

巨量數位影像管理與應用—— 以中研院史語所考古數位典藏計畫為例

溫子軍¹ 黃婷鈺² 林玉雲³

摘要

本文作者以專職影像管理者之角度出發，探討管理來自不同數位化來源之巨量數位影像，從影像檔案的生產、處理、備份等過程所面臨之問題與挑戰，到如何兼顧品質、資料應用以及長久保存等問題。

中央研究院歷史語言研究所執行之「考古發掘遺物、照片、記錄與檔案」數位典藏計畫至今累積超過 30 萬筆 10TB 以上之數位影像，且每年以約 3 萬筆 4TB 的數量持續增加。數位影像之生產來源多元，除了攝影師，尚包括資料掃描人員及外包廠商等，多方面的影像生產來源，帶來巨大的管理負荷。

歷史語言研究所考古數位典藏計畫自 2001 年之先導計畫至今已超過 10 年，所跨時間縱軸深，其間數位科技進展快速，加上數位影像因不同時期及需求，而有不同的版本。此外，隨著研究進展，考古遺物的影像亦隨之變動，例如甲骨綴合、成套器物等，種種原因產生不同階段的影像，影像版本的問題造成管理以及應用上種種挑戰。

此一龐大的數位資產，須考慮資料之正確與品質完好之外，更須兼顧典藏之完整性，備份的管理以及備份的頻率成為不可或缺的一環。檔案典藏的分類架構、跨平臺衍生的問題，在巨量資料的負擔下更需要有一套快速簡便的管理方法，以面對接踵而來之資料庫建置及各種影像應用需求。本文個案研究上述計畫管理此一珍貴且龐大的數位影像資產所面臨的挑戰，並將所提出的解決方案與美國「蓋提研究所」(The Getty Research Institute) 以及英國「考古資料服務網」(The Archaeology Data Service) 所提出之相關建議做一分析與討論。

關鍵詞：數位影像管理、影像編碼、數位影像品質

¹ E-mail: cgilm@mail.ihp.sinica.edu.tw

² E-mail: tingyu@mail.ihp.sinica.edu.tw

³ E-mail: hsyuyun@mail.ihp.sinica.edu.tw

前言

中央研究院（以下簡稱中研院）歷史語言研究所（以下簡稱史語所）執行之「考古發掘遺物、照片、記錄與檔案」數位典藏計畫，歷經 2001 年之數位典藏先導計畫（2001 年）、數位典藏第一期（2002-2006 年），持續至第二期（2007-2012 年），至今已邁入第 11 個年頭，共累積超過 30 萬筆 10TB 以上之數位影像檔案，且每年以約 3 萬筆 4TB 的數量持續增加。

影像內容包括考古遺物、田野影像（幻燈片、底片、照片）以及田野文件；媒體類型涵蓋玻璃板底片、正片、負片、X 光片、紅外線照片、拓片、線繪圖、數位照片、3D 環物影像等。此一龐大的數位資產，在管理上若無章法，並搭配與時俱進的設備，將成為一沈重的負荷。

本文個案研究中研院史語所「考古發掘遺物、照片、記錄與檔案」數位典藏計畫管理此一珍貴且龐大的數位影像資產所面臨的挑戰，並將其解決方案與美國「蓋提研究所」(The Getty Research Institute) 出版之《數位影像指南》(Introduction to Imaging)(Besser, 2003)⁴ 以及「考古資料服務網」(The Archaeology Data Service, 以下簡稱 ADS) 與 Digital Antiquity 出版之《最佳實務指引》(Guides to Good Practice)⁵ 所提出之相關建議做一比較與分析。

蓋提研究所致力於視覺藝術及其歷史的研究與推廣，其數位資源提供各國學者及民眾使用，並以藝術藏品數位化準則的研究與出版著稱。《數位影像指南》乃為協助館員、典藏經理等瞭解建

立數位影像資料庫所涵蓋的技術與程序，並提供數位影像存取、管理與保存的相關資訊；ADS 成立於 1996 年，是「藝術與人文資料庫」(Arts and Humanities Data Service, 以下簡稱 AHDS) 的主要貢獻者之一，提供考古學研究、學習與免費教學、高品質的數位資源，並長久保存、提昇及推廣考古數位資料，《最佳實務指引》主要針對欲建立考古數位資料集者，提供一個考古資料工作流程的基礎。前者為多元廣泛的博物館藝術品典藏所設想，後者為考古資料而設計，與本文個案研究性質相近，在兼顧一般性及主題性的前提下，是以做為本文與個案研究計畫之參照。以下就本文討論的重點，影像檔案之生產與應用、面臨的問題與挑戰，以及未來展望等議題，與蓋提研究所及 ADS 所出版的相關數位影像指南進行討論。

影像檔案之生產與應用

以考古資料範疇而言，數位化的對象大致可分為 3 種：考古遺物、田野影像以及田野文件。這些原件經過數位化，無論是數位攝影或是掃描，所產生的數位影像必須經過生產與衍生，才能夠達到應用的目的，以下分別說明之：

一、生產

數位化工作於執行之前即須訂立一套規格標準，此一標準須於品質與成本之間取得平衡。換句話說，所生產的影像品質越高，所耗費的成本例如執行物件數位化工作的時間、儲存容量及備份

⁴ http://www.getty.edu/research/publications/electronic_publications/introimages/intro.html（瀏覽日期：2012/11/30）。此指南為近 10 年前之出版品，蓋提研究所是否有更新的作法，因未見其更新之出版品無所知悉，因此僅以此版本做為參考。

⁵ <http://guides.archaeologydataservice.ac.uk/#section-Contents-GuidesToGoodPracticeTableOfContents>（瀏覽日期：2012/11/30）。

空間等也較多。蓋提研究所認為應儘可能在數位化的過程中取得最高品質的主檔 (Master File)，然而做為典藏的檔案是否為 RAW 檔⁶，則須視現有的數位資產管理政策、預算和存儲限制來決定⁷；ADS 亦提出須於檔案品質以及成本之間取得平衡，然而影像採用最高品質的規格會是較為理想的，ADS 建議採用不壓縮 TIFF 檔⁸做為典藏檔案，而非 RAW 檔⁹；史語所考古數位典藏則按原件的類型以及重要性，訂定不同的規格標準以及拍攝張數，除了採 RAW 檔做為典藏檔之外，亦轉出 TIFF 檔備存。蓋提研究所、ADS 以及史語所考古數位典藏計畫所提出的影像品質與規格整理如表 1。

