臺灣開放科學研究成果彙編

研究者資料實踐與研究資料基礎建設

主編
鄭瑋

編輯團隊（依姓名首字筆畫排列）
丁昱寧／高庭萱／黃郁雯／黃寶霈／楊郁如

目 次

關於編輯與作者　4

CHAPTER 1　緒論：資料密集視角的開放科學

13　壹、資料密集視角的開放科學：
　　研究資料開放與分享予科學界的意義
14　貳、認識研究資料與其生命週期
17　參、如何閱讀與誰適合閱讀這本書
21　參考文獻

CHAPTER 2　研究資料開放與再用實踐
——總論篇

25　壹、開放研究資料
29　貳、開放研究資料擇採模式（ORDAM）
45　參、章節小結
47　參考文獻

CHAPTER 3 研究資料開放與再用實踐——自然科學篇

51 壹、研究資料分享的正式管道：資料出版
52 貳、資料出版與資料論文出版生態
58 參、研究案例分享：臺灣自然科學學者資料出版的實踐與動機
70 肆、章節小結
72 參考文獻

CHAPTER 4 研究資料開放與再用實踐——人文與社會科學篇

77 壹、人文社科領域開放研究資料之實踐
78 貳、研究案例分享：犯罪學開放研究資料之實踐與再用
86 參、研究案例分享：中國文學學者研究之資料利用與移動
95 肆、章節小結：人社科學資料的豐富圖像與研究資料基礎建設之啓發
98 參考文獻

CHAPTER 5 研究通透度與可再現性

103 壹、研究通透度

111 貳、提升研究通透度與可再現性的國際組織與標準

122 參、研究通透度的實踐與利害關係人

125 肆、章節小結：
研究通透度與再現性為臺灣科研社群的共同使命

127 參考文獻

CHAPTER 6 新興科技輔助增強研究通透度

135 壹、研究再現性危機

137 貳、研究案例：
區塊鏈技術輔助研究通透度

140 參、為 BLT 打造良好的使用者體驗

150 肆、章節小結

152 參考文獻

153 結語

圖目次

31 圖一、影響研究者公開分享和使用研究資料的主要因素
57 圖二、資料論文出版生態圈
63 圖三、學者於資料論文投稿行為建立基礎之概覽示意圖
118 圖四、TOP Guidelines 的八項標準
144 圖五、調整前介面示意
144 圖六、調整後介面示意
145 圖七、BLT 研究紀錄儀表板
146 圖八、BLT 平台資料上區塊鏈頁面示意
147 圖九、BLT 平台權限管理頁面示意

表目次

66 表一、資料論文投稿行為之動機概覽
92 表二、資料摩擦力發生介面與資料移動種類
108 表三、各政府或學術單位對於不同 R-words 的定義示例（本書整理）
141 表四、原型設計歷程

關於編輯與作者

本報告由國立臺灣大學圖書資訊學系資料科學與資訊設計實驗室（DxLab）主持人，鄭瑋教授主編，內容為其研究團隊合力協助編輯與撰寫。

關於 DxLab（Data Science and Information Design Lab）

資料科學與資訊設計實驗室（DxLab）為鄭瑋博士所帶領的研究團隊，近年研究團隊成員致力於探索資料的永續發展、資料保護與社會議題。團隊致力於研究各學科領域中研究資料基礎建設與開放科學實踐的議題。同時，團隊也關注資料倫理、隱私保護、資訊安全意識推廣等相關課題。此外，研究範圍亦涉及公民科學的推展，企盼將科技參與與科學意識傳播至更多元的社會群體。

編輯與作者介紹

主編：鄭瑋　博士

鄭瑋現任職於國立臺灣大學圖書資訊學系。鄭瑋於美國匹茲堡大學資訊科學學院取得圖書資訊博士學位，研究興趣包含研究資料、資料隱私與保護、使用者資訊安全，以及再現性研究。鄭瑋的著作多發表於 PLOS ONE、Journal of the Association for Information Science and Technology (JASIS & T)、Journal of Informetrics 等資訊科學期刊。

執行編輯（依姓名首字筆畫排列）

◆**丁昱寧**　目前就讀國立臺灣大學圖書資訊學系研究所。研究興趣包含研究通透度、公民科學、科學傳播。

◆**高庭萱**　目前就讀國立臺灣大學圖書資訊學系研究所，研究興趣包含研究通透度、再現性以及資料分享。

◆**黃郁雯**　目前就讀國立臺灣大學圖書資訊學系研究所，大學時期取得國立臺灣科技大學設計系商業設計組輔系學位。研究興趣包含資訊設計、科學傳播、視覺修辭。

◆**黃寶霈**　目前就讀美國北卡羅萊納大學教堂山校區資訊學院博士班，畢業於國立臺灣大學圖書資訊學系研究所，研究興趣包含研究資料管理、資料分享與

出版、科學傳播。

◆楊郁如 目前就讀美國伊利諾大學香檳校區資訊學院博士班，畢業於美國卡內基梅隆大學碩士班、國立臺灣大學政治系與圖資系雙主修。研究興趣包含資訊倫理、資訊架構、隱私政策。

助理編輯（依姓名首字筆畫排列）

◆林語婕 目前就讀國立臺灣大學圖書資訊學系研究所。研究興趣包含知識管理、幽默研究、社會語言學等。

◆陳玟寧 畢業於國立臺灣大學圖書資訊學系與外文國語文學系雙主修。研究興趣包含科學傳播、數位修辭、隱喻研究等。

章節作者（依姓名首字筆畫排列）

◆石易儒 畢業於國立臺灣大學圖書資訊學系研究所。研究興趣為研究資料管理、資料公開與資料分享、開放政府資料與檔案學。

◆李建欣 目前就讀美國密西根大學安娜堡校區資訊學博士班，畢業於美國華盛頓大學資訊管理研究所、國立臺灣大學圖書資訊學系及政治學系。研究興趣包含公共衛生資料合作、社區健康資訊基礎建設、臨床研究資料共享等。

◆**林潔筠**　畢業於國立臺灣大學圖書資訊學系研究所，大學時期就讀政治大學中文系。研究興趣有資料分享與再用，以及知識管理。

◆**張韻琦**　畢業於國立臺灣大學圖書資訊學系研究所，研究興趣包含研究資料管理、資料治理、資訊設計。

◆**龔俐霏**　目前就讀國立臺灣大學圖書資訊學系研究所，亦於國家資通安全研究院人才培力中心擔任研究員。研究興趣包含資訊設計、使用者研究、可用式資安、桌上型兵棋推演執行計畫設計。

Chapter
1

緒論：資料密集視角的開放科學

章節作者：黃寶霈、鄭瑋
章節編輯：黃寶霈

章節地圖

- ◆ 資料密集視角的開放科學
- ◆ 認識研究資料與其生命週期
- ◆ 「如何閱讀本書」與「誰適合閱讀」

章節摘要

本章節從「資料密集」的研究視角來探討開放科學，闡述研究資料開放與分享對於科學界的重要意義。本章節帶領讀者透過認識研究資料的特性及其生命週期，讀者可以進一步了解資料在研究過程中扮演的角色。本章節接續介紹本書《臺灣開放科學研究成果彙編》的編排結構，以及預計面向之讀者，期待讀者透過閱讀本書，得以認識開放科學的概念與本土相關研究之進展。

壹、資料密集視角的開放科學：研究資料開放與分享予科學界的意義

開放科學（Open Science）係藉由公開研究成果、分享科研資源予社會大眾的行動，以致科學傳播的成效，常見的途徑包含研究資料分享與公開、研究過程與結果的通透性，以及由公民參與的科研活動等。

與開放科學並進的是資料密集典範的科學興起。在各研究領域於密集運算應用漸長下，科學發展已從過去的實驗、理論與運算等三波典範逐漸推進至第四典範（the fourth paradigm），即「資料密集」的科學（Hey et al., 2009）。受惠於電腦運算與儲存科技的進步，科學家自實驗器材、模擬、感測網路等極其多樣化的來源獲致各式各樣的資料，進入了上從研究素材蒐集、分析，下至研究成果發表、保存皆大幅仰賴資料的密集研究週期，進化到現今以「資料」為核心的階段。在這一典範下，即使是單一的實驗室也能迅速產出大量資料，強調了資料的品質與價值。

對科學的傳承來說，實踐透明且開放的研究資料便有其必要性，研究資料之開放與通透，成功反映一項科學是否具備足夠的可信度與課責性。理想中，國家應投入大量的資

金、人力與時間來改進不合標準的研究實踐、響應開放科學，如此才能系統性確保科研成果更加嚴謹、周全與可靠（rigorous, thoughtful, and dependable）；實際上，國家發展的進程之排序先後，以及資源資金總是有限的思量下，任何新的科研標準或是開放科學政策都可能成為國家預算或現有機構緊繃人力下的不合理負擔。鑑於國內外固有的學術評鑑制度往往直接影響學者個人的職涯進程，若開放科學與學術評鑑兩者脫鉤，則可能導致研究者在滿足學術評鑑績效之要求的同時，缺乏額外動力與資源去實踐研究資料的開放共享。因此，為了響應國際趨勢，鼓勵臺灣學者實踐開放科學，需要重新審視現有的評鑑與績效獎勵。唯有如此，才能在制度層面為研究者營造有利於開放科學實踐的環境，進而推動科學的傳承與創新。

《臺灣開放科學研究成果彙編》一書從資料密集典範的觀點出發，向科學讀者闡述研究資料在開放科學中的重要性，並揭示臺灣研究者在研究資料管理、分享、開放、出版與再用等實務面向上，所面臨的機會與挑戰。

貳、認識研究資料與其生命週期

開放科學的概念包含研究成果的公開、研究過程的透

明，以及公眾參與等面向。在這些面向中，「資料」扮演著關鍵的角色。本書特別聚焦於「資料」的視角，深入探討研究資料的特性與管理。不過，研究資料的樣態十分多元，依問題意識、方法論、資料獲取途徑、產製方式、學科領域、時空背景、人員判定等而有所差異，然其本質與重點皆環繞著「為支援研究所需」的目的，可被用於驗證研究結果。學者 Borgman（2015）加以整理美國科學委員會對於研究資料的分類後，提出以下四種資料類型：

1. **觀測型資料（Observational Data）：**觀測型資料是研究者利用儀器捕捉事實或現象發生以進行識別、標註或紀錄的資料。具體的例子包含自然科學領域的天氣觀測數據、社會科學領域的民族誌、訪談資料等。這些研究資料往往捕捉自特定時空脈絡，因難以再現而被認定為擁有最高的保存價值。
2. **電腦運算資料（Computational Data）：**電腦運算資料係執行電腦模型、模擬或工作流程中所產生的產物，最常見於物理與生命科學領域，然亦存於社會與人文科學中。例如在生命科學領域，科學家則可能利用電腦模擬蛋白質摺疊的過程，以了解其結構；如經濟學家可能利用電腦建立複雜的計量經濟模型，用以分析

與預測經濟現象；而語言學家則可能利用電腦進行大規模的語料分析，以探究語言的結構與演變。為確保這類型資料未來能再利用，除了模型本身，亦需要軟硬體與輸入資料步驟描述等紀錄。

3. **實驗型資料（Experimental Data）：**實驗型資料是在實驗的控制條件下，透過系統化的操作與測量實驗結果所產生的資料。這類資料在自然科學與社會科學領域中都有廣泛的應用。實驗過程中，研究者透過操控不同之實驗條件，收集接受不同操弄處理（treatment）的數據。實驗型資料的品質與可靠性，以及是否容易再現，往往取決於實驗設計的嚴謹性。
4. **紀錄（Records）：**不易納入上述三種類型，但關乎於現象或人類活動的資訊。

各種類型的研究資料，其自身的生命週期由研究者與資料間的互動所形塑而成。英國資料檔案館將資料生命週期劃分為六階段：資料創建、資料處理、資料分析、資料保存、資料近用、資料再利用（UK Data Archive, 2013），即資料因應學者的研究計畫而萌生，於產製後歷經描述、處理、分析、保存、近用到分享再利用各階段，每一階段間環環相扣，影響後續學者如何看待資料對自身的意義。

不過，值得一提的是，資料並非一個單純且具有固有本質的物件，資料的角色與資料的產生過程有關，其存在於背景脈絡中，由該背景與旁觀者的觀點賦予資料對應的意義，因此，有學者認為，比起說明「什麼是研究資料」的 What 問題，「資料在什麼時機可被視為研究資料」的 When 問題或許更為恰當（Borgman, 2015）。

參、如何閱讀與誰適合閱讀這本書

本書彙編了國立臺灣大學圖書資訊學系——資料科學與資訊設計實驗室（Data Science and Information Design Lab, DxLab）於 2018 至 2023 年在開放科學議題上的研究成果，這些成果來自於兩個國家科學委員會（National Science Council）的專題獎勵計畫：「邁向研究資料基礎建設：資料共享實務與資料之學術影響力探索」和「開放科學協助永續發展：加強科學成果再現性的共同設計」。研究成果環繞著開放研究資料與研究通透度這兩個主題，旨在推動開放科學永續發展（Open science for sustainability, OSFS）。這一理念強調，開放科學的實踐不僅能促進學術界內部的知識交流，也能加強學界與非學界、自然科學與社會科學之間的對話，從而為建設永續社會提供必要的科學支援。

本書主要分成二大部份，第一部分《研究資料開放與再用實踐》包含第二章至第四章，旨在為讀者建立臺灣學者響應開放科學，在資料分享、公開、出版與再用上的實踐。第二章的「研究資料開放與再用實踐——總論篇」首先揭示開放研究資料的概念，再深入介紹《開放研究資料適應模型》（Open Research Data Adoption Model, ORDAM），該模型自前人研究文獻歸納「開放研究資料」之影響因子，讓讀者全面識別學者們對於開放研究資料的共享和再利用之關注與主張。

第一部分之第三章與第四章則分別以自然生科與人文社會科學此二大領域為例，為讀者揭開本土學者對於開放研究資料的相關實踐。在自然生科篇中，黃寶霈揭示資料分享實踐方式之一的資料出版特性，並聚焦至資料論文此一種資料出版方式，探究臺灣自然科學與生物、醫學學者在資料論文上的投稿動機。而在人文社會科學篇中，石易儒與林潔筠分別從犯罪學領域與中國文學領域著手，討論人社學科推動開放科學的複雜性。前者分析犯罪學的研究特性，闡述從資料視角切入的研究流程，並揭示犯罪學研究者與其使用的研究資料互動中所蘊藏的信任、契約、價值等要素；後者則闡述人文資料的選擇與集合涉及學者本身的學識與背景經驗之特性，以及資料於不同學者或技術間的轉移下所帶來的潛在資料「摩擦力」。

本書的第二部份《研究通透度與可再現性》包含第五章和第六章，為國內讀者介紹目前國際開放科學領域研究之新興領域——研究再現性（reproducibility）。早期學術社群的觀點普遍認為，研究者僅需確保資料的可取用性（accessibility），即是具備通透度。然而，近期的觀點則認為，開放研究資料雖是提升研究通透度的必要條件，但隨著學術誠信的概念逐漸成熟，提高研究步驟、分析、推論研究成果之間的通透度，讓研究中的決策透明化，才能更有助於增強研究的可再現性。國際頂尖期刊如 Nature 與 Science 皆定期匯集有關「再現性」之文章專欄，亦於期刊政策中加入檢驗文章研究通透度的要點，使這一概念在開放科學的實踐中逐漸顯現出其重要性。研究通透度與再現性對國內圖書資訊學領域的師生也許較為陌生，不過今日國際學術傳播、開放科學中，這些議題受到前所未有的關注。

本書第五章「研究通透度與可再現性」由丁昱寧與高庭萱的研究引出研究通透度的概念，並介紹目前提升研究通透度與可再現性的國際組織與標準，以及學術傳播中協助實踐的利害關係人，以期藉由體制的建立，共同推動更為通透的科研發展。第六章「新興科技輔助增強研究通透度」則與讀者介紹如何以新興科技輔助開放科學實踐。龔俐霏與張韻琦設計與測試迭代一款以區塊鏈技術為後台的研究歷程紀錄工具。歷經使用

者需求蒐集、初始設計、使用者測試、設計迭代之過程，兩人的研究揭露了以區塊鏈技術整合於開放科學工具上的困難與機會。

本書力求使國內圖書資訊學系大學部高年級學生與各領域碩士班學生易於閱讀和理解。內容亦面向學術圖書館館員，以及對相關議題學術研究與從業感興趣的讀者。根據每章節內容，也在此列舉說明可能感興趣的對象讀者：

（一）有意認識資料密集視角下的開放科學概念之讀者

（二）希望更了解科研誠信與研究通透度內涵之讀者，以及想探究當代學術機構如何展現研究通透度、並期待了解使用科技方法輔助研究通透度的實踐案例（本書以生醫資料與區塊鏈技術之應用研究為例）者。

開放科學並非單獨個體即可達成的目標，而是需要多方群體一同理解與實踐才能彰顯其價值。衷心期待讀者能透過本書認識開放科學的核心概念，了解目前臺灣學界如何響應與現況又遇到何種挑戰，點燃彼此在開放科學上的更多思考火花與發展能量。

參考文獻

Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked world.* Cambridge, MA: MIT Press.

Hey, T., Tansley, S., Tolle, K., & Gray, J. (2009). *The fourth paradigm data-intensive scientific discovery.* Redmond, WA: Microsoft Research

UK Data Archive. (2013). Data life cycle & data management planning. https://www.ukdataservice.ac.uk/media/187718/dmplanningdm24apr2013.pdf

Chapter
2

研究資料開放與再用實踐——總論篇

章節作者：黃寶霈、鄭瑋
章節編輯：陳玟寧

章節地圖

- ◆ 開放研究資料
- ◆ 開放研究資料擇採模式（ORDAM）
- ◆ 章節小結

章節摘要

本章概述研究資料開放的相關概念與影響因素，探討研究資料分享（data sharing）與開放研究資料（open research data）的定義與理念異同。隨後，本章深入介紹開放研究資料擇採模式（Open Research Data Adoption Model, ORDAM），總結影響研究者公開分享和使用研究資料的主要因素，包括：研究者背景、內外部動機、技術資源、信任、社會性影響、法律與規範、資料特性等，這些因素既可能成為驅動力，也可能成為阻礙。奠基於此發現，學界將能更深入理解如何增加分享和使用資料機制的驅動力，減少阻力，以實現開放科學的願景。

壹、開放研究資料

資源的共享是推動科學研究發展的關鍵因素。隨著進入資料密集型科學典範，開放研究資料（open research data，簡稱ORD）係研究者將伴隨研究過程中所產製的資料公開，供其他研究者或甚至社會大眾自由存取與利用，其潛在價值與話題性逐漸升高。具體來說，生物醫藥領域的研究往往需要龐大的資料集，而透過 ORD，研究者可以共享資料，有機會加速疾病機制的探索和新藥的開發。這一點在 COVID-19 大流行期間尤為顯著：各國研究者通力合作，迅速分享病毒基因組、流行病學資料等關鍵資源，迅速推動了疫苗和藥物的開發進度。正是基於開放科學在抗疫中所扮演的角色，聯合國教科文組織（UNESCO）在2021年11月發表了一份開放科學建議書，且在其中強調，開放科學是實現聯合國可持續發展目標的關鍵，呼籲各國儘速制定相關政策，促進科學知識和資源互惠。

然而，值得注意的是「研究資料分享」（research data sharing）與「開放研究資料」這兩個概念，在文獻中有時會交替使用，但實際上存在著一定的區別：資料分享的範疇相當廣泛，它涉及將資料提供給第三方使用的各種情況，包

含研究者傳送資料給他人、將資料存放於公開的資料庫、在網站上公布資料，或是作為研究論文的附加材料等。這些分享的方式展現了從非正式管道到正式管道的連續性光譜（Borgman, 2012）。由此可知，資料分享可涵蓋任何向第三方釋出資料以供他方使用的形式。

至於 ORD，在定義開放研究資料前，可先討論開放資料（open data）之概念。根據開放知識基金會的定義，「開放」是一種免除限制之下的近用、處理與分享的資料實踐行為（Open Knowledge Foundation, 2021）。開放資料具有以下特徵：

- 可近用：資料與其授權所需的資訊必須以可讀取、解析的方式易於取得，並且能夠透過常見的軟體工具存取。
- 機器可判讀：資料必須以開放格式提供，即該格式的規格是公開於大眾，且在使用上並無任何收費或其他限制。故資料能以任何方式進行編輯、轉換或分析。
- 可授權分享：資料的使用或傳播遵循開放的授權條款，不受任何限制。

開放資料的概念比起資料分享，更重於無障礙地存取與散佈資料。學術界對開放資料的理解通常圍繞在資料可以在

不受規範的情況下被更廣泛的群體使用。例如，有研究者認為開放資料是由公共經費支持產生的資料，其使用或傳播不受限制，且不涉及隱私或機密問題（Janssen et al., 2012）。此外，開放資料也被視為一種鼓勵政府組織釋出公共服務產生或收集的資料的策略，這些資料特徵通常是非特定個人且具客觀性。同時，這股鼓勵開放資料的精神，也同時鼓勵擁有資料的組織提供不受到著作財產權限制的資料整合與利用活動（Hossain et al., 2016）。總的來說，資料開放的本質是為使資料的後續利用不受限制，常聚焦於政府單位為回應公眾經費的運用而釋放其資料上。

研究資料分享與開放研究資料亦在核心理念和應用範圍上有所不同。資料分享主要是研究者間的資料交流的集合稱呼，而開放研究資料則強調資料的無障礙共享，以促進學術研究和社會進步。儘管如此，單純的分享或開放研究資料並不保證資料的有效再利用。許多情況下，分享的資料可能不符合再利用的需要，如格式不一致或缺乏必要的說明，這對於整合和應用資料造成了額外的挑戰。

