

發展主計資料倉儲系統 推動資料治理與應用

行政院主計總處（以下簡稱主計總處）為推動資料治理，規劃建置主計資料倉儲系統，整合跨部門及跨系統數據，提供分析工具和自助平臺。透過資料盤點與定義，建構資料傳輸與加密、資料提供與軌跡留存機制，進行妥適保護與管理並提供加值分析使用，奠定智慧主計服務之基石。

謝明憲（行政院主計總處主計資訊處高級分析師）

壹、前言

近年來，各級政府機關均已著手推動資料之治理與加值應用，以保護資料，提升使用者體驗，並加速組織決策制定。為推動資料治理，主計總處規劃建置主計資料倉儲系統，落實資料盤點及分級，建構資料加密保護機制，及資料取用權限控管與存取軌跡紀錄等整體性規劃，確保機敏資料安全使用，並整合跨部門及跨系統數據，提高資料存取、查詢與分析之效率，提供分析工具和自

助平臺，支援決策分析。

貳、資料治理發展趨勢

由於巨量資料分析資訊技術蓬勃發展，公共政策產生許多實務應用案例，例如新冠肺炎（COVID-19）期間，我國口罩預購政策除前往實體藥局購買外，亦可使用自然人憑證或健保卡線上訂購；另口罩地圖係由政府釋出資料，民間專家提供視覺化平台服務；口罩實名制上線後，政府也針對銷售資料進行分析，藉由瞭解藥局周邊人口結構，微調口罩配

送量。這些案例展現結合跨部門資料介接、管理途徑與政策框架的巨量資料分析在政府變革的應用，背後須運作得宜的資料治理機制方能順利運作，同時體現資料驅動決策、公私協作新趨勢。

國家發展委員會（現移由數位發展部）2020 年提出「服務型智慧政府 2.0 推動計畫（110 年至 114 年）」，以「擁抱數位未來，打造開放與創新的智慧政府」為願景，強調政府必須迎接數位時代，以數位化方式營運政府各項業務，透

過資料法制化、塑造資料生態友善環境、建立需求導向之資料分析決策模式、深化新興科技應用之智慧服務等策略來落實智慧治理。其規劃核心理念為「以資料為骨幹，建構下一世代的公私協力治理模式，提高政府與民衆的信賴」，而資料治理實為推動智慧政府之核心引擎。

參、資料倉儲系統概述

「資料倉儲」(Data Warehouse, DW) 係由資料倉儲之父 Bill Inmon 於 1990 年提出，在「What is a Data Warehouse」書中舉出資料倉儲的資料蒐集有 4 種特性，根據這些特性使資料倉儲能將資料提供給決策管理系統進行處理：

- 一、主題導向 (subject-oriented)：集中與特定主題相關的資料，不只是目前營運資料。
- 二、經過整合 (integrated)：存放資料從不同來源合併，且維持一貫性。
- 三、依循時間變動 (time-variant)：可依特定時間點辨識儲存的資料。
- 四、不會流失 (non-volatile)：

資料隨時間累積增加，不會被移除，得以進行持續性觀察。

資料倉儲按資料範圍，可分成企業資料倉儲 (Enterprise Data Warehouse, EDW) 和資料超市 (Data Mart)。EDW 係包含整個企業的資料，由數個主題組成，例如客戶、產品、業務等面向，能用在決策支援，包括即時型資料及彙總過的資料。Data Mart 與 EDW 大致相同，惟 EDW 涵蓋整個公司的資料與人員，而 Data Mart 只包含特定範圍的資料，且使用者會鎖定特定工作群組的人員。

無論 EDW 或 Data Mart，都可作為資料探勘 (Data Mining) 和線上分析處理 (Online analytical processing, OLAP)¹ 等分析工具的資料來源，提供組織分析有價值的資訊，以利決策擬定及快速回應外在環境變動。

肆、主計資料倉儲系統規劃與建置

主計總處為推動資料治理與加值應用，規劃建置主計

資料倉儲系統 (下頁圖 1)，作業流程為先盤點各項資料來源，再經轉檔、加密、檢誤，建立各項資料之詮釋資料 (Meta Data)，完成資料字典，資料內容則經清理後匯入資料倉儲資料庫。存放於資料倉儲之資料，則可透過各種軟體程式或 Tableau、Power BI 等視覺化工具建構分析模型，或下載後供進一步分析應用，其重點分述如下：

