

環境資料及三維圖台之雲端整合應用

施奕良 常若愚 蔡坤龍 林立青

財團法人國家實驗研究院國家高速網路與計算中心

摘要

為因應公部門網路服務及運算雲端基礎設施之強韌，國研院國家高速網路與計算中心正持續建置三維空間資訊雲端基礎共構環境，作為公部門與學研間跨域數據治理與三維圖資呈現技術開發。此外，透過資料匯流與加速取用機制建構資料聯盟環境，期能擴展空間網格數據的共享。運用標準協定加速資料的傳輸、管理及授權使用，使環境感測資料或跨域分析之加值數據能達到永續儲存目標。以國內自主研發之圖台架構提供環境資料的導入及開發，使圖台可視化與分析模組能提供支援公部門之決策所需。因此有別於公部門各自建構之資料庫或展示系統，透過基於國網中心雲端整合架構，空間資料彙整並遵循OGC開放空間資料標準，能快速透過雲端圖台發佈成果，節省各自維護資料、產品重構等效能問題。另外，管理系統的授權機制更可保障數據使用的安全及可靠性。最後達到公部門各式服務的共構與共享，更有利推動公部門資料的公開使用以及擴大與學研的資料合作與開發。

關鍵字：雲端整合平台、資料聯盟、三維空間資訊共構

一、前言

資料收集目前正處於一個操作簡單及低成本的時代，在空間或時間上，大規模物理環境數據的收集將更能促進我們對地球環境的了解。若要從數據獲得觀察或分析個結果，則需要進行資料調查、收集、統一資料格式及資料組合等前期準備工作，以達到將文件轉換成數值檔案，進一步讓資料成為具可分析的資訊(Analysis-ready data, Wulder et al, 2016)。而地球環境的資料正以像素化或點雲(point cloud)的方式形成大數據的一環，特別資料性質包含時間與空間的網格(grid)資料，亦可稱為資料方塊(Datacubes)。從一維(t)的時間序列資料、到二維(x/y)的影像，甚至三維(x/y/t)或四維(x/y/z/t)包含時空分布的影像或是數值模式的數據，皆能被包含在資料方塊中，該名稱之由來為網格資料的時空連續性如方塊體的架構。

氣候觀測或數值模式資料的時空性質，及數據量的增長快速使得數據庫及數據的儲存方式需要更具彈性及效率的整合方案，才能加速資料轉換成ARD的過程，並促進數據分析或用的快速發展，讓使用者更專注於後續的研究與開發。然而數據庫若缺乏一致的標準，就有可能阻礙利用資料方塊進行氣候服務(Climate services, Hewitt et al., 2012)的危害。上述之標準包含資料庫間的相互操作、資料間的統一格式及資料傳輸間的共同服務標準等。並透

過「開放地理資訊系統協會」(Open GIS Consortium, OGC)所建立的各項標準促成上述資料庫間資料的流動、操作、共享等多目標應用。

基於數據前置處理與後端資料庫的共通連結與分享性，國家實驗研究院國家高速網路與計算中心(國網中心)提出一個雲端整合平台，針對地球環境數據的進行可於各資料庫間廣泛搜尋之連結操作，讓資料提供者與使用者間聯繫更加緊密。主要目標一為讓資料提供者獲得資料如何被使用及使用之目的，另一為讓資料使用者能夠有效及迅速的收集資料，讓資料具有統一的標準更有利於後續的科研應用。此目的更能讓資料提供者因應使用者的需求，而更精進其觀測或模式的改善等，使其更符合其科研使用或數據服務等目的。

除此之外，國網中心之整合平台為了連結公部門的環境數據庫與各公務機關及機構的使用者，亦將推展資料同盟的共構策略，以促進和科研機構之合作與開發目的。

二、雲端整合平台之架構

國網中心以OGC標準化之資料方塊系統(Rasdaman, raster data manager)(Baumann, 1992)為基礎建立資料儲存及分享環境，並以多維度的資