為解決這些資料分享與再用之間的問題，Wilkinson 等學者於 2016 年提出 FAIR 原則（Findable, Accessible, Interoperable, Reusable）為資料的有效再利用提供了指引（Wilkinson et al., 2016）。FAIR 原則自提出後，迅速成為科學資料分享與再用

的國際標準，且被廣泛應用於各學科領域。許多國家的資料政策、資料中心和研究機構都採用 FAIR 原則作為資料管理的依據。根據 FAIR 原則，資料應為可查找的（Findable），可近用的（Accessible）、可互操作的（Interoperable）、可再用的（Reusable）。這意味著，當資料與其後設資料（metadata）能夠於公開的網站或是資料典藏庫（data repository）[1]中被查詢到（符合可查找的特性），使用者能夠以特定方式下載資料（符合可近用的特性），其格式和說明應遵循學科領域之標準，以便於整合進其他研究者的研究實踐當中（符合可互操作的特性）。最後，這些資料應以可信賴的開放授權方式釋出，確保其可再用性（符合可再用的特性）。對於 FAIR 原則的更詳盡介紹，可見本書第五章。

儘管學術界對開放研究資料有明確的定義和原則來指引研究者們的資料實踐，如何在這些實踐中有效地達到目標，同時滿足利害相關者的需求和期望，依然是一項不小的挑戰。這要求研究者不僅要理解和遵守這些原則，還需要積極參與建立和維護一個支持資料共享與開放的生態系統。

[1] 資料典藏庫（data repository）是一儲存、管理、維護、加值數位物件並提供檢索功能的平台。

貳、開放研究資料擇採模式（ORDAM）

在開放研究資料的實踐中，一個普遍的觀點是資料分享能促進學術的進步。然而，實際上，部分研究人員對於分享資料持有顧慮。研究指出，時間與資金的不足，以及資料公開的權利問題，是大多數研究者不願分享資料的主要原因（Tenopir et al., 2015）。

不少研究者擔心，若自身研究成果在期刊獲得發表之「塵埃落定前」即分享資料，恐有競爭者搶先發表之，進而增添不必要的風險存在；再者，獨特且未公開的資料集有助於獲得研究資助，若資料過早公開，可能因吸引外界競爭，有減少本身獲得資助機會的風險；最後，有研究者也擔心公開的資料若有誤，會影響自身團隊的信譽。這些考量反映了學術界在推動資料分享與維護自我團隊利益之間的微妙平衡，使得研究者在決定公開資料的時機上，常感到猶豫不決。

從上述可發現，許多研究都探討了分享和使用開放研究資料的驅動和阻礙因子。然而，這些研究只揭示了部分的影響因素，學界仍需要能綜覽全貌的文獻整理。為了填補這個研究缺口，2020 年由荷蘭台夫特理工大學、瑞士蘇黎世理工大學與臺灣國立臺灣大學三位研究者，以系統性文獻回

顧方式，爬梳過去探究驅動與阻礙開放研究資料的32篇研究論文，在統整過往研究常使用於理解研究資料分享與使用的理論和模型外，更進一步識別與釐清影響研究者分享或使用他人資料的關鍵因素，旨在提供開放研究資料於釋出與利用上，探討並彙整更全面的影響因子（Zuiderwijk et al., 2020）。

他們的研究發現，這些影響研究者分享或使用他方資料的因素常具有一體兩面的特性，一方面既是驅動力，另一方面亦是阻礙，端看該因素呈現的特徵而定。舉例而言，「對資料生產者的信任」、「對其他研究者於測量方法上的信任」二者是促使研究者再利用他者所產製之研究資料的驅動因素，然而，當研究者對於其他資料生產者與研究者的信任程度較低時，便可能會降低其使用開放研究資料的意願，成為阻礙因素。由此可知，各因素並非純然作為一種驅動或阻礙，端看研究者如何感知該因素影響其意向的程度。

Zuiderwijk 等人根據深度文獻分析總結，影響研究者公開分享與使用研究資料的主要因素有11項，是為開放研究資料擇採模式（Open Research Data Adoption Model，ORDAM）（圖一）：

1. **研究者背景：**研究者所屬的學術領域、研究年資、年

圖一、影響研究者公開分享和使用研究資料的主要因素

圖片來源：（Zuiderwijk et al., 2020）

齡、性別以及所在的國家或地區等個人特徵。這些因素可能會影響研究者對資料共享的態度和行為。

2. **自願性質：**指研究者公開分享資料的動機是出於自願，還是受到外界的要求或規定。如：研究者在獲得專題研究獎助或期刊投稿時，是否必須將資料公開作為一項條件。
3. **內部動機：**指驅動研究者決定分享資料的內在因素，如個人的利他主義、對學術界發展的看法以及自身的性格特點等。這些因素反映了研究者的價值觀和動機。
4. **外部動機：**指外界因素對研究者分享資料決策的影響，如預期的成就表現提升、獲得酬賞以及引

文數增加等。這些因素體現了研究者對外部激勵（incentives）的回應。

5. **外部資源環境：**研究者所處的硬體和軟體條件，如可利用的科技、設備等。良好的外部資源環境能夠為資料共享提供便利。
6. **信任：**研究者對資料典藏庫的潛在貢獻者和使用者的信任程度。信任程度高意味著研究者相信其他人會恰當地使用和貢獻資料，從而促進資料共享。
7. **社會性影響：**指研究者所處的社交環境對其資料共享行為的影響，如同儕壓力、學術社群共享資源的風氣等。積極、鼓勵開放的社會性影響有助於一個學術領域形成資料開放的文化。
8. **努力成本：**指研究者認為分享資料所需投入的各種資源、心力，如金錢、人力、時間等。覺察到需付出的努力成本越多，研究者分享資料的意願可能越低。
9. **個人過去的經驗：**指研究者過去分享或再用資料的經歷對其後續行為的影響。正向的經驗能夠增強研究者對資料共享的信心和動力。
10. **立法、相關規定與政策：**指國家層面的法律法規和政策對研究者資料共享行為的規範和引導。完善的制度環境有助於營造資料共享的風氣。

11.**資料本質：**指資料本身的特性，如其敏感程度、所處的研究脈絡是否容易重建等。這些特性會影響研究者對資料共享的風險和收益的判斷。

上述的 ORDAM 擇採模式奠基於前人豐富的調查與研究成果，將國際相關文獻針對研究者資料分享與再用的動機與行為研究整合為一完整體系。這有利於未來相關單位設計激勵機制或創制使用者中心的資料服務。每一個因子說明如下：

一、研究者背景

研究者的個人背景是指，領域、經驗、年齡、性別和研究者所在之地理位置如何影響他們開放自己的研究資料之意願。首先，不同學科領域之間對於研究資料的分享之共識不一。有研究便認為，遺傳學、大氣科學和海洋學等學科便具有自由開放取用和強大資料典藏庫的傳統，而野生動物生態學、醫學和許多社會科學等學科卻沒有對應基底（Tenopir et al., 2011）。再者，開放研究資料的分享可能會受到某些組織文化、學科領域文化、支持資料分享文化、以及組織實踐方式所驅動或阻礙，而這些不同文化與實踐方式之間亦是相互影響連結的。

儘管由人口統計因素本身並不能解釋研究者於資料分享

與使用的行為，然不同的人口統計因素所發生的情況是有所差異的。例如，有些研究便指出，50 歲以上的研究者會有較大興致分享其研究資料，然對於 20 至 35 歲的研究者而言，則惦記於公開分享資料對自身的影響，而較傾向不對外分享（Schmidt et al., 2016）。另外，來自德國和加拿大的研究人員通常比來自美國或歐洲的研究人員更不願意分享研究資料（Enke et al., 2012），而 Zuiderwijk 等人（2020）推論，這種差異可能與各國社會經濟特徵、資料分享政策或文化差異有關，然尚須進一步的未來研究證實。

二、自願性質

受 UTAUT 理論的啟發，「自願性質」在 ORDAM 中，指的是研究資料公開的動機是否源自於外在的要求或義務。這種外在壓力主要體現於資料分享的面向。此因素探討研究資料的公開分享是否受到來自資助機構、政府部門或學術期刊等機構的政策影響。這些影響因素雖不同於法律規定的硬性要求，但仍然擁有一定程度的影響力，促使研究人員必須考慮與研究資料相關的處理和分享問題。在資料的使用層面上，開放研究資料政策的實施及研究人員所感受到的同儕壓力是推動他們使用開放資料的關鍵因素。在此自願性質的因素上，根據 Zuiderwijk 等人的研究，當研究者面臨一項特定

任務，而使得他們需要取得特定的資料時而無其他選擇時，對開放資料的需求便會顯著增加：儘管獲取資料可能存在困難，研究者仍必須想方設法去尋找且運用開放研究資料。

三、個體內部驅力與內在動機

為了深入探討研究者自身內在動機在開放研究資料分享與使用過程中的作用，研究者亦必須考慮個人性格特質、對開放研究資料之態度與信念，對於他們開放與利用資料的影響。增進現有領域知識和促進科學發展被廣泛認為是激勵研究者分享資料的內部動機。此外，研究者的個人性格特質也可能影響他們分享資料的意願。而 Fecher 等人（2015）指出，開放性、責任感、外向性、親和性和神經質等五大性格特質（Big five personality），亦可能促進或抑制研究者公開分享他們的資料。然而，過往的學術文獻探討了研究資料公開分享的驅動因素，而較少關注阻礙分享的因素。然而，一些研究者在分享過程中確實面臨諸多障礙，例如擔心資料僅被少數人重複使用、個人惰性，以及研究成果可能的商業化潛力等。這些問題雖然在文獻中有所提及，但尚未得到充分探討。

四、外部動機

研究者對於開放研究資料之分享與使用的外部動機往往與其預期成果有所連結。研究人員藉由公開分享研究資料以期能獲得接續可能發生的成果。這部分的實例包含，藉由公開分享研究資料，擴展研究人員自身的人脈網絡、增加跨學科的潛在交流機會，並增強合作及解決特定問題的潛力。此外，外部動機還涉及藉由公開研究資料，提高科學研究的效率、獲得學術認可和獎勵、提升個人影響力與知名度等方面。然而，在此因素中，也有研究者由於擔心失去對資料的控制、在特定研究領域的優勢、或是潛在的未來研究機會，而選擇不公開分享他們的資料。

除分享之外，個別研究者使用開放研究資料的預期表現因素則包含：感知資料集對自身有所用途、欲獲得新見解並推動科學前進、促進不同群體之間的協作、展開未被探索的主題、測試新的或替代的假設和分析方法，以及整合不同的資料以縮短研究過程等。其中，欲獲得新見解並推動科學前進會影響研究者如何感知資料的有用性，讓既有的資料能夠於新的情境中被應用、開展新可能，是促使開放研究資料更趨有效利用的一大要素。不過，使用他人的研究資料也有可能會面臨到既有資料的使用限制，如資料原本有缺陷、或是

嘗試過後無法創造出新研究而浪費時間，這是阻礙研究人員使用公開研究資料的顧慮之一。

五、外部資源環境

技術資源的充足或缺乏，亦會促進或阻礙研究者公開分享或使用研究資料。具體來說，基礎建設的可得性——如個別研究者周遭具備適當設計的基礎建設、便利的資訊系統與通訊技術，會影響研究者在分享其研究資料的動力。研究者 Wallis 等人（2013）便指出，擁有充足資金、人力、規模和基礎設施的研究人員更傾向於公開分享其研究資料。此外，大型資料庫的存在對研究人員公開分享資料也極為關鍵。如果缺乏合宜的資料儲存設施，資料就無法被妥善保存，更遑論是再被公開分享。這些儲存設施的持續運作也直接影響到研究資料能否被高效率地公開分享。

除了硬體設備的支持之外，充足的資金規劃與支持亦能有效激起研究人員公開分享其研究資料的意願。同時，當缺乏經費資源時，這往往形成研究人員在公開分享資料上的主要阻礙，此阻礙可能並非以對人員財務的直接損失出現，而是間接性的方式，例如，研究者在分享資料後便可能喪失了潛在因授權而獲得的收入。

六、信任

信任是影響開放分享研究資料的關鍵因素，既可以驅動研究者公開數據，也可能成為阻礙分享的障礙。在推動資料開放共享的過程中，信任涉及多個層面：同儕社群對研究結果的信任、開放資料使用者對個別研究者的信任、研究者對自己研究成果的信任，以及個別研究者對開放資料平台長期保存資料能力的信任。

具體來說，當研究者決定公開分享其研究資料時，增強了研究的可再現性，也提高了研究結果的可信度；這是因為其他研究者能夠更深入地了解研究過程的全貌，從而驗證或重複該研究。透過開放資料，研究者展示了其研究的透明度和誠信，有助於換得同行和公眾的信任。然而，研究者在開放資料的同時，也可能面臨失去對資料控制的擔憂。這些研究者可能擔心出現預期之外的資料再用情況，例如其他人可能會誤解資料的本質，進而導致資料的錯誤解讀或濫用。另一方面，研究者也可能憂慮，一旦公開了原本由自己獨家持有的資料，就失去了作為資料擁有者所享有的獨特資源優勢，從而在競爭中處於不利地位。總的來說，信任在開放分享研究資料的過程中扮演著關鍵角色。增進各方對研究成果、資料使用和平台運作的信任，是促進資料共享和開放科

學的重要基礎。

七、社會性影響

文獻顯示，社會影響力和歸屬感對研究者公開分享研究資料的決策有顯著影響。這些社會因素涉及廣泛的範疇，包括個人感受到的社會規範壓力、普遍的社會規範、學術出版要求而衍生的壓力、同儕的影響力、所屬之學術社群對資料分享的態度、全球對資料分享與保護的需求和關注，以及科學家與其所屬團體內的規範標準等。

研究文獻揭示了一個令人意外的矛盾：「研究成果的分享文化」竟成為阻礙資料公開的主要障礙之一。這種看似自相矛盾的現象背後有其深層原因：「研究成果的分享文化」指的是在高度重視文章發表的學術環境中，研究者往往認為，只要積極發表論文、與社群或公眾慷慨地分享「研究成果」就已足夠展示其對外分享研究成果的貢獻，從而忽略了分享原始資料的環節與必要性。這種情況導致了一種矛盾現象：越是強調論文成果發表的學術社群，研究者分享原始資料的意願可能越低（Sayogo & Pardo, 2013）。

在資料再用的情境下，過往文獻也有相似於公開分享資料的因素，如社會壓力、親近同儕的觀點、以及研究者對於資料再用和規範的正向態度，是推動研究人員使用公開分享

資料的因素。至於資料再用的阻礙，除了文獻提及同事的低社會影響力的因素之外，Zuiderwijk 等人總結（2020）認為，其他研究者不使用開放研究資料的觀點或認知、對資料再用的負面態度、研究主管不再利用資料等事由，皆是研究者再用資料上的阻礙，而此須在未來研究中進一步探究確認。

八、努力成本：覺察到需要付出的心力

公開分享研究資料被視為一種降低重複性資料收集工作的方法。避免重複蒐集相同的資料不僅能夠減少研究成本，同時也提高了資源的使用效率，這對於從事研究者來說是一項基於研究成本的實際考量。根據 Cragin 等人（2010），如果原始資料在研究過程中就已經完成處理和分析，而非在研究結束後還需要花費額外的時間和人力成本進行整頓的話，則研究者更可能願意公開分享其資料。再者，如果研究者並未親自參與資料的收集過程，例如這項工作是由其他同事或外部機構完成的，那麼他們公開分享資料的意願通常會更高。這其實不難理解：若公開分享資料所需要的投入和努力越多，這些成本和辛苦很可能提高研究者分享資料意願的門檻。換言之，如果研究者預期到分享資料會需要付出更多努力，他們可能就會有所顧慮，從而放棄將資料公開分享的想法。

關於開放研究資料的使用，當研究者能夠輕鬆存取並方便使用這些資料時，他們再次利用這些資料的意願自然會增強。然而，相對地，如果研究者預見到在資料再利用過程中需要付出較多的努力和心力，他們可能就會對利用這些開放分享的研究資料失去興趣。舉例來說，若資料散落在不同的地方，研究者則需要投入尋找、分辨及整合這些資料的工作，並且無法確定所投入的時間與資源是否最終能得到相應的回報。

九、研究者的過往經驗

研究者過往的正面經驗能夠提高其公開分享研究資料的意願。這些正面因素包括能夠接觸到研究資料管理相關的專家、有機會獲得資料管理諮詢、掌握資料管理技巧、理解並確實再用他人的研究資料，以及相信開放分享研究資料能夠有助於培養學生和新進研究人員。相對地，當研究者過去缺乏公開分享資料的經驗，或曾經歷負面的感受時，這些過往經驗可能會成為他們公開研究資料的障礙。Zuiderwijk 等人也歸結提出，研究者在公開分享資料方面的成功案例，可以成為鼓勵因素；另不論是個人經歷或聽聞的負面經驗，都可能導致研究者對於公開分享研究資料的意願降低。

在使用開放研究資料的動力，個體研究者則主要受兩大

因素影響：一為是否擁有先前的正面使用經驗，二為是否具備開放資料使用的相關能力。若研究者在過去已經熟練地運用開放研究資料並體驗到其效益，則在後續的使用過程中，他們將能更加得心應手，且節省了大量摸索時間。然而，缺乏使用經驗、對特定類型資料的不熟悉，以及技能不足，這些都可能成為使用開放資料的障礙。此外，培訓資料使用者所需的成本，也可能使研究者對於使用開放資料望而卻步。

十、法律與規範

在開放分享資料的環境中，法律和規範既可能促進也可能阻礙研究人員分享與使用研究資料。當存在明確且透明的資料分享政策，並且得到國家和地方政府在資料管理和規劃方面的支持時，尤其是在資料生命週期的每一階段都有適當的管理策略，這將有利於促進資料的分享。然而，法律上的權利和限制也可能成為障礙。當資料集涉及知識產權保護、敏感議題、個人隱私或商業秘密時，研究者就難以輕易地對外分享這些資料。此外，繁複的授權條款可能增加研究者的認知負擔。此外，當資料集是由多個組織共同創建時，這些組織間可能存在不同的安全等級、政策和法律要求，如何確保公開分享的資料符合所有組織的要求，是一個重要的考慮點。

在使用方面，依照學術社群對研究原創性的要求，法律

與規範並非推動研究者利用開放資料的正面因素。相反地，由於法律上的諸多限制，這種「法律瓶頸」（Haeusermann et al., 2017）包括了對資料敏感性、所有權挑戰、潛在隱私侵犯的擔憂恐會增加研究者再利用開放研究資料的摩擦。然涉及國家安全和商業秘密時，資料再利用變得更加複雜。因此，法律與規範主要作為一種防範機制，以確保資料再利用時的潛在風險得到控制，而非積極推動的力量。

十一、資料特性

不同學術領域採用的研究方法各不相同，由此產生的研究資料在類型、本質（機密性和敏感度）、以及使用方式上均有所差異。此外，各學科在資料管理實踐上的不同做法，如是否有效實施資料品質控管、是否採用獨特識別碼以區分不同資料集、以及是否附上適當的詮釋資料和描述文件，這些都可能影響資料的品質，進而成為促進或阻礙研究者公開分享或使用研究資料的關鍵因素。例如，當資料集以易於理解且具有互通性的格式呈現時，有助於提高研究者對該資料的分享和使用意願。然而，若資料集品質不佳、缺乏標準化或包含敏感內容，則可能阻礙研究者公開分享或再次利用該資料。

研究顯示，相較於質性資料，研究者更傾向於公開分

享量化資料（Cragin et al., 2010）。這種趨勢與質性資料中經常含有的個人隱私或敏感資訊有關，這不僅增加了去識別化或刪除資料的預期工作量，同時也為研究者在分享過程中增添了困難。在美國社會科學領域中，關於質性資料的分享，研究者通常更願意分享他們用於收集資料的研究工具，例如訪談大綱和含有實際問題的開放式問卷等。然而，由於涉及敏感問題，他們往往不願意公開原始資料本身（Jeng & He, 2022）。

綜合上述11項主要因素可見，研究資料的公開分享與使用是一個涉及多方面考量的決策過程。研究者在分享資料之前，不僅需要考慮處理成本和權利責任的分配，還需評估分享後可能帶來的個人利益以及潛在的競爭問題。與此同時，雖然不分享資料可能有助於最大化個人的研究成果，但也可能阻礙資料的再利用和潛在的合作機會。這項研究揭示了研究人員在分享和使用研究資料時面臨的多重矛盾和考量，為2020 年代之前的開放分享研究資料理論和研究樹立了一個重要的階段性里程碑。基於這些發現，學界能夠更深入理解如何增加分享和使用資料機制的驅動力，減少阻力，從而實現開放科學——開放取用的願景。

參、章節小結

Zuiderwijk 等人的發現，凸顯出營造有利於資料流通、分享與再用的學術環境，以及降低研究者面臨實務障礙的重要性。唯有在足夠的人力、資源與清晰的權利歸屬等「保障措施」皆到位的情況下，研究者才能放心地分享與公開資料，進而實現開放研究資料的潛在效益。這個結論呼應了前文所提及的觀點，亦即學術界需要在推動資料共享與維護個人學術利益之間取得平衡，方能創造一個促進開放科學的良性環境。

學術研究是一個複雜而深入的過程，涉及資金、期望、專業知識的結合，以及人類知識的進步。這種進步不僅表現在成果的累積，也體現在經驗和過程的知識傳承上。在當今資料密集型的科學時代，研究資料的開放分享已成為整體學術社群共同期待的價值觀與行動。為了實現研究資料開放的價值，不僅需要資料本身符合 FAIR（可查找、可近用、可互操作、可再用）原則，還需要學術界理解和共同承擔推動資料公開以及消除阻礙的做法。

