

生物多樣性資料整合 的共同願景

從TaiBIF平臺到資料庫共通查詢系統

撰文 | 劉璟儀 (TaiBIF 內容經理 / GBIF 亞洲諮詢顧問 / 通訊作者)
端木茂甯 (TaiBIF 執行長 / 中研院生物多樣性研究中心助理研究員)
李思賢 (TaiBIF 暨中研院生物多樣性研究博物館研發工程師)
張俊怡 (TaiBIF 產品經理)
龔明哲 (TaiBIF 研發工程師)
李金穎 (TaiBIF 研發工程師)
蔡思怡 (TaiBIF 資料管理)

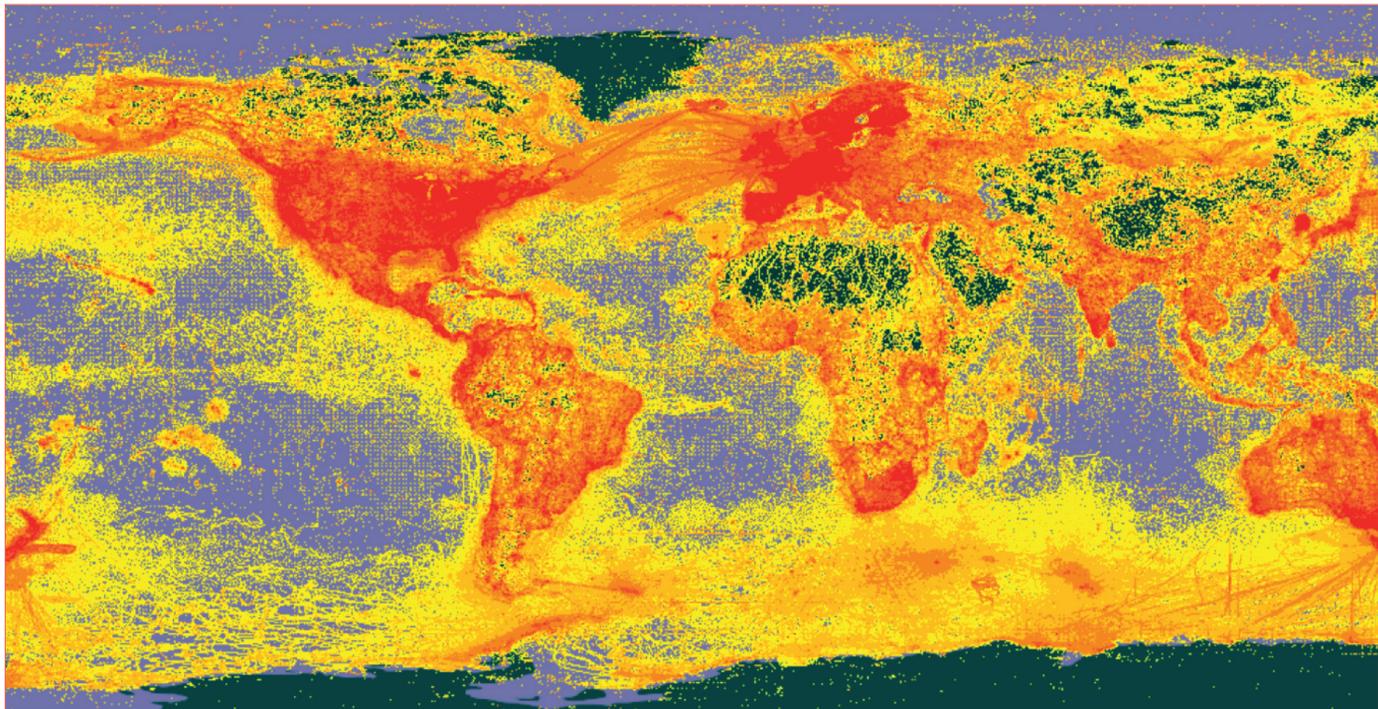
也許你知道臺灣的生物多樣性豐富度在亞洲名列前茅，那你知道臺灣也擁有亞洲第二大的生物多樣性開放資料平臺嗎？就讓我們來說說究竟過去如何建立起TaiBIF這樣規模的開放資料平臺，未來又將如何更進一步地發展吧！

