

國立臺灣大學文學院圖書資訊學系

碩士論文

Department of Library and Information Science

College of Liberal Arts

National Taiwan University

Master Thesis

從大學教師的觀點探討學術圖書館發展之資料度用服務

**A Research on How to Develop Data Curation Service
in Academic Library from Professors' Viewpoint**



秦韻涵

Yun-Han Chin

指導教授：陳雪華 博士

Advisor: Hsueh-Hua Chen, Ph.D.

中華民國 101 年 7 月

July 2012

誌謝辭

三年的時間說來也不短，其中真的有很多辛苦的部份，能夠順利走到畢業這一步，我最感謝的是天父上帝，總是垂聽我的禱告，也給我夠用的恩典，讓我能渡過許多難關，很多突發狀況也都有驚無險的解決，真的覺得今天能在這邊寫謝辭，是因著主的保守和看顧，讓我能走完這段研究所的日子。

非常感謝我的指導教授陳雪華老師，總在百忙之中給予我很多建議和教導，讓我不但能完成論文，也在當中學到許多寶貴的經驗；更感謝老師對我的包容，身為一個從理學院來唸碩士班的學生，當中有很多不習慣和不適應，沮喪和想放棄的時候也很多，謝謝老師的寬容與鼓勵，讓我順利畢業。

感謝我的父母，這三年來在各種不同方面給予我很多幫助。我的母親樂意聽我訴苦，幫我禱告，也時常鼓勵我，真的是我完成學業最大的動力；而我的父親在各種資源上幫助我，讓我唸完這三年的研究所，從不抱怨我從職場回到學校的決定，真的很感謝他們對我的幫助，希望我也能親口表達感謝之意。最後論文的衝刺階段，時常有來自弟弟的關心，讓我很感動，希望他明年也能順利畢業。

謝謝跟我同甘共苦的同學林玉，和給我許多幫助的雅文，雖然雅文已去國外唸書，但仍感謝她在我剛進研究所的階段給予我許多幫助，因為我真的很不習慣文學院在許多教學或作業上的模式，因此常麻煩她許多事。而在論文期間，常常和林玉一起討論，一起進行研究，而讓論文能有所進展，是辛苦但也美好的回憶，很高興今年跟林玉一起順利畢業。

最後想感謝一位對我而言十分特別的朋友，在我心情低落時陪伴我，想偷懶時督促我，不開心時安慰我，是我不可或缺的好夥伴；不只是畢業，許許多多其他的順境逆境，都有他的鼓勵與陪伴。為此我更加感謝主，讓我生命中遇到這麼多願意幫助我、願意為我代禱的人們，包括高雄榮耀教會和基督之家的小組員們，還有許多幫助過我的親戚朋友，願上帝祝福你們！

摘要

跨領域以及跨國研究的趨勢，使得研究資料的分享和保存成為學術環境中的新興議題，也是資料度用發展的依據；資料度用能促進研究資料的保存與組織，並提供檢索和取用，讓更多需要的使用者能夠持續的使用這些資料，以發揮其最大價值，減少資源的浪費。資料密集是目前學術研究的趨勢之一，因此研究過程中產生的大量資料如何被有效的保存、管理和分享是許多學術機構十分重視的問題；而研究者身為資料的主要生產者與使用者，其對資料度用的看法和需求，將是成功推動資料度用服務的關鍵。

本研究以臺灣大學的教師為對象，採用質性研究的訪談法，自 2011 年 8 月至 2012 年 5 月，共訪談了 12 位臺大教師，瞭解他們對於資料度用的看法，以及對於發展資料度用服務的需求，希望能夠作為臺大圖書館未來發展資料度用服務時的初步參考。

訪談結果的分析主要從資料來源與保存現況、資料之分享、資料之公開與遭遇的障礙以及對資料度用服務的期待與需求等 4 個層面進行討論和歸納，並根據研究結果，針對臺大圖書館、臺灣大學以及政府單位等 3 個層級提出推動及發展資料度用服務的建議。

本研究之結果為以下幾點：一、各國在資料度用之發展程度以英國和美國較具代表性，可作為國內學術圖書館發展資料度用服務之參考；二、研究資料之來源多樣化，整理方式不一；三、研究資料之分享以不影響個人研究為優先考量；四、對研究資料公開之態度較不積極；五、期待資料度用服務之發展能有實質回饋，並對研究有所幫助。

關鍵詞：研究資料、資料度用、使用者研究

Abstract

Sharing and preservation of research data attracts lots of attention in the academic environment because of the trend of interdisciplinary and transnational research. Data Curation provide a good solution to data sharing and preservation. Data Curation, which involves maintaining, preserving and adding value to research data, can help researchers to access and reuse the data as well as to maximize its value and reduce resource waste. How to preserve and manage large amounts of data effectively is a very important issue for the academic institutions because of the trends of data-intensive research. The researchers who produce and use the research data will be the main target of data curation service, so what researchers want and need is very important to the academic institutions which try to develop the data curation service.

This thesis investigates the professors' attitudes towards the concept of data curation and what they want or need for the data curation service . "Interviewing" is applied from the qualitative perspective and 12 of the professors of National Taiwan University were interviewed from the August, 2011, through the May, 2005. Following conclusions are derived by analyzing the interviewing results.

1. The development of data curation in United Kingdom and the United States can be the best practice for academic libraries in Taiwan which want to develop the data curation service
2. High diversity in the resource of research data.
3. The sharing of research data can not affect the current research.
4. Researchers' attitude toward data publication is negative.
5. Data curation service should provide feedback to their research.

Keyword: Data curation, research data, user study

目次

第一章 緒論.....	1
第一節 研究背景與動機.....	1
第二節 研究目的.....	4
第三節 研究方法.....	4
第四節 研究範圍與限制.....	5
第五節 名詞解釋.....	5
第二章 文獻探討.....	7
第一節 資料度用之意涵與重要性.....	7
第二節 歐洲地區之資料度用發展狀況.....	13
第三節 美加地區之資料度用發展狀況.....	22
第四節 使用者調查相關研究.....	31
第三章 研究設計與實施.....	49
第一節 研究方法.....	49
第二節 研究對象.....	50
第三節 研究流程.....	51
第四章 研究結果分析.....	55
第一節 受訪者背景與資料類型.....	55
第二節 研究資料之來源與保存現況.....	59
第三節 研究資料之分享.....	63
第四節 研究資料之公開與所遭遇的障礙.....	67
第五節 對臺大圖書館發展資料度用服務之期待與需求.....	74
第六節 綜合討論.....	79
第五章 結論與建議.....	86
第一節 結論.....	87
第二節 建議.....	94
第三節 未來研究建議.....	97
參考文獻.....	99
附錄一.....	104
附錄二.....	106

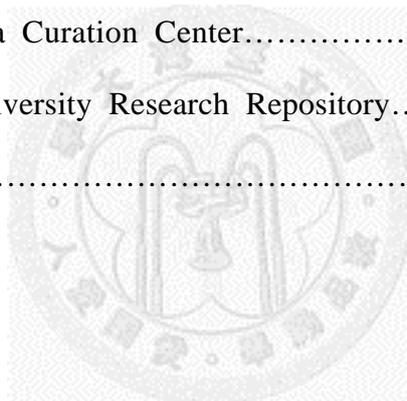
表目次

表 2-1	學術出版和研究資料之差異.....	10
表 2-2	國外使用者調查之研究結果分析.....	47
表 3-1	受訪者代號以及所屬科系和訪談時間.....	50
表 4-1	受訪者之科系和研究領域.....	56
表 4-2	受訪者資料類型分析.....	58
表 4-3	本研究與國外使用者調查研究之分析比較.....	85



圖目次

圖 2-1	Data Library and consultancy.....	18
圖 2-2	Edinburgh DataShare.....	19
圖 2-3	Research data management guidance.....	20
圖 2-4	EDINA.....	21
圖 2-5	DATA.GOV.....	23
圖 2-6	Research Data Management Service Group.....	27
圖 2-7	DataStaR.....	28
圖 2-8	Distributed Data Curation Center.....	30
圖 2-9	The Purdue University Research Repository.....	31
圖 3-1	研究實施步驟.....	54



第一章 緒論

現代學術研究的重要特徵包括跨領域合作和密集、大量的使用並產出資料，在這樣的新研究環境下，研究者們面對的不再單只是資料蒐集的問題，如何處理研究過程中的大量資料也開始受到重視，研究資料的保存、管理以及分享成為學術環境中不可不面對的新挑戰。

第一節 研究背景與動機

因著科技進步和研究上的需求，越來越多學術研究是以跨領域或遠距離合作的方式來共同解決新的科學問題，發展新的研究方式；而在這樣的合作模式中，大量的研究資料要如何被有效的組織並分享給研究團體中的其他人，成為研究過程中不可避免的環節之一。e-Research 之概念就是在這股新興的學術研究潮流下應運而生，其特徵是跨領域合作、大量且密集的資料使用和產出，重視資料的保存與便利的分享，以促進學術研究的發展；Jankowski(2009)指出 e-Research 是基於網際網路環境的學術研究，研究者以網路為基礎，並運用各種相關工具，打破國界和學科領域的限制，實現跨國或跨領域的合作研究模式。e-Research 的興起帶來研究思維上的突破，不同國家或領域之間密切合作的結果，產生了大量的研究資料；對研究團隊而言，安全的保存和有效的分享資料十分重要，而對其他的研究者而言，適度的開放使用這些資料，將使其更具價值，並有助於學術研究的進步。

除了研究型態的轉變之外，另一個推動研究資料公開的因素在於認為由國家補助而產生的研究，其成果也應為公眾所共享，亦包含在過程中產生的研究資料；基於這個想法，經濟合作開發組織 (The Organization For Economic Cooperation

And Development, 簡稱 OECD)於 2007 年提出了「公共贊助研究資料之近用原則與基準(Principles and Guidelines for Access to Research Data from Public Funding)」,其目的便在於促進研究資料的近用,尤其是由國家所贊助的計畫中所產出之資料,以期達到強化科學之交流與進步、刺激新研究的產生、發掘更多尚未被探討的研究主題等目的(OECD, 2007)。

研究資料的公開確實會對學術研究有所幫助,因此如何公開並讓其能被需要的人取用就成為必須面對和解決的難題。資料度用一詞翻譯自「data curation」,其意義為對學術研究或教育上有價值的資料進行主動且持續的管理,包括對資料的評鑑、選擇與組織等,以確保在其生命週期中能有效的被取得和使用(Shreeves & Cragin, 2008)。由於資料度用之對象是以數位化的資料為主,因此有一同義詞「digital curation」,英國數位度用中心(Digital Curation Centre, 簡稱 DCC)將其定義為「針對數位化之研究資料在其整個生命週期中進行維護、管理、保存和加值之行為。」(DCC, 2007)。資料度用包含了針對研究資料進行管理、保存及再利用之概念,強調的是在資料受到完善的保存與組織後,能夠提供檢索和取用,讓更多需要的使用者能夠持續的使用這些資料,以發揮其最大價值,並減少資源的浪費。資料度用的發展將能為 e-Research 下產生大量的研究資料找到合適的解決之道,是近年來相當受重視的議題。

資料度用的發展有許多層級,例如英國的數位度用中心(DCC)屬於國家級的機構,希望站在國家的高度,透過各種合作和研究,發展完善的策略,並推動資料度用的發展。若從學科領域的層級來看,有些學科在資料公開方面早已是慣例,因此發展了專門存放研究資料的資料庫,希望提供一個長久保存研究資料的解決方案,例如美國醫學圖書館(U.S. National Library of Medicine)所建置的 National Center for Biotechnology Information (簡稱 NCBI),就是生物科技領域中最具公信力和代表性的資料庫,其中有完整的基因和蛋白質結構等資料,是世界各國的研究者們主動上傳並註冊的,已成為生物科技及相關領域之慣例。

對於其他並無學科典藏庫的領域來說，機構層級的典藏庫是一個解決方案，故國外漸漸有學術圖書館將研究資料納入典藏範圍，例如英國愛丁堡大學的 Edinburgh DataShare 以及普渡大學的 Purdue University Research Repository，都是學術圖書館參與其中並推出資料度用相關服務的實例。Choudhury(2009)亦指出圖書館具有發展資料度用的潛力，透過將研究資料納入典藏範圍，擴充學術圖書館的館藏，館員則可扮演資料科學人員(data scientist)的角色，讓圖書館成為提供專業服務的資料中心，成為整個學術研究中重要的一環。

資料度用的觀念在國內並不普及，目前也尚未有學術圖書館提供資料度用服務之實例。2011年5月由臺灣大學(簡稱臺大)與教育部北二區區域教學資源中心圖書館所主辦的「e-Research：新世代學術研究之利器」研討會中，邀請伊利諾大學與賓州州立大學專家學者出席，並介紹國外學術圖書館發展資料度用服務之概況，應是國內首度對此議題有公開且共同的討論。臺灣大學圖書館(簡稱臺大圖書館)更以此為契機，成立了 e-Research 小組，成員包括圖書館館長與兩位副館長，學科館員亦參與其中，定期聚會並討論資料度用相關議題，瞭解國外學術圖書館在資料度用服務方面的發展，以評估在校內落實的可能性。

臺大圖書館在推動機構層級的資料度用時，首先會面對的就是同時身為資料生產者和使用者的大學教師們，他們的參與意願、對資料度用的看法和需求，都是能否順利推動資料度用服務的關鍵。國外大學在發展資料度用服務時，會先對校內的研究者們進行調查或訪談，瞭解他們的需求以及可能遭遇的障礙，作為推動資料度用的參考。因此本研究也希望透過對臺大教師的訪談，深入瞭解他們對於資料度用所代表的資料分享和公開等概念，抱持什麼樣的態度和看法，調查他們對於資料度用服務的期待和需求，以及目前學術環境中，是否存在不利於資料度用推廣的障礙，以期能夠作為臺大圖書館未來推動資料度用服務的參考；唯有協助教師們解決現存的障礙並提供符合他們需求的服務，才能讓學術圖書館在新的研究環境中成為助力，並扮演學術研究中不可或缺的角色。

第二節 研究目的

本研究旨在瞭解大學教師對於學術圖書館發展資料度用的看法和需求，由於考量到各機構的差異性與獨特性，故以臺大之教師為研究對象，探討學術圖書館資料度用之發展方向，並希望能提供臺大圖書館作為發展資料度用服務的參考。本研究的目的如下：

- (一) 探討國外大學之資料度用發展現況，並分析其資料度用相關之使用者研究及成果。
- (二) 調查臺大教師目前研究資料之來源及其保存
- (三) 訪查臺大教師分享研究資料之狀況。
- (四) 訪查臺大教師對於研究資料公開之態度與可能遭遇的障礙。
- (五) 分析臺大教師對於資料度用服務之期望與需求。
- (六) 提供研究成果作為學術圖書館、研究型大學及政府單位未來發展資料度用服務之參考。

第三節 研究方法

資料度用在國內屬於新興議題，大多教師對其認識不深，因此本研究採用訪談法，以期在對談的過程中加以解釋資料度用之意義，並瞭解臺大教師們之看法與需求。訪談對象之挑選採取方便取樣，以期能夠找到對此議題較有興趣之教師進行訪談；本研究使用半結構式的訪談，再根據受訪者的回答，轉成文字描述之訪談文本，並加以分析、詮釋。

為達研究目的，且因資料度用在國內幾乎尚未被討論，本研究首先蒐集國外相關文獻，以瞭解資料度用之意涵、重要性及發展現況，從中找出發展的較為完善的大學進行深入介紹；並彙整國外許多使用者研究，據以設計訪談大綱，進行訪談調查，最後將訪談結果加以整理歸納，提出結論與建議。

第四節 研究範圍與限制

本研究以臺大內從事研究的教師們為對象，希望能深入瞭解同時身為資料生產者和使用者的教師們，對於臺大圖書館發展資料度用服務有何看法，調查他們對於資料度用服務的期待和需求；以及目前學術環境中，是否存在不利於資料度用推廣的障礙，希望能做為未來臺大圖書館發展資料時的初步參考。考慮到臺大教師人數眾多，無法在本研究中涵蓋所有不同之學術領域，故訪談對象之挑選採用方便取樣，從有意願且容易接觸的教師開始訪談，再請受訪者推薦他們認為對此議題有興趣或有需求之教師，進行下一步的訪談。

由於資料度用屬於新興議題，國內文獻極為稀少，故本研究將以國外之文獻為主要分析之對象，且受到語言的限制，參考之文獻以英文為主；考慮到各國發展資料度用的進展不一，希望挑選目前發展較為健全，且有大學層級之資料度用發展的國家進行深入介紹，因此在各國之資料度用發展狀況方面，將以英國和美國為主，並以具代表性的大學為例進行探討。

第五節 名詞解釋

一、e-Research

e-Research 的特徵是跨領域合作、大量且密集的資料使用和產出，重視資料的保存與便利的分享，以促進學術研究的發展；Jankowski (2009)指出 e-Research 是基於網際網路環境的學術研究，研究者以網路為基礎，並運用各種相關工具，打破國界和學科領域的限制，實現跨國或跨領域的合作研究模式。

e-Research 和 e-Science、cyberinfrastructure 等詞意義相近，故時常被混用，e-Science 主要應用於歐洲，美國則較常使用 Cyberinfrastructure；以代表的意涵而言，e-Science 一般較為強調自然科學領域，而 Cyberinfrastructure 則是強調資

訊基礎建設對學術研究的幫助和影響。相較之下，e-Research 代表的是領域的擴張，除了自然科學，也涵蓋了社會科學、人文科學等領域，更具整體觀點和代表性(Harvey, 2010)。

二、資料度用

資料度用一詞翻譯自「data curation」，其意義為對學術研究或教育上有價值的資料進行主動且持續的管理，包括對資料的評鑑、選擇與組織等，以確保在其生命週期中能有效的被取得和使用(Shreeves & Cragin, 2008)。由於資料度用之對象是以數位化的資料為主，因此有一同義詞「digital curation」，英國數位度用中心(Digital Curation Centre，簡稱 DCC)將其定義為「針對數位化之研究資料在其整個生命週期中進行維護、管理、保存和加值之行為。」(DCC, 2007)。

資料度用包含了針對研究資料進行管理、保存及再利用之概念，強調的是在資料受到完善的保存與組織後，能夠提供檢索和取用，讓更多需要的使用者能夠持續的使用這些資料，以發揮其最大價值，並減少資源的浪費。

第二章 文獻探討

第一節 資料度用之意涵與重要性

一、資料度用的意涵

「資料度用」一詞翻譯自「data curation」，Shreeves and Cragin (2008)對其之定義為「對於在學術研究或教育上有價值的資料進行主動且持續的管理，包括對資料的評鑑、選擇與組織等，以確保在其生命週期中能有效的被取得和使用。」；而 Rusbridge et al.(2005)進一步的闡釋了「curation」一詞之意涵，認為其範圍不只是對於資料當下的保存和再利用，而是要對資料進行加值，使其透過適當的公開方式而更容易的被取得和使用。英國資訊系統聯合委員會 (Joint Information Systems Committee, 簡稱 JISC)的報告中，認為「curation」和過去圖書資訊學界使用的「archiving」及「preservation」等詞彙都有所不同，其包含的層面更廣，除了長久的保存之外，更要維持保存對象之完整性，並促進使用者對其之取用。以資料為對象時，則從資料被產生時就持續對其進行保存及管理，讓其能有效的被檢索、取用和再利用，也要注意資料的不斷更新，以確保其符合需求 (Lord, Macdonald, Lyon, & Giaretta, 2004)。

資料度用的概念在英國十分受到重視，並於 2004 年成立了數位度用中心 (DCC)，希望發展更完整的資料度用策略，以促進國內學術研究之進步。數位度用中心對於資料度用是採用「digital curation」一詞，其定義為「針對數位化之研究資料在其整個生命週期中進行維護、管理、保存和加值之行為。」(DCC, 2007)。Beagrie 認為在資料度用之議題上，很多詞彙都在持續的進化，或是在不同領域中有各自的常用詞彙，容易在溝通上造成混淆，而 digital curation 可用以代表科學界中常用的 data curation 和圖書資訊界使用的 digital preservation 等概念，而這

些詞彙多少都已留下一些刻板印象，例如不同領域對「data」的範圍界定有所不同；而大部份研究者會認為 digital preservation 是研究結束後的事情，是圖書館或出版者的責任，與他們無關。故選擇 digital curation 一詞除了能夠做為整合這些類似概念的橋樑，也能避免一些對於既有詞彙的刻板印象 (Beagrie, 2006)。總而言之，data curation 和 digital curation 所代表的概念基本上是相同的，只是在詞彙的選擇上會因領域的不同而有所差異，故在本文中一律將其譯為「資料度用」，避免更多因為翻譯上的不同而造成之混淆。資料度用包含了管理、保存及再利用之概念，強調的是在資料受到完善的保存與組織後，能夠提供檢索和取用，讓更多需要的使用者能夠持續的使用這些資料，以發揮其最大價值，並減少資源的浪費。

資料度用中的資料指的就是研究資料，不同領域中對於研究資料的範圍會有不同之界定，故以下針對較廣泛之研究資料定義進行闡述。依據 OECD 於 2007 年在其「公共贊助研究資料之近用原則與基準」中對研究資料的定義，研究資料指的是「作為科學研究主要來源之事實性紀錄，包括數據、文字紀錄、圖片和聲音等，並且這些紀錄能夠被用來驗證研究成果而受到學術社群所認可」，研究資料不包括用來做研究或分析的電腦設備、研究報告的草稿、未來的研究規劃、研究相關人員之間的對話記錄和一些實體物件，像是實驗用的老鼠、細菌、樣本等。(OECD, 2007)

根據美國國家科學基金理事會(National Science Board，簡稱 NSB)的分類，研究資料依其產生的方式可分為以下四大類：(NSB, 2005)

1. 觀察型 (observational)：包括測量值和調查報告等，大多具有特定時間、地點的特性，屬於具有是歷史性質且不可能重新蒐集的資料。
2. 電腦運算型 (computational)：來自於電腦模式或模擬之成果，要重製這些資料需要有相同的硬體、軟體等條件，很多時候只有模擬結果被保存。

3. 實驗型 (experimental)：大多來自於實驗進行中所紀錄下來的一切資料，雖然可以重新進行相同實驗，卻常有時間、人力和經費上的考量，因此仍具保存價值。

4. 紀錄型 (records)：指關於政府、商業活動、社會現象等相關紀錄，可應用於科學、社會科學、歷史研究等各方面。

英國的研究資訊網路(Research Information Network, 簡稱 RIN)認為研究資料的分類方法應不只一種，根據其產生之方式，研究資料可來自於科學實驗 (scientific experiments)、電腦模式或模擬(models or simulations)以及觀察 (observations)等三種不同的過程，若是根據資料被蒐集的原因不同，則可分為下列三類：

1. 為了特定研究計畫所蒐集的資料，在該計畫之外的情境下，這些資料可能不具太大的價值。
2. 為了特定學術團體、學科領域或跨領域所蒐集並用以輔助研究之資料。
3. 基本且能夠提供所有研究者都加以使用、參考的資料，像是化學結構、基因序列或是文字紀錄等。

而不同研究階段中亦會產生不同的研究資料，大致分為研究初始階段產生的「原始資料(raw data)」，通常來自觀察或實驗；另外則是將原始資料進行處理、分析或合併後所產生的「衍生資料(derived data)」。(RIN, 2008)

Kuula (2008)認為學術出版和研究資料是有所不同的，並對其進行比較，其對研究資料的描述較為完整，包括它可能並未被發表於研究成果中、牽涉到更多隱私或保密的問題、著作權和擁有權的界線較不明確、可能需要授權或特定的軟體才能使用等等；兩者的差別如表 2-1 所列。(Kuula, 2008)

表 2-1 學術出版和研究資料之差異

學術出版(publication)	研究資料(data)
代表了研究成果之資訊	不一定發表於研究成果中之資訊
使用者需用基本的軟體開啟	使用者可能需要特定軟體才能開啟
自行理解	需要更多相關資訊來幫助理解
不需要特殊的許可才能使用	使用前可能需要徵求同意
擁有權和著作權明確	擁有權和著作權常常不明確
學術社群可免費取用或是透過付費的方式取得	有不同程度的開放情況(使用權限上的限制不同)
被視為學術產出	目前尚未被視為正式的學術產出
已可被使用	需要額外的處理

資料來源：根據 Kuula, A., & Borg, S. (2008). *Open access to and reuse of research data: the state of the art in Finland*. Finland: Finnish Social Science Data Archive (FSD) 整理

綜上所述，研究資料極具多樣性，目前國際間尚無單一而完整之定義。本研究依據資料度用的意義，界定研究資料為研究過程中所使用和產出的資料，不包含硬體設備和實驗樣本，且以數位化之資料為主。

二、資料度用的重要性

學術研究會產出許多的研究資料，尤其在電腦與網路的普及後，許多未整理、公開的研究資料常被儲存在各個研究室、實驗室的硬碟中；Witt(2008)認為一個研究從資料的產生開始，隨著整理、組織和分析等過程到研究結果的產出為止，資料的量是不斷被濃縮和刪減的，尤其是到了研究結果發表的階段，通常是透過期刊文章或論文，所呈現出來資訊是最少的狀態，稱為「資訊瓶頸(information bottleneck)」；而這也是傳統上圖書館會參與的階段，透過保存、流通的方式來使這些學術成果被散佈和傳播。事實上，一篇學術文章中所包含的資訊量通常並不

足以提供他人重覆該研究，也幾乎不可能對該研究進行再次的驗證；對專家學者來說，他們只能用推測的方式去評估其原始產生的資料是否有其他潛力或用途，而對一般的使用者來說，更是無法進一步的了解該文章背後之資料的價值或其他應用的可能性。雖然在資訊瓶頸之外的資料可能會透過一些非正式的管道在學術團體中交流和分享，但和已發表的學術文章的能見度相比，這些資料被他人取用的機會顯得相形見绌。除了驗證研究成果的價值之外，資料的保存和分享也能減少資源上的浪費，避免在進行相關的研究時需要重新產出相同的資料；這些資料甚至能夠被不同領域的人所使用，進而促成不同學術領域的交流，深化學術研究的發展。總而言之，只要有新的使用者能夠取得並了解這些研究資料，它們就有可能在意想不到之處不斷創造新的價值。