料存取方式進行操作，示如圖1。其自1989年發展至今，已獲EarthServer及ECMWF等重要歐盟機構採用，其可支援各個資料方塊間對該格式資料的大量查詢，使得資料方塊中有相同性質的資料，可被有效的檢索與存取，因此形成平行化的資料同盟架構，使其具彈性的檔案操作便捷性，示如圖2。其擴展性及快速地檢索與資料運算之優點更能提供未來於科學、工程及其他領域中巨量資料環境的使用(Baumann et al., 2021)。

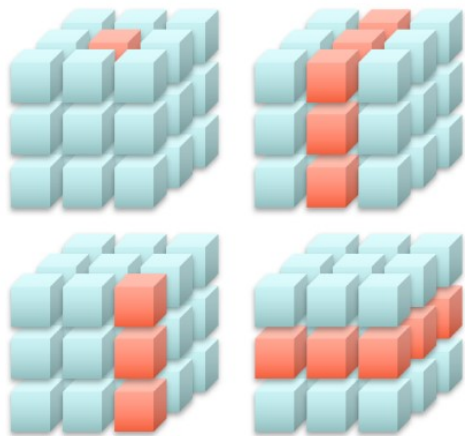


圖1 多維度資料存取方式

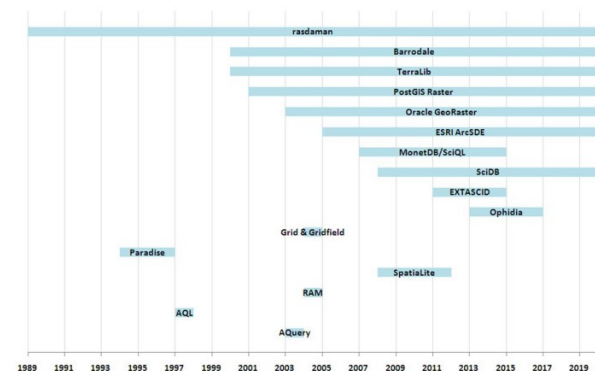


圖2 多維度資料庫(方塊)之發展(引用自Baumann et al., 2021)

以複數資料方塊的使用環境為例，當使用者欲計算影像植被指數(Normalized Difference Vegetation Index; NDVI)之最大值差異時，使用者節點並不擁有節點A及節點B之資料，如圖3所示。計算者的節點可同時傳送計算需求至上述兩個節點位置，並由各資料節點分別計算後傳送結果回使用者節點，並讓使用者直接計算A及B節點之差異值。

上述的使用情境揭示，資料分析端僅需發出執行命令，即可讓複數個資料儲存端進行平行計算後再回傳所需數值，以符合資料分析端的需求。而這樣技術條件可將各式環境巨量之資料方塊，拆分成

數個資料方塊的形式，而無損其資料連結與存取，其優點更能節省使用者端計算或檢索效率。有此可知，藉由系統標準化後，更能促進使用者廣泛具有相同資料性質的資料方塊，期能有效且深入的分析與應用，並產生較全面的結果而非單一時空範圍的見解。

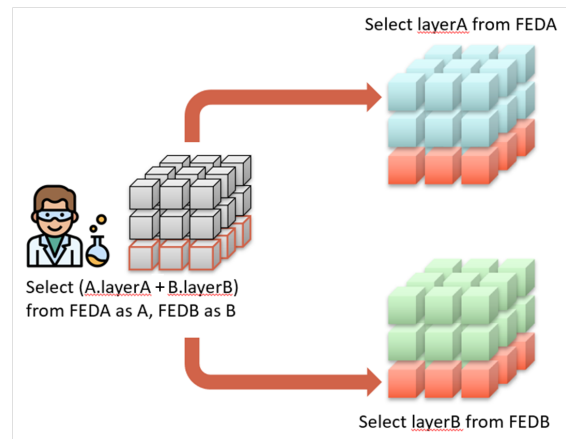


圖3 資料方塊間節點之分佈與檢索及存取

國網中心以標準化之資料方塊系統之基礎建構之資料服務平台(Data as a service)如圖4所示，其中OGC標準分為資料使用NETCDF、GRIB及TIFF格式，以及網頁標準通訊協定服務(如Web Map Service(WMS), Web Coverage Service(WCS)及Web Coverage Processing Service(WCPS)等)。該平台上之運作流程如下六項：

