

第十章

國家檔案數位典藏面臨的挑戰與發展方向

The Development Trend of National Electronic Archives and Its Challenges

陳昭珍

Chao -Chen Chen

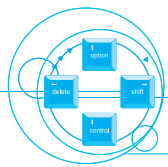
國立臺灣師範大學社教系圖書資訊學組 副教授

Professor, Graduate Institute of Library and Information Studies,

National Taiwan Normal University

壹、前言

檔案的典藏主要在提供全民原始(original)、權威(authentic)且一致的第一手文獻，這些文獻大部份來自政府機構。以美國國家檔案與文件署(National Archives and Records Administration，簡稱 NARA)而言，有三份最主要的原始檔案見證著這個民主國家的立國精神：即獨立宣言(Declaration of Independence)，憲法(Constitution)，人權法案(Bill of Rights)，每年有數百萬的民眾來自世界各地，到位於華盛頓特區賓夕法尼亞大道之國家檔案館(National Archives Building on Pennsylvania Avenue)參觀這三份文件（註 1）；此外，NARA 有十個總統圖書館，典藏著歷代總統各種手稿、文件等資料，在各州及華盛頓特區等地都有 NARA 的檔案館，典藏政府重要檔案。除了紙本及傳統的媒體之外，NARA 自一九七〇年收到第一份電子記錄（註 2），此後，紙本檔案並未減少，而電子記錄卻以等比級數不斷的成長，從柯林頓總統開始，總統的相關文件多了一種——即為數不少的電子郵件；自總統、議員、法官、到所有的聯邦機構，每天都在使用各種不同的電腦系統、數位媒體，及應用軟體處理他們的公務。而若就檔案的內容而言，NARA



所典藏的檔案：從法案到軍人個人檔案、從外交事務到個別意見調查、從鎮定劑的禁止宣告到美國的地形調查，可謂無所不包。在資訊時代，NARA 如何不因時間的移轉、技術的改變而持續長久的提供原始、權威、一致性的資料與服務給民眾？

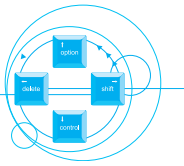
我國於一九九九年正式公布「檔案法」，並自二〇〇二年起正式成立檔案管理局(以下簡稱檔案局)，為政府資訊的公開邁開一大步。檔案局為積極建立機關檔案管理資訊化作業標準，提昇檔案管理效能，已訂定「機關檔案管理資訊化作業要點」，此要點明訂檔案管理系統應具備之功能、檔案儲存格式、公文欄位轉換格式、電子媒體規格、檔案目錄傳輸格式、機關檔案分類表傳輸格式、電子檔案格式。(註 3) 這份作業要點對於現階段檔案作業之處理提供了明確的規定，也相當符合目前的世界潮流，然而本文不擬從這些標準來探討國家檔案數位典藏之問題，而擬介紹 NARA 的電子記錄典藏(Electronic Records Archives，簡稱 ERA)、電子查詢計畫(Electronic Access Project，簡稱 EAP)及英國的檔案整合查詢計畫 (Access to Archives，簡稱 A2A)等計畫，供國內進一步思考數位典藏及網路存取國家檔案之相關問題與發展方向。

誠如 NARA 的 Kenneth Thibodeau 所言：“數位資訊衝擊著 NARA，其他機構也面臨相同的挑戰，但由於 NARA 獨特的角色與任務，以及政府計畫的多元性，使得此一挑戰更加艱鉅，然而，除非 NARA 能克服這些挑戰，否則在數位時代將無 NARA 這個機構。”(註 4) 我們深信這句話也適用於國內的情況。