Heidorn(2008)用「黑暗資料(dark data)」來形容那些未被適當的整理和保存的研究資料，這些未公開和分享的資料對潛在的使用者和學術社群來說近乎是隱形的，造成再使用上的困難甚至因此而遺失。Heidorn(2008)認為大型的研究計畫由於參與者眾，研究經費高且容易受到各方的關注，因此在研究資料的保存、組織和管理等各方面都設想的較為周全，有意者也較容易知道如何取得這些研究資料，並進行再利用；相對的，小型的研究計畫常因經費和人力的關係，在研究過程中就不曾好好的整理、組織相關的研究資料，研究完成之後，也不見得願意花費時間思考如何保存這些資料，或是如何與他人分享，造成資料的再使用率低，無法發揮其最大價值；而在現實的情境中，這些小型的研究計畫其實占了大多數，更讓資料流失的問題需要受到重視。

以 2007 年美國國家科學基金會(National Science Foundation, 簡稱 NSF)所贊助的研究計畫為例，總研究經費的前 20% 只有 254 個研究計畫，而其餘的 80% 之研究經費則贊助了 11,771 個研究計畫 (NSF, 2007)。Heidorn(2008)指出，這 254 個研究計畫應是大型或知名度較高的計畫，經費和人力都較為充足，而且可能會產出大量的資料需要儲存並管理，所以可以假設這些計畫有將其資料進行良好的

保存，其資料被取用的機會也較高。但相對的，剩下 11,771 個計畫應屬於小型計畫或是個人申請的計畫，很可能因為經費不足或人力問題，使這些計畫中的資料未被妥善的保存或管理，造成資料對整個研究環境而言是隱藏甚至流失的。這 11,771 個計畫大約占了總計畫數的 98%，即使個別的計畫中所產出的資料量不多，但加總之後相當可觀，這表示流失的資料量也不可忽視；這些資料就如同經濟學上長尾理論(long tail theory)的「尾巴」，熱門的資料雖然受到大部份人的關注，但這些位於尾巴的資料同樣會有其使用者，只要提高它們的能見度，就能讓更多潛在的使用者注意到並且再利用，因此仍有管理並保存這些研究資料的價值。Witt(2008)則指出所謂的 small science，也就是傳統上以個人的研究假設為出發點所進行的研究，成果通常來自對自己所產出的資料的分析，故資料量少而未受重視；但隨著科技的進步和研究型態的轉變，small science 也開始會有合作研究的情況發生，因此亦有資料度用的需求，而由於過去的資料未受重視，較缺乏發展資料度用的基礎，故需特別關注。

Goetz(2007)認為還有一種黑暗資料是來自於「失敗的實驗」，這並非是說研究本身的失敗，而是往往在成功之前有很多的嘗試，也產生了許多無法呼應研究假設的資料，這些資料並不會被發表在研究成果中，因此也不為人所知，故成為黑暗資料的另一來源。科學與科技的進步讓研究資料的數量快速增加，若不重視對這些資料的保存、管理和再利用，只會讓更多具有潛在價值的資料消失於黑暗之中，對科學的發展造成無法彌補的損失。Lord and Macdonald(2003)指出並非所有資料都有長久保存之價值，且其價值會因著時空和學科領域的不同而有所差異，因此數位資源的篩選、保存和度用之重要性會在越來越多的領域受到關注。美國國家科學基金理事會則認為數位研究資料本身就提供了科學研究上新的可能性，透過對大量資料的分析、整合，可能會帶來科學研究上嶄新的領悟，其潛在之效益值得更多的研究 (NSB, 2005)。研究資料是重要且極具價值的科學產出，如何透過資料度用讓其在未來發揮更大的貢獻，是必須積極面對的重要議題。

美國研究圖書館學會(ARL)於 2011 年於發表的報告「New Roles for New Times: Digital Curation for Preservation」中，認為在研究型態轉變與網路基礎建設快速發展的現況下，學術圖書館應以研究中不可或缺的角色為目標，積極轉型並發展出資料度用相關服務，培養具備相關知識的館員和專業的資料科學人員，進行更全方位的數位內容典藏，包括所處機構中產生的研究資料或其他一手研究資源等；並以過去數位內容出版的相關經驗，將這些數位的資料進行整理及保存，以期在未來能進行資料之出版或公開，促進研究的進步。

資料度用對許多國家而言都是新興議題，雖然受到關注和討論，但能將之付諸實行的國家仍在少數。為了呈現世界各國在資料度用方面的進展，並剖析其在大學層級之應用實例，以下將分析歐洲地區和美加地區之發展概況，並舉出資料度用發展較完整之大學為例，進行深入之介紹。至於亞洲地區的國家，如澳洲和日本，雖然有相關的國家級的典藏庫、學術研討會以及使用者研究，但因發展較晚，故缺乏大學層級之案例，其他各國則少有相關議題之討論，因此不再另外介紹其發展狀況，但仍會將使用者研究之部份進行整理和探討。

第二節 歐洲地區之資料度用發展狀況

一、歐洲整體之資料度用發展概況

歐盟委員會(European Commission，簡稱 EC)在 2010 年發表了一份報告「Riding the wave」，描述了歐洲地區面對研究資料快速增加、研究型態改變等狀況所採取的行動，並對發展過程中的重要事件進行回顧；最後則預測了 2030 年時會發生的情況，並點出了目前環境中阻礙資料度用發展的一些障礙，進而提出因應之對策。文中指出為了無縫的取得、使用、再利用資料，並兼顧資料品質和可信度，資料本身應被視為基礎建設的一環，而整合大量資料所帶來的衝擊和影響會大大的改變學術研究進行的方式，同時亦有助於社會、經濟、科技等各方面

的發展。(EC, 2010)

「Riding the wave」文中對 2030 年提出了很多期待，希望資料的產生者、使用者或是保管者都能充分瞭解資料分享之重要性；而各領域的研究者不但能夠取得需要的資料，也有能力評估並瞭解這些研究資料的意義，同時資料的提供者也能因此而受惠，進而讓更多人願意分享資料並將之儲存於可信賴的典藏庫中。對於政府提供資金的研究，則應將資料的再利用視為重要的一環，以發揮其最大價值，政府部門的資料也應更透明化，讓所有人、團體、企業因著能平等的取得這些研究資料，而進行自由的運用，創造更多價值或激發新的研究契機。為了達成這些願景，文中建議歐盟國家應立刻採取行動，開始國際之間的合作，在整個以資料為目標的基礎建設上達到共識，讓資料的儲存、檢索和使用等功能能夠有品質保障且具彈性，因此也必須要規劃足夠的預算來支持這些建設。在資料的品質和價值上，也應制定標準來幫助確保資料之品質，並評估其影響力。在推廣方面，需要訓練出一批新的資料專家，因此應有相關的課程放入學校教育體制中，同時亦能推廣相關之概念，讓大眾也能夠了解資料度用之重要性。文中最後強調沒有任何單一國家或機構能夠獨自面對資料度用這個新挑戰，因此國際間的定期討論是很重要的，應該要舉辦論壇或研討會，促進各國更加重視此一議題；同時並兼顧綠色科技的發展，以及相關政策的制定，讓未來科學研究的發展能夠有勞固的根基。(EC, 2010)

EC 的這篇報告顯示出整個歐盟國家已經開始注意到大量產生的研究資料所帶來的影響，若不盡早面對和處理，則可能演變成巨大的損失；相較之下，英國在更早之前就已經注意到這個問題，也在國內進行了許多討論和研究，是目前在資料度用方面文獻最多的國家。由於本研究的焦點在於大學教師對資料度用之看法，因此以下就英國在整個資料度用的發展過程中，大學如何參與其中以及發展之狀況進行介紹。

二、英國的大學之資料度用發展

英國很早就對資料度用的議題十分關注，其發展資料度用的過程大致上是由國家層級的機構開始發動，而在過程中和各大學的互動、合作都十分密切，因此要瞭解英國如何推動資料度用，就必需從國家型計畫或機構的角度探討大學在其中所扮演的角色。

英國於 2004 年 3 月 1 日成立之數位度用中心(DCC)由聯合資訊系統委員會 (Joint Information Systems Committee, 簡稱 JISC) 所贊助，是一獨立運作的研究型機構，希望藉由國家級機構的成立來面對資料度用所帶來的挑戰，而這個挑戰絕非單一機構或單一領域能獨自面對的，因此其參與之成員眾多，包括科技促進委員會 (Science and Technology Facilities Council, 簡稱 STFC)、英國巴斯大學(University of Bath)所主持的英國圖書資訊網絡辦公室 (United Kingdom Office for Library and Information Networking, 簡稱 UKOLN) 和英國愛丁堡大學、英國格拉斯哥大學、英國牛津大學等機構內的圖書館、計算機中心及研發單位。DCC 希望以國家級單位的角度來全面的推廣資料度用之概念，宣傳相關資訊以喚起各界對此議題之重視，並且開發各種工具、服務及制定標準，提供實務方面的資訊，讓更多高品質的資料有機會在未來創造新的價值 (DCC, 2011)。

DCC 從成立之初就有非常明確的階段性任務，第一階段是從 2004 年 3 月到 2007 年 2 月，組成一個專業團隊來探討並檢視英國高等教育之環境中所存在關於數位保存及資料度用方面的問題，團隊成員包括資料方面的專家、圖書館員、檔案方面的專家以及資料的創造者，也就是研究者們；在這個階段當中 DCC 也向外擴展其夥伴關係，納入高等教育以外的機構或是國際級機構為其合作對象。第二階段從 2007 年 3 月開始到 2010 年 2 月，團隊將其研究重心慢慢往學術社群移動，目的在於希望更多的瞭解研究者們在研究資料的整個生命週期中做了些什麼，需要的又是什麼，因此針對許多不同的學術領域進行了深入的個案研究。

(DCC, 2011)

第三階段從 2010 年 3 月到 2013 年 2 月，DCC 指出他們已發展足夠的工具和服務來協助資料度用，但他們不可能獨自管理所有英國的研究者所產出的資料，並且 DCC 自身並非資料典藏之單位，因此呼籲英國有能力管理或保存資料的機構、管理人來使用 DCC 的服務及工具，讓 DCC 幫助它們發展管理及保存研究資料的能力，以更全面的推廣資料度用，除了提供服務和工具之外，DCC 也提供教育訓練及最佳案例，甚至是更客製化的協助，希望能在實務工作上也能提供更多幫助。第三階段的核心成員是英國愛丁堡大學、英國巴斯大學和英國格拉斯哥大學，其中以愛丁堡大學居於領導地位，也在 JISC 的資金補助之下建立了提供全國高等教育相關機構典藏研究資料的資料中心「EDINA」(DCC, 2011; EDINA, 2011)。

愛丁堡大學在英國的資料度用發展過程中一直扮演十分核心的角色，其所建構的平台持續獲得 JISC 的贊助，也展現了初步的成果。愛丁堡大學於 1983 年成立了 Edinburgh University Data Library，目標在於協助愛丁堡大學內研究資料的取得、使用和管理，幫助研究者更有效率的進行研究。1995 年 JISC 將 Edinburgh University Data Library 指定為國家級學術資料中心，隨後於 1996 年，EDINA 此名稱正式發布，代表其已成為國家級的資料管理中心，服務對象不再侷限於愛丁堡大學內。(EDINA, 2011)

EDINA 的主要資金來源為英國的高等教育資助委員會(Higher Education Funding Council)之下的 JISC；其目標在於促進英國各大學及機構之研究產出，並提高教學和學習之效率，以此大方向為前提，進行各種研究計畫；同時也提供資訊服務，對象包括英國 90% 的大學以及 2/3 的學院。EDINA and Data Library 隸屬於愛丁堡大學中資訊服務(information service)部門之下，和大學圖書館為同一層級的單位，其所提供的服務內容也和圖書館十分類似，最大的不同點在於圖

書館的典藏以出版品為主，而 EDINA and Data Library 所蒐集並管理的是研究過程中所產生的資料。(EDINA, 2011)

EDINA 和愛丁堡 Data Library 的主要區隔在於 EDINA 的服務對象為全國的大學或研究機構，而 Data Library 則以愛丁堡大學之教職員及學生為主要對象。Data Library 所提供之服務包括「Data Library and consultancy」、「Edinburgh DataShare」以及「Research data management guidance」等，以下就此三項服務以及 EDINA 所提供之服務進行介紹：

(一)、Data Library and consultancy

此項服務之內容類似於圖書館的參考服務，目的在於向研究者提供協助，幫助他們尋找並獲得所需要的研究資料，若是在研究的過程中，遇到不會使用或分析這些資料的狀況，也可提供適當的幫助；若研究者們希望更有效的管理自己的研究資料，Data Library 的職員也可提供建議和方法。Data Library 以整合許多政府機關、學術研究專案所開放使用之研究資料的方式，提供資料之線上目錄，除了檢索之外，也可用學科類別、地區或資料之標題來進行瀏覽。研究者若想要使用這些研究資料，則必須先註冊，並遵守資料提供者的條件和限制，才可下載或使用這些研究資料。此項服務本身並未徵集研究資料，其最主要的功能是進行資源的整合，因此除了研究資料之外，也對英國國內及國際的各種網路資源進行了相當好的整理，使用者可以直接在網站上看到依學科分類的各種資料庫或網路資源，方便研究者搜尋相關資訊，另外也針對教學相關之網路資源進行整理，供研究者參考、使用。(The University of Edinburgh, 2011)

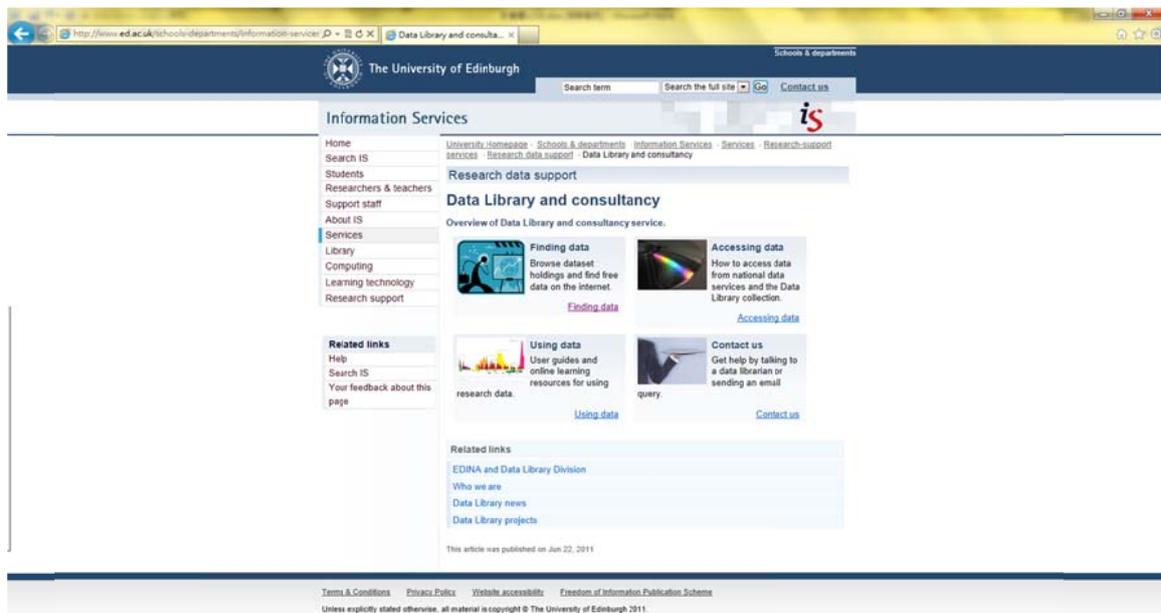


圖 2-1 Data Library and consultancy

(網址：

<http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/consultancy> 檢索日期：2011/12/20)

(二)、Edinburgh DataShare

類似於機構典藏之功能，但對象則是研究資料，以徵集愛丁堡大學內之研究資料，並進行資料之管理和保存為最主要的工作內容；並提供研究資料的檢索和瀏覽，可依照研究社群、資料創造者、資料之日期、標題以及學科領域等等進行瀏覽，亦不需事先登入。每一筆資料的標題之下都有關於該資料之後設資料，對其所屬領域、資料擁有者、日期等進行描述，也提供該資料之引用格式；而研究資料下載的部份，則依其檔案格式不同區分為多個檔案，讓使用者自行選擇需要的檔案下載。DataShare 於網頁上清楚解釋了典藏研究資料的各種優點以及它們的典藏政策，包括典藏資料的同意書、資料內容、後設資料欄位、取用權限等等都有詳盡的說明；下載資料並不侷限於愛丁堡大學內的使用者，而且是免費的，但每筆資料都有其授權內容，是資料的使用者必須遵守的。(The University of

Edinburgh, 2011)



圖 2-2 Edinburgh DataShare

(網址：<http://datashare.is.ed.ac.uk/> 檢索日期：2011/12/20)

(三)、Research data management guidance

以推廣研究資料的儲存與管理為目的，提高教職員們對於資料度用之認知，並鼓勵他們將資料存放至 Edinburgh DataShare。在研究資料之管理部份，由於許多研究者對於在研究過程中所產生的資料其實並不清楚該如何保存和管理，甚至不覺得保存這些資料是有價值或是必要的，網站中針對這類型的研究者提出說帖，首先強調保存和管理研究資料的重要性，也點出這些行為對研究者本身亦是有幫助的，希望他們能夠主動對自己所擁有之資料進行整理。第二步則是介紹研究資料的定義，並說明英國不同政府機關對於資料保存管理的政策。接下來開始對資料管理的實際步驟提出指引，讓研究者了解自己的研究資料是屬於什麼類型、資料的格式有哪些、資料可能被哪些學科的研究者重覆使用等，並列出一張資料管理清單，引導研究者更深入思考有效率的資料管理所應顧及的層面；對於資料如

何建檔和後設資料如何建立也有基本的說明。最後則提醒關於資料的儲存和備份所應注意的一些問題，讓研究者能重視資訊安全並確保資料不會流失。至於在研究資料分享方面，同樣也是先指出分享資料的優點讓研究者們了解，同時也舉出一些無法分享的狀況，例如牽涉到保密協定或隱私權等資料，是否能夠分享就需要進一步的確認；接下來則介紹如何分享資料，包括非正式的分管道，例如研究社群間的電子郵件、wiki 等等，以及正式的分管道，也就是將資料公開在學校或機構的網站上、國家級資料中心等方式。接下來則是對於資料之保護、取用權限、限制以及如何長久保存研究資料等部份進行說明。最後則是鼓勵研究者門可以將資料存放至 Edinburgh DataShare。除此之外，網站中亦匯整了許多資料管理相關課程和工具，讓有需要者能自行取用。(The University of Edinburgh, 2011)

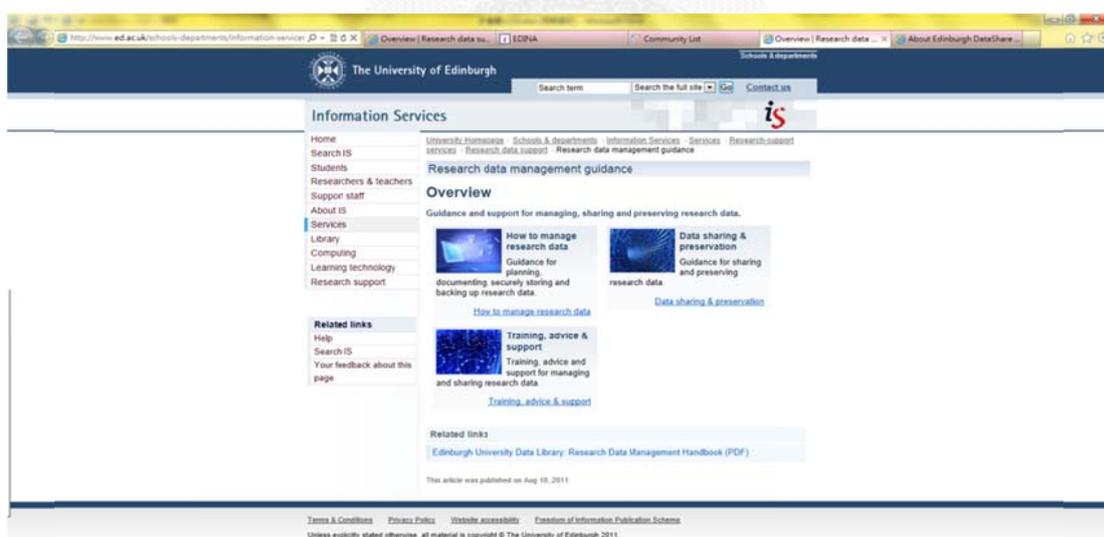


圖 2-3 Research data management guidance

(網址：

<http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/research-data-mgmt> 檢索日期：2011/12/20)

(四)、EDINA

相較於上述三個以愛丁堡大學為主要服務範圍之項目，EDINA 則是以提供全國性的資訊服務為目標，作為一個整合性的平台，搜集相關之網站供使用者取用。其所整合的資源可分為三大類，第一大類是多媒體資料和參考資源，包括影片、聲音、圖片等資料以及期刊目錄、學術文章等相關資源。第二大類則是資料和地圖，包括 GIS 資料、地圖以及許多提供研究資料之網站。第三大類則是給圖書館館員以及開發者使用的資源，提供資料典藏機構的網站以及相關之軟體，讓對資料度用有興趣或想要實行在自身機構中的館員及開發者，能夠快速的取得他們所需的資源。(EDINA, 2011)

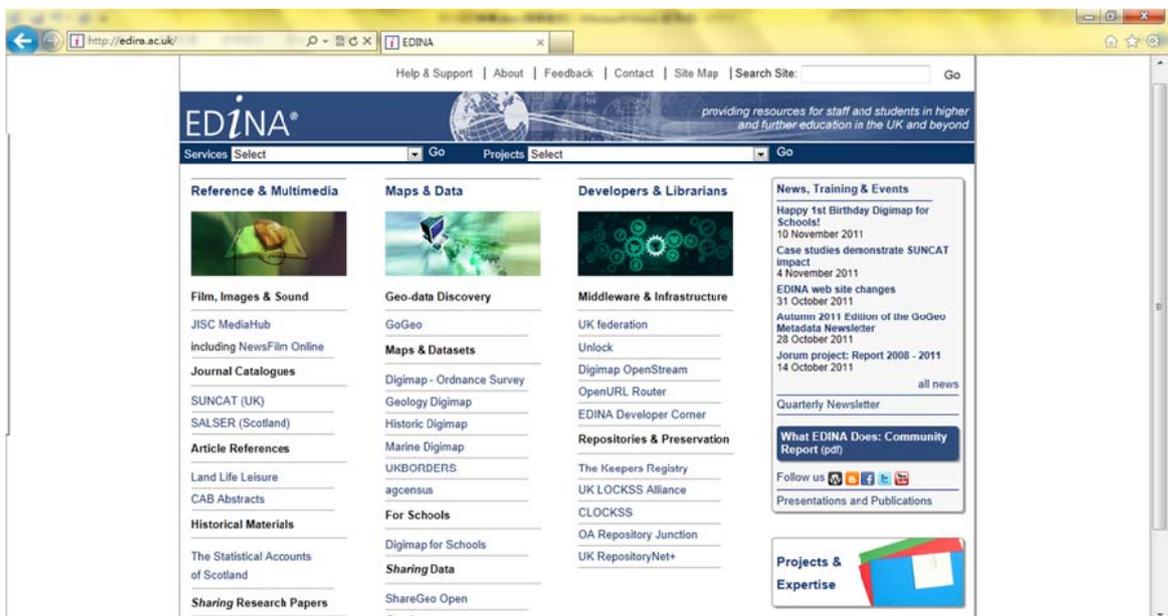


圖 2-4 EDINA

(網址：<http://edina.ac.uk/> 檢索日期：2011/12/20)

愛丁堡大學發展之資料度用服務相當完整，除了針對校內的研究者提供協助管理研究資料的服務之外，對資料度用的重要性和研究資料的價值也有非常清楚的說明，對於智慧財產權、保密協定等也都有初步之解釋，讓研究者了解在什麼樣的情況下可以公開、分享他們的研究資料，優點及該注意的事項有哪些等等，設想得十分周全。同時也提供專門的典藏庫，讓研究者之資料能夠存放。全國性

之服務方面，則著重於資源的整合，除了研究資料之外，也提供參考資源，對於研究之支援相當全面；並提供圖書館館員或開發者在發展資料度用服務時，可能會需要的軟體及現有的典藏庫，讓其他機構亦能發展合適的資料度用服務。

第三節 美加地區之資料度用發展狀況

一、美加整體之資料度用發展概況

歐巴馬總統上任後，大力推行美國聯邦政府資料的透明化，稱之為「Open Government」，將政府各單位所產出之資料的使用權開放給全體國民，並以此為目標，成立了 DATA.GOV 網站，讓使用者能夠方便且迅速的檢索、下載和使用這些政府所產出的資料。網站中匯整了政府各部會所產出的資料，並可依照資料所屬領域和類型進行檢索及瀏覽，資料量十分豐富；美國政府希望透過資料的開放和自由使用，為這些資料創造更多的附加價值，也鼓勵個人或企業開發以資料為基礎的新商機，例如各種網路應用程式(DATA.GOV, 2011)。加拿大在推廣政府資料透明化方面則進展較慢，雖然溫哥華(Vancouver)和納奈莫(Nanaimo)等都市都擁有自己的開放資料(open data)網站，但並無像 DATA.GOV 這樣整合所有政府部會資料之網站，在資料的管理和公開上都稍落後於美國。



圖 2-5 DATA.GOV

(網址：<http://www.data.gov> 檢索日期：2011/12/20)

除了政府以身作則，將資料開放給全體國民之外，美國在政策上的走向也對資料度用之推廣有很大的影響力。美國國家衛生研究院（National Institutes of Health，簡稱NIH）於2008年開始正式要求接受補助的研究計畫除了將發表的文章上傳至PubMed Central之外，也必需繳交符合規定格式的研究資料；美國國家科學基金會則於其2010年5月頒布的新政策中，規定所有於2011年1月18日之後到期或申請之計畫書都應包含1至2頁的「研究資料管理計畫」(data management plan)，其內容則至少包括以下幾點：(NSF, 2010)

1. 資料類型、範例、軟體、教材等，及其他任何在執行計畫過程中所產生之資料、物件；
2. 資料本身及其後設資料所採用的標準、格式和內容。若所處領域缺乏相關標準或規定，則需提出因應的方法；

3. 資料分享和取用之政策，包括對資料的隱私權、保密協定、安全性、智慧財產權或是其他相關權益；
4. 資料再利用、再傳播及產生衍生性產品的相關規定和政策；
5. 對於研究資料和其他研究產出的長期保存及取用之規劃。

雖然美國國家科學基金會並未要求提出申請的計畫書一定要將其研究資料上傳至資料庫或資料典藏單位，但此一政策顯示了國家對於研究資料管理及保存的重視，也促使所有向美國國家科學基金會申請計畫的研究者必須更加重視自己所產出的研究資料，並思考如何對其進行更有效的管理，在資料度用的發展上提供了無與倫比的助力，也彰顯了美國推廣資料度用的決心。以此政策為契機，許多大學開始思考在研究過程中如何給予研究者們更多的幫助，以及如何有效幫助研究者們管理、分享他們的研究資料，讓資料度用之概念和相關服務在美國的大學中受到更多重視。