中研院史語所考古數位典藏計畫與蓋提研究所以及 ADS 針對影像品質與規格的取決有 2 點不同之處：第一為典

藏主檔的格式。史語所考古數位典藏計畫主要典藏 RAW 檔（本文稱為 A 檔，參見表 2），並另外轉出未壓縮 TIFF 檔（本文稱為 B 檔，參見表 2），除非讀取及處理 RAW 檔的軟體授權僅限於一臺電腦（例如：X 光科學影像分析儀——影像讀取機 (FUJIFILM FLA-5000) 規格詳見附件一），則必須轉出未壓縮 TIFF 檔做為 A 檔典藏，以減少 RAW 檔無法讀取的風險。而蓋提研究所以及 ADS 建議典藏未壓縮 TIFF 檔，相較之下，史語所考古數位典藏計畫須耗用較大的典藏空間，除了因上述計畫具備較大的典藏容量條件之外，採用此一作法為考量考古資料，尤其是遺物的脆弱性，不宜高頻率的調出拍攝，RAW 檔為相機或掃描器所接收最原始的影像資訊，保存最多資訊，且沒有經過任何修飾，較便利於

表 1. 影像品質與規格比較表

項 目	蓋提研究所	ADS	中研院史語所考古數位典藏計畫
原件類型	博物館藝術品藏品	考古遺物、文件、甚至於雷射掃描 (laser scanning)、近景攝影測量 (close-range digital photogrammetry)、影音資料等	考古遺物、田野影像、田野文件、紀錄片
影像品質	儘可能取得最高品質	須在影像品質與檔案大小間取得平衡	按原件的類型與重要性，訂定不同的規格標準（詳見文中討論）
影像規格	應儘可能捕捉最高解析度和最大採樣深度	儘可能取得最高規格的影像	拍攝：以數位相機／機背最大的像素做擷取 掃描：根據原件類型而有不同規格（詳見文中討論）
典藏主檔	建議典藏未壓縮的 TIFF 檔	建議典藏未壓縮 TIFF、DNG 檔	典藏 RAW 檔，但亦典藏轉出之未壓縮 TIFF 檔
遺物拍攝張數	無	無	精品及重要標本以數位機背拍攝多張 一般標本採數位單眼相機拍攝 1 至 2 張

⁶ http://www.adobe.com/digitalimag/pdfs/understanding_digitalrawcapture.pdf（瀏覽日期：2011/06/11）；
<http://www.digitalpreservation.gov/formats/fdd/fdd000241.shtml>（瀏覽日期：2011/06/01）。

⁷ http://www.getty.edu/research/publications/electronic_publications/introimages/capture.html（瀏覽日期：2012/11/30）。

⁸ <http://www.digitalpreservation.gov/formats/fdd/fdd000022.shtml>（瀏覽日期：2011/06/01）。

⁹ http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_2（瀏覽日期：2012/11/30）；
http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_3（瀏覽日期：2012/11/30）。

後續參數調整，以達最佳化影像品質，可減少因重拍或重掃而造成的遺物或原件損壞。此外，亦典藏未壓縮 TIFF 檔（B 檔）可方便數位影像之應用，例如印刷、出版、展覽等。第二，蓋提研究所以及 ADS 建議盡量擷取最高品質影像為準則，史語所考古數位典藏計畫則按原件的類型與重要性而訂定不同的規格標準¹⁰以及拍攝張數；考古遺物乃根據其重要性分為精品、重要標本及一般標本三大類；精品及重要標本以數位機背做擷取，並拍攝多張，一般標本採數位單眼相機拍攝 1 至 2 張，影像規格皆以數位機背／相機之最大像素做擷取。田野底片或照片之掃描則以印刷輸出 A3，300dpi 做為基準，文件類採 1：1，300-600dpi 掃描（設備規格參見附件一）。此作法可減少人力以及時間成本的耗用，並兼顧數位影像品質以及後續影像應用之需求，且此方式亦可於典藏空間及品質之間取得平衡。

二、衍生

由於各種設備有其專有之 RAW 檔格式（見附件一），加上初始產生的 RAW 檔影像雖保留原始資訊，一般軟體卻無法支援，致使無法進一步處理與應用，因此需要將 RAW 檔轉為一般通用格式，例如 TIFF、JPEG¹¹等衍生圖檔。蓋提研究所提及現今影像大多透過網路傳輸，JPEG 與 GIF 檔是較為普遍的應用格式，且大部分的瀏覽器都支援這些規格，使用者不需要再下載任何檢視或解壓縮軟體¹²；ADS 則未對衍生檔案規格提出建議。史語所考古數位典藏計畫則衍生不止一種圖檔規格，且針對不同用途衍生出專屬規格，其衍生方式及規格詳述如下：

數位影像自初始 RAW 檔（A 檔）轉出 TIFF 檔（B 檔）之後，再以 Adobe Photoshop 批次降階轉出較小之檔案（C、D、E 檔），如表 2 所示。

B 檔，即副主檔（Sub Master File）（徐

表 2. 史語所考古數位典藏計畫影像衍生檔案種類及用途¹³

種類	格式	說明	用途	大小
A 檔	RAW、TIFF	最原始產出的檔案	典藏	800MB 以下
B 檔	TIFF	轉自 A 檔 色彩深度 16bit 的 TIFF 檔，品質 ≤ RAW 檔	典藏、衍生其他檔、修圖	800MB 以下
C 檔	TIFF	轉自 B 檔 色彩深度 8bit，A4 大小的應用檔	影像資料庫 ¹⁴ 、出版印刷	30MB 以下
D 檔	JPEG	轉自 C 檔 色彩深度 8bit，C 檔的 JPEG 檔	網頁及驗收文件留存	2MB 以下
E 檔	JPEG	轉自 D 檔 色彩深度 8bit	索引圖清單（內部使用）	50KB 以下

