

學術圖書館與資料應用

報告人：曾品方

2012/08/03

於臺大圖書館

大綱

- 資料度用之意義與重要性
- 學術圖書館發展資料度用之理由
- 學術圖書館發展資料度用之策略
- 館員在資料度用的角色
- 資料度用個案
- 結論與建議

資料度用之意義

- data curation 或是digital curation
- data：研究資料或稱數位研究資料，資料來自於研究者蒐集、觀察或創造，研究者透過資料分析後產出研究結果。
- 度用（curation）源自於拉丁文curare：照料（care for），研究資料如果沒有善加蒐集和照料，就無法提供取得和再次使用(Buneman, Müller & Rusbridge, 2009)。
- data curation包括度藏和使用，故譯為「資料度用」
- 資料度用是針對數位化的研究資料加以維護、保存和加值，建置值得信賴的數位典藏庫，以利現在和未來的使用（DCC, 2010）。

e-Research vs. data curation

- 資料度用是因應e-Research如何保存維護、公開取用大量的資料集，資料度用之發展與e-Research的興起息息相關。
- 為支援e-Research的發展，除了需要網路基本建設、網格運算等相關技術之外，也需要一套資料度用的方法，用以評估、蒐集、組織、保存和應用研究資料，俾便資料的再使用，以支援學者驗證或開創研究。

資料度用 vs. 數位保存

- 資料度用一詞往往包含數位保存、數位典藏的概念，但其實仍有所差別
- 資料度用是對於整體學術社群的深層思考，包括研究者、出版者、度用者、資料管理者以及使用者之間的緊密合作，資料度用不只是數位保存和數位典藏，更涵蓋研究社群與資料資源之間的密切互動（Beagrie, 2006）。
- 數位保存是資料度用的一部份，資料度用除了保存與維護數位資料之外，更強調資料增值、主動管理、建置值得信賴的資料典藏庫以確保未來的使用（Harvey, 2010）。

資料度用的重要性

- 資料氾濫

- 過去數十年來，科學研究透過精密的儀器、電腦化模擬、計算式模組產出大量研究資料。

- 資料流失

- 數位資料本身的特質、學者對於資料的認知、缺乏保存策略。

- 資料效益

- 驗證研究、啟發研究、創新研究。

學術圖書館發展資料度用之理由

- 長期以來，圖書館從事學術出版品的蒐集、典藏、組織，以提供取用服務，但對於研究歷程中最基本的元素：**資料**，卻沒有一套妥善的管理機制（Borgman, 2007）。
- 圖書館具備**資訊組織、資訊取用、館藏管理與機構典藏**的專業和經驗，可以作為資料度用的基礎能力（Witt, 2008）。
- 對於可感知的、明確的、具體的資訊，圖書館長久以來的徵集方法、資訊組織和典藏策略已然足夠，但是面對**數位化研究資料的不明確、不穩定、多變化**等特徵，圖書館面臨的是不同屬性的挑戰（Ogburn, 2010）。

學術圖書館發展資料度用之策略

- 主動參與研究歷程
 - 圖書館與研究者的夥伴關係是推動資料度用的重要驅力。
- 整合校園內外各項資源
 - 校內與校外既有資料，建置研究資料入口網站。
- 培訓練與重用館員
 - 館員是資料度用的執行者而非經紀人，有必要增進館員知能。
- 建立合作關係
 - 考量整體社群的需求而不是個別的工作流程或系統。
- 執行合作計畫
 - 機構內跨單位、機構之間、多機構的合作。
 - 透過度用服務更能彰顯圖書館和學術社群的合作關係。

館員在資料度用的角色

- 館員是內容徵集和所有權的諮詢顧問
- 館員是教師的教學夥伴
- 館員是數位研究資料的產生和使用之觀察者
- 館員是資料度用系統的建置者
- 館員是研究資料的產出者和傳播者
- 館員是組織的設計者
- 館員是合作網絡的創建者和參與者