二、美國的大學之資料度用發展

Sustainable Digital Data Preservation and Access Network Partners (簡稱 DataNet)計畫是美國在發展資料度用方面一個相當重要的里程碑，許多大學和圖書館都參與其中，並且扮演重要的角色；而美國普渡大學和美國康乃爾大學不但參與在 DataNet 計畫中，且此兩所大學在資料度用方面的服務皆被列入 ARL2010 之年度報告中(Soehner, Steeves, & Ward, 2010)，十分具前瞻性且各有特色，值得作為本研究探討之對象，以下分別就這三者進行介紹。

(一)、DataNet 計畫

美國國家科學基金會於 2007 年提出了 DataNet 計畫，其目標在於為科學之發展與教育提供一個保存、管理大量研究資料之資料度用平台，並透過各領域典藏庫之建置，發展資料分享與再利用之服務和工具來達成目的。這個為期五年的

補助計畫挹注了高額の資金，共開放五個申請名額，補助資金總額之上限則高達 10 億美金；2008 年底宣布了兩個通過申請的計畫，分別是 Data Conservancy (簡稱 DC)和 Data Observation Network for Earth (簡稱 DataONE)，預定於 2009 年中開始執行 (NSF, 2007)。以下分別就這兩個計畫之內容、目標，以及研究圖書館如何參與其中做簡單的介紹。

DC 的計畫主持人是美國約翰霍普金斯大學薛瑞登圖書館(Sheridan Library)的副館長，而美國伊利諾大學香檳分校及美國洛杉磯加州大學的研究圖書館也都參與其中。其目標在於整合現存的系統和標準，使其有更好的互通性，並透過模組化、資料的儲存、基本的概念框架等方式，讓跨領域的資料能夠進行整合和管理 (Gold, 2010)。DC 計畫中研究圖書館扮演了重要的角色，同時也有許多不同領域的專家學者參與其中，希望能以使用者為中心打造出一個系統雛型，讓研究資料透過完善的保存和管理，在學術研究上展現出其新的價值。DC 以大氣、海洋、土地、生態和人類活動等領域的相關資料做為初步整合的示範目標，以展現這些大尺度且跨領域的資料之複雜性，因此合作對象也包含許多大型的相關學術研究計畫，希望在發展系統的同時也能提供學術團隊即時的回饋。DC 展現了研究圖書館在學術環境中的新價值，也希望能以此做為研究圖書館發展資料度用之藍圖，促使研究圖書館檢視自身在資料度用上所能扮演的角色或是能夠提供的協助，甚至讓資料度用的發展成為研究圖書館的資金來源，使相關服務能夠長久持續 (DC, 2009)。

取得 DataNet 補助資金的另一個計畫為 DataONE，其目標在於建立分散式且永續性的網路資訊基礎建設(cyberinfrastructure)，其對象是各種地球相關之觀測資料，希望讓這些資料能夠被完整的描述並且容易檢索和取得；DataONE 認為它們所建置的是一個虛擬的資料中心，透過各種工具、軟體甚至是後設資料的整合，讓資料的分享和檢索都變得更迅速。DataONE 也透過最佳範例、線上管理方案之指引和案例、影片檔等方式，來幫助研究者們學習如何更有效的管理自己

的研究資料，讓未參與大型研究計畫的研究者們也能因此受益。使用者可以直接在 DataONE 的網頁上使用關鍵字進行資料的檢索，而檢索結果可依照資料提供者、所屬研究計畫、所屬研究單位等層面進行篩選；DataONE 的重點是與地球相關的觀測資料，因此也提供結合 Google Map 的介面，讓使用能透過地圖上的區域來檢索資料。研究圖書館和圖書資訊相關人員也並未在 DataONE 計畫中缺席，計畫主持人本身是美國新墨西哥大學(UNM)圖書館 e-Science 部門的主任，主導團隊中則有多位來自圖書資訊界的相關人員，DataONE 亦呼籲圖書館可多加利用它們所提供的各項工具及資源，讓圖書館成為計畫擴展服務對象的節點 (Gold, 2010)。除了學術界之外，DataONE 認為教育界和一般大眾都可以透過這些觀測資料來了解自己所居住的環境正在產生怎樣的變化，其服務的對象是非常廣泛的，所蒐集的資料自然也是跨國、跨領域且包含各種尺度的，因此各種標準的轉換和發展更多為資料加值的技術都在 DataONE 的發展目標中。(DataONE, 2009)

(二)、美國康乃爾大學

為因應美國國科會的新政策，康乃爾大學於 2010 年底開始規劃如何為校內的研究者提供新服務，以幫助校內研究者順利的寫出研究資料管理計畫，並更好的保存、管理自己的研究資料。在規劃此服務之初，先根據研究資料管理計畫所應包含之內容，歸納成出幾個項目，來分析校內原本就提供的相關服務和實際需求服務之間的落差；進行校內現有服務的調查後，發現許多已經有提供的服務由於單位之間缺乏合作，因此效率較低，也不容易被大部分研究者所知；領域方面涵蓋的不夠完全則是另一個問題，例如後設資料的標準或是領域典藏庫等，都並非所有領域都有專屬的標準和典藏庫可供研究資料的保存和管理。較大的服務缺口則出現在長期的管理方面，由於目前對於研究資料長期保存所需的資金和人力都沒有充足的證據或研究，因此基本上是完全無法預估的；而研究資料在保存的

過程中，若需要經歷典藏庫和機構之間的移轉，其在保護或規劃上，也沒有明確的政策，因此這部分或許有賴學校更多的宣導或制定相關規範。(RDMSG Planning Group, 2010)

根據上述調查之結果，康乃爾大學成立了一個虛擬的服務團隊「Research Data Management Service Group (簡稱 RDMSG)」來扮演協調、整合和窗口的角色，主要之工作內容就是將校內目前已提供的服務進行整理，並建置網站，讓有需要的研究者可以透過此已網站提供的訊息和聯絡資訊，一次就取得所有需要相關服務，而不需要透過各單位一個一個的進行查找(Research Data Management Service Group, 2010)。美國許多其他大學也漸漸開始提供類似的服務，可見政策的影響力是非常深遠的，同時也讓研究圖書館有很好的切入點，提供更多研究者需要的服務。

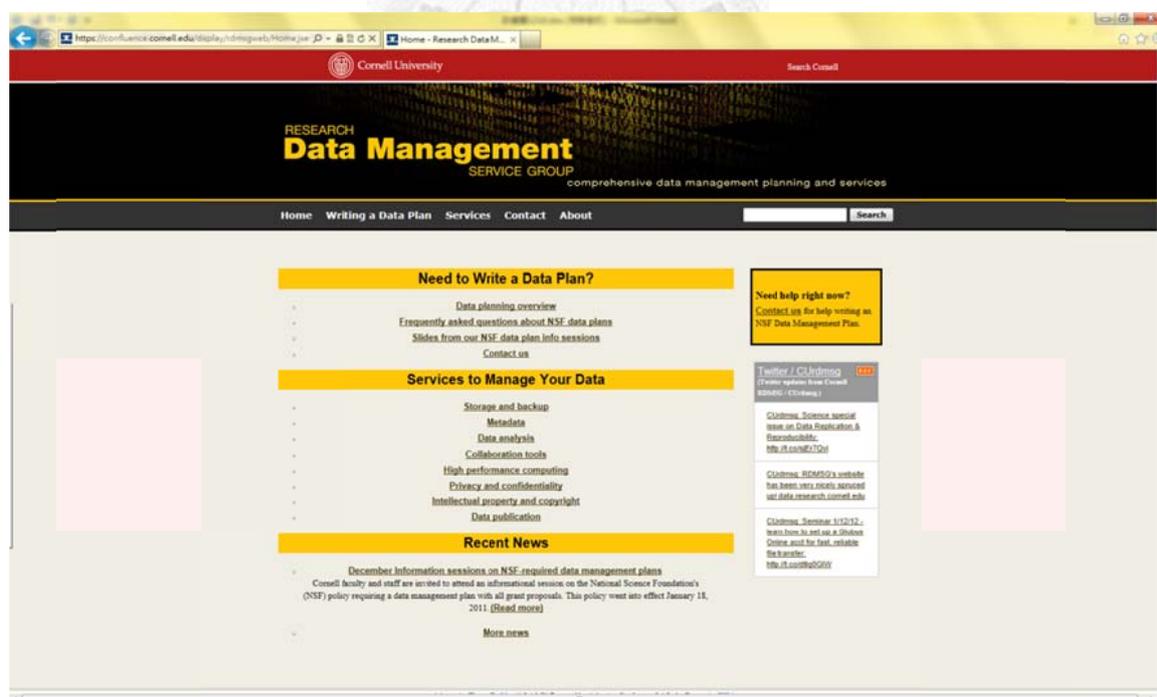


圖 2-6 Research Data Management Service Group

(網址：<https://confluence.cornell.edu/display/rdmsgweb/Home> 檢索日期：2011/12/20)

由康乃爾大學的 Albert R. Mann 圖書館所管理的 Data Staging Repository (簡稱 DataStaR)，是以在研究過程中支援研究資料在研究者或團隊之間的分享為目的，並希望促進研究者們在研究完成後，將研究資料和其後設資料放入永久性的典藏庫中，以供更多使用者再利用。DataStaR 是一個儲存研究資料的「暫時性典藏庫」，其功用在於讓研究者在研究過程中，能以 DataStaR 為平台來和自己的合作對象或相關人員分享研究資料，研究者可自行決定將資料分享給誰，在研究階段中是可以不公開的，取用權限上可由研究者決定。而在研究資料上傳及分享的過程中，DataStaR 可自動產出最低限度之後設資料，再透過圖書館員的協助，依領域典藏庫的不同要求，人工鍵入更多所需的後設資料，在每個研究階段中逐步產生後設資料，以便未來能快速的將其存入永久性的典藏庫中，而 DataStaR 則保存 metadata 和連結到該領域典藏庫之路徑，方便使用者日後的檢索。(DataStaR, 2011)

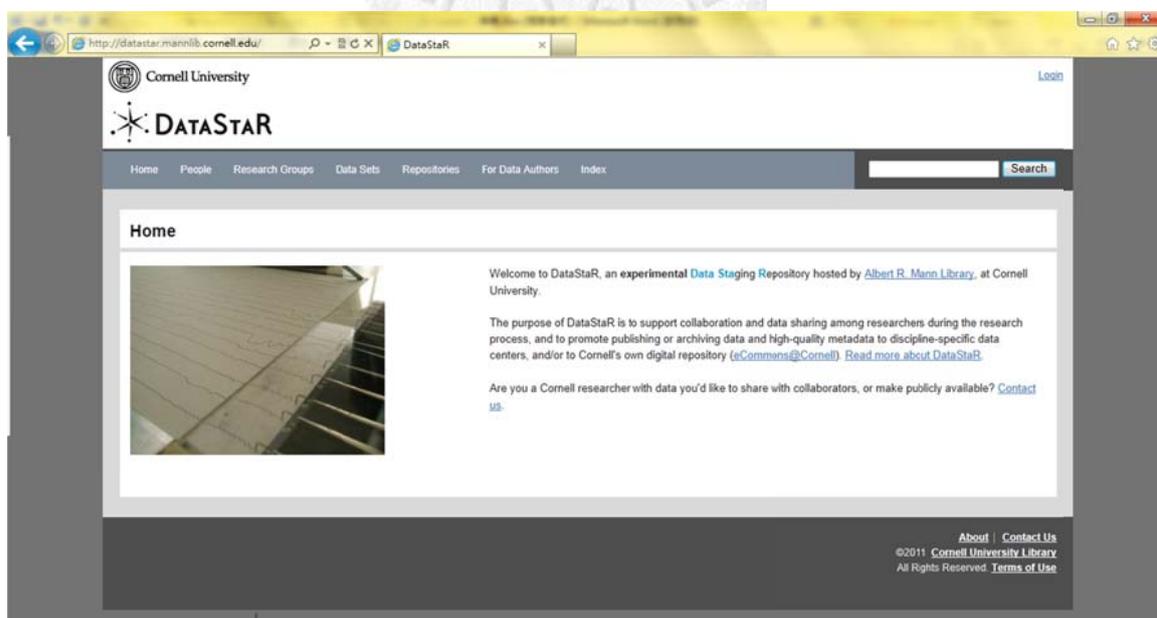


圖 2-7 DataStaR

(網址：<http://datastar.mannlib.cornell.edu/> 檢索日期：2011/12/20)

康乃爾大學所提供之服務以校內的研究者們為主要對象，以幫助研究者們撰

寫出合適的研究資料管理計畫為目標，積極整合校內現有的服務，並分析服務不足之處，作為後續改進的方向。以此服務為出發點，讓大學圖書館更有機會在研究的早期階段就參與其中，並了解該研究所產出的資料，進而推廣 DataStaR 此一典藏庫，讓研究者們一步步的將資料整理、上傳，並建立後設資料，最後再依研究者的意願將其資料存放至合適的資料庫或機構典藏中。康乃爾大學之資料度用服務規劃相當完善，而且在 NSF 政策的趨使下，研究者使用這些服務的機會將大幅提升，有利於資料度用之推廣與發展。

(三)、美國普渡大學

為了有效的對校內研究者們所擁有的研究資料進行調查和整理，普渡大學成立了分散式資料度用中心(Distributed Data Curation Center，簡稱 D2C2)，希望藉由和校內研究者間的合作推廣資料度用。普渡大學並和美國伊利諾大學合作，對校內的研究者進行深度訪談，並將這些訪談進行歸納和整理後，發展出一套資料度用檔案(Data Curation Profile)作為調查與瞭解研究資料現況之工具。(Witt, Carlson, Scott Brandt, & Cragin, 2009)

Data Curation Profile 之工具組中包含了使用者指引 (User Guide)、訪談手冊 (Interviewer's Manual)、訪談工作表 (Interview Worksheet)以及模組 (Template)等四個部份，讓使用這個工具組的訪談人員能快速建立訪談之架構，且每一個訪談段落都可對應到模組中，所以只要將每次訪談之內容依照此工具組所建議的方式填入模組，即可完成屬於該研究資料的 Data Curation Profile。將現存的研究資料建立各別的 Data Curation Profile 將有助於研究資料之管理與保存，也能夠分享研究者的需求讓所有相關人員瞭解，並用以評估目前提供資料相關服務的狀況。該計畫目前已完成普渡大學、伊利諾大學香檳分校與加州大學聖地牙哥分校針對生物、交通運輸和大氣相關領域中，共 13 個研究者的檔案，以這些檔案為基礎，可作為在自身機構發展資料度用服務之指引，並了解不同領域的資料類型和面對資料分享之態度與需求。(D2C2, 2011)

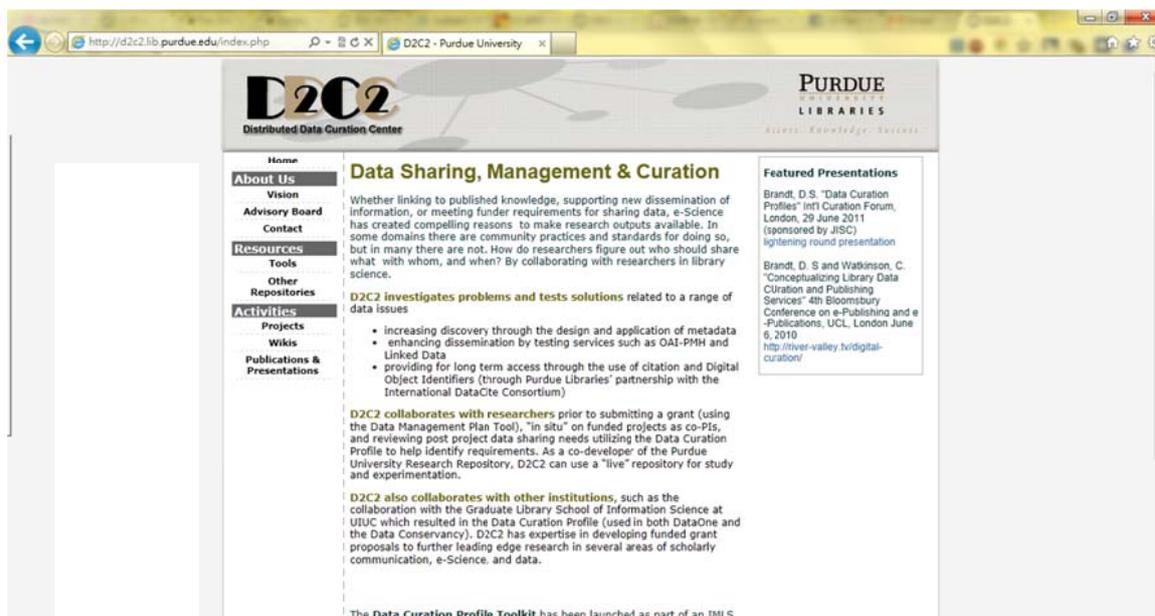


圖 2-8 Distributed Data Curation Center

(網址：<http://d2c2.lib.purdue.edu/> 檢索日期：2011/12/20)

除了針對校內研究者的需求進行了解之外，普渡大學於 2011 年中上線的 Purdue University Research Repository(簡稱 PURR)則提供線上的資料管理與儲存服務，是一個讓普渡大學的研究者們能夠線上分享研究資料之平台，並以成為校內產出之研究資料的機構典藏為目標，持續發展此平台的各項功能與服務。PURR 服務之提供主要是透過校內的圖書館群和 IT 部門進行合作，IT 部門發展平台功能和軟體開發，而圖書館的服務重點在於協助研究資料管理計畫的撰寫以及 PURR 中研究資料的挑選和管理，並會為研究資料加上最合適的後設資料，以符合該領域中常用的標準；PURR 也重視資料的長期保存及安全性，在隱私權、智慧財產權、保密協定等方面亦有相關政策，未來由 PURR 提供之資料都會加上數位物件識別號 (Digital Object Identifier, 簡稱 DOI)，讓 PURR 中的資料在透過網路傳播的同時，也保障資料擁有者的權益(PURR, 2011)。由於此一服務上線尚不滿半年，其公開的研究資料僅有一筆，但從其所提供的服務和對於智慧財產權的重視，可推測其將成為普渡大學推廣資料度用的重要服務。

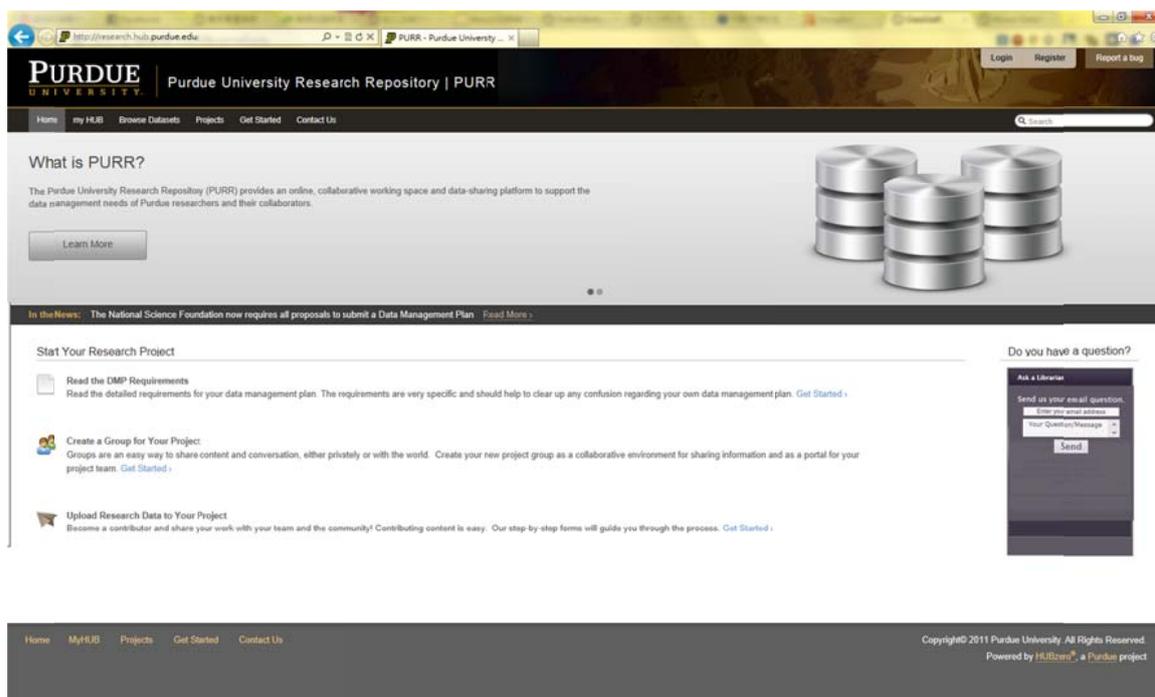


圖 2-9 The Purdue University Research Repository
(網址：<http://research.hub.purdue.edu/> 檢索日期：2011/12/20)

普渡大學在發展資料度用的過程是循序漸進的，首先透過訪談來了解校內研究者的需求以及目前管理、保存研究資料之狀況，並發展出一套完整的訪談工具，讓其他對資料度用有興趣的機構也能使用。PURR 之服務則同時包括了協助研究資料管理計畫的撰寫和分享研究資料之平台，它同時也是以研究資料為對象的機構典藏，學術圖書館在這其中扮演了服務之主要提供者的角色，也讓館員在研究過程中提供更多支援，發展資料度用應是未來圖書館服務的重心之一。

第四節 使用者調查相關研究

學術研究者身為研究資料最大的提供者，同時也是使用者，他們對於資料度用所抱持的態度、期望以及需求，皆會影響到服務提供的方式和內容，因此研究機構或大學在提供相關服務之前，大多會針對機構內的研究者進行初步調查，瞭

解他們的需求以及研究資料管理、保存之現況。以下針對各國在使用者方面的研究進行介紹和分析。

一、英國 DCC 針對不同領域之個案研究

英國的 DCC 對研究資料的保存、管理和再利用等方面投入許多心力，也進行了很多相關研究；其中為了深入了解不同領域研究者在面對研究資料的保存和分享時，其態度和想法上的差異，而針對人文領域、社會科學領域、生命科學領域和理工領域進行了總共 16 個個案研究，其中兩個個案因為研究人員的健康因素而無法完成，因此用了一個較舊的個案來替代，最後將此 15 個個案其進行整合及比較，發現即使是同一領域下的不同學門，也會對於研究資料之保存、管理及分享反映出截然不同的態度和需求，下面就各領域進行深入的討論。(Key Perspectives, 2010)

(一)、人文領域：

此領域只有單一個案-經典文學，而資料分享對此學門之研究者而言並不陌生，大部份都會將其資料公開於機構典藏或領域典藏庫中；經典文學之研究者大多有自己專精所在，因此較少會碰到彼此間的較量或競爭，故他們對於資料分享的態度大多較為正面，而用其他研究者的資料來輔助自己的研究亦是此領域的慣例。對於經典文學領域的研究者而言，資料的保存是沒有期限的，只要還有人可能會對他們的研究對象產生興趣，相關的研究資料就應該被保存下來，研究資料的保存及能夠被取用在研究者心中是和期刊論文或其他學術出版品同等重要的。以英國的情況而言，大型的研究專案往往會對資料的長期保存提供資金，但私人或慈善機構的經費贊助則不一定會有長久的資金去保存資料；至於提供服務的典藏庫，則是以 AHDS (Arts and Humanities Data Service) 最常被使用，但此一典藏庫的經費補助已於 2008 年 4 月到期，雖然目前仍能檢索過去所典藏的資料，但對研究者而言，少了此一機構的服務必會對未來在相關資料的檢索和蒐集上造成

障礙或困難，資料能夠保存之處的選擇也因此減少。

(二)、社會科學領域：

此領域中有 4 個個案研究，分別是農村經濟與土地利用(rural economy and land use)、公共衛生科學(social and public health sciences)、社會互動研究(social studies of interaction)以及建築學(architecture)。此一領域中研究資料常有的共同特徵是與資料提供者之間有保密協定或是對其隱私之保護，這對研究資料的分享或公開造成很大的障礙。以社會互動研究此學門當中的影片資料為例，影像的拍攝在了解人際互動的相關研究中是不可或缺的，伴隨而來的就是道德和法律上的考量，要公開這些影片可說是相當困難，不過為了能持續使用或再分析這些資料，往往會有較長期的保存。公共衛生科學的研究者們其研究資料的來源常是國家型計畫中的縱貫式調查(longitudinal survey)或是世代研究(cohort study)，這些資料本身可能是公開的，但研究者以這些原始資料進行的分析或處理後，並不常將其結果公開分享。

農村經濟與土地利用是針對該研究計畫去進行個案研究，此計畫中的參與者來自社會科學和自然科學的領域，更特別的是他們得到英國國家資料庫(UK Data Archives, UKDA)在資料管理及分享的協助。參與此計畫的研究者因此對於資料分享的方法和其價值有更多的了解，同時亦發現研究者本身對資料公開的價值、優點等認知不足，是造成他們不願意花時間在整理資料並將其公開或典藏的原因之一。對建築學的研究者而言，相較於資料本身，其演變、思考或設計的過程更為重要，因此若要保存資料，則需要將其整個發展的脈絡都加以紀錄並保存，以現階段的狀況而言可說是有相當程度的困難。

對社會科學領域的研究者們而言，研究資料無法公開或是只能在有條件的情況下公開有許多原因，首先是資料本身牽涉到個人資訊，因此有必要保密並尊重其隱私權，尤其當中有醫療相關資料時，其取用的權限更是需要嚴格的控管。長

期的調查或觀察研究需要被研究者的同意，因此研究者們往往不願意冒險去做任何會降低參與意願的事。另一方面，許多研究者擔心的是自己的資料被不正確的使用或解釋，進而造成誤解。研究計畫的出資者也有他們的考量，在確定該研究資料的價值被發揮的淋漓盡致之前，他們不一定願意公開這些資料讓其他人佔便宜。有些資料則是以近似於付費資料庫的方式存在，只有在持續支付費用的情況下才能使用，研究者通常不會想免費的公開他們買來的資料，尤其在這些資料很貴的情況下更是如此。