同時，制定相關的資料分享開放政策有助於建立標準，指引學術社群向前邁進。儘管在臺灣目前只有調查研究被強制要求分享其研究資料，但各學術領域已逐漸意識到國際對

於研究資料分享的重視趨勢，並逐步形成相應的共識與措施。相關的基礎設施建設也在進行中，例如中央研究院的資料寄存所（depositar）資料典藏庫，基於 FAIR 原則，鼓勵國內研究者妥善管理和保存研究資料，以實現資料公開分享和再利用的願景。臺灣學界還需要對此議題有更多的認識、實踐、討論與反思，而可以期待未來更為通透交流與互信的學術環境。

參考文獻

Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology, 63*(6), 1059-1078. https://doi.org/10.1002/asi.22634

Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data sharing, small science and institutional repositories. Philosophical Transactions of the Royal Society of London. *Series A: Mathematical, Physical, and Engineering Sciences, 368*(1926), 4023-4038. https://doi.org/10.1098/rsta.2010.0165

Enke, Thessen, A., Bach, K., Bendix, J., Seeger, B., & Gemeinholzer, B. (2012). The user's view on biodiversity data sharing — Investigating facts of acceptance and requirements to realize a sustainable use of research data. *Ecological Informatics, 11*, 25-33. https://doi.org/10.1016/j.ecoinf.2012.03.004

Fecher, Friesike, S., & Hebing, M. (2015). What drives academic data sharing? *PloS One, 10*(2), e0118053. https://doi.org/10.1371/journal.pone.0118053

Haeusermann, T., Greshake, B., Blasimme, A., Irdam, D., Richards, M., & Vayena, E. (2017). Open sharing of genomic data: Who does it and why? *PloS One, 12*(5), e0177158–e0177158. https://doi.org/10.1371/journal.pone.0177158

Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of Organizational Computing and Electronic Commerce, 26*(1-2), 14-40. https://doi.org/10.1080/10919392.2015.1124007

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management, 29*(4), 258-268. https://doi.org/10.1080/10580530.2012.716740

Jeng, W., & He, D. (2022). Surveying research data-sharing practices in US social sciences: A knowledge infrastructure-inspired conceptual framework. *Online Information Review, 46*(7), 1275-1292. https://doi.org/10.1108/OIR-03-2020-0079

Open Knowledge Foundation (2021). Open definition. https://opendefinition.org/od/2.1/en/

Sayogo, & Pardo, T. A. (2013). Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data. *Government Information Quarterly, 30*, S19–S31. https://doi.org/10.1016/j.giq.2012.06.011

Schmidt, B., Gemeinholzer, B., & Treloar, A. (2016). Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. *PloS One, 11*(1), e0146695–e0146695. https://doi.org/10.1371/journal.pone.0146695

Tenopir, C., Allard, S., Frame, M., Birch, B., Baird, L., Sandusky, R., Langseth, M., Hughes, D., & Lundeen, A. (2015). Research data services in academic libraries: Data intensive roles for the future? *Journal of escience librarianship, 4*(2), e1085. https://doi.org/10.7191/jeslib.2015.1085

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: practices and perceptions. *PloS One, 6*(6), e21101-e21101. https://doi.org/10.1371/journal.pone.0021101

Wallis, Rolando, E., & Borgman, C. L. (2013). If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PloS One, 8*(7), e67332-e67332. https://doi.org/10.1371/journal.pone.0067332

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*, 160018. https://doi.org/10.1038/sdata.2016.18

Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PloS one, 15*(9), e0239283. https://doi.org/10.1371/journal.pone.0239283

Chapter
3

研究資料開放與再用實踐
——自然科學篇

章節作者：黃寶霈
章節編輯：黃郁雯

章節地圖

- ◆ 研究資料分享的正式管道：資料出版
- ◆ 資料出版與資料論文出版生態圈
- ◆ 研究案例分享：臺灣自然科學學者資料出版的實踐與動機
- ◆ 章節小結

章節摘要

延續第一章介紹的研究資料生命週期，本章將焦點放在「資料發現與近用」的階段，以一較正式的公開分享方式——資料出版（data publication）為主，逐步介紹資料出版的特性、管道，以及其相關新興學術出版品——資料論文（data paper）。章節後段的研究案例分享並以臺灣自然科學領域為例，深入探索學者投稿資料論文的動機，與資料出版的驅動力。

壹、研究資料分享的正式管道：資料出版

資料出版（data publication）在學界仍未有精確定義，且學界普遍將資料出版的概念基於一個隱喻——即將所謂的「資料釋出」比擬為「出版」。所謂的「資料釋出」，係原本可能受限於特定單位或研究者團隊個人使用的資料，經過適當的策管（curation）與描述後，釋出給更廣泛的受眾使用。這個釋出的過程讓資料能夠被更多人取得、分析與再利用。傳統意義上的出版（publication）是一種將文件向公眾公開（made public）的行為，且出版後的產品需有讓他人閱讀的意圖。若將這個概念應用於資料，即將出版的對象從常見的文字、圖像等資訊內容，擴展為各種型態的資料，則出版可被初步解讀為一種將資料向公眾公開傳播予以閱覽的行為（Borgman, 2007; Kling & McKim, 1999）。

然而，前人文獻中可再萃取出對於資料出版的進一步理解：Lawrence 等人（2011）提出「Data Publication」與「Data publication」兩種概念（「P」與「p」字母的大寫與小寫型態）。字母小寫 p 所指涉之資料出版，僅指涉將資料放到網站上的行為，而 P 字母大寫之資料出版（後稱為大出版）則反映更為宏觀的期待：「讓資料盡可能在網站上被永

久近用（To make data as permanently available as possible on the Internet）」（Lawrence et al., p.7）。因此，大出版不僅需要提供可信賴、易於理解的格式與內容，也需要確保資料集具備完整與不變動的特性，並需具備足以支援後續他方研究的詮釋資料（metadata），藉此彰顯資料的科學價值與重要性。綜上，除了資料需向第三方公開的特性外，資料出版也能更進一步被解釋為：一種對資料能永久被他人使用的承諾。

資料出版的公開性與對詮釋資料的重視，能改善資料的發現取用、互操作性與再利用的困境，是開放科學、研究誠信與現代學術交流的重要基礎（Bloom et al., 2015; Reilly et al., 2011）。此外，資料出版也能作為一種對資料提供者的反饋，提供其相對應的學術聲譽，使資料集與其提供者獲得學術社群信任與承諾（Borgman, 2007; Kratz & Strasser, 2015）。這種歸屬聲譽的展現，特別適用於後續內容中會提及的一種特定資料出版類型——資料論文（data paper）。不過，在進入討論之前，以下再更進一步整理與認識資料出版的類型。

貳、資料出版與資料論文出版生態

一、資料出版的類型

資料出版依其資料主體性由低到高，可主要分為三種類

型：以文章補充方式出版、經資料典藏庫（data repository）出版、以文章主體方式出版（Costello et al., 2013; Kratz & Strasser, 2014; Mayernik et al., 2015）。

（一）**以文章補充方式出版：**將資料視為附屬物，作為傳統研究文章的附註資訊，主要目的為支援研究重現與分析，而非對資料本身的再利用。

（二）**經資料典藏庫出版：**此種資料出版方式將資料視為主體，由資料典藏庫提供資料描述文件。描述形式為從單一網頁連結到標準化的控制格式，引用方可直接近用或申請資料集連結。

（三）**以文章主體方式出版：**將資料視為主體，由研究人員撰寫對資料集的描述資訊，並提供連結資料集位置或近用方法。此描述文件會經過同儕審查流程，採用傳統期刊文章方式出版。常見的呈現方式為資料論文[1]。

由上述可知，後兩種資料出版方式較能將資料視為出版

[1] 資料論文（data paper）是此種出版品最為廣泛被認識的稱呼，不過，期刊可能會為反映不同特殊目的、以及對這種出版品的特定理解，而出現不同命名方式，如：資料庫論文（database paper）、資料集論文（dataset paper）、資料注釋（data note）、原始資料文章（data original article）等（Candela et al., 2015）。

之主體，而非僅為研究成果附屬品。儘管在多數與資料出版相關的研究中，「經資料典藏庫出版」方式更常被文獻提及（Costello, 2009; Klump, 2017; Klump et al., 2006），但「以文章主體方式出版」更適切與地體現了資料出版的意涵：這種方式將資料作為文章描述之主軸，賦予其優先性，並讓資料歷經正式出版流程而產製出成果。奠基於此，接下來本章接續將介紹以文章主體方式出版的常見形式——資料論文。

二、資料論文

資料論文是以資料作為文章主體的代表出版品，它們被視為一種從外附加於資料典藏庫的描述性文件，提供「關乎資料的事情（what）、地點（where）、原因（why）、方法（how）、人（who）」（Callaghan et al., 2012, p.112）。一篇資料論文中通常包含兩大區塊：一是提供傳統學術交流的資訊，如：文章標題、作者、摘要、關鍵字與參考文獻等；另一則是關乎資料的詮釋資料，如：資料出處、範圍、格式、再利用潛力等。目前尚無統一的資料論文標準化描述格式，但基本原則為不提供資料分析或結論（Candela et al., 2015; Chavan & Penev, 2011; Kratz & Strasser, 2014; Pouchard et al., 2014; Schöpfel et al., 2019）。

根據學術傳統的出版流程，資料論文與資料期刊（data

journal）的關係，恰如同研究論文之於學術期刊的關係一個別的資料論文被彙集成資料期刊（data journal）。根據所收錄的資料論文數量占總內容量的比例，這些資料期刊可分為兩大類：一類是將資料論文作為主要或唯一收錄範疇的專門型資料期刊（pure data journal）；另一類則是同時收錄一般學術文章與資料論文的混合型資料期刊（mixed data journal）。

Schöpfel 等人（2019）針對於專門型資料期刊的調查指出，這些期刊多數出版不及二十年，大多來自科學、科技、工程與數學（STEM）領域，特別是其中之生物醫學領域。他們發現，目前擁有最多資料論文數量的期刊是出版商Elsevier於 2014 年出版的《Data in Brief》[2]，這個期刊與Elsevier旗下的另一個期刊《Atomic Data and Nuclear Data Tables》[3]、以及出版商 Springer Nature 的《Scientific Data》[4]三者，出版了專門型資料期刊中半數以上的資料論文。此外，多數的專門型資料期刊為開放近用（Open Access, OA），同時延續傳統學術期刊盛行之封閉式同儕審查（closed peer review）機制，這種相對嚴謹的同儕審查機制能確保稿件與資料的品質（Schöpfel et al.,

[2] 《Data in Brief》於2019年已累計收錄3628篇，累計至2024年6月24日已出版10,285篇。

[3] 《Atomic Data and Nuclear Data Tables》於2019年已累計收錄832篇，累積至2024年6月24日已出版1,236篇。

[4] 《Scientific Data》於2019年已累計收錄1520篇，累積至2024年6月24日已出版3,745篇。

2019）。

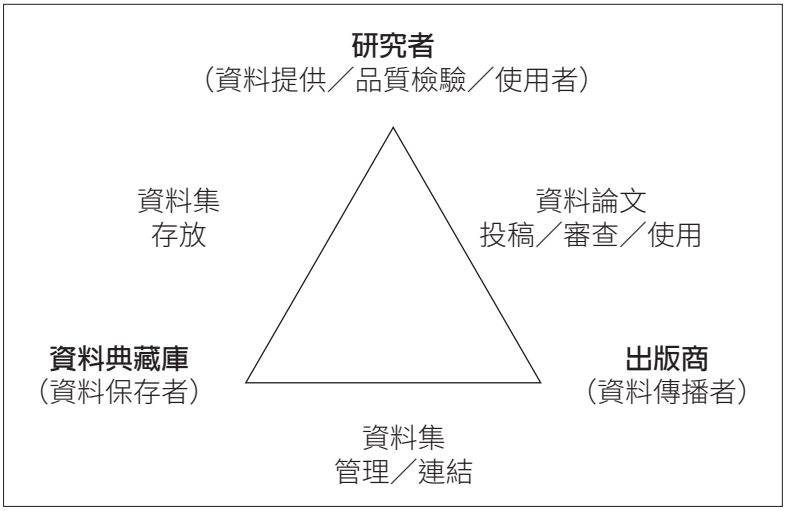
總結來說，資料論文除了對資料集進行描述、經同儕審查對資料品質的認證外，其論文形式也更讓學界去理解與評價學術成果。資料論文的種種特性都使其所承載的資料獲得更多關注，得以發揮影響力與價值，獲得學術社群的信任（Kratz & Strasser, 2015）。

三、資料論文出版生態圈

以資料論文這種資料出版類型為例，出版流程可能依不同領域而有所差異。除了科研政策、研究資助機構、資料中心與圖書館補助開放近用之文章處理費政策等因素會帶來影響外，資料出版之體制主要仰賴研究者、出版商與資料典藏庫三者間的緊密搭配。依圖二所示，研究者將研究中產製的資料進行格式編輯、並且添加必要的描述性資訊，再來將處理後的資料放入資料典藏庫中，取得連結儲存位址的數位物件識別碼（Digital Object Identifier, DOI）後，並撰寫對應的資料論文投稿至資料期刊。在收到稿件後，期刊出版商會安排領域專家對資料論文進行同儕審查。通過審查的資料論文將被正式出版，而其所描述的研究資料也將透過資料典藏庫提供給其他研究者再利用。

圖二之資料出版流程體現了研究者、出版商與資料典藏

庫三者的密切協作。研究者負責產製研究資料並撰寫資料論文；出版商負責組織同儕審查並出版資料論文；資料典藏庫則負責儲存和提供資料的長期近用。三者的通力合作，支持並促進了資料出版這一新興的學術實踐。



圖二、資料論文出版生態圈

圖片來源：本章作者整理

由此可知，作為資料擁有、提供者的學者是資料出版機制的原動力，主宰了資料生命週期的走向，從資料創建、處理、分析、保存到出版等，同時掌管了資料價值與品質的檢驗程序。學者對資料的觀點與行動將會影響後續其他再利用

與管理資料的利害關係者（Wallis & Borgman, 2011）。另一方面，作為資料傳播者的出版商在資料出版上的定位，則是建立資料分享的正式管道（Lin & Strasser, 2014），提供對應宣傳方式以增加學術社群對資料、資料分享與出版的意識，進而促進資料研究通透性。最後，資料出版需奠基於有效的資料管理與穩固的數位基礎建設之上（Bloom et al., 2015），作為提供資料長期近用承諾的資料典藏庫便是基礎建設中的必要存在，支撐資料傳遞與再利用，可謂資料可得性的重要基石（Borgman, 2015）。

在了解資料出版的生態圈後，接續便以資料論文為例，從資料出版原動力的學者視角切入，探索學者發表資料論文的動機與觀點，有助理解資料出版機制是如何被啟動並維持動能。

參、研究案例分享：臺灣自然科學學者資料出版的實踐與動機[5]

在多數前人研究中，關於資料分享動機的討論常環繞著「尋找正確誘因」的議題，其中，資料論文這種正式出版品與其可能伴隨而來的引用機會，被學界廣泛認為是激勵學者

[5] 內容摘錄自黃寶霈（2023）未出版之碩士學位論文

分享研究資料的重要因素（Klump, 2017）。然而，資料論文的發表能否有效激發學者分享其研究資料的意願，仍有待進一步探究。

本章節研究案例中，進一步探究學者選擇資料論文此新型態出版方式進行投稿的原因，以及探討驅動或激勵其投稿行為的因素。透過對投稿動機的分析，本案例將探究資料論文的潛在動能與阻礙困境，從而作為學界與相關利害關係方認識資料出版驅動力的開端，並指引未來訂定或調整相關制度面的參考方向。

一、研究方法簡述

（一）研究對象與執行流程

本案例之研究對象聚焦於廣泛的自然科學領域學者[6]，採立意抽樣的方式篩選出曾於2015至2022年間，於《Data in Brief》、《Scientific Data》二本資料期刊中發表過資料論文的臺灣作者[7]，且其須為第一作者或通訊作者。最終，共招募到

[6] 學者的研究工作屬於自然現象的觀察與實證，如生命科學（生物學）、物理科學（含物理學、化學、地球科學、天文學）等學科領域。同時，倘若他方領域的臺灣資料論文作者有意接受訪談邀約時，本研究亦不將其排除，而是視之為一種補性的觀點。

[7] 由於學者投稿期刊不一定會被接受與出版，研究者難以掌握曾投稿過資料論文的臺灣學者數量，引此，便改以曾發表過資料論文的臺灣作者作為替代母體，並將研究關懷聚焦於廣泛的自然科學領域學者上，採立意精神抽樣。

16 位來自廣泛自然科學領域之受訪者。

本研究採取半結構深度訪談法，待所有正式訪談完成後，研究者便整理資料並予以編碼：首先進行客觀的開放編碼，再向上歸納出主題概念、主軸編碼，以形成對資料論文投稿行為之概覽。同時，研究者以自我決定論框架檢視主題概念，演繹凝聚出動機元素的組成。研究所採用的研究工具有二：

訪談前問卷：為確認受訪者資格符合研究目的所需，並更充分了解受訪者背景資訊而設計。問卷內容主要包含三大部分：研究領域與出版資料的經驗、所有資料論文的投稿經驗與其承載之資料集內容、最後一次投稿資料論文的動機輪廓。

訪談大綱：為指引訪談過程，本研究奠基於文獻回顧與前導研究[8] 成果設計訪談大綱。大綱內容聚焦於學者投稿資料論文的動機與觀點，主要分為四大區塊：研究領域、一般研究論文的投稿動機、資料論文的投稿動機、以及對資料論文的觀點。

（二）主要參照框架：自我決定論

為盡可能勾勒出行為背後複雜動機的全貌，研究採用自

[8] 研究者於2021年，先透過半結構訪談法探詢3位不同領域、研究生涯階段的學者，了解學者在投稿資料論文的驅力與決策行為。

我決定論（Self-Determination Theory，以下簡稱 SDT）作為一種指引框架，提供研究者認識動機類型與層級的理論基礎，於此設計訪談前問卷中關乎動機的題項；也在資料分析時，作為指引研究者檢視、理解動機元素與組合的框架。

SDT 認為，人的行為源自於兩個驅動力：內在興趣以及將外在規範內化[9]後所形成的價值觀。人擁有將自我意識整合的傾向，這種傾向是來自於三種人固有且基本的心理需求：自主性（autonomy），個體感知到自身行為是源於自我意識；勝任感（competence），個體在與環境互動中感知到個人能力得以實行與展現；關聯感（relatedness），個體感知到與他者連結並獲得歸屬感。奠基於心理需求，SDT 將動機劃分為兩種：由個體固有興趣所驅動的內在動機（intrinsic motivation）；以及仰賴於獲得其他外部成果或損失規避的外在動機（extrinsic motivation）。另外，SDT 以無動機（amotivation）指稱個體缺乏行動意向的狀態，此可能源於個體對行為缺乏勝任感、無法認定行動或產出成果有所對應價值等（Deci & Ryan, 2002）。

[9] 內化是一種將外在價值、信念或行為規範轉化入自身的過程，涉及個體被其社會脈絡所同化，以及內部相關價值觀的協調，是個體整合內在傾向的一種表現（Ryan et al., 1985），在SDT假設裡，內化能讓人們基本心理需求得以滿足（Ryan & Deci, 2017）。

二、研究成果

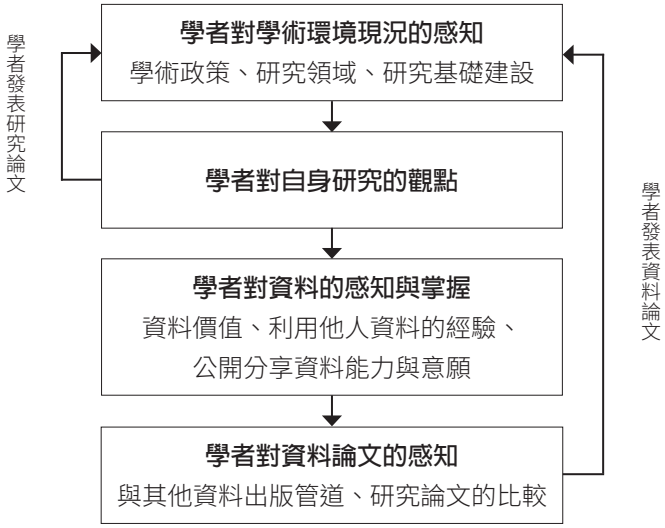
（一）資料論文投稿行為之建立基礎

研究發現，學者的資料論文投稿行為是來由層層建構下的認知信念所驅使。最基層便為學者感知自身所處的學術環境背景為一規範或支持體系，此基底並從外部影響學者後續對於其研究信念、資料價值、以及資料論文的認知。這些思維型塑出一個環環相扣的體系，以其涵蓋的面向大小可切分出：學者對自身研究的觀點、學者對資料的感知與掌握、學者對資料論文的感知等（圖三）。

1. 學者對學術環境現況的感知

學術環境現況主要是由學術政策、學者所屬的研究領域、以及研究資料基礎建設所組成。學術政策建構了受其規範區域的整體學術環境與研究發展進路；政策效力則直接影響該區研究領域及其中學者的研究行動，例如，當國家學術政策訂定指標時以期刊排名作為評斷學術表現依據的原則[10]，便會促使機構去規範學者成員產出對應績效，而學者則以發表期刊論文回應此重視量化績效的政策要求，如

[10] 當期刊被列為 SCI 期刊（Science Citation Index，科學引文索引）、SCIE（Science Citation Index Expanded，科學引文索引擴展版）、SSCI（Social Science Citation Index，社會科學引文索引）後，該期刊的期刊學術影響力係數（Journal Impact Factor）位於特定學科領域的排名狀態。