¹⁰ 國立故宮博物院張志光先生亦建議影像數位化規格的標的應以原稿種類與數位化方式而訂立（張志光，2006）。

¹¹ http://www.digitalpreservation.gov/formats/fdd/still_fdd.shtml（瀏覽日期：2011/06/01）。

¹² http://www.getty.edu/research/publications/electronic_publications/introimages/capture.html（瀏覽日期：2012/11/30）。

¹³ 史語所考古數位典藏計畫所衍生的以上幾種規格檔案乃因計畫執行實際需求而訂立，並同時將其他應用需求一併考量。

¹⁴ 史語所考古數位典藏計畫之影像資料庫為中研院資訊科學研究所所開發之「多媒體中心」。

明景，2011），乃為衍生其他 C、D、E 檔，即取用檔 (Access File) (徐明景，2011) 而產生，本計畫為達到更好的影像品質，因而修圖採用此一規格。由於 C 檔主要乃為出版所需而衍生，而出版品大多不超出 A4 規格，又 TIFF 檔廣泛應用於出版業界，印刷業對其支援性高，因此 C 檔以此為基準，並可直接應用於影像資料庫¹⁵，便於線上檢視、管理與下載；D 檔主要為網頁而衍生，選擇採用 JPEG 壓縮圖檔以利網路傳輸速度效能，復加上瀏覽器對其支援性較佳所致。D 檔規格檔案小且仍具備一定的影像品質，目前以此檔提供驗收文件留存，同時也是「考古資料數位典藏資料庫」對外提供查詢瀏覽的規格。E 檔製作為索引圖清單，提供研究者參閱。

如此衍生出數種不同規格的影像，其好處在於當需求用途不同時可立即使用，不需要動用初始 A 檔，尤其是需求量大時（例如：線上資料庫之呈現、圖錄出版、支援展覽等），因而於應用時，可節省人力及時間的耗費，且不需考慮 A 檔格式、色彩描述檔（色彩管理用之 ICC Profile¹⁶）及檔案大小問題而造成的種種限制。然此作法亦有其缺點：第一增加儲存空間所需的容量，第二增加影像管理以及流程的複雜度。此方式較適用於對影像品質要求高，且具備足夠人力及儲存容量條件的單位或計畫。

三、應用

所有數位影像其產生的目的即為應用，若影像無法應用，此影像的生命週期 (information life cycle)¹⁷ 恐將中斷。

數位影像產生後，經過調校，最後根據不同的應用需求而衍生適合的規格。中研院史語所考古數位典藏計畫所產出之衍生影像檔之應用以 C 檔與 D 檔占最大宗，其中尤以資料庫為主，本計畫建置之考古資料數位典藏資料庫影像 (<http://ndweb.iis.sinica.edu.tw/archaeo3/System/pages/searchSP.jsp>) 即開放 118,798 張提供一般民眾檢索與瀏覽，每年平均上站人數皆在 15,000 人次以上。此外，還應用於論文、教科書、雜誌、圖錄等出版印刷、電視節目製作、文創商品、舞臺演出等，自 2009 年至今授權之相關影像已超過 3,100 張。少部分超出 C 檔或 D 檔規格之影像需求，例如展覽所需之牆面大圖，則以 B 檔輸出（見圖 1），E 檔則主要做為機構內部使用之索引圖清單，另亦提供網頁使用。

詳細衍生檔規格應用之實例參見表 3。

史語所考古數位典藏計畫所產出之衍生影像檔應用範圍廣泛，可說明設立



圖 1. 「商王武丁與后婦好：殷商盛世文化藝術特展」展覽牆面大圖應用實例¹⁸。

¹⁵ 中研院資訊科學研究所開發之「多媒體中心」會將上傳之 C 檔批次降階轉為 JPEG 檔，以利史語所考古數位典藏計畫「考古資料數位典藏資料庫」(http://ndweb.iis.sinica.edu.tw/archaeo2_public) 之線上檢索及呈現。

¹⁶ <http://www.color.org/iccprofile.xalter> (瀏覽日期：2011/06/27)。

¹⁷ 資訊生命週期指的是資料從產生、保護、讀取、遷移、存檔、應用到回收的過程，有其發展週期。
<http://www.dlib.org/dlib/january00/01hodge.html>。

¹⁸ 影像由丁瑞茂先生攝影、提供。

表 3. 中研院史語所考古數位典藏計畫衍生檔與應用

應用項目		衍生檔規格	實 例
展覽	牆面大圖	B 檔	中研院史語所文物陳列館館內展示 商王武丁與后婦好：殷商盛世文化藝術特展 臺灣人類學百年特展
	說明圖	C 檔	
資料庫		C、D、E 檔	考古資料數位典藏資料庫 http://ndweb.iis.sinica.edu.tw/archaeo3/System/pages/searchSP.jsp 數位典藏與數位學習聯合目錄 http://catalog.digitalarchives.tw/
文創商品		C 檔	義美食品公司——手工薄片禮盒：紋鑑款 祥龍股份有限公司——十三行雕花杯盤
舞臺演出		C 檔	漢唐樂府南管古典樂舞團——殷商王 · 后
出版印刷	圖錄	C 檔	中研院史語所——殷墟出土器物選粹 故宮博物院——商王武丁與后婦好：殷商盛世文化藝術特展圖錄 日本教育圖書株式會社——書道 I、書道 II 中州古籍出版社——中原文化大典 · 文物典 · 古文字卷
	論文	C 或 D 檔	Huang, Ting-Yu and Lin, Yu-Yun, 2010. The application and presentation of the group-level cataloguing function in the Archaeological Artifact Database. <i>Museology Quarterly</i> , 24(3): 63-78.
	教科書、參考書	C 檔	南一書局、翰林出版事業股份有限公司
	雜誌	C 檔	經典雜誌
電視節目製作		D 檔	公共電視臺——發現十三行 大愛電視臺——史前踏浪人 日本 NHK 電視臺——The Culture of Chinese Characters in Asia 上海東方電視——文化電視專題
影片		D 檔	中研院史語所 80 週年紀錄片 故宮博物院——漢字書法
網頁		D、E 檔	世紀考古大發現：商王大墓重現 http://proj1.sinica.edu.tw/~dmuseum/ 中央研究院數位典藏資源網 http://digiarch.sinica.edu.tw/
索引圖清單		E 檔	內部使用