對建築學領域的研究者而言，檢索或取得他們需要的研究資料在現階段而言是困難的，原因之一是缺乏較好的標準和後設資料，造成檢索上的障礙；另一方面，對於設計的思考過程、創意和想法等部份都難以紀錄下來，造成資料在脈絡上不夠清楚，增加了再利用的困難。影片資料也面臨類似的問題，缺少被普遍使用的後設資料，不同學術社群、團體所採用的註解方式也有很大的差異，因此檢索和再利用這些影像資料是困難的。有些研究者希望自身所處的機構能為影片資料的長期保存和管理提供解決之道，至少讓他們有存放檔案的空間，但目前各機構的典藏庫少有提供相關的服務。相較之下，公共衛生科學的資料大多有其資金來源可供長期的保存和管理；每隔一段時間會針對資料進行審查和再檢視，以確認是否有繼續保存的價值。不過對於較小型的研究計畫或個人的研究來說，其資金來源便無法確保研究資料的留存，相較於大型、長期的研究計畫，這些資料可能會被認為相對的不太重要，且研究者大多認為研究資料之引用在學術上的回饋不夠直接，缺乏相關的獎勵機制，因此降低了他們付出時間和心力去整理並保存研究資料的意願。

(三)、理工領域：

此領域共有 4 個個案研究，分別是天文觀測學、氣象學、結晶學和針對英國巴斯大學機械工程系的創新設計與製造研究中心(Innovative Design and Manufacturing Research Centre)所進行的個案研究。上述學門對於資料願意分享的

程度都各有不同，對天文觀測學的研究者們而言，資料的分享可說是他們研究中的一部份，他們透過天文望遠鏡取得影像、光譜和光變曲線(light curves)後，通常會花上一年的時間進行分析和處理，接下來這些資料就會被公開給大眾，資料的公開在天文觀測學的研究者間是不成文的規定和默契，也已經行之有年。一手的觀測資料通常會存放於國家級的典藏庫或是各地研究中心的典藏庫中，經過處理的資料雖然大部份由研究者個人保存，但若是基於研究需要向他們索取，大多都會同意讓其他研究者使用，這些研究者認為資料的分享不只有助於整個研究社群的進步，也讓他們得到學術聲望上的回饋。

氣象學的研究者們根據自己研究專長的不同，其對於研究資料的分享態度亦不相同，以電腦程式模擬氣候的研究者們並不習慣分享他們的模式或其產生的資料，這些資料通常儲存在他們的研究室或個人電腦中；相較之下，進行氣象觀測或是發展海洋模式的學者則較常分享他們的資料。英國的 British Atmospheric Data Centre (BADC) 負責保存氣象領域相關研究資料，而其出資單位 Natural Environment Research Council (NERC)也規劃在未來提出資料公開的相關政策，來幫助氣象領域中資料的分享，解決目前難以尋找、取得相關研究資料之情況。

對於結晶學研究者而言，英國的主要研究經費來源是 Engineering and Physical Sciences Research Council (EPSRC)，由於其並無關於研究資料之管理和分享的相關規定或政策，因此由其出資進行的研究計畫很少會將其資料公開。其它的經費的來源可能是民間機構或公司，因此資料會因保密協定的緣故而無法公開，故對結晶學領域來說，可能有一半以上的研究資料都未被分享，自然也不易取得。針對巴斯大學的創新工業設計研究中心之個案亦發現資料的分享並不普遍，其資金來源常是民間公司是主要原因之一，設計相關的資料會被視為商業機密加以保存，甚至不一定能和研究中心的其他人分享；在沒有保密協定的情況下，研究者們面對研究資料分享的態度仍有很大不同，有些人願意分享，但也有些人認為不會有其他人對他們的研究資料有興趣或進行再利用。

(四)、生命科學領域：

生命科學領域中研究資料近年來以驚人的速度不斷成長，產出大量資料的研究計畫也不斷增加，讓生命科學成為一個相當資料密集的領域；即使是在這樣的背景下，有些生命科學中的學門對研究資料的分享依照採取相當保守的態度。以神經系統學門中的一個個案為例，此研究計畫的猜參與者認為要對自己所產出的研究資料有全盤的掌控權，並沒有主動公開的意願，雖然他們也認為在研究成果發表後，這些資料可以提供給領域中相關的研究者，但這也表示要等上很長一段時間才有可能取得。神經心理學的研究者們也反應出類似的想法，他們不願意分享資料的最主要原因在於資料本身的保密協定，另一方面，他們認為在不了解研究假設、研究內容和限制的情況下，其他人不一定能夠正確的理解這些資料的用途，進而造成再利用或闡述上的錯誤。神經心理學中有很多影像資料，這些資料常有足以辨識出個人的特徵，因此在使用上需要特別小心，自然也不可能隨意的將之公開分享；除了隱私上的問題之外，此學門的研究者們也非常重視研究本身的價值，資料分享被認為可能會影響他們的學術競爭力。

相較於上述兩個持保守態度的學門，生命科學領域中也有許多學門是非常積極的公開或分享他們的研究資料。以一個研究小鼠基因的研究專案為例，他們最大的產出就是一個公開小鼠的基因資料庫，而相關學門的研究者們，例如分子生物學和發展生物學，也都很樂意並主動的把他們的研究資料存入此一資料庫中。另一個對資料分享非常積極的個案是基因組學，在此學門中，許多提供研究經費的單位都對研究資料的分享和保存管理持正面態度，其中英國的生物科技和生物科學研究委員會(Biotechnology and Biological Sciences Research Council)直接要求向它們申請經費的研究計畫都要附上處理研究資料之計畫(data plan)。基因組學的研究者們將他們的資料存放於大型且公開的資料庫已是慣例，甚至有些期刊會要求研究提出他們已將資料存於資料庫的相關資訊，在期刊文章或研究成果中註明資料存放之處也是大多數研究者會主動做的事，甚至有些研究者會在產生

資料後，立即透過個人網站、部落格等方式進行分享。

系統生物學的研究者們亦非常重視資料的分享和再利用，他們的研究中不但會產出大量的資料，也經常會將現有的資料進行分析、比較，試圖從其中發現新的研究成果，最常使用的資料庫是基因和蛋白質之資料庫，有些大型的研究團隊甚至每天將其資料上傳至相關資料庫。不過除了基因資料有其固定可上傳的資料庫之外，此學門中亦經常產生的顯微鏡影像資料就顯得無處可歸，由於目前尚無大型且政府資助的資料庫可典藏這些影像資料，因此大多會保存在自己的電腦或研究室中，儘管如此，有些研究者仍會在網站上透露自己擁有的相關資料並歡迎其他研究者索取，對他們而言，資料的分享可說是研究生命周期中的一個部份。

從上述個案中可發現每個學科都有其不同的文化，因此面對研究資料的分享、保存和管理等議題，研究者們的態度和行為都有顯著的差異。對天文觀測學、海洋觀測與模擬、基因組學、系統生物學等學門而言，研究資料的分享與保存早已是研究生命週期的一部份，因此這些學門也大多擁有國家級或國際級的大型資料庫，供他們使用並保存相關的研究資料，除了英國的 British Atmospheric Data Centre (BADC) 之外，美國的海洋暨大氣總署 (NOAA) 和國家大氣科學研究中心 (NCAR) 等都是十分具權威性的資料庫。生物領域中基因相關之學門有幾個非常重要的資料庫，像是 Protein Data Bank (PDB) 和 National Center for Biotechnology Information (NCBI) 等，都是國際級的大型資料庫，收錄的資料筆數相當可觀。

社會科學領域中，有許多國家或國際機構會定期的進行大型或長期的調查，這些縱貫式調查 (longitudinal survey) 或是世代研究 (Cohort Study) 等等常是社會科學領域研究者的重要研究資料來源，英國最具代表性的資料庫是英國國家資料庫 (UK Data Archives, UKDA)，而美國則是由密西根大學所營運的 ICPSR (Inter-University Consortium for Political and Social Research)，這些資料庫

也會徵集社會科學研究者們所產出的研究資料，可惜的是並非所有使用這些資料庫的研究者都會將其研究成果主動提供。以台灣的情況而言，中研院調查研究專題中心的學術調查研究資料庫也是以徵集社會科學領域的研究資料為其任務，收錄了政府調查資料、長期主題型調查和追蹤調查、國際合作調查等等，內容相當豐富，是目前國內社會科學領域中相當重要的資料庫。社會科學的研究者們比其他領域都更容易遇到因為隱私權而無法分享研究資料的問題，質性研究中進行的影片拍攝或是田野調查等，也許會因為碰觸到一些敏感議題而無法分享。

相較之下，有些領域目前的資料分享和保存都尚不成熟，但原因則各自不同，結晶學的研究者常與企業合作，因此其研究成果和研究過程都可能因商業上的考量而無法公開，甚至不能分享給同領域的研究者；神經心理學的研究者們除了考慮到研究資料可能透露隱私訊息之外，也擔心資料會被錯誤解釋；建築或設計相關學門則是因為其最有價值的部份是創意和構思之過程，目前仍無合適的保存或分享方式。這些障礙讓上述之學門亦較為缺乏大型或跨國的資料庫，在資料效用方面仍有許多進步和發展的空間。

二、歐盟的 PARSE. Insight 大型跨國調查

由 EC 所主導的 PARSE. Insight 研究專案，於 2008 至 2009 年間針對學術研究者、資料管理者和出版相關者進行跨國的大型調查，目的在於瞭解各國對於數位研究資料的保存、管理之現況，以及學術社群對這些資料進行長久保存的需求和意見。此研究採用問卷調查法，問卷透過學會、聯盟、期刊訂購名單等管道以電子郵件的方式發送，共回收 1,840 份問卷，其中學術研究者占 1389 份，資料管理者占 262 份而出版相關者占 178 份；後兩者的回覆率遠較學術研究者低，且由於資料管理者和出版相關者的總人數難以估計，因此無法確定此回覆率是否具有代表性；故單就學術研究者所回覆之問卷進行分析，則包括歐盟、美國、澳洲、日本以及其它國家都有學術研究者回覆此一問卷，其中歐盟國家占 44%，美國占

33%，其他國家則占 23%。若用領域來分，則以理工領域(physical)的比例最高，占 33%，第二是科技領域(Technology)，占 14%，再來則是生命科學(life science)和社會科學(social science)領域，分別占 13%和 11%。研究結果歸納出研究者們認為長期管理並保存資料的最重要的前三個原因，依序為：(EC, 2009)

1. 現存資料的再分析、再利用；
2. 可供日後之驗證；
3. 能夠刺激科學的進步及研究的深度。

由調查結果可知，研究者們認為將研究資料保存、管理的原因在於分享、公開這些研究資料後對學術社群的益處，因為唯有研究資料的分享才能促進再利用、研究的驗證和深化未來的研究發展，這說明了研究者們對於資料的分享普遍抱持正面態度。

雖然研究資料的長期保存、管理和分享有許多優點，但在目前的學術環境中卻不普遍，可見其中必定存在一些障礙導致研究資料並未普遍的被分享。EC 針對研究資料之保存和分享所遭遇的障礙分別進行調查，發現研究者認為現今學術環境中資料保存的障礙，排名前三的項目依序為：

1. 因為硬體、軟體等問題造成資料無法被取用；
2. 資料現在的保管者或是計畫經費來源或許無法長久維持；
3. 使用者可能無法瞭解資料真正的意義或用途。

從排名第一和第二的項目可以發現，經費對研究者而言是一個讓他們擔憂的問題，硬、軟體沒有更新或資料格式老舊造成無法取用，反映出人力的問題；而計畫經費的不穩定，讓研究者們為了爭取下一個經費來源，而無法將研究完成後的研究資料進行更多的整理或組織，造成在分享和保存上的困難；而第三個項目則突顯資料描述的重要性。

關於分享研究資料的顧慮，法律問題和資料誤用並列第一名，有 4 成以上的

研究者都對這兩個因素感到擔憂，第三名和第四名則為資料格式不相容以及缺乏良好的資訊基礎建設，可見除了研究者本身意願之外，科技的因素也會對資料分享造成影響；調查中也問到研究者們是否曾經需要其他人的研究資料卻無法取得，發現有一半以上(53%)的受訪者都有這樣的經驗。此研究也調查了研究者們的資料類型和保存現況，發現大部份研究者都擁有多種格式的資料，office 文件占 94%，網路為基礎的資料和圖像並列第二，都是 79%，而資料的總大小則大多落在 1GB 至 1TB 左右。研究者目前的資料保存以工作上使用的電腦為主，占 80%，隨身型的資料儲存裝置則為第二，占 66%，第三和第四名則是機構中的伺服器 and 家中電腦，都占 50% 以上，可見將研究資料放於典藏庫或資料庫中並非常態，並且有 84% 的研究者都表示沒有對資料的管理有進一步規劃。

三、澳洲政府的 DART 研究專案

為了更深入了解如何提供研究過程中所產出之研究資料的度用服務，Denison、Kethers 與 McPhee (2007) 等人在澳洲政府所支持的 Dataset Acquisition, Accessibility and Annotation e-Research Technologies(DART) 計畫中，針對研究者們在研究過程中，面臨關於資料管理或分享等方面的問題進行研究。DART 是一由多所大學所合作的國家型計畫，其關注的焦點是如何支援合作研究的過程，並促進對研究資料之取用，特別是針對由政府所補助之研究計畫；希望藉由 DART 的進行以瞭解多領域研究資料的蒐集和管理、資訊基礎建設的發展性和永續性、研究過程中智慧財產權的問題以及新形態的學術出版模式等。為達成上述目標，DART 有三大示範型計畫，分別是「bioinformatics and X-ray crystallography」、「climate research」和「digital history」，皆為跨領域的合作研究。此研究的進行方式則是派三位資訊管理人員，分別以一周一天的頻率參與三個不同專案團隊的討論和研究，了解其工作流程、資料管理的方法並將這些記錄予以保存，希望藉由親身參與的方式，蒐集一手的觀察資料，也讓研究者能針對資料管理方面的問

題提出意見和想法，同時藉由與資訊管理人員的互動，改進團隊中的資料管理程序。

除了觀察法之外，此研究亦進行了 15 個單獨訪談，面談的對象並不限定於 DART 之成員，其中 6 個是來自 DART 計畫之外；主要希望了解研究人員的研究內容、所使用的資料、目前的資料管理方式、有哪些相關需求、在資料分享上是否遇到障礙、希望得到協助的項目等問題。雖然 DART 在規劃時已考慮到領域之差異性，而提出三大不同領域的示範計畫，但研究結果呈現之重心仍在於相同點的歸納，針對不同領域的需求差異並未深入說明。由本研究之成果可發現研究者們對於資料分享的考量和需求有下列 3 點：

1. 由值得信任的機構來提供服務：

經費充足與否十分受到研究者們的重視，因為這關係到資料是否能夠被長久的保存和取用，所以研究者們傾向將資料交給自己能信任的機構來保存及管理。換言之，提供資料度用服務之機構本身是否具公信力，會影響到研究者們提供內容的意願。

2. 資料的安全性以及法律相關問題：

法律問題是許多研究者們的主要考量，很多研究者非常關心誰會使用他們的研究資料，除了著作權或智慧財產權上的考量外，也擔心資料的錯誤闡釋會造成研究上的謬誤；另外有些學者因考慮到出資者有其競爭對手，認為他們的研究資料不能為競爭對手所知，甚至連後設資料都須保密，因此非常重視資料的保密以及安全性。

3. 重視立即的回饋：

研究者們希望在資料提供的同時享受到立即的回饋，由負責資料度用之單位協助處理授權問題，讓他們無後顧之憂，並提供更迅速且有效的典藏工具，來協助他們管理研究過程中所產生的各種資料，節省他們的時間，讓工作效率提高而

非花更多時間在提供內容上。立即的回饋對研究者們而言是重要的需求，也提供了更大的誘因促使研究者們管理並保存其研究資料

本研究的最後提出了圖書館所能扮演的角色，因為越來越多的研究是以跨領域或跨機構的方式進行，使得資料的所有權變得模糊不清，而圖書館則可以站在中立者的角色，進行資料的蒐集和典藏。圖書館過去在各種資料分類架構上的經驗，也能有效的運用在研究資料的管理上，幫助研究者們建立長久保存資料的模式，並提供相關的知識和經驗，因此圖書館可說具有舉足輕重的地位。

四、英國牛津大學針對校內研究者的訪談

Martinez-Urbe (2007)為協助英國牛津大學深入瞭解研究人員對研究資料度用服務之需求，於2008年5月至6月間訪談37位牛津大學的研究人員。該訪談以研究的生命週期，包括經費的申請、資料的蒐集、資料的分析和處理、研究成果的發表等，做為訪談的基礎架構，以瞭解研究者在整個研究過程中的哪些環節會需要什麼樣的服務。該研究結果顯示，大多數的研究者都認為在研究資料管理的層面上，仍有許多潛在的服務可幫助他們更有效率的進行資料管理。研究者最重視的三大需求為：(Martinez-Urbe, 2007)

1. 安全、容易使用的大量資料儲存系統，且能夠有效控制資料的分享。研究者們十分重視資料的安全，這當然包括硬體上的維護，不但要保證資料的權限控管，也要讓其能夠長久的保存，不會因為硬體或經費問題而造成資料的流失。
2. 針對未有專門的國家級機構進行資料典藏的學科領域，提供長期維護的資訊基礎建設。有些領域擁有發展完善的典藏庫或資料庫，因此資料不怕無處可去，但也有些領域並沒有這種學科導向的典藏庫可供存放資料，大學應對這些領域的研究資料之保存和管理有所規劃。
3. 在資料的管理上給予實際建議和更具體的作法。雖然研究者也瞭解資料

管理的重要，但由於經費和時間上的限制，他們並不會思考如何整理資料以進行更長久的保存，大部份都以自己的方式為資料下註解或歸檔，但也有人表示幾年之後，就連自己都不懂當初所下註解的意義。因此研究者需要的是具體的方法和幫助，讓他們能有效的整理研究資料，並使其能被其他人理解和使用。

該研究同時指出，研究人員的研究資料大多儲存在自己的個人電腦中，大部份的研究者較願意在自己熟悉的學術社群中透過一些非正式的管道來分享他們的研究資料，例如用電子郵件和個人網站等方式進行資料的分享；一旦資料不斷累積，保存和分享就會越來越困難，故應即早讓他們知道要如何有效的整理並分享、保存這些資料。此外，有些資料本身涉及隱私權、保密協定等問題，而研究者大多對這些法律問題缺乏深入的瞭解，讓他們對資料的分享、再利用感到困擾和疑慮，是目前發展資料度用的障礙之一。(Martinez-Urbe, 2007)

五、美國普渡大學針對大學中研究者的訪談

Cragin, Palmer, Carlson and Witt(2010)等人為了調查大學中的研究者們在什麼樣的情境下願意和什麼人分享哪些資料，訪問了來自 12 個不同領域的共 20 位研究者，這些研究者們在各自的領域中從事會產生、使用大量資料之研究，或是會產生許多不同類型、格式之資料的研究，因此對於資料度用之議題較有興趣，也比較會有相關之需求，透過圖書館之學科館員和這些研究者取得聯絡，再進行兩階段的訪談，第一階段的重點在於了解研究者目前的研究範圍，進行中的研究計畫或是曾經執行會產出較多研究資料之計畫，並透過半結構式的訪談大綱來了解研究者們對於資料度用服務有何期待和需求；第二階段的訪談則根據第一階段的成果形成更為客製化的問題，深入了解資料類型、是否願意公開或分享、資料之管理和描述等項目，再將之成果進行分析。(Cragin et al, 2010)

此研究的成果分析共有下列 4 個面向：(Cragin et al, 2010)

1. 可被分享之資料類型：

研究者們認為可被分享的資料應該是那些有學術價值或是能夠被再分析、再使用的，因此越是經過適當的整理之資料，他們分享的意願越高，並且對於資料的分享或公開大多持正面的態度。以產出資料之類型而言，大多是影像、表格和資料庫等，很多研究者在同一研究中會產生不只一種類型的資料；60%的研究者認為從資料的公開是需要緩衝期的，時間則從 1 個月到 5 年不等，也有些人認為這些資料不該被公開，只願分享給少數自己信任的研究夥伴。很多研究者提到共同作者的問題，認為資料公開或分享與否，以及時間點等問題都不是他們自己可以決定的。

大多數的研究者都認為自己的研究資料是有長久保存的價值，其中 13 位研究者認為資料至少要保存 10 年以上，而這種資料大多是屬於觀察型之資料，具有和其他資料相互比較之價值，或是和未來產出的其他資料一起進行整合、分析。有些領域的研究者提出對於特定資料之典藏庫的需求，因為他們所處的研究領域人數很少，在資料的分享上相對困難，也不容易了解或取得其他人的資料，因此希望能有研究社群取向之典藏庫或保存、分享資料之平台，以促進該研究社群中之資料分享和資訊交換。

2. 私有與公開之界線

資料之分享其實可分為不同之形式，一種是將資料提供給索取者、合作的研究者或是認識的人；另一種則是將資料直接公開，不論是透過網頁或資料庫，讓更多人能直接檢索或瀏覽其內容。在這兩種資料的分享方式間，大多數研究者的需求是在資料的提供上，很多人提到將資料攜帶去參加研討會的困難，或分享之空間不足等問題，但若論到直接將資料公開，則抱持保留的態度，一方面是因為有些資料有持續深入研究及發表成果的價值，另一方面則是很多人擔心自己的資料會被錯誤的解釋或使用不恰當。

3. 公開給大眾

研究資料若被公開的原因通常有下列三種：首先是個人因素，例如需要其他研究者的協助或建議；第二則是該研究領域之慣例，必須將資料公開讓他人都能使用；最後則是來自於資金補助單位或所身處之機構的要求。對研究者本身而言，研究資料的公開牽涉到許多因素，包括資料如何流動，是分享給單一的對象或是同時有多人能使用；自己保有多少對資料的掌控等等，因此即使有外在的因素要求研究者們公開他們的資料，公開的程度仍需要在研究者自身對資料公開的掌控、意願和外在的回饋或壓力之間取得平衡點。以本研究中的研究者而言，資料的公開是一種評估和衡量損益之過程，必須考慮到對象和自己的熟悉程度、是否值得信任，以及整理資料所需之成本和複雜度、完成度等等；有些研究者則表示完全不考慮公開或分享給自身學術社群以外的人。

4. 避免資料之誤用

資料的誤用對大多數研究者而言，是在研究資料之分享上一個不可忽視的大障礙，也是研究者們最常拿來解釋他們之所以不分享資料或是僅限於分享給少數人的原因。雖然不一定有資料被誤用的經驗，但對此一因素的擔憂確實造成研究者們對於資料公開或分享之意願降低。此研究中的一位研究者也講述他的親身經歷，也就是另一位研究者用他仍在分析中的研究資料發表了相關成果，而他在分享自己的資料之初完全沒有想到會有此種情況發生，讓他從一個積極的資料分享者，變成一個保守者，只分享給同一研究團體中的人，並在結果發表後才公開資料。另一位研究者則認為自己的研究資料若保存在公開的資料庫中，可能會被營利性質的團體挑選其中對他們有利的部份，當作行銷工具或其他商業用途，而這部份目前並無任何協訂或條款能限制非學術之團體或組織對資料的使用。

除了資料的誤用之外，對於資料的錯誤理解或闡釋也是研究者們所擔心的，有些研究者認為如果資料分享的對象無法正確理解這些資料所代表的意義或用途，則他們不會想分享給對方；錯誤的引用也屬於誤用之範圍，有一位研究者分

享他的經驗，表示曾經看過一個研討會海報上的寫明是引用他的資料，但事實上並不是，甚至他看不懂那個海報中的研究內容，但這樣的錯誤引用讓他無從辯解或申訴。從這個研究中可看出資料誤用的類型多樣化，沒有單一的解決方案可適用於所有情境；再加上不同的研究者會有不同的顧慮和承受風險的程度，使得資料誤用之議題更加複雜，是機構在發展資料度用時需要特別重視的。

六、使用者相關研究之整理與分析

綜合以上使用者調查的相關研究可發現，研究者們對於研究資料度用的態度與需求是調查的重點所在，其中對研究資料度用的態度又可大致分為資料分享或公開的原因、所遭遇的障礙等兩個面向。另一方面，對於研究者管理及保存研究資料現況的瞭解也十分重要，這有助於提供資料度用服務單位掌握目前機構內研究資料分佈的狀況，瞭解研究者處理這些資料的習慣，對於發展出符合需求的系統將有所幫助。因此根據使用者調查的文獻整理出的四個重要的面向分別為資料保存現況、研究資料分享或公開的原因、研究資料分享或公開會遭遇的障礙，以及對資料度用服務的需求；故上述四個面向將是本研究訪談所關心的重點所在，以期在研究完成後能和國外的研究成果進行比較。

從資料保存的現況來看，研究者們所使用的資料其格式很多樣化，有文件、圖檔、資料庫等等，而這些資料主要保存在研究者工作上或個人電腦中，也沒有固定的整理或歸檔方式，仍以個人習慣或使用方便為主。以保存、分享資料的原因來看，研究者們對研究資料的重要性和獨特性皆有一定程度的認知，在能夠再利用這些資料的情況下，他們是有意願保存這些資料的，而不論是再利用或驗證，都表示這些資料會有一定程度的分享，可見研究者們並不排斥分享他們的資料，甚至樂觀的認為分享將有助於該領域之研究的發展和進步。除了研究者本身的意願之外，研究領域的慣例以及出資單位的要求等外部因素，也是資料被公開或分享的重要原因。

若從分享或公開資料之障礙來看，可發現目前的學術環境對於研究資料的保存並未給予太多正面的幫助，由於缺乏學術上的回饋和經費有限等狀況，使得研究者們傾向花時間進行新的研究，而非整理舊有的研究資料。另一方面，研究者們有對於自己的資料會如何被使用感到擔憂，若使用者不了解資料產生過程中的研究情境，可能會誤解或誤用這些資料；由於硬體或軟體方面的問題，造成資料的遺失或不完整，亦可能降低資料之可信度；而像是版權、所有權、著作權、保密協定等法律問題，對很多研究者而言並不會去深入了解，若是牽涉到與個人相關的隱私權，問題就更為複雜，故加倍的令他們不想花時間去面對和處理。

對於資料度用之需求方面，研究者們都認同研究資料保存之重要性，故認為應先加強缺乏領域典藏或大型資料庫的學科的資料的保存，但同時他們也期待能有一個穩定而可信賴的機構來提供相關服務，呼應了他們對於經費持續性和資料安全性的顧慮。在服務層面上，研究者們希望是容易使用且能讓他們享受立即的回饋，不是要求他們付出多餘的心力和時間，而是給予實際的幫助和建議，以提高效率、促進研究。

由於本研究採用訪談法，故研究成果比較的對象也以同樣採用訪談法的普渡大學之研究、牛津大學之研究以及 DART 研究專案等三個使用者調查為主；將上述 3 個研究之成果依據使用者調查的 4 個面向進行整理，如表 2-2 所示，以供後續與本研究之訪談結果進行比較。

