

數位典藏建置與應用

吳錦範

國立臺東大學圖書資訊館

FRANK@NTTU.EDU.TW

大綱

- 前言
- 建置
- 數位典藏網站
- 應用
 - 商業增值
 - 中小學教學
 - 學術研究
- 結語

前言

- **數位典藏 (digital archive)**
 - 係指將有保存價值之實體或非實體資料
 - 透過數位化方式 (攝影、掃描、影音拍攝、全文輸入等)
 - 並加上後設資料 (Metadata) 的描述
 - 以數位檔案的形式儲存

資料來源: [維基百科](#)

- 數位化的目的? (賴忠勤,2007)
 - 資料保存?
 - 減少原件取用的損害?
 - 取用與分享?

- 典藏什麼東西?
 - 所有館藏?
 - 選擇主題?
 - 臺東大學圖資館：兒童文學、南島文化

我國數位典藏的歷程

(典藏→ 資料詮釋 →脈絡分析→AI? (未來))

- 數位博物館專案計畫
 - 1998.09 ~ 2001
- 數位典藏與數位學習國家型科技計畫 (國科會)
 - 1999.09 ~ 2000.12 (籌備期)
 - 2002.01~2006 (第一期)
 - 2007(第二期)
 - 2008~2012 (第三期)
 - 參與單位：故宮博物院、國家圖書館、國立歷史博物館、國立自然科學博物館、臺灣省文獻委員會、臺灣大學和中央研究院
- 國家文化資料庫計畫 (文建會)
 - 2008 ~
 - 挑戰2008：國家發展重點計畫
- 數位典藏邁向數位人文
 - 2012 ~
 - 文本資料分析
- AI (未來)

6

數位典藏建置

數位典藏工作流程架構

(蔡永橙、黃國倫、邱志義等)

- 採集：
 - 依據不同典藏品的類別，有系統、有計畫的累積蒐集，作為數位化之素材。
- 後製：
 - 運用資訊技術將蒐集之媒體資料轉換為可儲存、處理及編輯之形式，並加以組織及分類；
 - 進行資料描述記錄，使用者查詢時索引工具。
- 典藏：
 - 藉由整體性的典藏規劃，以建構合適資訊系統；
 - 透過系統運作達到數位資料保存及管理功能。

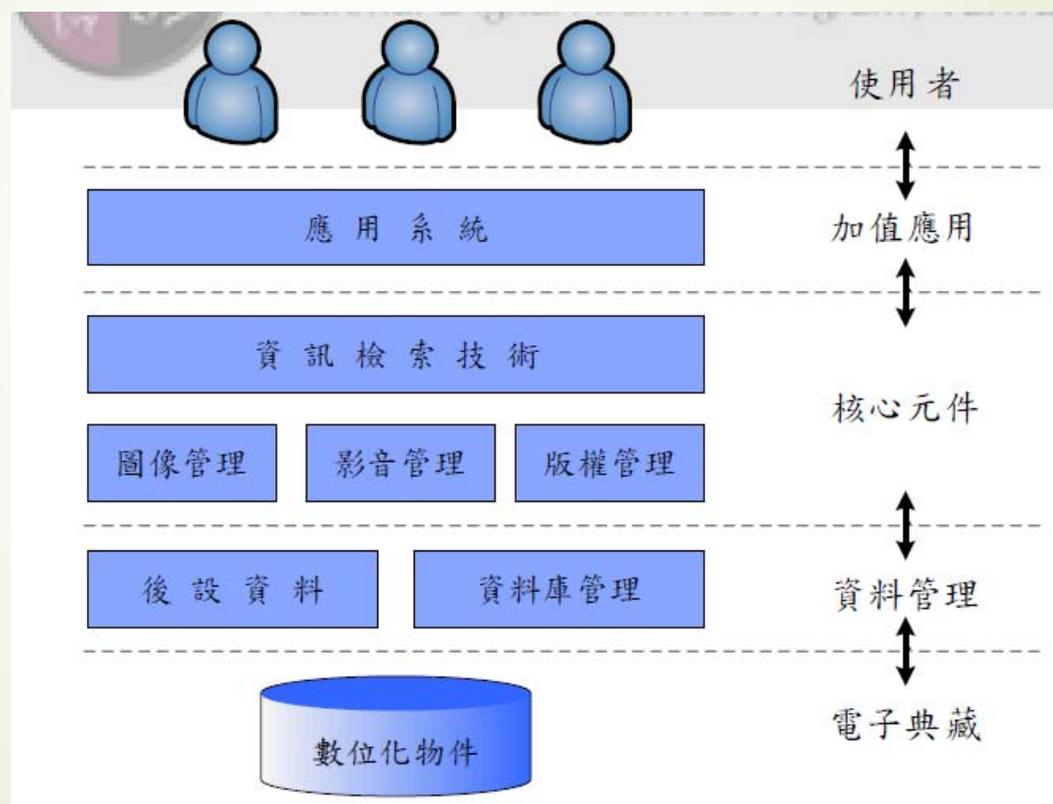
數位典藏工作流程架構(續)