全球生物多樣性資訊的崛起

我們很早就知道，幾世紀以來人類的經濟發展造成了全球環境及生態的衝擊，雖然時至今日，這些衝擊仍未止息，但其實自1992年開始，世界各國便意識到人類的永

續發展和生物多樣性的保育工作刻不容緩，因此於地球高峰會上正式簽署了生物多樣性公約 (Convention on Biological Diversity, CBD)，期望透過近兩百個締約國制訂國家層級的保育政策，以共同行動來保育全球的生物多樣性及永續利用自然資源。

然而，光是靠這樣的理念和熱忱是遠遠不夠的，如要制訂出更有效且全面的保育策略，便需要足夠的科學數據來評估及支持。於是，在那個還未有成熟的大數據概念的時代，聯合國經濟合作暨發展組織 (OECD) 下的生物多樣性資訊小組即建議：「國際間需要建立可以自由存取生物多樣性資料、



全球生物多樣性資訊資料分布圖。© OpenStreetMap contributors, © OpenMapTiles, GBIF

資訊的機制，藉由提供可靠正確的科學證據，可為經濟、社會帶來益處並促進永續發展。」(GBIF.org)，全球生物多樣性資訊機構 (Global Biodiversity Information Facility, GBIF) 便背負著如此的期望，在2001年正式誕生，並由多國政府共同簽署合作備忘錄 (MOU) 加入成為會員並設立節點 (Nodes)，將「提供大眾無論何時何地均能公開且自由存取有關地球上各種生物的資料。」的宗旨奉為圭臬。

全球生物多樣性資訊的開放與流通，憑藉著GBIF網絡中各會員國及相關組織的努力，在廿年來有極為顯著的進展，不僅累積

了龐大的資訊基礎，更反映出全球的生物多樣性變化趨勢和指引政策決策方向，也讓生物多樣性資料的應用發揮更大的研究價值，成為全球最全面、最開放、最具指標性的生物多樣性資訊存取共享平臺。截至2021年底，已有超過1,700個資料發布者，將約64,000個資料集 (Datasets)、超過19億筆的物種出現紀錄資料 (Occurrence)，透過GBIF的資料平臺開放給全球的資料使用者，而這些開放資料也化作一個一個物種的點位，逐漸匯集成一幅全球的生物多樣性地圖。

資料生命週期



傳統及理想的資料生命週期及管理流程。©TaiBIF 臺灣生物多樣性資訊機構

► Data availability in Taiwan



Data availability

Total data available for selected taxonomic groups in Taiwan



Mammals = Class Mammalia
Birds = Class Aves
Bony fish = Superclass Osteichthyes
Amphibians = Class Amphibia

Insects = Class Insecta
Reptiles = Class Reptilia
Molluscs = Phylum Mollusca
Arachnids = Class Arachnida

Flowering plants = Phylum Magnoliophyta
Gymnosperms = Superclass Gymnospermae

Ferns = Phylum Pteridophyta
Mosses = Phylum Bryophyta
Sac fungi = Phylum Ascomycota
Basidiomycota = Phylum Basidiomycota

截至2022年1月，臺灣發布至 GBIF 平臺的資料類群分布概況。©GBIF 全球生物多樣性資訊機構

TaiBIF成立的契機

過去臺灣生物多樣性研究領域風氣保守，研究觀念也較為傳統，少有開放資料的概念，多數的研究人員往往將自己的研究資料存放於自己的電腦及硬碟中，若未有任何國內或國際期刊的研究發表，那些原始數據便會永遠鎖在電腦之中，不見天日；大眾及政府部門也不易了解生物多樣性領域的發展及具體成果，甚至視學術領域為象牙塔，不利於公眾及媒體溝通，也讓社會經濟發展與生物多樣性保育間的衝突不時上演，最後犧牲的總是生態環境。這樣的困境讓一些相關部門及專家學者開始意識到，唯有整合國內生物多樣性資訊，建立一個能即時且完整取得生物多樣性資料的平臺，才有機會提供足夠的資訊，讓開發與生態達成雙贏的局面。