貳、資訊時代國家檔案數位典藏面臨的挑戰

在美國國家檔案與文件署一九九七年至二〇〇七年的發展策略一文中指出，該局在資訊時代所面臨的挑戰主要有：

1. 紙本氾濫的情況在任何時代都不會停止；
2. 需要更多更好的空間來保存聯邦記錄；
3. 電子記錄之典藏面臨極大的挑戰；



4. 有越來越多的電子記錄需要保存；
5. 使用者越來越期待直接由網路上免費列印資訊；
6. 上述問題的解決方法，尚有諸多議題待研究。

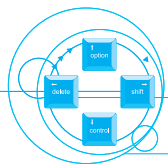
由上述挑戰，我們可以得知檔案管理不只需處理傳統的媒體記錄，也需將傳統記錄轉為數位記錄，此外，當代檔案更以新的媒體型態—電子形式，不斷的被生產出來，而這些電子檔案來源不一，所生產的軟體相當多元，因此檔案管理者面臨著如何將這些資源長久、正確保存的挑戰。保存電子檔案最大的挑戰在於：第一是數位化之後的檔案如何表現出原來實體檔案之間的關係，也是案件間彼此的關係；第二是如何將儲存於當代媒體、軟體、硬體的電子檔案，不因時間的轉移、技術的變遷、媒體品質的衰竭而長久保存到後世，以下茲分別說明之。

(一) 記錄與文獻的差異

保存檔案最大的責任乃在將權威性的記錄典藏並傳給下一代人。記錄是實際業務的文獻彙集，常有人將文獻視作記錄，如信件、報告等，但事實上記錄與文獻不同，記錄可以是任何形式的文獻，但記錄更重要的是要表現出其與實際活動的關連，如果只保存文獻，而未將關連性保存下來，記錄仍算是遺失了。因此，檔案學家一致認為記錄必需保存的屬性有三種：內容(content)、結構(structure)、情境(context)，簡言之，檔案必須按照其原來的順序加以保存。如果原件是數位形式，它也必須以數位形式將其內容、結構及情境保存下來，而且必須是長久的保存，可惜，到目前為止，並無有效的方法可以將三種屬性都保存下來，並保證該電子資源可以長久典藏。

(二) 數位資源保存不易的原因

數位資料不易長久保存的最大原因，乃在於資訊科技典範的轉移。所謂典範轉移是指某學科理論及作法產生結構性(基礎性)之重大改變，例如資料庫的設計從網路式資料庫、階層式資料庫轉移到到關連式資料庫或物件導向資料庫，這種轉移往往需要重新設計整個資料庫，以便轉移到新的典範。(註 5)而在資訊科技領域中，典範會改變的，包括軟硬體、網路架構等各種層面，且轉移的速度相當快；除了典範的轉移會造成大變動外，軟體系統的發展通常以三至五年一個週期，在新理念、新技術一點一滴的突破下，迫使典藏單位或一般使用者必需經常去“更



新(refreshing)”媒體，或”轉移(migration)”系統。

此外，數位媒體的壽命比起其他媒體而言，更屬易碎易損易變質的媒體，根據推測，磁片的壽命為五至十年，光碟的壽命則從十至一百年不等，但是由於讀寫技術或處理倍數的改變，事實上可以取用存在光碟中的資料之壽命，可能更短。

除了軟體、儲存媒體不易長久保存外，呈現介面也是電子資源的一部份，不同的呈現及查詢介面，都代表著對資料結構化的處理，表現資料庫內容的不同觀點，所以如果原來的觀點不能再建構，也就失去了對資料原來的組織觀點。綜合言之，數位典藏不易的原因主要在於：

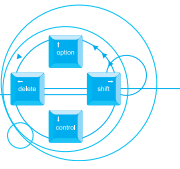
(1) 電子媒體的損壞或無用

前面曾經提過，電子媒體極易損壞，即使是壽命較長的光碟，也大概只有十至一百年的壽命，再加上軟硬體的生命週期頂多也只有三年到五年，原來的材料、設備、軟體、趨動程式等就從市場消失，而造成數位資料的無用。所以，一般而言，系統或資料管理者都已有共識，就是要做媒體的更新(稱為refreshed)，以及要在技術不見之前要做資料的轉移(migrated)。問題是無論媒體的更新或資料的轉移都是相當費力、費錢的事，當資料量少的時候還無所謂，如果數位資料已成為知識主要的儲存形式，這種金錢與人力的負擔恐非所有的單位都能負荷。此外，技術的轉移往往無法事先預測，使得我們很難在技術轉移前就做好資料轉移的動作。而即使數位資料轉移過去了，原來的詮釋資料(metadata)是否也能無誤的轉過去，並正確的結合在一起，都嚴重的挑戰著我們。

