

美國檔案知識庫發展經驗之研究

The Development of the Archival Knowledge Database in the United States

邱子恆

Tzu-Heng Chiu

通識教育中心副教授兼圖書館副館長
臺北醫學大學

Associate Professor, Center of General Education,
Associate Direct of the Library,
Taipei Medical University

蔣以仁

I-Jen Chiang

醫學資訊研究所副教授
臺北醫學大學

Associate Professor, Graduate Institute of Medical Informatics,
Taipei Medical University

【摘要】

檔案的目的在為未來創造知識，各國政府為了能成為二十一世紀的主宰，無不卯足全力，對過去的歷史檔案做有效的管理，為後人留下紀錄以供查閱，並為政府施政提供參考資料；此外，除了尋求現代科技保存政府檔案之外，更希望能發掘檔案中的隱性知識，使得這些歷史事蹟的里程，成為尋求最佳脈絡的解決方案。本文介紹美國檔案知識庫之發展經驗，包括其主管機構國家檔案暨文件署之組織架構，以及其主導的「電子檔案典藏計畫」，並在文末提出三點建議，期望藉由研究美國之相關經驗，尤其是其建立檔案知識庫的願景、目標、發展策略與需求規劃，以及其致力提供更多更完善之服務的目標，提供我國檔案主管單位寶貴的經驗，做為規劃發展檔案知識庫相關計畫時之參考。

【Abstract】

The purpose of archiving is to build up a knowledgebase for future reference. In order to play the leading roles in the 21st century, countries around the world are all effectively maintaining their historical archives; not only does it preserve records for future inquiry, it also will provide guidelines for political reference information. Besides, by archiving records using modern technology, we also expect to explore the tacit knowledge, set up milestone from these historical deeds, so that we can seek for the best solution more efficiently. This article presents the development of the United States archival knowledge databases, including the organization of its authority – the National Archives and Records

Administration (NARA), and its "Electronic Records Archives (ERA)" project. At the end of this article, after observing the United States archiving system, particularly on its vision, goal, strategy, and planning, as well as its efforts of providing more and better service, the authors propose three suggestions; and hope this article can provide our National Archives Administration some valuable recourses and references as we plan on our archiving and related developments.

關鍵字：知識管理；檔案知識庫；美國國家檔案暨文件署；電子檔案典藏計畫
Keywords: Knowledge management；Archival knowledge database；NARA；ERA

壹、前言

管理學大師彼得杜拉克（Peter F. Drucker）曾指出，開發中國家首要的經濟目標為知識創造力，誰先掌握誰就能統領二十一世紀的經濟；而審視人類的歷史，也不難發現哲人偉大的創見，都是我們用來建設今天、造就未來的一面明鏡，指引我們最佳的出路與方向。而這些知識的來源，就是人類經年累月留存下來的文化結晶，被記錄成文字，也就成爲所謂的「顯性知識」。

政府機關之知識的傳承來自於檔案，其不只反映政經社會的脈動，更具有歷史的使命，其所累積的量決不亞於網際網路上可供公開瀏覽的網頁。然而，這類檔案多數未經電子化，經歷的時間越久，將如同善本書一般，面臨保存的問題，因此電子化處理爲現今解決檔案留存最重要的不二法門。除了尋求現代科技保存珍貴的政府檔案之外，更重要的是要能發掘檔案中的「隱性知識」，藉以尋求檔案間的關連性，使得這些歷史事蹟的里程，成爲尋求最佳脈絡的解決方案。

各國政府爲了想成爲二十一世紀的主宰，無不卯足全力，對過去的歷史檔案做進一步有效的管理；除爲後人留下紀錄以供查閱外，並爲政府施政提供參考資料，以成爲後世之鏡。有鑑於此，世界各國的國家檔案主管機構，莫不以發展檔案知識庫爲重點。例如美國國家檔案暨文件署（National Archives and Records Administration，簡稱 NARA）爲讓檔案資料能結合現代化資訊科技，委託美國大學與研究機構進行資料數位化之研究；而歐洲地區如英國、德國及瑞士等亦不惶多讓。

NARA 的 Modern Records Programs 之主任 Lowell 認爲檔案的目的，是在爲

未來創造知識，他主張「檔案是機構的資產，其不但支援該機構現今業務上與法律上的需求，同時也是建立滿足未來目標的知識管理系統之基礎。」（註 1）所以 NARA 計畫於 2022 年處理完成約 347 Petabytes 的電子檔案資料（含文件、影音等），美國國會因此編列了一億美元的預算，而布希總統也另外於 2006 年編列了三千六百萬美元資助。面對如此龐大的文件檔案資料，已非傳統人力或傳統資訊技術所能處理，因此各國紛紛嘗試藉助現代化資訊技術來探求其可行性，NARA 也因此與麻省理工學院進行技術合作，執行資料檔案大規模數位化儲存與知識的整合管理，以爲後世留下歷史見證與借鏡。（註 2-3）

本文介紹美國檔案知識庫之發展經驗，包括其主管機構之組織架構，以及其電子檔案典藏計畫（Electronic Records Archives，簡稱 ERA），之後提出三點建議，期望藉由研究美國之相關經驗，尤其是其建立檔案知識庫的目標、願景、發展策略與需求規劃，以及其致力提供更多更完善之服務的目標，提供我國寶貴的經驗，做爲規劃未來發展方向之參考。

貳、美國國家檔案暨文件署

（NARA）之組織架構（註 4）

美國國家檔案暨文件署隸屬於聯邦政府，爲聯邦獨立機關（Federal Independence Agency），相當於我國部會層級之機關，直接向總統負責。該署是美國公民、公務員、總統、國會及法院確保其取得記載美國公民權利、聯邦公務員行爲及國家經驗的根本證明文件之便捷管道。其組織架構如圖 1 所示，業務職掌如下：

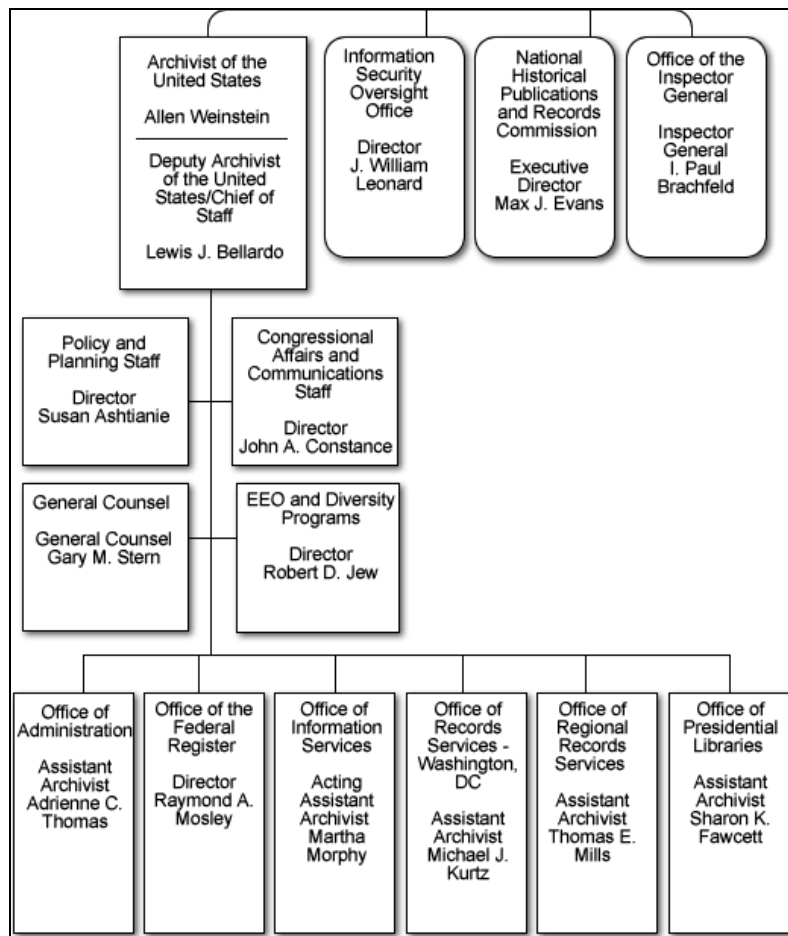


圖 1 美國 NARA 組織架構圖

資料來源：“Organization Chart, About the National Archives,” The U.S. National Archives and Records Administration, <<http://www.archives.gov/about/organization/index.html>> (27 September 2007).