資料度用個案

- 英國愛丁堡大學Edinburgh DataShare
- 美國麻省理工學院DSpace@MIT
- 美國普渡大學的D2C2(Distributed Data Curation Center)和PURR(Purdue University Research Repository)

Edinburgh DataShare



The University of Edinburgh

[University Homepage](#) [IS Homepage](#) [Research Support](#)

[Contact Us](#) [Login](#)

Information Services



[Edinburgh DataShare](#) > [Community List](#)



Search Edinburgh DataShare

Go

[Advanced Search](#)

Browse

[Edinburgh DataShare](#)
[Research Communities](#)
[Date Issued](#)
[Data Creators](#)
[Titles](#)
[Subjects](#)

My Account

[My Exports](#)

What is Edinburgh DataShare?

Edinburgh DataShare is an online digital repository of multi-disciplinary research datasets produced at the University of Edinburgh, hosted by the [Data Library](#) in Information Services.

Edinburgh University researchers who have produced research data associated with an existing or forthcoming publication, or which has potential use for other researchers, are invited to upload their dataset for sharing and safekeeping. A persistent identifier and suggested citation will be provided.

Upload Your Data



- [How to deposit](#)
- [Benefits of deposit](#)

Research Communities in Edinburgh DataShare

Select a community to browse its collections.

- [Information Services \(IS\)](#)
- [School of Engineering](#)
- [School of GeoSciences](#)

Spotlight



New to Edinburgh DataShare - a survey of technology use and attitudes of University of Edinburgh undergraduates.

Further Information

- [Overview](#)
- [Service Background](#)
- [Our Definitions](#)
- [Service Policies](#)

Latest Items



- [A survey of undergraduate technology use and attitudes \(27 Feb 2012\)](#)
- [Collaboration Tools Survey \(13 Feb 2012\)](#)

125%

愛丁堡大學發展資料度用的策略

- Edinburgh DataShare是英國資料分享計畫的一部份，隸屬於Information Service之Data Library
- 以DSpace建置的資料典藏庫，另有機構典藏：Edinburgh Research Archive（ERA）
- 研究人員以Edinburgh DataShare提交研究資料，每一筆完整的紀錄：詮釋資料、資料檔和文件檔，可以自行設定取用權限，並取得一個永久的URL。
- 在管理上，由上而下制定研究資料的**管理政策**；在技術上，由下而上根據研究者的需求，設計資料儲存**格式**、處理**流程**和相關**工具**（Rice & Haywood, 2010）。
- 圖書館並不是資料度用的主要單位，而是協同單位。

Search DSpace@MIT

- Search DSpace@MIT
 This Collection

[Advanced Search](#)

Browse

All of DSpace@MIT

- [Communities & Collections](#)
- [By Issue Date](#)
- [Authors](#)
- [Titles](#)
- [Subjects](#)

This Community

- [By Issue Date](#)
- [Authors](#)
- [Titles](#)
- [Subjects](#)

My Account

- [Login](#)
- [Register](#)

Links

[Home](#)

The Big Data Problem

Search within this community and its collections:

[Advanced Search](#)

Our MIT GeoSpatial Data Center, which consists of the MIT Auto-ID Laboratory (including the MIT Geonumerics Group, et al), MIT Center for Grid Computing, and MIT Intelligent Engineering Systems Laboratory, has a rich heritage and culture of advancing geospatial capabilities. We have had extensive large scale simulation experience to address the "Big Data Problem" that often plagues traditional

off-the-shelf open source GIS platforms, and we have collaborated with Sandia National Lab, Los Alamos National Lab, and the U.S. Department of Homeland Security National Infrastructure Simulation and Analysis Center (NISAC) on these large scale simulations.

The integrated mission for our MIT labs is to deploy a next generation open source geospatial platform, which provides more granular analytic capabilities and faster processing for complex queries by utilizing parallel processing capabilities (grid computing, GPU, and multicore). Our entire MIT Team and our collaborative partners are fervently dedicated to achieving this mission.

