

NATIONAL TAIWAN UNIVERSITY LIBRARY

臺灣發展研究資料管理政策之芻議

劉應琳

國立臺灣大學圖書館

2012.8.3



簡報大綱

- 前言
- 國外對於研究資料的管理
- 國內對於研究資料的管理
- 國內發展研究資料管理政策之芻議
- 結語



當今的研究趨勢

● 資料量遽增

- 隨著科技進步，研究設備（衛星、感測網等）分秒帶回資料類型的轉變；原生數位（born digital）的資料量遽增。
- 資料氾濫（data deluge）的時代。

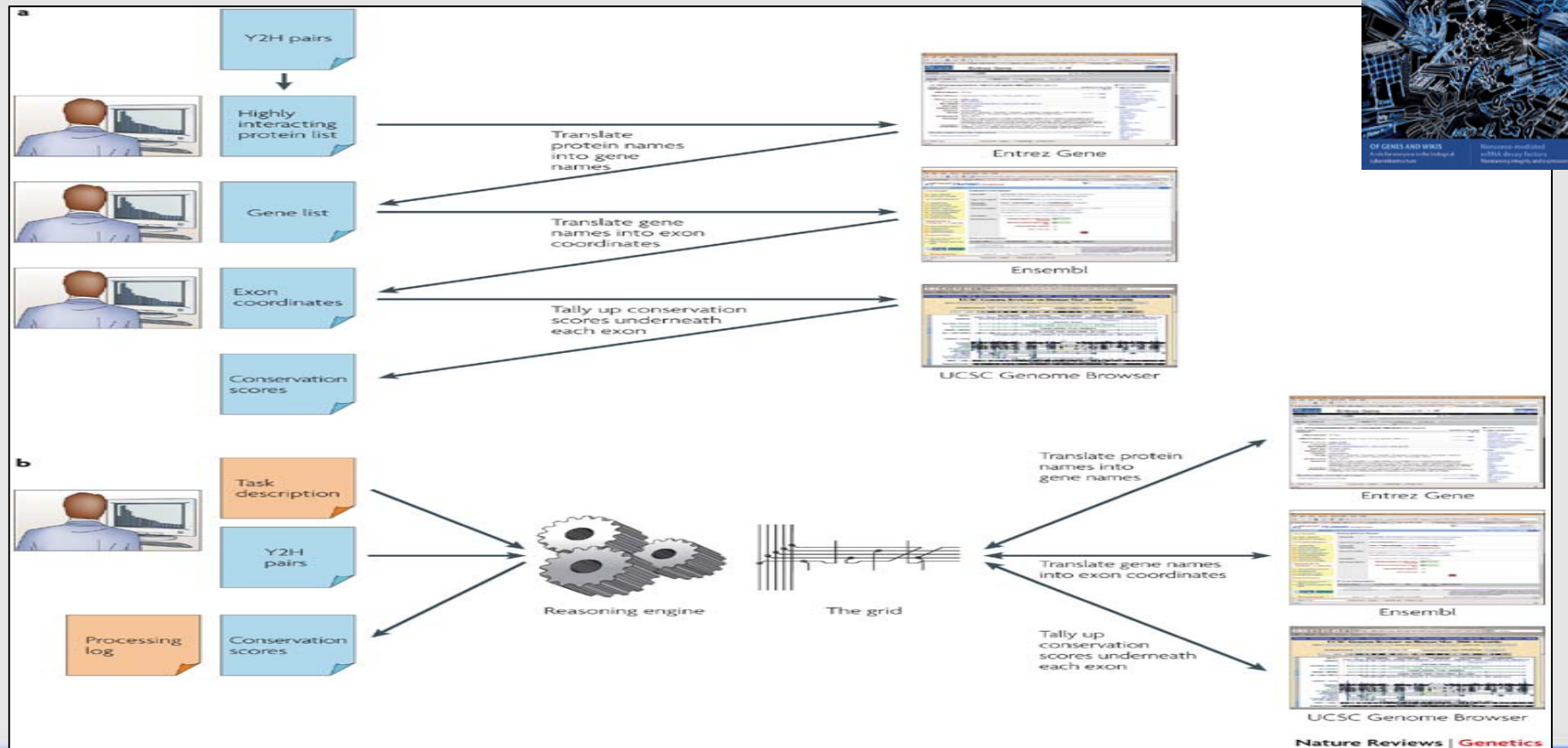
● 跨國與跨界合作研究

- 跨國或跨界合作引領的研究方向已成為新契機。
- 這些全球合作的科學家或研究人員多來自不同的學科背景，在線上搜尋資料時，往往會在不同的工具與傳輸協定之間感到困惑。