- | | |
|---|--|
| <p>一、資訊安全監督司 (Information Security Oversight Office)：負責對總統提出有關政府網路安全管理的建議及報告。</p> <p>二、國家史料出版品暨文件委員會 (National Historical Publications and Records Commission)</p> <p>三、稽查司 (Office of the Inspector General)：負責改善 NARA 作業績效，並防止機關內部作業流弊發生。</p> | <p>四、政策暨計畫室 (Policy and Planning Staff)</p> <p>五、國會公共事務暨公關室 (Congressional Affairs and Communications Staff)</p> <p>六、法律顧問室 (General Counsel)</p> <p>七、公平任用與多元計畫室 (EEO and Diversity Programs)</p> <p>八、行政事務司 (Office of Administration)：負責行政事務工</p> |
|---|--|

作，包括：預算控制、空間安全維護、國家檔案管理信託基金等事務。

- 九、聯邦目錄司 (Office of the Federal Register)：負責出版總統文件、聯邦政府公報、法律、行政命令和規章條例等。
- 十、資訊事務司 (Office of Information Services)：負責 NARA 資訊作業計畫。
- 十一、華盛頓特區文件事務司 (Office of Records Services – Washington, DC)：負責華盛頓特區聯邦政府的文件鑑定、登錄、保存、描述及應用工作，下設檔案應用、典藏及現代文件管理等三部門，並下轄國會檔案中心，提供國會檔案服務。
- 十二、區域文件事務司 (Office of Regional Records Services)：負責各區域文件中心的規劃及管理，下轄全國 16 個文件中心。
- 十三、總統圖書館管理司 (Office of Presidential Libraries)：負責領導各總統圖書館業務規劃及執行，下轄 10 個總統圖書館及 2 個總統資料整理計畫。

NARA 爲了推動電子檔案典藏計畫 (ERA)，有長達五年以上的規畫期，因此特別成立了專案辦公室，其下主要分爲「程式支援」(Program Support Division) 和「系統工程」(Systems Engineering Division) 兩大部門，並包含其他的內控單位，負責整個計畫的執行與管控。此一專案辦公室負責處理八項主要業務，以支持 ERA 的最最佳開發：(註 5)

- 一、組織變遷管理 (organizational change management)
- 二、通訊 (communications)
- 三、環境狀態管理 (configuration management)
- 四、專案管理 (program management)

- 五、品質管理 (quality management)
- 六、研究 (research)
- 七、風險管理 (risk management)
- 八、系統工程 (systems engineering)
- 九、藉由專案辦公室有效的與承包商和合作夥伴建立溝通的運作模式，使 ERA 計畫能照預期目標順利執行，並配合美國政府電子化政府的推行，不致失焦。

參、美國國家檔案暨文件署 (NARA) 之檔案知識庫

NARA 的館藏涵蓋美國立國二百多年來的發展歷史，並管理包括國家檔案館、總統圖書館 (典藏歷代總統之各種手稿、文件等資料)、及在各州與華盛頓特區的 NARA 檔案館 (典藏政府重要檔案)，同時也是部分政府文件 (如聯邦公報與法規、美國法典) 的出版單位。(註 6) 依據 NARA 的變遷管理主管 (Change Management Officer) Fynnette Eaton 的看法，檔案爲「美國公民的權利、聯邦官員的作爲、國家經驗」(註 7)，因此 NARA 的首要任務與功能，就是保存這些必要的檔案，所以需要處理各式的資訊，諸如在國會立法過程中的記錄、法院案例、軍隊營運與人事記錄、各式研究資料等；而典藏這些資訊的媒體類型則包括紙本文件、照片、海報、圖片、動畫、聲音、視訊記錄、縮微膠片、其它微縮照相格式、地圖、空照圖、製圖、聯邦政府出版刊物、圖書館資料、以及電子檔案等。

檔案的價值在於其留存了組織性、功能性、決策性、程序面、操作面、及其他開創性的活動之實證，就因其具有實證的價值，值得進行檔案保存，以提供業務或法源上的服務，產生有價值的資訊，以進行知識分享。(註 8) 電子化檔案具有建立、控制、標示、保存、及

存取等重要功能，為長期保存與提供檔案之必要方式，因此為檔案的保存價值提供了加值的效果。然而電子化的過程將面臨：範圍（scope）、數量（volume）、變異性（variety）、與複雜性（complexity）等重大挑戰。如何指導政府單位管理檔案記錄，以促進任何形式的檔案記錄在被處理、選擇與典藏過程中的品質和效率，並且能永存保存，以能隨時隨地供民眾、學者與政府官員取用，誠為 NARA 的一大挑戰。

以檔案的數量來計算，美國在柯林頓政府時期，聯邦政府有 6 Terabytes（其中含四千萬封信件）、州政府有 0.5 Terabytes（其中含二千五百萬封交際文件）、國防部有 8 Terabytes（每年產生五千四百萬件影像）、以及 44 Terabytes 的戶政資料，且這些檔案都混雜不同的資料格式與多種資料內容性質，包含資料庫、地理資訊、照片等，成為數位化典藏的重大挑戰。在資訊時代，隨著政府部門對數位技術使用的增加，此挑戰轉變為 NARA 如何以數位化型式典藏各種類型的記錄，發展成可用於任一種聯邦政府之電腦平台的多元應用服務，同時能提供發掘與傳遞的管道，以將這些資產遞送到有興趣的人手中。為確保檔案典藏，並提供各種類型的電子檔案之取用，NARA 進行了一系列檔案知識庫的發展計畫，包括有「檔案管理先導計畫」（Records Management Initiative，簡稱 RMI）、「電子檔案管理計畫」（Electronic Records Management，簡稱 ERM）和「電子檔案典藏計畫」（Electronic Records Archives，簡稱 ERA）等。（註 9）

RMI 旨在重新設計整個聯邦政府的檔案管理，使其更具效率。ERM 為全美 24 個「電子化政府先導計畫」（E-government initiatives）之一，其目標在發展以民眾為導向的電子化政府與減少重複的系統建置。NARA 為「電子

化政府先導計畫」的主持單位，該計畫的內容分為信件管理（correspondence management）、全機構之電子檔案管理（enterprise-wide ERM）、電子資訊管理標準、以及永久電子檔案轉換等四部分。在電子資訊管理標準上，其採用 DoD 5015.2（美國國防部，全名為 Department of Defense）制定的標準做為所有政府機構的標準。另外，在檔案格式上也有考慮到標準化，如採用 XML 語法、Dublin Core 詮釋資料欄位等。ERA 計畫則是提供實作 RMI 與 ERM 結果所必須的資訊科技、資訊架構、與技術性的解決方案，其目標在於典藏並提供電子檔案的取用。ERA 系統主要包括輸入、儲存、輸出三大功能，並發展協助自動輸入、檔案目錄描述與參考服務的工具。同時，該系統需具擴張性，可以接受未來可能產生之任何格式的檔案記錄。由於 ERA 提供了 RMI 與 ERM 規劃成果的實作，因此在研究 NARA 之檔案知識庫的發展時，筆者就將焦點放在 ERA 計畫。

肆、電子檔案典藏計畫（ERA）之願景與目標

ERA 的主要目標，在促使 NARA 能保存並存取任何於聯邦政府體系下所產生的電子檔案。資訊科技的導入是必然的趨勢，而就保存而言，電子檔案長久典藏的問題之所以受重視，主要可歸納為以下五點：（註 10）