圖三、學者於資料論文投稿行為建立基礎之概覽示意圖

其中一位受訪者便提及，一個人的 paper 就像是分數、像是 credit。學者亦深受其所處的研究領域之慣習所影響，例如研究資料公開分享的風氣。此外，其他學術發展支持系統與利害關係者（如期刊出版商、資料典藏庫），亦影響著學者對於研究或其研究成果產製的作法。

2. 學者對自身研究的觀點

學者對學術環境帶來的支持與規範感知將影響學者接續對自身研究的觀點產生。此觀點主要展現於對自身研究成果

的期待，以及隨之衍生出的對應研究成本。舉例來說，為完成具研究價值—特殊性、原創性、應用性、可再現性—的研究，學者也必須投注對應的時間、金錢等成本，因此學者必須思考如何利用現有資源將研究效益最大化。學術環境下，學者發表研究論文的驅動力常是為了符合外在規範的要求，用於尋覓或鞏固職位、協助升遷，是一種為生存所需的「實際考量」，也是為了回應經費來源的一種社會責任與義務。這種種現實面的思量，便成為學者從事研究與決策發表行為的基石，影響其後續研究資料和出版模式決策行為。此外，學者對於發表研究論文的考量點，也將作為其後續比較資料論文投稿效益的基準。

3. 學者對資料的感知與掌握

一般而言，學者通常會以能闡述其研究概念的研究論文作為引用素材；而當學者將關懷視角由研究概念轉而聚焦至研究資料上時，對資料的感知與掌握便將影響其行為。首先，對資料的感知包含學者對研究資料價值的認定，以及自身利用他人資料的經驗二者。而對於資料的掌握，則是指學者公開分享資料的能力與意願，包含其是否為資料所有權人、分享行為是否會損及自身利益等考量。學者對於資料的感知與掌握，將作為其後續如何理解與認識資料論文概念的

基礎。

4. 學者對資料論文的感知

學者通常是透過外部的引介而初次接觸資料論文這種新型態的學術出版品。例如，在投稿研究論文時，受到期刊編輯的投稿邀約、或來自於同儕或合著者的提議。此外，學者亦有可能藉由網路搜尋、學術交流、或查看他人的研究資料或研究等途徑，認識到此類型學術出版品的存在。

學者與這些管道的互動與探索將形塑出其對資料論文的初步印象，如其中一位受訪者提到：「我覺得（資料論文）是一個很神奇的東西，跟一般的文章概念是不一樣的，那我老實說，一開始我也不太相信（有這種出版品的存在）」。此外，學者也會將資料論文與研究論文進行比較，嘗試理解二者區別與特性，作為後續投稿決策的依據。

（二）資料論文投稿動機組成元素

本研究採用 SDT 提及的六大動機相關元素概念（自主性、勝任感、關聯感、內在動機、外在動機、以及無動機）作為框架，嘗試探討資料論文投稿行為的動機組成，整體概覽如表一。

表一、資料論文投稿行為之動機概覽

心理需求或動機元素	內容項目
自主性 (autonomy)	學者於執行研究上如自身所願
	學者具備公開資料的權力
	學者能自行決定資料出版的管道
勝任感 (competence)	學者能付出投稿所需的對應成本
	學者能符合期刊對於稿件的要求
關聯感 (relatedness)	學者欲奠基於受惠於研究社群的經驗，回報他方
	學者欲奠基於學術交流討論與互助精神上，協助他人
內在動機 (intrinsic motivation)	學者欲完成自身的階段性工作
外在動機 (extrinsic motivation)	學者須符合所屬機構單位對自身職位的要求
無動機 (amotivation)	學者消極被動滿足他人提議

1. 自主性

首先，學者對資料論文的投稿與後續作業是奠基於對研究執行的自主意識上，即研究能否如自身所願般地進行，此關乎學者如何看待自身角色定位與研究慣習。在資料論文作為一種公開分享資料的管道下，學者常見的考量點是自身是否具備資料公開權力，以及資料公開的時間點。而資料型

態或學者所處的情境，也會影響學者如何選擇資料出版管道（受訪者 P08）。

2. 勝任感

學者對資料論文投稿的勝任感體現在他們認為「自身可完成投稿作業」的自信上。這種自信包含兩個面向：一是學者認為自己能夠付出投稿所需的成本（如：時間、金錢、心力）、是否能符合期刊對稿件的要求。

學者對投稿所需相對成本的認定，與其感知的期刊要求難易度密切相關：當學者認為自身有能力可完成稿件規範之要求，對所需成本也有較好的掌握時，便容易對順利完成資料論文投稿歷程產生自信。對部分受訪者而言，準備資料論文是相對游刃有餘的工作，可能只是將研究「重新包裝」（受訪者 P11），投稿所需成本較低，能較輕易產生勝任感。

3. 關聯感

數名受訪者提及其心中常存「站在巨人肩膀上」的思維（受訪者 P01、P16），其投稿動機在於推進學術發展、社群間的相互幫助。如受訪者 P14 提及，先前接收他人資料的經驗促使他產生反饋社群的想法，當他手上有一批具價值的研究資料時，也期望研究能量得以回饋學術群體。此外，關聯感也體現於學者對於資料能被他方再利用的期待上，例如不

必重製相同的研究資料，可幫助提升學術社群彼此交流的透明度。

4. 內在動機

本研究發現，個體的內在動機比外在動機更難辨識。在動機摻雜多方因素的狀況下，難以判斷學者表述的投稿動機中是否也包含「單純獲得個人樂趣」的內在動機。因此，本研究中僅針對 Vallerand 等人（1993）提出的「為完成的內在動機（intrinsic motivation to accomplishments）」[11] 進行討論，此指學者欲藉由資料論文投稿發表作為完成階段性工作的證明。學者投稿資料論文時常是為了將手邊任務告一段落，作為研究工作中的「逗點」（受訪者 P04）。學者得以從中獲取成就感，或是藉由和審稿者的互動討論中獲得正向肯定之感（受訪者 P01）。

5. 外在動機

在外在動機方面，學術機構對於職位晉升的要求直接影響學者提交資料論文的決策。資料論文是否能被認可為學

[11] Vallerand等人（1993）將內在動機區分成三種類型：為了解的內在動機（intrinsic motivation to know）、為完成的內在動機（intrinsic motivation to accomplishments）、以及為體驗刺激的內在動機（intrinsic motivation to experience stimulation）。其中，為完成的內在動機關注於試圖超越自己、達成或創造某事物。

術成就，進而影響學者是否選擇進行投稿，成為一項關鍵考量。從一系列訪談中得知，目前在國內的學術界，資料論文尚未被普遍認定為可以用於職位晉升或反映學術表現的有效成果（受訪者 P05、P06、P11、P15、P16）。由上述可知，資料論文的投稿並不僅是考慮資料分享公開與否的單點思考，而會受到機構規範、評鑑與升等體系的牽引：「我覺得最主要原因還是回到升等，因為你知道發 paper 已經不容易了，我還要去發一篇是對升等沒有幫助的」（受訪者 P06）。不過，儘管無法直接帶來績效成果，具有論文形式的資料論文仍有間接的其他效益，例如可在相關研究領域上「claim territory（宣告領土）」（受訪者 P04），並能被視為學術發表的正式紀錄之一；或當資料論文有其對應的研究論文且彼此相互引用時，資料論文便協助提升研究論文的被引用數（citation）與能見度。

6. 無動機

最後，本研究對無動機的討論主要聚焦在「當個體無法認定其行動或無法產出成果具備對應價值」的情境。舉例來說，當學者受到他人指引而進行投稿，自身並無太多自主想法時，便會傾向依附他人主張，被動接受提議（受訪者 P15）。儘管此學者自身已擁有投稿經驗，然而外在或內在動

機卻仍不存在此個體行為中。

肆、章節小結

資料論文的發展被視為一種促進學術社群資料分享、交流、再用的激勵管道，期望幫助提升研究通透度。本研究從臺灣自然科學學者的資料論文投稿動機出發，概覽學者之資料論文投稿行為建立基礎，並解析動機元素，藉以探索此一資料出版機制的驅動力。研究結果發現，學者於資料論文的投稿決策的念頭並非一夕之間冒出，而是從大環境到單點事件，在層層影響與利弊考量下漸變形成。

整體來說，學者投稿資料論文的動機與發表研究論文仍然相似，皆為回應學術交流的本質、並擔負身為研究者的責任。然而，目前資料期刊並未認列為SCI期刊，在當前學術政策無法將其作為學者績效表現紀錄時，資料論文便僅存宣告研究事實、增加學術能見度與影響力的間接效益，與能獲得實質績效分數的研究論文相比，仍缺乏強驅動力。部分資料論文投稿驅動力甚至帶有一種依附意涵，是為了提升其對應研究論文的能見度與影響力而起。

總結而言，若要將資料論文轉化為一種可實質激勵資料分享的機制，仍有賴學術體系制度面的響應，建立採認資料

論文作為學術績效表現的政策準則，以正式肯認資料出版的學術價值。如此一來，資料分享與出版便不僅是屬於少數領域中的美德與共識，而是能作為一種普遍被認可與得以實踐的行為，由此激勵更多學者參與資料的分享，促進學術知識的進步與創新。

參考文獻

黃寶霈（2023）。*從自然科學學者資料論文投稿動機觀察其學術傳播之角色*〔未出版之碩士論文〕。國立臺灣大學。

Bloom, T., Dallmeier-Tiessen, S., Murphy, F., Austin, C. C., Whyte, A., Tedds, J., Nurnberger, A., Raymond, L., Stockhause, M., & Vardigan, M. (2015). Workflows for research data publishing: Models and key components. https://doi.org/10.5281/zenodo.20308

Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. MIT Press.

Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked world*. The MIT Press.

Callaghan, S., Donegan, S., Pepler, S., Thorley, M., Cunningham, N., Kirsch, P., Ault, L., Bell, P., Bowie, R., Leadbetter, A., Lowry, R., Moncoiffé, G., Harrison, K., Smith-Haddon, B., Weatherby, A., & Wright, D. (2012). Making data a first class scientific output: Data citation and publication by NERC's environmental data centres. *International Journal of Digital Curation, 7*(1), 107-113. https://doi.org/10.2218/ijdc.v7i1.218

Candela, L., Castelli, D., Manghi, P., & Tani, A. (2015). Data journals: A survey. *Journal of the Association for Information Science and Technology, 66*(9), 1747-1762. https://doi.org/10.1002/asi.23358

Chavan, V., & Penev, L. (2011). The data paper: A mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics, 12 Suppl 15*(S15):S2, . https://doi.org/10.1186/1471-2105-12-S15-S2

Costello, M., Michener, W., Gahegan, M., Zhang, Z., & Bourne, P. (2013). Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology & Evolution, 28*(8), 454-461. https://doi.org/10.1016/j.tree.2013.05.002

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. University of

Rochester Press.

Kling, R., & McKim, G. (1999). Scholarly communication and the continuum of electronic publishing. *Journal of American Society for Information Science, 50*, 890-906.

Klump, J. (2017). Data as social capital and the gift culture in research. *Data Science Journal, 16*(5). https://doi.org/10.5334/dsj-2017-014

Klump, J., Bertelmann, R., Brase, J., Diepenbroek, M., Grobe, H., Höck, H., Lautenschlager, M., Schindler, U., Sens, I., & Wächter, J. (2006). Data publication in the open access initiative. *Data Science Journal, 5*, 79-83. https://doi.org/10.2481/dsj.5.79

Kratz, J. E., & Strasser, C. (2015). Researcher perspectives on publication and peer review of data. *PloS ONE, 10*(2), e0117619. https://doi.org/10.1371/journal.pone.0117619

Kratz, J., & Strasser, C. (2014). Data publication consensus and controversies. *F1000 Research, 3*, 94. https://doi.org/10.12688/f1000research.3979.3

Lawrence, B., Jones, C., Matthews, B., Pepler, S., & Callagham, S. (2011). Citation and peer review of data: Moving towards formal data publication. *International Journal of Digital Curation, 6*(2), 4-37. https://doi.org/10.2218/ijdc.v6i2.205

Lin, J., & Strasser, C. (2014). Recommendations for the role of publishers in access to data. *PLoS Biology, 12*(10), e1001975. https://doi.org/10.1371/journal.pbio.1001975

Mayernik, S., Callaghan, S., Leigh, R., Tedds, J., & Worley, S. (2015). Peer review of datasets: When, why, and how. *Bulletin of the American Meteorological Society, 96*(2), 191-201. https://doi.org/10.1175/bams-d-13-00083.1

Pouchard, L., Barton, A., & Zilinski, L. (2014). Data narratives: Increasing scholarly value. *Proceedings of the American Society for Information Science and Technology, 51*(1), 1-4. https://doi.org/10.1002/meet.2014.14505101088

Reilly, S., Schallier, W., Schrimpf, S., Smit, E., & Wilkinson, M. (2011). Report on integration of data and publications. https://www.stm-assoc.org/2011_12_5_ODE_Report_On_Integration_of_Data_and_Publications.pdf

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in*

motivation, development, and wellness. The Guilford Press.

Ryan, R. M., Connell, J. P., & Deci, E. L. (1985). A motivational analysis of self-determination and self-regulation in education. In C. Ames & R. E. Ames (Eds.), *Research on motivation in education: The classroom milieu* (pp.13-51). Academic Press.

Schöpfel, J., Farace D., Prost, H., & Zane, A. (2019). Data papers as a new form of knowledge organization in the field of research data. *Knowledge Organization, 46*(8), 622-638. https://doi.org/10.5771/0943-7444-2019-8-622

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senecal, C., & Vallieres, E. F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: Evidence on the concurrent and construct validity of the Academic Motivation Scale. *Educational and Psychological Measurement, 53*(1), 159-172. https://doi.org/10.1177/0013164493053001018

Wallis, J. C., & Borgman, C. L. (2011). Who is responsible for data? An exploratory study of data authorship, ownership, and responsibility. *Proceedings of the American Society for Information Science and Technology, 48*(1), 1-10. https://doi.org/10.1002/meet.2011.14504801

Chapter
4

研究資料開放與再用實踐
——人文與社會科學篇

章節作者：林潔筠、石易儒
章節編輯：丁昱寧

章節地圖

- ◆ 人文社科領域開放研究資料之實踐
- ◆ 研究案例分享：犯罪學開放研究資料之實踐與再用
- ◆ 研究案例分享：中國文學學者研究之資料利用與移動
- ◆ 章節小結

章節摘要

本章節旨在探討人文與社會科學領域中研究資料開放與再使用的議題，特別是透過犯罪學及中國文學兩大學術領域的案例分析，深入了解這些領域在資料處理和學術研究中的特殊性與挑戰。在社會科學領域，本章以犯罪學為例，分析犯罪學的研究特性，學者從蒐集、使用到分享資料的研究流程，並揭示犯罪學者與其使用的研究資料互動中所蘊藏的信任、契約、價值等要素。人文學科方面則以中國文學為例，闡述人文資料的選擇與集合涉及學者本身的學識與背景經驗之特性，以及資料於不同學者或技術間的轉移，所可能帶來的資料「摩擦力」，進一步討論了開放科學於人文學科推動的潛在複雜性。

壹、人文社科領域開放研究資料之實踐

面對提倡開放科學，建立公開透明的科學環境之脈絡下，資料的開放可供他人檢驗研究的可再現性，避免捏造、竄改等有違學術倫理之事。此外，開放研究資料亦可促成其他研究者再用資料進行研究探索、節省研究成本。

在社會科學方面的資料再用研究中，林奇秀與賴璟毅（2017）調查了臺灣的社會、政治、教育、經濟、心理領域學者的資料再用行為。在人文方面，隨著科技發展，原本看似靜態的人文資料也逐漸地與科技互動，運用數位人文的技術展現資料的新樣態。

考量到社會科學與人文學科資料的特殊性，以及對於這些資料的特徵與使用行為需要更深層次理解的迫切需求，本章節致力於社會科學資料實踐議題的深度探討，特別是透過犯罪學的場域來分析該領域資料的獨特性質，分析揭示了資料開放與再使用在犯罪學代表社會科學領域推行時，可能遭遇的挑戰與發展機會，也將探討這一過程中的潛在阻力（石易儒，2021）。另外，在人文學科資料的探討方面，本章將呈現原創之研究，研究團隊與中國文學學者的訪談，深入了解文本資料在不同情境及媒介轉換過程中可能出現的變異，

以及這些變異如何對資料的開放性與再利用性造成影響（林潔筠，2020）。這一部分的分析旨在揭示人文學科中資料處理的複雜性，並提供關於如何有效管理這些資料以促進學術共享的見解。透過對社會科學與人文學科中資料開放與再使用的兩個案例，本章旨在建立一個全面的理解框架，指出人社領域中資料處理的共通挑戰與機遇。此外，藉由探討這些領域特有的資料特性與再使用的潛在障礙，本章提供了對於促進跨學科研究與資料共享的策略與建議，為學術研究的進步開闢新徑。

貳、研究案例分享：犯罪學開放研究資料之實踐與再用[1]

一、犯罪學研究範疇與資料特性

犯罪學著重關懷社會中的犯罪現象和衍生的知識，並且發展出一套可經驗證的通則、原理，以解釋與分析犯罪現象和背後的成因，探討個人或環境因素會如何影響犯罪者的行為（Sutherland & Cressey, 1978；林山田等，2020）。

犯罪學作為社會科學範疇底下的一門學科，其研究標的為真實世界中複雜的犯罪現象和犯罪問題，因此犯罪學研

[1] 內容摘錄自石易儒（2021）未出版之碩士學位論文

究多採行實證研究方法來探討「犯罪」事實與成因，隨之衍生的資料在性質上也呈現出繁複、多變的樣貌。此外，犯罪學與其他社會科學領域的學科一樣，其蒐集與處理資料的方法偏向從真實世界中獲取資料，並以系統性的實徵研究方法來建構領域知識。因此，若想要一窺社會科學研究資料的特性，以「犯罪學」作為探索基礎，則能更加清楚掌握社科研究資料自哪些研究方法產製而來，以及呈現出來的資料型態又是怎樣的樣貌。

承接前述，犯罪實證研究主軸將圍繞在「犯罪人」之行為與影響其行為的社會環境因素，根據林山田等人（2020）提出犯罪實證研究採用的方式包括個案研究、追蹤研究、犯罪統計研究以及控制組研究等方式，犯罪學者可採取其中一種或結合數種方式進行研究，並且透過問卷調查、訪談、參與觀察或非侵入性方法（如官方統計資料、個人紀錄性資料）來獲取研究資料（許春金等，2016）。這些藉由不同方式蒐集而來的犯罪研究資料，又可大致劃分為「一手資料」與「二手資料」，一手資料為研究者與「被研究者/研究客體」實際互動（如訪談、調查或觀察）產製而出的資料；二手資料則為利用他人基於特定目的蒐集的資料，像是官方統計資料、他人蒐集的調查資料等。

二、臺灣犯罪學領域研究資料開放的實踐情形

（一）犯罪研究資料之蒐集

臺灣的犯罪研究資料多半散落在各政府機構、相關組織的所轄業務範圍之下，從司法相關機關、衛生福利部、國教署等皆有學者所需的犯罪相關資料。然而，這些官方掌管的資料未必詳細地公開於官方網站或資料庫中，因此學者在蒐集與利用上會面臨一些困難，必須透過其他途徑來取得可供分析使用的資料，主要可透過人際交情、向相關機構交涉溝通，以及在與政府建立合作或委託研究關係三種管道。具體來說，犯罪研究尋覓受訪對象的過程相當仰賴人脈關係的幫忙，像是執行犯罪團體（俗稱幫派）研究，就需要學者認識的刑警和監所單位協助引介受訪者；此外，因許多犯罪資料未公開於官方網站上，或者所公開的資料不夠精細難以作為研究分析所用，因此學者也必須透過私下的交涉與溝通，再發以正式公文流程來取得更細部的研究資料。最後，學者在與政府建立合作或接受委託研究的前提下，在蒐集相關研究資料時，能經由該委託單位拿到內部更豐富且詳盡的資料執行研究。然而，學者雖可透過上述方式取得資料，但未必皆能順利蒐集到所需資料，例如學者訪談監獄受刑人時，受限於管理方式，造成研究人員難以篩選樣本（如：因某項罪名

服刑之受刑人）發放問卷進行研究；又或是政府機關對研究人員的信任度不足，擔心研究結果會影響機關聲譽（如：機關釋出的數據，因學者研究未盡妥善而有批評聲浪），亦會造成機關在提供資料上偏向保守的態度，使許多犯罪資料未能公開予研究人員或大眾所用。

（二）犯罪研究資料之使用

學者最常運用的犯罪研究資料可大致分為官方（如政府統計數據、卷宗文書）以及非官方（政府委託計劃案衍生的訪談或問卷資料）兩種類型。在犯罪學研究中，學者大量運用政府官方蒐集而來的犯罪資料，但這些資料的產製目的為因應公務而生，不必悉數符合學術研究用途，也因此存在資料品質良莠不齊、資料精確性不足的問題，而學者欲使用資料時，也會面臨不同資料庫間的資料互通性、跨部會資料間的串接難題。最終，使得學者若要使用官方的數據資料，往往需要耗費不少功夫清理資料、以及透過不同方式（如遺漏值的差補）讓資料能更加符合研究分析用途。除此之外，學者也會經由自身額外蒐集資料（如問卷調查、訪談），去填補、克服官方資料不足之處（像是資料處理程序不明、欄位資料不夠齊全等問題），拓展研究議題的深度，並且更加貼近自身研究的資料使用需求。

（三）犯罪研究資料之分享、公開與再用

犯罪學者在分享或公開研究資料的經驗上，受到「人際交情關係」影響其分享的意願，學者傾向分享資料給熟識者，像是學生、同事、研究團隊成員或者學界內友人。而未有分享資料經驗的學者則認為在建立「合作關係」的前提之下，也能進行資料分享的行為。此外，談及分享或公開自身研究資料的顧慮，受訪學者主要的擔憂為資料遭到誤解或誤用、資料濫用以及保護受訪者隱私與個資等考量，因此才會選擇將資料分享給信任的熟識者，或有研究合作關係的對象，確保研究資料能夠被妥適的處理以及正確的詮釋，同時保護具機敏性質的資料。

