



開放資料流程

柯智仁

2023生物多樣性資料發布與應用工作坊

關於我

鳥類調查
族群監測
野生動物聲音辨識
生物分布資料開放與應用

小柯（柯智仁）
助理研究員
經營管理組
特有生物研究保育中心





台灣生物多樣性網路

<https://www.tbn.org.tw>



SCAN ME

野生動物與植物的時空分布
開放資料流通應用平台(TBN)

特有生物研究保育中心 台灣生物多樣性網絡團隊成員



柯智仁

計畫主持人
- 總機
- 動物調查



張慧玲

協同主持人
- 標本館
- 植物調查



俞佑錚

資料管理員
- 系統規劃
- 模式建立



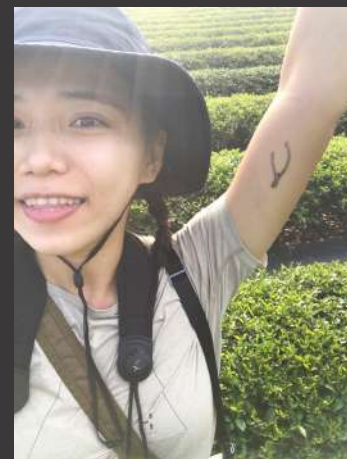
倪旻萱

平台經理
- TBN主站台
- 公民科學



吳俊毅

資料分析師
- API
- 生物地圖



陳佑真

內容經理
- 開放資料
- 資料視覺化

張藝鴻



系統工程師
- 全系統開發
- Drupal

探尋臺灣野生生物的足跡

Exploring the footprints of wildlife in Taiwan

找資料

Data

找物種

Species

看資料視覺化

Visualization

當公民科學家

Participation

19,439,580

紀錄

Occurrence

1,577

資料集

Dataset

27,940

物種

Species

12,907

下載次

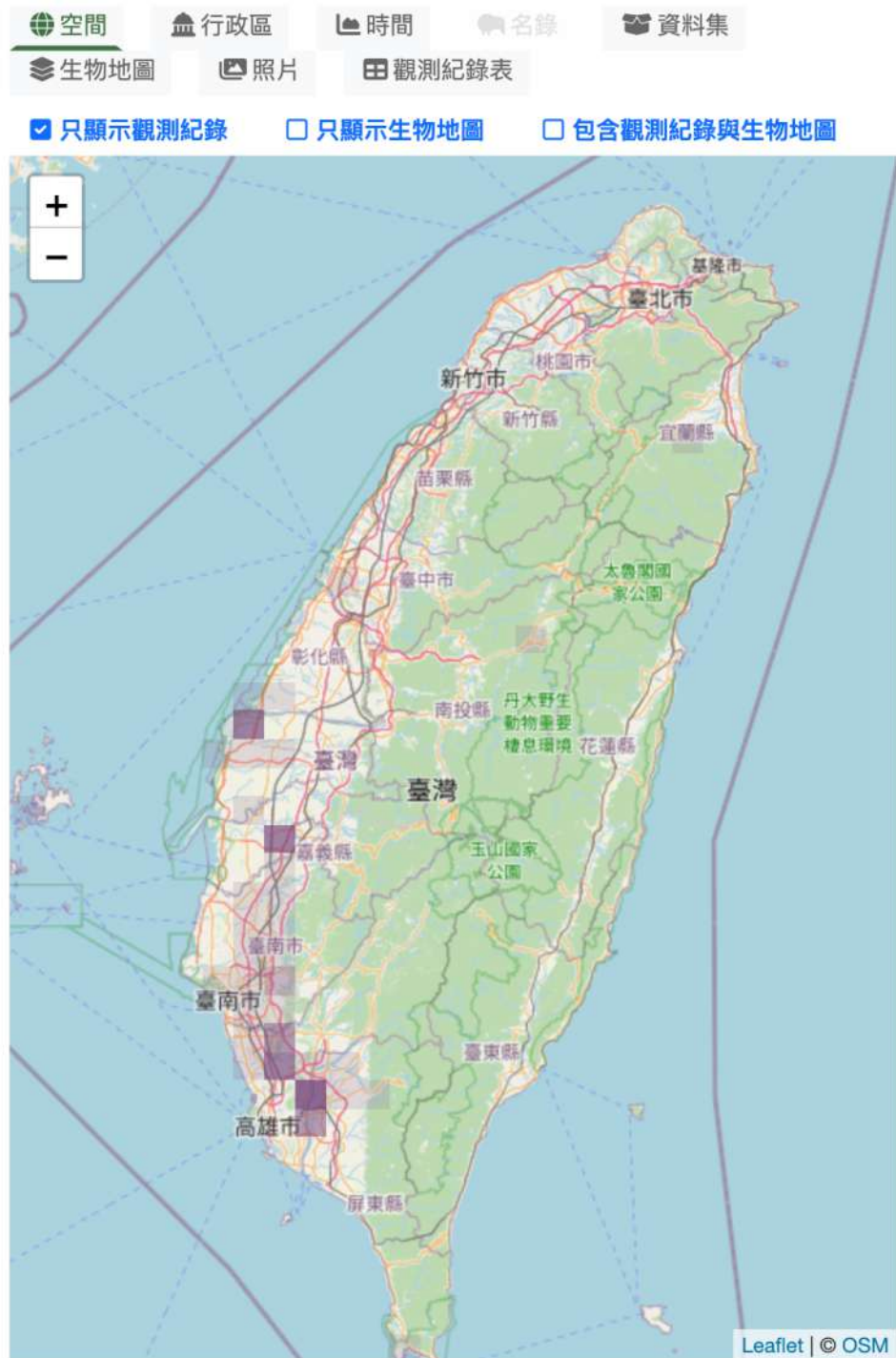
Download

TBN的主要功能:

1. 查某物種的分布位置
2. 查某地區的物種清單
3. 把資料下載打包回家

功能1 範例： 關注物種 草鴉的分布位置

查某物種的分布位置

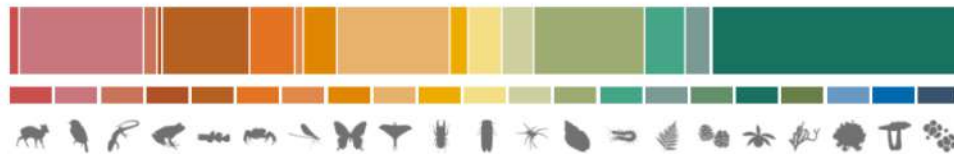


功能2範例： 台東池上附近 5KM的物種清單

查某地區的物種清單

下載名錄

2,468物種



類群

- 哺乳類 (23)
- 鳥類 (261)
- 爬行類 (29)
- 兩棲類 (9)
- 魚類 (185)
- 蝦蟹類 (94)
- 蜻蛉類 (20)
- 蝶類 (70)
- 蛾類 (235)
- 甲蟲類 (39)
- 其他昆蟲 (72)
- 蜘蛛類 (66)
- 蝸牛與貝類 (235)
- 其他無脊椎 (87)
- 蕨類 (50)
- 裸子植物 (4)
- 被子植物 (965)
- 石松類 (5)
- 苔蘚 (3)
- 真菌類 (14)
- 原藻類 (2)

魚類 (185)