因此，臺灣在GBIF於2001年成立之初，

便率先由當時的國家科學委員會（國科會）與其簽署MOU成為初始會員，並於中央研究院成立「臺灣生物多樣性資訊機構」（Taiwan Biodiversity Information Facility, TaiBIF），不僅為推動國內生物多樣性資料整合及流通的單位，更是臺灣與國際生物多樣性資訊連結的重要窗口。除了推展GBIF的策略目標，引進國際共通的資料標準與工具、資料授權與開放的概念，更建立了生物多樣性資訊基礎建設，提供可以同時自由開放資料及利用生物多樣性資料的入口網站，串連起理想的生物多樣性資料生命週期。

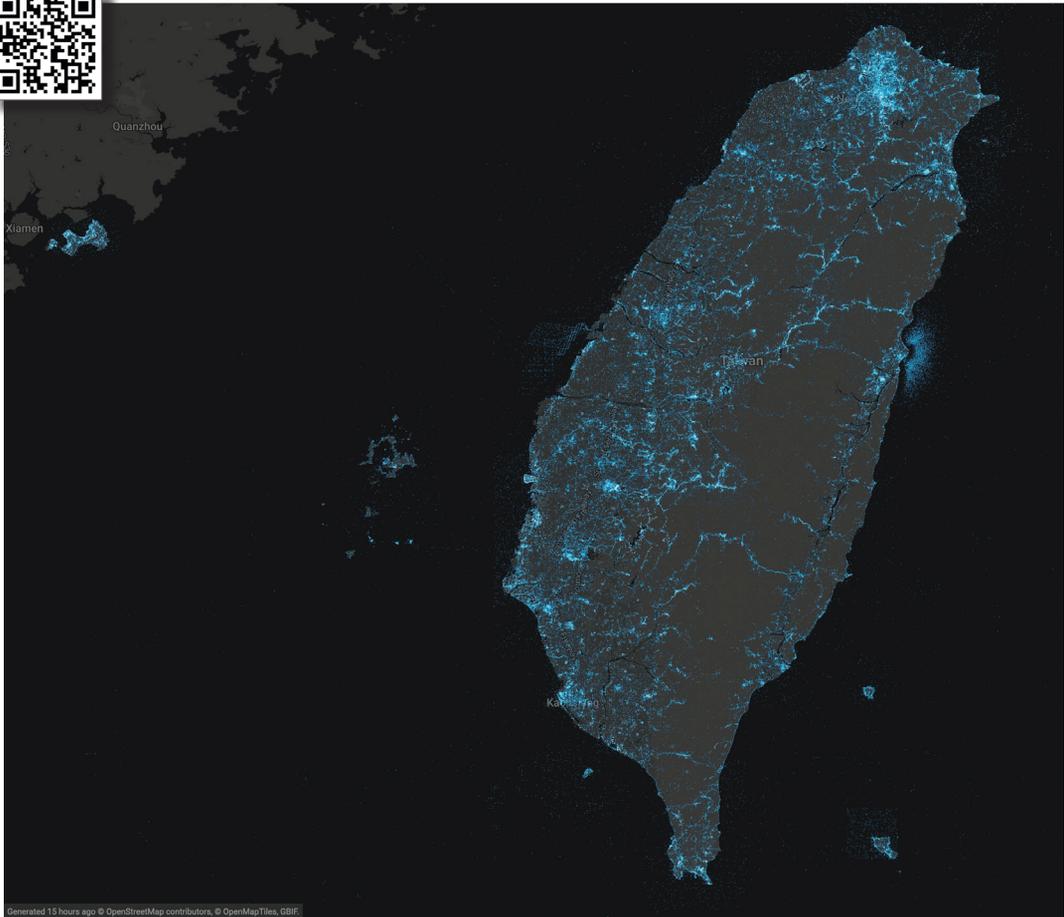
開放資料代表的意義

TaiBIF自設立以來，便在國內持續地提倡將生物多樣性資料透過GBIF平臺開放的概念，但是，這些數字究竟代表的意義是什麼？首先來看看在TaiBIF開放的資料都是什



自左至右的QR Code分別為：

- GBIF 的資料發布年度報告 (https://analytics-files.gbif.org/country/TW/GBIF_CountryReport_TW.pdf)
- eBird 的資料集 (<https://www.gbif.org/dataset/4fa7b334-ce0d-4e88-aaae-2e0c138d049e>)
- 林試所的植物標本 (<https://www.gbif.org/dataset/eb7681c5-5c9c-4e28-954f-f328991c7004>)



臺灣開放到GBIF的出現紀錄分布地圖。©OpenStreetMap contributors, ©OpenMapTiles, GBIF.