(2) 電子文獻無法獨立存在，必須與軟體結合在一起

數位資料之所以「存在」是因為有軟體可以去存取及顯示，二者唇齒相依。而最能讀取、解釋及顯示數位資料的當然是原始程式。然而這些程式的可用性又和儲存媒體及媒體讀取程式有關。所以，從典藏的角度而言，誰也不知道該如何用過時的原始程式去讀取未來的媒體，或用未來的軟體去讀過時的媒體。

為解決上述問題，NARA 在其一九九七年至二〇〇七年的發展策略一文中，將數位記錄的典藏列為第一優先，並提出數位記錄典藏(Electronic Records



Archives，簡稱 ERA)計畫，以建立未來的檔案典藏機制。

參、ERA 計畫簡介

ERA 是美國國家檔案與文件署自 1997 年以來所進行的計畫，此計畫規模相當龐大，也是 NARA 寄與厚望，希望藉此計畫得知數位檔案長久典藏之道，以下即詳細說明此計畫之目的及進行方式：

一、ERA 計畫之目的與進行方式

ERA 計畫之目的，乃在保存任何形式的電子記錄，並提供給現在及未來的使用者存取。到目前為止，ERA 仍是個研究計畫，且共有六個主要的合作計畫形成 ERA 計畫的核心，他們分別是：

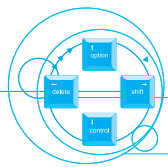
(一)開放檔案資訊系統參考模式(OAIS Reference Model (註 6))

開放檔案資訊系統參考模式 (Open Archival Information System，簡稱 OAIS)是由太空資料系統諮詢委員會(Consultative Committee on Space Data Systems)所開啓，而由 NARA 做先導的計畫，此計畫主要在明瞭任何要長期保存資訊之系統所應具備的功能與單元，目前 OAIS 模式已成為國際標準草案。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(entities)、功能(functions)、資料流程(data flow)及行政活動(administrative activities)一個高階的架構。

(二)電子系統長久典藏權威記錄國際研究計(InterPARES) (註 7)

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)是一個國際性的合作計畫，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，主要的目的乃在探討下列問題：

- (1) 研究不同類型之電子記錄權威性之典藏需求為何。
- (2) 定義保存記錄之選擇原則與實務。



- (3) 探討電子記錄之處理、輸入、輸出、控制及典藏機制。
- (4) 評估各種可能的技術。
- (5) 發展典藏之原則與標準。

Inter PARES 的研究主要架構在 OAIS 的基礎之上，它將以 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(DOCT) (註 8)

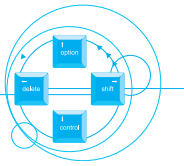
除了 OAIS 及 InterPARES 外，ERA 也參加了美國國防部與專利標準局合作的分散物件處理測試計畫 (Distributed Object Computation Testbed，簡稱 DOCT)計畫，此一計畫關心的是永久保存記錄事先的建立、溝通、管理，以及高效能的電腦環境。DOCT 的研究後來由聖地牙哥高速電腦中心 (San Diego Supercomputer Center，簡稱 SDSC)加以實作，並為 NARA 進行物件一致性典藏計畫 (Persistent Object Preservation，簡稱 POP)，NARA 最終的目標不只在能夠實驗出長久保存權威記錄的技術，還必需能將技術轉移，使得每個機構都可建立電子記錄典藏，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(NPACI) (註 9)