- 一、**規模(Scale)**：數位形式的政府資訊數量快速增長，且其可能的成長總量難以預估。
- 二、**多樣性(Diversity)**：數位格式具多樣性，因此在處理政府的電子檔案時，實際上是面對無限種類的數位格式。
- 三、**複雜性(Complexity)**：需處理的資料形態日趨複雜，除數值性的資料外，尚包括網路與關聯式資料庫、

地理資訊系統、各式網頁、虛擬實境、以及更多的複合式資料形態。檔案記錄的特質在於其“意圖”與“未遭篡改”，因此必須確保其真確性，因此對於複雜的電子檔案而言，必須設法保存其原始的內容、背景、結構、與行為等特質。

四、**持久性(Durability)**：軟硬體、資料格式與數位媒體易於過時作廢，因此必須發展出獨立於特定軟硬體的電子檔案典藏方案。

五、**變化性(Change)**：檔案的核心功能在於為現在與未來傳遞過去所發生的事實，然而電子檔案無法獨立存在，必須依賴特定的媒體，但資訊技術不斷更新變化，必須考量未來的使用者如何在使用最新的技術下，仍能夠發掘、傳遞與處理過去的記錄。

為回應上述問題的重要性，NARA 在其 1997-2007 年的發展策略中，將電子檔案的典藏列為第一優先，並於 1998 年提出了 ERA 計畫，以建立未來的電子檔案典藏機制。簡言之，ERA 計畫的目的乃在保存任何形式的電子檔案，並提供現在及未來的使用與存取。依照 NARA 之規劃，ERA 將於 2007 年開始進入實作階段。

美國檔案學家 John W. Carlin 曾提到「ERA 將切實的保存與提供使用者取用任何類型的電子檔案，使這些檔案獨立於特定的軟體與硬體，如此 NARA 在未來仍可以完成它的任務。」所以 ERA 是一個虛擬典藏知識庫的建立計畫，其願景為「提供一種綜合的、有系統的、且動態的方式，不論聯邦政府透過何種特定的軟硬體皆能確實典藏並提供取用任何類型的電子檔案」。(註 11) 為達成此願景，NARA 採取了下列行動：(註 12)

一、擔任電子檔案典藏之創新過程的領

導者。

二、與其他聯邦政府的工作夥伴溝通協調，發展方針與技術性指導原則，以促使電子檔案的建立與管理。

三、透過研究夥伴的協助，發展與維護技術性的功能，以記錄、典藏、描述、取用與適當的處理政府電子檔案。

四、管理具一致性、全國性與具水準的系統，以為聯邦政府電子檔案長久典藏之用。

五、在 NARA 記錄中心系統中，發展能夠管理聯邦機構電子檔案的功能。

六、確保任何人、在任一時間、從任一地方，皆可取用最好的工具以發掘並使用 NARA 所典藏的記錄。

七、NARA 的工作人員需有能力使用在每一個生命週期的電子化工具。

八、傾聽使用者的需求，並致力滿足、同時也尋求其回饋。

ERA 系統是一系列的功能與服務，讓 NARA 和其他政府機構與大眾能夠透過網際網路，隨時進行政府檔案記錄的管理、傳送、發掘和整合檢索。而此系統複雜之處，在於 NARA 規劃讓它能包括整個檔案記錄管理的生命週期，ERA 系統能藉由使用這些記錄被創作與建立過程中之活動的資料和資訊，支援所有檔案記錄的管理，同時也能實際地處理這些電子檔案。理想上希望不受限於現存之技術瓶頸，希望能隨著資訊技術的成長，而使檔案的保存與存取更達於預期，並善用資訊技術，在最佳效能與使用者服務上提供完美進展。

美國至今尚未對檔案知識庫有明顯具體的大型專案計畫，因此美國創新科技發展基金 (Advanced Foundations for American Innovation) 資訊與網路 2004 總統預算 (FY2004) 賦予 ERA 另一項重要使命，即是對檔案管理科技進行重點研究，包括不同檔案格式的相容互通

性 (interoperability)、索引協定、使用者介面；資料管理、儲存、與認證；網路通訊瓶頸 (networking bottlenecks)；以及長期典藏保存等研究議題。這些議題是目前各個領域的數位圖書館最迫切需要解決的技術，如此才能夠將這些現有知識之價值極大化，人類不再只能從這些知識看見已知的發現，而是可以從中發掘出更多潛在的、更有顯著進展的知識。(註 13) 因此 ERA 的重要研究發展方向，就是知識的發掘與分享，也就是如何利用資訊科技與資訊系統，與知識管理的理念相結合，而 ERA 就是美國 NARA 在此方面具體的實踐。

伍、電子檔案典藏計畫 (ERA) 之發展策略

ERA 之重要使命為支持 NARA 對電子檔案的策略性目標、豐富以生命週期管理之電子檔案知識庫、以及充份應用資訊技術以為檔案管理發展之所需。

(註 14) 茲說明 ERA 之計畫架構與策略架構如下。

一、計畫架構

ERA 發展策略成功的關鍵在於和其他機構與產業的合作。NARA 的基本策略並非在於發展獨特的技術以符合檔案或記錄管理的需求，而是希望在新技術中尋求發展電子化政府、電子商務、與下一代資訊建設的方案。有鑑於此，ERA 之合作夥伴與計畫即依此原則，以六個主要的合作計畫形成 ERA 計畫的核心，其分別為：(註 15)

- (一) 電腦科學與資訊技術領域的合作案－開放檔案資訊系統參考模式 (Open Archival Information System Reference Model, 簡稱 OAIS)：是由美國太空資料系統諮詢委員會 (Consultative Committee on Space Data

Systems) 所開創，而由 NARA 做先導的計畫，主要在研究任何需要長期保存資訊之系統所應具備的功能與單元，目前 OAIS 模式已成為 ISO 國際標準。事實上，OAIS 是一個參考模式而非實作指南，對 ERA 計畫而言，OAIS 提供了實體、功能、資料流程及行政活動的一個高階架構。

- (二) 檔案科學、記錄管理與資訊科學領域的合作案－電子系統長久典藏權威記錄國際研究計畫 (International Research on Permanent Authentic Records in Electronic Systems, 簡稱 InterPARES)：為一國際性的合作計畫，此計畫包括了十個北美、歐、澳、亞洲地區之國家檔案單位的代表，組成七支研究團隊，主要的目標在探討下列問題：
 - －研究何為不同類型之電子檔案權威性之典藏需求；
 - －定義保存檔案之選擇原則與實務；
 - －探討電子檔案之處理、輸入、輸出、控制及典藏機制；
 - －評估各種可能的技術；
 - －發展典藏之原則與標準。

InterPARES 的研究主要是建構在 OAIS 的基礎之上，它以 OAIS 為基礎，建立一個正式的保存模式，以便將記錄作權威性的保存。

- (三) 分散物件處理測試計畫 (Distributed Object Computation Testbed, 簡稱 DOCT)：除了 OAIS 及 InterPARES 外，ERA 也參加了美國國防部與專利標準局合作的 DOCT 計畫，此計畫關心的是永久保存檔案事先的建立、溝通、管理，以及高效能的電腦環境。DOCT 的研究結果後來由聖

地牙哥高速電腦中心 (San Diego Supercomputer Center, 簡稱 SDSC) 加以實作, 並為 NARA 進行物件一致性典藏計畫 (Persistent Object Preservation, 簡稱 POP), NARA 最終的目標不只在能夠實驗出長久保存權威記錄的技術, 還必需能將技術轉移, 使得每個機構都可以建立電子檔案典藏, 因此, POP 必須是可重複且每次的結果都要一致的機制。