表 2-2 國外使用者調查之研究結果分析

使用者調查之面向	國外使用者調查的研究結果
資料保存現況	<ol style="list-style-type: none"> 1. 所擁有的資料格式多樣化 2. 資料儲存處以個人或工作用電腦為主 3. 資料整理、歸檔以個人經驗或使用方便為主
研究資料分享或公開的原因	<ol style="list-style-type: none"> 1. 研究的驗證以及資料的再利用等學術上的幫助 2. 研究領域中的慣例

	<ul style="list-style-type: none"> 3. 個人研究上的需要，對象通常為熟識的其他研究者，或研究團隊內部分享 4. 出資單位或所處機構的要求
研究資料分享或公開會 遭遇的障礙	<ul style="list-style-type: none"> 1. 資料誤用或誤解 2. 著作權或智慧財產權等法律問題 3. 經費或人力不足 4. 硬體、軟體等科技層面之因素
對資料度用服務的需求	<ul style="list-style-type: none"> 1. 值得信任的資料度用平台 2. 在資料的取用或整理上有立即回饋 3. 資料的安全性 4. 先針對有需求的領域提供服務



第三章 研究設計與實施

本研究運用訪談法，訪問臺大教師對於資料度用之看法與需求，並瞭解臺大教師目前如何管理及保存研究資料，希望能提供臺大圖書館作為未來發展資料度用之參考。本章節次安排首先是研究方法，其次為研究對象，最後是研究流程。

第一節 研究方法

由於目前國內尚未有學術圖書館提供資料度用系統，資料度用之概念亦不普遍，故為深入瞭解研究者的看法，同時考慮到大多數的研究者對於資料度用之瞭解可能有限，本研究採用質性研究的訪談法，以期在對談的過程中加以解釋資料度用之意義，透過深入的對談來瞭解研究者的看法與需求。本研究採用半結構式訪談，希望能多聽取受訪者們的意見，並根據訪談時的互動，觀察臺大教師們關心的重點，再透過進一步的訪問，來瞭解老師們的想法。

訪談法是一種有目的的談話，研究者運用口語敘述的形式，針對特定對象收集與分析有關的資料，以便對研究的現象或行動有全面式的瞭解。由於研究者本身即是訪談法的主要研究工具，在訪談過程中除了具備「問問題」的技巧之外，還需輔以積極的傾聽與仔細的觀察，研究者本著開放的態度與彈性的原則，創造出一種自然、舒適的情境，讓受訪者在被尊重的互動中，儘量表達內心的看法，共同建構研究現象的本質與內涵，進而透過詮釋分析的過程，歸納出研究主題的脈絡與意義。(潘淑滿，2003)

第二節 研究對象

本研究以臺大教師為訪談之對象，由於資料度用在國內屬於新興議題，大多教師對其瞭解不深，且考慮到臺大教師人數眾多，無法在本研究中涵蓋所有不同之學術領域，故訪談對象之挑選採取方便取樣，從有意願且容易接觸的教師開始訪談，再請受訪者推薦他們認為對此議題有興趣或有需求之教師，進行下一步的訪談。為了儘量涵蓋不同之學術領域，原則上求取異質的受訪者，且以受訪者本身會產出大量的研究資料者為佳，共訪談 12 位臺大之教師；訪談時間約在 40 至 80 分鐘左右不等，基於保密原則，以 A 至 L 的代號來代表本研究中的受訪者，12 位受訪者的所屬科系和訪談時間如表 3-1 所示。

表 3-1 受訪者代號以及所屬科系和訪談時間

代號	科系	訪談時間(分鐘)
A	人類學系	60
B	社會工作學系	48
C	地理環境資源學系	55
D		50
E	大氣科學系	82
F		79
G	地質學系	63
H	應用力學研究所	50
I	工程科學及海洋工程學系	41
J	生物環境系統工程學系	68
K	生化科技學系	77
L		49

第三節 研究流程

本研究主要分為文獻搜集與分析、研究執行和研究分析與結論共三階段，研究實施之步驟詳見圖 3-1

一、文獻搜集與分析階段

本研究先根據擬定之研究主題行文獻蒐集，由於國內並無相關文獻，因此主要透過 Library and Information Science Abstract(LISA)和 Google Scholar 等進行網路資源和期刊文獻之檢索；也蒐集相關的研究會資訊、會議論文等，以瞭解國外的研究趨勢及關注焦點，例如美國的 Association of Research Libraries(簡稱 ARL)不但持續關注資料度用、e-Research 等議題，並有專門的網路頁面整理相關的資源和文獻，也給予圖書館在落實方面的建議或實施步驟等，對本研究而言極具參考價值。除此之外，國外提供資料度用之大學和學術圖書館的網站也是重要之資訊來源，以瞭解資料度用落實的現況。進行文獻的分析和整理後，進而提出研究問題和研究目的，並以臺大教師為研究對象，建立研究之架構。

在文獻蒐集整理之階段，發現資料度用之落實可分為許多層面，有些國家是從國家的高度，進行政策的制定或是成立國家級典藏庫；另外則是在機構的角度，對機構內之研究資料首先進行整理，再蒐集研究資料並建立典藏庫。不同層級之資料度用發展的狀況都各有不同，而本研究著重在瞭解臺大教師對於資料度用之看法和需求，並期望能作為臺大圖書館未來發展資料度用之參考，因此在文獻的選擇上，以大學或學術圖書館所發展出的資料度用為主，找出發展較為全面之國外大學為案例，進行介紹與分析。同時蒐集、整理使用者研究之文獻，瞭解國外的研究者們對於資料度用之態度與需求，作為本研究完成後，研究成果分析與比較之依據。

二、研究執行階段

本研究以普渡大學 D2C2 所提供的 Data Curation Profiles Toolkit (D2C2, 2011)，以及牛津大學針對校內研究者進行使用者研究時所使用之訪談大綱 (Martinez-Uribe, 2007) 為基礎，將之翻譯並向臺大圖書館的學科館員們請益和討論後，擬定適合訪談臺大各領域教師之訪談大綱，主要的目的在於瞭解臺大教師們對資料度用目前之資料來源、保存方式，對於研究資料的分享和公開之想法及遭遇的障礙，以及對於資料度用服務的需求和期待。

訪談大綱完成後，先於 2011 年 7 月中旬訪談三位不同領域之臺大教師作為本研究之前測，在訪談過程中釐清每個题目的表達方式是否恰當，並聽取老師們對於訪談大綱的意見，進而修改題目呈現的方式、用語等；且由於原本的訪談大綱是以某單一研究計畫為出發點，深入瞭解其中的資料產出以及各個研究階段中對資料的處理或分享、公開等議題，其中有許多細節，花費的訪談時間也很長。前測的三位受訪者們認為目前國內並無資料度用的相關服務，並且此觀念屬於新興議題，可能沒辦法討論太多的細節，因此在前測結束後，除了簡化訪談題目，讓受訪者更容易瞭解之外，並希望以受訪者過去普遍性的經驗為討論的基礎，不再侷限於單一的研究計畫，讓受訪者能更自由的表達他們的想法和需求。附錄一和附錄二為本研究中使用的訪談說明及訪談大綱。

本研究於 2011 年 8 月開始，透過臺大圖書館學科館員的聯繫與陪同，初步訪談了 2 位臺大地理系之教師，並請老師們推薦會對此一議題有興趣或需求者，作為後續聯繫之對象；接下來的 10 位受訪者則於 2012 年 4 月 5 日至 5 月 23 日之間進行訪談，每位受訪者的訪談時間約為 40 分鐘至 80 分鐘左右不等。

三、研究分析與結論階段

本研究之訪談重點有四點，首先是調查受訪者們目前所使用的研究資料之來

源，以及如何組織、管理自己產出的研究資料，以瞭解研究資料之保存現況；第二點是研究資料的分享，瞭解受訪者們對於資料分享的經驗；第三點是資料公開，重點在於瞭解目前學術環境中，對資料的公開存在哪些障礙；最後則是受訪者們對於資料度用服務的期待和需求。訪談成果將根據上述四個面向進行分析，並和研究方法相似的國外使用者研究之結果相互比較。



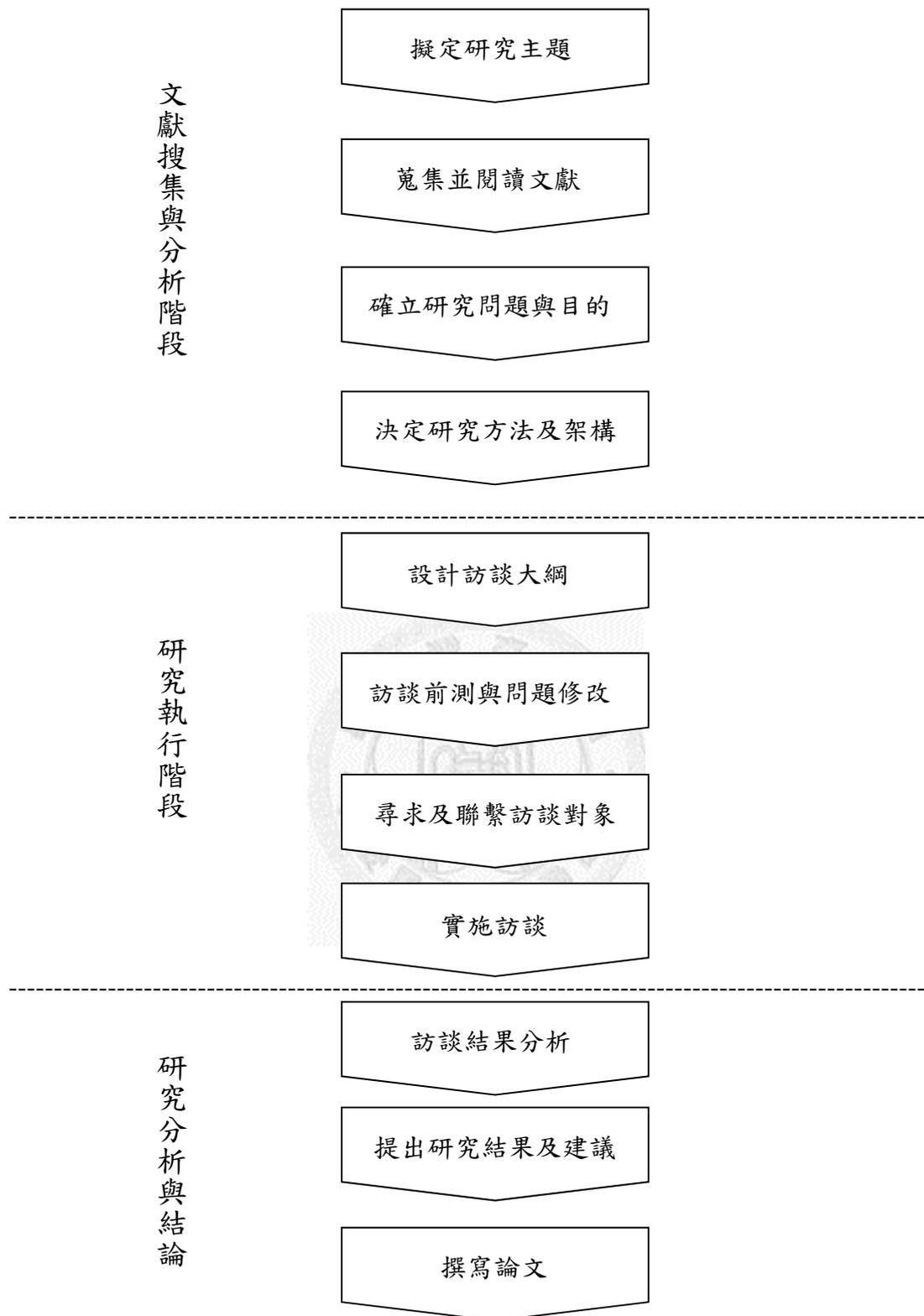


圖 3-1 研究實施步驟

第四章 研究結果分析

資料度用關心的是讓資料受到完善的保存與組織後，能夠提供檢索和取用，讓更多需要的使用者能夠持續的使用這些資料，以發揮其最大價值，並減少資源的浪費，因此發展資料度用服務的重點在於如何對研究資料進行管理、保存及再利用。由於資料度用在國內屬於新興議題，且目前國內並無由學術圖書館所提供之資料服務可供受訪者參考，而從訪談過程中亦可知臺大的教師們對資料度用的概念十分不熟悉，大部份受訪者都沒有聽過「資料度用」一詞。故本研究希望藉由訪談的過程，解釋資料度用的意涵和優點，並瞭解受訪者對於此一服務的看法和需求。

本研究透過訪談的方式，總共有 12 位臺大的教師接受訪談，並將訪談結果分成幾個對發展資料度用服務而言十分重要的面向來討論。第一節先簡單介紹受訪者之背景，並對他們的研究資料類型進行分析；第二節為資料的來源與保存，瞭解資料的特質和保存的現況；第三節為資料之分享，瞭解影響受訪者們分享資料意願的各種因素，第四節為資料之公開與所遭遇的障礙，瞭解在推動資料度用過程中可能面臨的困難，第五節為對臺大圖書館發展資料度用服務之期待和需求，可作為未來發展相關服務時的參考；第六節為綜合討論，總結上述五節並和文獻中相似的使用者調查進行比較。

第一節 受訪者背景與資料類型

一、受訪者背景

本研究共訪談 12 位受訪者，分別來自臺大的 6 個學院 9 個科系，但即使是同一科系，受訪者的研究領域亦各不相同。表 4-1 是本研究中 12 位受訪者的代

號以及他們所屬的科系與研究領域。

表 4-1 受訪者背景

代號	學院	科系	研究領域	職稱
A	文學院	人類學系	考古學	助理教授
B	社會科學院	社會工作學系	老人照顧、身障者就業	副教授
C	理學院	地理環境資源	移民研究	教授
D		學系	水文模式	助理教授
E		大氣科學系	氣候模擬	教授
F			大氣測計	副教授
G		地質學系	斷層調查與研究	教授
H	工學院	應用力學研究所	醫學工程	教授
I		工程科學及海洋工程學系	水下聲	教授
J	生物資源暨農學院	生物環境系統工程學系	環境永續發展、汙染源	教授
K	生命科學院	生化科技學系	食品營養	教授
L			生物化學	副教授

本研究中之受訪者以性別來看，女性有 3 位，男性則有 9 位；以職稱來看，以教授占最多數，共 7 位，副教授為 3 位，助理教授則為 2 位。以所屬單位來看，以理學院最多，共 5 位，工學院和生命科學院次之，各為 2 位，其他學院則各有 1 位受訪者。

二、資料類型

根據美國國家科學基金理事會(National Science Board, 簡稱 NSB)的分類, 研究資料依其產生的方式可分為觀察型、電腦運算型、實驗型和紀錄型等 4 大類 (NSB, 2005); 以下簡述 NSB 資料分類之定義, 並將本次的 12 位受訪者依照其所使用及產出的資料類型進行分析。

(一) 觀察型 (observational)

包括測量值和調查報告等, 大多具有特定時間、地點的特性, 屬於具有歷史性質且不可能重新蒐集的資料。(NSB, 2005)

受訪者 A 和 G 大多以實地考察的方式取得所需之研究資料, 其中包括儀器所取得的數據和現場觀測之紀錄、照片等, 屬於觀察型資料。受訪者 F、I 和 J 之資料來自於架設之儀器或感測器而得的數據或聲音檔, 亦屬於觀察型之資料。

(二) 電腦運算型 (computational)

來自於電腦模式或模擬之成果, 要重製這些資料需要有相同的硬體、軟體等條件, 很多時候只有模擬結果被保存。(NSB, 2005)

受訪者 D 和 E 之研究主要是透過電腦程式得出所需之數據, 其中亦會用到觀測資料作為基礎, 但仍以電腦運算型的資料為主。

(三) 實驗型 (experimental)

大多來自於實驗進行中所紀錄下來的一切資料, 雖然可以重新進行相同實驗, 卻常有時間、人力和經費上的考量, 因此仍具保存價值。(NSB, 2005)

受訪者 H、K 和 L 之研究是透過實驗, 因此其研究資料是以實驗數據和紀錄為主, 受訪者 H 也會有照片資料產出。

(四) 紀錄型 (records)

指關於政府、商業活動、社會現象等相關紀錄, 可應用於科學、社會科學、

歷史研究等各方面。(NSB, 2005)

受訪者 B、C 之研究資料來自於政府機關所保存的許多紀錄，或是自己的問卷調查，屬於紀錄型之資料。

根據上述分類，本研究中 12 位受訪者之資料類型分析如表 4-2 所示，雖然各類型的資料都有受訪者，但人數皆不多，因此較無法就資料類型歸納出具代表性之結論。

表 4-2 受訪者資料類型分析

資料類型	人數
觀察型 (observational)	5
電腦運算型 (computational)	2
實驗型 (experimental)	3
紀錄型 (records)	2

第二節 研究資料之來源與保存現況

受訪者的所使用或產出的研究資料，除了在類型上各有差異之外，資料本身的來源也各不相同，而不同的資料來源對後續研究資料的分享和公開都有所影響；另一方面，對研究資料進行資料度用的難易度，則和受訪者目前如何保存他們的研究資料息息相關，因此本節先就受訪者們的資料來源以及保存進行瞭解。

一、研究資料來源

根據訪談的結果可發現受訪者的資料來源主要有三種，分別是政府機關、自身產出和大型資料庫，顯示受訪者在研究中所使用的研究資料不一定是完全原創，常是從它處取得後，再進行分析和研究。

(一)、從政府機關取得

政府機關是很重要的研究資料來源，有 7 位受訪者提到他們所使用的資料來自於政府機關，包括氣象局、環保署、水利署、勞工局、健保局等等。而資料產生的方式則是透過量測或是調查，受訪者 D、E、F、G、J 和 I 使用的是政府機關透過儀器所取得的量測資料，以數據和圖片為主。

「一般來說資料來源是公家機構，……公家機構才會有長期的蒐集資料，學校的研究者不太可能有錢去做長期的蒐集資料，……像是台灣的水文資料，是水利署每天派人去蒐集的，……台灣幾十條河流，又要每天去，大概就要上千人，可以想像是多大一個工程，……要很龐大一筆錢去做，所以公家機關才可能去長期做。……像台灣有 6、7 個地震儀，要花錢買的，放的土地當然也是政府，這些資料是屬於氣象局的，……所以說這種長期的資料都是政府機構的。」(G:36-43)

受訪者 B 常會使用由政府單位進行調查後所產生之調查資料，例如勞動狀況調查、健康狀況調查等等，這些調查往往每隔幾年會做一次，因此資料量相當豐富而且龐大。受訪者認為這些資料品質和數量都相當好，唯一的缺點要花很多時

間在資料的索取，造成研究上時間的壓力。

「台灣的調查資料庫算是很不錯，以國際來講算是水準很好的，……有一定的規範，而且是有經驗的政務官在維持，有連貫性，……樣本數很大、變項數很多，……研究在作相關的統計或分析時，都覺得水準蠻好的。」(B：38-40)

(二)、自行產出研究資料

透過自己進行實驗、調查、量測來產出研究資料的受訪者則有 9 位，其中 5 位同時也會使用政府機關之資料，分別是受訪者 B、D、F、G 和 I。另外 4 位受訪者則以自己產出的研究資料為主。受訪者 A 的資料產生方式是透過親自實地考察和量測後，再進行記錄和整理；受訪者 C 的資料產生方式是透過訪談；受訪者 H 的資料產生方式是透過對血液樣本進行拍照，再對照片進行分析；受訪者 K 的資料產生方式則是透過實驗對各種食材進行檢驗並產生實驗數據。

「成都平原地區發現了一些城，大概 10 座，但我們覺得應該還有很多小型遺址，被埋在地底下，是肉眼看不到的，……所以我們用了很多比較科學，比較先進的方法，像是配合各種鑽探技術、磁力探測、雷達等等。……每隔 20 公尺鑽一個孔，每個孔都有一張數據的表格，一天就大概 100 張以上的表格。……其他也有照片阿、現場紀錄阿，很多不同資料。」(A：16-26)

移民研究議題範圍廣，這幾年的訪談紀錄 90% 都有打字並存進電腦……，我研究的區域主要是加拿大、澳洲、紐西蘭、美國等地的臺灣移民，……透過長時間的接觸和他們建立關係，並透過訪談或是問卷來瞭解他們的生活狀況。……這些資料當然不可能再次產出一樣的，……我認為很有價值，如果有人要拿這些資料去做分析，我也滿高興的。(C：32-35)

「我們是用實驗的方法，我們拿病人的檢體，去探討、去跟醫生討論，……像是針對癌細胞的轉移，……會從原來的位置跑到血管去，……會產生一些實驗資料，一些量測到的數據，像是血液中有多少癌細胞；還有一些是影像的資料，主要還是影像資料為主，是最基礎的，……譬如說一個圖檔，像是對血液樣本拍照，但要看解析度多高，還有範圍多大，……有時候影像資料很大，很占空間。」(H：6-11)

(三)、以大型資料庫中的資料進行研究和分析

受訪者 L 的資料來源是以大型的基因和蛋白質資料為主，由於分子生物學領域的特性是以先註冊該基因或該蛋白質結構者，其發現會受到學術界之認同，因此一旦有新的發現，研究者們都會主動上傳到幾個大型且知名的國際級資料庫中，同時也能使用資料庫中的資料進行分析和研究。

「我們所有的研究者除非是為了專利的目的，要不然我們都無條件的把我們的 data 丟進去，原因是你不丟進去別人就丟進去，你丟進去至少還保有起碼得 credit，……除非你覺得這些發現後面會有很有價值的產出，不然一定會先送進資料庫；……送進去後會產生一組號碼，…就是在資料庫中註冊了這資料。」 (L：12-16)

二、 研究資料保存

研究資料保存的問題可分為兩個層面，第一個是檔案的大小，一旦檔案太大或是成長非常快速，則在資料的傳輸和儲存上都會遭遇困難，而資料的格式亦會影響到檔案的大小。另一個問題是資料的整理，長期累積的資料若未經適當的整理，則在進行資料度用時，就必須花費更多人力和時間去統整這些研究資料，甚至需要具學科背景的專業人士的幫助。上述兩個因素皆對資料度用的難易度造成影響。

(一)、檔案大小與資料格式

最多受訪者提到的格式是數據，但受訪者們皆表示格式不會只有一種，圖片、表格、照片、筆記、聲音檔等都是會產出的資料格式。受訪者 H 的研究資料以照片為主，這些高解析度的照片會快速消耗硬碟空間；受訪者 I 則是以聲音檔為研究資料，由於原始的音檔可能錄製了很長的時間，因此檔案很大，但並非全部都有研究價值。

「我們的資料量很大，……看你錄多久阿，錄久了當然檔案很大，……有些原始檔真的大到恐怖，真的會想典藏嗎？……我是建議典藏一些有意義的聲音片段吧，一些原始檔中剪出來的 clip，例如有一些生態

相關的聲音。」 (I：15-18)

數據資料本身檔案並不大，因此有些受訪者不認為儲存空間對他們而言有困難，但也有例外的情況，受訪者 E 和 F 所使用的資料雖然是數據，但由於氣象觀測數據產生的頻率很高，因此檔案數量增加的速度十分驚人，對儲存空間的需求量很高。

「我在做的氣候研究，...就要用很多過去的觀測資料，還有很多過去的模式模擬出來的關於過去的氣候的資料，或是未來氣候的資料，...自己也會有模式去跑資料出來，所以真的要用到很龐大的資料，可能是幾十 TB。...所以現在對研究氣候的人來說，模擬出來的資料要儲存在哪裡真的是很大的問題。...像我們前陣子買了 100T 的硬碟，現在已經剩下 20 幾 T 而已了，應該也很快就會被填滿。...所以資料儲存真的變成一個很昂貴的議題，...大型計畫的投資說不定要到幾千萬。」 (E：2-13)

(二)、資料整理現況

受訪者都表示目前資料皆保存在自己的研究室或電腦中，通常不會主動刪除，有些會燒成光碟，或用硬碟進行備份，至於整理資料的工作則大多教給學生或助理去做。除了受訪者 L 表示會將基因或蛋白質等研究資料上傳至大型資料庫中之外，其它 11 位受訪者並無將研究資料公開或上傳至資料庫等經驗，皆是以自行保存為主。受訪者 K 指出資料長期累積以及整理方式不固定的結果就是資料十分零亂；而且整個教育體系中都沒有教導研究資料的重要性，學生對資料的整理自然沒有概念，整理方式因人而異，老師們也不一定有時間去確認整理的方式是否恰當，因此實驗室中的資料要從頭開始把它整理好真的很困難。

「因為我們一直沒有人力和資源去把資料的管理建立起來，.....另外就是學生的老習慣，因為都沒有特別的重視或去訓練，學生也不知道有資料管理的需求，所以大部份都是看學長姐怎麼做，他們就跟著怎麼做，所以這其中變異性就非常大，.....我也是到 5、6 年前才慢慢開始想把實驗室裡的資料規格化，.....還要花很多時間說服學生，.....因為長期以來，(學生)都覺得 result 比較重要，中間的 quality 不重要，所以這樣的基本訓練真的很不足。」 (K：112-121)

第三節 研究資料之分享

資料度用的目標之一是資料的再利用，要達成此一目標則必須透過資料的分享或公開，讓資料擁有者以外的人知道該資料的存在，或是取用所需的資料；故受訪者對資料分享的意願和經驗，皆會對資料度用的推廣有所影響。根據訪談結果，大部份的受訪者對研究資料分享抱持正面的態度，也表示有分享之經驗，而影響分享意願的因素則可歸納如下：

一、分享研究資料之對象以熟識者為主

大部份受訪者對於資料之分享是建立在一定程度的信任和瞭解的基礎上，對於完全不認識的人則持較為保留的態度；有 10 位受訪者表示若是認識的人、同一研究團隊或具名望的其他研究者向他們索取資料，他們較願意分享，但也有兩位受訪者仍然認為有更多其他需要考慮的問題，因此持保留態度。若是非學術界的其他人士，例如商業界，則全部受訪者皆表示較不願意分享他們的資料。

受訪者 A 近年參與過一個大型的考古學調查研究，由於範圍大且研究進行的時間長，因此常會有一些不同領域的研究團隊以短期參與的方式，加入此調查研究，因此受訪者 A 對資料分享抱持開放態度，認為曾經參與的團隊都可以獲得他們所需的資料，也樂於彼此分享，但同時亦指出這種情況對考古學界來說並非慣例。

