定彰化縣、雲林縣及屏東縣特定鄉鎮具規模養豬場列為沼氣發電優先輔導示範區，做為施政規劃之依據。

肆、結語

因應大數據及物聯網的時代，我們經過過去 3 年的努力，完成異質性資訊系統整合，進而加以運用，提升統計資料跨界運用價值。105 年我們以農業大數據為基礎，參考荷蘭統計局的架構，建置 9 大領域淺顯易懂、敘事式的互動式關聯統計圖表，並於本（106）年 2 月正式上線以提供更多面向的統計服務。未來我們將配合休

閒農業與農產業群聚規劃，以地理資訊系統等展示技術，建置重要農產品生產地理資訊分布，期藉由地理資訊之空間分布所呈現資訊，持續協助農業施政。❖