當今的研究資料—生物資訊學領域

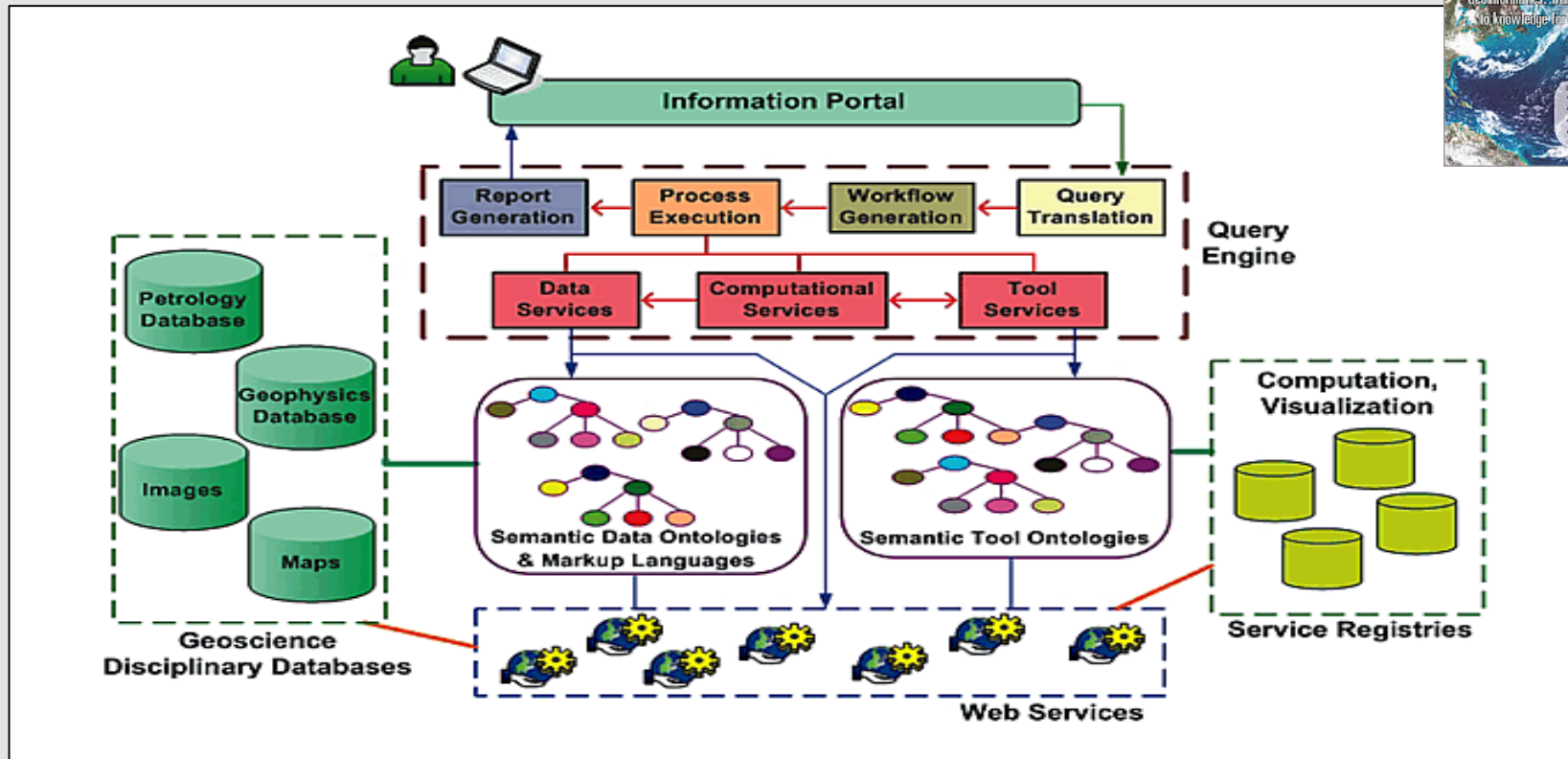
The process of **bioinformatics** research now and in the future



Stein, L. D. (2008). Towards a cyberinfrastructure for the biological sciences: progress, visions and challenges. *Nat Rev Genet*, 9(9), 678-688.

當今的研究資料—地理資訊學領域

Geoinformatics: transforming data to knowledge for geosciences



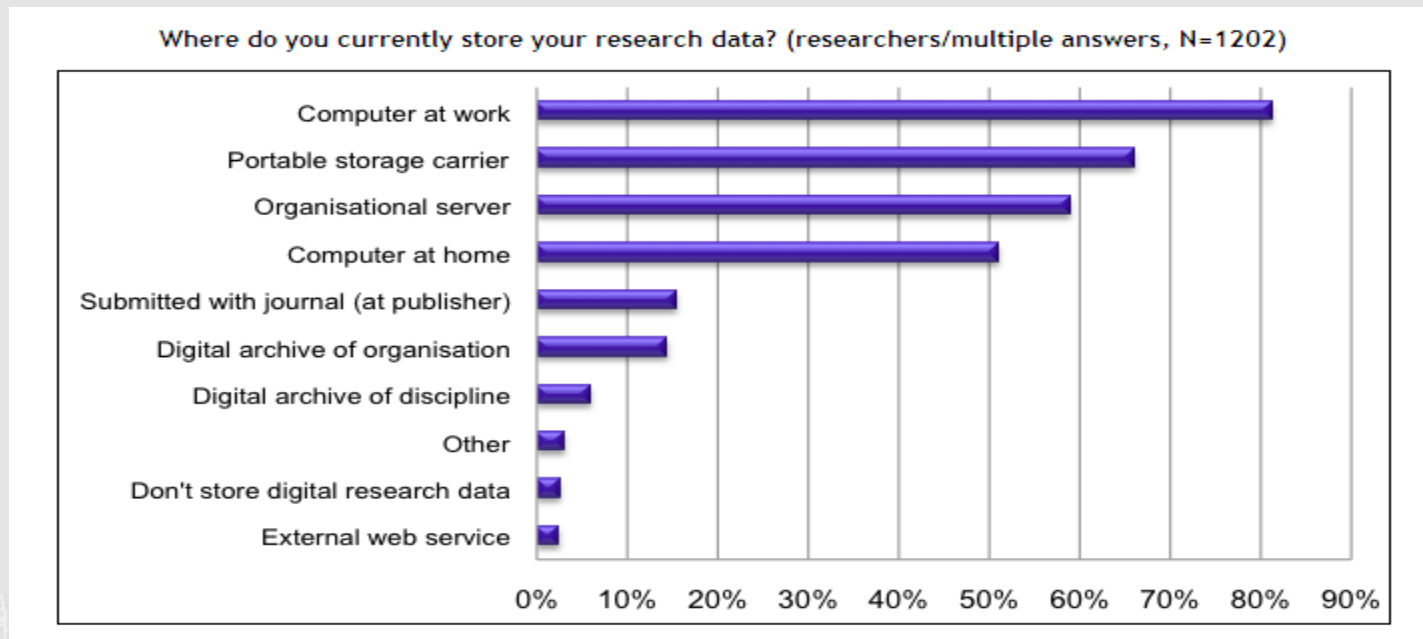
Sinha, A. K., Malik, Z., Rezgui, A., Fox, D. L., Barnes, C. G., Lin, K., . . . Zimmerman, H. (2010). Geoinformatics: Transforming data to knowledge for geosciences. *GSA Today*, 20(12). doi: 10.1130/GSATG85A.1