麼類型，又是來自哪裡。

資料包含的物種類別

根據GBIF的資料發布年度報告，毫無意外地，開放的資料中涵蓋最大宗的是動物的出現紀錄，其中又以鳥類的資料量最多，占全部動物的89%，主要來自公民科學eBird的資料集；植物其次，以被子植物的資料量

最多，占全部植物的81%，主要來自林業試驗所的植物標本。

開放資料的來源

生物多樣性開放資料的產生來源主要有4大類：(1) 公民科學、(2) 自然史典藏（如博物館的標本）、(3) 生態調查資料（如個人研究資料）、(4) 政府計畫資料（如林

務局、海保署等機關執行或委託的監測計畫），其中以公民科學資料為最大宗，占約87%，包括鳥類（eBird、BBS臺灣繁殖鳥類大調查）、蛾類（慕光之城）、路殺社、兩棲類（兩棲類保育志工調查）、海洋生物（珊瑚礁總體檢）等等。

這樣的結果顯示出，越容易或越能就近被觀察到的生物，出現紀錄的資料也越容易累積；而生物多樣性資料的累積，除了研究人員外，更需要大眾的參與，才能成就如此龐大的資料量，也才有機會進一步提供政府作為生物多樣性保育政策決策的參考依據，不論是用於開發案的環境影響評估、保護區的設置規劃，還是外來入侵種的防治，甚至是人畜共通疾病傳播的控制，都能透過這些資料所呈現的生物圖資，找出特定物種的熱點分布範圍及其棲地資訊，進而發揮其大數據的最大利用價值。

資料量在亞洲區坐二望一

近年來藉由推廣生物多樣性資訊的概念及價值，並在科技部、農委會及中央研究院的長期支持下，臺灣透過TaiBIF平臺及部分全球公民科學資料集（eBird, iNaturalist）發布到GBIF的物種出現紀錄資料量已超過1100萬筆，包含了17個資料發布單位，75個資料集，涵蓋了超過5萬種的物種，資料量為GBIF亞洲區網絡中的第二名，僅次於印度。雖然資料量與印度有所落差，但臺灣的資料所涵蓋的物種類群分布範圍較廣，不

僅限於鳥類等熱門生物類群；若再考量土地面積及人口的差異，臺灣實為亞洲表現最亮眼的國家，也因此GBIF秘書處副主任Tim Hirsch在2019年第26屆GBIF理事會上，特別點名臺灣是近幾年生物多樣性資訊發展的典範之一。

然而，國內仍然有許多生物多樣性資料分散在不同的資料庫中且尚待整合。在四類主要來源的生物多樣性資料中，自然史典藏和政府計畫資料通常會因計畫或個別單位管理所需而建立自己的資料庫。個別資料庫的優點是可以保有不同生物多樣性資料的特色，各單位在資料管理上也較為便利即時、成本較低，但長期下來，容易衍生各個資料庫間的資料內容重複、重工、資料格式及物種學名不一致等不利於開放再利用的議題，也顯得資料較為雜亂無章，未能被有效管理。

最常見的問題，莫過於許多政府機關窗口、民眾、研究人員或學生在搜尋資料時，會發現想要的資料在許多資料庫中都有類似的內容，但卻無法確認哪個資料來源最為可信、正確無誤；或是每當要探討特定議題時，則需自行整合多個來源的資料，才能獲得完整的參考依據，整個過程既費時又費力，不僅難以立即找到符合需求的資料庫，也阻礙了生物多樣性領域的推廣及發展。