多樣之衍生檔案規格可滿足各式應用所需。

的來源多、版本多以及備份相關問題三大類，詳述如下：

面臨的問題與挑戰

中研院史語所考古資料典藏豐富，計畫執行時參與的人員眾多，由於資訊科技的推陳出新，設備的種類或機型多樣，累積 10 年的工作成果，所形成的數位資產相對龐大，也因而面臨種種管理上的困境，究其原因可歸納為數位檔案

一、來源多

史語所考古數位典藏計畫執行時間長，設備更新、參與人數眾多，造成影像產出來源多元，且產量大，不易管理。多種設備所產生的檔案格式各異，須分別以其所屬專用軟體處理，因此管理者須瞭解各種軟硬體操作，再加上軟硬體版本更新所造成的種種問題，增

加了管理者的負擔以及所耗用的時間，例如：IMACON FLEXTIGHT 949 掃描器和 HASSELBLAD IXPRESS 528C 數位機背原可共用舊版 RAW 檔專用軟體 FlexColor，然而新版軟體 Focus 只支援機背，若以 Focus 處理過的影像，其設定無法相容於 FlexColor。此外，由於數位攝影的特性所致，不同的攝影師即使使用相同設備，卻因參數設定依人而異¹⁹，無法求得一致，亦加重品質管理（尤

其是色彩管理）之負荷。本計畫人員人數、設備參見附件一。

由於影像生產來源多元，而且量大，為求品質一致性，必須有一套管控流程（張志光，2006），進行檢查與校對，蓋提研究所及 ADS 雖未明確建議一套管理流程，但蓋提研究所亦建議必須設立步驟以減少不同操作者及設備的差異性²⁰。史語所之考古數位典藏計畫巨量影像管理流程參見圖 2，並簡述如下：

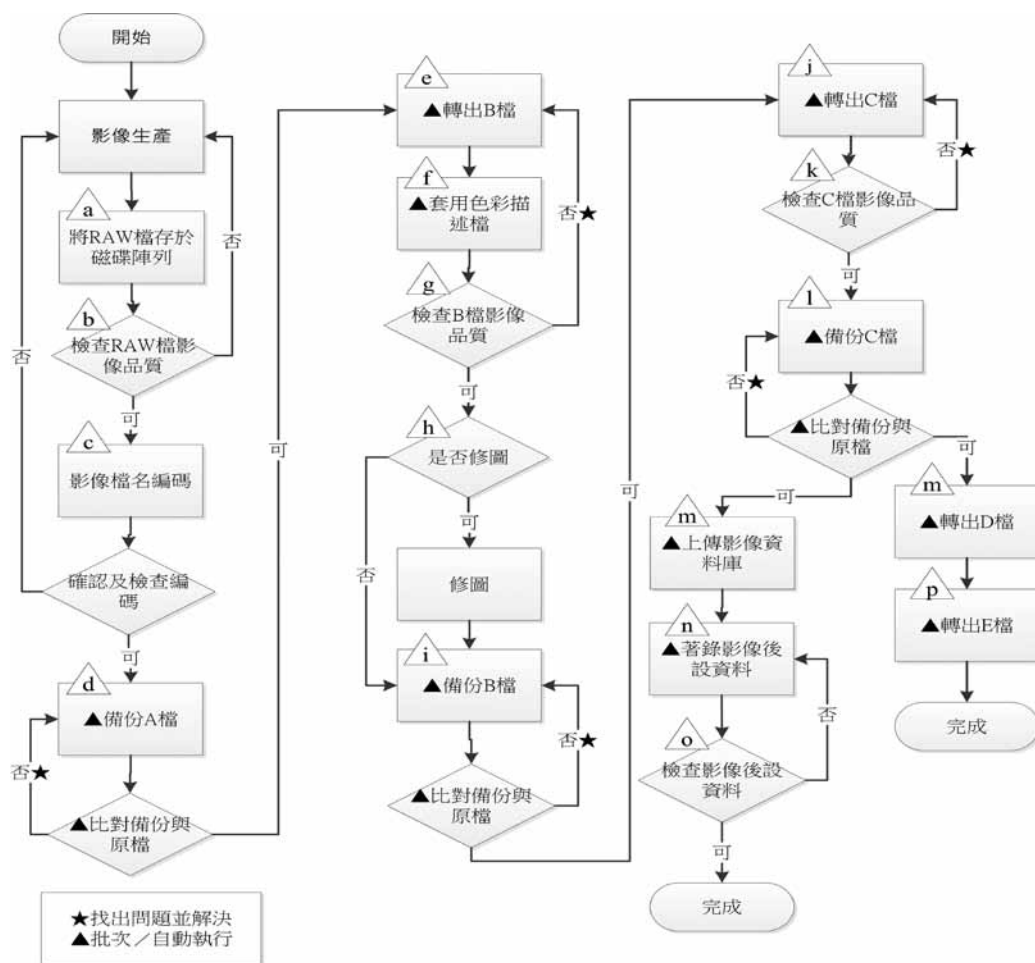


圖 2. 中研院史語所「考古發掘遺物、照片、記錄與檔案數位典藏計畫」巨量影像管理流程圖。

¹⁹ 以掃描進行數位化者設有一套參數（見本文「影像檔案之生產與應用」）。

²⁰ http://www.getty.edu/research/publications/electronic_publications/introimages/quality.html（瀏覽日期：2012/11/30）。