2000 年 NARA 又加入由美國國家科學基金會 (National Science Foundation，簡稱 NSF)合作資助的國家先進電腦研究合作計畫 (National Partnership for Advanced Computational Infrastructure，簡稱 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 (Persistent Object Preservation)。

(五)總統電子記錄處理作業系統(PERPOS) (註 10)

總統電子記錄處理作業系統 (Presidential Electronic Records Processing Operational System，簡稱 PERPOS)計畫主要在探討、評估及發展先進的資訊科技以處理電子記錄，這是一個以技術轉移為目的的合作計畫，合作的對象為：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，而此實證研究主要著重於布希政府以來之總統記錄的處理，自一九九八年起，此研究的重點主要在於某位總統下台後，該政府所留下的大量數位檔案應如何



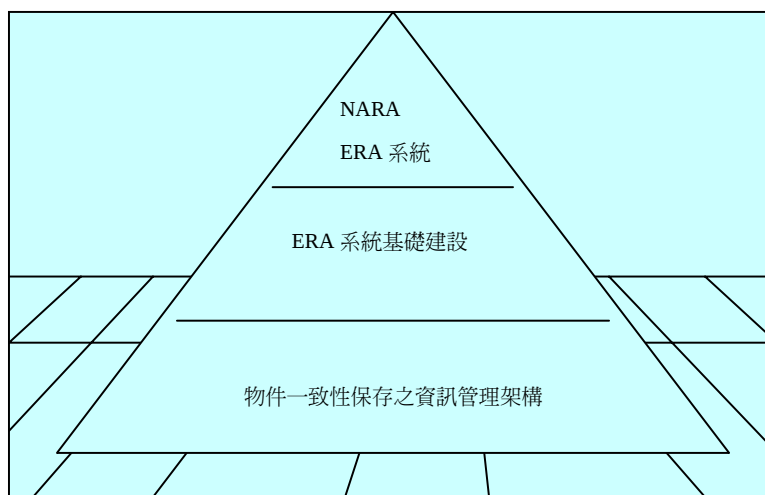
找尋及過濾。

(六)檔案管理者工作台計畫(Archivist's Workbench Project)

另一個技術轉移計畫是由國家歷史出版與記錄委員會 (National Historical Publications and Records Commission, 簡稱 NHPRC) (註 11) 補助 SDSC 的檔案管理者工作平台計畫(Archivist's Workbench project), 此計畫主要目的乃在將 POP 的結果規模化, 使之適用於較小的機構, 如州或大學檔案等。(註 12)

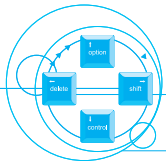
二、ERA 的架構與概念模式

ERA 是一個持續性的計畫, 此計畫令人佩服之處, 在於它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, 才要開始發展系統, 因此到現在為止, ERA 並未發展系統。雖然此系統尚未發展, 但經過一連串的研究, ERA 已建立起發展策略及以 OAIS 為基礎的系統架構, 其發展策略如圖一所示:



圖一 ERA 發展策略圖

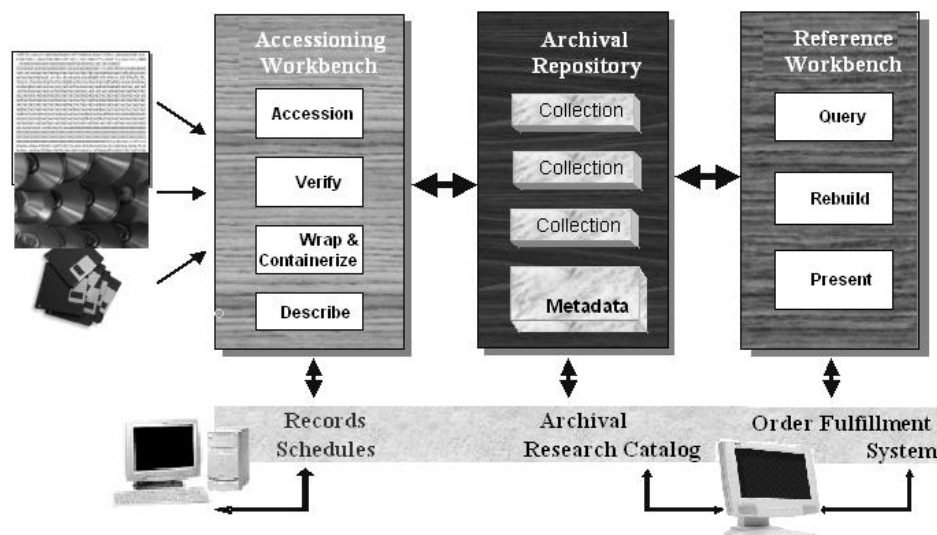
圖一內容主要包括四個部份: 第一個部份主要在以既有的基礎科技 (圖中方塊部份)來支援電子政府、電子商務及相關研究, 這些基礎科技主要包括: XML 家族、及整合與分散等技術; 第二個部份主要在基礎技術下發展可以長久儲存及傳遞資訊的資訊管理架構, 這部份的研究可適用於所有的數位典藏機構, 如圖書館、博物館、檔案館等, 如上述的 POP 之研究內涵; 第三層架構主要針對各種不同檔案



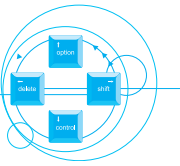
處理之特殊問題進行研究，NHPRC 的補助，NPACI，Inter PARES 的研究屬於這一層；最後一層也是將花最少力氣的一層，主要在針對 NARA 的需求而進行的研究，NARA 的需求主要緣起於其儲存國家檔案、總統圖書館、聯邦政府記錄管理、立法及法規管理等之任務。

在上述策略下，NARA 的解決方案會是如何呢？誠如前面提到的，ERA 將架構在 OAIS Reference Model 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及顧客。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)、提供使用者查尋並傳遞資訊 (dissemination)。在 ERA 中，則將此三種功能轉變成三個虛擬工作台(virtual workbench)：即登錄工作台 (Accessioning Workbench)，典藏工作台(Archival Repository)，參考工作台(Reference Workbench)。如圖二所示：