The MIT GeoSpatial Data Center is currently collaborating with:

- [Harvard Kennedy School's Center for Public Leadership \(CPL\)](#)
- University of Memphis' Department of Earth Sciences, GIS

Our integrated MIT research laboratories include:

- [MIT Auto-ID Laboratory](#)
- [MIT Center for Grid Computing](#)



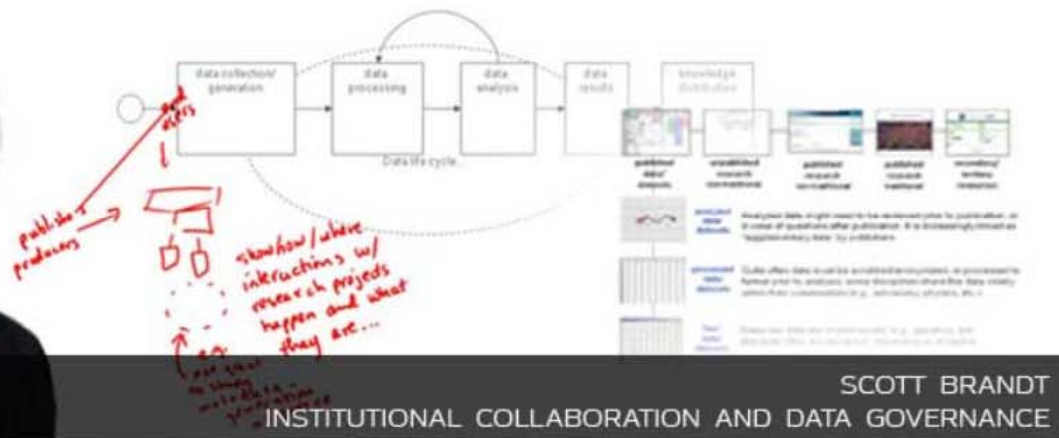
THE MIT GEOSPATIAL DATA CENTER
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
MIT Auto-ID Lab
MIT Center for Grid Computing
MIT Intelligent Engineering Systems Laboratory

MIT發展資料度用的策略

- **基礎建設**：採用原DSpace作為資料典藏庫，獲得NSF DataNet贊助，與校內的教師、資訊服務和科技部門合作，持續提昇DSpace的各項功能。
- **工作小組**：2005年由學科館員組成資料倡議小組，學習資料度用之作法，探討如何回應校內教師對於資料管理的需求，以及如何提供資料管理的服務。
- **學科館員**：分析資料管理的需求、協助製作資料管理計畫書、教導學生使用研究資料、蒐集與傳播資料集、研究資料的上傳、下載、客製與產出等相關事宜。

Gabridge, 2009

Distributed Data Curation Center



Distributed research. Distributed data. Distributed approaches. Distinct solutions.

The “data deluge” has created problems for researchers in a range of domains (from e-Science to Digital Humanities), and library science has been recognized by the NSF, NEH and others as having a critical role in helping researchers, institutions and organizations address it. The **Distributed Data Curation Center** investigates issues and problems related to making

普渡大學發展資料度用的策略

- 普渡大學**圖書館**為主要負責單位，合作方案的層次
 - **校內**：圖書館與校內電算中心合作
 - **校外**：普渡大學和伊利諾大學香檳分校（UIUC）合作共同探討研究人員在何種情境下願意公開資料
- D2C2提供資料度用相關之工具、參考文件、個案內容、資料管理政策等，但未具備資料提交、查詢、下載功能。
- 2011年推出**Purdue University Research Repository (PURR)**，提供一個線上合作的資料典藏庫，具備提交和查詢研究資料的功能，可支援全校研究人員撰寫資料管理計畫書。

What Does My Data Management Plan Need to Address?

There are many things that need to be included in your data management plan. Watch our video tutorials, read the step by step instructions, or view a completed data management plan to help you get started. You can also click the link below to start now.