研究資料大多存放何處？

依據歐盟（European Union）贊助執行的 PARSE.Insight 計畫研究結果：

Q: 全世界跨領域1,200名科學家如何處理研究產出資料？

A: 研究資料分散，80%只存在研究人員自己的電腦中。



PARSE.Insight consortium. (2010). PARSE.Insight: Science Data Infrastructure Roadmap.

研究資料可以度存於何處？

- 研究資料的度存，依功能性主要可分為三種類型：
 - **研究型資料庫** (Research data collections)：
一個（或以上）研究計畫的產物，
對於資料處理與管理的程度有限。
 - **資源/社群型資料庫** (Resource or community data collections)：
主要在服務特定或單一的研究社群，
一般很難預測可以經營或維持多久。
 - **參考型資料庫** (Reference data collections)：
主要目的在於服務大規模學術社群。

研究資料可以度存於何處？（續）

- **研究型資料庫（Research data collections）：**
 - **Earth Sciences:**
FLOSS (Fluxes Over Snow Surfaces) Project
<http://www.eol.ucar.edu/isf/projects/floss/>
 - **Biology:**
Ares Lab yeast intron database
http://compbio.soe.ucsc.edu/yeast_introns.html
 - **Economics:**
University of Virginia conducted online game-like programs
<http://people.virginia.edu/~cah2k/reserach.html>



研究資料可以度存於何處？（續）

研究型資料庫（**Research data collections**）：
資料未必公開，且計畫結束後可能不復存在

Oops! This link appears to be broken.



Did you mean: people.virginia.edu/~cah2k/research.html

Additional suggestions:

- Go to virginia.edu
- Go to people.virginia.edu
- Go to sitemap www.virginia.edu/sitemap.html
- Go to people.virginia.edu/~cah2k/
- Search people.virginia.edu for [~cah2k research](#)
- Search on Google:



研究資料可以度存於何處？（續）

- **資源/社群型資料庫（Resource or community data collections）：**
 - **Earth and Space Sciences:**
NASA's Earth Science Enterprise (ESE) – ten discipline-specific data centers- Distributed Active Archive Centers (DAACs)
<http://nasadaacs.eos.nasa.gov/about.html>
 - **Biology:**
The Maize Genetics and Genomics Database (MaizeGDB)
<http://www.maizegdb.org/>
 - **Ocean Science:**
Ocean Drilling Program Database (Texas A&M University)
<http://www-odp.tamu.edu/database/>



研究資料可以度存於何處？（續）

資源/社群型資料庫（Resource or community data collections）：
資料庫分散，且多需社群權限才可利用

[Moratorium](#) | [Layout](#) | [Depths](#)

Username

Password

Database

Data Request System

About Moratorium Access

[Press keypad **Home** to return to the top of this form.]

To give researchers reasonable time to analyze and publish scientific results, the data collected on each cruise of the Joides Resolution is subjected to a one-year moratorium. In this moratorium period, access to the coring and cruise data is restricted to those persons who participated on that cruise. Upon expiration of the moratorium period, all of the coring data and associated scientific data for that cruise become publicly available.

The access code you specify above is temporarily recorded by your web browser. The access code remains in effect until you exit your web browser. After exiting your browser, you will need to re-visit this page to re-establish access to moratorium data. To protect your personal moratorium access code, be sure to exit your browser when you are done downloading data.

If you do not wish to change your moratorium access configuration at this time, press the **Cancel** button. You will be returned to the page you were viewing previously.

Moratorium Access Code

Moratorium access is controlled by a moratorium access code. This code has been assigned for every person who has ever made a sample request of the various ODP core repositories. If you do not know your moratorium access code or wish to change it, contact JanusWeb@iodp.tamu.edu.

The access code is case sensitive. Check the **Caps Lock** state on your keyboard if you have any trouble accessing moratorium data.



研究資料可以度存於何處？（續）

- **參考型資料庫（Reference data collections）：**

- **Space Science:**

NASA's Earth Observing System
Data and Information System (EOSDIS)