- *Ablabys macracanthus* 大棘帆鰭鮪
- *Abudefduf vaigiensis* 條紋豆娘魚
- *Acanthopagrus pacificus* 太平洋棘鯛
- *Acanthopagrus schlegelii* 黑棘鯛
- *Acanthopagrus taiwanensis* 臺灣棘鯛
- *Acanthurus dussumieri* 杜氏刺尾鯛
- *Acanthurus mata* 後刺尾鯛
- *Acanthurus xanthopterus* 黃鰭刺尾鯛
- *Acropoma japonicum* 日本發光鯛
- *Alepes djedaba* 吉打副葉鰷
- *Aploactis aspera* 絨皮鮪
- *Apogonichthyoides niger* 黑似天竺鯛
- *Arothron hispidus* 紋腹叉鼻魷
- *Arothron manilensis* 菲律賓叉鼻魷
- *Arothron reticularis* 網紋叉鼻魷
- *Atherinomorus lacunosus* 南洋美銀漢魚
- *Atropus atropus* 溝鰷
- *Bleekeria mitsukurii* 箕作布氏筋魚
- *Bodianus bilunulatus* 雙帶狐鯛
- *Bothus mancus* 蒙鯉
- *Brachysomophis cirrocheilos* 鬚唇短體蛇鰻
- *Callionymus curvicornis* 彎角
- *Callionymus planus* 扁
- *Callionymus scabriceps* 粗首鰻
- *Caprodon schlegelii* 許氏菱齒花鮨
- *Carangoides armatus* 甲若鰺
- *Carangoides ferdau* 平線若鰺

功能3範例： 開放資料一鍵下載 直接寄到你的信箱

把資料下載打包回家

範圍	空間範圍=121.21489954646678,23.08897500770274 半徑5公里 統計至種階層 只包含觀測紀錄
資料時間	2023-02-04
授權條款	政府資料開放授權條款－第1版 。個別資料授權情形，請見個別資料內標示
使用條款	請閱讀本站 使用條款 ，下載資料即表示您同意該條款內容。
保留期限	本站保留下載與檔案連結一年（至2024-02-04）。如有延長需求或因學術出版引用而有永久保留需求，請 聯絡我們 。
檔案格式	匯出之資料為csv格式，你可以用excel等試算表軟體開啟。 如有開啟困難，請參閱 csv亂碼解決教學
備註	檔案為離線產生，處理完成後，系統會寄下載資訊到您輸入的電子郵件信箱。 如未收到信件，請檢查您的郵件設定，如仍未收到信件，請 聯絡我們 。

請輸入您的電子郵件信箱 *

請再輸入一次同樣的信箱 *

下載

為什麼要開放資料？

促進研究調查資料 的應用價值

最開始的初衷

只是覺得過去的資料躺在硬碟裡很可憐
尤其如果硬碟莫名壞掉的時候那種悵然所失的感覺…

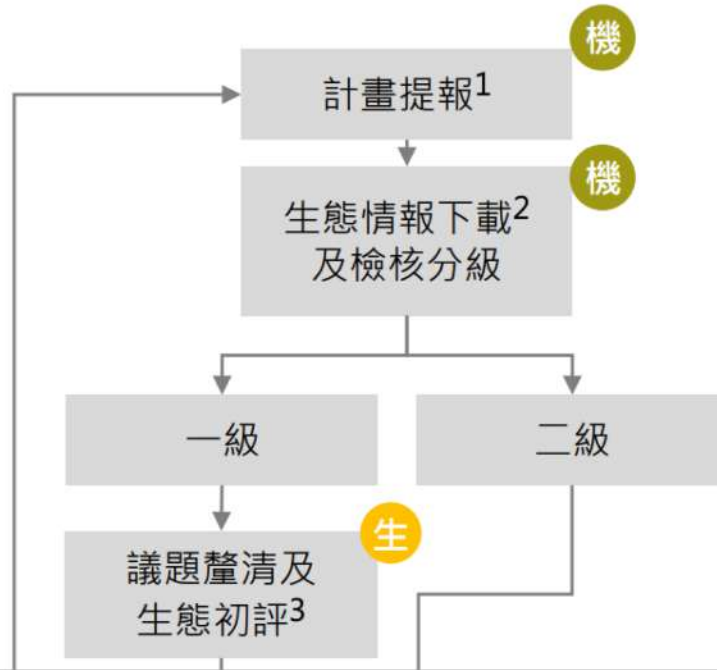
開放資料的正面效益

1. 讓你的資料成為全球的生物多樣性知識的一環
contribute to global knowledge about biodiversity
2. 讓你與其他資料擁有者及研究者之間迸發新的合作契機
revealing new opportunities for collaboration among data owners and researchers
3. 讓你在蒐集與典藏生物多樣性資料的工作成就被看見與重視
properly credited for your work to create and curate biodiversity data
4. 讓你追蹤辛苦數位化的資料被使用與引用的情形
trace usage and citations of digitized data
5. 讓你滿足計畫結案時需將資料開放的經費支助條件
make data freely accessible at the end of a project

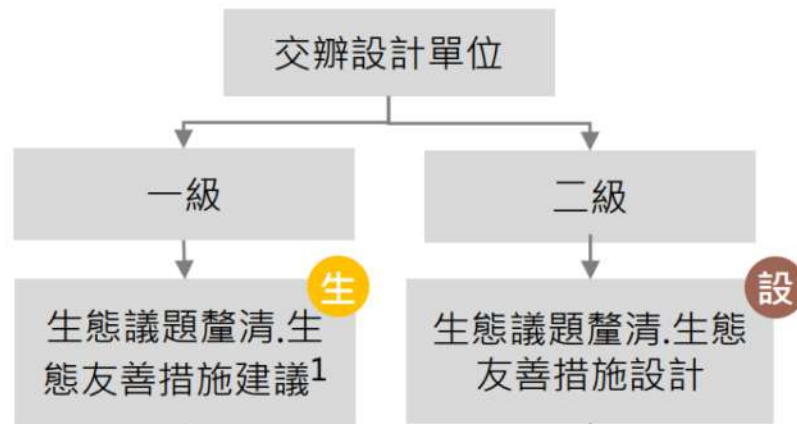
成為全球生物多樣性知識的一環

- 就能成為保育與永續發展的解方。
- 舉例：
 - 行政院公共工程委員會規定政府補助50%+的公共工程案要做生態檢核（主要為彌補面積小不用做環評的小型工程案，如溪流整治）
 - 行政院農業委員會水土保持局（水保局）會將工程場域於事先分級，被歸類為一級檢核區的會在施工流程中加入生態初評、生態友善措施建議、生態監測等作業程序。
 - 生態檢核的分級依據的是工程案有範圍內有沒有敏感性物種分布（如保育類或紅皮書物種）。
 - 有沒有敏感物種的分佈的依據是符合FAIR資料原則的開放資料。

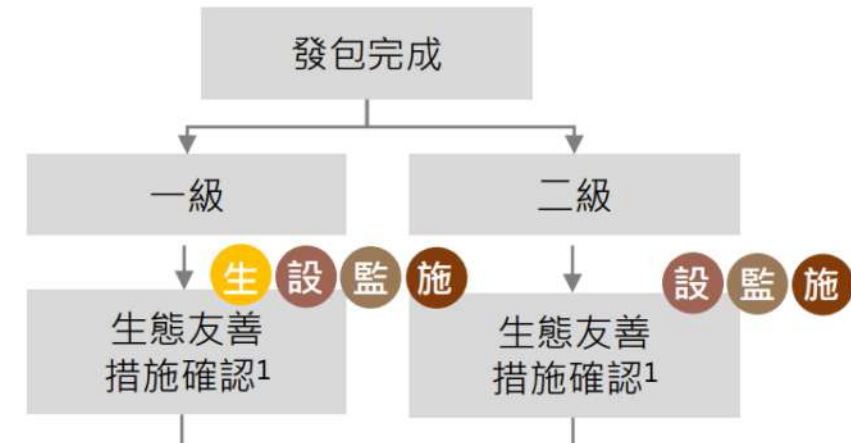
2-2 提報審議階段



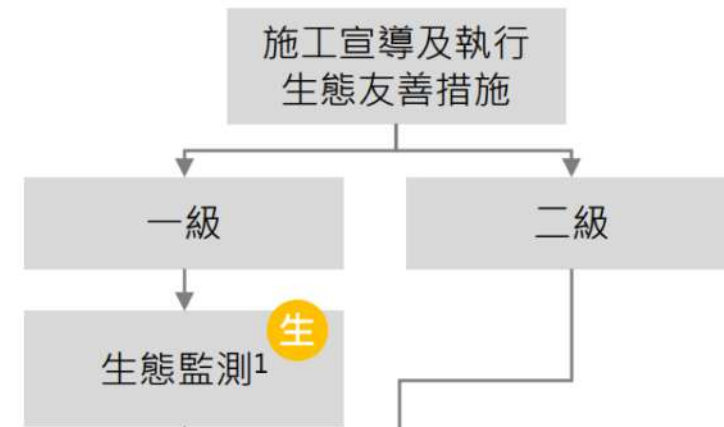
2-3 設計階段



2-4 施工階段(開工前)