- a. 存入：經不同設備數位化後的圖檔以 RAW 檔（A 檔²¹）格式存入磁碟陣列。
 - b. 檢查 A 檔影像品質：以各 A 檔專用的影像軟體檢查影像品質，如有問題則請相關人員重新數位化。
 - c. 檔名編碼：將通過品質檢查的 A 檔進行影像檔名編碼²²。為避免檔名重複編碼，會參照影像資料庫中已有的相關圖檔資料。
 - d. 製作 A 檔備份：將編好檔名的 A 檔存入兩處磁碟陣列及 LTO 磁帶（備 2 份）。備份後，比對備份檔和原檔的檔名及大小是否相符，如不相符則找出原因排除後再重行備份。
 - e. 轉出 B 檔：完成編碼的 A 檔個別以其專用的影像軟體批次轉出相同尺寸 16 位元色彩深度的 TIFF 檔——B 檔。
 - f. 套用色彩描述檔：將 B 檔以 Adobe Photoshop 批次套用為其製作的色彩描述檔 (ICC Profile) 作色彩管理。
 - g. B 檔影像品質檢查：以 Adobe Photoshop 檢查轉出的 B 檔影像品質，若有問題則找出問題來源並重新轉檔。
 - h. 修圖：判斷是否需要修圖，如有需要，即以此檔作修圖處理。
 - i. 製作 B 檔備份：修完的 B 檔和不需修的 B 檔皆備份至磁碟陣列及 LTO 磁帶（備 2 份）。備份後，比對備份檔和原檔的檔名及大小是否相符，如不相符則找出原因排除後再重行備份。
 - j. 轉出 C 檔：B 檔以 Adobe Photoshop 批次轉出 A4 尺寸 8 位元色彩深度的 TIFF 檔——C 檔。
 - k. C 檔影像品質檢查：以 Adobe Photoshop 檢查轉出的 C 檔品質，若有問題則找出問題來源並重新轉檔。
 - l. 製作 C 檔備份：將沒問題的 C 檔備份至磁碟陣列。備份後，比對備份檔和原檔的檔名及大小是否相符，如不相符則找出原因排除後再重行備份。
 - m. 上傳 C 檔至影像資料庫並另轉出 D 檔：上述 C 檔分別傳至影像資料庫及以 Adobe Photoshop 批次轉出 A4 尺寸 8 位元色彩深度的 JPEG 檔——D 檔，作為驗收文件。
 - n. 著錄影像後設資料：將影像後設資料著錄於資料庫，完成後，系統自動連結圖檔以及影像後設資料。
 - o. 檢查影像後設資料並對照連上之影像：檢查後設資料以及連結上的圖檔，如有問題則找出原因並修正，或依需求重做。
 - p. 轉出 E 檔：以 Adobe Photoshop 批次以 D 檔轉出 8 位元色彩深度的 JPEG 檔——E 檔，作為索引圖清單。
- 數位化工作可視為大量生產系統的一種，此影像管理流程之設立規範實際執行操作步驟以及產出影像之管控流向，即便於計畫結束以後，機構內零星產出其他數位影像，亦能達到品質一致之目的。

二、版本多

由於所跨時間縱軸深，產出的影像版本多，史語所考古數位典藏計畫以影像編碼²³區分不同影像版本²⁴。影像版本產生的類型以及管理方式如下：

²¹ 本文所稱之 A、B、C、D、E 檔規格及說明請見表 2。

²² 影像編碼原則請見計畫入口網站：<http://archeodata.sinica.edu.tw/h2/h2d.htm> 之「數位影像檔規格及命名」。

²³ 影像編碼原則請見計畫入口網站：<http://archeodata.sinica.edu.tw/h2/h2d.htm> 之「數位影像檔規格及命名」。

²⁴ 英國 ADS 雖特別提到版本控制是極其重要的一環，然而其版本控制指的是多人同時處理同一個檔以及一個檔在多道處理程序下而產生的不同版本。（http://guides.archaeologydataservice.ac.uk/g2gp/CreateData_1-0（瀏覽日期：2012/11/30）），與本文多版本的情況不同，故不納入討論。

(一)前後期拍攝

1. 不同時期：遺物或遺址於不同時期所進行的拍攝。例如：銅爵 (R001070) (圖 3、4)，由於不同時期的拍攝具有紀錄性質，因此所有版本皆予以保留。

2. 不同目的：依研究或出版所需，

同一件遺物依不同目的而產生的不同版本影像，例如：乳丁紋簋 (R002069) (圖 5、6)，由於目的不同，不同版本所強調之處亦相異，因此所有版本皆予以保留。



圖 3. 早期拍攝²⁵。



圖 4. 數位典藏計畫 2002 年拍攝。

²⁵ 圖 3 到圖 12 之圖像皆引自中研院史語所考古資料數位典藏資料庫，分列如下：http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=R001070BBWA_6 (圖 3)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=R001070BCWK (圖 4)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=R002069BCWA (圖 5)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=R002069BCWA_2 (圖 6) http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=R026012CCWA (圖 7)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=ZR026012BCWA (圖 8)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=R015078_02CCWA (圖 9)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=R007132CCWK (圖 10)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=ZR007132BCWA (圖 11)
http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=ZR007132BCWZ_1 (圖 12)

(二)因研究進展而產生的版本

1. 甲骨綴合：甲骨經過研究不斷拼合，而不同階段的拼合影像，造成多版本問題，例如：帶卜辭龜腹甲(R026012)(圖7、8)，因此類影像皆同屬研究進展之紀錄，故全部予以保留。

2. 成套器物：考古遺物隨著研究進展，而發現成套關係之遺物，其分別有部件、單件及成套拍攝的不同版

本，例如：銅鑾(R007132)(圖11)，因田野發掘時已斷開，因而分別登錄為R015078_02(圖9)和R007132(圖10)；經研究，得知此二件應合為一件，是為銅鑾的一種，為西周早期鑾的前身，並與另一件(原登錄號為R007133及R015078_01)合為一套(圖12)(黃銘崇，2007)。此類影像亦皆屬研究進展之紀錄，故全部予以保留。



圖 5. 數位典藏計畫。



圖 6. 強調紋飾(圖錄出版)。



圖 7. 過去綴合。



圖 8. 新近綴合。



圖 9. 研究前之變鈴。



圖 10. 研究前之鈴身。



圖 11. 研究後單件。



圖 12. 研究後成套。

(三) 重複拍攝

同一地點或同一件器物於同一時間、同一方向重複拍攝之影像，例如水蛙堀遺址 TP2 探坑西牆幻燈片，共拍攝 2 張如圖 13、14。

影像經過數位化後只挑選最好的一張建檔，供資料庫及其他影像應用所需，但其他圖檔皆予以典藏備份。

總結版本多元的影像取決原則，凡影像具備歷史或研究脈絡者則全數予以

保留；若原件為減低紀錄遺失之風險而產生的重複影像，則選擇品質最好的一張轉製為 B 檔與其他衍生檔，其他予以典藏保留。

三、備份相關問題

此一龐大數位資產，為其安全考量，製作備份勢所必行。蓋提研究所建議影像應儘快進行備份，而 ADS 則建議最佳作法為每週和每月於異地備援²⁶。史

²⁶ http://www.getty.edu/research/publications/electronic_publications/introimages/longterm.html ;
http://guides.archaeologydataservice.ac.uk/g2gp/CreateData_1-0