<http://nasadaacs.eos.nasa.gov/>

- **Biology:**

The Protein Data Bank

<http://www.pdb.org>

- **Social Science:**

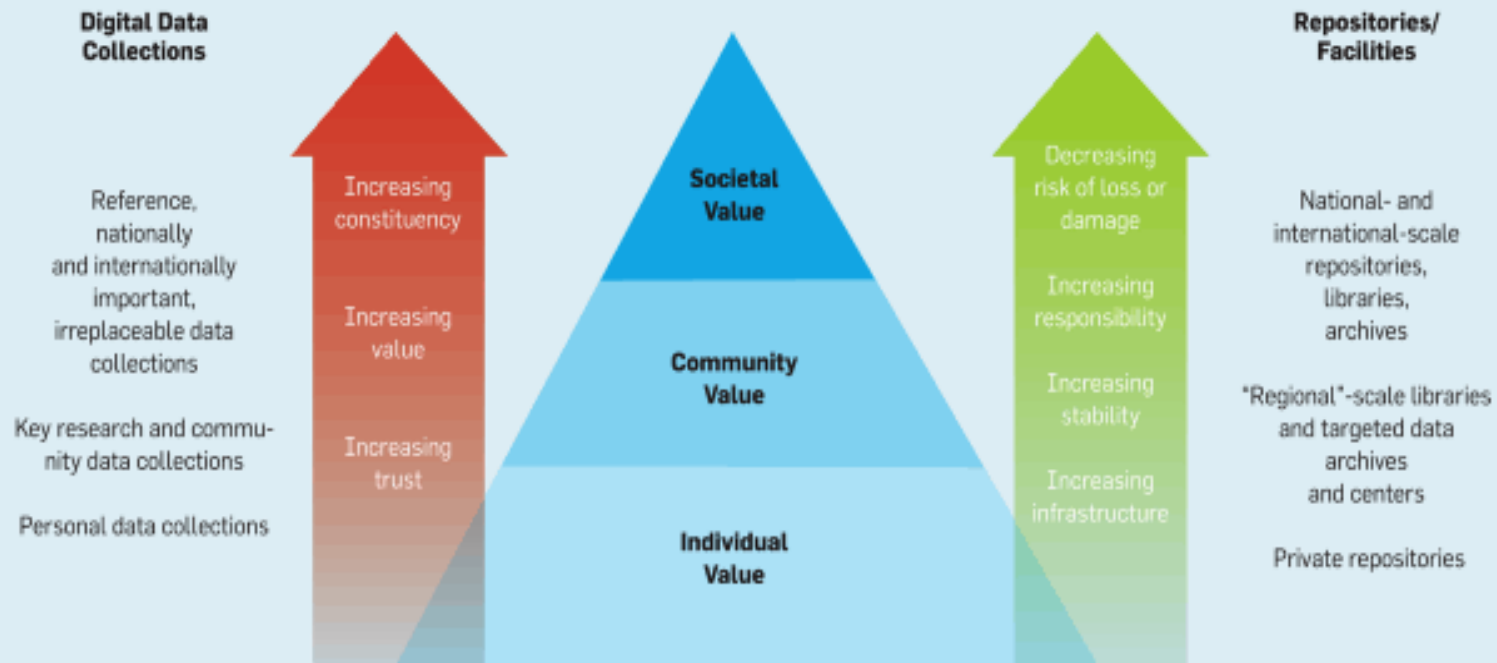
The Panel Study of Income Dynamics (PSID)

<http://psidonline.isr.umich.edu/>



研究資料建議度存在…

研究產出資料建議應度存在「**國際/國家級的參考型資料庫**」中，不僅可降低遺失或毀損的風險，並將大為提升後續價值。



Berman, F. (2008). Got data?: a guide to data preservation in the information age. *Commun. ACM*, 51(12), 50-56. doi: 10.1145/1409360.1409376

研究資料的管理現況

- 未必都有完善保存機制：策略、政策、技術、經費等。
- 難以整合：部分領域即使各自維護有相關資料庫，未必公開，或都有提供健全的資料描述（**identifier**、**metadata**）、利用資料的軟體工具，以及採用與國際接軌的度藏標準，也因此於整合與開放上增添難度。
- 缺乏跨資料庫平台：遑論管理與服務人員。
- 不易取得：相關領域研究人員若要利用公部門經費補助產出的研究資料，非但不易取得，甚至反而必須付費透過營利性的資料庫公司取得。



研究資料：建議開放

- 資料共享有助於科學的應用：激發創造力。
- 資料共享讓資源使用更具效率：降低成本、讓資源使用最大化、更正錯誤分析、減少人體研究的負擔。
- 共享有助於科學社群氣候的形成：將提升研究人員致力於資料的可信度與生產力。
- 共享是一種責任與需要：避免重複研究帶來的花費；研究資料量大，無法納入成果報告中。

Fischer, B. A., & Zigmond, M. J. (2010). The Essential Nature of Sharing in Science. *Science and Engineering Ethics*, 16(4), 783-799. doi: 10.1007/s11948-010-9239-x

Schroder, P., & Uhler, P. F. (2008). Open Data for Global Science. In B. Fitzgerald (Ed.), *Legal Framework for e-Research: Realising the Potential*. Australia: Sydney University Press.

NATIONAL TAIWAN UNIVERSITY LIBRARY

國外對於研究資料的管理



國外相關單位的政策與策略

一. 國際層級

二. 國家政府層級

三. 研究經費補助單位層級

四. 研究社群層級

五. 出版社層級

六. 機構層級



一、國際層級

- 為推動開放研究資料提供強而有力的基礎
- 相關指標性宣言與文件
 - ▶ 馬克斯普朗克協會 Max Planck Society (2003)
Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (簡稱柏林宣言)
 - ▶ 經濟合作開發組織 OECD (2004)
OECD Principles and Guidelines for Access to Research Data from Public Funding
 - ▶ 聯合國教科文組織 UNESCO (2004)
Policy Guidelines for the Development and Promotion of Governmental Public Domain Information
 - ▶ 歐盟 European Union (2010)
Riding the Wave: How Europe can Gain from the Rising Tide of Scientific Data