- (四) 國家先進電腦研究合作計畫 (National Partnership for Advanced Computational Infrastructure, 簡稱 NPACI): 2000 年 NARA 加入由美國國家科學基金會 (National Science Foundation, 簡稱 NSF) 合作資助的 NPACI 計畫, 在此計畫中, NARA 支持其他的研究計畫, 以發展其物件一致性典藏計畫。
- (五) 總統電子檔案處理作業系統計畫 (Presidential Electronic Records Processing Operational System, 簡稱 PERPOS): 主要在探討、評估及發展先進的資訊科技以處理電子檔案, 這是一個以技術轉移為目標的合作計畫, 合作對象為美國陸軍研究實驗室 (U.S. Army Research Laboratory) 及喬治亞科技研究所 (Georgia Tech. Research Institute)。對 POP 而言, PERPOS 提供實證的研究結果, 而此實證研究著重於布希政府以來之總統檔案記錄的處理, 自 1998 年起, 此研究的重點主要在於某位總統卸任後, 該政府所留下的大量數位檔案應如何查尋及過濾。
- (六) 檔案管理者工作台計畫 (Archivist's Workbench

Project): 也是一個技術轉移計畫, 由國家歷史出版與記錄委員會 (National Historical Publications and Records Commission, 簡稱 NHPRC) 補助, 此計畫主要目的乃在將 POP 的研究結果規模化, 使之適用於較小的機構, 如州政府或大學檔案等。

二、策略架構

ERA 是一個持續性的計畫, 此計畫並不急於建立資訊系統, 而是先與其他機構合作, 進行一連串的研究, 等研究結果成熟穩定之後, 才開始發展系統。經過一連串的研究, ERA 已建立起發展策略和以 OAIS 為基礎的系統架構。此策略架構主要包括金字塔式的四個部分: (註 16-17)

- (一) 第一層 - 為最底層, 在以既有的基礎科技來支援電子政府、電子商務及相關研究, 這些基礎科技包括 XML 家族、整合與分散等技術。
- (二) 第二層 -- 在基礎技術下, 發展可以長久儲存及傳遞資訊的資訊管理架構, 如上述的 POP 之研究內涵; 這部分的研究可適用於所有的數位典藏機構, 如圖書館、博物館、檔案館等。
- (三) 第三層 - 此層架構主要針對各種不同檔案處理之特殊問題進行研究, 前述的 NPACI、InterPARES 等研究屬之。
- (四) 第四層 -- 是最後一層, 也將是花費最少力氣的一層, 針對 NARA 的需求進行的研究, 而 NARA 的需求主要緣於其儲存國家檔案、總統圖書館、聯邦政府記錄管理、立法及法規管理等任務。

在前述策略下, ERA 將架構在 OAIS

參考模式基礎之上，OAIS 參考模式定義了長久典藏資訊的通用系統架構，並假設該系統的資訊是由外界所建立，同時也將傳遞資訊給外界的使用者。就內部而言，OAIS 包括三個功能：將資訊攝取進入系統（ingest）、將資訊長久儲存（storage）、提供使用者查詢並傳遞資訊（access）。（見圖 2）在 ERA 中，則將

此三種功能轉變成三個虛擬工作台（見圖 3）：獲取工作台（Accessioning Workbench），負責將檔案記錄導入典藏庫；典藏工作台（Archival Repository），負責保存檔案記錄；參考工作台（Reference Workbench），負責提供檔案記錄的查詢與檢索。（註 18）

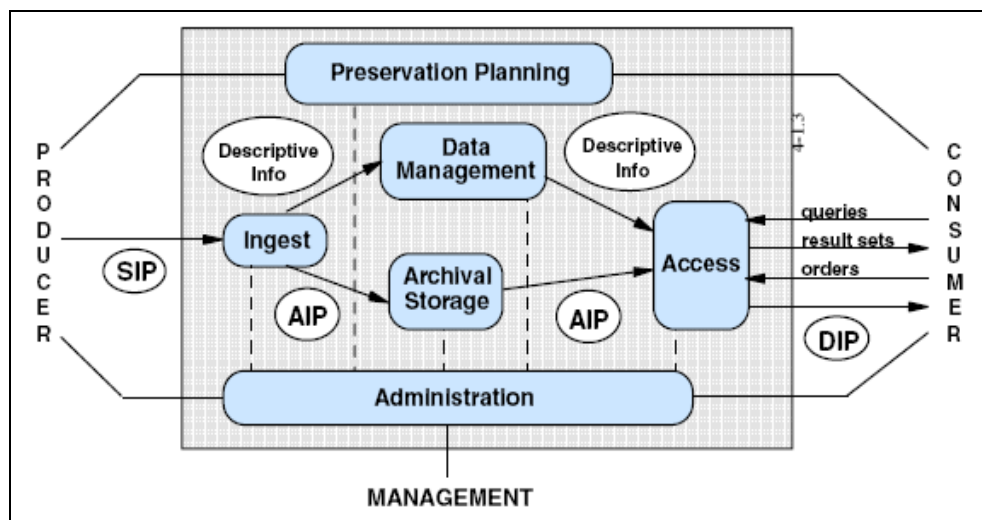


圖 2 OAIS 功能架構圖

資料來源：“Reference Model for an Open Archival Information System,”
<http://public.csds.org/publications/archive/650x0b1.pdf> (27 September 2007).

虛擬工作台的概念有幾個特點：每個工作台皆設計有嵌入式功能以處理該部分時常發生的工作。例如獲取工作台具備核對傳遞檔案記錄至 ERA 典藏者的身份(即圖 3 中之 Verify 功能)，同時必須能封裝欲儲存至典藏工作台的檔案記錄(即圖 3 中之 Wrap and Containerize 功能)，而參考工作台則必須能至典藏工作台取得檔案記錄，重新建立檢索結構並依序排置(即圖 3 中之 Rebuild 功能)。每個工作台皆需設計有能配合需求，促使特殊用途工具的應用之彈性。例如必須能接受不同媒體的輸入。由於 NARA

常接收不同媒體格式的電子檔案，但在機構寫入檔案和轉換媒體間常有時間上的落差，導致 NARA 可能得讀取過時的媒體。工作台之間由中介軟體做連結，讓整個系統不受某一時間點所使用之特定資訊技術的影響。例如某一工作台的軟體或硬體元件被取代後，整體系統功能只需在中介軟體處做修正以維持工作台之間的互通。整個虛擬工作台為功能性導向，各工作台皆有其任務與功能，而三個工作台使用共通的技術架構，在 ERA 內相互支持。（註 19）

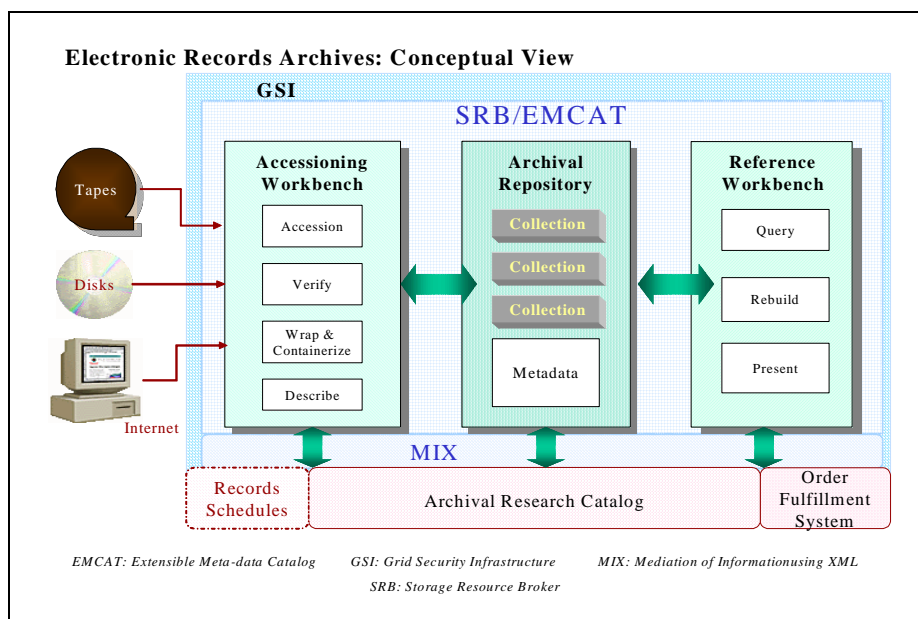


圖 3 ERA 系統架構圖

資料來源：K. Thibodeau, "Building the Archives of the Future," D-Lib Magazine, February 2001, <<http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>> (18 September 2007).