圖 13. 水蛙堀遺址 TP2 探坑西牆(一拍)²⁷。



圖 14. 水蛙堀遺址 TP2 探坑西牆(二拍)²⁸。

語所數位典藏計畫則是依進度不定期隨時進行備份工作²⁹，將 A、B、C 檔備於本端（史語所）備援 2 處（磁碟陣列和 LTO 磁帶）及異地備援 1 處（中研院計算中心）。由以上 3 例可歸結一重點：備份工作必須及早且定期執行。

史語所考古數位典藏計畫之巨量影像於備份時面臨的諸多挑戰可歸結為幾點，討論如下：

（一）巨量檔案的存取

巨量的影像檔案於備份時須有一方便迅捷的管理架構，以提高檔案的存取效率，此架構亦關係到此一巨大數位資產之永續管理，簡明、具邏輯性的架構

方能經得起物換星移的考驗。ADS 亦根據考古資料的特性，建議影像要有符合特定專案的合理結構。檔案結構可以反映某特定遺址或遺跡或田野調查或攝影測量的不同例子³⁰。

由於考古資料的產生是考古學家於遺址進行調查之後進行發掘，在發掘時透過系統化的探方發掘方法，而產生一個個發掘單位（即探坑），進而發現地層中的人類遺留——遺跡與遺物，因此針對考古資料特性的存取架構，較為理想的方式，應是依考古系絡關係，將影像按其所屬遺址、遺跡、發掘單位分層，最後再按不同的原件類型區分，如圖 15。

根目錄	第一層目錄	第二層目錄	第三層目錄
考古資料	遺址	遺跡及發掘單位	遺物
			田野文件
			田野影像
	其他		

圖 15. 考古系絡備存架構。

²⁷ 感謝劉益昌先生提供影像。

²⁸ 圖像引自中研院史語所考古資料數位典藏資料庫 http://ndweb.iis.sinica.edu.tw/archaeo2_public/Include/GetImage_FileName.jsp?type=2&filename=S0802SWKPT2BCWA_17。

²⁹ 處理完成之影像直接備份於磁碟陣列，LTO 磁帶之備份則累積 200GB 之影像量方進行備份。

³⁰ http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_3

然而考古資料數位影像的存取架構仍須視各機構之資料情況及需求而定，尤其是田野文件及田野影像，有的因時代久遠，原件來源資訊不明確，恐無法依上述架構歸類。由於資料現況不同，史語所考古數位典藏計畫依大陸考古資料與臺灣考古資料而有不同的備存架構，見圖 16。

大陸考古資料變動不大，且經數次整理，脈絡清楚，除了遺物按照入藏登錄號設列，以便於典藏單位管理以及影像申請，田野文件與田野影像則先按遺址再按原件類型，再次按遺跡與發掘單位分層置放，其未按圖 15 之考古系絡備存架構之理想設計，乃因在實際操作上，考古系絡備存架構較為耗時，第一是操作步驟較多：考古系絡備存架構需要耗費的步驟比史語所現行之存取架構為多，尤其是單次存取的影像數量越大，操作步驟則越多，例如：要提取某遺址下墓 1 到墓 7 的田野文件及田野影像各一張，則史語所考古數位典藏計畫的架構要花 16 個步驟，而考古系絡的理想架構則須花費 21 個步驟。第二是上層目錄資料量較下層多：由於一個遺址的遺跡及發掘單位數量往往遠遠超過原件資料類型（遺物、田野文件、田野影像），

換句話說，遺跡及發掘單位的目錄夾數量比原件資料類型資料夾多，按考古系絡設計的架構會造成管理者於進入第二層目錄時，即耗費許多時間讀取目錄夾，而後才能進入第三層，此乃未考慮存取設備的性能所致；史語所考古數位典藏計畫大陸考古資料採行之架構將上層目錄量減至較低，以省去讀取的時間，因此在檔案存取，尤其是須跨遺址存取檔案時，可較快到達欲進入的目錄夾。因此，此備存架構無論是立基於永續管理或實務操作面，都可能是一個更為理想的方式。

臺灣考古資料多為仍在研究發掘以及調查的狀態，變動性高，主要研究者即為發掘者，因此架構第一層先按發掘者分，較符合檔案存取現況，並便利發掘者使用資料；第二層則按遺址分，再次按原件類型各有不同的分類架構，田野文件按原件資料夾名稱，而田野影像乃依發掘次分，此為依照臺灣考古資料現況而設計之架構，雖不完全按考古系絡，卻是管理此批資料最方便的模式。

（二）備份品質管理

由於製作備份的過程中會產生新舊版本、傳輸失敗、硬體損壞而造成檔案

大陸考古資料	遺物（以登錄號分，每千號一個資料夾 ³¹ ）		
	遺址	田野文件（按墓號／坑號分）	
		田野影像（按墓號／坑號分）	
	其他（其他類型影像）		
臺灣考古資料	發掘者	遺址	遺物（以登錄號分，每千號一個資料夾）
			田野文件（按原件資料夾名稱分）
			田野影像（按發掘次分）
		其他（遺址不明之資料）	

圖 16. 史語所考古數位典藏計畫備存架構。

³¹ 遺物以千號登錄號為一個資料夾之分類架構主要考量存取效能的問題，若分類號距過多，會對伺服器之效能造成極大之負擔，於本地端的管理者亦會耗費過多時間於清單之取得。

不一致的問題，為確保備份檔案品質，備份管理勢在必行，蓋提研究所提出兩種核對備份資料的方式：第一為定期校核檔案的位元數（檔案大小），第二採用校驗和 (Checksum) 核對³²。史語所考古數位典藏計畫兩種核對方式皆採行，除了備份檔案大小的核對，還另外增加檔名核對，以確保各處備份檔案內容一致，而校驗和僅用於長久典藏端之備份核對。因此，更理想的方式應還須將校驗和核對用於檔案上傳備份之初。然受限於人力以及備存處權限，無法進行。