另外，學者們於再用研究資料的行為模式上，多仰賴自己或研究團隊蒐集的一手資料（涵蓋政府官方資料的蒐集、下載），同樣地，在沒有特別交情的情況下，較少再用他人研究資料的經驗。不過，也有學者提出從研究主題的發想乃至於後續的資料蒐集活動，才是展現個人學問的不可或缺的研究實踐行為。另一方面，犯罪學者也基於擔心資料精確性，以及自己未參與研究設計、資料蒐集和清理過程，對資料的品質會存有疑慮，因此在研究資料的再用意願上較低。

三、犯罪研究資料中的信任、契約和價值

本研究案例揭示出「信任」、「契約」和「價值」三大核心要素，浸染於犯罪學者與研究資料的互動歷程之中，現說明如下：

（一）信任

1. 資料品質之信任

犯罪學者再用他人的研究資料時，會考量目標資料集本身的「精確性」和「資料處理程序（是否可信）」兩個面向，舉例而言，精確性涉及訪談資料與受訪對象真實經驗的吻合程度、資料建檔過程的完整性；而資料處理程序則涵蓋研究者個人或政府機關蒐集、清理、初步分析資料的過程是否處理妥當且透明公開。

2. 人際信任

「人際信任」關係也影響著學者蒐集、分享和再用研究資料的過程，學者在蒐集資料時，仰賴人際關係取得未公開的數據，或者獲得引介受訪者的機會，於分享研究資料階段，也會考量對方是否具備足夠的專業領域訓練與相似研究興趣，同時也受到友誼、師生關係、業務交情或過往合作經驗的影響，決定分享資料的對象和所分享內容的詳盡程度。

（二）契約

1. 合約規範

政府委託研究計劃案或合作備忘錄的簽署屬於合約規範的一環，合約的存在可以劃分清楚政府與學者間提供、使用資料的應盡義務與保障。而研究資料「所有權」和「使用權」也會在合約中一併加以規範，像是學者對結案後的資料進行改編或出版的權利，而這些合約內容也會影響學者分享研究資料的意願與態度。

2. 法律制度

學者在取得、蒐集具機敏性質的犯罪資料時，會受到法律制度的影響而必須倚靠其他方式獲取資料以執行研究，舉例而言，為保障兒少權益，政府得塗銷少年犯的前科資料（少年事件處理法第 83-1條），因此，學者只能由其他管道蒐集所需資料。此外，相關的個資隱私保護、政府犯罪資料的申請程序與使用原則，也都在法律制度的範疇之內。

3. 倫理道德

學者基於自身道德，為保護易受傷害族群與受訪者個資等理由，針對資料進行去識別化才願意分享，或持有更保守態度不願分享帶有受訪個資的研究資料；另外，也受到研究倫理委員會（Research Ethics Committee，簡稱 REC）或機構

審查委員會（Institutional Review Board，簡稱 IRB）等審查條件約束，必須依循當時送審的受訪者知情同意、研究資料的保密與銷毀程序等文件，不得任意將所蒐集的資料分享給他人。

4. 學者默契

此處學者默契指的是學者彼此之間「默示的契約」，犯罪學者在蒐集、使用、分享和再用研究資料時，彼此間基於瞭解並尊重其他學者蒐集、處理資料背後所付出的精力，因此通常倚靠自己一手蒐集資料，較無向他人索取資料甚至再用資料的經驗。

（三）價值

學者耗費心力一手蒐集、清理資料的種種活動，例如訪談資料清理與編碼、處理統計量化數據的遺漏值等，皆屬於為資料賦予價值的「加值」行為。而學者在分享、公開研究資料時，傾向在充分將資料的發表價值「利用完畢」後，才較有分享或公開資料予他人再用的意願。然而，犯罪學領域內非常經典的資料再用案例，Sampson 與 Laub 再用格魯克夫婦的研究資料而有了新的研究成果，並因此獲得斯德哥爾摩犯罪學獎這項殊榮（Stockholm Prize in Criminology, as known as “Nobel Prize” in Criminology），似乎又提醒了學術社群——

其實資料價值具有存續性、可在不同時空背景下發揮獨特的意義。

參、研究案例分享：中國文學學者研究之資料利用與移動[2]

人文學科是以研究人類文化和人類記錄的研究領域，辨識潛在可利用的資料作為研究證據。其中，資料如何經發現、聚合並淬鍊出與研究主題連結的詮釋邏輯，付諸於以人文學者自身角度切入；將資料與資料間的關係相扣、瀏覽並理解於各領域之資料內涵，從而將其轉變為人文領域研究意義的「資料」（Borgman, 2015）。

以下將分三部分撰述人文領域研究資料之展現，首先概括性闡釋出人文研究特性，再來則是探究人文資料的移動與摩擦力，最後為人文資料摩擦力與資料價值的疊加。

一、人文研究特性

人文研究性質主要以文本分析等質性資料資源進行研究。相對於量化資料來說，質性資料產生方式多來自於深度訪談、閱讀眾多文件，或是以觀察人事物與現象，再注入詮

[2] 內容摘錄自林潔筠（2020）未出版之碩士學位論文

釋者自身的經驗與研究意識，無法以數字進行統一表示特徵的描述性資訊且往往擁有研究者視角之主觀性。具體體現上，質性資料也因物件描述的深淺、擷取的要素、意會程度的不同而產生架構上的差異；例如：具結構的、半結構式的以及無結構式的資料構造。

「資料再現（reproducibility）」一詞從自然科學、社會科學領域漸受關注（Baker, 2016），承接上述人文科學在研究材料上，對文化涵義可因發散、聚焦來回轉移而產生多個有效答案，尋求超越於本質文字上的意義；因此，在文本與文本之間、經驗不同的互動下可能使資料的再現變得困難。再者，人文資料於集合、分析上具有較高度的資料使用專一性，歸因於研究者自身之研究背景資訊與經驗的獨一性強，其更是人文資料再用阻礙因素之一（Mauthner & Parry, 2009）。不過，正因人文研究常賴以研究者自身經驗與思想，注入對資料集的洞見；因此，資料必須得到保護：以概念性再現方法驗證先後研究結果的相對可靠性根據（Peels & Bouter, 2018）。

二、人文資料的移動與摩擦力

Edwards 等人（2011）首次正式提出並解釋了「資料摩擦力（data friction）」這一概念，他們將資料摩擦比喻為於

介質（surface）之間移動發生的相變與轉換；而此變化可以說是一種阻力。其緣由在於摩擦力表示著將統計數據轉換為資訊和知識所涉及的鬥爭與掙扎，資料每一次的移動都會耗損一些成本與努力，例如：資料遞交的過程中可能會因一些誤解、疏失或是不相容等阻力而產生亂碼或丟失（Edwards, 2010）。人文學領域資料因應新科技的使用與跨領域合作，資料付諸於領域學門的跨越與傳播交流也伴隨資料摩擦力的產生，抑或存在著人類與科技間的摩擦力（Taylor et al., 2014）；但卻也正是在跨領域的合作下，使得部分第一手文獻記載在面臨資訊不充分的困境與意境時，透過現代科技工具的應用得以補足闕失，還可能促使新的主題與延伸性研究產生（Jaton & Vinck, 2016）。

因此，Bates 在 2018 年重新審視「資料摩擦力（data friction）」的概念，他認為資料摩擦力的生成因有著使用者發現資料，並與資料互動、轉讓，進而幫助社會行為者進行資料的辨別、檢查，從而採取對該資料應用的適當作法，實際應用則體現在資料跨越各領域之間的價值轉變與移動過程。而了解資料移動過程間的變動、疊加應用與最終呈現，可為制定國際規範政策之各層面內涵提供重要參考，例如：資料使用聲明書的建立需求。

人文資料於原始轉變為加值移動過程，資料摩擦力便

將顯然的在資料移動與轉移中發生，產生能量並使資料內容性質發生變化；因此，「資料本身移動」的狀態與種類梳理有其必要性。資料移動為被動過程，是以資料經由被發現後離開原本因靜止不動的位置，從而被理解、消化、吸收與應用；而依照移動後的資料產製結果，本研究將人文資料以11種資料移動方式論述，分為資料內容性質變化與否，以及資料狀態間因時間因素具相承與連結移動等三種面向探討之，如下表二所示。

（一）資料性質不變：學者經由移動資料後，資料本身內容性質尚無任何明顯變化；舉如數位文檔的形式轉換、空間等移動，又如資料表象梳理的表格、楷化[3]等整理。

（二）資料性質改變：資料經由學者移動後並消化、理解與詮釋，資料內容性質開始產生價值上的加值變化；舉如：版本間內容理解、文本翻譯策略、句讀、點校加值等資料內容變化。

（三）資料移動的過渡、相承與穿梭：同一資料可能因著受訪者經驗的疊加、研究目的與情境的轉移

[3] 楷化：是繁體俗字的一種，將較不符合現今所閱讀方式的中文字體進行字型上的轉換（來源：教育部異體字字典）。

（repurpose），或是與不同想法、領域的接觸，資料即在這不同時空之間來回穿梭與再用（reuse）。

資料在經由使用時的轉移現象，以及發生移動時資料的內裡與表象呈現，包含著物理形式、位置轉移、狀態內容等變化。資料移動不一定伴隨著摩擦力產生，但資料摩擦力發生的契機存在於資料移動過程，則是最明顯的呈現；其中，又可分為資料在「點、線、面」等介質上的移動體現：

1. 資料在點：在人、資料、知識、基礎建設、組織、空間等之間轉換
2. 資料在線：在時間軸的進程的過渡、相承與穿梭
3. 資料在面：在學科、格式、思考層面的淬煉

奠基於資料於點線面之間的移動，進一步延伸以資料「內部（local）摩擦力」與「外部（global）摩擦力」具體化移動之表現。

（一）「內部摩擦力」為一種以學者為中心，在資料移動並加值的過程中所發生的摩擦力表現。資料加值（value-added data）即為資料自移動並使用後所

產生的附加價值，而資料價值與形式則是學者依憑個人研究慣習與需求，展現出資料內容與表面的變化；而無論資料性質轉變與否，摩擦力於無形當中也成為了資料過渡間的產物。人文研究中探究歷代間詞語的意義轉換與詮釋差異，對於查詢研究所需資料之首要任務，則是須辨認並整理「在現代理解下之意義於該朝代的用詞為何」；而此類關鍵詞的掌握有賴學者長期經驗的累積，一旦關鍵詞被確立，便是開啟古典與現代間的一把鑰匙，研究視野也隨之開拓。

（二）「外部摩擦力」為當資料投放至公開的場域後，經由學者在發現並取得資料利用時發生的情境中，所牽涉之環境、科技（如數位化）、社群等外部複雜因素造就下產生摩擦力；意即影響自身願意分享資料之因素，與自身再用他人資料的影響契機。舉如資料的跨領域或是翻譯上的移動摩擦，體現至人文研究則展現在文本性質翻譯與古籍內容加值；此移動承受的資料摩擦，會是在於需了解不同版本間的內容差異，更甚於牽涉至語言之間的轉換時，須了解翻譯相關理論與專業，如何吸收與消化仍有一定困難度。

表二、資料摩擦力發生介面與資料移動種類

資料摩擦力發生介面 資料移動種類	 點	 線	 面
(一) 物質	基礎建設之載體的運用所付出的努力	基礎建設的更迭，需了解不同的載體變化	(無發現)
(二) 地理空間	付出努力蒐集卻與自身期待的資料呈現不平衡	(無發現)	(無發現)
(三) 表象梳理	不同資料使用者間出現梳理的差異性，以及質的衡量很難判斷	(無發現)	(無發現)
(四) 外部領域	陌生的資料內容與發現過程的努力	時間的承載會形成多種不同類型的資料	困惑於蒐集的資料範圍
(五) 脈絡梳理	不同資料使用者間出現梳理的差異性，以及質的衡量很難判斷	對於文本詮釋隨著時間承載造成思想變化的可能；版本的陸續增加	學科間理解的差異，形成學術文化間的碰撞
(六) 文學性質翻譯	不同資料使用者存在著翻譯目的上的差異	翻譯概念的詮釋隨情境與時間承載而改變	翻譯比較文學等外領域知識碰撞

資料摩擦力發生介面 資料移動種類	 點	 線	 面
（七）古籍內容加值	不同資料使用者對版本的加注與理解差異	資料可能會為了符合時代理解而改變	（無發現）
（八）回溯性	（無發現）	（無發現）	（無發現）
（九）時光語料	（無發現）	（無發現）	（無發現）
（十）研究情境	資料使用者的經驗改變	資料範圍的無限擴增	跨領域合作的理解誤差
（十一）研究資料再用	不同資料使用者間的理解差異	資料原始樣態若無妥善控管，恐造成持續誤解	（無發現）

（註：表格中之描述，為資料移動於不同介質處所可能發生的摩擦力現象；其中「無發現」表示資料移動過程間，摩擦力於本研究中並無明顯發生的證據顯現。）

三、人文資料摩擦力與資料價值的疊加

資料本身價值的展現有賴於人類的發現與利用，而資料利用的過程則會因附加的多種元素形塑下產生不同樣貌。人文資料加值種類形式多樣，分類明確並部分移動具有承接

關聯；例如：初步進行資料檔案整理（表象梳理移動），再則是進行內容性質轉移，與時間過渡的資料利用等多樣移動加值之變化。人文資料經由多方移動的脈絡下所產製的摩擦力，展現出資料獨特且專一的豐富樣貌，漸而發散出多元的意象與世界觀。

再來，則是關注於研究資料的「選擇」與集合概念，其背後已蘊含著學者自身學識累積與經驗判斷；它是一項人文研究的起始點，更同時是智慧結晶產出的基石。經由學者的研究歷程可以發現人文資料移動過程的複雜性，資料的價值一再的被予以疊加，造就著資料的目的性愈趨強烈與專一，資料再利用的廣泛程度也漸縮小。然而，知識普及化亦是多數學者認為的學術資源最終使命之一，學術不只是將研究成果有所展現，也應於學術研究過程中，對於資料的應用與取得更加便利與開放；但也不免思考部分人文資料種類（如：出土文獻），雖存在著能夠驗證過往並挖掘潛藏未知理論，卻也因資料本身脆弱、稀有與獨特性出現部分典藏機構有意把持資料，而被世人、學者所忽略，甚至放棄理解。人文資料於數位化過程，須回歸至關注於開放科學環境下，秉持資料內容通透性的原則，包含公開透明的內容涵括，提供完整呈現並利用之（Lyon, 2016）。

肆、章節小結：人社科學資料的豐富圖像與研究資料基礎建設之啓發

從〈犯罪學開放研究資料之實踐與再用〉與〈中國文學學者研究之資料利用與移動〉兩個研究案例中，可看到人文社科學研究資料的特殊性，以及本國學者面對研究資料開放的潛在議題。

犯罪學探討真實世界中複雜的犯罪現象和犯罪問題，研究資料時常涉及研究對象的個人隱私，為了保護受訪者個資或防止資料遭人誤解濫用，學者對研究資料開放的觀點容易持保留態度，認為須確保符合法律制度、合約規範以及倫理道德，並傾向將資料分享給信任的對象。在資料的再用方面，犯罪學學者則認為運用他人研究資料涉及對資料品質的信任，且該領域學者間存在著一種默契，基於尊重他人研究所付出的心力，通常傾向自行蒐集一手資料，較缺乏再用他人資料之經驗。

中國文學學者的研究從文本分析出發，過程中以個人的知識背景與生命經驗，融入與其他文本間的互動，使得相同的文本經不同人的選擇、集合與解讀後，產出別具新意的詮釋觀點。正是因學者間對人文資料的運用與理解的角度不盡

相同，資料於各個學者或媒介的使用之間產生轉換、價值疊加，過程中也可能產生對資料的誤解、疏失等所謂的資料摩擦力，使得資料的開放與再用有所侷限。

本章嘗試以犯罪學與中國文學兩個領域，來探究國內學者所產製之社會科學與人文科學的研究資料特性，思考在提倡開放科學的環境下，若要建立人文社科學的研究資料基礎建設，可能需顧及資料隱私、資料品質，又或是資料經個人的價值判斷取捨、疊加後的可再用性，綜合考量為人文社科學術環境帶來更為開放、通透、具誠信的研究之有效做法。

從〈犯罪學開放研究資料之實踐與再用〉與〈中國文學學者研究之資料利用與移動〉這些研究成果中，可見到人文社會科學研究資料的獨特性及面對資料開放時所遭遇的挑戰：臺灣犯罪學學者在研究實踐上，反映了如何在維護隱私、法律與倫理框架下，謹慎地對待研究資料的開放；而中國文學學者的訪談中則揭示了文本資料在不同學者手中，經過選擇、集合與解讀後，如何產生新的價值與意義，同時也指出了資料移動過程中的摩擦力問題。

本章的兩個研究案例對於建構人文與社會科學之研究資料基礎建設有所啟發：研究資料基礎設施除了需齊備資料典藏庫之功能性、資料安全等，還必須加入有效的近用控制機制，使學者能夠有更高的自主性來控制自己的資料。此外，

尚須建立一套既能保障資料隱私與倫理，且國家政策、學術社群需建立起促進研究資料共享與再利用的機制。

在處理人文領域中的文本資料方面，研究資料基礎建設應當提供對後設資料與研究資料的記錄進行版本控制，以及支援語料資料中的語義標注功能，旨在降低資料轉移過程中可能出現的「摩擦力」，從而促進學者間的相互理解與溝通。

參考文獻

林山田、林東茂、林燦璋、賴擁連（2020）。*犯罪學*（修訂六版）。三民。

林奇秀、賴璟毅（2017）。臺灣社會科學學者資料再用行為之研究。*圖書資訊學研究，11*(2)，95-138。

許春金、楊士隆、周愫嫻、鄭瑞隆、沈勝昂、陳玉書、黃翠紋、謝文彥、黃富源、戴伸峰、曾淑萍、賴擁連、邱獻輝、董旭英、林明傑、孟維德、許華孚、黃蘭媜、郭佩棻……侯崇文（2016）。*刑事司法與犯罪學研究方法*。五南。

石易儒（2021）。*臺灣犯罪學開放研究資料及典藏庫之實踐*〔未出版之碩士論文〕。國立臺灣大學。

林潔筠（2020）。*原始資料至加值資料：人文資料移動中的摩擦力*〔未出版之碩士論文〕。國立臺灣大學。

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature, 533*, 452–454. https://doi.org/10.1038/533452a

Bates, J. (2018). The politics of data friction. *Journal of Documentation*, *74*(2), 412- 429. https://doi.org/10.1108/JD-05-2017-0080

Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked world*. Cambridge, MA: MIT Press.

Edwards, P. (2010). *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: MIT Press.

Edwards, P. N., Mayernik, M. S., Batcheller, A. R., Bowker, G. C., and Borgman, C. L. (2011). Science friction: Data, metadata, and collaboration. *Social Studies of Science, 41*(5), 667-690.

Jaton, F. & Vinck, D. (2016). Unfolding frictions in database projects. *Revue d'anthropologie Des Connaissances, 11*(4), 489-504. https://doi.org/10.3917/rac.033.0489

Lyon, L. (2016). Transparency: the emerging third dimension of Open Science and Open Data. *LIBER Quarterly, 25*(4), 153-171. https://doi.org/10.18352/lq.10113

Mauthner, NS & Parry, O. (2016). Qualitative data preservation and sharing in the social sciences: On whose philosophical terms? Australian Journal of Social Issues, 44, 291-307. https://doi.org/10.1002/j.1839-4655.2009.tb00147.x

Peels, R. & Bouter, L. (2018). The possibility and desirability of replication in the humanities. *Palgrave Communications, 4*(95), 95-98. https://doi.org/10.1057/s41599-018- 0149-x

Sutherland, E., & Cressey, D. (1978). *Criminology (10th ed.)*. Philadelphia: Lippincott.