POP 架構將 ERA 虛擬工作台緊密的結合，典藏的物件可以是任何需要被典藏的數位資訊。其將物件轉換為一致的形式，提取物件的內容、背景、結構和外形特徵等資訊，將其嵌入到一個正式的模式中，例如個別的檔案記錄可根據 DTDs 模型化，集合性的檔案記錄則可使用 DTDs 或 XML schemas，而複雜的館藏性檔案記錄則可採用 XML Topic Maps。考量軟硬體易過時作廢的問題，透過 POP 架構，其將檔案物件轉化為獨立於任何特定技術外的中性格式，如此不僅利於長久保存，亦便於日後更先進的技術上市後，可以對檔案物件直接進行管理、檢索、取用與傳遞。(註 20)

歸納而言，NARA 對於 ERA 計畫的策略為：

(一) 和其他機構與產業合作，研究與解決重要的典藏議題。

- (二) 定義電子檔案生命週期之需求條件。
- (三) 以主流技術尋找解決方案。
- (四) 追隨下一代的國家資訊基礎建設。
- (五) 與美國政府的整體資訊科技方案緊密配合。

陸、電子檔案典藏計畫 (ERA) 之系統需求規劃

一、ERA 系統之規劃

NARA 於 1998 年開始定義其系統需求，2000 年開始做相關概念的探索，於 2003 年公佈 ERA 系統的系統需求書 (RFP)，2004 年底選定由 Lockheed Martin Corporation 及 Harris Corporation 兩家廠商競爭建置 ERA 系統雛型，2005 年底選定了 Lockheed Martin 公司，簽約

成爲 ERA 系統的開發廠商，並於 2007 年開始使用該系統，預定於 2011 年完成所有的功能，全面運作。(見圖 4、圖 5) 按照 ERA 的發展計畫，全案自 1998 年開始進行相關研究，規劃要歷時 13 年才能全部完成以爲運用，其發展進程大致可分爲以下四個階段：(註 21)

- (一) 1998-2000 研究階段 -- 由 NARA 尋求政府財源支持及尋找研究合作夥伴，透過研究結果發表檔案數位保存初始概念後，NARA 才正式成立 ERA 計畫專案辦公室。
- (二) 2001-2003 計畫發展階段 -- 研發 ERA 的系統需求，並提出系統

需求書的設計與相關標準。

- (三) 2004-2007 設計與建置 ERA 之實作初始階段 -- 初期選擇二家承包商各自針對 ERA 系統進行設計，一年後擇優其中 1 個團隊(即 Lockheed Martin Corporation) 進行後續的建置工作，至 2007 年需完成初步實作系統。
- (四) 2008-2011 實作完成階段 -- 爲實作的第 2 階段，採逐年增加系統功能，期能以先進的技術完成 ERA 的各項需求，預計花費 5 年的時間逐步發展，完整可用的系統預計於 2011 年建置完成。

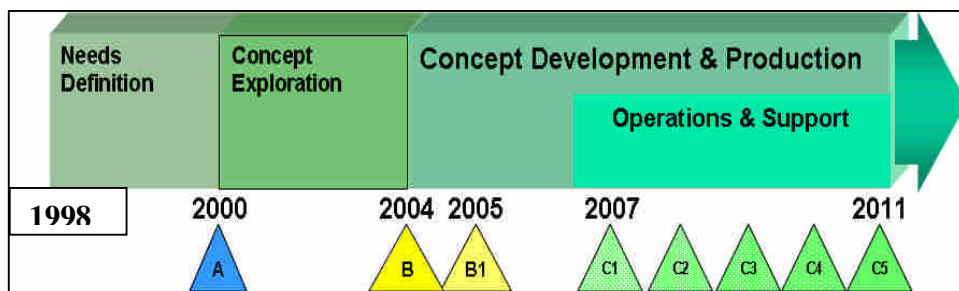


圖 4 ERA 開發時程規劃圖(I)

資料來源： K. Thibodeau, “Where in the World is ERA?”
<http://www.nccaiim.org/Downloads/NCC0501.ppt> (27 September 2007).

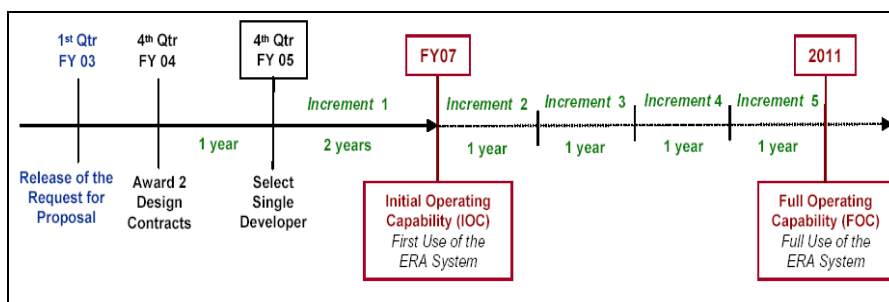


圖 5 ERA 開發時程規劃圖(II)

資料來源： F. Eaton, “Electronic Records Archives: Building the Archives of the Future: an Update,”
<http://www.archives.gov/records-mgmt/presentations/eaton.ppt> (27 September 2007).

而規劃中的 ERA 系統，應包括以下 的功能與範圍：(註 22)

- (一) 能處理所有檔案在各生命週期的執行。
- (二) 能建立、儲存與檢索所有的檔案。
- (三) 能處理與儲存電子檔案。
- (四) 能追蹤非電子檔案的處理程序或位置。
- (五) 能與其他系統連接，以提供其相關的生命週期管理資訊。(如圖 6 所示)
- (六) 不需提供將非電子檔案轉換為電子格式的功能。
- (七) 導入非電子檔案轉換為電子格式後的結果。
- (八) 能保證轉入後的電子檔案不受資訊技術改變的影響，仍能被取用。
- (九) 依處置協議之規定處理電子檔案。
- (十) 能執行電子檔案取用與傳佈的限制。
- (十一) 能儲存未分類、具敏感性、和依 Top Secret / Sensitive Compartmented Information (SCI) 分類的電子檔案。

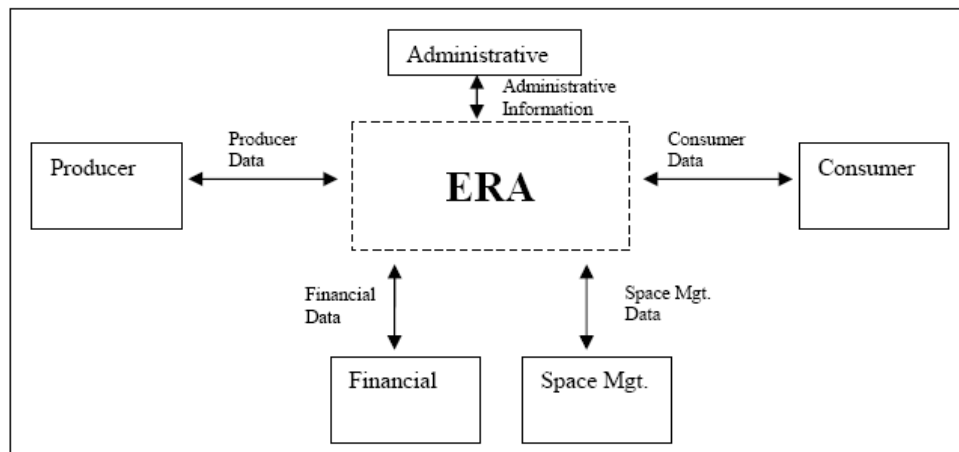


圖 6 ERA 與外部系統介面示意圖

資料來源：“ERA Requirements Document,” <<http://www.archives.gov/era/pdf/requirements-amend0001.pdf>> (27 September 2007).