(三) 跨平臺問題

數位影像自產出到備份，須跨 Macintosh、Windows、Unix、Linux 等作業系統，由於各系統對於檔名與編碼之限制不同，而造成存取檔案時之接受性的歧異，例如空格在 Windows 可被接受，但若在 Unix 系統則會視為斷句符號，因此影像編碼須考慮前述限制。此外，由於史語所考古數位典藏計畫典藏內容以中文為主，然而各作業系統皆以英文做為主要設定，因此以英文做為檔名較不易衍生其他問題，而這也導致在備份分類架構中必須將資料夾或檔名名稱譯成英文的狀況，然此一作法卻又造成名稱過長的困境（一般作業系統檔名長度限制不超過 256 字元），若僅使用簡寫則又不易辨識。目前史語所考古數位典藏計畫暫以簡寫處理，並另列清單標明其所代表的完整名稱。

(四) 設備更替快速

資訊設備世代交替的更新速度極快，所購買的設備大多過了保固期即已

停產且不一定有零件可供維修，為了確保資料不因設備之問題而造成資料毀損，每過幾年即須做備份資料的搬遷，造成管理上一大負荷。ADS 亦提到此一問題，蓋提研究所則建議，為長遠計，如果可能，每年或每兩年須定期檢查檔案的完備性及儲存媒體的穩定性，5 年左右即準備轉置計畫，但這都會加重管理負荷³³。此外，若將新舊軟硬體設備同時並存使用，則可能產生兼容性及效能的問題。以史語所考古數位典藏計畫為例，早期的備份設備為 LTO1 100GB 的機型，後來更新至 LTO2 200GB 的機型，雖然新機型可以向下相容 100GB 的帶子，但以新機型讀取舊機型的帶子則會耗費較多時間。

(五) 異地備援距離不足

目前史語所考古數位典藏計畫的異地備援主要使用中研院計算中心所提供的備援空間，雖然中研院計算中心也另外於臺灣大學天文數學館設置備援中心，然二者同屬於臺北盆地內，以全球的視野來看，如此的距離明顯不足，尤其臺灣是一個天災頻仍的地區，如能有更遠距的備援空間，可將區域性的災害問題造成資料損失的風險降至更低。

(六) 維護成本高昂

維持高品質的影像檔案勢必付出高成本的代價，所占用的實體、虛擬空間多，軟硬體設備、網路品質要求高，耗費的人力、時間、電費極鉅，維護網路安全、設備折舊等等問題也是一筆龐大的支出，史語所考古數位典藏計畫目前以印刷輸出需求作規範（見表 2），在

³² http://www.getty.edu/research/publications/electronic_publications/introimages/quality.html（瀏覽日期：2012/11/30）。

³³ http://www.getty.edu/research/publications/electronic_publications/introimages/longterm.html（瀏覽日期：2012/11/30）。

品質和成本之間取得平衡，雖然儲存單位成本不斷下降，但為了維持其資料的可靠性，相關成本仍然很可觀。

備份檔案的製作管理至為必要，卻也耗費人力時間甚多，發現問題檔案的後續處理亦極為費時繁瑣。備份多，檔案損失風險雖然降低，相對維護成本卻是提高的，以目前的科技而言，資料風險與維護成本成為無法分割此消彼長的因與果。

結論及未來展望

巨量數位影像定要有一套管理流程，以確保影像品質，並須具備一套有效、簡明的管理方法及架構，方能將數年來投資龐大人力、物力、金錢與時間所產出的成果，快速而便捷的儲存、找尋及提取，以供產出影像的後續應用。

史語所考古數位典藏計畫提出 10 年來摸索得出之經驗與方法，尚有諸多問題待改善，簡述如下：

一、自動化流程

由於所採行的影像品質管控流程繁複，共有 22-23 個步驟，雖然其中有半數步驟已經採自動化處理，然而卻無法全程連貫，例如：轉檔→檢查轉檔影像品質→備份→比對備份。未來若能朝向全套流程自動化管理，將可省去更多人力及時間的投注，並減少人為疏失。

二、RAW 檔格式

時下 RAW 檔格式分歧，各有專用之處理軟體，因此管理者必須安裝數種不同軟體並花費時間熟悉操作，方能處理所生產的影像。雖然市面上已有

軟體（例如：Adobe Photoshop 的插件 Camera Raw）可供處理幾種 RAW 檔格式，但影像讀入後的參數會與拍攝時的原設參數有異，造成色彩管理的困難。將來若有更多更成熟的跨格式軟體提供選擇之外，儘快訂立 RAW 檔的國際性規格，使各 RAW 檔軟體能夠支援轉存，此問題方能得到根本的解決。

三、人力成本

追求高品質的影像須付出龐大的人力成本，影像品質的管控有些無可避免必須以人工判別，例如轉檔程式或拍攝過程造成細微的影像瑕疵，恐僅能以人眼辨識，若能有程式協助檢驗，將可減輕負擔。

四、永續經營

國家數位典藏計畫於 2012 年結束，10 年來所產出的影像除須製作詳細明確的目錄以及清單（包括每一張影像的典藏路徑、檔名、副檔名、檔案大小）之外，此一龐大數位資產後續的管理、維運以及永續發展問題，尚須建立可持續維護管理與應用之機制。

誌謝

本文之完成首先要感謝中研院計算中心給予巨量影像儲存以及管理上之支援；並感謝本計畫「數位知識總體經營計畫：考古發掘遺物、照片、記錄與檔案」同仁給予許多寶貴意見，以及溫子雲先生協助英文摘要潤稿；而本計畫主持人陳光祖先生總是給予十分的支持是驅策本文的動力，我們致以最深的謝意；最後並感謝 2 位匿名審查者的寶貴意見。