2-5 施工階段(施工期間)



資料來源：
生態檢核HANDBOOK
水保局 (108年)

負責單位

- 機 工程機關
- 生 生態團隊
- 設 設計單位
- 監 監造廠商
- 施 施工廠商

野溪整治暫停工 屏東內文溪是稀有「黃尾弓蜓」棲地 水保局將會勘

◎ 2022年05月24日

環境資訊中心記者 劉庭莉報導

[5/24 17:00 更新：水保局回應]

屏東縣獅子鄉內文溪，是稀有的蜻蜓目成員「黃尾弓蜓」唯一族群數量穩定的棲息地。台灣蜻蜓學會發現，當地正在進行野溪整治工程，部分河段河道挖深，「被挖的地方，我們已知黃尾弓蜓最穩定的棲息地。」目前水保局台南分局下令暫時停工，預計下週日）舉行會勘，協調相關補救措施，將傷害減到最輕。



內文溪河道旁植被遭到破壞。圖片來源：台灣蜻蜓學會提供

生態檢核資料庫沒有黃尾弓蜓 水保局：會再納入資料庫中

由水保局台南分局主辦的「內文一號橋下游野溪整治工程」，主要因內文溪橋樑既有護岸基礎已經掏空、河道轉彎處河床逐年刷深，在地居民擔心豪大雨來臨時，護岸崩塌、橋樑損壞，或河水氾濫影響附近農田。工程預算來自前瞻基礎建設計畫經費，今年3月10日開工、原定10月完工。

水保局台南分局21日表示已暫時停工。副分局長祝瑞敏今受訪表示，25日於恆春鎮公所的港口溪生態流域座談會，將一併討論內文溪工程，台灣蜻蜓學會也受邀參加；5月30日暫定召開會勘，除台灣蜻蜓學會，也邀請獅子鄉公所、陳情人、生態學者一同協調補救措施，將傷害減到最輕。

中央的新建公共工程，依法需做生態檢核。而為了避免踩雷，水保局2018年起建置「集水區友善環境生態資料庫」，區分出一級跟二級生態檢核區。工程進行前，承辦人員會將工程座標輸入生態資料庫中，系統會依據搜尋到的物種、棲地、生態環境是否為敏感區位等，自動進行生態檢核分級。

「資料庫一定是慢慢累積，才會越來越完整，」祝瑞敏坦言，因生態資料庫中資訊不足，不清楚有黃尾弓蜓存在，所以當初沒有邀請蜻蜓專家來參與，「如果生態學者或環境團體有提供更詳細的資訊，會再納入資料庫中。」

迸發新的合作契機

- 資料的發布促成了全球資料集的整合，更容易讓志同道合且擁有彼此需要的知識的人找到彼此
 - 舉例：
 - 蛾類採集者 x 氣候變遷學者 x 資料科學者
- <https://www.nature.com/articles/s41467-019-12500-2>

[nature](#) > [nature communications](#) > [articles](#) > article

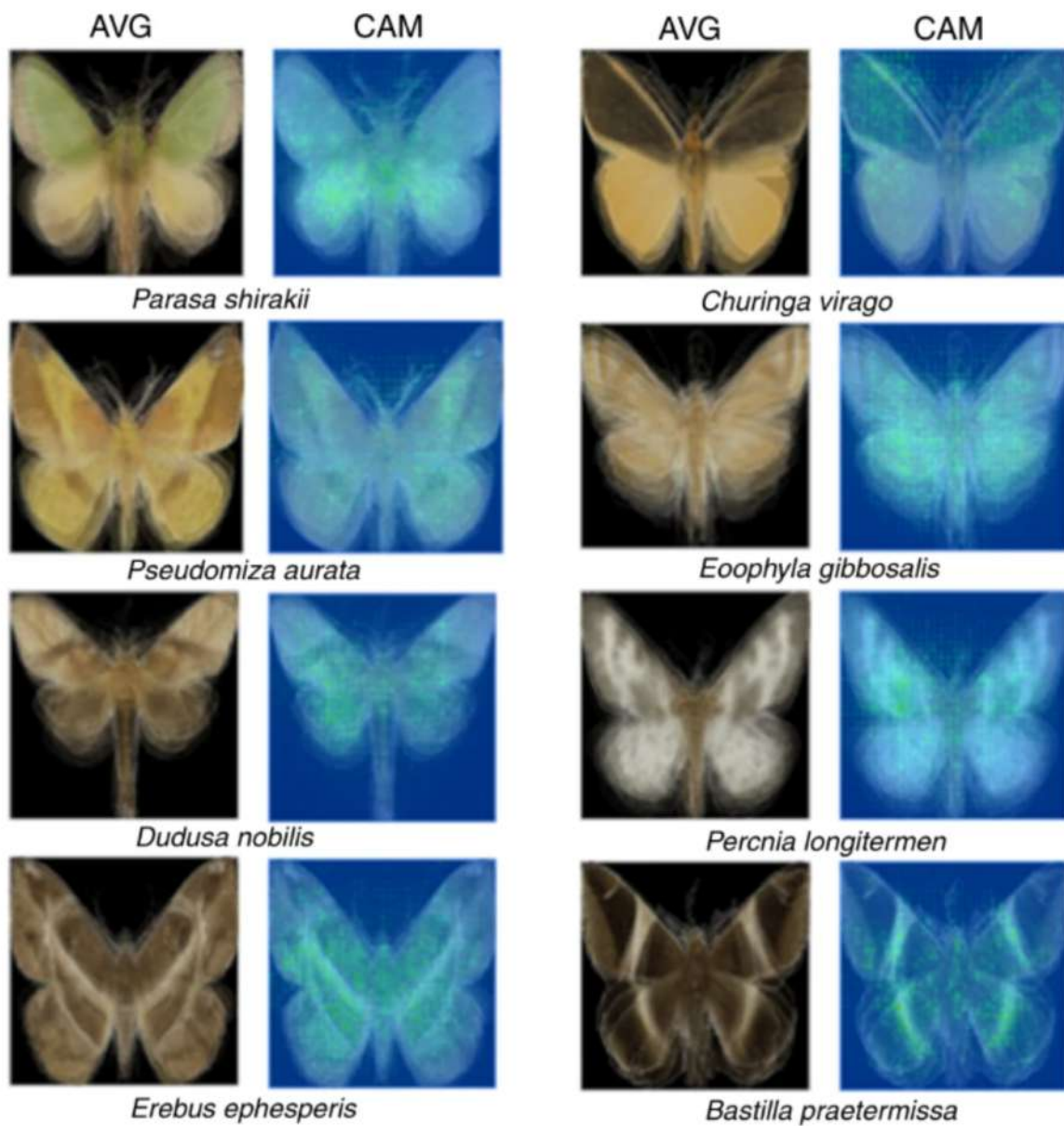
Article | [Open Access](#) | [Published: 07 October 2019](#)

Artificial intelligence reveals environmental constraints on colour diversity in insects