- 應用加值：
 - 應用資訊技術，將數位化之素材結合專業人員之描述註解，加以編輯組合成各類網路服務或資訊產品。

- 傳播：
 - 利用網際網路或電子媒體傳達給使用者。

- 使用與呈現：
 - 設計多種資訊存取方法，提供使用者快速取得所需之資訊。

數位典藏系統模組架構



資料來源：[數位典藏技術導論](#)

技術議題 (蔡永橙、黃國倫、邱志義等)

- 人機介面：
 - 利用軟、硬體設施及存取工具，提供操作環境與互動式介面。

- 通訊網路：
 - 提供網路連線，讓使用者可遠距離使用。

- 自然語言處理：
 - 透過自然語言語法、語意、自動翻譯等技術，瞭解使用者需求，提供最佳服務品質。

技術議題(續)

- 資訊壓縮：
 - 利用壓縮技術節省儲存空間，並提供快速的傳輸與存取服務。

- 資訊安全認證：
 - 利用浮水印 (Watermarking)、數位簽章 (Digital Signature) 及加密 (Encryption) 等技術，提供智慧財產權保護機制。

- 整合服務：
 - 利用智慧型代理伺服器 (Intelligent Agent) 與資料探勘技術，整合資訊蒐集、管理及傳播功能，提供自動學習能力的數位典藏系統。

數位化開始之前

➤ 清點

- 傳單、海報、信件（傳真）、信封、明信片、卡片、文物、模型、教具、幻燈片、底片、照片、錄影帶(VHS、V8、Hi8)、錄音帶、CD、...

➤ 資料類型

- 照片、文件、影音、立體文物、地圖、手稿、...

➤ 資料整理

- 分類、記錄標題、相關人物、時間、空間、關聯藏品、群組組件、主題、關鍵詞、說明、備註等資訊；權威資料

➤ 著作權盤點

- 公共財、取得授權、...

數位化設備工具

- 掃描器
 - 文件、照片、手稿、....等
- 底片掃描機
 - 正片、負片
- 3D照像機
 - 立體文物數位化
- 類比影音資料轉檔設備
 - 錄音帶、錄影帶、...
- OCR
 - 文本
-

數位化方式

- 自建
 - 從資料清點、整理、掃描等全部由單位自行處理
 - 問題：人力、設備、專業能力、...等
- 委外
 - 單位較省事
 - 問題：經費、招標程序
- 混合
 - 上述二者混搭
 - 問題：如何分工、權責劃分、招標程序

系統平台

- 委外開發
 - 客製化系統
 - 問題：結案後的維護、升級等

- 自行開發
 - 使用Open Source
 - 如：[AtoM](#)、[Archivematica](#)等
 - 自行開發系統
 - 問題：單位資訊能力

資料儲存

- 資料安全(硬體)
 - 具容錯功能
- 存取效率
- 備援
 - 資料備份321原則
 - 至少備份三份
 - 使用兩種以上不同形式
 - 其中一份備份要存放異地(異地備援)

檔案典、閱分離原則

➤ 典藏

- 高解析度 (如：文字300dpi)
- 不壓縮或不失真的壓縮
- 國際標準檔案格式 (如: tif格式)

➤ 線上瀏覽

- 中、低解析度 (如：72dpi)
- 支援網頁瀏覽的檔案格式(如:gif、jpg)

詮釋資料(Metadata)

- ▶ 何謂詮釋資料?(又稱：後設資料、元資料)
 - ▶ Data about Data
- ▶ 數位典藏常用的詮釋資料-都柏林核心集 (Dublin Core, DC)
 - ▶ 是一組簡單而有效的核心元素集
 - ▶ 僅15個元素
 - ▶ 因應各類型資料不同，會有所調整