ERA Virtual Workspace Model

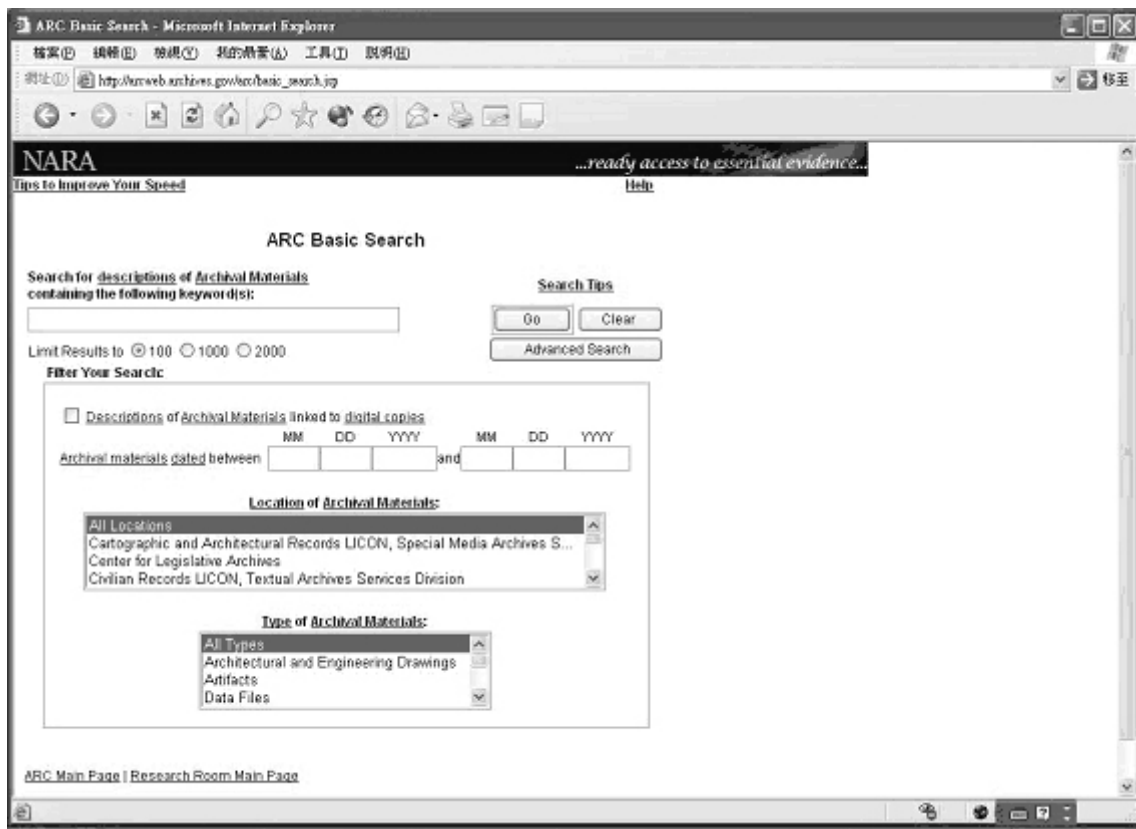


圖二 ERA 系統架構圖



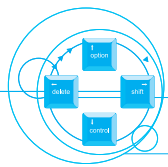
肆、電子查詢計畫(EAP)簡介

爲了便於民眾查詢美國國家檔案與文件署的文獻，NARA 的發展策略中還有一個重要的計畫，即 EAP（註 13），此計畫主要在建立美國國家檔案與文件署所有典藏機構之館藏虛擬聯合目錄，目前已有一個雛形系統美國國家檔案與文件署檔案資訊定址系統(NARA Archival Information Locator，簡稱 NAIL)可供民眾查詢，現已改稱 ARC。查詢畫面如圖三所示：



圖三 ARC 查詢畫面

此外，將部份重要及普及的手稿、照片、錄音資料、地圖、圖、及其他文獻數位化也是 EAP 計畫的一部份，目前已有一二四、〇〇〇件數位檔案可由 NAIL 查得，此計畫中也訂定了文字、照片、地圖與製圖、圖片等資料數位化時之永久檔、下載檔及提示檔之檔案格式。



伍、英國的 A2A 計畫

在歐洲國家，有關電子檔案典藏或檢索的相關計畫或系統也很多，其中英國的 A2A（註 14）是一個全國性的合作計畫，相關作法頗值得我國借鏡，茲說明此計畫內容於下：

A2A (Access to archives)計畫是英國國家檔案審議會(National Council on Archives)於 1998 年所提出之” Archives on Line: The Establishment on a United Kingdom Archival Network”理想的實現，此計畫主要之目標乃在建立虛擬國家檔案目錄，讓民眾可由此系統查到國家、區域及地方，從十二世紀到二十世紀的檔案。

A2A 建立後，主要的效益在於：

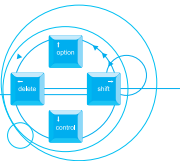
1. 民眾可由一個入口網站查詢所有的檔案；
2. 採用通用的標準來描述檔案；
3. 可讓民眾從家裡、圖書館、教室、辦公室等任何地方查詢檔案，提升檔案受重視的程度；
4. 讓不同地方的檔案典藏管理者可在此環境下一起工作，及服務其使用者；
5. 透過將既有的紙本目錄回溯轉換(retrospective conversion)，建立起檔案的經濟規模 A2A 的架構是由以公共文書館(Public Record Office)為基礎的計畫群所設計，所採用的標準則包括：

(1)資料描述標準：General International Standard for Archival Description (ISAD(G)), 第二版, Ottawa, 2000。

(2)名稱控制詞彙標準：International Standard for Archival Authority Records (Corporate, Personal and Family)，USAAR (CPF), Ottawa, 1996。

(3)個人、地名及團體名稱建立規則 Rules for the Construction of Personal, Place and Corporate Names, (National Council on Archives)，London, 1997。

(4)主題控制詞彙標準：IFLA Thesaurus.