Learn More



In the News: The National Science Foundation now requires all proposals to submit a Data Management Plan [Read More >](#)

Start Your Research Project



Read the DMP Requirements

Read the detailed requirements for your data management plan. The requirements are very specific and should help to clear up any confusion regarding your own data management plan. [Get Started >](#)



Create a Group for Your Project

Do you have a question?

Ask a Librarian

Send us your email question.

Enter your email address

Your Question/Message

Send

個案比較

- **發展背景**：國家型計畫的支持
- **機構典藏**：皆已建置機構典藏，但在應用上卻有涉入深淺的差異
- **合作模式**：內部與外部的合作策略
- **圖書館和館員的角色**
 - Edinburgh DataShare是全校資訊服務的一環，由副校長帶領跨部門的專責單位，圖書館僅為協同單位
 - DSpace@MIT以學科館員為主力
 - D2C2由跨學科研究館員作為推動資料度用的聯絡人

結論與建議

- **精進館員知能以勝任資料度用任務**
 - 館員在資料度用的角色，幾乎涵蓋整套資料度用的發展程序
 - MIT學科館員成立資料計畫小組；普渡大學任用跨學科研究館員
- **實施需求調查以奠定資料度用基礎**
 - 愛丁堡大學：根據研究人員需求，規劃儲存格式、處理流程和工具
 - MIT圖書館：館員和研究人員非正式的對談
 - 普渡大學圖書館：訪談研究人員，建置Data Curation Profile
- **規劃合作策略以推動資料度用服務**
 - 對象：電算中心、系所、研究者、學術圖書館社群以及研究經費提供機構等
 - 合作策略不只是建置資料典藏庫，更應關注各領域研究人員的研究情境，才能在多變的環境中發展健全的合作關係

參考書目1

- 陳雪華 (2010) 。 e-Research與圖書館知識服務 。 上網日期：2012年2月3日，檢自<http://www.lib.tku.edu.tw/libintro/pub/990513-2.pdf> 。
- Beagrie, N. (2006). Digital curation for science, digital libraries, and individuals. *The International Journal of Digital Curation*, 1(1), 3-16.
- Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, Mass.: MIT Press.
- Buneman, P., Müller, H., and Rusbridge, C. (2009). Curating the CIA world factbook. *The International Journal of Digital Curation*, 3(4), 29-43.
- Choudhury, G. S. (2008). Case study in data curation at Johns Hopkins University. *Library Trends*, 57(2), 211-220. Retrieved August 30, 2011, from <https://jscholarship.library.jhu.edu/handle/1774.2/34023>.
- DCC. (2010). About the DCC. Retrieved July 14, 2011, from <http://www.dcc.ac.uk/about-us/history-dcc>.
- Gabridge, T., (2009). The last mile: Liaison roles in curating science and engineering research data. *Research Library Issues: A Bimonthly. Report from ARL, CNI, and SPARC*, 265, 15-21. Retrieved June 20, 2011, from <http://www.arl.org/bm~doc/rli-265-gabridge.pdf>.

參考書目2

- Harvey, R. (2010). Digital curation: A how-to-do-it manual. New York: Neal-Schuman.
- Ogburn, J. L. (2010). The Imperative for data curation. Portal: Libraries and the Academy, 10(2), 241-246.
- Rice, R., and Haywood, J., (2010). Research data management initiatives at University of Edinburgh. Paper presented at the 6th International Digital Curation Centre Conference, Chicago, IL. Retrieved September 03, 2011, from http://edina.ac.uk/presentations_publications/IDCC-rice-haywood.pdf
- Walters, T., and Skinner, K. (2011). New roles for new times: Digital curation for preservation. Washing, DC.: Association of Research Libraries.
- Witt, M. (2008). Institutional repositories and research data curation in a distributed environment. Library Trends, 57(2), 191-201.
- Witt, M., and Carlson, J. R., (2007). Conducting a data interview. Libraries Research Publications. Paper 81. Retrieved June 01, 2011, from http://docs.lib.purdue.edu/lib_research/81.

報告完畢
請多指教