其他重點

- 遵循國際標準
 - 資料交換
 - 數位物件檔案轉換
 - 語文字碼
 - Unicode
- 永久保存
 - 硬體 (升級、搬遷)
 - 檔案
- 著作權
 - 財產權 (重製權、公開傳輸權、公開播映、...)、人格權

著作權議題 (謝銘祥, 2008)

- 典藏前 (著作權盤點)
 - 數位化的法源基礎
 - 是否還受著作權法保護?
財產權消滅,但例外:文化資產保存法、原住民智慧創作保護條例等
 - 原住民智慧創作保護條例：
 - 原住民族傳統之宗教祭儀、音樂、舞蹈、歌曲、雕塑、編織、圖案、服飾、民俗技藝或其他文化成果之表達
 - 採登記制
 - 著作權之保護與權利歸屬狀態加以釐清
 - 重製權

著作權議題(續)

- 數位化後(加值應用)
 - 數位化成果在法律上之地位
 - 委由外部單位(個人或企業)進行數位化工作，承攬契約中應明定數位化成果**權利歸屬於典藏機構**，以保障典藏機構之權益，避免不必要爭議。

數位化建置成本 (賴忠勤, 2007)

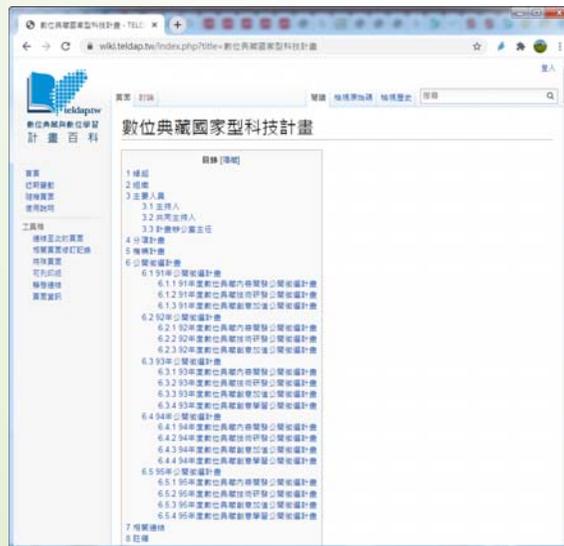
SN	項目	單價	單位
1	影像掃描	11	頁
2	影像檔轉存	?	批
3	Metadata建檔及校對	100	冊
4	圖書裝訂及修復作業	400	冊
5	儲存與備份媒體製作(DVD、HDD)	?	份
6	學者專家出席、審查、交通、差旅費	2000	人次
7	工讀費	25000	人/月
8	雜支	?	式
9	數位典藏系統	?	套
10	伺服器主機及儲存設備	?	套
11	隱性成本	?	

23

數位典藏網站

數位典藏國家型科技計畫

➔ <https://wiki.teldap.tw/index.php?title=%E6%95%B8%E4%BD%8D%E5%85%B8%E8%97%8F%E5%9C%8B%E5%AE%B6%E5%9E%8B%E7%A7%91%E6%8A%80%E8%A8%88%E7%95%AB>