「我們那時候其實採取了一個很 open 的態度，就是各國的人或團隊或學生，只要願意參與，我們就讓他們來，……所以參與者超過 10 個不同國家，單位的話更多，20、30 個左右不同單位或團隊都來，……地表調查大部份資料是比較開放的，……很多學生來參與，我們也都讓他們能拿資料去寫論文或發表。」 (A：24-35)

「我自己去用別人 row data 的情況就很少，若是被處理過的資料就更不會用，……考古學的東西有很多個人的思考脈絡在，別人的資料不見得有用，每個人有自己的解釋。……這個圈子裡也有些人對自己的

資料很保密，所以不太會給別人。」(A：102-105)

受訪者 H 表示若有其他研究者對研究資料有興趣，則可以和對方進行討論或意見的交換，但不一定會願意把原始的數據直接讓對方使用。

「如果是發表過的東西那沒關係，別人來要的話我們是樂意分享，.....因為譬如說期刊都會要求附上的圖有一定的解析度，所以(對方)應該也不會再想要更原始的圖，.....所以一般來說，我們的分享不是資料本身，而是背後的解讀、意義或是一些想法阿，.....可以討論一些對研究的看法。」(H：26-29)

受訪者 I 表示分享資料給自己不熟識的人帶有一定程度的風險，因為被分享者或許會將資料再分享給第三者，而第三者不一定知道資料真正的源頭，因此就算註明了資料出處，也是不正確的出處。

「我後來就覺得不是很好，.....因為這資料我給你，但又有別人要用，那這人是跟你拿，這樣第三者就不知道其實資料是我給的。.....所以我如果要給，就會講清楚是只給你用，不是讓你再給別人的，.....所以還是要熟悉的人比較有可能給吧。」(I：51-53)

若是不屬於學術界或政府單位，受訪者則表示不會將資料給對方。受訪者 F 都認為資料分享應限於學術研究，受訪者 E 則指出資料不應用於商業用途。

「資料的話主要還是針對研究社群，.....我覺得國內目前的普遍想法是不能去用在商業用途。」(E：100)

「不是學術界的人如果來問我，對資料有興趣，我會說 no，因為你不是學術圈的人，而且我跟你不熟，所以還是有防備之心，.....不知道你的動機是什麼，而且我對你沒有約束力，不知道你會把資料拿去做什麼」(F：79)

二、分享研究資料之時間點在研究發表後

除了受訪者 L 之外的 11 位受訪者皆認為分享研究資料的時間點應在研究完成並發表之後，主要是考慮到對自身學術研究進行的影響。受訪者 F 認為所分享的資料必須是在自己的研究已經完成之後，研究仍在進行中或認為這些資料還有

更多的研究價值時，並不想將其分享給其他人。

「學者間有跨國性的串聯，申請一個特別實驗的觀測計畫，……例如針對東沙島附近有非常密集的觀測和調查，……可能是一個月，create 出一筆很特殊的資料，……我們的認定是說好不容易得到的資料怎麼能隨便給別人呢。……所以我們在形成這樣的團隊時，大家會互相約定，團隊之內是彼此信任的，會分享資料，但團隊之外，頂多就是說明計畫內容，……為了保護大家投資的心血，所以會有默契，讓資料有 3 到 5 年的保護期，這段時間資料不能給別人，……大家要 sign 一個文件，要參與的話就要同意這個協議。」 (F：10-19)

受訪者 L 認為在分子生物學領域中，將蛋白質或基因之結構在資料庫中註冊後就等於是公開分享其研究成果，而且要越快越好，因此不一定會等到研究完成，而是會搶先上傳部份成果，但並不代表研究者們會毫無保留的分享其正在進行中的研究，還是會視情況而有所保留。

「希望有快速的取得和適當的保護，……這是人性啦，還是會想有點保留，……國外那些 data center 是要花很多錢去維護的，……但最後的結果是好的，因為讓資料是彼此競爭多於保護，……不過對我們個人而言，我們還是會希望有點保護，至少在某個期限之內。」 (L：42-45)

三、資料處理的程度會影響分享意願

資料產生後，可能需要先經過篩選和校正後才能使用，因此受訪者認為資料被處理的程度和資料的品質都會影響他們分享的意願，也是他們不願意分享給不認識的研究者的原因之一；尤其是自行產出觀察型資料和實驗型資料的 8 位受訪者，認為資料的品質對他們而言很重要，不希望因為分享品質不佳的資料而受到其他研究者質疑。受訪者 F 認為完全不經處理的資料其品質是尚待檢驗的，因此資料若要分享，一定要跟對方說清楚這些資料目前的處理狀況。

「資料收回來要考慮一些過程中雜訊、誤差的問題，……先想到的問題是這資料的品質我安不安心，在資料的品質還沒有保證的時候，給人家也只是害人家，……所以拿到資料後，都要花一段時間去確認資

料的品質，……不然資料錯誤的話，不管是自己用或別人用，都很糟糕。」(F：20-23)

「以作科學的人來說，我資料要給你的話，就要告訴你這資料處理到什麼程度，或是根本沒處理，……因為是你要拿去用的，所以自己要注意一點。」(F：62)

受訪者 J 則認為最原始的資料不一定最有價值，很可能其他人根本不知該如何使用，反而是將資料進行初步的整理或解讀之後，這些資料才會更容易被他人所理解和使用。

「比如說做國土規劃的人，他會想要污染資料，因為這對國土的規劃是很重要的，……但他要用的資料其實不能太 raw data，因為在專業上他也無法去作判斷，……這樣的人可能會來找我，我知道他們的需求後，我可以很快的提供給他一些真正有用的資料，……不一定是像學術研究的成果那樣精細的資料，……因為(污染)不是他的專業，而且資料太細或量很多，所以 raw data 對他不見得真的有幫助。」(J：55-61)

受訪者 I 的研究資料來自於台灣附近海域中感測器所錄下的聲音，原始檔案可能錄製了很長時間，檔案很大而且真正有意義的聲音不一定多，因此一定要經過過程式的篩選才能截取出有用的片段，因此受訪者 I 認為分享原始檔不一定有意義或價值。

「其實氣象局有幾個 sensor 是放在西海岸那邊，會錄下很多聲音，……但我覺得應該不會有人想聽原本的錄音吧，錄很長的話，會很多雜訊在裡面阿，可能也有很多空白或背景噪音，……其他人來用的話，或許聽不出什麼重點吧，我們自己用的資料也是需要很多過濾或處理，甚至寫程式去跑等等。」(I：42-44)

四、研究團隊的態度將決定資料分享的對象或程度

參與過研究計畫的受訪者皆表示資料的分享與否並非一人能決定，有 7 位受訪者有相關經驗，尤其是分享的對象和分享到什麼程度等細節，常要經過討論才能有共同的決定。受訪者 A 參與在大型的考古調查計畫當中，有好幾個計畫主

持人，因此即使本身對資料的分享抱持開放的態度，仍然需要和其他計畫主持人討論或告知，並非個人可以決定之事。

「考古團隊中的同學是美國人，……後來他去哈佛大學教書，……再找了很多其他人，北大的老師、成都考古所、我們以前的指導老師……，我們一群人從 2004 年到去年，在成都平原進行了很長時間的調查。……後來就發現一個很大的問題，因為我們單計畫主持人就 6、7 個單位，內部的資料分享就很辛苦。……基本上團隊內都很開放，如果是外面的人來要資料，那就會問一下其他人，……就算是禮貌性的也要告知一下。」 (A：21-33)

五、資料的所有權造成無法分享

部份受訪者認為資料是否分享給別人並非他們能夠決定，而是應該由出資或委託的單位來判斷，尤其是使用政府機關之資料和接受政府委託計畫的 7 位受訪者，皆認為分享應和資料來源或出資單位討論，甚至是由這些單位來決定。受訪者 B 就表示若有其他研究者向他們索取資料，則他們會請對方直接和出資或委託單位聯絡並取得同意，或是請他們直接向擁有該資料的單位索取。

「我們是受委託者，職訓局是委託者，……如果在契約的內容有說要保守相關的機密，當然在資料的使用上會很小心，……其他人要來索取這個資料的話，就請他直接去問委託單位，反正我最後會把資料都給委託單位，委託單位同意的話，如果在資料的使用上出了什麼問題，那就是委託單位自己負責；……那如果是我自己就同意給資料，我是受委託者，若是拿資料的人在使用上有什麼問題的話，就變成我要負責。」 (B：27-31)

第四節 研究資料之公開與所遭遇的障礙

資料的分享和公開存在著程度上的明顯差異，在訪談的過程中，幾乎所有的受訪者都表示他們有資料分享的經驗，對象和程度則因人而異；但若提到是否能將資料公開，讓更多人知道這些資料的存在，甚至是取用資料，則受訪者們的態

度都較為保留，原因也各有不同。顯示在目前的學術環境中，從資料的分享到公開之間存在著許多障礙，而這些障礙有可能造成資料度用在推動上的困難。以下就受訪者對研究資料公開的態度以及研究資料公開會遭遇的障礙等兩點進行分析與討論。

一、受訪者對研究資料公開的態度

(一)、認為期刊上公開發表之資料已足夠

大部份受訪者認為資料公開的程度是以期刊文章為準，因為發表文章時會附上相關數據來支持自己的研究成果，這些資料是公開且可被檢驗的，同時也是和其他研究者間交流並瞭解彼此研究的主要管道，透過期刊文章瞭解其他研究者是否擁有自己有興趣的資料，再進一步判斷是否要請對方分享資料。除了受訪者 L 由於領域中的慣例而主動公開基因或蛋白質等資料之外，其他受訪者皆認為目前資料公開的程度已經足以讓他們知道自己所需的資料要如何取得，對於是否進一步公開更多的資料則認為並非必要。

受訪者 H 表示在醫學工程領域，發表之期刊文章都會要求要附上一定程度的研究資料，包括數據、照片、圖表等，以供審核研究成果是否值得信任，因此這些研究資料等於是已公開之資訊，但沒有意願主動公開更多相關資料。

「(期刊)會要求充份揭露研究數據，因為你要佐證你的研究阿，.....所以 paper 被公開，當然這些數據也都被公開，.....不過真正放在期刊的大概是 10%左右的資料吧，是精挑細選過的。」 (H：40-46)

「如果發表過是沒有問題的，但沒發表的話，就有可能有些隱私的問題。.....資訊主要還是從期刊獲得，如果對資料有興趣的話，(我)就寫信問問那個研究者這樣，.....目前主要是透過這樣的私下交流吧，我覺得應該是足夠。」 (H：25-32)

(二)、擔心研究資料被誤用或對其理解不正確

本研究中有 6 位受訪者提到對於資料被誤用或理解錯誤的顧慮或經驗，因為資料一旦公開，誰會使用或怎麼使用都不易掌控，因此在仍有疑慮的情況下，他們並不贊成將資料公開。

受訪者 J 認為如何解讀資料很重要，若是將原始資料公開可能會被其他人有目的或不正確的解讀，尤其是非學術領域的人，若是片面的收集符合自己主張或支持自己論點的相關資料後，將其公開給大眾或媒體，很容易誤導一般民眾對某些議題的認知，因此認為資料公開的價值不一定完全是正面。

「我覺得比較糟糕的是先有立場在去找資料，.....因為資料都已經在那邊，那如果先有立場，再去刻意蒐集一些支持我論點，或是對我立場有利的資料，.....再刻意去做出一些結論，.....這是現在比較多發生的，也是比較困擾的一點。因為資料都在那邊，不管是不是政府資料，.....所以有些人可以去篩選對他有利的資料來支持他的結論，.....這是我比較擔心的問題。.....如果是一些比較激進的團體去做這樣的事，那很多人其實不知道這些資料被蒐集的背景。」(J：97-107)

受訪者 H 則是會擔心其他人對他們的資料不能理解或是解讀錯誤，造成學術研究上的問題。

「沒發表過的資料就丟出去的話，怕是人家解讀錯誤，.....因為我大部份是圖檔，那別人來看的話怕是有些誤判的情況，那這樣我是不是有些責任，心裡多少會覺得怪怪的。.....所以應該是我先發表，列出我對資料的解讀，那其他人就可以瞭解這個資料的意義，.....也能當作彼此討論的基礎，我覺得這樣比較恰當。」(H：33-38)

(三)、對現行體制缺乏信任感

本研究中有 4 位受訪者認為現行的體制並不重視研究資料，也不利於資料的公開，對於有升等壓力的研究者們而言，資料公開將是一種負擔。受訪者 K 認為資料的公開是有意義的，但目前的學術環境中長期缺乏對資料的重視，而國家也沒有對資料的分享或公開有制度化的規範，造成很多資料都只存在於研究者們

的實驗室中；另一方面，目前並沒有一個由國家經費長期支援的平台或系統，可供資料公開之用。若是沒有長遠的規劃和制度性的規範，研究者們不可能放心的交出研究資料。

「所以最大的問題是政府從來不管理資料，所以當然就沒有資料管理的系統，如果國科會有這樣的系統，我們當然就可以去用阿，……國科會沒有長遠的眼光，一開始就沒考慮到這些資料的價值，……他們的系統也一直在改阿，沒有前瞻性，經驗也都沒有傳承……國家每年都花錢出去，回來的就是幾頁報告，這樣就夠了嗎？……其實如果有一個系統在那邊，我們就會知道資料要怎麼去處理，不會每個人都用自己的 format。」 (K：32-40)

「我們非常需要的是制度性和規劃，因為私下的分享是一定有的，但要真的大規模的(資料)交流一定需要更好的規範，……讓大家都瞭解彼此的資料，才會把資料貢獻出來。」 (K：87-89)

受訪者 G 認為資料的品質應該要經過適度的審查，而審查也該由值得信任的單位或專家來進行，如果蒐集來的資料沒有審查，其品質就有疑慮，即使公開這些資料也可能沒人想使用，甚至是不敢用。

「即使蒐集了很多資料，也開放給大家用了，……問題是要如何品管你的資料，……有些人比較小心，會 check 資料的可信度，知道要怎麼用，可是學生的話可能搞不清楚，就直接用了。……圖書館能建立審查機制嗎？沒有審查的話是不是可信度就不夠，資料來源和處理過程是不是也該紀錄下來，……如果都沒有這些審查或紀錄的話，可能別人看到資料也不敢用。……資料蒐集單位應該要負起審查的責任，去檢視資料的品質，不然收一堆資料在那裡，沒人敢用的話，就白忙一場了吧。」 (G：22-26)

二、研究資料公開會遭遇的障礙

(一)、與出資或委託單位之間的協定

全部的受訪者皆有受委託而進行研究計畫的經驗，因此在計畫中所產生的資料是否能公開，需要考慮到出資單位或委託單位的立場和意願，因此每一個研究計畫所面臨的狀況皆不相同，需要一一檢視和討論，難有一體適用的情況。

受訪者 G 常受到政府單位的委託進行台灣的地質調查和斷層研究，每個委託單位會對資料有不同的要求，有些是簽訂了保密協定，而不能公開；有些單位雖然不會特別要求，但受訪者 G 仍然認為資料公開的程度應由委託單位決定，避免造成不必要的困擾。

「我們資料給他們(政府單位)後，他們將來要怎麼應用或開放到什麼程度，那是他們要決定，這有牽涉到行政問題，……開放也是有限，也不會無限制的全部開放；……像是地質敏感區的研究，這種資料要很小心，……受委託者的成果到底品質怎樣、資料正不正確，是政府單位要負責審查，確定沒問題了，才能公告出來。……如果沒審查，這資料可信度如何不知道，沒人能保證。」 (G：15-23)

受訪者 I 會接受軍方的委託進行水中聲音的相關研究，其資料屬於軍事機密，因此不能公開。

「如果你們要做這種典藏，……那軍方的資料我真的完全不能給，那個保密的規範很嚴格。……資料全都會給軍方，他們大概也不會開放吧，有些可能算是軍事機密，……或許像反潛的研究之類的。」 (I：28-29)

受訪者 L 認為有商業價值的資料往往會和廠商之間有協定，既不能分享也不能公開。

「我剛有說我們會無條件送進(資料庫)去，但其實也不見得，如果有商業價值的話，我們就考慮不送出去了。……像是水稻的研究計畫，很多資料就沒有那麼公開，……因為它大部份雖然是拿政府的錢，但也有種子公司的錢，……研究結果的商業價值很高。……就算我有參與這種計畫，拿到相關的研究資料，應該也不行公開。」 (L：19-22)

(二)、資料來源對研究資料之公開有所規範

主要使用政府機關資料進行研究的 7 位受訪者皆提到資料來源將對資料能否公開有很大的影響。臺灣許多政府單位都有很多資料可供研究使用，但條件或限制各不相同，例如氣象局的資料需要付費取得，因此不適合免費公開；而受訪者 E 所使用的資料部份來自水利署的雨量或水文資料，這些資料水利署明確的要求不能將資料分享給其他人。

「比如說我們拿到水利署或農委會的雨量資料，可是它就只准許我們自己使用，不准許我們再給別人，.....我們可以拿來做研究，研究成果可以發表，但原始資料不能公開。.....(政府)可能也是比較傳統的或保護的心態，覺得資料為什麼要給人，.....有些(資料)則是要收費，行政上有些相關規定等等。」 (E：15-20)

受訪者 A 所使用的資料雖以自行產出為主，但由於所參與的考古調查計畫位於中國四川省，而中國政府對地圖的保密十分嚴格，因此和中國地圖相關的資料或是將其應用於 GIS 的資料都無法公開。

「其實我們一開始吃了很大的虧，因為中國對他們的地圖管制非常嚴格。.....單單為了想要一個比較精密的地圖，就花好大心力，.....也不只是要錢，很多時候錢也買不到，要到處找關係，.....但我們後來還是都數位化，因為想整合在 GIS 裡。.....所以 GIS 資料會比較有問題，因為我們拿到的管道也是比較偏，而且中國那邊的態度是不開放，.....所以不是我們不想給，而是不能給。」 (A：14-19；90-93)

(三)、研究倫理的限制

研究倫理的限制是資料無法公開的原因之一，尤其是研究對象是人的時候，其規範相當嚴格，本研究中有 3 位受訪者有相關經驗。受訪者 B 接受政府單位委託進行跟身障者職業相關的調查研究時，常會碰到研究倫理的問題，如果調查對象很特定，那麼就算是匿名，根據資料中的許多特徵，還是有可能得知其身份，因此有個資和倫理的考量所以不能公開或分享。

「雖然我們都會匿名啦，但有的時候如果調查的對象是某特定機構，或是某種特殊的身心障礙的話，.....即使名字、身份證字號都不寫，但因為人數很少，範圍小，.....如果真的有心想查的話，對象是誰可說是呼之欲出。.....因為現在個資法去年又修了，現在更嚴了，.....我們自己要資料其實有時也很痛苦，.....大家(研究者)因此都很謹慎，怕會出什麼問題，.....像是資料的外洩或身份曝光等等。」 (B：30-37)

受訪者 C 透過問卷和訪談的方式來了解移民適應當地環境或就業等問題，因此個資的部份造成其資料無法公開。受訪者 C 認為他的資料都必須先非常小心的整理和匿名處理後才有可能分享或公開。

「訪談、問卷之類的，.....有些受訪者我都認識.....，是因為信任我才會跟我談，.....所以這些研究都屬於質性研究，一定要考慮到倫理問題，很多資料都需要 double check，把敏感性的資訊拿掉才可以放出來，.....像是姓名或外號什麼的，有可能讓別人猜出受訪者是誰.....，因為很多受訪者都是要匿名才願意接受訪談的，而且每個研究都有它的脈絡，所以訪談資料若要經過重新整理，那我覺的整理的過程中，研究者需要在場，一一確認內容或敏感資訊等等.....，這會是一個很大的工程，很辛苦。」 (C：52-60)

受訪者 H 的研究資料來自病人的血液樣本之照片或數據，醫學相關的研究資料受到倫理委員會的嚴格規範和保護，同時需徵詢患者的意願和取得同意，故無法公開。

「會有一個 IRB，是一個倫理審查，那是很嚴謹的，醫院阿、病人的檢體阿，都要經過這個倫理審查的申請，.....就是病人的檢體要經過一定的規範，衍生出來的資料當然也要符合倫理審查的規定。」 (H：36-39)

「有時候想要某些病人的資料，並人可能就不願意給你，.....有時候想要多一點血液樣本，有些人會不想給阿，.....所以一定都要經過病人和醫院的同意，不同意的話，我們也是跟他們說謝謝，就沒辦法啦。」 (H：87-89)

(四)、智慧財產權、著作權等法律問題

法律問題對所有受訪者而言都並非他們的專長，對相關法令的瞭解也不深，自然會有所顧慮和擔憂。使用政府機關資料的受訪者認為自己並無資料的所有權，因此任意公開可能會觸法；而接受委託研究計畫的受訪者則可能簽下保密協定或相關條約，而無法公開研究資料，因此對所有受訪者而言，法律問題都是資料公開一定會遭遇的障礙。

受訪者 F 認為資料的公開會牽涉到智慧財產權或是其它法律問題，由於這部份並非自身的專長，因此不敢貿然公開，而且資料一旦公開則其它人要如何使用就無法掌握，也不知道是否會產生其他問題，因此公開的意願不高。

「智慧財產權的問題，……像是我簽了一個保密協定或是跟智慧財產權有關的協定，……我自己的瞭解是這個資料我是能拿來做其他研究，但不能原封不動的把成果或資料有給別的委託單位，……我目前是沒受到規範是說不能再分析或再利用，但這是我個人遇到的狀況和理解。我覺得一般人對於會不會觸犯法律或是違反智慧財產權其實不清楚，尤其我又不是唸法律的，真的不知道，…所以確實是有顧慮，不知道會不會違法，如果哪天被告了要怎麼辦。」 (F：68-79)

受訪者 D 指出資料經過多重加值後，著作權的歸屬問題目前並沒有被清楚規範，自己也不是很瞭解，因此對這樣的資料公開會有疑慮。

「有些研究的地圖來源是我向政府機構或 GPS 地圖公司購得的，我針對這些地圖去做加值、分析，這個東西可以聲稱是我自己的嗎？……譬如說前一個人把內政部的數值資料辛苦的轉成圖檔，我只是將五都改制後縣市的邊界改掉而已，就聲稱這是我的圖，這樣是合理的嗎？」 (D：78-83)

第五節 對臺大圖書館發展資料度用服務之期待與需求

由於大部份受訪者並無公開資料的經驗，臺大目前亦無資料度用相關之服務或平台，因此本研究在訪談過程中，先假設臺大圖書館未來會發展資料度用服務，並以此為基礎和受訪者們進行討論，希望瞭解受訪者對資料度用服務的期待和需求，並將之整理成對臺大圖書館的期望以及對資料度用平台的需求兩個面向。

一、對臺大圖書館的期望

(一)、清楚界定資料收錄之範圍

本研究中有 4 位受訪者對資料收錄之範圍提出疑問或建議，有些是認為並非所有資料都有保存的價值，有些則是認為資料處理之程度和品質需要考慮，才能讓收錄的資料能被使用。受訪者 D 表示自己使用的部份資料是自行架設感測器或觀測站而得，通常是在氣象局或水利署等單位沒有架觀測站的地方，且為了特

定的計畫或研究來進行資料的收集，所以往往是小範圍且限於特定某段時間的片段資料，故是否會有其他人需要這樣的資料，或是單純因為觀測資料的不可重製性而保存，諸如此類資料收錄範圍和基準的問題，是需要首先釐清的。

「我有架一些 sensor 去偵測某些地方的雨量，這些都是中央氣象局沒有設置觀測站的地方.....但是它資料的時間是片段的.....就是研究計畫結束，它就結束了。」 (D：16-20)

受訪者 G 認為不同的資料所需的處理和校正不同，而原始資料和處理過的資料何者較有價值，則依資料特性或使用者而有所不同，因此每筆資料的收錄都要詳加考慮。

「很多實驗室一出來的資料其實還不能用，需要經過一些修正和加工，才會變成研究中能用的 row data。.....那你們(圖書館)到底要收集哪一塊的資料，是一開始最原始的嗎？還是經過一些初步處理的呢？.....最原始的數據有很多校正要去做，所以要真正的專家才有辦法用，.....一般使用者應該還是有處理過的數據比較有用，.....但是不同領域差異很大，有些人是要用最原始的數據，有些人是喜歡處理過的，用起來比較方便。」 (G：60-72)

(二)、訓練具學科專業背景的人協助資料整理

本研究中有 3 位受訪者希望圖書館能培養專門的人才來協助資料的整合與保存，若能具學科知識或參與在研究過程中則更佳。受訪者 K 認為實驗過程中所使用的控制組是一種很適合公開的研究數據，因為研究對照的基準點本來就是應該被人所知的，這樣不同研究單位的研究結果才能互相比較，互相瞭解對方的研究結果是基於什麼樣的基準點而得；而且控制組只是一個基準，不會因為公開而對研究有所影響。不過由於每次研究中所使用的控制組都不見得相同，不同研究者、研究單位也都有自己的控制組，所以要能夠整理、歸納這些數據，需要同時具有食品科學和圖書資訊背景之專業人士，參與在整個研究過程中，才能對資料的本質和如何整理有深入了解，進而真正讓這些資料顯出價值。

「像我想到，最需要互通的是 control group，是基準的那一組，.....這才是實驗室之間能夠互相比較的基準。.....圖書館如果能蒐集到很

多這種 control group 的資料，那將來就很容易知道和比較說這個研究領域中，大家都用什麼變項，這個變項的來源是什麼，這對我們做實驗的人來講，就可以知道很多訊息，像是彼此誤差的來源等等。(K：9-13)

「我覺得你們的館員或是圖資系學生的訓練，就可以開始來跟我們搭配，瞭解研究室中有哪些資料，從頭開始釐清，慢慢去建立一個 model 出來，……建立跨領域的合作，彼此瞭解，這樣這個服務才會可行。」(K：39-41)

(三)、政策和經費之支援

本研究中有 3 位受訪者認為明確的政策、規範，以及實質上的經費支援才是現階段最需要的，也是推動資料度用最快速、最有效的方法。受訪者 D 提到自己曾經在中央研究院協助建製台灣生物多樣性資訊入口網，系統建製之初需要聯絡該領域的研究者們，請他們提供自己所擁有的資料，讓系統建製人員匯整並上傳，研究者們也都很樂意協助；不過一旦請他們開始整理相關資料並交出時，大多數人卻表示研究都來不及做了，哪有多的時間去整理以前的資料，造成效率低落，系統中的資料量也不足。系統建製團隊因此和出資單位進行協調，將該計畫的部份經費撥給研究者們，讓他們有義務整理並交出手中的研究資料給該系統。

「我們那時候就是開放計畫申請，就是問那些老師願不願意，就是老師手上的資料要不要趁這個機會，等於就是我們付老師錢，然後老師就是趁機整理整理，然後給我們上傳。」(D：94-95)