二、ERA 之系統需求

ERA 系統在設計時，考量了下列六項原則：(註 23)

- (一) 基礎建設獨立：電子檔案典藏架構能獨立於任何其使用的軟硬體。
- (二) 模組化：當工作量與技術改變時，插入元件可被取代，並對既有元件之影響降至最低。
- (三) 規模化：可容納成長與管理不同
- 大小的儲藏庫，以及持續增加的檔案。
- (四) 擴展性：可處理未來新的電子檔案類型。
- (五) 綜合性：提供所有檔案類型在生命週期管理流程的支援。
- (六) 彈性：能讓 NARA 依其使用者需求，發展電子檔案服務，並能實作改進其業務流程。再者，ERA 系統應能與其他機構的各式系統

連結以進行檔案的轉換。

在前述六項原則之下，**ERA** 發展出其具體的系統需求，茲舉要如下：(註 24)

(一) 檔案管理

- 1、系統能管理檔案的配置。
- 2、系統能提供工作流程管理。
- 3、系統能提供描述功能。
- 4、系統能管理權威來源。
- 5、系統能提供管理檔案生命週期資料的功能。
- 6、系統能提供範例檔案的功能。

(二) 典藏

- 1、系統能提供管理樣版的功能。
- 2、系統能提供典藏電子檔案的功能。
- 3、系統能提供電子檔案整理的功能。

(三) 檔案儲存

- 1、系統能儲存資產。
- 2、系統能提供自動的媒體管理。
- 3、系統能依循開放的標準來處理媒體。

(四) 安全性

- 1、系統能顧及電子檔案的安全性。
- 2、系統能提供自身的安全性保護。

(五) 攝取進入(Ingest)

- 1、系統能接受所有種類的電子檔案。
- 2、系統能提供轉換電子檔到至 ERA 的功能。

(六) 取用

- 1、系統能提供檢閱資產內容的功能。

2、系統能提供編輯資產內容的功能。

3、系統能提供檢索資產內容的功能。

4、系統能提供取用資產內容的功能。

(七) 使用者介面

1、系統能提供客制化的使用者介面。

(八) 管理機制

- 1、系統能提供使用者限制的功能。
- 2、系統能提供使用者協助。
- 3、系統能提供和使用者溝通的功能。
- 4、系統能提供維護事件之 log 記錄。
- 5、系統能提供報表功能。
- 6、系統能提供系統管理功能。
- 7、系統能提供運算管理功能。

(九) 系統特徵

- 1、系統能管理使用者註冊服務。
- 2、系統能提供服務管理。
- 3、系統能符合或超越指定的績效需求。
- 4、系統能符合或超越指定的可得性需求。
- 5、系統能依 NARA 的企業架構做設計。

就美國政府而言，**ERA** 系統將促使 **NARA** 在未來能達成其任務，對於其他聯邦機構同樣亦具有其直接價值，主要的效益是在促進 **NARA** 改善對其它機構的服務，當各機關需要各單位的電子文件時，由於技術的先進，政府相關單位電子文件的獲取性可相對地提高。就政

府以外的單位之效益而言，ERA 系統將成為刺激各機關在未來創新發展數位資訊運用的主因之一，更進一步促使數位資訊於全世界的傳播更為便利。

柒、電子檔案典藏計畫 (ERA) 之現況

自從 NARA 選定標準後，於 2005 年底與 Lockheed Martin Corp. 簽約，將於 2007 年到 2011 年間陸續完成 ERA 系統之建置，預算共計約三億三千八百萬美元。(註 25)

ERA 系統於 2007 年開始運作，並可進行調整，其建置程序歷經多年的規劃，足見美國對於此項工程的重視。從初期研究、業務分析、辨識檔案需求、驗證現有系統、策略規劃等階段，現在已進入階段 F：設計檔案系統，即是進入實作的階段。(見圖 7) 整個 ERA 系統是由政府機構、國會、總統、以及捐贈者處轉入檔案，而一般大眾、企業機構、政府機關、以及聯邦政府可以由其參考工作台 (Reference Workbench) 取得資訊、滿足需求。(見圖 8)

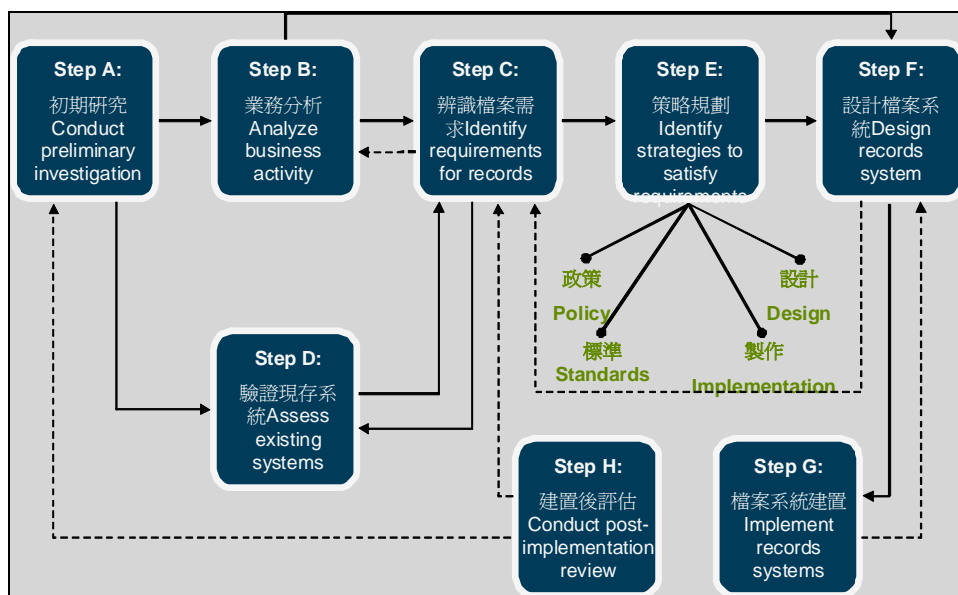


圖 7 ERA 系統之建置程序

資料來源：F. McGovern, “Development a Standard-Based Records Management Problem,”
 <<http://www.aiim.org/chapters/william penn/2005-12-13%20AWP%20Presentation%20-%20Records%20Mgmt%20Std.s.ppt>> (27 September 2007).

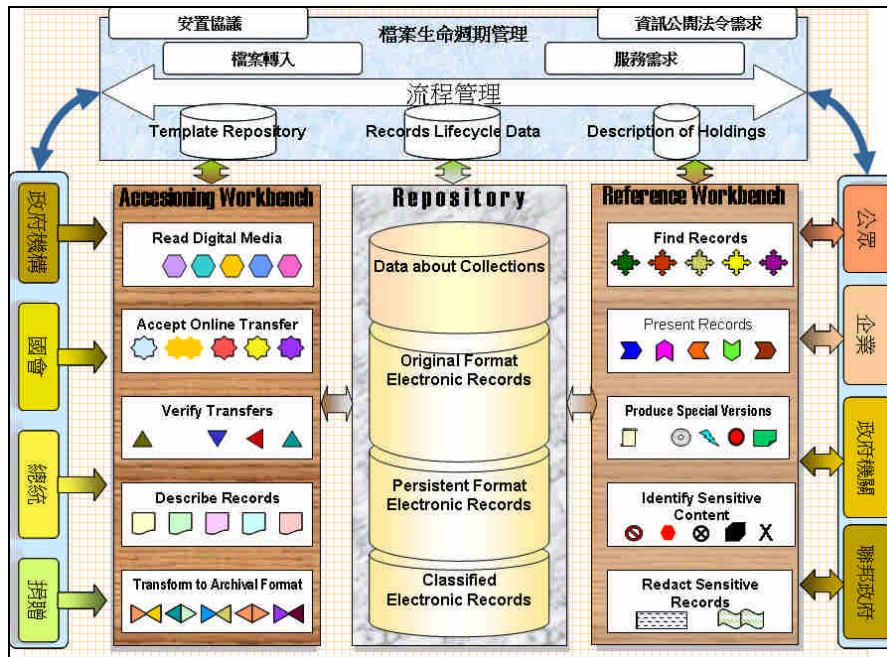


圖 8 ERA 之系統環境

資料來源： K. Thibodeau, “Where in the World is ERA?”
 <<http://www.nccaiim.org/Downloads/NCC0501.ppt>> (27 September 2007).