2020東區館合年會



109/11/20

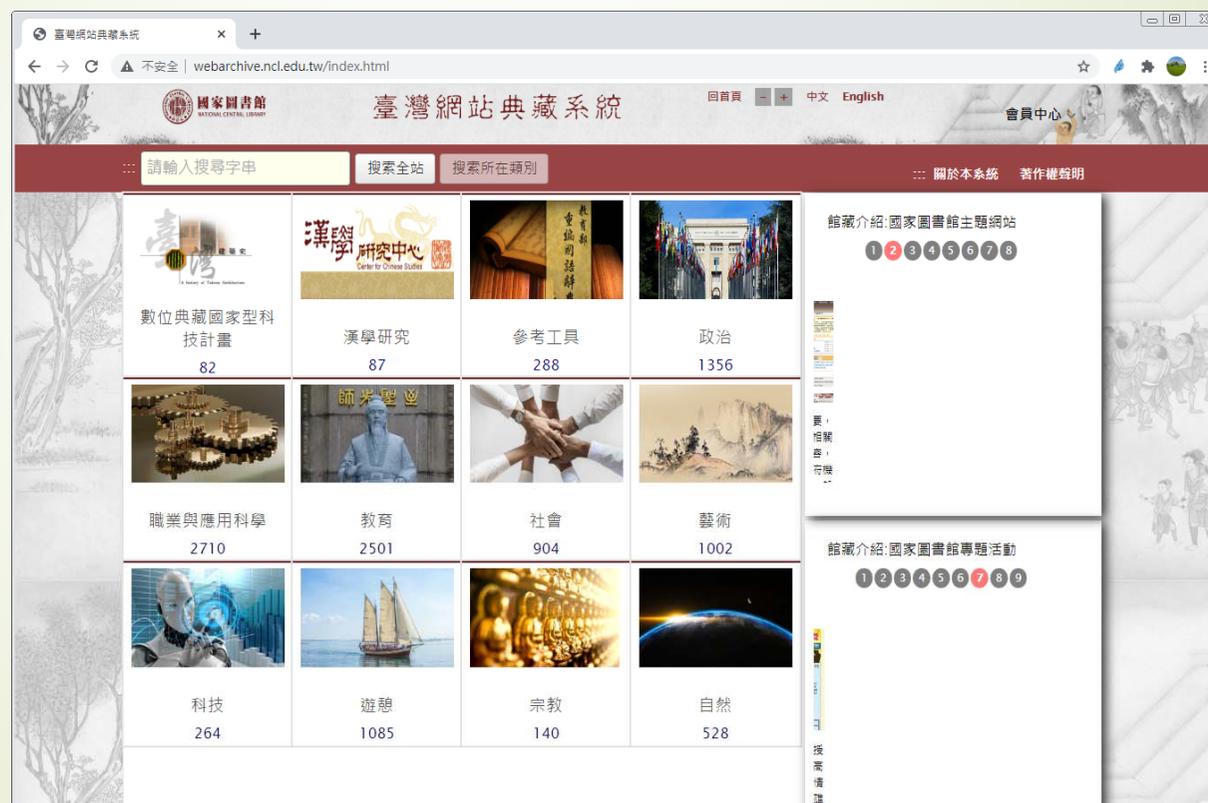
國家圖書館：古籍與特藏文獻資源

<http://rbook.ncl.edu.tw/NCLSearch/>



國家圖書館：臺灣網站典藏系統

(<http://webarchive.ncl.edu.tw/>)



文化部國家文化資料庫

(<http://nrch.culture.tw/>)

➤ 國家文化記憶庫

- 將各文化機關（構）典藏文物或管轄之業務相關文化元素，予以系統化蒐集整理，提供便利的創新文化資訊務，供大眾使用。
- 共分：台灣歷史、台灣美術、台灣工藝、台灣傳統藝術及台灣文學等五大主題。