二、國家政府層級

- 從國家研究體系架構的角度，指示資料共享服務的發展目標與方向。
- 案例：
 - ▶ 英國：

聯合資訊系統委員會JISC（Joint Information Systems Committee）致力於提供一個公開取用的國家級資料中心，英政府責成 Research Information Network（RIN）提出了一系列原則，可作為操作指南。
 - ▶ 澳洲：

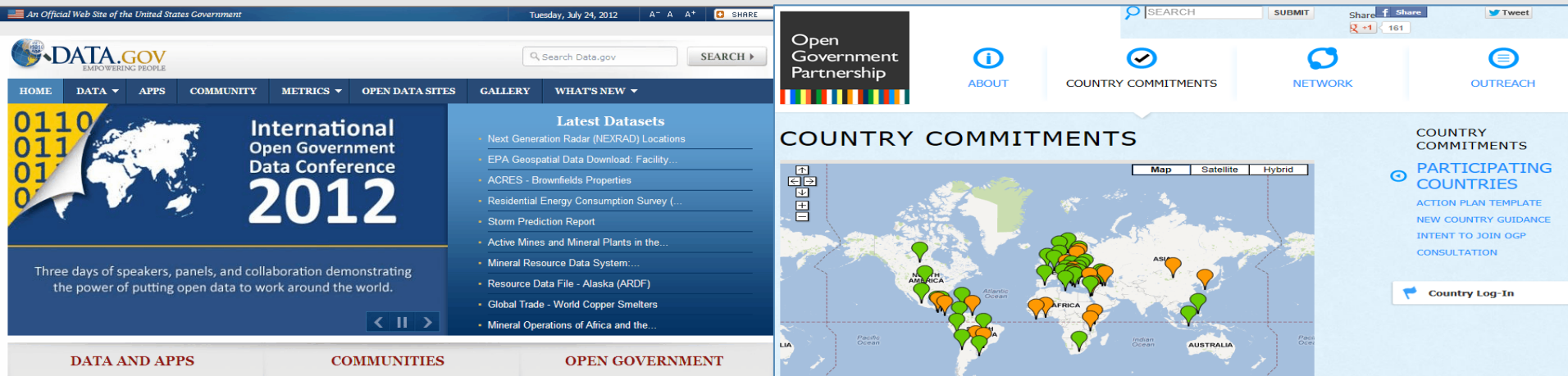
2004、2006年政策報告中指示澳洲應有研究資料共享空間（Australian Research Data Commons）的資訊架構，以確保資料共享與利用。
 - ▶ 美國：

公開政府計畫（Open Government Initiative）

二、國家政府層級（續）

- 美國歐巴馬總統一上任即以開放資料（Open Data）放在優先位置，致力於推動開放政府計畫（Open Government Initiative），期能達成目標：
 - 公開政府將可促進歸責性（accountability），進而提高效能。
 - 提升施政透明度（transparency），以確保資訊公開能被最大程度地利用。
 - 政府應在大眾參與之下，建立政策與規劃。
- 成果：Data.Gov 入口網站（2009 年~）

Open Government Partnership（開放政府夥伴關係聯盟 2011.9~）



The screenshot displays the Open Government Partnership website interface. At the top, there is a navigation bar with 'SEARCH', 'SUBMIT', and social media sharing options. Below this is a main header with 'Open Government Partnership' and navigation links for 'ABOUT', 'COUNTRY COMMITMENTS', 'NETWORK', and 'OUTREACH'. The central content area features a 'COUNTRY COMMITMENTS' section with a world map showing various countries marked with colored pins. To the right of the map, there is a sidebar with 'PARTICIPATING COUNTRIES' and links for 'ACTION PLAN TEMPLATE', 'NEW COUNTRY GUIDANCE', 'INTENT TO JOIN OGP', and 'CONSULTATION'. On the left side of the page, there is a 'Latest Datasets' section listing various data sources like 'Next Generation Radar (NEXRAD) Locations' and 'EPA Geospatial Data Download: Facility...'. At the bottom of the page, there are three main categories: 'DATA AND APPS', 'COMMUNITIES', and 'OPEN GOVERNMENT'.

Open Government Partnership. (2012). Country Commitments: United States Retrieved July 25, 2012, from <http://www.opengovpartnership.org/countries/united-states>

二、國家政府層級 (續)

- 成果：Open Government Dashboard 2012 (since 2010.4)

KEY:

- Meets Expectations
- Progress Toward Expectations
- Fails to Meet Expectations

Agency	High-Value Data	Data Integrity	Open Webpage	Public Consultation	Overall Plan	Formulating the Plan	Transparency	Participation	Collaboration	Flagship Initiative
Agency for International Development	■	■	■	■	■	■	■	■	■	■
Department of Agriculture	■	■	■	■	■	■	■	■	■	■
Department of Energy	■	■	■	■	■	■	■	■	■	■
Department of Education	■	■	■	■	■	■	■	■	■	■
Department of Commerce	■	■	■	■	■	■	■	■	■	■
Department of Defense	■	■	■	■	■	■	■	■	■	■
Department of Health and Human Services	■	■	■	■	■	■	■	■	■	■
Department of Homeland Security	■	■	■	■	■	■	■	■	■	■
Department of Housing and Urban Development	■	■	■	■	■	■	■	■	■	■
Department of the Interior	■	■	■	■	■	■	■	■	■	■
Department of Justice	■	■	■	■	■	■	■	■	■	■
Department of Labor	■	■	■	■	■	■	■	■	■	■
Department of State	■	■	■	■	■	■	■	■	■	■
Department of Transportation	■	■	■	■	■	■	■	■	■	■
Department of the Treasury	■	■	■	■	■	■	■	■	■	■
Department of Veterans Affairs	■	■	■	■	■	■	■	■	■	■