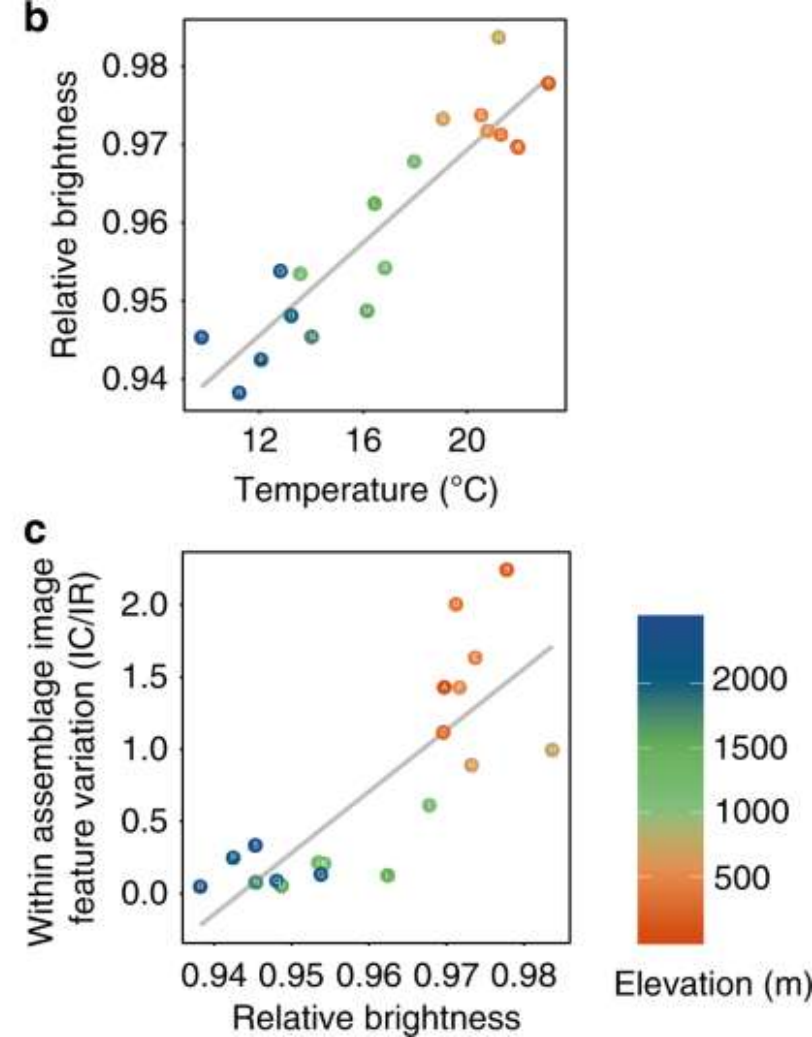
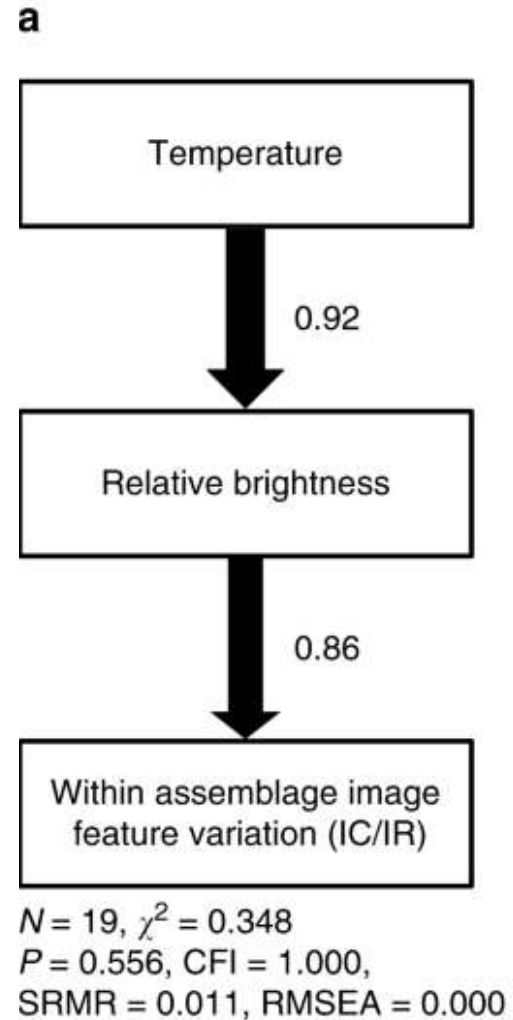
[Shipher Wu](#), [Chun-Min Chang](#), [Guan-Shuo Mai](#), [Dustin R. Rubenstein](#), [Chen-Ming Yang](#), [Yu-Ting Huang](#),
[Hsu-Hong Lin](#), [Li-Cheng Shih](#), [Sheng-Wei Chen](#) ✉ & [Sheng-Feng Shen](#) ✉

[Nature Communications](#) **10**, Article number: 4554 (2019) | [Cite this article](#)

8271 Accesses | **14** Citations | **48** Altmetric | [Metrics](#)



Class activation mapping generated localized discriminative visual features of arbitrarily selected images of eight species of moths. CAM represents class activation mapping, which is a method of localizing the discriminative image regions⁵⁹. AVG shows the mean RGB of each pixel from images of all individuals of a species



蒐集與典藏資料的工作成就被看見與重視

- 透過好的詮釋資料 (metadata) 撰寫
- 彰顯資料搜集與典藏中，不同角色的個人工作成就
- 提升發佈資料的機關能見度 (寫成看起來很厲害的KPI)
- 若再撰寫成資料論文 (data paper)，將更可透過學術引用，促進你的相關工作成就在學界的認可
- (請大家期待第三天的資料論文課程)

追蹤資料被使用與引用的情形

- 特別是在GBIF，他對於資料集的引用與追蹤做得很好
- 可以透過追蹤他人如何使用你所搜集或典藏的資料啟發研究靈感
- 是很棒的KPI（而且KPI統計都還是GBIF幫你做的）
- （請大家期待第三天的資料應用課程）

滿足結案時需將資料開放的經費支助條件

- 提升計畫的公益性
- 促成研究成果的可再現性
- 越來越多經費補助的申請條件包含開放資料
- 或至少明確規劃與交代「資料最終放置位置」
 - （通常不是個人電腦）
- 研究期刊對於期刊論文所採用的資料應該要open access的期待也日益提升。

- 舉例：
 - 國科會自然科學與永續研究發展處-永續發展整合研究
 - 112 年度計畫申請及議題：
 - 「為促進研究計畫之科學資料產製品管、資料保存及研究成果資料分享。鼓勵於研究計畫納入「研究資料管理方案(Data Management Plan, DMP)」」
 - DMP需要明訂交代後續資料分享的方式。
 - DMP所需經費最高可佔核定額度總數的10%。
- (請大家期待第三天的DMP課程)

當生物多樣性遇見開放資料



現地
採集者

觀察現象
與記錄第一手
資料

資料
管理者

管理建檔
資料以協助知
識典藏

研究
發表者

建構理論
進行驗證
論文發表

策略
擬定者

評估資訊
考量情境
定下決定

一線
工作者

科學教育推廣
保育措施實作
生物多樣性

不論你的角色是何者
你都可以參與開放資料的貢獻與應用

X, Y 座標 ; 地點描述

如何將帶有地理空間資訊的
野外調查資料發佈到網路上，
以擴大資料的應用價值？

合作機會
發表機會
表揚機會

TaiBIF IPT




GBIF

INTEGRATED PUBLISHING TOOLKIT^(IPT)

free and open access to biodiversity data

資料發佈工具

TaiBIF IPT >>> GBIF



GBIF

INTEGRATED PUBLISHING TOOLKIT (IPT)
 free and open access to biodiversity data

email

登入

申請帳號





繁體中文






首頁

關於

此 IPT 上維護的公開資源

過濾器:

標誌	名稱	組織	資料集類別	資料集子類別	紀錄筆數	最後修改	上次發布	下次發布
--	The Taiwan Breeding Bird Survey Data	Taiwan Endemic Species Research Institute	Sampling event	--	423,119	2018-03-27	2018-03-27	--
	Alien plant presence dataset from the point-radius plot surveys in 2010-2015 in Taiwan	Taiwan Forestry Research Institute	Sampling event	--	4,879	2016-09-23	2016-09-23	--
	台灣花蓮沿海賞鯨活動鯨豚觀察紀錄 (The observation records from the activities of whale and dolphin watching in the offshore of Hualien, Eastern Taiwan)	未註冊	Sampling event	--	5,247	2016-11-28	2016-11-28	--
--	2015 北部各核能電廠附近海域之生態調查 魚類	未註冊	Sampling event	--	160	2016-06-28	2016-06-28	--
--	Digital Archives of Taiwan Malacofauna Database	Taiwan Biodiversity Information Facility (TaiBIF)	出現紀錄	標本	7,637	2017-04-18	2014-03-25	--
	A Dataset of Deep-Sea Fishes Surveyed by Research Vessels in the Waters around Taiwan	Taiwan Biodiversity Information Facility (TaiBIF)	出現紀錄	--	3,653	2014-12-22	2014-10-28	--
	A dataset from bottom trawl survey around Taiwan	Taiwan Biodiversity Information Facility (TaiBIF)	出現紀錄	觀察	3,529	2014-05-22	2014-03-25	--



SEARCH DATASETS | 3 RESULTS

ALL OCCURRENCE CHECKLIST SAMPLING EVENT METADATA

[Taiwan Moth Occurrence Data Collected From Social Network](#)

Occurrence dataset

此一資料集為特有生物研究保育中心(Taiwan Endemic Species Research Institute, TESRI)應用社群網站-臉書(Facebook)成立社團-「暮光之城-蛾類世界」(<https://www.facebook.com/groups/EnjoyMoths2/>)，邀集對蛾類有興趣的網友加入，並鼓勵加入的網友分享蛾照及拍照地點和日期的資料，特生中心的蛾類科普及推廣團隊則...

Published by Taiwan Endemic Species Research Institute

183,075 occurrences

[The Taiwan Breeding Bird Survey Data](#)

Sampling event

The Taiwan Breeding Bird Survey (BBS Taiwan) is a national monitoring project with the aim of establishing a breeding bird population index for conservation. It was established in 2009 as a citizen sc...

Published by Taiwan Endemic Species Research Institute

384,325 occurrences



[The Taiwan Roadkill Observation Network Data Set.](#)

Occurrence dataset

We describe the wildlife road kill data set collected by citizen scientists from the Taiwan Road Observation Network (TaiRON). This data set includes 46,416 geospatially referenced occurrence points o...

Published by Taiwan Endemic Species Research Institute

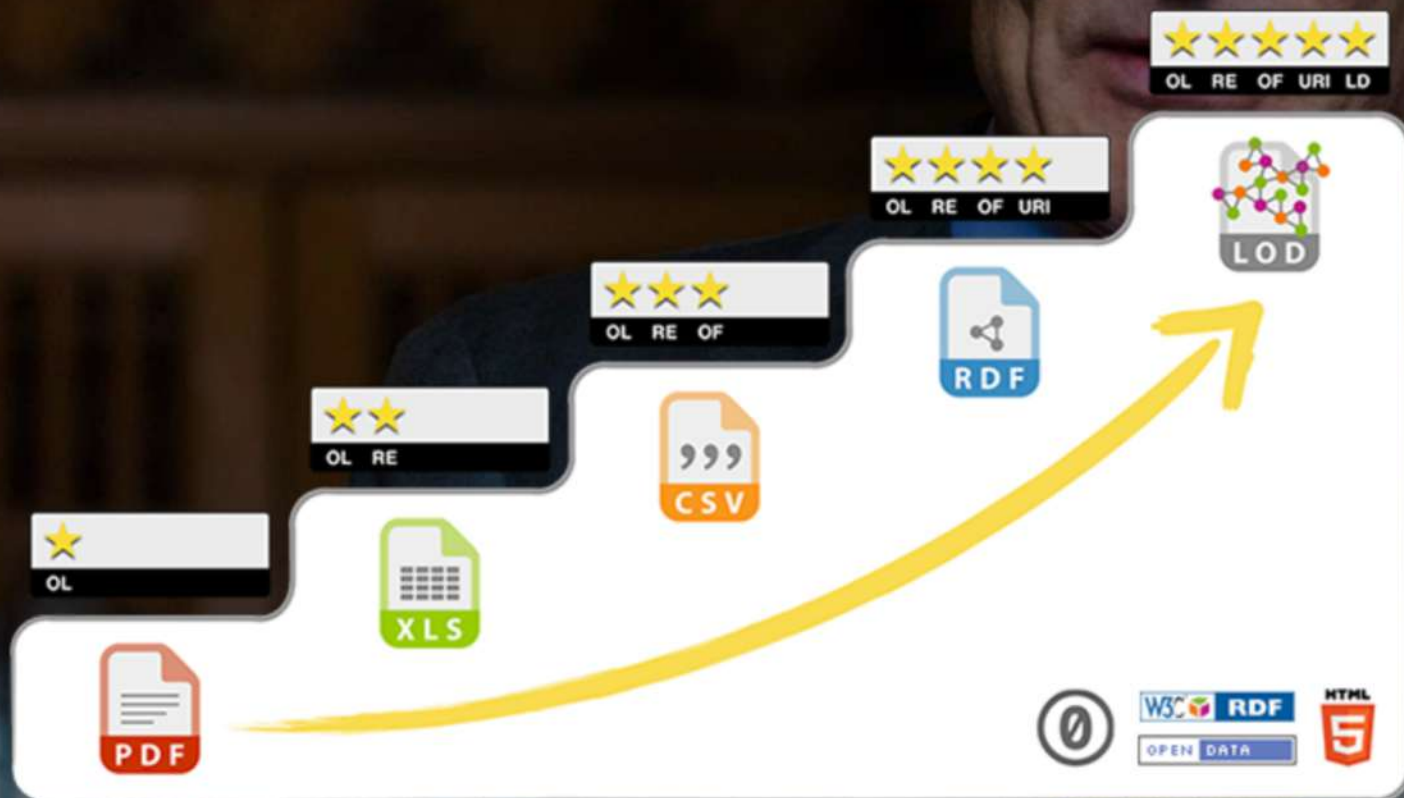
<https://www.gbif.org>



什麼是這場工作坊
要談的「開放資料」？

五顆星 ★ 開放資料

全球資訊網（World Wide Web）發明者和鏈結資料的創始者，提姆·柏納-李 (Tim Berners-Lee) 建議了一個開放資料五顆星的分類架構 在此，我們提供在每一顆星中每一步驟的範例且解釋這些步驟的成本和效益：





採用開放授權，讓手上的資料（任何資料格式）可以在網路上取得¹



讓這份資料能以結構化的方式取得（例如用 Excel 取代掃描的表格）²



使用開放格式取代專屬格式（例如用 CSV 取代 Excel）³



使用 **固定網址** 來表示資料，使其它人可以連結到資料在資料網絡中的位置⁴



鏈結你的資料到其它資料，以提供資料之間的脈絡關係，例如兩份資料間的相等關係（owl:sameAs）。⁵

CC0 (公眾領域貢獻宣告)

CC BY (姓名引用標示)

CC BY-NC (姓名引用標示+非商業使用)

補充說明：

以臺灣來說，「政府資料開放授權條款」在定義上等同於CC BY，所以採用「政府資料開放授權條款」釋出的資料可以無痛改成在GBIF上以CC BY釋出。

FAIR Data Principle

科學資料FAIR原則

但資料不僅僅是有開放授權就夠，更完整的定義是資料要符合四個條件：

Findable、Accessible、Interoperable、Reusable

SCIENTIFIC DATA

Amended: Addendum

OPEN

SUBJECT CATEGORIES

» Research data

» Publication

characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*[#]

FAIR原則是根據2016年於Nature的系列期刊Scientific Data內發表的文章，已被科學界廣為採用。這四個原則有很詳細的定義，請見後表：

[doi: 10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

Box 2 | The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1 the protocol is open, free, and universally implementable
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with detailed provenance
 - R1.3. (meta)data meet domain-relevant community standards