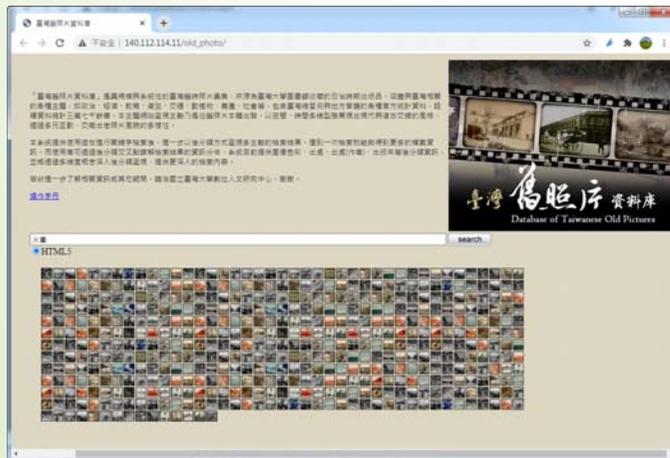


文化部國家文化資料庫

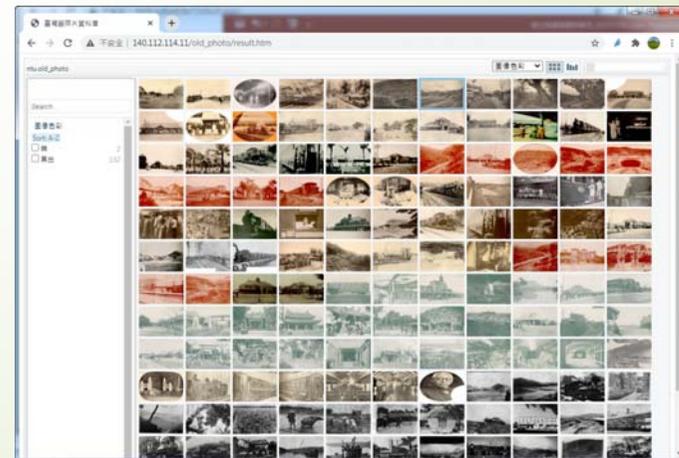
--台灣舊照片資料庫

(http://140.112.114.11/old_photo/)

- 建置單位：國立臺灣大學圖書館
- 源為臺灣大學圖書館收藏的日治時期出版品，涵蓋與臺灣相關的各種主題，如政治、經濟、教育、衛生、交通、動植物、農產、社會等
- 詮釋資料總計三萬七千餘筆



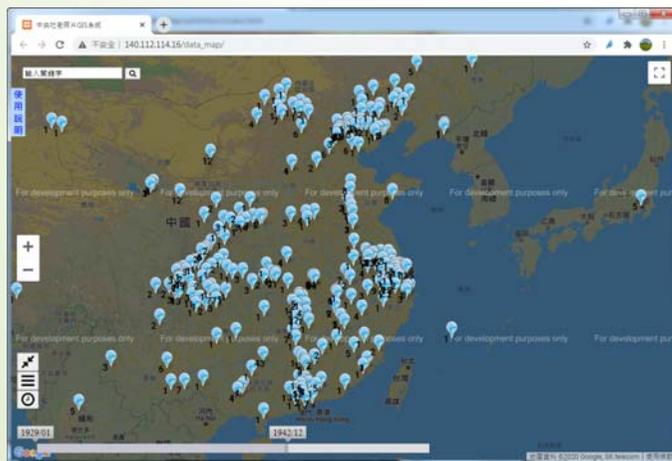
2020東區館合年會



109/11/20

文化部國家文化資料庫 --中央社老照片GIS系統 (http://140.112.114.16/data_map/)

- 建置單位：中央社
- 結合GIS功能



31

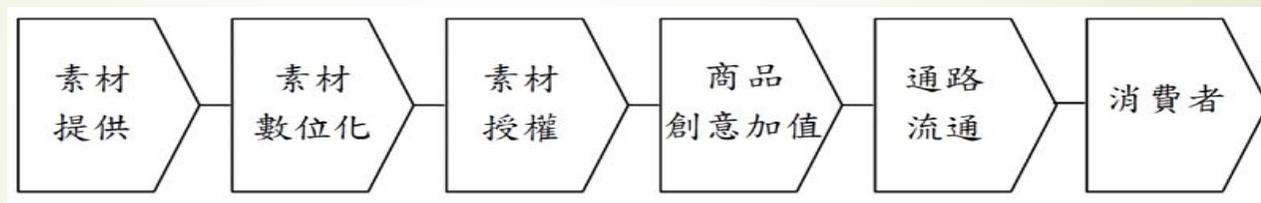
數位典藏應用

商業增值

應用

--商業增值(謝銘祥, 2008)

➤ 商業 (產業增值)



資料來源: 蘇鈺絜 (2006)