Taylor, A., Fisher, J., Cook, B., Ishtiaq, S. and Piterman, N. (2014). Modelling biology - working through (in-)stabilities and frictions. *Computational Culture: A Journal of Software Studies, 4*. http://computationalculture.net/article/modelling-biology

Chapter
5

研究通透度與可再現性

章節作者：丁昱寧、高庭萱、李建欣、鄭瑋
章節編輯：林語婕

章節地圖

- ◆ 研究通透度
- ◆ 提升研究通透度與可再現性的國際組織與標準
- ◆ 研究通透度的實踐與利害關係人
- ◆ 章節小結

章節摘要

研究再現（reproducibility）近年為國際學術傳播、開放科學社群中熱門的議題，且此議題與各學術領域皆為相關。國際頂尖期刊如Nature與Science皆定期匯集有關「再現性」之主題文章與期刊論文，其編輯委員會亦多次發表編輯意見，闡述自身期刊對於再現性的積極實踐（Amaral & Neves, 2021; Berg, 2018; Vinson, 2021），以彰其支持開放科學之態度與立場。然而，綜觀國內，專以「可再現性」之概念與應用出發的專題研究仍在發展中。

本章簡介了研究通透度的定義與相關概念，如「R-words」、研究可再現性，闡釋研究通透度與開放科學間的關聯，並且以學術界現正面臨的再現性危機延伸，向讀者介紹國際組織和學術機構所發展出的提升研究通透度之原則標準與重要實踐等現況趨勢，以及促進研究通透度的舉措與利害關係人，提供讀者綜覽以對開放科學中的研究通透度議題有初步的認識。

壹、研究通透度[1]

一、何謂研究通透度

研究通透度（research transparency）它強調研究者應該以誠實、開放的態度，描述與分享研究過程，目標為增加研究的通透度。為了達到這個目標，研究者需要清楚地闡述資料蒐集的過程與分析方法，並且揭露研究的執行細節。這樣做不僅能夠增加研究的可信度，也有助於其他研究人員理解和驗證研究結果。

資訊學與生物領域學者 Lyon 是最早一批將研究通透度引入開放科學之學者。Lyon 於 2016 年提出開放科學的三面向，包含開放近用（open access）、參與（participation）、研究通透度（Research transparency），說明開放科學的體現應包含：科學研究成果對眾人的開放取得、公眾皆能參與研究進行的過程，以及研究資料具備公開與透明的特徵（Lyon, 2016）。

其中，研究通透度不僅是在「成果公開」的基礎上要求研究產品的釋出，其終極目標為提升研究再現性。因此，研究通透度在實踐上，更加重視在研究中的重大事件（critical event）紀錄、採取研究方法的理由，以及研究者在規劃一份

[1] 本章內容摘錄自高庭萱（2024）未出版之碩士學位論文

研究時，所設定的假設與參數、及更改實驗步驟的決定……等。總的來說，在實踐上，研究通透度也同樣注重研究資料的可得性，不過會更加強調實驗程序（research protocol）的方法細節，使研究的可複製性與再現性提高。

在研究具備通透度的前提下，研究人員得以向他人揭示其研究結論所根據的資料、理論與方法，使其他學者得以近用與分析其發現，評估該研究的描述與推論依據是否具說服力，以增進學者間的交流對話，亦有助於改善研究品質，促進學科方法、理論、資料蒐集的進步（Moravcsik, 2014），甚而解決研究信任與誠信的問題，提高研究的可再現性（Moore, 2023）。此外，若蒐集的資料對外公開可用，亦有機會使國際間的研究者重複利用這些資料，深入挖掘數據潛藏的其他科學潛力（Knottnerus & Tugwell, 2016）。

為了更有效地理解與評估研究的通透度，有些學者試圖拆解研究通透度在不同研究階段所需考量的角度，或是提出研究通透度的模型，供研究人員更容易理解該如何在研究過程衡量並提升自身研究的通透度。Tuval-Mashiach（2017）認為研究通透度發生於兩個階段，第一階段是研究者需在研究過程中不斷的自我反思，意識到自身採取每個步驟的原因，以及各決策可能對研究帶來哪些影響。第二階段則是在研究結果公開報告時，研究者需將自身的見解與反思清晰地傳達

給讀者。對此，Tuval-Mashiach 提出了旨在協助研究者提升通透度的通用框架，研究者在執行研究的所有階段包括規劃、資料蒐集、分析與撰寫，皆需持續反思自己：

- 「在研究中做了什麼（what）」
- 「研究是如何進行的（how）」
- 「為什麼會這麼做（why）」

這些反思不僅能夠加深研究者對自身研究的理解，也能夠讓他人更清楚地看到研究過程中每個決策背後的思考邏輯。同時，這些反思也形成紀錄，對外部的檢視來說，可作為稽核軌跡（audit trail），供他人檢驗研究者與研究問題及資料間的通透度。

其他的學者也將研究通透度拆分為不同面向，如 Moravcsik（2014）則是將研究通透度分成資料通透度、分析（analytic）通透度、產製（production）通透度等三個維度。「資料通透度」係指使讀者能夠近用支持該研究主張的證據或數據，藉以評估資料是否有被研究者正確解釋與分析。「分析通透度」為確保讀者能夠近用關於資料分析的資訊，了解研究者如何達到特定結論的推斷過程。「產製通透度」則是使讀者能夠近用研究者研究過程中，所有可能的選擇，以及最終所選的特定證據、論點、方法之原因，來評估研究者的資料蒐集與論點的產出，是否具有選擇偏見。

由上述可知，通透度於研究各階段皆有其重要性，且可拆分為不同面向。研究者不僅需要在規劃與執行研究時，經常反思自身資料的蒐集與選擇原因，確保最終導出的結論有所依據，也需盡可能將研究的相關資料與分析過程供他人檢驗，以提升研究的通透度，促進嚴謹的學術發展與交流。

二、R-words與再現性危機背景

研究通透度之所以受到重視，與「R-words」及其中的「可再現性」息息相關，以下詳述 R-words 概念與可再現性危機背景。

（一）各形各貌的R-words

學者於 2010 年代後期指出學術社群中「再現性危機」的存在（Baker, 2016），諸如科學可再現性（reproducibility）、可複製性（replicability）與可重複性（repeatability）等「R」開頭的字彙（R-words）亦逐漸受到矚目。在國際學術社群的共識中，R-words 的核心概念為「驗證一個科學研究能夠在相同或不同的條件下，再次產生相同的研究結果」。然而事實上，即便是在同一學科領域或部門之中，個別 R-words 的定義有時並不完全一致，甚至對個別 R-words 有不同的類型區分，如 Gómez 等學者（2014）從過往文獻中辨識軟體工程領域對於實驗可複製性（replicability）的認知，主要區辨出三種不

同的可複製性的類型：複製實驗與原實驗的相似度是否完全一樣、某些面向仍具相似性、只有理論架構一致等；而在牽涉跨學科、跨部門或不同研究方法（如是否涉及電腦運算）時，各利害相關群體對 R-words 的解釋往往存在更大的歧異（Peng & Hicks, 2021; Plesser, 2018）。

隨著相關概念的廣泛應用，更多 R-words 的變體亦隨之於學術討論中現蹤。具體而言，R-words 的示例還包括 reusability（可重用性）、re-analyzability（可重新分析性）、re-runnability（可重新運行性）、rigor（嚴謹度）、robustness（強健度）等。

本小節開頭所提及的三個字例是較常見的 R-words，如美國國家科學基金會（National Science Foundation, NSF）、國家衛生研究院（National Institutes of Health, NIH）以及多數與電腦運算相關的學科多使用 reproducibility（可再現性）一字，而 replicability（可複製性）一字則較易在社會科學中看見，如經濟學等。

略舉數例，表三呈現的是美國國家標準暨技術研究院（National Institute of Standards and Technology, NIST）、國家科學基金會（National Science Foundation, NSF）以及國際電腦機械學會（Association for Computing Machinery, ACM）等指標性政府單位或學術團體對表列三個 R-words 的解釋（ACM, 2020,

2021; Bollen et al., 2015; Taylor & Kuyatt, 1994）。

表三、各政府或學術單位對於不同 R-words 的定義示例（本書整理）

	Reproducibility（可再現性）	Replicability（可複製性）	Repeatability（可重複性）
NIST	研究中的測量原理、方法、儀器、場所、參照標準、觀察者等條件（conditions）皆可不同。	未提及	研究在以下配置可得相同結果： - 相同團隊 - 相同測量流程及系統 - 相同操作條件
NSF	研究在以下配置可得相同結果： - 相同流程（procedures） - 相同材料（materials）	研究在以下配置可得相同結果： - 相同流程 - 不同資料 - 相似分析工具	未提及
ACM	研究在以下配置可得相同結果： - 不同團隊 - 相同測量流程及系統 - 相同操作條件 - 相同或不同場所皆可	研究在以下配置可得相同結果： - 不同團隊 - 不同測量流程及系統 - 不同場所	研究在以下配置可得相同結果： - 相同團隊 - 相同測量流程及系統 - 相同操作條件 - 相同場所

綜而觀之，可以發現這幾個組織的定義存在細微的差異：首先，在研究的「構件」方面，各組織對於再現研究時所需保持一致的要素有不同的看法。有些組織強調研究方法、實驗環境、使用資料等因素的相似性，而其他組織則允許這些要素存在一定程度的差異。

各組織對於再現研究與原始研究之間的差異程度也有不同的定義。有些組織要求，再現研究應該與原始研究幾乎完全相同，而其他組織則認為即使存在一定程度的差異，只要關鍵的研究「構件」相同，就可以視為成功的再現。

最後，在研究再現的執行者方面，有些組織強調由原研究團隊自行進行再現，以確認研究結果並提高對研究發現的信心；而其他組織則鼓勵由外部團隊進行再現，以檢驗原團隊的成果，並發現可能存在的問題。

（二）不分領域的再現性危機

雖然科學界已普遍達成共識，認為研究過程與成果應該朝向更加開放和更通透的方向發展，但在實際執行上仍面臨著困難和挑戰。首先，私人利益的考量可能會阻礙研究資料的公開分享。例如，在生物醫學領域，研究可能涉及專利申請，因此研究者可能不願意公開詳細的研究細節，以保護其潛在的經濟利益。在分享過程中，製作共享資料所需的額

外時間和精力投入，也可能成為研究者的一大負擔。最後，公開分享的研究資料不一定具備可用性，導致執行再現研究時往往不一定能產生有效的研究結果（Hardwicke et al., 2018; Kidwell et al., 2016）。

在上述討論的基礎上，可以發現儘管科學界對於開放和透明研究的重要性已有共識，但在實踐中仍存在著諸多障礙。這些障礙不僅限於研究者個人的考量，如私人利益和額外工作量，還體現在公開分享的研究資料的可用性不足，導致再現研究難以取得有效結果。

然而，這些問題並非僅存在於少數學科領域，事實上，眾多學科領域陸續發現研究再現的結果不佳，以致產生正視「再現性危機」的呼籲（Baker, 2016）。如經濟學領域中的再現研究，發現即使作者公開研究資料和原始碼，也不能保證再現研究可以成功複製相同結果（Vlaeminck, 2013）。此外，Stagge 等（2019）實際複製水文學領域的期刊論文時，發現研究的可再現性程度很低，問題包含研究中提及可用的研究材料，實際上可以被取用的數量極少；以及研究中公開的數據與原始碼，能夠成功計算出相同結果的比率極低。亦有 Nüst 等人（2018）以地理科學領域進行再現研究，發現歐洲地理資訊科學年會（Association of Geographic Information Laboratories

in Europe, AGILE）中的期刊會議論文，總體的可再現比率不高，背後原因諸如研究者對於研究資料和原始碼的公開意願低落，多數僅願意公開其研究方法和結果等。

科學價值在於提供有用且受信任的知識成果，倘若研究資料無法展現其透明與開放，則無法被驗證其真實性，亦無法執行可再現性研究（Breznau, 2021; Stanovich, 2019）。當研究者意識到自身的研究內容將受到外界的檢驗與再現，會促使他們提供實際可用與可信任的資料集，並減少不誠實行為的發生（Ishiyama, 2014）。因此，提升研究資料的通透度及可再現性不只作為科學態度的展現，也能夠導正科學研究過程中所出現的不公正或投機取巧的行為。

貳、提升研究通透度與可再現性的國際組織與標準

隨著學術社群對研究的通透度、再現性，以及研究資料共享與開放等概念愈趨重視，相關議題的指引與標準在國際間便應運而生。本節為讀者介紹影響學術單位及研究人員增強其研究之通透度的重要組織沿革及其催生出的標準。

一、國際組織

學術界為了引導機構制定研究資料管理的相關措施，建

立了各項原則。例如：由歐洲地區萌芽的 FAIR 原則，為歐洲科研環境建立資料取用與使用上的核心價值；而北美地區也有美國國家衛生研究院以資料分享、管理和取用為核心，所研擬的各項策略和設備。

（一）FAIR 原則：奠定歐洲科學社群資料共享的基石

以克服資料取得與應用上的障礙為目標，於 2014 年荷蘭萊頓所舉辦的「Jointly Designing a Data Fairport」研討會中，首次提出了 FAIR 原則的草案內容，並於 2016 年由 Wilkinson 等人在 *Scientific Data* 期刊中發表了 FAIR 原則的核心概念及相關內容（Wilkinson et al., 2016）。FAIR 原則的重點包含資料應具備 Findability（可查找）、Accessibility（可近用）、Interoperability（可互操作）以及 Reusability（可再用）四大核心原則。實踐 FAIR 原則並推動研究資料達到 FAIRness 的理想狀態，得以促進研究資料間的獲取與再使用，如資料發布者能夠藉由採用 FAIR 原則，協助完善資料取用與管理上的基本要素與條件，使得研究資料能夠被有效取用與使用，以支持知識的發展與創新（Wilkinson et al., 2016）。

FAIR 原則提出後，旋即獲得國際學術界的普遍支持。如地球科學領域因面臨觀測數據的大量增長、及取用他人觀測資料的困難，體認到資料於共同取用和使用上的重要性。當

FAIR 原則提出後，地球與太空科學資料出版聯盟（Coalition for Publishing Data in the Earth and Space Sciences, COPDESS）就以其作為出版資料時的最高指導原則。Stall 等人（2019）認為促進地球科學領域的資料達到 FAIRness 可以協助社會性資訊與服務的提供，像是接收世界各地氣象資料、使用地球科學相關資料建立全球定位系統，及藉由各地觀測到的地震數據以緩解和驗證地震災害等。

除此之外，FAIR 原則的釋出也為國際機構提供更多可遵循的標準與發展方向。例如，歐洲開放科學雲（European Open Science Cloud, EOSC）[2] 視 FAIR 原則為重要基礎，期望為歐洲的科研環境建構一個符合 FAIR 原則的網路與資訊服務（European Open Science Cloud, n.d.）。此外，荷蘭與德國更是成立「GO FAIR 國際支援與合作辦公室（The GO FAIR international support and coordination office, GFISCO）」，來協助推動歐洲開放科學雲的發展。GO FAIR 計畫由 FAIR 的四個核心概念，發展出相匹配的研究設備與研究流程，促進全球研究資料的可查找、可近用、可互操作和可再用，為學術環境建構一個開放和包容的研究生態系統（The GO FAIR

[2] 歐洲開放科學雲（European Open Science Cloud, EOSC）是歐洲委員會（European Commission）於2016年以提供全歐洲的研究人員、機構、公司與人民開放的資訊使用環境為目標所建立的一項計畫。

international support and coordination office, n.d.）。而 GO FAIR 計畫也由歐洲地區擴展其影響，陸續延伸至美洲、亞洲與非洲等地區，促進全球研究資料間的共享與通用，減少研究資料往來時的隔閡與限制。

（二）政策推動的資料共享

除了歐洲地區外，在研究資料的共享以及對通透度的重視與發展上，美國於 2013 年也由白宮科學與技術政策辦公室（White House Office of Science and Technology Policy, OSTP）制定了一份說明文件，內容敘述在可執行且符合法律和政策的情況下，社會大眾、企業和學界可以自由取用由聯邦資助的科學研究所產生的研究資料（Holdren, 2013）。該文件釋出後，美國許多研究機構陸續響應資料分享與通透度之概念，並建立相關策略與計畫。例如，美國國家衛生研究院（National Institutes of Health, NIH）[3]以促進資料共享與研究通透度的發展為目標，於 2016 年制定資料管理和共享策略（NIH Policy for Data Management and Sharing and Supplemental Information）。這份計畫以廣泛蒐集大量志願者的想法為基

[3] 美國國家衛生院成立於1887年，隸屬於美國衛生及公共服務部，是一個重要的生物醫學研究機構。該機構使命在於探討與生物系統相關的行為和知識基礎，並以增強人類健康、減輕疾病和其他造成之障礙或困難等為目標。

礎，並透過反覆與志願者來回討論內容來修訂草案。最後，美國國家衛生研究院於 2020 年 8 月釋出最終版本的資料管理和共享政策（National Institutes of Health, 2020），內容討論資料長期的保存價值，以及長期性的管理策略等。

除此之外，為鼓勵生醫領域的研究與服務符合 FAIR 原則，美國國家衛生研究院於 2015 年建立了一個「共用元素數據資料庫」平台（NIH Common Data Elements Repository, CDE Repository），鼓勵研究人員使用通用之資料元素，以提高生醫研究資料之間的準確性、一致性和互操作性，將臨床與研究資料更有效地相互整合（National Institutes of Health, n.d.）。研究人員可以在該平台中，相互分享生醫數位物件（digital objects）以及研究所需的元素。最後，美國國家衛生研究院也為其所補助的研究計畫產生的大量數據與資料，進行現代化的管理與應用發展。其機構內的資料科學策略辦公室（Office of Data Science Strategy）於 2018 年發布了一份「資料科學策略計畫」，以支持高品質的生醫研究資料基礎建設、促進資料生態圈的發展、開發先進的資料管理、分析和視覺化工具、加強生醫科學的人力資源開發，與制定合適的資料政策等，期望生醫領域的研究資料得以被永續保存與使用，促進科學的開放與永續發展。

二、學術機構

除了政府機構對促進研究資料的通透、使用、管理與保存的重視之外，許多國際性的學術機構也加入支持研究資料朝向開放與透明的行列，並提出相關的期刊政策與規範。

（一）期刊編輯通透度條約（JETS）

以激勵資料分享和通透度為目標的國際組織——DA-RT（The Data Access and Research Transparency Statement, DA-RT）[4] 於 2010 年成立，期望組織內的成員維護學術誠信，並投入研究通透度與再現性相關議題之發展，進而建立一個多元化的研究社群。

DA-RT 為資料分享和研究通透度做出的貢獻如：2014 年建立的期刊編輯通透度聲明（Journal Editors Transparency Statement, JETS，後續以 JETS 稱之），規範期刊應該要求研究者盡可能提供研究資料內容，以提高研究資料的可取用性及通透度。詳細條列如下：要求作者公開其研究中所運用的資料，並確保這些資料被存放在可供取用且受信任的資料庫、

[4] DA-RT 屬於美國政治學協會旗下的一個特設委員會。DA-RT 由來自於不同學校、機構的研究人員、期刊編輯、專業學者所組成，其成員結構分散，為組織提供不同面向的貢獻（The Data Access and Research Transparency Statement [DA-RT], 2015）。

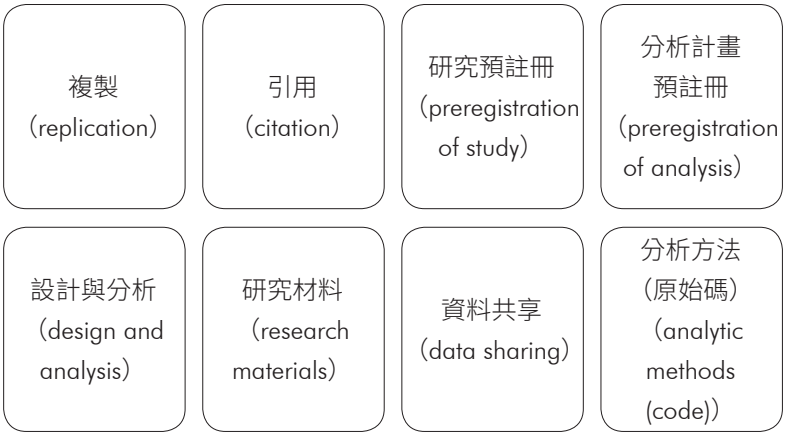
研究中的分析過程與策略應該被公開展示且妥善保存、維持一致的資料引用規範，確保資料引用的內容正確性與可使用性等（DA-RT, 2015）。自 JETS 提出後，陸續有許多國際期刊發表採用 JETS 條約的聲明，且由於 DA-RT 組織來自政治學科領域，因此採用 JETS 的期刊也多數來自政治領域。

（二）TOP Guidelines：八項標準與三個層級

另一方面，除了官方組織和國際組織之外，國際間的學術社群也積極推動研究通透度與再現性。國際間推廣研究通透度與再現性的社團與組織中，以 Nosek 等（2015）提出的資料通透度與開放性促進標準（Transparency and Openness Promotion Guidelines, TOP Guidelines，後續以 TOP Guidelines 稱之）[5] 為先驅者，自 2015 年建立至今，TOP Guidelines 已是一項目前被學術期刊與機構廣為接納與採用的標準。該標準以促進研究成果的開放與通透為目的，提出八項標準（Nosek et al., 2015），而根據八項標準的內容，又可將其歸納成三個部分討論，第一組敘述複製研究和資料引用的標準，第二組說

[5] TOP Guidelines由透明度和開放性促進委員會於2014年11月在維吉尼亞州夏洛茨維爾的開放科學中心（Center for Open Science in Charlottesville, Virginia）舉行會議，討論「期刊程序和出版政策」的激勵標準。該標準由領域權威、期刊編輯、獎助機構，以及社會科學和行為科學領域中的學科專家組成之委員會，建立期刊實踐開放共享的標準，希望能改變當前的學術制度，制定出各領域皆可使用的標準，提升學術界對於研究通透與再現性的重視。

明研究與分析計畫的預註冊，最後一組則是討論研究資料、材料與分析方法等（見圖四）。各項標準的概述如下：



圖四、TOP Guidelines 的八項標準

圖片來源：本書整理

首先，**複製**（replication）和**引用**（citation）兩項標準關注研究人員在實踐再現性科學過程中，所花費的時間和精力。舉例而言，複製標準提及研究的可被複製對於驗證研究結果的價值，也提到期刊對於複製研究的接納與發表狀態；而引用標準關心期刊文章對於資料、原始碼和研究材料等的引用程度。

接著，有兩項標準與研究之預註冊[6]（pre-registration）有關，分別是**研究預註冊（preregistration of study）**與**分析計畫的預註冊**（preregistration of analysis）。預註冊是一種透過過預先釋出研究內容，而確保研究依照計畫正確進行的投稿方式。在預註冊的文件中，會記錄資料蒐集方法、資料清理規則和分析方法等細節，透過預先為研究制定重要決策與釐清研究問題，幫助減少研究過程可能出現的偏差與不實報告。在研究預註冊標準中，說明了研究過程中需要被預先登錄的內容，同時亦確保預註冊文件被記錄在公開的系統中，能夠被其他人檢視；而分析計畫的預註冊與研究預註冊標準的內容相近，但更著重於研究中對於資料分析計畫的提前紀錄。