資料及其說明有在**可搜尋**
且**可下載**的公開儲存庫

找得到

F
indable



拿得到

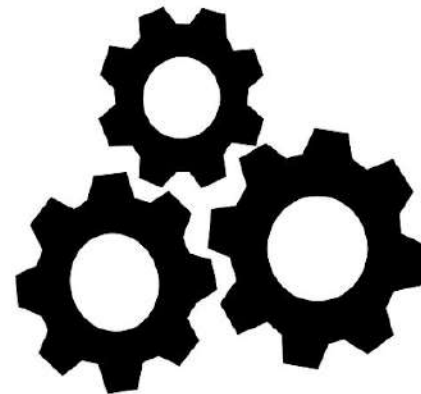
A
ccessible



資料說明與**格式**
有依據資料標準

看得懂

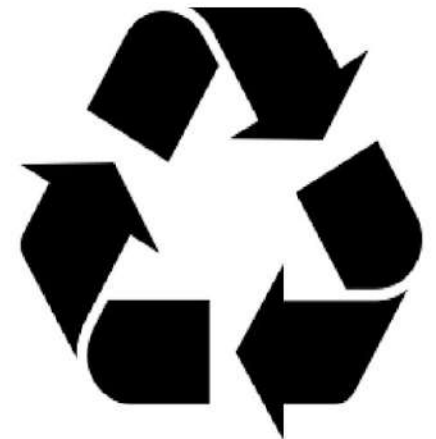
I
nteroperable



資料以可信賴
的**開放授權**釋出

可再用

R
eusable



符合 FAIR Data Principle 並且以
CC0 或 CC BY 或 CC BY-NC 或
釋出的資料

這場工作坊聚焦的開放資料的明確定義為：

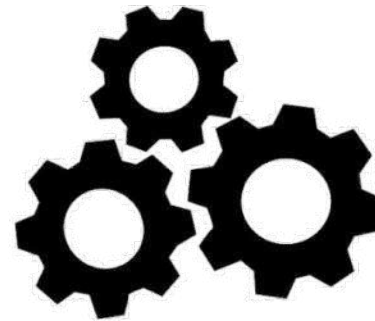
F_{indable} A_{ccessible} I_{nteroperable} R_{eusable}



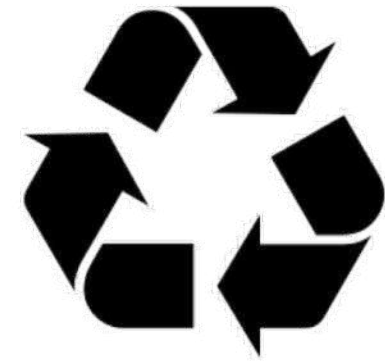
找得到



拿得到



可互用



可再用

可被搜尋 =
有標籤跟有辨識號

可被取得 =
有公開的取得位置

看得懂 =
公用與流通的語言

不會被告抄襲 =
符合著作權相關法律

不論你是誰、你在哪、任何時候，可以不受限的取得資料
並以CC0、CC BY、CC BY-NC使用的就是開放資料！

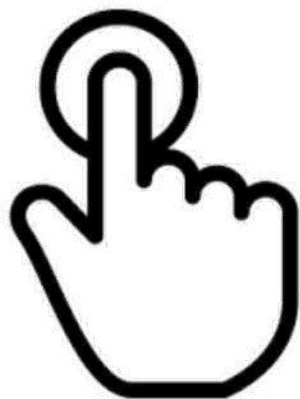
什麼是理想的 開放資料流程？

就是幫助我們達成前述各種開放資料的條件的SOP

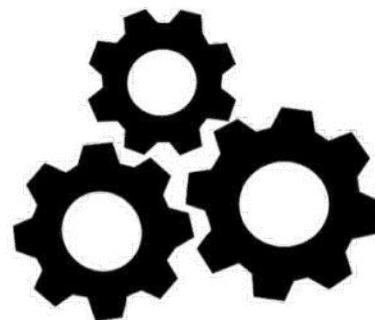
用FAIR理解開放資料的to-do list



Findable
找得到



Accessible
拿得到



Interoperability
可互用



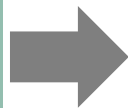
CC BY-SA 4.0 SangyaPundir

Reusable
可再用

在公開的地方
介紹你的資料
詮釋資料



在公開的地方
放上你的資料
xlsx, csv, ...etc.