參考文獻

- 徐明景，2011。數位化工作流程指南：影像資料。臺北：數位典藏拓展臺灣數位典藏計畫。
- 張志光，2006。數位典藏計畫影像品質管理之探討。第五屆數位典藏技術研討會。
- 黃銘崇，2007。商代的鑾鈴及其相關問題，古今論衡，17: 3-40。
- Archaeology Data Service and Digital Antiquity, 2011. Guides to good practice. Retrieved November 30 2012 from <http://guides.archaeologydataservice.ac.uk/#section-Contents-GuidesToGoodPracticeTableOfContents> .
- Besser, H., 2003. Introduction to imaging (revised edition). Revision edited by Sally Hubbard and Deborah Lenert. Los Angeles, CA: Getty Publications. Retrieved November 30 2012 from http://www.getty.edu/research/publications/electronic_publications/introimages/intro.html .
- Fraser, B., 2004. Understanding digital raw capture (revised 2004/09/04). Retrieved June 11 2011 from http://www.adobe.com/digitalimag/pdfs/understanding_digitalrawcapture.pdf .
- Hodge, G. M., 2000. Best practices for digital archiving: An information life cycle approach. D-Lib Magazine, 6(1). Retrieved December 15 2012 from <http://www.dlib.org/dlib/january00/01hodge.html> .
- International color consortium, Introduction to the ICC profile format. Retrieved June 27 2011 from <http://www.color.org/iccprofile.xalter> .
- Sustainability of digital formats planning for library of congress collections, Camera raw formats (Group Description). Retrieved June 01 2011 from <http://www.digitalpreservation.gov/formats/fdd/fdd000241.shtml> .
- Sustainability of digital formats planning for library of congress collections, JPEG. Retrieved June 01 2011 from http://www.digitalpreservation.gov/formats/fdd/still_fdd.shtml .
- Sustainability of digital formats planning for library of congress collections, TIFF, Version 6.0. Retrieved June 01 2011 from <http://www.digitalpreservation.gov/formats/fdd/fdd000022.shtml> .

作者簡介

- 溫子軍：中央研究院史語所「數位知識總體經營計畫：考古發掘遺物、照片、記錄與檔案」助理。
- 黃婷鈺：中央研究院史語所「數位知識總體經營計畫：考古發掘遺物、照片、記錄與檔案」助理。
- 林玉雲：中央研究院史語所「數位知識總體經營計畫：考古發掘遺物、照片、記錄與檔案」協同主持人。

附件一：史語所考古數位典藏計畫影像產生來源表

人 員	人數	類 型	型 號	掃描尺寸或像素	格 式
計畫掃描人員、外包廠商	5	大圖掃描器	CONTEX FSC3040	寬 1016mm	TIFF
		高階半滾筒式掃描器	IMACON FLEXTIGHT 949	35mm to 12*25cm film	FFF、TIFF
		高階平臺式掃描器	CREO EVERSMART SELECT A3	A3	TIFF
		平臺式掃描器	EPSON EXPRESSION 1640XL		
			EPSON GT-15000		
			EPSON GT-12000		
			EPSON EXPRESSION 10000XL		
			MICROTEK SCANMAKER 9800XL		
			UMAX POWERLOOK 2100XL		
			EPSON PERFECTION 1240U PHOTO	A4	
		EPSON PERFECTION 4870 PHOTO			
計畫攝影師、外包廠商、計畫助理	7	掃描式數位機背	PHASE ONE POWERPHASE FX+	10,500 pixels*12,600 steps	TIFF
			KIGAMO 8000+	8,000 pixels*11,320 steps	
		數位機背	PHASE ONE P45+	3,900 萬像素	EIP、IIQ、TIFF
			PHASE ONE P45		
			HASSELBLAD IXPRESS 528C	2,200 萬像素	FFF、TIFF
		數位單眼相機	FUJIFILM FinePixS3Pro	1,200 萬像素	TIFF
			NIKON D80	1,020 萬像素	NEF、JPEG
			NIKON D70s	600 萬像素	
		數位相機	SONY DSC-F717	500 萬像素	JPEG
			SONY DSC-F707		
			SONY DSC F505V	300 萬像素	
NIKON COOLPIX990					
實驗室技術人員	3	X光科學影像分析儀——影像讀取機	FUJIFILM FLA-5000	A4	IMG
		實體顯微鏡	ZEISS Stemi SV 6	依擷取設備而異	TIFF
		繪圖人員	5	繪圖及影像處理軟體	ADOBE PHOTOSHOP
向量繪圖軟體	ADOBE ILLUSTRATOR			AI	

The Management and Application of Massive Numbers of Images: A Case Study of the Archaeological Data Project, Academia Sinica

Chee-Kuan Wan* Ting-Yu Huang** Yu-Yun Lin***

Abstract

From the point of view of the professional image manager, the author discusses the problems and challenges of managing large amounts of digital images from different sources. These problems and challenges arise during image production, processing and file backup, as well as in the areas of image quality control, application and preservation.

The Archaeological Data Project of the History and Philology Institute, Academia Sinica has collected over 300,000 digital images requiring more than 10TB of storage space. Each year, 30,000 images requiring 4TB of space are added. Sources of digital images include photographers, data scanning personnel and outsourcing vendors which have created a huge management overload.

The Archaeological Data Project was established in 2001. Since then, digital technology has advanced at a rapid pace, leading to the development of different versions of digital images. In addition, with continuing research, there have been changes in the types of images of archaeological relics that are acquired, such as those following the rejoining of oracle bones or complete sets of artifacts.

In addition to the quality and accuracy of the data, completeness is an important issue for managing substantial digital assets. File backup frequency and administration have become essential components of image management. File classification structure of the collection and cross-platform problems need to be addressed to meet the demands of database building and application. The Archaeological Data Project is used as a case study in this paper to examine the above-mentioned challenges and to compare and discuss solutions provided in the digital imaging guidelines of the Getty Research Institute of the U.S. and the Archaeology Data Service of the U.K.

Keywords: management of digital images, image coding, digital image quality

* Assistant, Digital Archives of Archaeological Data, Institute of History and Philology, Academia Sinica; E-mail: cgilm@mail.ihp.sinica.edu.tw

** Assistant, Digital Archives of Archaeological Data, Institute of History and Philology, Academia Sinica; E-mail: tingyu@mail.ihp.sinica.edu.tw

*** Manager, Chinese Archaeology Collection and Co-director, Digital Archives of Archaeological Data, Institute of History and Philology, Academia Sinica; E-mail: hsyuyun@mail.ihp.sinica.edu.tw