一站式資料庫共通查詢系統

要整合國內所有的生物多樣性資料，



TBIA 資料庫共通查詢系統首頁設計。© TBIA 臺灣生物多樣性資訊聯盟

單靠TaiBIF小規模團隊的一己之力是遠遠不夠的，也因此，臺灣生物多樣性資訊聯盟（Taiwan Biodiversity Information Alliance, TBIA）的成立為這樣的困境帶來一絲曙光。TBIA為由中研院生物多樣性研究中心（TaiBIF團隊）、特有生物研究保育中心（臺灣生物多樣性網絡TBN團隊）、林務局、海洋保育署、林業試驗所、營建署（國家公園及濕地資料庫團隊）等國內收集、管理及利用生物多樣性資料的主要公家機關，於2021年9月29日共同簽署MOU正式成立，最大的目標之一便是透過跨單位的合作，設計一套可行的資料流通整合模式，善用各單位專長互補長短的精神，串連國內主要的生物多樣性資料庫，以建構去中心化

（De-centralized）形式的國內生物多樣性資訊網絡。於是，一站式的資料庫共通查詢系統的概念便因應而生，並在林務局經費支援下，由TaiBIF著手設計規劃及開發，預計最晚於2023年前正式上線。

TBIA資料庫共通查詢系統的概念，與TaiBIF網站平臺及其他目前檯面上的資料庫網站的功能和定位不同，並非取代各個資料庫提供個別出現紀錄資料查詢，也沒有倉儲與管理的服務，而是以資料庫為單位，提供使用者各資料庫共通欄位的生物多樣性資訊，亦可連結至原資料庫獲得更詳細的資料，如此便能讓使用者清楚了解哪些資料倉儲於哪些單位的資料庫中，也可保留各資料庫的資料特色。可以想像這個系統是一個生

物多樣性資料的線上購物平臺，各家的資料庫單位就像商家，可以展示他們有什麼樣的商品（資料），消費者（資料使用者）可以用關鍵字去搜尋有興趣的商品，也可以直接到該商家的網站去瀏覽。

而此系統除了提供各資料庫的物種出現紀錄資料，亦規劃了自然史典藏資料庫的查詢入口，未來將串連國內各大博物館及標本典藏單位的標本資料，並藉此合作機會建立更大的生物多樣性資訊社群網絡。同時，此系統也將成為TBIA聯盟對外聯繫的主要媒體管道，除了將持續更新發布聯盟內夥伴單位的生物多樣性資訊相關新聞及活動外，也將上傳TBIA聯盟針對不同議題所撰寫的文件及教育資源（如生物多樣性資訊行動倡議、生物多樣性領域資料標準、敏感物種資料開放作業原則等），以供生物多樣性資訊相關單位下載及參考。

一站式的資料庫共通查詢系統設計聽起來很美好，有望可以符合不同受眾的需求，但要成就這樣的概念著實不易，最具挑戰性的部分並非設計規劃網站架構，而是系統後端要整合各單位來源的資料庫。除了要面臨各個資料庫的資料量龐大，匯入後能否維持一定品質的搜尋速度的挑戰，還要解決各個資料庫資料格式的相容性，以及資料是否維持一定的品質等議題。若能完成這部分，即是為國內生物多樣性資訊進展立下新的里程碑，也等同於解決了多年來複雜且難以被處理的問題。

願景的拼圖尚需大家共同完成

TaiBIF一路走來，跨越了重重阻礙，歷經不同團隊成員的努力及累積的豐碩成果，不僅漸漸改變了生物多樣性領域的資訊合作模式，不再是各自為政，也開啟了跨單位的溝通對話，建立了友善且具熱忱的社群氛圍。然而我們與夥伴單位的努力只是小小一部分，尚未全面擴及到臺灣各地及各生物多樣性領域，中央單位的合作串連固然重要，但更需在地單位等第一線資料收集者的共同參與，才能讓這樣的願景藍圖更趨完整。相信待TBIA資料庫共通查詢系統上線後，大家更能看見TBIA聯盟的合作模式下所產出的成果，也將能吸引更多政府部門、學術單位，甚至民間團體加入，讓臺灣生物多樣性資訊的發展超越過去20年來的進程，最終達到讓所有人都能自由、即時取用任何生物的資料的願景，促進生物多樣性的保育及永續利用，成為亞洲甚至全球的典範之一。