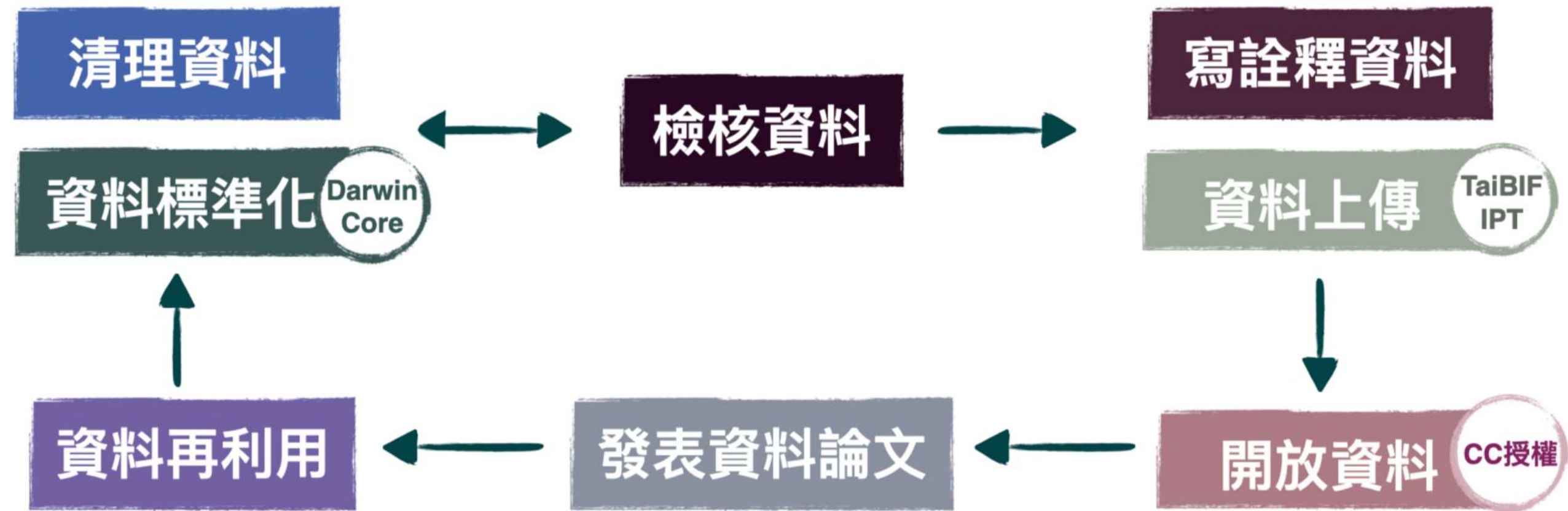
照公開的標準
整理你的欄位
Darwin Core



選適當的開放法律授權
宣告引用或使用方式
Creative Commons

這些全部，都可以在 TaiBIF IPT 上帶你一步一步作到好！

理想的開放資料流程 = IPT的資料發布流程！



取得此壁報



開放資料：如何成為環境決策的協作參與者

HOW TO OPEN DATA ON TaiBIF IPT

#我想開放資料，該如何開始？

02 DATA MANAGEMENT

管理規劃



好的開放資料前期的管理規劃，可謂該資料集完成了一半。



生物調查紀錄方法與需求

確認生物調查計畫的作業方式、作業流程、紀錄內容、紀錄格式、事後分析方法與資料用途。



目的開放資料平台

各平台的資料開放格式、管理方式、及資料授權選擇皆有異，可以根據資料集定位做開放資料平台的選擇。此壁報將以TaiBIF IPT為例。



欄位標準的對應

根據開放資料平台所規範的匯入資料集格式，設計記錄用的欄位表單，並請留意必填欄位、欄位填寫內容標準與參考。例：Darwin Core。

資料清理



手邊的資料若已經備齊，便可著手執行資料上傳前的最後一環，資料的清理。



保留原始資料及資料真實性

備份原始資料後進行清理，多版本控制以保留原始資料。



表單格式整理

維持欄位名於表單最上方列、不合併欄位、值與單位分開等。



欄位標準的對應

核對欄位名以及填入值是否吻合開放資料平台所制訂標準。



有效學名比對

確認學名可被辨識，實用工具如：NomenMatch。



地理座標格式與範圍驗證

核對座標格式，並且座標座落於調查紀錄合理範圍。



日期格式與區間驗證

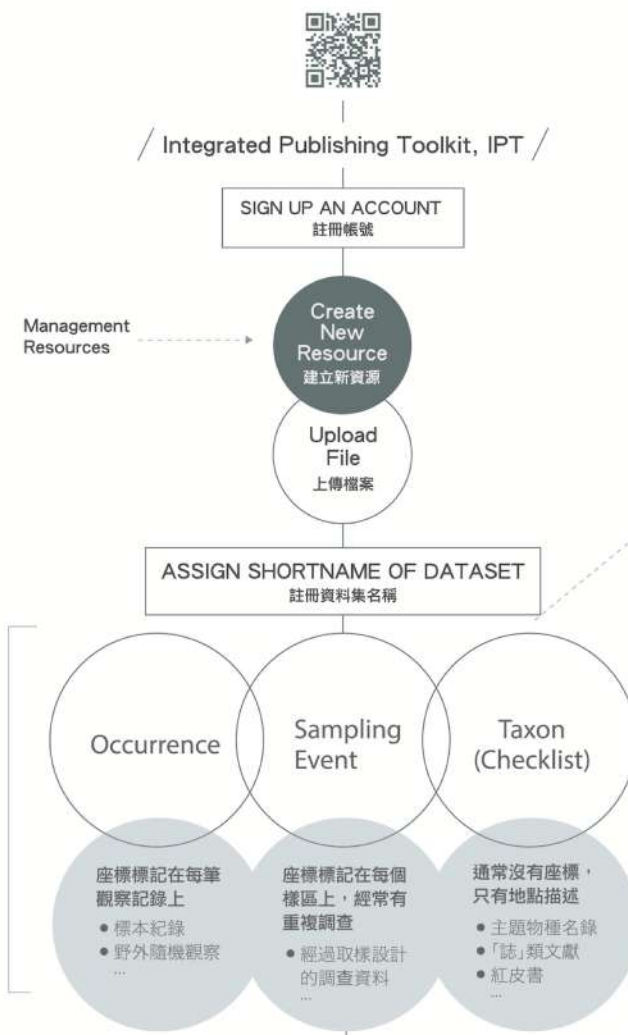
核對日期格式，並且日期座落於調查紀錄合理區間。



空值排除與定義

確認資料空值並非資料填寫缺漏、錯誤。

03 PUBLISH DATA ON TaiBIF IPT



無論是個人、研究單位、民間組織等，如果您認同資料開放的理念，都可以透過TaiBIF IPT成為貢獻資料的一員。TaiBIF為GBIF的台灣節點，TaiBIF IPT為台灣節點的發布工具。TaiBIF IPT提供資料提供者的資料管理及發布服務，可透過此工具於雲端儲存個人或組織的生物多樣性資料開放至TaiBIF及GBIF平台。

TBN收錄所有源自GBIF的資料集，一旦於IPT發布並且註冊（register）成功，將可於TBN頁面被搜尋及下載應用。

#我應該放上哪些資料欄位？

參考必填欄位

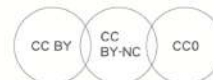
Occurrence	Sampling Event	Checklist (Taxon)
occurrenceID basisOfRecord eventDate individualCount locality	eventID eventDate samplingProtocol	taxonID scientificName

Darwin Core quick reference guide

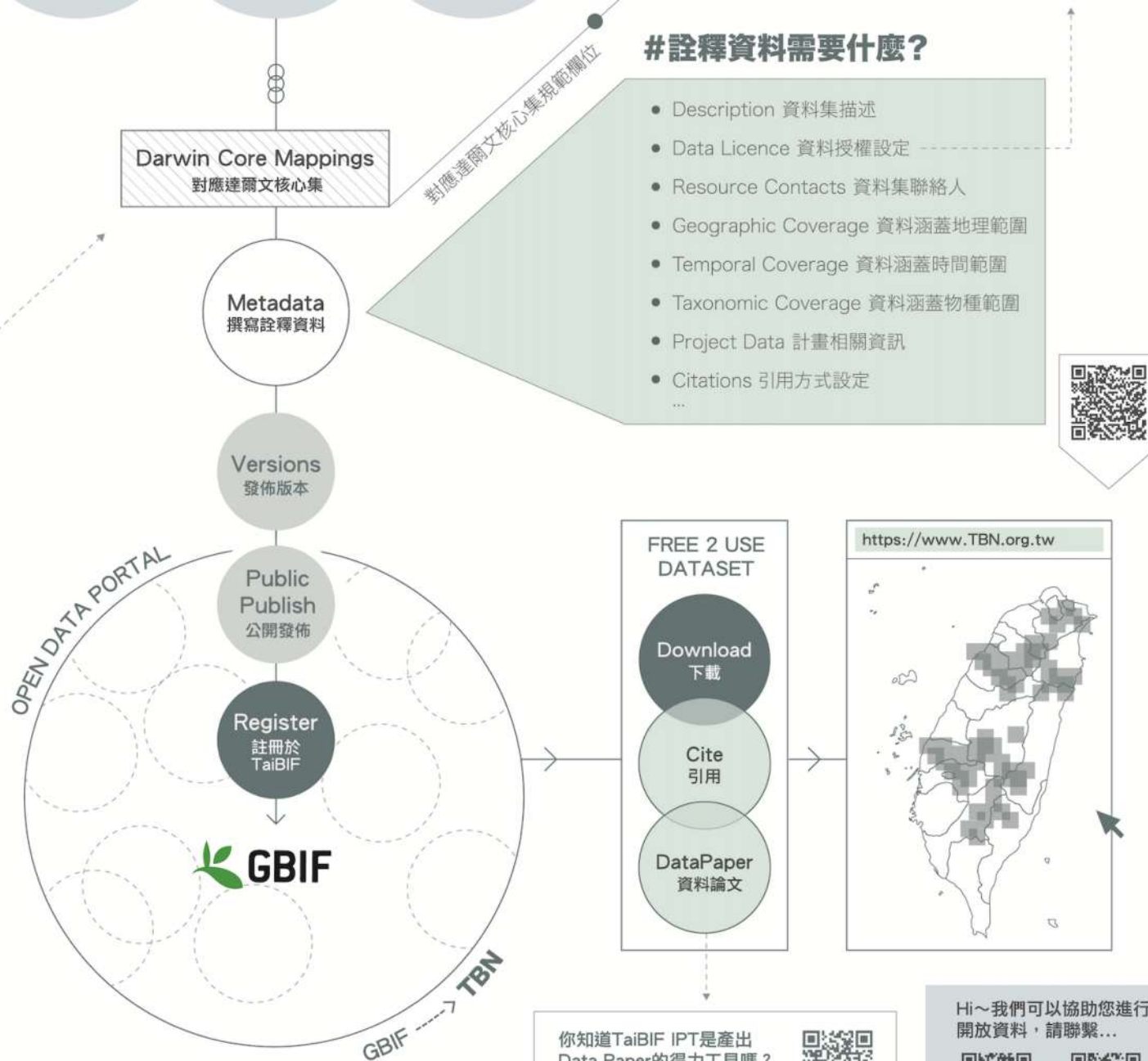
更多欄位



Choose
TYPE (Core)
選擇核心集類型



取得此壁報



你知道TaiBIF IPT是產出
Data Paper的得力工具嗎?