為能永久完善的保存電子檔案，NARA 亦定義檔案轉錄的格式，以利檔案之保存，而 PDF/A 正是 NARA 所選擇的機讀式電子檔案儲存格式。PDF/A 為 PDF 的子集，其符合 ISO 19005-1 之標準，具有與裝置無關、字型內含等優點，但強調文件須無法加密，相信此舉是為了避免加密易造成資料因密碼之疑忘疏失而無法解讀。(註 26)

檔案系統的發展沿革是有脈絡可循的，過程上從電子文件管理 (Electronic Document Management, 簡稱 EDM) 進

入到電子檔案管理 (Electronic Records Management, 簡稱 ERM)，再衍生成電子檔案保存 (Electronic Recordkeeping, 簡稱 ERK)，之後進入現在所謂的電子檔案典藏 (Electronic Records Archive, 簡稱 ERA)，最後終將成為知識管理 (Knowledge Management, 簡稱 KM)。(註 27) 傳統的檔案管理與電子檔案管理的明顯不同在於電子化的程序，其保留下值得儲存的檔案，而在自動化的程序下，檔案可自動的進行篩選，進而決定檔案本身的生命週期 (見圖 9-10)。

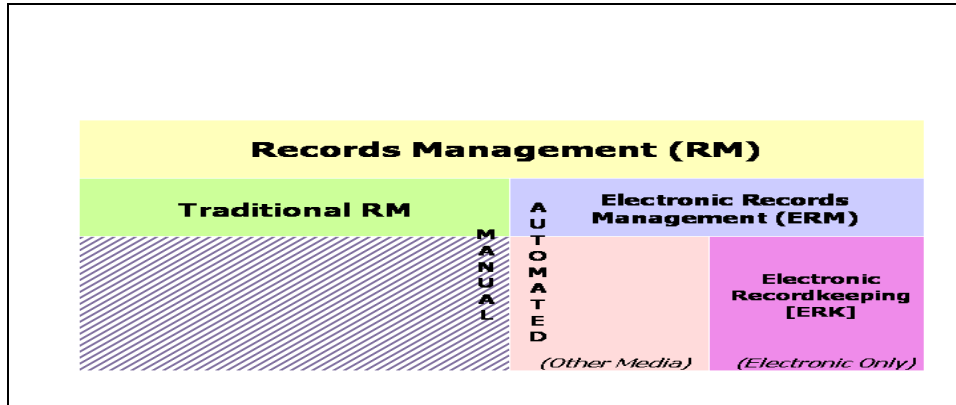


圖 9 檔案管理示意圖

資料來源：“Context for Electronic Records Management (ERM),” <<http://www.archives.gov/records-mgmt/initiatives/context-for-erm.html>> (18 September 2007).

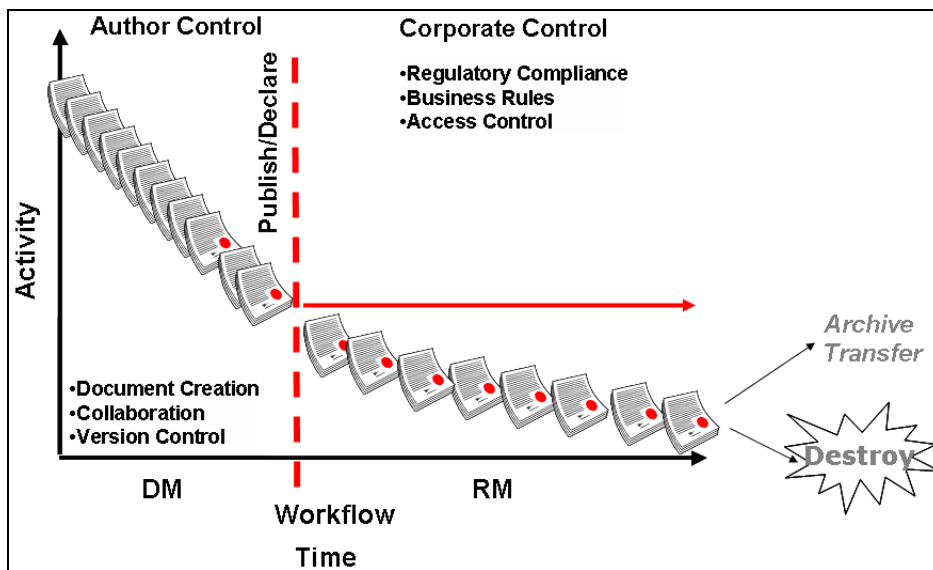


圖 10 檔案之生命週期

資料來源：Gunter Sandmann. 2006 年 6 月演講之投影片” Document / Record Lifecycle” <http://www.add.si/uploads/izobrazevanja/prezentacije_meridio/Meridio_GuenterSandann.ppt> (27 September 2007).

整個 ERA 系統可謂從電子與非電子檔案的上傳開始，各個機構自行上傳檔案、進行機器可讀式作業、存檔、註

釋資料的建置與萃取、自動化歸檔、到公眾閱覽檔案，其將於 2011 年透過資訊科技達成，隨著過程的進展，也將引入

新的或再行研發適用的資訊科技。

捌、結語與建議

美國檔案知識庫的發展以 NARA 主導的一系統相關計畫為主軸，包括前文提及的 RMI、ERM、和 ERA 等等，其中 ERA 是前兩個研究成果的具體實作，因此 ERA 是研究美國的檔案知識庫時之重點。為了推動電子檔案典藏計畫 (ERA)，NARA 特別成立了專案辦公室，而 ERA 的發展時程有 13 年之久，自 1998 年的開始進入研究階段，目前已進展到第三階段：系統建置之初始期，預計逐年增加系統功能並應用最新科技，在 2011 年正式建置完成。由本文可知，美國為了建置一個不受限於特定軟硬體之電子檔案知識庫，整個計畫歷時之久、資源投入之多，此一經驗實在可以做為我國的藉鏡。筆者認為國家級的檔案知識庫之開發，是需要周密且長時間、並且延續性的發展，而不是以“短打式”的一年或幾個月為期來建置雛型系統即可完成，在整個計畫的管考方面，也應成立專案團隊來負責，如此才能建置出可以永續發展、與時並進的檔案知識庫。筆者最後提出三點建議，以供國內主管單位規劃檔案知識庫時之參考。

一、重視檔案典藏之資訊安全標準

檔案典藏的安全與知識的授權，是檔案典藏管理的一個重大項目。美國 NARA 以 ERA 貫通整個檔案典藏的體系，但美國的 Government Accountability Office (簡稱 GAO) 曾針對其部分缺失提出稽核 (註 28)，其中重要的一項就是資訊安全問題。GAO 在其 2003 年的通報中指出 ERA 的資訊安全問題，認為 NARA 所提出的方案並未參考任何 IEEE 所制定的標準。所以這也是我國在

規劃檔案管理保存與知識共享時，所需嚴謹面對的課題。目前我國在此方面是依循檔案管理局提出的「資安生活化」、「資安系統化」、「資安關聯化」、「資安普及化」、「資安價值化」與「資安人文化」之思維與作法，參照 ISO27001:2005 的國際資訊安全標準認證要求為規範。

(註 29) 資訊安全標準分為 ISO27001:2005 及 ISO17799:2005，後者再分為 BS7799-1:2005 與 BS7799-2:2005；而 ISO17799:2005 於 2007 年 4 月後更名為 ISO27002。簡單的說，ISO27001:2005 的內涵就如同其標題「Information Security Management -- Specification With Guidance for Use」，是定義如何做好維持、監督、防範資訊死角等，保障資訊安全，以免除資訊漏洞；而 ISO27002 則是定義實現資訊安全防護的資訊作為，前者可說是管理面，而後者是作業面。因此筆者建議我國檔案管理局，除了規範管理的標準外，亦應訂定作業與處理程序之流程標準。