此外，分析計畫的預註冊證明了確認性研究與探索性研究之間的區別，即所謂的假設檢驗與假設產生的研究區別。透過為確認性研究和探索性研究之間的區別增加透明度，可以加強這兩項研究的可再現性（Asendorpf et al., 2013;

[6] 預註冊並非註冊研究報告（pre-registered reports）。註冊研究報告在一些響應開放科學之期刊中，作為一種投稿文體。作者在研究發想、設計研究方法和分析計畫後，先將研究問題、研究分析參數等資料撰寫為註冊研究報告，期刊收件後會直接以註冊研究報告的形式發送同儕審查。若通過同儕審查，即為原則性接受（in principle acceptance, IPA）。倘若作者後續的研究過程按照原訂計畫進行，撰寫完研究結果的稿件會再進行一次同儕審查，以待確保研究步驟與先前註冊報告是否有落差。此時，若研究步驟悉數一致時，無論成果的好壞、假設是否成立，期刊都會遵循 IPA，原則性刊登稿件。

Wagenmakers et al., 2012）。

最後，有四項標準解釋了開放與通透度對於科學研究過程的重要性。**設計與分析**（design and analysis）標準的訂定強調提高研究過程的通透度、降低研究設計或分析方法的模糊或不完整；**研究材料**（research materials）標準注重研究中所提供與研究方法有關的所有要素及材料的公開；**資料共享**（data sharing）標準則激勵作者在信任的資料庫或典藏庫中分享其研究資料，例如Dataverse、政治與社會研究大學校際聯盟（Inter-University Consortium for Political and Social Research, ICPSR）、開放科學框架（Open Science Framework, OSF）等；**分析方法（原始碼）**（analytic method (code）標準中則關注用於研究的統計模型或虛擬原始碼的開放使用。

為使各領域的期刊都能夠採用此標準，上述的八項標準皆已經過謹慎的字詞裁量，不同的標準之間亦可相互補充。此外，為了因應學科領域的多元與差異性，透明度和開放性促進委員會在每個標準之下又定義了三個層級（Nosek et al., 2015）：層級一（level 1）是該標準最低的採用方式，對期刊而言幾乎沒有採用上的困難；層級二（level 2）較上層更為詳細與限縮，但同時也留意盡可能不增加編輯者或出版者的成本；層級三（level 3）是最嚴格的標準，此標準在實務層面可能就會為期刊帶來執行上的難處。例如，在分析方法（原始

碼）共享標準（analytic methods (code) transparency）中，層級一規範作者必須在研究內容中提及原始碼是否可用以及在何處可用；層級二則要求期刊將共享之原始碼存放在可信任的資料庫中，並檢查研究內容中是否具有共享原始碼的有效連結；層級三最為嚴格，規定研究使用之原始碼除了要放置於可信任資料庫之外，研究分析結果還要能夠在研究發表前能被獨立複製。

TOP Guidelines 的制定，為學術界提供了研究通透度與再現性的實踐管道。透過建立與實踐期刊政策，期刊可以作為推動通透度的重要助力，提升研究人員對於研究通透與再現的敏感度，並展現期刊對研究誠信之重視（Baker, 2016；Simmons et al., 2011）。學術界也積極採用 TOP Guidelines 標準，以展現對於科學開放與通透的重視，如美國心理學會（Association of Psychological Science, [APS]）便認為 TOP Guidelines 是一套改進學術出版和研究實踐的具體指南。此外，多個國際性的出版機構，如：Springer Nature、BioMed Central、Science 等，在 TOP Guidelines 釋出後隨即進行各項標準的運用裁量，並且發布聲明表示對採用 TOP Guidelines 之積極實踐，以及對資料通透與開放的重視。現今，越來越多的期刊已將 TOP Guidelines 內的各項標準納入期刊政策的考量，並要求作者積極實踐，且認為採用 TOP Guidelines 對於促進研

究內容的透明、品質和可信度皆有幫助。

參、研究通透度的實踐與利害關係人

在學術界推動研究通透度的過程中，期刊編輯扮演著關鍵的角色。實踐研究通透度除了是研究者對科學社群與讀者的責任，在學術出版過程中，亦有其他角色可對提升研究通透度發揮影響力，共同協助與監督研究者落實更為通透的研究，以下介紹各種在學術傳播過程的利害關係人，能如何透過體制的建立，來共同推動整體的學術研究朝向更為通透的方向發展。

一、獎助機構

研究獎助單位可透過制定政策，要求研究人員將其研究過程透明化，分享與研究結果相關的所有資料，以促進其贊助的研究通透度（Gennetian et al., 2020）。例如歐盟的 Horizon 2020 計畫著重於卓越的科學研究、產業領導力與應對社會挑戰的主題，獲得該計畫資助的研究者，必須使他們的研究資料符合 FAIR 原則，供他人可查找（Findable）、近用（Accessible）、可互操作（Interoperable）和可再用（Reusable）（Basile et al., 2023）。又如世界衛生組織（World

Health organization, WHO）、歐洲核子研究組織（European Organization for Nuclear Research, CERN），以及美國國家衛生研究院（National Institute of Health, NIH），於 2022 與 2023 年起皆要求研究人員提交資料管理與分享計畫，讓研究資料可被他人取用（何明誼，2023）。這些政策皆有助於規範研究人員開放可供他人驗證與再現研究的資料，提高科學界整體的研究通透度。

二、期刊編輯

期刊編輯同樣可透過制定規範與指南，要求研究者遵循相關規定以促成研究通透度。例如歐洲政治科學期刊（European Political Science, EPS）與美國語言聽力協會（American Speech-Language-Hearing Association）旗下的期刊，皆以 Nosek 等（2015）提出的 TOP Guidelines 為基礎，訂定一套要求研究者在資料取用與通透度方面的遵循指南，促進期刊發展開放科學（Basile et al., 2023; Schroeder et al., 2023）。

此外，同儕審查的形式轉變，也是期刊促進研究通透度的重要手段之一。在學術傳播的過程中，研究內容的真實與可行性會透過專家的角度來謹慎評估，以確保研究的誠實與信用；也透過審查者的檢閱與評論，確保有瑕疵的內容不會被輕易發布，而這項以專家角度來對研究品質進行評估的程

序在學術界稱為同儕審查（peer review）。它為期刊編輯提供是否收錄投稿文章的判準，也為科學家提供一個發表前糾正錯誤的機會，是學術交流與出版過程中重要的一環。一般常見的執行方式依審查過程中審查者和作者身份的披露程度來劃分，可以分為單盲（single-blind）和雙盲（double-blind）同儕審查[7]。而此兩項機制又因為審查者毋須揭露身份，使得審查結果容易產生審查者的偏見、審查動機不足與內容剽竊等問題。基於對審查結果的不信任與質疑，以及網路科技發展下的載體變革，學術界出現許多新穎的作法，如：預印（pre-publication）同儕審查、出版後（post-publication）同儕審查以及開放式同儕審查（open peer-review）（黃慕萱、嚴竹蓮，2016）。

其中，開放式同儕審查打破傳統單、雙盲審查的運作機制，藉由揭露作者和審查者的身份和審查內容，增進審查過程的通透度，強化學者與審查者兩者間的課責性，降低單、雙盲審查產生的偏見、內容剽竊和攻擊性評論等問題，敦促雙方進行正向的學術交流。

無論期刊採用何種類型的同儕審查，都是藉由評鑑研究

[7] 在單盲審查中，審查者知道作者身份，而作者不知道審查者身份，被視為學術出版中傳統且最常見的審查模式（Ware, 2011）。而在雙盲系統中，為避免審查者單向發生的個人偏好或缺失，審查者和作者之間彼此都維持匿名。

內容來確保其符合一定程度的可靠性，以維護科學價值。故隨著學術界對於研究通透度更加重視，審查制度也為了避免人為因素的操縱和不當行為的發生，逐漸轉以透明公開的方式執行。

三、大學與研究機構

研究人員所處的研究機構亦是影響研究者的行為與研究通透度的關鍵。大學可以為研究生與博士後研究員提供研究通透度與研究可再現性的課程，培訓其支持開放科學實踐（Krishna & Peter, 2018）。而大學的教職員聘用、晉升與終身聘用評估，若獎勵透明地發表所有研究結果並公開共享資料、原始碼、程序與其他研究材料，亦有助於提高研究者的意願採取更為透明的研究實踐（Moher et al., 2018）。

肆、章節小結：研究通透度與再現性為臺灣科研社群的共同使命

回顧國際間對無論是政府機構、學術單位抑或民間組織，都逐漸將研究資料的管理與共享視為重要發展重點，視科學研究的通透度與再現性為重要科學內涵。國際間無論是學術機構或是國際組織都紛紛響應相關政策與計畫的制定與

執行，為學術界提供研究通透度與再現性的實踐管道，期望藉由期刊與學術機構作為推動這些概念的重要助力，來提升研究人員對於研究通透與再現的敏感度，並展現期刊對研究誠信的重視（Baker, 2016; Simmons et al., 2011）。然而，回顧臺灣科研領域現況，雖然對資料開放與再現的意識逐漸抬頭，但相關政策與組織仍未發展成熟；此外，國內學術期刊對於研究通透度和再現性的相關標準也尚未採用。面向國際社群對於研究通透度與再現性的討論熱潮，我國未來也應積極正視國內學術環境的現況，探討國內學者、研究社群與組織對於研究通透度與再現性的認知與看法，以及採用提升研究通透度與再現性相關的政策考量。

本章簡介研究通透度的定義與可再現危機的背景，說明研究通透度對改善研究可再現性之重要性，並介紹促進研究通透度的國際組織與標準，顯示研究通透度議題於國際間的受重視程度，最後則是闡釋學術傳播過程中，可能影響研究通透度的各種角色，體現出提升研究通透度的責任需由多方人員共同努力，從相關規範制定到執行監督，皆是落實研究通透度的重要關鍵。當研究人員視研究資料的開放共享、通透度與再現性為一項習以為常的科學態度，且願意實際在研究過程中致力執行，改善過去研究中容易產生之問題研究行為，便有助於維護科學誠信，提高研究結果供他人檢驗與利

用之根據，延續科學研究之永續發展。

參考文獻

何明誼（2023 年 7 月 19 日）。*讓資料共享成為責任　簡介國際近期的科研資料開放政策*。科學月刊。https://www.scimonth.com.tw/archives/6512

黃慕萱、嚴竹蓮（2016）。同儕審查的起源，研究現況與展望。*圖書資訊學刊，14*(1)。https://doi.org/10.6182/jlis.2016.14(1).041

Asendorpf, J. B., Conner, M., de Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M., van Aken, M. A. G., Weber, H., & Wicherts, J. M. (2016). Recommendations for increasing replicability in psychology. In A. E. Kazdin (Ed.), Methodological issues and strategies in clinical research (pp. 607-622). American Psychological Association. https://doi.org/10.1037/14805-038

Association for Computing Machinery (ACM). (2020). Artifact review and badging - Current. https://www.acm.org/publications/policies/artifact-review-and-badging-current.

Association for Computing Machinery (ACM). (2021). Artifact review and badging - Version 2.0. https://www.acm.org/publications/policies/artifact-review-badging.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature, 533*, 452-454. https://doi.org/10.1038/533452a

Basile, L., Blair, A. & Buckley, F. (2023). Research transparency and openness. *European Politcal Science, 22*, 177-181. https://doi.org/10.1057/s41304-023-00424-x

Bollen, K., Cacioppo, J. T., Kaplan, R. M., Krosnick, J. A., & Olds, J. L. (2015). Reproducibility, replicability, and generalization in the social, behavioral, and economic sciences: Report of the Subcommittee on Replicability in Science Advisory Committee to the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences. Arlington, VA: NSF.

Breznau,N. (2021). Does Sociology Need Open Science? *Societies, 11*(1), 1-25. https://doi.org/10.3390/soc11010009

DA-RT (2015, June). About DA-RT. https://www.dartstatement.org/about

European Open Science Cloud (EOSC) (n.d.). What the European Open Science Cloud is. https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc_en

Gennetian, L. A., Tamis-LeMonda, C. S., & Frank, M. C. (2020). Advancing Transparency and Openness in Child Development Research: Opportunities. *Child Development Perspectives, 14*, 3-8. https://doi.org/10.1111/cdep.12356

GO FAIR. (n.d.). GO FAIR Initiative. Retrieved May 29, 2022 from https://www.go-fair.org/go-fair-initiative/

Gómez, O. S., Juristo, N., & Vegas, S. (2014). Understanding replication of experiments in software engineering: A classification. *Information and Software Technology, 56*(8), 1033-1048. https://doi.org/10.1016/j.infsof.2014.04.004

Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., Mohr, A. H., Clayton, E., Yoon, E., Tessler, M. H., Lenne, R. L., Altman, S., Long, B. & Frank, M. C. (2018). Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal Cognition. *Royal Society open science, 5*(8). https://doi.org/10.1098/rsos.180448

Holdren, J. P. (2013). Memorandum for the heads of executive departments and agencies: Increasing access to the results of federally funded scientific research. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Ishiyama, J. (2014). Replication, research transparency, and journal publications: Individualism, community models, and the future of replication studies. *PS:Political Science & Politics, 47*(1), 78-83. https://doi.org/10.1017/S1049096513001765

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L. S., Kennett, C., Slowik, A., Sonnleitner, C., Hess-Holden, C.,

Errington, T. M., Fiedler, S., & Nosek, B. A. (2016). Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLoS biology, 14*(5). https://doi.org/10.1371/journal.pbio.1002456

Krishna, A., & Peter, S. M. (2018). Questionable research practices in student final theses–Prevalence, attitudes, and the role of the supervisor's perceived attitudes. *PLoS One, 13*(8). https://doi.org/10.1371/journal.pone.0203470

Knottnerus, J. A., & Tugwell, P. (2016). Promoting transparency of research and data needs much more attention. *Journal of Clinical Epidemiology, 70*, 1-3. https://doi.org/10.1016/j.jclinepi.2016.01.007

Lyon, L. (2016). Transparency: The Emerging Third Dimension of Open Science and Open Data. *Liber Quarterly: The Journal of European Research Libraries, 25*(4). https://doi.org/10.18352/lq.10113.

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biology, 16*(3). https://doi.org/10.1371/journal.pbio.2004089

Moore, L. (2023, January 23). *Why transparency in open research is more important than ever*. Wiley. https://www.wiley.com/en-us/network/publishing/research-publishing/open-access/why-transparency-in-open-research-is-more-important-than-ever

Moravcsik, A. (2014). Transparency: The Revolution in Qualitative Research. *PS: Political Science & Politics, 47*(1), 48-53. https://doi.org/10.1017/S1049096513001789

National Institutes of Health (NIH) (2020). Final NIH Policy for Data Management and Sharing. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

National Institutes of Health (NIH) (n.d.). The CDE Repository. https://cde.nlm.nih.gov/home

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck,S., Chambers, C. D., Chin,G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B.,

Humphreys, M., & Yarkoni, T.(2015). Promoting an open research culture. *Science, 348*(6242), 1422-1425. http://doi.org/10.1126/science.aab2374

Nüst, D., Granell, C., Hofer, B., Konkol, M., Ostermann, F. O., Sileryte, R., & Cerutti, V. (2018). Reproducible research and GIScience: An evaluation using AGILE conference papers. *PeerJ, 6*, e5072. https://doi.org/10.7717/peerj.5072e5072.

Peng, R. D., & Hicks, S. C. (2021). Reproducible Research: A Retrospective. *Annual Review of Public Health, 42*(1), 79-93. https://doi.org/10.1146/annurev-publhealth-012420-105110

Plesser, H. E. (2018). Reproducibility vs. Replicability: A Brief History of a Confused Terminology. *Frontiers in Neuroinformatics, 11*, 76. https://doi.org/10.3389/fninf.2017.00076

Schroeder, S. R., Gaeta, L., El Amin, M., Chow, J. C., & Borders, J. C. (2023). Evaluating Research Transparency and Openness in Communication Sciences and Disorders Journals. *Journal of speech, language, and hearing research. JSLHR, 66*(6), 1977-1985. https://doi.org/10.1044/2022_JSLHR-22-00330

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Stagge, J. H., Rosenberg, D. E., Abdallah, A. M., Akbar, H., Attallah, N. A., & James, R. (2019). Assessing data availability and research reproducibility in hydrology and water resources. *Scientific data, 6*(1), 1-12. https://doi.org/10.1038/sdata.2019.30

Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Lehnert, K., Nosek, B., Parsons, M., Robinson, E., & Wyborn, L. (2019). Make scientific data FAIR. *Nature, 570*(7759), 27-29. https://doi.org/10.1038/d41586-019-01720-7

Stanovich, K. E.（2019）。*這才是心理學！*（楊中芳譯；三版）。臺北市：遠流出版事業股份有限公司。（原著出版於2019年）

Taylor, B., & Kuyatt, C. (1994). Guidelines for evaluating and expressing the uncertainty of NIST measurement results 1994 Edition. https://doi.org/10.6028/NIST.tn.1297

Tuval-Mashiach, R. (2017). Raising the curtain: The importance of transparency in qualitative research. *Qualitative Psychology, 4*(2), 126-138. https://doi.org/10.1037/qup0000062

Vlaeminck, S. (2013). Data management in scholarly journals and possible roles for libraries–Some insights from EDaWaX. *Liber Quarterly, 23*(1), 48-79. http://hdl.handle.net/10419/106609

Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. *Perspectives on psychological science, 7*(6), 632-638. https://doi.org/10.1177/1745691612463078

Ware, M. (2011). Peer review: Recent experience and future directions. *New Review of Information Networking, 16*(1), 23-53. https://doi.org/10.1080/13614576.2011.566812

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., Santos, L. B. S., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C., Finkers, R., …, Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, *3*(1), 1-9. https://doi.org/10.1038/sdata.2016.18

Chapter
6

新興科技輔助增強研究通透度

章節作者：黃寶霈、龔俐霏、張韻琦、鄭瑋
章節編輯：黃郁雯

章節地圖

- ◆ 研究再現性危機
- ◆ 區塊鏈技術輔助研究通透度
- ◆ 為 BLT 打造良好使用者體驗
- ◆ 章節小結

章節摘要

在當代科研環境中，研究再現性危機（reproducibility crisis）成為學術界亟需面對的挑戰。為提升研究通透度並促進再現性，本章承接上一章對於研究通透度的討論，將焦點轉向新興科技在輔助研究通透度設計方面的實際應用案例。本章將為讀者深入介紹基於區塊鏈技術的 BLT（Blockchain for Lifecycle Transparency）框架，串連開放科學工具並實現資料的分享與提升個別研究的通透度。

BLT 框架係利用區塊鏈去中心化的本質設計，能自動驗證研究過程並產生不可篡改的紀錄，研究者得以依據各式需求自訂工作流程與智慧合約。藉由持續精進 BLT 平台並推動其在學術出版上的應用，展現其在提升研究通透度和再現性方面的潛力，為學術界提供新興科技輔助增強研究通透度的解決方案。

壹、研究再現性危機

承接上一章對於研究通透度的討論，當代重視的研究再現性，除了關注研究能否被其他團隊再現，而是更進一步探討科研結果的真實性——即研究成果是否如科學家所宣稱的那般可靠。

當科研結果不具有再現的可能，學術界便陷入了所謂的「再現性危機」當中，不只是浪費當初各方投注研究的資源，後續作者亦會受到他方的質疑與批判，甚至損害整個學術界的信譽。最廣為人知的研究再現性計畫案例之一，是 2015 年由開放科學中心（Center for Open Science）一群學者所執行的心理學領域研究，他們重複執行 100 個已發表的心理學研究，發現只有 36% 的研究具有統計顯著性，這與原始團隊達到 97% 統計顯著性的研究具明顯的差異（Open Science Collaboration, 2015）。類似的情況也出現在臨床前研究中，據估計約有一半的研究無法被重複驗證，每年因此而耗費約 280 億美元（Freedman et al., 2015），相當於 9,000 億新台幣（以 1 美元兌 32 新台幣匯率計算）。

這些發現令人震驚，同時也引發了對研究過程中可能存在的問題：研究者在研究歷程中似乎遇到多種妨礙後續研究

再現的情況。透過對研究的各個階段進行剖析，可以發現研究者可能面臨不同的挑戰：

一、研究早期設計與發想階段：可能存在假設不明確、樣本量不足等問題

研究者未能控制自身的偏見、選擇錯誤的方法、逕行挪用他人想法，甚至建構出最初根本就不存在的研究，如荷蘭心理學家 Diederik Stepel 詐欺案便是著名的例子，造假出 55 篇論文，由其指導的 10 篇學位論文亦使用假的研究數據（Enserink, 2012）。

二、研究中期資料蒐集與分析階段：可能出現 p-hacking 數據篩選偏差、測量誤差等情況

研究者可能竄改資料、捏造不存在的資料或受試者同意書、蒐集與控制資料的品質不佳、為達成理想具顯著性的 p 值而進行資料捕撈行為（p-hacking 與 data dredging）。

三、研究晚期成果撰寫與出版階段：可能存在選擇性報告、出版偏誤等現象

研究者可能對研究成果選擇性報告而發生出版偏誤，或是不將研究資料、研究計畫分享，使得研究的設計與實施細

節難以被外人所知。在發表階段，亦可能面臨發表偏差，即同儕審查傾向收錄有正面或顯著結果的研究成果。

現在的學術研究，研究過程中常被描繪成一帆風順、毫無阻礙的「完美研究」，但這與真實世界的情況大相徑庭——研究者在實際工作中往往會遇到各種挑戰和困難。因此，更重要的議題是，如何幫助研究者擺脫「完美研究」的迷思，如何協助他們克服研究過程中的實際障礙，以及如何在面對這些挑戰的同時，仍能確保研究的通透度。在推動研究再現與通透度之提升時，亦須需要建立一個更加開放、包容的研究環境，鼓勵研究者坦誠面對研究中的負面結果或是阻礙，並提供必要的支持和資源，以維護科學研究的誠信和可信度。