1. 具有上述格式之環境資料經由WCS匯入資料方塊中，轉為地理內涵(geographic coverage)，包括幾何位置(geometry)和屬性值(attribute values)(李等人2006)。
2. 使用者端藉由WMS的地理資料傳輸服務，讓使用者端向資料提供端發送資料請求，可即時產生地理資料影像並回傳給使用者端。因此，上述兩資料立方以共通的標準語言進行一致性的溝通，建立地理資料圖像化的流通與共享。
3. 若資料提供者端允許使用者下載其檢索的時空範圍，亦可以WCPS進行資料擷取、處理及分析功能。圖2之檢索及存取需求亦以上述標準，讓使用者端獲得分析所得之資訊。
4. 為了促進各公部門之資料流動，國網中心於此架構中建立授權機制，讓公務機關或部門可透過資料使用者授權，進行敏感性及限制性資料的使用，並以此保護資料安全。並讓資料提供及使用雙方皆可依據授權保護機制，讓機制工具成為資料的使用助力。

- 此外，亦可藉由架構於國網中心雲平台之三維空間資訊系統，進行資料可視化與分析模組擴充，讓公部門能將其資料立方之成果進行網頁發布、行動軟體的多元應用及提供應用程式介面進行資料的公開使用等功能。
- 上述1至5項內容可針對複數資料立方建構資料同盟目標，使資料的檢索或擷取的時空範圍能更加廣泛，提升科研成果的可靠性及應用性。



圖4 雲端整合平台架構圖

三、資料同盟之共構策略

由於Rasdaman系統具有連結資料節點，在任何節點上皆可接收查詢檢索，傳送並組合其運算結果至查詢節點端。因此Rasdaman系統更容易佈署在公部門的資料節點，以適應大量及快速之資料操作，並提高系統的性能與可擴展性。因此在資料庫的內部連結操作及資料交換上，各資料庫如同盟般在共同原則及框架上分享資料，並對任一同盟節點提供來自於不同節點資料庫之查詢及檢索服務。

在此資料同盟的策略下，歐盟資料方塊同盟(The European Datacube Federation)於2020年成立，同盟中CODE-DE(<https://code-de.org/>)為其重要之夥伴，其為Sentinel衛星系統及數據產品之重要入口。同盟之成員皆可通過其資料方塊共享架構分享其衛星資料，並支持聯合國永續發展目標與區域治理於氣候變遷、防災及氣候服務等各領域的發展(Baumann and Team, 2020)，使用者亦無須任何程式編碼的工作，即可透過與同盟的節點以OGC統一標準格式獲取數據。

因此國網中心參考上述Rasdaman系統及其資料方塊同盟的架構，建構公部門之資料同盟資料方塊。政府自2015年起訂定了相關資料開放及使用法規，並成立政府資料開放平臺(<https://data.gov.tw/>)，至今政府開放資料超過4萬項資料集，下載量超過1千萬人次。全球開放資料排名，台灣蟬聯全球第一，其中「空氣品質」、「地理圖資」、「天氣預報」等地球環境資料則尚未完

全符合外界期望(陳梅英，2017)。然而資料開放與資料管理面臨以下兩點問題：

- 標準共通格式不足：資料開放平臺中的資料格式以CSV格式為最多，XML格式的資料次之。網頁標準資料(如RDF、OWL)則小於1%。資料結構化仍需藉由共通的標準加以提升(許志義等人，2019)。
- 資料管理不透明：因法規限制如涉及隱私、機密或國土安全等，加上政府內部的監督管理並無適當的工具去規範使用者之目的。加上政府內部自行認定資料的公益性質，亦無法管控使用者之使用目的，以致產生可能的業務訪查問題，因而大部選擇不開放相關資料(郭琇真，2020)。