White House. (2012). Open Government Dashboard Retrieved July 24, 2012, from <http://www.whitehouse.gov/open/around>

三、研究經費補助單位層級

- 研究經費補助單位，對於研究產出的管理範圍，已經擴大至「資料」類型。
- 案例：
 - ▶ 英國：
 - 英國研究委員會 (Research Councils UK, RCUK)
 - 維康信託基金會 (Wellcome Trust)
 - ▶ 美國：
 - 國家衛生研究院 (National Institutes of Health, NIH)
 - 國家科學基金會 (National Science Foundation, NSF)
 - ▶ 澳洲：
 - 澳洲研究院 (Australian Research Council, ARC)
 - 國家健康與醫療研究院
(National Health and Medical Research Council, NHMRC)

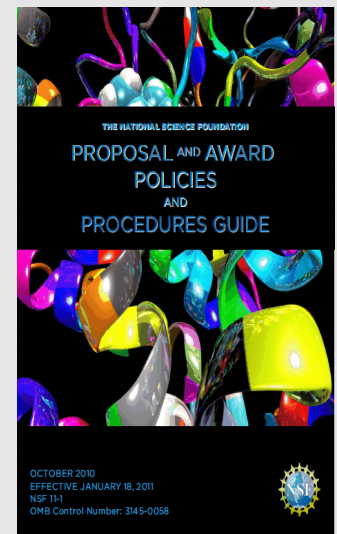


三、研究經費補助單位層級（續）

- 美國國家科學基金會 (NSF) 已經修訂獎助補助辦法與申請說明

(Proposal and Award Policies and Procedures Guide , PAPPG) :

要求所有在2011年1月18日之後向該會提出的研究計畫申請書，在申請補助經費階段，就要附上一份不須超過兩頁的資料管理計畫 (data management plan) ，說明將如何取用、典藏以及開放共享計畫產出資料。



三、研究經費補助單位層級（續）

NSF--Data Management Plan (DMP)

- 必須說明該研究計畫在傳佈與分享研究成果上，業已遵守NSF的政策。
- 每位處長或學門召集人或有針對該領域社群提出特別的要求，於撰寫計畫書之前務必確認詳閱各處室網頁的相關說明。
- 資料釋出政策（**data sharing policy**）：
鼓勵計畫主持人在不增加經費負擔之下、於合理的時間範圍內，將包括以NSF經費補助下產出或採集而來的樣本等一手資料，釋出給其他研究人員利用。

四、研究社群層級

- 一般學術社群多會參考該領域研究相關倫理與常規，某些研究社群已經相當重視資料管理議題，尤其是大量產出資料，且依賴資料進行研究的學術領域。（如生物資訊學、地理資訊學等）
- 研究領域的屬性與文化，可帶動研究人員自發性公開資料的風氣。
- 各學術領域案例
 - ▶ 國際地球觀測組織（Group on Earth Observations, GEO）
 - ▶ 人腦計畫（Human Brain Project）
 - ▶ 國際核苷酸序列資料庫合作計畫（International Nucleotide Sequence Databases Collaboration, INSDC）

五、出版社層級

- 出版社對於作者發表文章中相關研究資料的政策，多收錄於期刊徵稿說明文件中。
- 在50種高被引期刊的資料政策中，88%的期刊會要求作者公開資料。
 - 附上可於必要時提供資料的聲明。
 - 放在公開典藏庫。
- 案例：Nature期刊出版社 (<http://www.nature.com/authors/policies/availability.html>)

authors & referees

Search go Advanced search

authors & referees > Policies > Availability of data & materials

Site content

Homepage

Policies

- l. Publication ethics
- l. Bioethics
- l. Availability of data & materials
- l. Peer-review policy
- l. Embargo
- l. Corrections
- l. License to publish
- l. Feedback

Author resources

- Peer review
- Nautilus blog
- Peer-to-peer blog

Site map

Availability of data and materials

Availability of data and materials

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. Therefore, a condition of publication in a Nature journal is that **authors are required to make materials, data and associated protocols promptly available to readers without undue qualifications.** Any restrictions on the availability of materials or information must be disclosed to the editors at the time of submission. Any restrictions must **also** be disclosed in the submitted manuscript, including details of how readers can obtain materials and information. If materials are to be distributed by a for-profit company, this must be stated in the policies should contact the chief editor of the journal (or the chief biology/chief physical sciences editors in the case of *Nature*). In cases where editors are unable to resolve a

SUBMIT NOW

ADVERTISEMENT

a nature conference

2012 Nature / Institute for Genomic Medicine Symposium

Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H., & Ioannidis, J. P. A. (2011). Public Availability of Published Research Data in High-Impact Journals. *Plos One*, 6(9), e24357.

六、機構層級

- 許多研究機構已致力於推動開放研究產出，如機構典藏（IR）的建置，但重心仍在文獻類型的研究成果，尚未普及至研究資料。
- 隨著國家或研究經費補助單位對於研究資料的關注，各個機構也開始配合發展自身的資料管理政策。
- 多納入學校機構內的**研究發展**或**資訊服務**內推動。
- 案例：英國愛丁堡大學