詳情請見



Hi~我們可以協助您進行
開放資料，請聯繫...



TBN

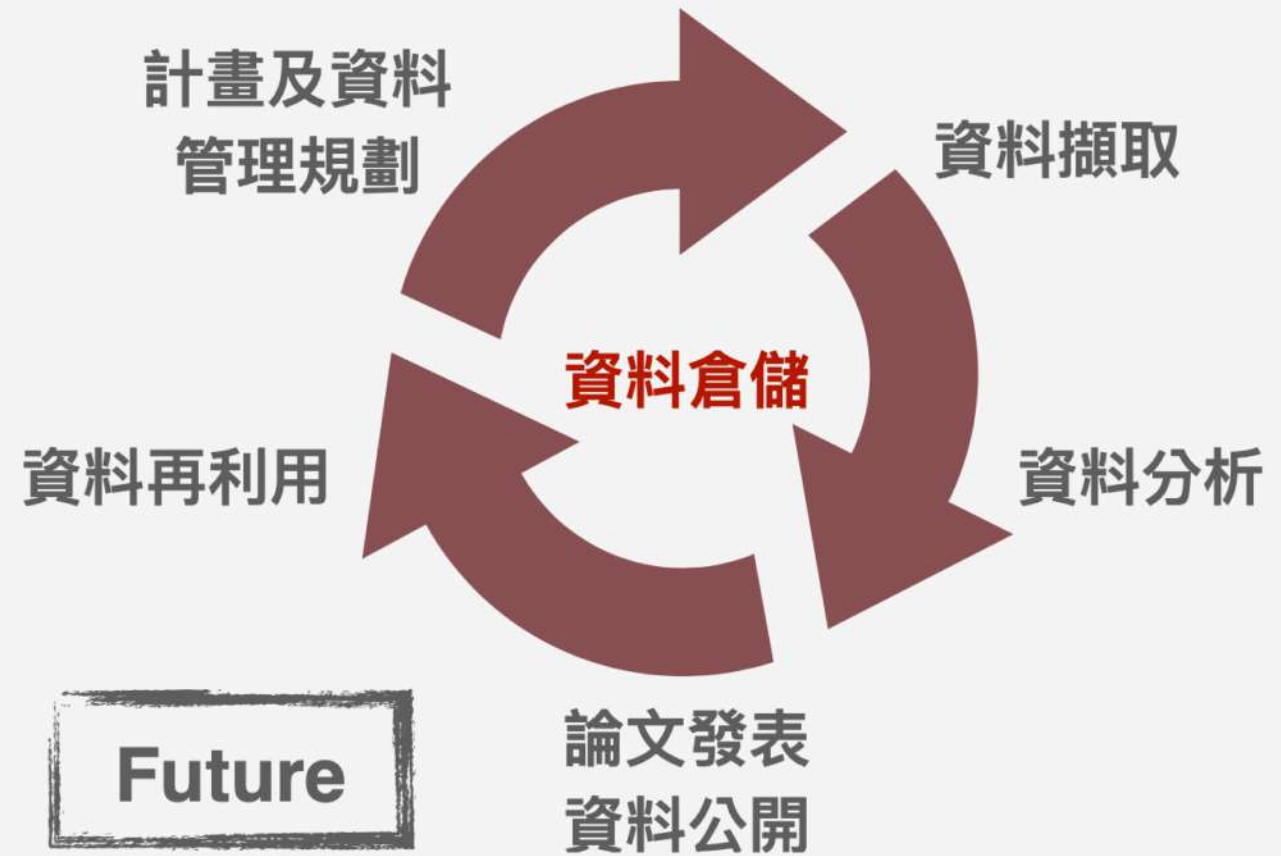


TaiBIF

結語

以開放資料圓滿
資料生命週期

資料生命週期



計畫與
資料管理規劃

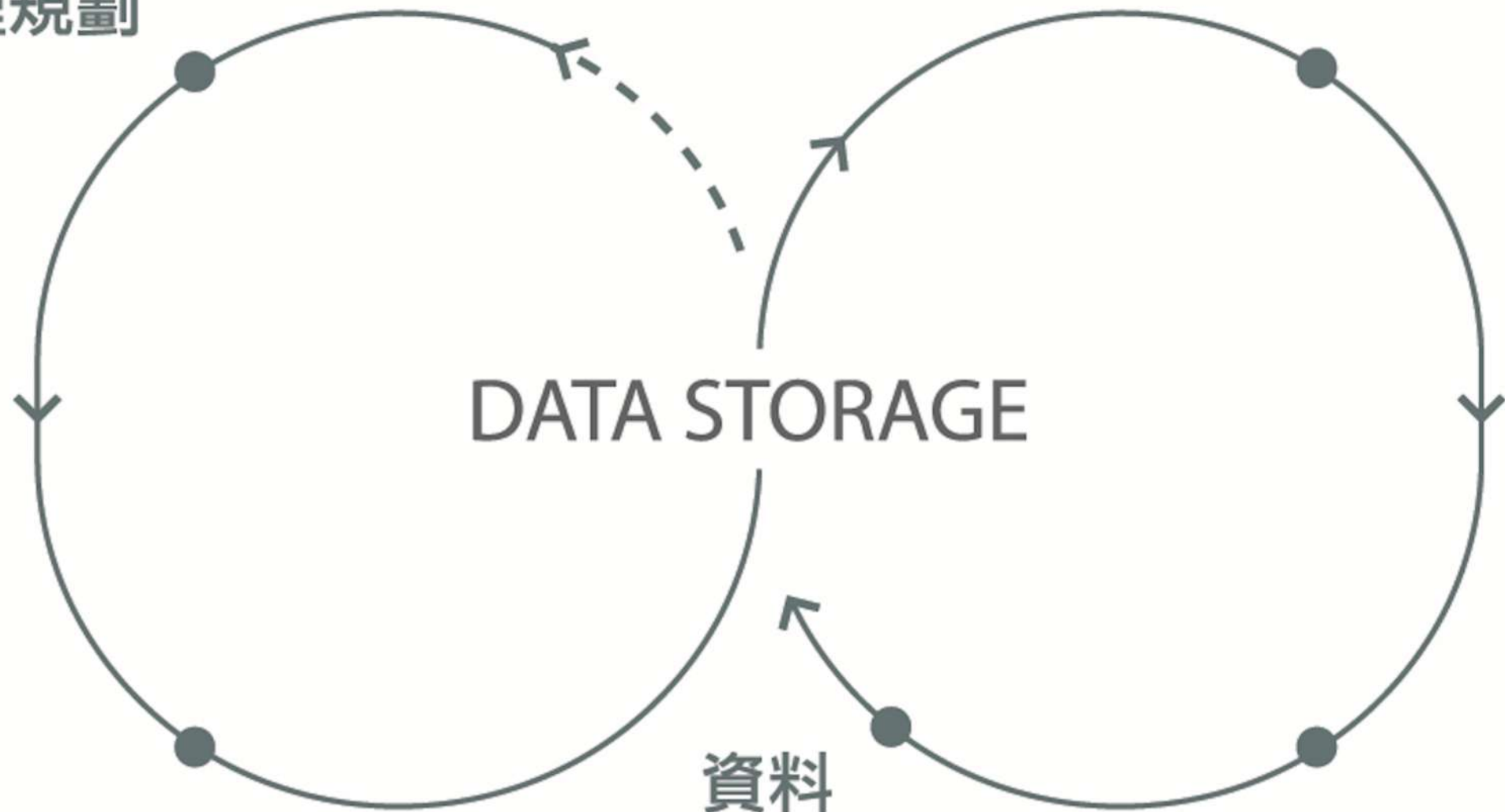
資料分析

DATA STORAGE

資料蒐集

資料
再利用

資料開放



Q & A

~ 開放資料流程 ~