貳、研究案例：區塊鏈技術輔助研究通透度

為了協助研究者與學術界提升研究過程的通透度，區塊鏈（blockchain）技術因其去中心化與不可竄改的特性，近年來開始被應用於科研環境中。

區塊鏈能確保資料的真實性與完整性，並以分散式儲存提升資料的永久保存。透過區塊鏈，研究者能將研究數據、

同意書、研究計畫等重要文件加以加密並紀錄於鏈上，確保其不受竄改，上鏈的資料將具備時間戳記，並能透過智慧合約，將整個研究流程與需求條件加以程式化並自動驗證。此外，區塊鏈的公開透明特性，亦使得研究過程中的各項行為均留下不可「耍賴」的證據。

為此，本研究案例提出 BLT（Blockchain for Lifecycle Transparency）框架，一個建立於以太坊（Ethereum）區塊鏈上的去中心化平台，目的在於連結既有的開放科學工具，打造一個共通的介面與後台，自動化驗證研究歷程中各階段文件與行為間的語法關係（syntactic context）。例如驗證研究者是否先取得受試者同意書再進行臨床試驗、或是確認資料集是否如期開放等。

BLT 的生態系統涵蓋三大實體：去中心化後端、開放科學工具及使用利害關係人。其中，去中心化後端包含儲存平台與運算平台，所有的工作流程操作都在此執行並儲存對應的交易紀錄。開放科學工具可藉由 BLT 定義的應用程式介面（API）與去中心化後端互動，如 Open Science Framework、FigShare 等現有工具可透過嵌入一段 JavaScript 程式碼來與 BLT 的平台以 API 方式串接。使用利害關係人則涵蓋研究計畫主持人、共同研究人員、研究參與者、期刊編輯、出版商、獎助機構等，他們可透過支援 BLT 的開放科學工具，將其研究

相關活動記錄於區塊鏈上。

BLT 的設計具有三大特色：

一、去中心化信任（Decentralized trust）

BLT 透過區塊鏈底層技術，將研究過程中產生的文件與數據記錄於區塊鏈上。這些記錄具備不可竄改的特性，使研究者與驗證者無需依賴單一中心機構，即可直接驗證研究過程的完整性。這種去中心化的信任機制可提高研究的可信度。

二、讓研究者掌控透明度（PI-controllable transparency）

BLT 賦予研究者對研究透明度的掌控權：研究者可自行決定將哪些研究過程公開透明、何時公開、對誰公開，而非盲目追求極致透明。然而，一旦研究者決定揭露資訊，其透明化過程將在區塊鏈上留下不可篡改的證據。這種設計可以平衡研究透明度和保護知識產權的需求，同時也考慮到了研究者對於過早公開研究內容可能被他人搶先發表的擔憂。

三、自動化驗證（Automated syntax validation）

BLT 的智慧合約功能允許期刊或資助單位為研究流程設

定特定條件，例如：期刊或資助單位可透過智慧合約，針對研究流程設定特定條件（如資料需在某個時間點開放），系統能自動驗證研究者是否滿足這些要求，大幅減輕人工審核的負擔，提高效率並降低人為錯誤的可能性。

BLT 架構的設計目標是作為各開放科學工具的通用媒介。無論是現有工具還是未來開發的新工具，都能透過 BLT 定義的應用程式介面（API）將使用者的研究活動記錄於區塊鏈上。這種設計確保了研究者在選擇工具增強自己開放科學工具的自由度，同時也使這些工具能透過 BLT 相互連接，並自動驗證研究活動的語法正確性。最終，BLT 可幫助研究團隊塑造一個完整且可驗證的研究生命週期，大大提升科研活動的透明度和可信度。

參、為 BLT 打造良好的使用者體驗

一、BLT 的原型設計歷程

新興技術從開發到廣泛應用，良好的使用者體驗為不可忽視的關鍵。如何提升創新技術的易用性與資訊設計，進而增加使用者的使用意願，為本研究案例致力探討之議題。本章節將詳述 BLT 平台之設計歷程、原型成果以及在推廣 BLT

平台過程中的潛在挑戰。為了系統性地展示 BLT 平台的設計過程，BLT 之原型設計歷程如表四所示。本研究案例採用 Garrett（2011）提出之使用者經驗要素（The Elements of User Experience）作為主要介面設計流程，下文將以該框架作為敘述脈絡，說明研究者如何提升 BLT 平台的使用者體驗。表四依照結構層、骨架層和表層三個階段，歷程涵蓋了從初步的流程圖繪製，到最終的網站建置和區塊鏈錢包串接，呈現 BLT 框架從概念到實現的完整過程。

表四、原型設計歷程

	結構層	骨架層				表層
內容	繪製流程圖（Flow chart）	繪製框架圖	繪製低擬真原型	驗證原型	製作高擬真互動原型	建置網站並串接區塊鏈錢包
目標	釐清平台功能與建置頁面	建構頁面網站元素之擺放位置	基於框架圖具體化網頁元素	檢驗當前設計是否滿足使用者需求，且可帶來良好體驗	檢驗平台使用者體驗之流暢度	
使用工具	miro	Adobe XD	Figma	易用性訪談	Figma	Vue.js

（一）將技術價值具象化——製作原型

1. 策略層（Strategy）

於策略層階段，研究者或產品設計師須釐清使用者需求，並建構產品創建之目標。基於生醫工作坊之研究資料，研究團隊以親和圖（Affinity Diagram）的方式探尋學者於資料分享與資料管理面對的困難，發現學者進行資料分享與管理之需求有三，分別為：（1）希望有更好的學術產出（2）希望能最大化資料分享的效益（3）希望能更有效的管理資源（包括經費、人力等）。

2. 範圍層（Scope）

在釐清學者需求後，研究團隊於範圍層依據需求凝聚平台功能。基於上述需求，研究團隊導入區塊鏈技術，基於上述提及之設計考量，研究團隊將各價值具象化形成對應之三大功能：研究資料儀表板、使用者自主上鏈研究紀錄、研究團隊權限管理。

3. 結構層（Structure）

於結構層階段，研究團隊利用 Miro、Figma 等設計工具協助設計互動流程，包括使用者若選擇要將研究紀錄上鏈，其後續頁面轉換、系統邏輯判斷，皆於此階段設計完成。

4. 骨架層（Skeleton）

骨架層則為介面與資訊設計之階段，研究團隊先後利用 Adobe XD、Figma 作為介面設計軟體，研究者藉由一頁式網站（landing page），讓使用者能夠快速了解 BLT 平台之功能與好處，並鼓勵其註冊試用。在持續疊代三版原型後，研究團隊邀請我國北部國立大學之六位專家學者進行專家諮詢，以了解介面設計是否符合其需求，詳細之專家諮詢請見下文（二）。

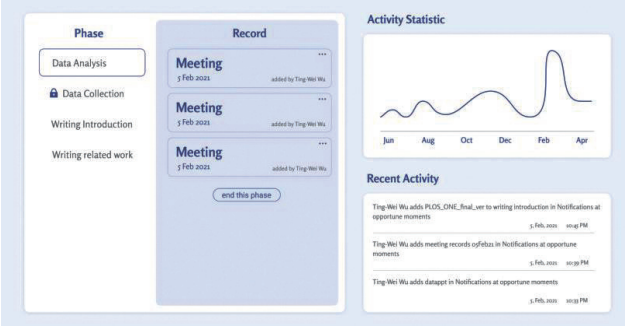
5. 表層（Surface）

介面設計的最終目標為發展成可以使用的網頁，亦為表層之目標。研究團隊於國立臺灣大學資訊工程學系網路安全實驗室之研究生合作，實作團隊繪製之介面，並實際具備將研究紀錄上鏈之功能，最終版的介面示意圖請見（三）。

（二）框架圖與低擬真原型驗證——專家諮詢訪談

進行高成本的表層網頁開發前，為了驗證骨架層的產出滿足使用者的需求，研究團隊於 2022 年 4 月邀請我國北部國立大學之六位專家學者進行專家諮詢，六位學者所屬領域為生物醫學、基礎科學、應用科學與社會科學等領域。每場專家諮詢大約 90 分鐘，共分為兩個主題進行訪談。第一部分詢問學者日常的研究資料紀錄行為，第二部分則邀請學者試用介面原型，並於任務操作後給予反饋。經由專家諮詢訪談，

研究團隊發現學者對於外接之虛擬貨幣錢包感到安全疑慮，由於該錢包之中譯說明不精準，故讓學者感覺不信任，學者亦提及平台應具備服務條款、免責聲明的資訊，方能增加平台之可信度。針對介面元素，學者對於鎖頭圖標（icon）之意義感到疑惑（圖五），設計者原希望藉由鎖頭呈現該階段已經被鎖定，無法變更的狀態，故研究團隊於後續疊代更改圖標（圖六）。



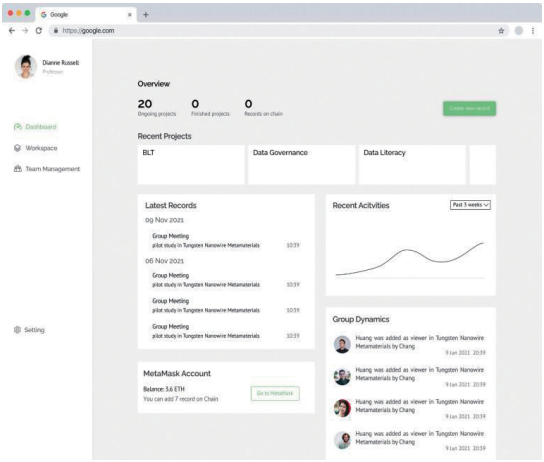
圖五、調整前介面示意

圖六、調整後介面示意

（三）設計成果

1. 研究資料儀表板

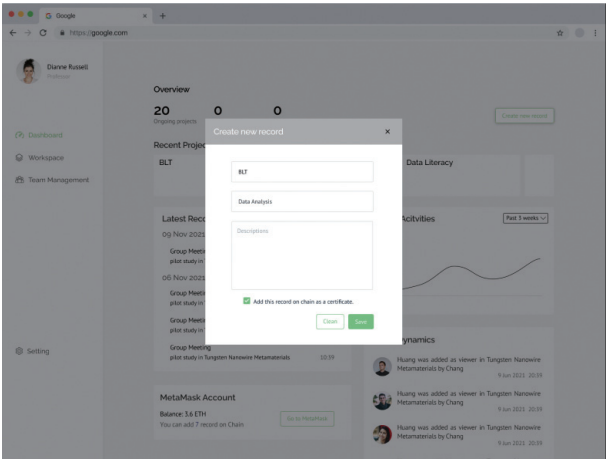
學者可利用儀表板查看最近有進行資料更動的專案，並藉由檢視最近新增的研究紀錄，快速掌握研究進度與團隊成員動向（圖七）。



圖七、BLT 研究紀錄儀表板

2. 使用者自主上鏈研究紀錄

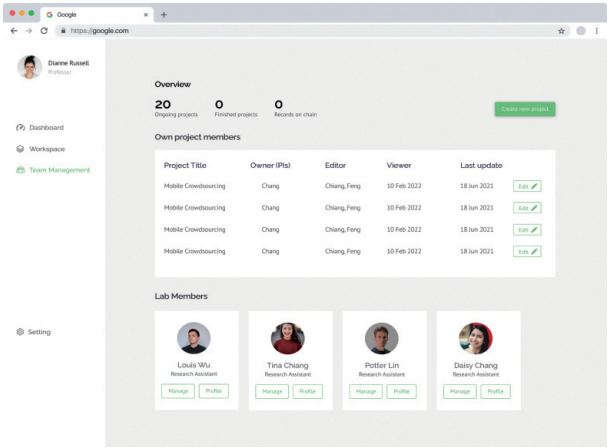
新增完紀錄後，學者可以選擇是否要將研究紀錄上傳至區塊鏈，若學者勾選「要將資料上傳至區塊鏈」，系統則會查驗其帳戶餘額是否足夠，若足夠就會將資料上傳至區塊鏈，若不足則會引導其至虛擬貨幣錢包儲值（圖八）。



圖八、BLT 平台資料上區塊鏈頁面示意

3. 研究團隊權限管理

在生醫領域工作坊與骨架層的專家訪談中，有多位學者提到研究團隊之成員權限管理為其關心的研究資料管理議題。故於 BLT 平台中，研究團隊提供了權限管理的功能，計畫主持人可以將團隊成員設定為三種權限，分別為：擁有者、編輯者、檢視者（圖九）。透過此設定，可以有效管理團隊內部研究資料安全，降低計畫主持人團隊管理負擔。



圖九、BLT 平台權限管理頁面示意

二、新興科技輔助研究通透度之設計挑戰

（一）區塊鏈技術的限制

1. 需串接第三方金流產品

區塊鏈技術在進行資料上鏈時需要消耗大量的計算資源，因此使用者（即欲上鏈者）需要支付相應的運算費用，這在區塊鏈領域中被稱為燃料費（gas fee）。BLT 使用以太鏈作為區塊鏈基礎，該鏈以虛擬貨幣以太幣（ETH）作為燃料費的交易單位，而虛擬貨幣需要有對應之第三方錢包與帳戶方能使用。對應 BLT 將研究資料記錄上鏈之功能，由於使用者於使用服務前皆須擁有第三方虛擬錢包，故該第三方虛擬錢包之服務體驗皆可能成為使用者使用 BLT 之推力或阻力。

誠如前文提及，研究團隊於專家諮詢中發現，使用者可能因為感知到不良的錢包平台安全性與使用者體驗，在尚未使用 BLT 本體服務前即放棄使用，亦可能影響其使用 BLT 將研究資料上鏈之意願。若使用者不願意將研究資料上鏈，則有違 BLT 平台欲提升研究通透度之初衷，故第三方虛擬錢包之服務易用性與使用者體驗，為區塊鏈服務於推廣時需謹慎考量並選擇之環節。

2. 帶有預防價值之創新傳播困境

BLT 平台可以幫助研究者藉由區塊鏈的技術，為自身的研究資料提供客觀、可信任的證明，此功能之重要目的為保護研究者之研究成果，並為未來可能發生的潛在風險提供保護（例如，研究被質疑資料真偽、研究是否正確操作等），亦稱為預防性創新（preventive innovation）。預防性創新之推廣困境在於創新的相對優勢（relative advantage）較難被使用者感知，導致使用者採用新技術的意願降低（Rogers, 2003），於 BLT 的情境中，BLT 具備之功能較難被使用者察覺好處與實際效益，需等潛在風險發生時方能彰顯其效果。針對此挑戰，研究團隊試圖引入訊息框架（Message Framing）之概念提升使用者之使用動機。訊息框架概念經常用於廣告行銷領域，透過將同一問題以不同邏輯方式呈現達成影響決

策之效果（Kühberger, 1998, p. 150）。研究團隊將訊息框架結合介面設計，於 BLT 首頁分別以正面與反面方式呈現使用BLT的好處，期望了解哪種訊息框架方式較能提升區塊鏈應用創新之使用者使用意願。

（二）研究通透度推廣的挑戰

1. 研究通透度與方便性需平衡

如前文所述，BLT 可供學者管理研究進度、新增研究紀錄與管理研究團隊之權限，乍看之下功能與一般研究資料管理平台無異，然而其可透過區塊鏈建立憑證、並與開放科學工具進行串接，乃為此創新之特點。因此，除了鼓勵使用者採用平台紀錄研究資料，研究團隊更鼓勵使用者基於維持或提升研究通透性，將研究資料上傳至區塊鏈留存憑證。然而，將研究資料上傳於區塊鏈可能會延伸出若干成本，如使用者需於介面執行較多步驟新增資料至區塊鏈、需花費經費支付上傳區塊鏈之燃料費，上述成本皆可能成為使用者使用平台之阻力。因應此挑戰，研究團隊可藉由減少頁面步驟、將操作流程更順暢的串連，以減緩區塊鏈上鏈較多步驟帶來的不方便感，藉以降低推力。

2. 正面激勵鼓勵研究者更多展現通透度行為

如何鼓勵學者展現研究通透度之行為，為各國政府、

學術單位努力思考的目標，除了利用 TOP Guildlines 等指引呼籲、政策強制規範外，研究團隊持續思考是否能藉由工具使用者體驗設計鼓勵學者展現研究通透度行為。具體而言，團隊期望透過正面訊息反饋方式，告知使用者其正在提升研究通透度，賦予其使命感，進而提升操作動機，如透過盔甲動畫具象化學者將研究資料上鏈之行為，代表提升研究通透度亦為保護研究成果之行為。此外，亦可以導入遊戲化設計（gamification），為學者的研究通透度行為提供激勵因素，讓研究通透度形成正向循環。Rogers（2003）提出的創新擴散理論亦提及，提升介面的可試性（triability）亦可以增加使用者的使用意願，透過讓使用者可以輕鬆地嘗試工具功能，讓使用者感知到工具的效益，如提供試用代幣，讓學者可以體驗將研究上傳至區塊鏈，進而提升其持續使用之意願。

肆、章節小結

本章為讀者介紹 BLT（Blockchain for Lifecycle Transparency）框架，目標在透過將區塊鏈技術應用於研究生命週期的各個階段，以提升個別研究者或研究團隊，進行研究過程中的通透度，最後促進科學研究的再現性。BLT 框架的核心理念是將研究工作流程的操作過程與互動紀錄儲存於區塊鏈中，

藉此實現研究過程的存證與自動化驗證，並產生不可竄改的「過程證據」。此外，BLT 框架允許研究者依據各自的需求、期刊方與政策方的要求，藉由智慧合約，自訂研究工作流程，或符合預先設定的條件。藉由這些設計，BLT 框架能有效緩解或偵測各類型的再現性威脅與學術不當行為，為研究誠信提供確鑿證據。

儘管 BLT 框架具有增強研究通透度的功能性，推動學術社群廣泛採用仍面臨若干挑戰。區塊鏈技術及其原理對大多數研究人員而言仍相對陌生。研究者需要投入相當時間來理解其運作機制和潛在效益，這可能成為初期採用的一大障礙。其次，部分研究者可能擔心過度透明會洩漏重要研究發現或影響論文發表機會，這種顧慮亦反映了學術界在開放性和保護知識產權之間的平衡難題。推動 BLT 平台的普及化有賴學術界各利害關係人的共同參與，僅僅依靠新興科技工具的研發，推廣之力度恐嫌不足。唯有技術創新與開放科學文化之形塑並行，新興科技才有機會成為支持科研永續發展的關鍵基礎設施。這意味著不僅需要持續完善BLT等技術工具，還需要學術界在開放科學之激勵機制、政策、標準等方面做出相應的調整和革新。

參考文獻

Enserink, M. (2012). Final report on Stapel also blames field as a whole. *Science, 338*,1270-1271. https://doi.org/10.1126/science.338.6112.1270

Freedman, L. P., Cockburn, I. M.,& Simcoe, T. S. (2015). The economics of reproducibility in preclinical research. *PLoS Biology, 13*(6), e1002165. https://doi.org/10.1371/journal.pbio.1002165

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science, 349*, aac4716. https://doi.org/10.1126/science.aac4716

Garrett J. J. (2011). *The elements of user experience: User-centered design for the web and beyond* (2nd ed.). New Riders.

Rogers E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Kühberger A. (1998). The Influence of Framing on Risky Decisions: A Meta-analysis. *Organizational behavior and human decision processes*, 75(1), 23-55. https://doi.org/10.1006/obhd.1998.2781

結語

本研究成果報告書彙編了 2018-2023 年，國立臺灣大學資料科學與資訊設計實驗室在開放科學議題相關的研究。本書聚焦於研究資料的生命週期和管理，向讀者呈現了開放科學在資料密集典範下所展現的多種議題。

開放科學的核心理念是促進科學知識的積累和發展。透過開放科學的實踐，研究者在研究過程中所產製的資料得以被外部團隊重複使用和驗證。這不僅體現了研究資料的永續價值，也彰顯了資料共享對於科學進步和社會發展的重要性，這不只進一步促進了科學、科技與社會的互動，亦增強科學研究成果被社會接受的可信度，有助於科學研究更好地服務與應用於各角落。

開放科學與永續科學之間亦存在密切的關聯。在這兩個領域中，科學成果再現性（reproducibility）扮演著關鍵角色。透過促進開放研究資料、增強研究通透度與再現性，開

放科學為實現「有責任的科學」鋪平了道路。這種實踐推動科學朝向永續發展：開放科學行動促進了自然與社會、學界與非學界間的對話，從而為永續社會提供科學服務。

開放科學與永續科學之間存在著緊密的關聯。展望未來，期望在全球提倡開放科學的潮流下，政府與學術社群能展開相關之獎勵機制之設計，透過持續推動開放科學實踐，使得利害相關者都覺察且認同再現性的重要性，才有更多機會能實現「有責任的科學」，確保科學研究的永續發展。

臺灣開放科學研究成果彙編：

研究者資料實踐與研究資料基礎建設

主　　編／鄭　瑋

執行編輯／丁昱寧、高庭萱、楊郁如、黃郁雯、黃寶霈

助理編輯／林語婕、陳玟寧

章節作者／石易儒、李建欣、林潔筠、張韻琦、龔俐霏

出版策畫／國立臺灣大學資料科學與資訊設計實驗室

106 臺北市羅斯福路4段1號 臺大圖書資訊學系

電話：+886-2-3366-2954

信箱：wjeng@ntu.edu.tw

製　　作／秀威資訊科技股份有限公司

114 台北市內湖區瑞光路 76 巷 69 號 1 樓

電話：+886-2-2796-3638

信箱：publish@showwe.com.tw

出版日期／2024 年 7 月

定　　價／350 元

版權所有・翻印必究　All Rights Reserved

Printed in Taiwan