同盟架構以公部門為對象之優勢可適度的解決資料法規之障礙，雲端整合平台對各資料方塊具有授權使用工具，將可以供透明的管理方式(如記錄資料方塊之資訊流動歷程、限定時間及使用範圍之WCS授權啟動)，讓資料提供與使用雙方在此工具的基礎下，進行資料共享使用。

再者，雲端整合平台則是在OGC統一標準格式中進行各資料方塊之共享服務，經由資料結構化後，可促進資料的多元應用並藉此激發出不同的創意，進行創造多重效益。

圖5為資料同盟情境使用，同盟資料方塊中有氣象單位產製降雨產品，因防救災之需求水利單位透過資料方塊取得歷史或即時之降雨資訊做為淹水模式之輸入值，進行獲得歷史或即時之淹水深度值。防災管理單位於同盟中將可獲得降雨及淹水兩種圖層資訊，以管理淹水災害之使用範圍，如NDVI植被指數或農政單位之種植使用區分等資料。上述三種資料將可利用行地理資訊演算法，計算災害影響區域的受災程度或社會經濟之產值影響等數值。

不管使用歷史資料或即時資料，於同盟中將能加速的促進資料流動過程，並產生增值應用資料圖層。在這樣的循環及持續的資料流動，基礎資料(如氣象單位之環境資訊)將可被多次的利用，轉化成為不同目的使用的資訊，新的資料因此產生多重的價值。與現有的政府資料開放平臺相比，整合平台之資料/資訊的流動將更有其目的性，資料提供者更能夠後續分析其資料之被使用範圍與目的，以滿足使用者之需求流向，或藉此流向進而產生多元的數據項目及種類以促進使用意願。使用者亦能依上述之流程，回饋公部門其加價值之產品，即以相同的OGC標準匯入至同盟資料方塊中，並讓資料提供者了解數據使用的限制與需求，如圖6所示。

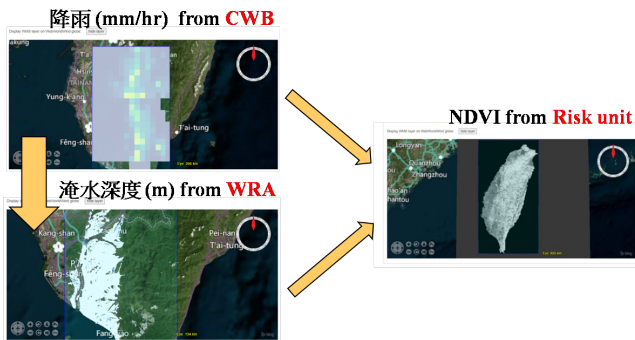


圖5 資料同盟共構情境

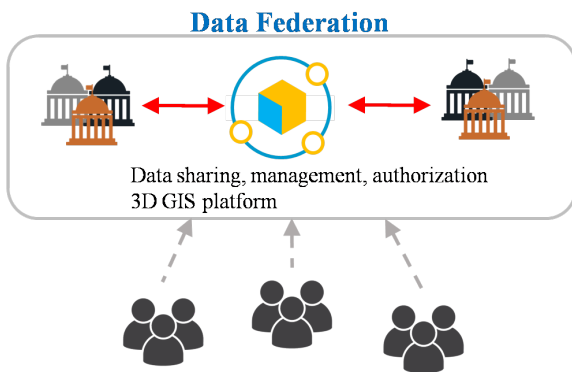


圖6 科研機構與公部門之資料同盟

四、結論

同盟資料之循環最重要的目的將是促使資料不僅是提供之單一流向，而是創造資料的多目的流向，進而有機會產生創新加值之應用數據。然而此一流程有賴於科研機構與公部門共同合作，吸引具資料轉化並符合公部門需求應用之科研成果，提升公部門政策執行(如防災決策、國土安全及社會福利等)的重要依據。