➤ 增值應用

- 保存與典藏目的外的各種衍生商品
- 如：動畫、遊戲；拍電影、電視；民生用品、皮件、衣飾等

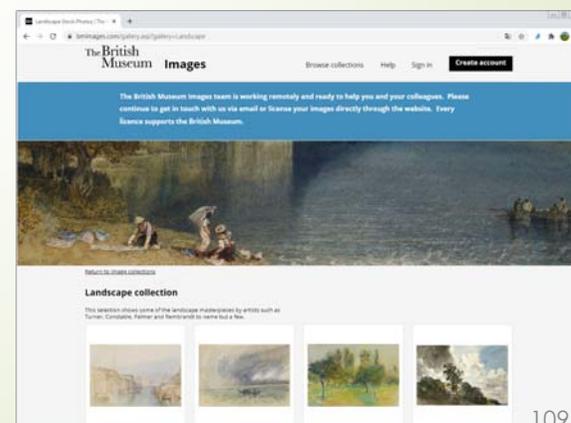
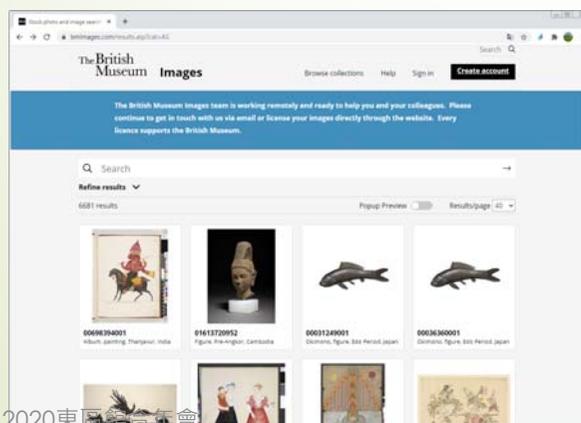
➤ 因應資訊公開法

- 典藏：高解析度
- 瀏覽：中、低解析度

應用

--商業增值(續)

- 英國倫敦大英博物館 (<https://www.bmimages.com/>)
 - 1753 年成立，館藏文物總計八百萬件以上
 - 增值服務目前可大致分為影像授權以及應用增值商品化
 - 服務包括搜尋、購買、授權與供下載較高解析度
 - 與部份廠商協力合作，授權廠商增值應用商品



應用

--商業增值(續)

- 國立故宮博物院
 - 圖像授權模式
 - 圖像會加上隱性浮水印
- 國家圖書館
 - 數位化標的
 - 珍藏古籍文獻、地方文獻及期刊報紙等
 - 數位化的利用
 - 國圖內使用及教育、學術性質為主之利用
 - 合理使用的空間
 - 先天上的限制
 - 典藏內容都是他人著作，必須先取得權利人的授權

典藏成果之應用問題 (謝銘祥, 2008)

- 典藏單位未必擁有權利
 - 典藏品來源多元，可能只有保管權及物權，未必有權從事授權
 - 典藏品著作權保護期間仍未屆滿；來源年代久遠，找不到權利人或不清楚原先訂定之權限範圍
- 素材特殊性導致授權加值疑慮
 - 典藏品牽涉肖像權的人物照片、涉及隱私權的私人信件、甚至政治敏感等問題，已超出著作權所考量的範圍
- 創用CC 授權的運用困難
 - 典藏資源並非每一筆權利狀態都很清楚，在一個整合文字、圖片、影音等龐大資料的資料庫中，無法在資料庫首頁一體宣告適用創用CC 條款