受訪者 K 認為由政府出資或委託的研究，其產出的資料應視為公共財，研究者們其實是有義務要好好整理這些資料，並讓這些資料成為未來研究的基礎。然而目前的學術環境中缺乏對資料的重視，很多資料都是由研究者自行保管，就算有繳交出資或委託的政府單位，政府單位也不見得會重視這些資料，或是讓這些資料能被其中的相關研究使用，造成資源的浪費，很多有價值的資料也因此被埋沒。

「我認為很公平的一件事，既然是國家經費支持的(研究)，也不應該想什麼回饋或不回饋的，……這應該是一個對等的，拿了國家的錢，

這些資料就放到合理的資料庫，讓別人能應用。……那如果你資料放不出來，這資料的品質是你自己就沒信心嘛，不能讓人檢驗，……資料的分享代表其實可以讓人檢驗的。……後來很多美國的期刊都要求你把資料提供出來，讓其他讀者是可以去檢驗的，……這樣的要求受到大家的認同，……好的東西本來就會有人用，那用你的資料就會引用你阿，等於你的研究就能不斷被引用，就是你的 credit 了。」(K：25-33)

二、對資料度用平台的需求

(一)、提供簡單易懂的介面，節省上傳時間

由於國內目前並無資料度用系統或介面能讓受訪者參考，因此對介面提出實際建議的受訪者較少；在介面設計上，有 2 位受訪者希望能以簡單易懂為原則，避免花費很多時間去學習，造成負擔。受訪者 F 認為如果真的要發展資料度用，並鼓勵研究者們自行上傳資料，則介面一定要簡單易懂並容易使用，因為研究者們並不會想為了上傳或分享資料而多花時間去整理自己的研究資料，如果還要花時間去熟悉介面和系統的話，意願就更低了，因此最好是將介面和上傳的流程都能簡化，讓研究者很容易就能分批上傳，不需要花太多精力去整理資料。

「這樣的平台我很歡迎，但讓我有點疑慮的是我到底要花多少時間搬東西上去，然後好不好搬，……舉例來說，學校常叫我們填一些東西，有時候欄位很多，真的很頭痛。……另外像是大氣和其他領域的資料怎麼整合，有沒有一個全部都適用的格式在，這格式好懂嗎？……對我來說，關鍵點是 upload 越容易、越簡單，當然越容易達成。」(F：100-106)

(二)、讓資料提供者瞭解資料使用的情況，並建立有彈性的授權機制

部份受訪者希望知道自己的資料被什麼人使用，或是下載次數等資料的使用情況，因此認為資料度用平台應提供一些相關的統計數據，對他們而言也是一種回饋，本研究中有 3 位受訪者提出希望能瞭解資料使用狀況之需求。

「定期的告訴我有多少人下載資料，回報一下使用狀況等等，……或許將來可以設計下載次數多的，就給他多一點(儲存)空間，當作回

饋；……甚至是在教學研究或是升等上面有些加分，有一點實質的鼓勵，……這樣才會有誘因。」(J：146-150)

受訪者 F 表示資料度用平台應該要有兼具彈性和自動化的授權機制，最好是讓想要下載資料的使用者先表明自己的身份和資料的用途後，透過系統自動發送授權申請至資料的提供者端，讓資料提供者勾選是否授權，系統便能判斷能否開放給該使用者下載資料。同時兼具自動化和彈性的授權機制，可以讓研究者自行判斷是否讓對方使用資料，又不用花太多時間回覆電子郵件，省下許多麻煩。

「NASA 的全球觀測計畫……這個系統的好處是每個觀測點是誰負責都有公開，這個觀測點的 PI 是誰，e-mail 是什麼都看得到，……資料是可以 download 的，但你點下去時它會跳出一個視窗說，這資料是誰的，應該先徵詢同意再 download 或使用。……又或是比如說有其他研究者或學生想要資料，只要簡單描述一下是什麼系的或要做什麼研究，再透過這系統讓我看到，……或是跳出一個選項來，我只要回答好或不好，勾選之後，對方就能開始下載，……對我來說，等於是多一層保護。」(F：109-119)

(三)、以政府機關所擁有的資料作為資料整合的基礎

雖然只有 2 位受訪者建議在資料的收錄方面，以政府的資料為基礎，但在本研究中，使用政府機關資料的受訪者多達 7 位，因此若是將政府資料整理的更好，應能對研究者有所幫助。

受訪者 G 認為目前的學術環境中並無資料公開的習慣，因此若要針對每個研究者都去問有沒有意願或資料能夠提供其實很困難，僅有事倍功半之效，所以應該先由目前擁有最多資料之處下手，也就是從政府單位開始進行初步的接觸和合作，例如透過提供儲存空間、專業人力、協助資料的規劃和管理等服務，取得和政府機關合作之契機，將其所擁有的資料備份到台大。未來有人需要資料時，將可透過台大的專門窗口為其提供取用的服務；對政府單位來說，可以省下信件往來、應對、審核、篩選、提供資料等許多流程和時間，一旦這樣的合作模式建立起來，雙方也能夠彼此信任之後，就能進一步討論資料的開放等議題，而個別

的研究者也較會有意願將自己的資料上傳到這個已運行一段時間且值得信任的平台中。

「總圖去面對政府機構，就是提供儲存空間，不公開沒關係，就簽個協定什麼的，由總圖去幫忙審核，看哪些人可以申請，……這是總圖的資產，或許幾十年後這資料可以公開了，那對研究會很有幫助。……所以不要從小的、分散的開始做，因為很複雜，格式都不一樣，不知道要怎樣整合，品質你也沒辦法保證或審核，……倒是例如說很多政府單位還有很多資料是紙本，圖書館就去找他們談，幫他們電子化和保存，……先做了再去談說能不能開放或是開放給誰。……其實也是幫政府單位的忙，幫它們整理資料，因為它們自己也處理不好，還有很多其它業務，也沒空去好好弄；……一旦時間久了，資料也累積起來，就是重要資產，或許就能開放了，……有很多資料在那邊，其他個別的研究者才有可能把資料丟進去，你們(圖書館)也跟研究者們比較好談，畢竟已經有很多資料為基礎。」 (G：105-120)

第六節 綜合討論

本節將上述研究成果進行統整，並歸納出發展資料度用服務的 5 個重點；最後則和國外相似的使用者研究進行比較。

一、資料來源及收錄範圍是資料度用需優先思考的議題

從研究結果可以發現資料來源牽涉的範圍相當廣泛，首先是關係到資料的所有權，本研究中大部份受訪者所使用的資料都並非完全由自己的研究中產出，而是會使用到由政府單位或是其它大型資料庫所提供的資料；這代表他們只有資料的使用權而非所有權，因此資料的分享或公開與否，並不能由受訪者全權決定。因此臺大圖書館若是希望能收錄校內的研究資料，就必須先瞭解每個研究者的資料之來源，以及資料來源對資料度用的態度或立場；尤其是政府機關，常常會對資料有不同的限制，若一個研究中用到許多不同政府機關的資料，要如何處理這些不同政府機關對資料的要求或限制，將會是一個很大的問題，也將是資料度用服務必須面對的障礙之一。

資料來源的另一個影響層面則是收錄範圍，很多資料或許已經存在於其他資料庫中，或是由政府單位所擁有，因此這些資料是否仍需由臺大圖書館收錄是必須考慮清楚的；如果完全不收錄這些資料，則根據這些資料所衍生出的研究資料是否能收錄，又或者是否會因此而使得整個研究的呈現不完整，進而造成對資料的理解不正確，都是資料來源對收錄範圍的影響。除此之外，根據本研究中受訪者的經驗，很多政府機關所擁有的資料都未開放，可能需要花時間或金錢，甚至有時需要人脈才能取得，因此從資料度用的觀點來看，這些資料雖然有被保存，但在資料的開放程度上明顯不足，是否能達到再利用或驗證研究之效果也不得而知，因此雖然政府單位是資料的擁有者，資料也已經被留存，但資料度用是否能讓這些資料更公開，或許也是在思考收錄範圍時的一個重要方向。

收錄範圍也必須考慮到資料本身的特質，例如本研究中的受訪者提及資料品質及資料處理程度的問題，究竟是收錄最原始產出的資料比較合適，還是收錄已經初步處理過的資料比較方便使用，可能會因學科領域以及資料的使用者而有所差異；然而本研究的受訪者分散於不同學科領域中，因此無法就特定領域的收錄範圍做出結論，僅能看出不同領域的受訪者對資料的收錄範圍的確有不同意見，是臺大圖書館在推動服務前需要更深入瞭解的。

二、資料的分享多屬私下交流，對資料公開的態度則較為保留

從本研究中受訪者的經驗可發現目前學術環境中的資料分享是以私下交流為主，且大部份的受訪者都認為資料分享的對象必須是自己熟識的人或是研究團隊中的其他研究者，並不希望自己的資料被學術圈以外的人使用，而分享的時間點通常是在研究完成後。大多數受訪者都認同期刊文章是很重要的交流管道，公開於期刊上的研究資料足以讓他們瞭解別人的研究，若是對資料有進一步的需求，則會採用私下聯絡的方式，因此進一步的公開研究資料似乎多此一舉。這代表受訪者目前能夠接受資料公開的程度就是以已經發表過的研究為主，因為這些研究已經完成，資料公開與否並不會傷害他們身為學術研究者的利益；可看出在思考資料公開此議題時，主要是以保護自身的學術研究為出發點，而並非考慮到資料公開對整體學術研究的益處，因此不希望將還有研究價值的資料公開。

除了學術上的保護之外，還有許多造成受訪者們態度保守的外部因素，例如研究團隊中其他人的立場，讓資料的分享或公開有時並非受訪者可自己決定。另一方面則是研究的出資或委託單位，不一定樂於看到相關的研究資料被分享或公開，例如受訪者提到有商業價值的資料，或是軍事機密等級的資料等等。即使出資或委託單位沒有明確的規定資料能否分享或公開，部份受訪者由於不希望引起爭議，因此傾向將決定權交給出資或委託單位。從這些外部因素可看出資料度用的複雜性，必須要考慮到每筆資料都有其差異，並不是說某研究者的資料或是某領域的資料就一定適合資料度用這麼單純。

三、資料誤用、研究倫理以及法律問題為國內外研究者的共同擔憂

資料誤用的問題根據本研究的訪談結果，大致可分為兩個層面，首先是對資料的理解，若對資料的理解不正確，就會造成學術研究上的問題，受訪者認為這或許是資料公開的負面影響。另一個層面則是不能讓資料被使用在錯誤的地方，例如有些受訪者認為不能用於商業用途，或是被用於誤導大眾等。上述兩個資料誤用的面向亦在普渡大學的使用者調查中被提及，普渡大學的研究中還有受訪者提到錯誤的引用也是資料誤用的形式之一。資料誤用的問題受到國內外研究者們的普遍重視，也是多數研究者不願分享或公開資料的主要障礙。資料誤用的形式多樣化而且牽涉的層面廣泛，因為資料一旦公開，其如何使用或被誰使用就會變得難以控管，而不同的研究者願意承受的風險程度也不相同，因此沒有單一的解決方案可適用於所有情境。

研究倫理是因為有些資料關係到個資或隱私權等敏感議題，因此本研究中的受訪者認為公開並不恰當，而且國內的個資法有相關規範，資料的公開可能會為他們的研究帶來困擾。英國DCC的個案研究中，社會科學領域的研究者們提出他們對資料公開或分享的擔憂主要是來自於對個資的保護，尤其當中有醫療相關資料時，其取用的權限更是需要嚴格的控管。法律問題包括著作權、智慧財產權等，本研究的受訪者認為法律領域並非自己的專長，因此並不瞭解資料的公開是否會觸犯相關法令，因此不敢貿然公開研究資料。澳洲的DART研究專案中，受訪的研究者們也對著作權、智慧財產權等問題有所疑慮，是造成他們不願公開資料的原因之一。牽涉到研究倫理的資料或許並不適合資料度用，可在界定收錄範圍時

優先排除；而法律問題則可由提供資料度用服務的單位主動幫助研究者們解決，例如由臺大圖書館提供協助，釐清資料的著作權或智慧財產權，這也代表需要具備法律專業的人力參與其中。

四、提供資料度用服務之單位必需具備良好的基礎實力

雖然國內目前並無機構層級的資料度用服務，因此需要具備的資金或人力等尚無法明確評估，但從本研究中受訪者對資料度用服務的期待和需求中可發現，提供此服務的機構或單位是否具有一定程度的基礎實力非常重要。先從資料度用平台的方面來看，受訪者希望介面的設計簡單易懂，同時要有彈性的權限控管機制，這代表需要專業的人力來進行平台設計，也要有足夠的資源來支持平台的開發；而資料檔案大小會因人而有極大差異，資料格式也不一，因此硬體設備也必須相當完善，不然無法處理大量的資料；同時受訪者希望能有專業的人來協助他們整理資料，最好是有學科背景並不時參與研究過程，才能對資料有深入的瞭解。上述條件都必須在有足夠的經費和人力等條件下才有可能達成，因此提供資料度用服務的單位一定要有良好的基礎實力做後盾，才有辦法面對許多的資料提供者不同的需求，並提供合適的服務。

另一方面，本研究中的受訪者認為目前的學術環境對資料不夠重視，長期累積的結果就是在整個體制或研究流程上皆不利於資料的公開，而且目前並沒有建立值得受訪者們信任的平台，在不瞭解平台的運作或資料審核機制等細節的情況下，受訪者對資料的公開缺乏信心，因此提供資料度用服務之單位必需具公信力，並和研究者之間建立良好的互動與信任，對他們的資料有一定程度的瞭解，才能得到研究者們的信任，並放心的將資料公開於此平台上。臺大圖書館長期累積的經驗和人力在國內大學中都是首屈一指，而臺大本身的經費亦十分充足，因此由臺大圖書館來提供校內的資料度用服務應是可行之道，也可作為其它大學或機構發展資料度用之參考。

五、政策的制定和經費的補助將是資料度用能否成功的關鍵

從本研究的訪談結果可發現受訪者認為現行的體制和學術環境並不利於資

料度用的推廣，一方面缺乏相關的政策或規範，導致整個研究過程中，都沒有重視資料的整理；而在申請研究計畫時，國科會也並未對資料有更多的要求，結果的呈現僅以結案報告為主，不一定會要求提交研究過程中的資料，因此很多寶貴的資料都被並未開放給其他人使用。即使是有心整理並公開資料，目前也沒有共通的規範可供依循，或是可供使用的平台，導致大部份的資料都存在研究者自己或工作用的電腦中。本研究的受訪者認為資料度用需要有政策的支持，讓整個研究環境有體質上的轉變，才能真正讓研究過程中的資料受到重視；同時亦指出，由國家經費產出的資料原本就應該回饋給所有的研究者，而非一人獨占；這樣的想法和 OECD 提出的「公帑贊助研究資料之近用原則與基準」不謀而合，許多國家也是基於此而對研究資料的公開有相關規定。美國國科會近年來亦對申請計畫經費的研究者有所要求，必須對研究過程中的資料提出 1 至 2 頁的「研究資料管理計畫」，作為計畫書審核的項目之一。政策的強制性將是推動資料度用不可或缺的動力，也是成功的關鍵。

六、本研究與相關研究之比較

由於本研究採用訪談法，故研究成果比較的對象也以同樣採用訪談法的美國普渡大學之研究、英國牛津大學之研究以及澳洲 DART 研究專案等三個使用者調查為主；國外使用者研究之成果主要參考第二章的表 2-2，再配合本研究之成果的 4 個分析面向進行整理，以利於和本研究成果進行比較。

美國普渡大學之研究訪談了 20 位研究者，英國牛津大學訪談了 37 為研究者，而澳洲 DART 研究專案則訪談 15 位研究者；相較之下，本研究共 12 位受訪者，樣本數較小。在研究成果方面，本研究和上述 3 個使用者研究的共同重點在於呈現受訪者間的共通點，而對於領域差異則較少提及或分析。根據本研究之成果分析，歸納為資料的來源與保存、資料之分享、資料之公開與所遭遇的障礙以及對資料度用服務之期待和需求等 4 個面向，和國外使用者研究之成果皆整理歸納如

表 4-3 所示。

表 4-3 本研究與國外使用者調查研究之分析比較

國外使用者調查之面向	本研究之訪談結果	國外使用者調查的研究結果
資料的來源與保存	<ol style="list-style-type: none"> 1. 研究資料並非完全自身產出 2. 資料格式和檔案大小因人而異 3. 儲存於個人或工作用電腦為主，資料整理工作大多教給學生或助理 	<ol style="list-style-type: none"> 1. 所擁有的資料格式多樣化 2. 資料儲存處以個人或工作用電腦為主 3. 資料整理、歸檔以個人經驗或使用方便為主
研究資料之分享	<ol style="list-style-type: none"> 1. 對象以熟識的其他研究者為主 2. 時間點在研究完成後 3. 需考慮資料品質或處理程度 4. 研究團隊或資料所有權等因素，使資料的分享與否有時並非個人能決定 	<ol style="list-style-type: none"> 1. 有助於研究的驗證以及資料的再利用等學術上的幫助 2. 是研究領域中的慣例 3. 個人研究上的需要，對象通常為熟識的其他研究者或研究團隊內容的分享 4. 出資單位或所處機構的要求
資料之公開與所遭遇的障礙	<ol style="list-style-type: none"> 1. 以期刊為主要的資料公開管道 2. 擔心資料誤用且對現行體制缺乏信心 3. 研究倫理、法律問題 4. 出資/委託單位或資料來源對資料公開的限制 	<ol style="list-style-type: none"> 1. 資料誤用或誤解 2. 著作權或智慧財產權等法律問題 3. 經費或人力不足 4. 硬體、軟體等科技層面之因素
對資料度用服務的期待和需求	<ol style="list-style-type: none"> 1. 清楚界定資料收錄之範圍 2. 訓練具學科專業背景的人協助資料整理 3. 政策和經費之支援 4. 平台的設計應簡單易懂，並有彈性的授權機制 5. 以政府機關所擁有的資料作為資料整合的基礎 	<ol style="list-style-type: none"> 1. 值得信任的資料度用平台 2. 在資料的取用或整理上有立即回饋 3. 資料的安全性 4. 先針對有需求的領域提供服務

從歸納整理之結果可看出國內外研究者在資料的保存上最具共通點，皆是以個人或工作用電腦為主要的儲存資料處，而資料的格式也都相當多樣化；而本研究之結果發現資料的整理大多交給學生或助理，是國外使用者研究未提及。

在資料的分享方面，國外使用者調查提出一些正面觀點，認同資料的分享或公開有助於學術研究的進步，而本研究之受訪者則比較多考慮到對自身研究的影響，因此對資料的公開持較為保守的態度。資料公開的障礙則有較多的共同點，例如資料誤用、隱私權、著作權等法律相關問題，都是國內外研究者們共同的擔憂，也是資料度用在推動過程中的一大挑戰。

最後在對資料度用的需求方面，國外研究者重視資料的安全性，也提出領域的差異，認為應從有需求的領域開始進行；而本研究的受訪者則對介面的設計提出建議，並認為應從擁有大量資料的政府單位作為收錄資料的基礎。共通點應在於提供資料度用的單位必須具備一定的實力和公信力，以保障資料的安全並開發設計適合的平台，同時有足夠的資金與人力為後盾，瞭解特定領域的需要，或是和政府單位交涉協調，並順利推動資料度用服務。

第五章 結論與建議

資料度用包含了針對研究資料進行管理、保存及再利用之概念，強調的是在資料受到完善的保存與組織後，能夠提供檢索和取用，讓更多需要的使用者能夠持續的使用這些資料，以發揮其最大價值，並減少資源的浪費。資料度用的發展將能為 e-Research 下產生大量的研究資料找到合適的解決之道，是近年來相當受重視的議題。資料度用中的資料指的就是研究資料，不同領域中對於研究資料的範圍會有不同之界定，且目前國際間尚無單一而完整之定義。本研究依據資料度用的意義，界定研究資料為研究過程中所使用和產出的資料，不包含硬體設備和實驗樣本，且以數位化之資料為主。

資料度用的重要性可從資訊瓶頸以及黑暗資料兩個觀點來看。資訊瓶頸是指研究被發表的階段是資料量最濃縮的階段，過程中許多研究資料都並未被公開，因此也無法用於驗證研究結果，或是被其他人使用。黑暗資料則是指研究環境中，由於經費或人力不足等因素，而未被適當保存和整理的研究資料，故對研究者以外的人而言，這些資料是近乎隱形而無法發揮再利用的價值。為了透過資料度用解決目前研究資料未公開或無法被取用的問題，國外漸漸有學術圖書館將研究資料納入典藏範圍，例如英國愛丁堡大學的 Edinburgh DataShare 以及普渡大學的 Purdue University Research Repository，都是學術圖書館參與其中並推出資料度用相關服務的實例。本研究希望透過對臺大教師的訪談，深入瞭解他們對於資料度用所代表的資料分享和公開等概念，抱持什麼樣的態度和看法，調查他們對於資料度用服務的期待和需求，以及目前學術環境中，是否存在不利於資料度用推廣的障礙，以期能夠作為臺大圖書館未來推動資料度用服務的參考。

本章主要根據文獻探討以及訪談結果，歸納出本研究的結論，並對不同層級的資料度用以及未來研究方向作出建議。

第一節 結論

根據本研究的文獻探討以及訪談結果，可歸納出下列 5 點結論。

一、各國在資料度用之發展程度以英國和美國較具代表性，可作為國內學術圖書館發展資料度用服務之參考

資料度用對許多國家而言都是新興議題，因此各個國家發展的程度和重點都有所不同，本研究著重於大學層級的資料度用狀況，因此以歐洲和美加兩個地區為重點，再進一步聚焦到英國和美國，探討這兩個國家目前發展資料度用之狀況，最後則找出具代表性的大學，以瞭解大學層級的機構如何發展資料度用。

相對於歐洲其他國家，英國很早就對資料度用之發展相當關注，故成立了相應的國家級機構「數位度用中心(DCC)」，此機構由聯合資訊系統委員會(Joint Information Systems Committee，簡稱 JISC)所贊助，是一獨立運作的研究型機構，希望藉由國家級機構的成立來面對資料度用所帶來的挑戰。DCC從2004年成立至今，以每三年為一階段，一步一步達成其階段性任務，除了探討目前學術環境中資料度用所遭遇的問題之外，也深入了解學術社群和研究者們對資料度用的看法與需求，並且不斷擴展其夥伴關係，讓整個學術體系在資料度用的推廣上連成一氣，各自扮演適當的角色。(DCC, 2011)

愛丁堡大學在英國的資料度用發展過程中一直扮演十分核心的角色，並在JISC的資金補助下，成立了提供全國高等教育相關機構典藏研究資料的資料中心「EDINA」，目標在於促進英國各大學及機構之研究產出，提高教學和學習之效率，並以此大方向為前提，進行各種研究計畫；同時EDINA也提供一個整合研究資料的平台，提供多媒體資料和參考資源、資料和地圖等相關研究資源的取用。除了提供全國性服務的EDINA之外，愛丁堡亦針對校內研究者們提供多樣化的服務，不單提供研究者們各種關於研究資料的諮詢，協助其取得所需之資源，也給予研究者關於資料管理方面的建議，讓他們更有效率的整理自己的研究資料；並推廣資料度用之概念，強調其優點以及對研究的幫助，提高研究者們對資料度用的認同。若研究者們希望分享或公開自己的研究資料，則提供Edinburgh

DataShare作為平台，讓研究者們能透過此一平台進行資料的分享和取用。(The University of Edinburgh, 2011)

美國政府成立了DATA.GOV網站，讓使用者能夠方便且迅速的檢索、下載和使用政府單位所產出的資料；除了以身作則，開放各部會所擁有的資料之外，也針對研究計畫制定了相關規範，在美國國家科學基金會於2010年5月所頒布的新政策中，規定所有於2011年1月18日後到期或申請之計畫書都應包含1至2頁的「研究資料管理計畫」(data management plan)，清楚描述研究過程中所產生的資料，其類型、格式及長期保存的規劃，若是資料有牽涉到智慧財產權、隱私權或保密協定等法律問題，也應註明清楚(NSF, 2010)。在政府本身資料開放和新政策的制定之下，美國許多大學在資料度用方面都有快速的發展。

康乃爾大學和普渡大學皆推出了針對校內研究者的資料度用服務和保存研究資料之平台。康乃爾大學透過整合校內現有的服務來幫助研究者們規劃並撰寫其研究資料管理計畫(RDMSG Planning Group, 2010)；同時建立了DataStaR平台，作為研究過程中，資料臨時的存放處，讓研究者們能透過此一平台進行研究資料之分享，並在研究過程中就有效的逐步整理和上傳資料，研究完成後便能快速的將資料併入合適的典藏庫或資料庫中(DataStaR, 2011)。普渡大學則透過訪談的方式，深入了解整個研究周期中，研究資料如何被產生、處理、保存、使用、分享等等與資料度用息息相關之議題，且在完成訪談後建立了一套資料度用檔案(Data Curation Profile)作為調查與瞭解研究資料現況之工具(Witt, Carlson, Scott Brandt, & Cragin, 2009)；並於2011年推出了Purdue University Research Repository(簡稱PURR)之平台，以成為校內產出之研究資料的機構典藏為目標，提供線上的資料管理與儲存服務(PURR, 2011)。

愛丁堡大學、康乃爾大學和普渡大學在發展資料度用服務的過程中，學術圖書館都參與其中，不論是作為服務提供的單位，或扮演規劃服務內容、發展相關系統的角色，都顯示出圖書館在典藏與管理各種出版品的相關經驗，對資料度用之發展有所幫助；因此對國內的學術圖書館而言，發展資料度用服務應能作為未來的目標之一。

二、研究資料之來源多樣化，整理方式不一

根據訪談的結果可發現受訪者的資料來源主要有三種，分別是政府機關、自身產出和大型資料庫，顯示受訪者在研究中所使用的研究資料不一定是完全原創，常是從它處取得，這代表在考慮到資料的分享、公開等資料度用相關議題時，受訪者們不一定握有對自己研究資料的完全主控權；即使是自身產出的研究資料，也要考慮到出資或委託單位對於資料的要求，受訪者不見得能夠決定是否分享或公開。因此資料來源會影響資料的所有權，同時也牽涉到出資和委託單位，顯示資料度用不只要考慮到研究者本身的意願，也應了解其出資或委託單位的立場，以及其他資料來源本身所造成之限制。