二、擴展檔案典藏之徵集內容

如前文所述，NARA 負責管理國家檔案館、總統圖書館、以及在各州與華盛頓特區的 NARA 檔案館，其同時也出版聯邦公報、法規、美國法典等政府出版品。但是其檔案知識庫中除了以上資源之外，亦將 2005 年 1 月 20 日前所有聯邦機關的 Web 檔案予以複製歸檔，其以 Heritrix 之 Crawler 開放原始碼系統，將 General Services Administration 所提供之 982 個開放與第二層限制等級的“.gov”，“.mil”網頁也典藏入庫。根據統計，光是從 2004 年 10 月 14 日至 2004 年 11 月 19 日短短一個多月的時間，就蒐集了將近 6.5terabytes 的資料 (分屬 50,000 多個政府機關的約 7,500 個網頁) (註 30)，由此可知網路上的資料的龐

大與複雜。美國如此大費周章儲存官方網頁資料，就是因為這些網頁傳遞最直接的政府資訊給民眾，亦是政府施政與決策的紀錄，故蘊藏豐厚的知識。因此筆者建議我國政府在建置檔案知識庫時，各級政府機關的網站與網頁，亦是值得思考的知識庫補充資料。

三、建置分類功能之查詢檢索

美國 NARA 所建置的 ERA 系統，將與美國政府的入口網站 First Government (<http://www.firstgov.gov/>) 整合，而此網站所採用的搜尋機制即是一種具即時分類分群功能的搜尋引擎，能根據查詢的詞彙所列舉出來的文件之內容，即時的與其他相關聯的重要關鍵詞組合，產生所列舉文件的分類分群。此一過程完全是即時性的，主要目的在避免人工分類作業的繁複與時間的耗費；更重要的是，為使即時分類分群的結果符合語意，協助使用者洞悉所得，該系統在分群前使用了控制字彙的辭典做為基礎。(註 31) 具有分類功能的查詢檢索，正也是我國建置檔案知識庫時已規劃的項目，從美國的經驗中得知，控制字彙的建立是最重要的關鍵，雖然相較於英文，中文的斷詞技術比較困難，然而就公文與政府用詞而言，定義出特定控制字彙相對來說是比其他領域容易且可能的。因此筆者建議我國在建置檔案知識庫時，應該對中文公文裡能掌握區隔的人名、地名、機關、與時間等詞作規範，俾使將來產生檔案詮釋資料時，能做有效的控制。

(誌謝：本文為檔案管理局委託筆者進行之「英美檔案知識庫發展經驗之分析」研究計畫 RDEC-NAA-資-095-005 的部分研究成果，特此誌謝。)

註釋：

1. “NARA’s New Records Management Workshops,” <http://eros.usgs.gov/government/records/Fall2004.pdf> (27 September 2007); “RACO West 2004,” The U.S. National Archives and Records Administration, http://www.archives.gov/records_management/training/raco_west_2004.html#presentations (27 September 2007).
2. D. Talbot, “The Fading Memory of the State, the National Archives Struggles to Ensure that Endangered Electronic Records will be around for as long as the original declaration of independence,” *Technology Review* (Cambridge, Mass.) 108 (July 2005): 44-49.
3. S. Garfinkel, “MIT’s DSpace Explained, Electronic Repositories Stretch to Meet Scholars’ needs,” *Technology Review* (Cambridge, Mass.) 108 (July 2005): 50.
4. “U.S. Government Manual: Browse the 2006-07 Edition,” GPO Access (22 August 2006) <http://www.gpoaccess.gov/gmanual/browse-gm-06.html> (23 September 2007).
5. 同前註。
6. The U.S. National Archives and Records Administration, <http://www.archives.gov/index.html> (18 September 2007).
7. “Repository design and creation: a case study, Trusted Digital Repositories for Cultural Digital Heritage,” Electronic Resource Preservation and Access Network, http://www.erpanet.org/events/2003/rome/presentations/presentation_Eat

- on.pdf> (18 September 2007).
8. “The 18th Annual Records Administration Conference (RACO),” The U.S. National Archives and Records Administration, <<http://www.archives.gov/about/speeches/2006/05-09-06.html>> (23 September 2007).
 9. K. Thibodeau, “Preserving Electronic Records: Developments at the National Archives and Records Administration,” <<http://www.archives.gov/era/pdf/thibodeau-040617.pdf>> (18 September 2007).
 10. 同前註。
 11. “ERA Vision Statement,” The U.S. National Archives and Records Administration, <<http://www.archives.gov/era/about/vision.html>> (18 September 2007).
 12. 陳昭珍，〈政府出版數位資源之管理、服務與典藏〉，《研考雙月刊》29卷2期（民91年4月）：6-18。
 13. “Networking and Information Technology Research and Development: Advanced Foundations for American Innovation,” National Coordination Office for Networking and Information Technology Research and Development <<http://www.nitrd.gov/pubs/bluebooks/2004/04BB-final.pdf>> (18 September 2007).
 14. 同註8
 15. 同註9
 16. 陳昭珍，〈國家檔案數位典藏面臨的挑戰與發展方向〉，《檔案季刊》1卷1期（民91年3月）：61-68。
 17. 同前註，頁66。
 18. K. Thibodeau, “Building the Archives of the Future,” D-Lib Magazine, February 2001, <<http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>> (18 September 2007).
 19. “ERA Requirements Document,” <<http://www.archives.gov/era/pdf/requirements.pdf>> (18 September 2007).
 20. 同註18
 21. K. Thibodeau, “Where in the World is ERA?” <<http://www.nccaiim.org/Downloads/NCC0501.ppt>> (27 September 2007).
 22. 同註19
 23. “ERA Concept of Operations (ConOps),” <<http://www.archives.gov/era/pdf/concept-of-operations.pdf>> (18 September 2007).
 24. 同註18
 25. M. Waltz, “ERA: Electronic Records Archives,” The Center for Research Libraries, August 16, 2006, <<http://www.crl.edu/content/DigArc/DigArc2/ERA%20profile.pdf>> (27 September 2007); “Congress Approves FY 2006 Budget of \$338 Million for the National Archives,” The U.S. National Archives and Records Administration, <<http://www.archives.gov/press/press-releases/2006/nr06-26.html>> (27 September 2007).
 26. “PDF/A: A new Standard for Long-Term Archiving,” PDF/A Competence Center, <http://www.pdfa.org/doku.php?id=pdfa:en:pdfa_whitepaper> (18 September 2007).
 27. “Context for Electronic Records Management (ERM),” The U.S.

- National Archives and Records Administration, (27 March 2000) <<http://www.archives.gov/records-management/initiatives/context-for-erm.html>> (18 September 2007).
28. “Information Management: Acquisition of the Electronic Records Archives is Progressing,” The U.S. Government Accountability Office, <<http://www.gao.gov/new.items/d05802.pdf>> (23 September 2007).
29. 檔案管理局全球資訊網，〈局長的話〉，(民96年7月24日)。
<<http://www.archives.gov.tw/Publish.aspx?cnid=131>> (民96年9月18日)
30. “2004 Presidential Term Web Harvest,” The National Archives and Records Administration, <<http://www.webharvest.gov/collections/peth04/>> (23 September 2007); “A Selected List of Web Page Presentation and Archiving Projects,” ResourceShelf, (30 September 2006) <<http://www.resourceshelf.com/2006/09/30/european-archive-foundation-launches-free-digital-library-list-of-other-web-archives/>> (27 September 2007).
31. “What is Clustering?” Vivisimo, <<http://vivisimo.com/html/clustering-about>> (23 September 2007); J. N. Hoover, “The Ultimate Search Engine,” InformationWeek, <<http://www.informationweek.com/story/showArticle.jhtml?articleID=201202986>> (27 September 2007).