37

數位典藏應用

中、小學教學

應用

--中小學教學

- 數位典藏
 - 人文
 - 民國活動
 - 期望文化
 - 教育



無法連上這個網站

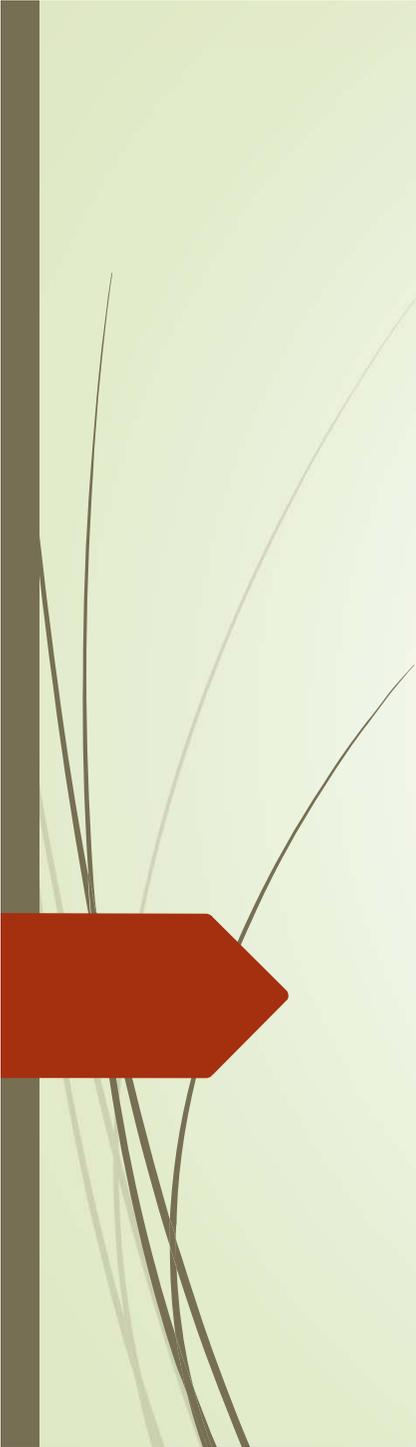
idatp.moe.edu.tw 的回應時間過長。

建議做法：

- 檢查連線狀態
- 檢查 Proxy 和防火牆
- 執行 Windows 網路診斷

ERR_CONNECTION_TIMED_OUT

體育五大領域
活動設計比賽」
推廣國家重要



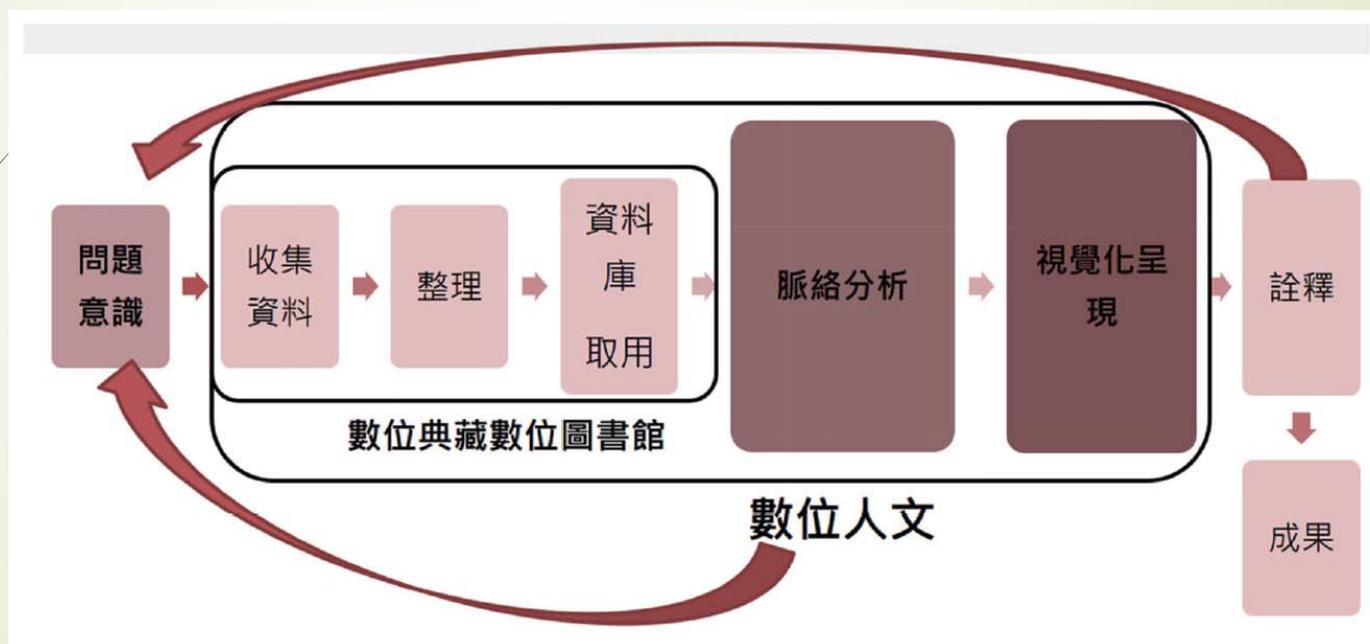
數位典藏應用

學術研究--數位人文

何謂數位人文？ (項潔, 2018)

- 結合大量數位資料，運用資訊科技，來從事人文研究之跨域整合工作。 (項潔, 2011)
- 唯有借助數位科技方能進行的人文研究 (項潔, 2011)
- 應用數位資訊技術於數據資料與文本資料的研究議題,特別是議題的結構、脈絡與趨勢。 (陳光華, 2017)
- 利用數位科技來輔助人文學者做研究。 (柯皓仁, 2017)
- 是一個資訊技術與人文不斷互動的過程。
- 透過資訊工具來協助人文研究者進行研究歷程中的「脈絡分析」與「視覺化呈現」的需求。
- 跨領域的結合資訊技術與人文研究，建置符合檔案特性，以及研究者需求的數位人文工具。

何謂數位人文? (續)