資料保存的問題主要分為兩個層面，第一個是檔案的大小，若是資料量非常龐大，或是增加的速度極為快速，則提供資料度用服務的單位在硬體設備上就要有充足的準備，才能讓資料提供者能夠放心將資料儲存在系統中；同時也要考慮到資料的傳輸，如果網路速度不夠快，就可能在取用上造成困難。本研究中有幾位受訪者提到資料量龐大是他們面臨的困境之一，因此若資料度用服務能為這個困境找到出口，或許會成為讓受訪者願意採用此服務的契機。另一個問題則是資料的整理，從訪談中可發現受訪者所擁有的研究資料主要還是根據當時的研究計畫進行分類，而且常常是由學生和助理去管理這些資料；然而資料整理的方式因人而異，一旦碰到學生畢業或助理離職等情況，整理或歸檔的方式就可能隨之改變，因此這些資料雖然被保存，但其整理之方式卻常缺乏一致性，而受訪者也表示沒有時間一一去確認資料被整理的方式是否恰當。這個現象顯示出目前研究環境中研究資料的重要性並未被彰顯，研究者們忙著進行新研究，而沒有時間、人力去整理舊有的研究資料，只能任其埋沒在研究室中。資料長久的累積和整理方式缺乏一致性都會造成資料度用的困難，這些情況會讓研究者們需要花費更多的時間或人力去整理並判斷該資料是否適合進行資料度用，額外的付出對研究者而言是一種負擔，因此提供資料度用服務之單位必須要有能力幫助研究者解決資料

整理等問題，這也表示需要具備一定程度的學科知識。

三、研究資料之分享以不影響個人研究為優先考量

對於資料的分享，所有受訪者都表示有資料分享的經驗，而影響分享意願的因素主要有 5 個，首先是對分享對象的瞭解或熟識程度，受訪者表示如果是同一個研究團隊，則資料的分享當然不成問題，若是認識的其他研究者，也大多會分享給對方，但有時則視情況而定，例如資料的出資者若要求保密，那就無法分享；但也有受訪者表示即使是認識的人也不想分享，因為不知道他會不會再將資料分享給別人。如果是不熟的人，則他們大多傾向不分享，更不用說是非學術界的其他人士。第 2 個因素是分享的時間點，受訪者都認為還在研究中或是仍有研究價值的資料他們不會想分享，畢竟目前學術環境中仍是以所發表的研究成果，來給與研究者評價，因此辛苦得到的資料一定要發揮最大價值，必須等到研究告一段落，才有可能分享給其他研究者。

第 3 個因素是資料的處理程度，受訪者認為大部分的原始的資料仍應要有人為的確認，才能用於研究，所以最原始的資料不見得適合分享，也不見得真正對其他研究者有所幫助，這表示資料在分享前都要經過一定程度的處理，當受訪者對其資料品質放心時，才有可能分享給其他人。第 4 個因素是研究團隊的態度，由於很多研究是由不同單位甚至不同領域合作，因此資料的分享與否也必須要跟團隊中的其他人討論，並取得共識，非一人可以決定。最後的因素是資料的所有權，由於很多受訪者會使用政府機關之資料，故受訪者認為這些資料的所有權並不在他們，而是原本的資料來源，資料分享與否的決定權自然也在於原本的資料來源。

從上述的 5 個因素可以看出受訪者在面對資料分享議題時，仍是以自身的研究為主要考量點，而較少想到資料的分享對整體學術環境所帶來之影響和進步；

因此推動資料服務的首要條件應是讓研究者們瞭解資料度用的優點，並積極解決目前在資料整理和分享上所遇到的問題，才能讓研究者們有動力參與其中。

四、對研究資料公開之態度較不積極

研究資料的公開在目前的學術環境中並非常態，因此必須瞭解其中的原因和存在的障礙，唯有解決這些障礙，才有可能順利推動資料度用服務。從訪談中可以發現幾乎所有受訪者對於資料的公開有許多顧慮，又可分為受訪者自身的態度和資料公開會遭遇的障礙等兩個面向。

受訪者面對資料公開的態度整體而言是較為保守的，相對於資料分享，有資料公開經驗的受訪者很少，大多數的受訪者都認為最好的資料公開的管道是期刊文章，因為所附上的數據或圖表等資料已足以支持他們的研究；若是對該研究有興趣，也可以透過主動聯絡的方式來取得進一步的研究資料，因此公開更多資料對受訪者而言似乎是多此一舉。資料誤用的問題也常被提及，大致可分為兩個層面，首先是對資料的理解，若對資料的理解不正確，或許會造成學術研究上的問題，受訪者認為這或許是資料公開的負面影響。因此對提供資料度用之單位而言，研究資料的整理、註解或後設資料的設計等非常重要，除了應有清楚的規範之外，也要得到原本資料提供者的認可，才能讓想要使用資料的人能清楚瞭解資料的意涵。另一個層面則是不能讓資料被使用在錯誤的地方，例如有些受訪者認為不能用於商業用途，或是被用於誤導大眾等，這方面就比較難以處理，即使在資料的取用上有嚴格控管，也不一定能夠完全避免這些情況發生。在現行的制度下公開研究資料也讓受訪者很不放心，若是沒有一個穩定且具公信力的機構來建立資料度用的平台，並制定統一的規範來整合各種來源之資料，則受訪者們仍對資料的公開缺乏信心，沒辦法安心交出他們寶貴的研究資料。

除了受訪者本身態度保守之外，很多外在環境的限制也會造成資料公開的障

礙，包括資料來源或出資、委託單位對資料的限制、研究倫理、法律問題等等，這些因素讓受訪者覺得資料公開與否並非與他們的個人意願有關，因此也較不會去進一步思考資料公開的優點或是對整體研究環境的幫助。這是提供資料度用服務之單位必須面對的困境，也是應優先解決的問題；例如透過和出資或委託單位的主動協調，達到對於資料度用的共識。著作權或智慧財產權等法律問題，也可以由供資料度用服務之單位協助釐清；至於牽涉到研究倫理等敏感議題的資料，則可在訂定資料收錄範圍時就先排除，避免引起不必要的困擾。

五、期待資料度用服務之發展能有實質回饋，並對研究有所幫助

受訪者所提出的期待與需求大多和上述提及目前學術環境中資料度用的障礙有所關聯，又可分為對臺大圖書館的期望以及對資料度用平台的需求兩個面向來討論。由於臺大圖書館目前並未有資料度用服務，因此本研究的訪談過程中是先假設臺大圖書館希望發展資料度用服務，並以此為基礎和受訪者們進行討論；因此部份受訪者明確的對臺大圖書館做出建議，包括清楚界定資料收錄之範圍、訓練具學科專業背景的人協助資料整理以及政策和經費之支援等三點。

資料收錄之範圍對臺大圖書館而言應是最根本的問題，必須考慮的重點首先是資料來源，因為有些資料可能已經被收錄在政府單位或其他資料庫中，那麼這些資料該不該由臺大圖書館再重新收錄，或是部份收錄以增加該研究的完整性等問題，都是值得思考的。另一方面，牽涉到智慧財產權、著作權、個資、隱私權等敏感議題的資料是否適合收錄，也應在制定收錄範圍時一併考慮，或是由臺大圖書館主動協調並取得授權，或是不收錄這些資料，都需要清楚規劃。

部份受訪者指出研究資公開前應該要有適當的整理，但目前的經費和人力都有限，因此希望能由同時具備圖書資訊專業以及學科背景的人來協助，最好是能參與研究過程，進而深入瞭解資料，才能真正將資料匯整，並讓其能夠被他人使用；而目前臺大圖書館並無這樣人力，可以給予研究者們研究資料整理上的幫助，

因此要發展資料度用服務也需要考慮到目前人員的狀況，可能需要更多的相關訓練來為此服務打下基礎。

政策和經費的支援對於能否在目前的學術環境中順利的推動資料度用十分關鍵，政策所代表的是強制性，表示研究者有義務整理他們的研究資料，也必須要適度的公開這些資料讓更多人能使用；經費則是給研究者們一個整理資料的動力，讓他們在研究的過程中，能夠重視資料的整理與保存，同時也不能再以經費或人力不足為藉口。若臺大圖書館希望能推動資料度用服務，可能需要臺大從政策面去改變才是最有效的，例如將研究資料的分享或公開也作為考核的標準之一。另一方面，由於許多研究者的研究經費來自於國家，因此若能在國家層級就針對資料制定政策，例如在申請計畫的同時就要求計畫書中必須對資料的整理、保存與公開有所規劃，將能快速提升研究者們對資料的重視，並且主動整理研究資料，以作為未來資料度用的基礎。

對資料度用平台的需求則包括提供簡單易懂的介面、瞭解資料使用的情況，並建立有彈性的授權機制以及以政府機關所擁有的資料作為資料整合的基礎等三點，可作為未來臺大圖書館在系統設計方面的參考。由於很多受訪者都會使用到政府機關的資料或是接受政府單位的委託，因此對受訪者而言，與其讓他們將資料整理並提供出來，先將政府機關中的資料進行整合，同時提供更方便的取用，才對受訪者而言有真正立即的好處。一旦政府的資料完成整合後，對提供資料度用之單位而言，也有更多的籌碼能和想取用這些資料的研究者們討論研究完成後之資料提供等問題；同時這也代表提供資料度用之單位有能力去處理大量的資料，所發展的平台也有一定的成熟度，研究者們提供資料的意願才會隨之上升。

第二節 建議

本節根據研究結論，分別針對學術圖書館、研究型大學以及政府單位在發展資料度用服務或推廣資料度用等方面提出建議。

一、對學術圖書館發展資料度用服務的建議

(一)、培養資料度用服務的專業人才

圖書館過去在整理各種出版品的經驗對資料度用服務的發展是有利的基礎，由其是近年來越來越多的電子出版品或資料庫，增加了圖書館面對數位化資料的機會，對資料度用之發展也有了硬體或技術上的基礎；但更進一步的專業人才培育仍是不可或缺，尤其是具備法律知識或學科背景的人才，都將在資料度用服務中扮演重要角色。唯有這些專業的人才，才能協助解決智慧財產權、著作權等法律相關問題，並深入瞭解不同學科的研究資料，發展出合適的整理、註解方式，降低研究者們對資料公開的擔心和疑慮，並真正讓資料能被更多人使用。

以臺灣大學為例，臺大的機構典藏亦由圖書館提供，和校內研究者間已建立聯絡管道，因此在後續的資料度用方面可利用現有之資源和人脈，和研究者有更多互動，以得到他們的信任和認同；在此良好的基礎上，臺大圖書館若能在團隊中廣納更多具備不同專長的人才，相信能提供更完整的服務。

(二)、深入瞭解不同領域的資料特性，提供客製化服務

本研究雖然涵蓋不同領域，但也造成每個領域中的受訪者人數不多，不足以代表該領域的普遍特質；尤其在訪談過程中發現即使是同一科系的教師，也會因研究專長的差異，而對資料度用有截然不同的看法和需求。因此領域的差異將是臺大圖書館在發展資料度用過程中不可忽視的重要基礎資訊；不但應該要有針對校內教師的大規模調查，以瞭解他們普遍對資料度用的認知和相關需求，更需要

從中找出值得特別關注的領域，例如有資料分享需求卻缺乏相關資料庫或平台的領域，可從這些領域開始進行初步的接觸，深入瞭解他們的資料特性和需求後，設計雛型系統供其使用，並得到更多的建議和回饋，亦可作為發展其他領域之資料度用時的參考。

二、對研究型大學推動資料度用的建議

(一)、加強宣導資料度用概念，發展相關訓練課程

資料度用屬於新興議題，尤其在本研究中可發現所有的受訪者對資料度用一詞都感到陌生，表示確實有針對此議題加強宣導的必要；除了提高校內教師對資料度用的認知外，也應強調其優點和對整體學術進步的幫助，讓教師們在思考資料的公開與分享時，不再單單以個人的研究利益為出發點，而是考量到資料對大環境的幫助，讓資料成為學術研究中的基礎建設，進而強化研究的深度與廣度，學術研究的進步才會更快更好。

建議研究型大學在面對資料度用時，不應單只視為圖書館或少數單位的責任，而是應從整體的觀點進行規劃，例如參考康乃爾大學的作法，整合校內現有的單位，建立提供服務的單一窗口，一次性的幫教師們解決研究相關的問題。大學也可規劃相關的訓練課程，培育專業人才，甚至是發展相關的學程，讓學生們也能對資料度用有初步的概念；由於很多教師的資料都交給助理或學生整理，相關的訓練課程或學程規劃將是資料度用成功推動的重要基礎。

(二)、給予實質的鼓勵或回饋

對校內教師們而言，整理舊有的資料並確認其是否能公開將是大工程，因此適當的鼓勵和回饋是落實資料度用的重要動力，光只有觀念上的改變並不能促使教師們將之付諸實行，唯有在制度上給予回饋才是讓他們有所行動的誘因。建議研究型大學可將研究資料作為升等或評鑑的項目之一，而不是單單只看所發表的

期刊文章，因為一筆好的資料也能對學術有很大的貢獻，故應將其列入評鑑的考量。對於有些資料量很大的教師來說，儲存空間或許會造成經費上的壓力，大學可透過提供儲存空間的方式鼓勵其公開資料，越多公開就能得到越多空間，相信對教師們而言也是不錯的誘因。

除了正面的給予鼓勵和回饋之外，亦可在制度面上進行改善，例如編列屬於研究資料的專有預算，讓各系所有義務去整理現有的資料，並對其公開或分享的可行性進行評估，一旦有固定的預算和制度，就不能再以沒有人力或經費為理由，逃避整理或公開資料的責任。

三、對政府單位推動資料度用的建議

(一)、政策的制定

美國國科會在其申請研究經費的計畫書中，增加了「研究資料管理計畫」此一項目後，許多大學都開始針對資料度用提供相關服務，不管是協助研究者撰寫研究資料管理計畫，或是建立實質的資料度用平台，都有快速的進展，顯示出國家政策的制定是資料度用發展的重要關鍵。政策所代表的是強制性，隨著政策的制度，在整個資料度用的發展上才会有更整體的眼光和規劃，畢竟國家層級的高度和機構明顯不同，應該要更多考慮到資料度用對整個學術環境的幫助，並有全盤性的規劃。

建議國科會應抓住資料度用議題方興未艾的時機，明確的對未來的研究經費申請做出新的規範，讓所有申請者一律平等的在計畫書中闡述會產出的研究資料，以及他們對資料管理的規劃。對所有的學術機構或研究者而言，政策是最明確的基準，許多資料度用面對的障礙，也能夠因著政策的存在而有判斷的標準和基礎；故唯有政策的制定，才是全面推廣資料度用的成功之道。

(二)、經費的支援

經費和政策是一體兩面的，單有政策雖具強制性，但現有的人力或資源不足等問題仍然存在，如同本研究中部份受訪者所提及，資料檔案大小會是一個經濟上的負擔，故經費的支援是必需的。學術體系中應規劃出固定的經費來支持研究資料的保存和公開，尤其對很多機構而言，或許真的沒有足夠的資金去進行機構層級的資料度用，若是國家能有相關預算的規劃，就能評估不同學科領域或機構的需求，建立國家級的資料中心，幫助研究資料的保存、管理和公開。有經費的支援代表也能有人力去對資料進行長期的維護和審核，以確保其品質，建立真鄭有價值且可長可久的資料度用服務。

第三節 未來研究建議

一、以量化方式針對臺大教師進行全面的問卷調查

由於資料度用議題在國內並不普遍，因此本研究選擇訪談法，以期在研究中直接和受訪者解釋資料度用的意涵，但也因此使得受訪者人數不足以代表整體學術環境中的狀況，故建議未來之研究可參考本研究之結果，制作問卷並進行大規模的問卷調查，較能得知臺灣大學的教師們普遍對於資料度用的看法和需求，並以分析統計的方式，掌握整體的狀況或比較不同學科間的差異，將對推動校內的資料度用有所幫助。

二、對特定學科領域之教師分別進行相關研究

本研究雖已包含不同學院和科系，但限於受訪者人數，並不足以代表該學院或該科系的整體狀況，且在研究進行的過程中，亦發現即使是相同科系的教師，也會因為研究領域的不同，而對資料度用有不同的看法和需求。建議未來的研究可針對臺灣大學中的各科系進行更深入的訪談，並透過足夠的訪談人數，找出同

科系教師們關於資料度用的相同點和相異點，也可在不同科系之間進行比較，將有助於發展出真正符合需求的資料度用服務。

三、開發資料度用之雛型平台以進行好用性評估

由於臺大圖書館目前並無資料度用的平台和相關服務，故本研究中是以假設未來會有相關服務為基礎，和受訪者進行討論，故很多細節或許無法在本研究中被發掘。建議未來可根據本研究之結果，設計出一個資料度用平台之雛型，並以此為討論的基礎，讓臺大教師能實際的使用，從中得到回饋，瞭解在平台的設計有何進一步的需求，並評估授權機制、資料上傳、表格填寫等細節，以發展出合適且好用的系統，作為資料度用之基礎。



參考文獻

潘淑滿(2003)。質性研究：理論與應用。臺北市，心理。

Beagrie, N. (2006). Digital curation for science, digital libraries, and individuals. *The International Journal of Digital Curation*, 1(1).

Choudhury, S. (2009). Rethinking Scholarly Communication: Building Data Curation Infrastructure. Retrieved March 19, 2012, from <http://www.it.utah.edu/leadership/research/ciday/2009/notes/choudhury.pdf>

City of Vancouver (2011). Retrieved Dec. 20, 2011, from <http://data.vancouver.ca/>

Cragin, H. M., Palmer, L. C., Carlson, R. J., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Phil. Trans. R. Soc. A* 368(1926), 4023–4038

Data Conservancy (2011). Retrieved Dec. 20, 2011, from <http://dataconservancy.org/>

DATA.GOV (2011). Retrieved Dec. 20, 2011, from <http://www.data.gov/>

Data Library and consultancy (2011). Retrieved Dec. 20, 2011, from <http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/consultancy>

DataONE (2011). Retrieved Dec. 20, 2011, from <https://www.dataone.org/>

DataStaR (2011). Retrieved Dec. 20, 2011, from <http://datastar.mannlib.cornell.edu/>

Denison, T., Kethers, S. , & McPhee, N. (2007). Managing the Soft Issues in E-Research: A Role for Libraries? *Australian Academic & Research Libraries* 38(1).

Digital Curation Centre, DCC (2010). Retrieved Sep. 26, 2011, from

<http://www.dcc.ac.uk/>

Distributed Data Curation Center, D2C2 (2011). Retrieved Sep. 26, 2011, from

<http://d2c2.lib.purdue.edu/>

EC (2010). *Riding the Wave: How Europe can gain from the rising tide of scientific data*. European Union: e High level Expert Group on Scientific Data. Retrieved Sep. 26, 2011, from

<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

EDINA(2011). Retrieved Dec. 20, 2011, from <http://edina.ac.uk/>

Edinburgh DataShare (2011). Retrieved Dec. 20, 2011, from

<http://datashare.is.ed.ac.uk/>

Goetz, T. (2007). Freeing the dark data of failed scientific experiment. *Wired Magazine*, 15(10). Retrieved Sep. 26, 2011, from

http://www.wired.com/science/discoveries/magazine/15-10/st_essay.

Gold, A. (2010). *Data Curation and Libraries: Short-Term Developments, Long-Term Prospects*. Retrieved Sep. 26, 2011, from

http://digitalcommons.calpoly.edu/lib_dean/27/

Harvey, R. (2010). *Digital curation : a how-to-do-it manual*. New York : Neal-Schuman Publishers Inc.

Heidorn, P. B. (2008). Shedding light on the dark data in the long tail of science. *Libr. Trends* 57, 280–299.

Jankowski, N. (2009). The Contours and Challenges of e-Research. In Nicholas Jankowski (Ed.). *e-Research: Transformation in scholarly practice*, pp.3-31. New York: Routledge.

Key Perspectives. (2010). *Data dimensions: disciplinary differences in research data sharing, reuse and long term viability*. SCARP Synthesis Study. Digital Curation Centre, Retrieved Dec. 20, 2011, from <http://www.dcc.ac.uk/scarp>

Kuula, A., & Borg, S. (2008). *Open access to and reuse of research data: the state of the art in Finland*. Finland: Finnish Social Science Data Archive (FSD)

Lord, P. & Macdonald, A., Lyon, L., & Giaretta, D. (2004). From data deluge to data curation. Proceedings of UK e-Science All Hands Meeting, August 31-September 3, 2004, Nottingham.

Lord, P., Macdonald, A. (2003). *Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision*. UK: The Digital Archiving Consultancy Limited

Martinez-Uribe, L. (2009). Digital Repository Services for Managing Research Data: What Do Oxford Researchers Need? *IASSIST Quarterly* 31, 3(4).

Nanaimo Data Catalogue (2011). Retrieved Dec. 20, 2011, from <http://data.nanaimo.ca/>

National Institutes of Health, NIH (2011). Retrieved Sep. 26, 2011, from <http://www.nih.gov/>

National Science Board, NSB. (2005). *Long-Lived Digital Data Collections*. Retrieved Sep. 26, 2011, from <http://www.nsf.gov/pubs/2005/nsb0540/>

National Science Foundation, NSF. (2011). Proposal Preparation Instructions. Retrieved Sep. 26, 2011, from http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp#dmp

Organization for Economic Co-operation and Development, OECD. (2007). OECD

Principles and Guidelines for Access to Research Data from Public Funding.

Retrieved Sep. 26, 2011, from <http://www.oecd.org>

PARSE. (2009). PARSE. Insight: INSIGHT into Issues of Permanent Access to the Records of Science in Europe. Retrieved Sep. 26, 2011, from <http://www.parse-insight.eu>

Purdue University Research Repository, PURR (2011). Retrieved Oct. 24, 2011, from <http://research.hub.purdue.edu/>

RDMSG Planning Group (2010). *Meeting Funders' Data Policies: Blueprint for a Research Data Management Service Group (RDMSG)*. USA: Cornell University

Research Data Management Service Group (2010). Retrieved Dec. 20, 2011, from <https://confluence.cornell.edu/display/rdmsgweb/Home>

Research Information Network, RIN. (2008). *To share or not to share: Publication and quality assurance of research data outputs*. Retrieved Sep. 26, 2011, from <http://www.rin.ac.uk/>

Research data management guidance (2011). Retrieved Dec. 20, 2011, from <http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/research-data-mgmt>

Rice, R. & Haywood, J. (2010). Research data management initiatives at University of Edinburgh. In the 6th International Digital Curation Centre Conference, 6-8 December, 2010. Chicago, IL. Retrieved Sep. 26, 2011, from http://edina.ac.uk/presentations_publications/IDCC-rice-haywood.pdf

Rusbridge, C., et al. (2005). The Digital Curation Centre: A vision for digital curation. In: proceedings of Local to Global Data Interoperability – Challenges and

- Technologies, 20-24 June 2005, Sardinia, Italy. Retrieved Sep. 26, 2011, from <http://eprints.erpanet.org/82/>
- Shreeves, S. & M. Cragin. (2008). Introduction: Institutional repositories: Current state and future. *Library Trends* 57 (2), Fall 2008. Retrieved Sep. 26, 2011, from <http://www.ideals.illinois.edu/handle/2142/10679>
- Walters, T. & Skinner, K. (2011) . *New Roles for New Times: Digital Curation for Preservation*. Retrieved Sep. 26, 2011, from http://www.arl.org/bm~doc/nrnt_digital_curation17mar11.pdf
- Witt, M. (2008). Institutional Repositories and Research Data Curation in a Distributed Environment. *Library Trends*, 57(2), 191-201.
- Witt, M. & Carlson, J. R. (2007). Conducting a data interview. *Libraries Research Publications*, 81. Retrieved June 01, 2011, from http://docs.lib.purdue.edu/lib_research/81
- Witt, M., Carlson, J., Scott Brandt, D., & Cragin, M. H. (2009). Constructing Data Curation Profiles. *The International Journal of Digital Curation*,3(4)

附錄一

訪談說明

本研究目的在於瞭解臺灣大學圖書館若將研究資料納入典藏範圍，並提供保存、管理與分享之服務，校內教師對於此新服務之看法與需求。茲簡述本研究之背景與相關名詞。

研究背景：

資料度用(data curation)意指針對研究資料進行管理、保存及再利用之概念，強調的是在資料受到完善的保存與組織後，能夠提供檢索和取用，讓更多需要的使用者能夠持續的使用這些資料，以發揮其最大價值，並減少資源的浪費；而資料度用之優點包括現存資料的再分析、再利用；可作為驗證研究成果之依據；刺激科學的進步及研究之深度；避免重覆研究所造成的資源浪費等等。

資料密集是目前學術研究的趨勢之一，因此研究過程中產生的大量資料如何被有效的保存、管理和分享是許多學術機構十分重視的問題。英國的牛津大學和美國的普渡大學，訪談了許多研究者或教師，以瞭解他們對於研究資料保存、管理和分享的看法和需求，作為發展校內資料度用服務之重要參考；而資料度用也確實讓資料的取用更簡單迅速，研究者們可以專注於學術研究，提高研究品質與效率。國立臺灣大學(以下簡稱本校)自創校以來，為臺灣最負聲望的高等教育學府，多年來引領臺灣學術研究的風潮與脈動，累積豐富多元的學術研究成果。為擴大本校學術研究的影響面和能見度，希望了解校內教師對於研究資料的保存、管理及分享的看法與需求，作為臺大圖書館未來發展資料度用服務之重要參考。

名詞解釋：

研究資料：研究資料極具多樣性，目前國際間尚無單一而完整之定義。本研

究界定研究資料為研究過程中所使用和產出的資料，不包含硬體設備和實驗樣本，且以數位化之資料為主；可能包含**觀測數據、量測數據、實驗數據、地圖/圖層、基因序列資料、透過問卷或訪談蒐集而得的第一手資料**等。



附錄二

訪談大綱

1. 請簡單介紹您的研究領域，以及您近年來關注的研究主題。
2. 您的研究過程中，若是需要使用資料或數據，您通常會如何取得？例如資料庫、政府機關、其他學者等。
3. 您研究過程中，會使用和產生哪些類型的資料？例如數據、表格、圖檔、電腦程式等。
4. 您目前如何儲存和管理研究過程中所產出的資料？例如存在個人電腦、光碟、網路硬碟等。
5. 您是否願意和其他人分享您的研究資料？您認為所分享的資料其價值為何？例如可供再利用、避免重覆研究等。
6. 您的研究資料是否曾上傳國、內外任何資料庫或其他平台？
7. 臺大圖書館希望能夠協助校內教師更完善的保存研究過程中所產出的研究資料，並促進其再利用，使研究的影響力能夠更擴大，請問對於這樣的一個願景所發展的資料度用系統，您的看法及參與意願為何？
8. 請問您覺得是否有任何回饋機制(credit)可以增加您公開研究資料的意願？例如正式被引用於他人研究中、成為國科會認定的研究成果等。