而科研機構亦可投過公部門的整合雲端平臺之資料同盟架構，有效的檢索與取得應用研究所需的所有資訊，即同盟環境下公部門資料授權且安全的存取資料，而無須逐一向各部門索取資訊。而同盟中各種類資料將以GIS圖層的呈現，讓使用者進行不同時空需求之檢索，以擷取在此需求下之所有資訊。利用資料同盟的方式，提升資料檢索與取得的效能。

此舉與歐盟資料方塊同盟所述求之資料共享目的相同，其差異僅在公部門實施此一資料流動過程。未來政府資料法規更能藉由此一網頁標準化工具，對民眾公開相關數據。讓網頁共同格式成為公部門與一般使用者間的橋梁，提升開放資料之發展。

參考文獻

李德財, 林峰田, 張欽隆, 張藝鴻, 莊庭瑞, 鄧東波, 2006:“都市地理資訊共享機制之建立以台北市經驗為例”, 第八屆海峽兩岸城市地理資訊系統學術論壇, 香港中文大學, 330-341.

許志義, 王筑著, 柳育林, 許懷元, 2019:“政府資料開放與資料管理”, 公共行政學報, 56, 131-162.

郭琇真, 2020: 官民合作! 開放資料平台再進化 政府不能籠統回, 聯合新聞網, 2020年01月14日, 資料引自 https://vision.udn.com/vision/story/12425/4287493?from=udn-relatednews_ch1010

陳梅英, 2017: 全球開放資料評比台灣蟬聯全球第一, 自由時報電子報, 2017年07月20日, 取自: <http://news.ltn.com.tw/news/politics/breakingnews/2102658>

Baumann, P., 1992:” Language Support for Raster Image Manipulation in Databases;”, Proc. Int. Workshop on Graphics Modeling, Visualization in Science & Technology. Darmstadt, Germany.

Baumann, P., Dumitru, A., Mercariu, V., Misev, D., & Rusu, M., 2013:” Breaking the big data barrier by enhancing on-board sensor flexibility”, In Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, 32-36.

Baumann, P., Misev, D., Mercariu, V., & Huu, B. P., 2021:” Array databases: concepts, standards, implementations”, Journal of Big Data, 8(1), 1-61.

Baumann, P., & Team, R., 2020:” The EarthServer Global Datacube Federation”, IGARSS 2020 IEEE International Geoscience and Remote Sensing Symposium, 3192-3194.

Hewitt, C., Mason, S. & Walland, D., 2012:” The global framework for climate services”, Nat. Clim. Change 2, 831-832.

Wulder, M., et al., 2016:” The global Landsat archive: status, consolidation, and direction”, Rem. Sens. Environ, 185, 271-283.

Integrated Cloud Services of Environment Data and 3D GIS Platform

Shih Y.-L., Chang J.-Y., Tsai K.-L., Lin L.-C.

National Center for High-Performance Computing, National Applied Research Laboratories

Abstract

In response to the content of the Forward-looking Infrastructure Development Program, the government intends to strengthen public sector network services and cloud computing infrastructure, the National Center for High-Performance Computing (NCHC) has built a cloud integration platform composed of cross-domain data management and visualization service. In order to expand the provision of key services in the public sector, the management system of data collection and storage, and a 3D GIS platform are constructed. In addition, this integration platform follows data standardization to boost the transmission, management, and authorized use of data so that environmental data or value-added data can achieve the goal of sustainable storage. With the self-development framework of 3D GIS, the cloud integration platform will provide support of visualization and analysis modules for the decision-making needs. In contrast to the current database and/or demonstration system constructed by each public sector, our platform provides an integrated solution that allows the data management system to directly transfer data to GIS layers by OGC geospatial web standards. This procedure can increase the performance of data management, publish and display. Besides, the authorization mechanism of the management system can ensure the safety and reliability of data usage. Finally, the cloud integration platform can achieve the co-construction and sharing of various services in the public sector, which is more conducive to promoting public data sharing and expanding cooperation and development with the private sector.

Key words: Integrated Cloud Services, Data management, 3D GIS