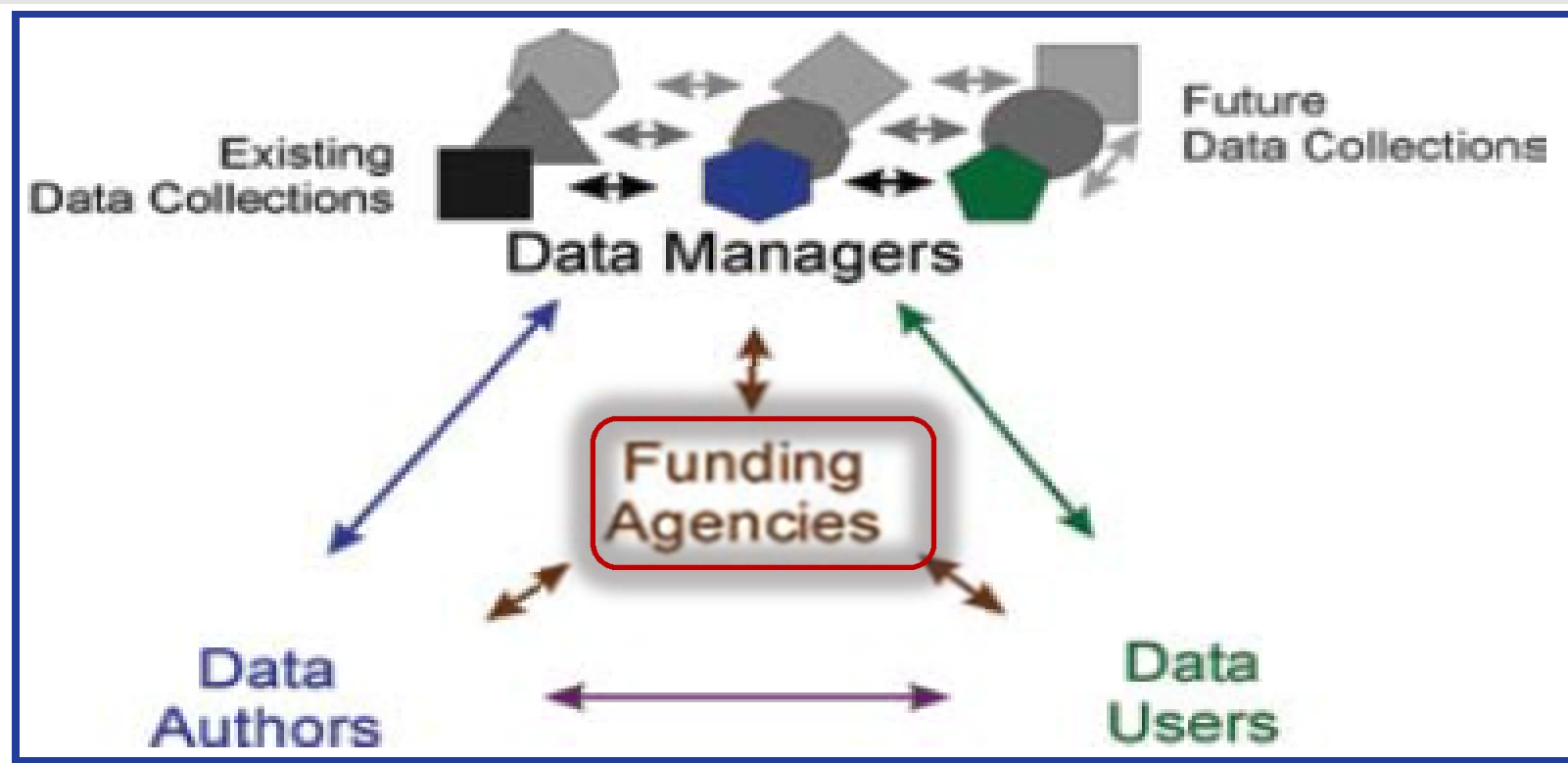
小結

- 各國資料管理政策沒有一定的模式，可能與補助研究經費作法差異有關。
- 研究資料的管理，可歸納為以下三種層次：



推動面：研究經費補助單位扮演關鍵角色

Structure of the Digital Data Collections Universe



US National Science Board. *Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century*. September 2005. p.18

NATIONAL TAIWAN UNIVERSITY LIBRARY

國內對於研究資料的管理



國內相關單位的政策與策略

一. 國家政府層級

二. 研究經費補助單位層級



一、國家政府層級

- 相較之下，國內開放政府資料尚處於起步階段。
- 臺北市政府於2011.9推出「臺北市政府公開資料平台」。
- 行政院科技會報辦公室於2012.1.18舉行「公開資料增值推動策略會議」，討論相關議題，包括以商業模式開放政府資料增值。

Data.Taipei
臺北市政府公開資料平台

[Apps.Taipei](#) [網站地圖](#)



[回首頁](#) | [關於我們](#) | [資料目錄](#) | [AppStar入圍作品](#) | [開發指南](#) | [最新消息](#) | [常見問答](#) | [使用規範](#)

OPEN DATA
臺北市政府公開資料平台

資料搜尋：



作品展示

最新資料

熱門資料

二、研究經費補助單位層級

國內目前多屬個別研究計畫層次，可歸納為三種類型：

1. 研究經費補助產出之研究資料必須公開

- 未特別針對數位研究資料規範，但現行辦法有其涵括性。(如文化資產保護法)

2. 研究經費補助計畫目的即為建置公開資料庫

- 自然科學領域：海洋反射震測資料庫、

臺灣大氣化學與環境資料庫

- 生物科學領域：整合臺灣生物多樣性資訊網站

3. 以徵集資料為主的資料中心

- 中研院調查研究專題中心：學術調查研究資料庫 (SRDA)