資料來源: [從數位典藏到數位人文](#) (項潔, 2018)

國家圖書館： 通用型古籍數位人文研究平台

(<http://ming.ncl.edu.tw/>)

- ▶ 明人文集為文本，將其數位化成果全文轉置而成，結合政大社資中心開發之各項數位分析工具，建置出一個既符合人文學者研究需求，又能提供大眾古籍內容分析的平台



台灣歷史數位圖書館

(<http://thdl.ntu.edu.tw/>)

➤ 資料蒐錄

- 集合台灣史一手史料之資料庫，提供全文檢索、詮釋資料 (metadata) 檢索等功能。本資料庫目前包含「淡新檔案」、「明清臺灣行政檔案」與「古契書」三文獻集，有全文資料逾十萬筆。

➤ 研究工具

- 參考工具，將研究者常查閱的工具書數位化，便於線上查索；
- 檢索分析工具，針對THDL內的資料，提供進階的功能及更深入的觀察；
- 關係探勘工具，利用自動方式重建文件之間的關係。

數位人文的應用

DocuSky數位人文學術研究平台

(https://docusky.org.tw/DocuSky/index2_20181204.html)

➡ 玄奘取經路線圖



Docu上第002：玄奘取經路線圖



46

結語

結語

- 數位化只是第一步，應用、加值才有價值
- 數位典藏，要先評估標的及目的
- 著作權議題不可忽略
- 必須考慮永續經營
- AI導入，未來發展值得觀察

48

謝謝聆聽!

Q & A

參考資料

- 中央研究院。《數位典藏與數位學習國家型科技計畫》。URL: <http://teldap.tw/Introduction/introduction.html>
- 維基百科。數位典藏。URL: <https://zh.wikipedia.org/wiki/%E6%95%B8%E4%BD%8D%E5%85%B8%E8%97%8F>
- 吳明德 (2005)。小學教師網際網路教學資源尋求及使用行為之研究。教育資料與圖書館學，42(4)，481-498。URL: <http://readopac3.ncl.edu.tw/nclJournal/GetPDF?fid=A05013302&jid=83300030&eid=da0648067f516851dbfb8052995cb88f>
- 蘇鈺絮(2006)。數位典藏產業價值體系與關鍵因素之研究。未出版碩士論文，國立台灣大學商學研究所，台北。URL: <http://dx.doi.org/10.6342/NTU.2006.00301>
- 數位典藏國家型科技計畫。《數位典藏技術導論》，2007年。URL: <https://ebook.iis.sinica.edu.tw/download.htm>
- 賴志勤(2007)。數位典藏建置劃與管理。佛教圖書館館刊，第45期，6-18。URL: <http://www.gaya.org.tw/journal/m45/45-main1.pdf>
- 謝銘洋、趙義隆、陳曉蕙 (2008)。數位典藏之保護與授權加值應用相關法律問題探討。藝術教育研究，第16期，77-106。URL: <https://www.gnae.world/wp-content/uploads/0sub/16/16-3.pdf>
- 項潔、涂豐恩(2011)。導論—什麼是數位人文。《從保存到創造：開啟數位人文研究》，臺北：國立臺灣大學出版中心，2011，頁9-28。URL: http://thdl.ntu.edu.tw/thdl_pdf/book1.pdf
- 科技部人文及社會科學研究發展司。《數位人文研究》，2017年01月。URL: <https://web.most.gov.tw/hssnq/ebook28/mobile/index.html>
- 陳光華(2017)。序文。《數位人文研究》，科技部人文及社會科學研究發展司，2017年，頁3-4。URL: <https://web.most.gov.tw/hssnq/ebook28/mobile/index.html>
- 柯皓仁(2017)。從數位典藏到數位人文—從數位圖書館與數位典藏說起(上)。URL: <https://www.youtube.com/watch?v=ubYSqCfwrkE>
- 項潔(2018)。發刊詞：從數位典藏到數位人文。數位典藏與數位人文，第1期，Pi-v。URL: <http://www.digital.ntu.edu.tw/hsiang/pdf/%E7%99%BC%E5%88%8A%E8%A9%9E%EF%BC%9A%E5%BE%9E%E6%95%B8%E4%BD%8D%E5%85%B8%E8%97%8F%E5%88%B0%E6%95%B8%E4%BD%8D%E4%BA%BA%E6%96%87>