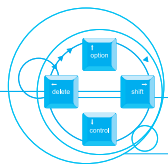
目前在 A2A 網站上共可查詢 117 個檔案機構的資料，其查詢畫面如圖四：



圖四 A2A 查詢畫面

陸、結語

面對資訊時代的需求，各國國家檔案都有相關的數位化或線上查詢計畫，本文特別以 NARA 的 ERA 計畫為例，主要希望借由其爲了長久、權威性典藏數位檔案所做的努力與深入研究，以提醒國內相關單位重視此問題的複雜度與困難度，而借由 EAP 及英國的 A2A 計畫，也讓我們知道在分散式的環境下，必需提供整合性的檢索，才能滿足使用者對於資訊系統查詢的需求。國內無論檔案電子目錄、檔案數位化或國家數位典藏計畫，目前皆處於初期發展階段，而 NARA 等的作法與方向非常值得我們借鏡。此外，數位檔案典藏不應自外於一般數位圖書館、數位典藏相關技術，而應積極參與或結合國內相關計畫或機構，進行各樣的研究與



測試，方能儘早解決數位檔案典藏所面臨的挑戰，提供便利、有效的檔案服務，以及將國家檔案正確無誤的永久保存給後代子孫。

【原刊載於檔案管理局出版之檔案季刊（九十一年十二月）第一卷第四期】

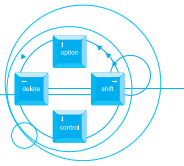
註釋：

註 1. National Archives and Records Administration. “ Ready Access to Essential Evidence: the Strategic Plan of the National Archives and Records Administration 1997-2000.” <<http://www.nara.gov/nara/vision/nara2000.html>>

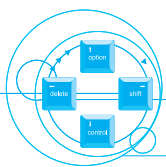
註 2. Kenneth Thibodeau. “ Building the Archives of the Future: Advances in Preserving Electronic Records at the National Archives and Records Administration.”
<<http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>>

註 3. 檔案管理局。機關檔案管理資訊化作業要點(核定本)及附表。

註 4. Kenneth Thibodeau. “ Building the Archives of the Future: Advances in Preserving Electronic Records at the National Archives and Records Administration.”
<<http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>>



- 註 5. Kenneth Thibodeau. “ Building the Archives of the Future: Advances in Preserving Electronic Records at the National Archives and Records Administration.”
<<http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>>.
- 註 6. ISO Open Archival Information System Standard.
<<http://ssdoo.gsfc.nasa.gov/nost/isoas>>
- 註 7. Inter PARES Program, <<http://www.interpares.org>>
- 註 8. R. Moore et al. “Collection-Based Persistent Digital Archives, D-Lib Magazine, March 2000, Volume 6 Number 3 (part 1)
<http://www.dlib.org/dlib/march00/moore/03moore-pt1.html> (part 2)
<<http://www.dlib.org/dlib/april00/moore/04moore-pt2.html>>
- 註 9. NPACI and SDSC Online, <<http://www.npaci.edu/online/v4.24/urbino.html>>
- 註 10. PERPOS (Presidential Electronic Records Pilot Operations System)
<<http://perpos.gtri.gatech.edu/>>
- 註 11. NHPRC <<http://www.nara.gov/nhprc/>>
- 註 12. Kenneth Thibodeau. “ Building the Archives of the Future: Advances in Preserving Electronic Records at the National Archives and Records Administration.”
<<http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>>
- 註 13. National Archives and Records Administration. “ The Electronic Access Project 1996-2001>
- 註 14. Welcome to A2A database <<http://www.a2a.pro.gov.uk/>>



檔案資訊資源管理