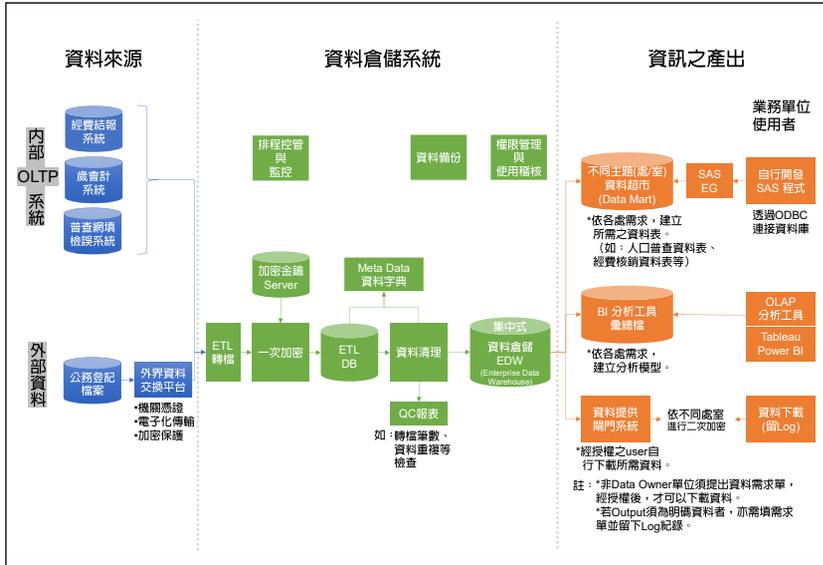
一、以「資料來源」範疇，進行資料盤點與定義

為推動資料治理，首先組織須進行資料盤點，並視資料應用需求，進行資料萃取及轉換，以及資料儲存結構及格式等設計。另為利資料加值與應用，必須提供資料使用者簡單易用且標準的資料定義，包含資料檔及欄位的說明等，並且持續性維護資料字典的正確性 (下頁圖 2)。

因此資料盤點與定義是很重要的起步，主計總處在系統建立初期，必須剖析資料來源，包括內部線上交易處理 (Online transaction processing, OLTP)²

專題

圖 1 主計資料倉儲系統



資料來源：主計總處主計資訊處。

圖 2 主計資料倉儲系統－資料字典

The screenshot shows the '資料字典' (Data Dictionary) interface. At the top, it says '行政院主計總處 資料字典'. Below the header, there are navigation tabs: A.資料字典查詢, B.資料字典維護, C.上場作業, D.資料匯出, Y.歸主目錄, Z.設定.

The main content area shows a search for 'WD1006 檔案資料'. The search results are displayed in a table format:

系統類別:	彙編資料	檔案名稱:	DVD_ADDR	中文名稱:	正規化郵政地址檔
資料性質:	DVD-彙編彙編彙編	檔案類別:	View	輸入頻率:	年
管理單位:	郵政管理處資料發展科	管理單位:	郵政管理處資料發展科	擁有單位一:	郵政管理處資料發展科
擁有單位二:		擁有單位二:		檔案資料內容:	戶籍檔、村里戶籍檔之來源庫中的地址
備註:	永久	資料倉庫管理單位:	年	資料提供:	~ 20240119
資料類別:	內政部				

資料來源：主計總處主計資訊處。

系統如經費結報系統、歲會計系統及普查檢誤系統等，另亦蒐集主計總處業務需用來自各政府機關的外部資訊等，決定收納資料類型與範圍，以及相關前端系統的資料類型與取得方式，制定一致化的資料清理標準，去除重複資料並決定資料更新機制與頻率，透過資料字典，提供業務單位確認資料倉儲收納資料情形。

二、以「資料倉儲」機制，鏈結資料傳輸與加密保護

主計總處在外部資料來源部分，過去多以光碟、可攜式儲存媒體或電子郵件方式取得，倉儲系統規劃透過跨機關間資訊單位的技術協作，調整為由來源機關提供線上傳輸，並搭配傳輸及資料雙重加密機制，以確保傳輸資料的安全。

經過流程再造，主計總處的操作人員自外部取得資料時，已將原須先儲存於個人電腦或其他儲存媒體，再上傳至資料倉儲平台的程序簡化，改由直接透過系統與來源機關接之線上傳輸機制取得資料，

解密後立刻再依倉儲系統加密金鑰，針對可識別欄位（如身分證字號）進行一次加密，確保資料妥善保護。而內部線上交易處理 OLTP 系統則循前述過程自動化進行資料的擷取轉換載入（extract, transform, load, ETL）³ 轉檔程序，經一次加密後轉入倉儲系統。由 ETL 程式轉檔及一次加密程序後存入 ETL DB。再結合資料字典的定義進行資料清理，並進入資料倉儲系統（EDW）。

ETL 程式由排程控管模組設定並監控執行，且於資訊產出予使用單位時，透過權限管理與使用稽核模組控管資料合法使用及軌跡紀錄的保存。

因此，有了強化傳輸安全性及機敏資料加密保護，除確保機敏資料安全外，亦有助於創造各機關間公務資料相互加值與合作的新契機，有效推升資料應用價值。

三、以「資訊產出」管理，滿足資料提供與軌跡留存需求

經盤整業務需要，規劃主計資料倉儲系統資料提供模式

三種，供業務單位選擇使用：

- (一) 透過開放資料庫互連（Open Database Connectivity, ODBC）介面，以軟體程式連接資料庫讀取資料。由於主計總處現行多採 SAS 程式處理資料，將預先依各使用資料單位需求建立資料超市，供特定領域分析使用。
- (二) 運用線上分析處理 OLAP 工具（如 Tableau、Power BI 等）存取資料彙總資料，並進行視覺化設計，提供視覺化圖表及下鑽

功能（Drilldown），彈性支援決策管理。

- (三) 使用資料提供閘門系統，經授權之使用者可申請下載資料檔案後，自行運用資料及分析（圖 3）。

無論透過任一模式，權限管理及使用稽核至為關鍵，使用資料均須經由主計總處電子表單申請程序，並敘明使用期限，經業務單位主管核准後，資料使用者才能撰寫程式讀取資料分析，或自行下載資料使用。其中取得的資料識別欄位均經加密處理，可進行勾稽分

圖 3 主計資料倉儲系統－資料提供閘門



資料來源：行政院主計總處主計資訊處。