國內對於研究資料管理的落差

- **欠缺政策：**

國內尚未有國家/研究經費補助單位層級的全面性研究資料管理政策。

- **欠缺誘因：**

目前的相關努力多屬學術服務，並未列入學術表現的評鑑機制內，研究人員缺乏釋出資料的充分誘因，多無意願申請計畫建置資料庫或進行整理資料工作。

- **欠缺支援：**

國內研究與經費贊助結構，未能於體制面技術支援計畫結束後的資料管理，遑論永續經營的國家/機構級資料中心。



NATIONAL TAIWAN UNIVERSITY LIBRARY

臺灣發展研究資料管理之芻議



建議

- 政府：正視研究資料的重要性，參考國際間作法，積極研擬推動臺灣相關政策與執行細則。
- 經費補助單位：召集或補助各領域相關學者專家，組成專案小組，瞭解研究學門範圍研究資料現況——
 - ▶ 調查已補助經費計畫產出之研究資料現況：
是否已有建置相關資料庫、維護與管理單位等。
 - ▶ 瞭解學門領域範疇內產出之研究資料現況：
是否已有相關資料管理政策或詮釋資料標準等。
 - ▶ 獎勵與規範雙管齊下：
 - 獎勵：建立研究體系實質獎勵制度，如資料引用次數納入學術評鑑範圍，強化研究人員公開研究資料的動機。
 - 規範：受補助計畫經費產出的研究資料必須提出資料管理計畫。

建議

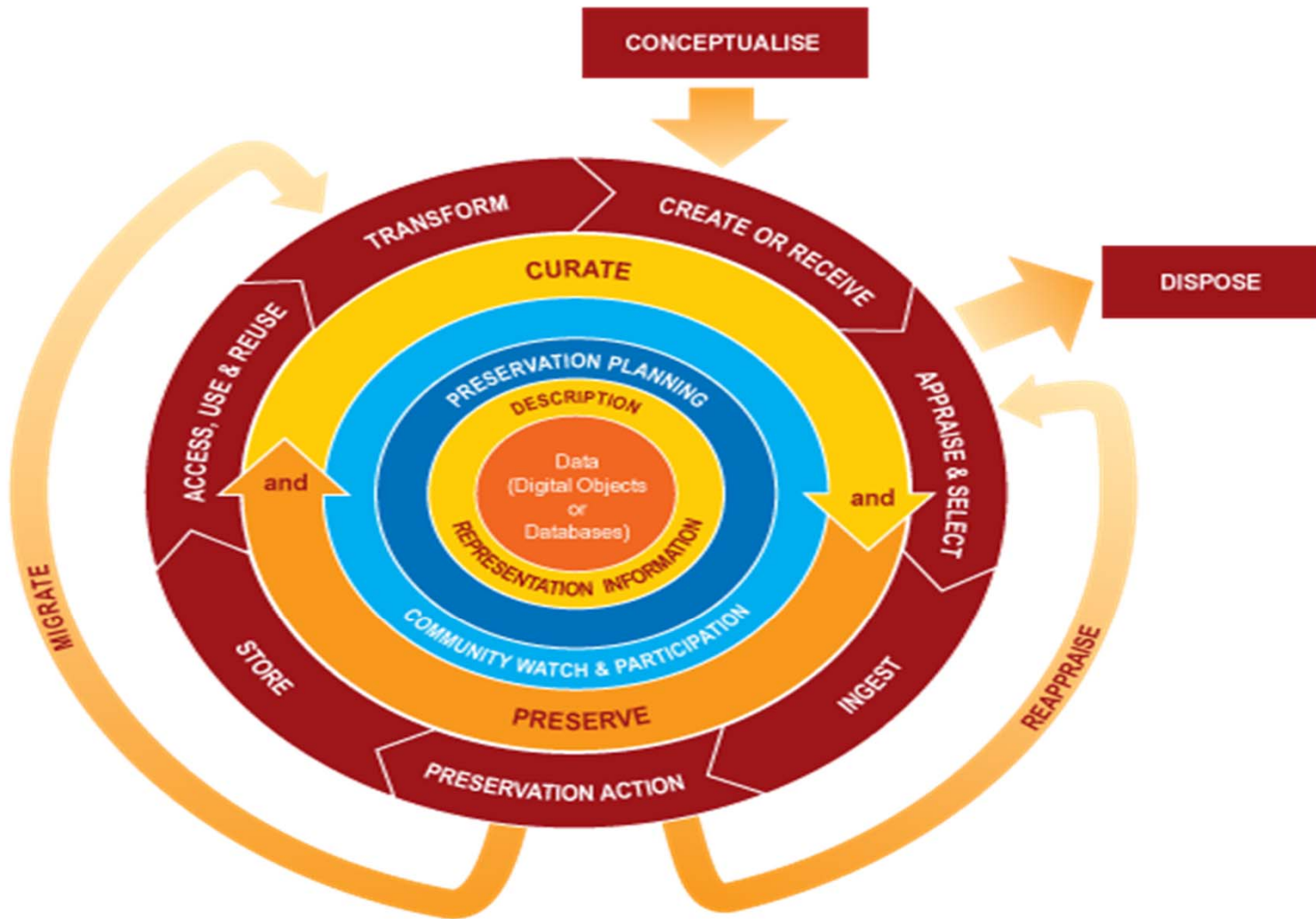
- 研究人員：

配合政策同意公開與開放研究資料，建立資料管理計畫，確認採用相關資料管理標準，並將資料存放在永續經營的典藏庫中。

- 機構單位：

支持機構內執行研究資料管理，如設置資料中心。

從研究發展或資訊服務體系下，研擬提供研究人員於資料自產出至再利用過程中，典藏（**preserve**）與度用（**curate**）等相關技術與服務的協助。





謝謝聆聽 敬請指教

圖片出處：<http://www.e-news.com.bd/2010/12/21/bangladesh-maldives-will-not-go-under-the-sea/water-circle-hands/>