

檔案管理局

電子檔案長期保存技術服務與監造案
國外電子檔案長期保存技術及運作現
況分析報告

中華民國 97 年 12 月

目 錄

壹、 前言	1
貳、 電子檔案管理策略及實施現況.....	2
一、 檔案管理體制.....	2
二、 檔案管理策略.....	3
三、 國外實施現況.....	6
參、 電子檔案長期保存相關技術之應用現況.....	10
一、 相關技術	10
二、 國外應用現況.....	14
肆、 結論與建議.....	49
一、 短期（98年）建議.....	49
二、 中期（99~100年）建議.....	52
三、 長期（100年以後）建議.....	52

圖目錄

圖 3-1 DOS Management Console	15
圖 3-2 英國國家檔案局積極性長期保存循環.....	16
圖 3-3 Active Preservation Management Console	17
圖 3-4 The Technical Registry PRONOM.....	18
圖 3-5 識別工具 DROID	19
圖 3-6 確認工具 JHOVE 測試結果.....	20
圖 3-7 The ERA 時程	22
圖 3-8 ERA 架構圖.....	23
圖 3-9 澳洲國家檔案局轉置工具 Xena 測試結果.....	37
圖 3-10 中國大陸中國科學院長期保存系統架構圖一	43
圖 3-11 中國大陸中國科學院長期保存系統架構圖二	43

表 目 錄

表 3-1 常見電子檔案格式..... 40

壹、前言

由於我國國家檔案管理制度發展較其他先進國家晚，過去多年有關檔案管理業務之國際合作交流，多係機關基於本身需求，各自作業。因此有必要藉此機會進行國外電子檔案長期保存工作比較研究，以探索檔案管理保存工作之共同規律，達成相互了解，並能正確地處理與評價國內外檔案管理保存工作的差異與關係。

首先，檔案管理之策略鑲嵌於檔案管理體制，檔案管理體制所指的是一個國家管理全部檔案及相關檔案工作，因此本規劃報告將由國外電子檔案管理策略及發展現況開始展開；再者，檔案管理局以電子檔案長期保存為重心，以建立政府資訊優質服務，永續保存政府電子檔案為宗旨，因此接下來主要探討國外電子檔案長期保存相關技術問題及其現況，以作為我國後續發展規劃之參考依據；最後，進行對我國電子檔案長期保存技術之影響分析，並提出具體建言。

貳、電子檔案管理策略及實施現況

檔案管理之策略鑲嵌於檔案管理體制，檔案管理體制所指的是一個國家管理全部檔案及相關檔案工作的管理方式和組織制度；包括設定檔案管理機構並協調彼此關係、劃分權限等機制，以此決定一個國家用什麼方式來管理自己的檔案和檔案工作。此外檔案管理體制的設定也取決於兩種因素：一個國家檔案管理的結構形式以及對於檔案管理業務範圍的認定。

一、檔案管理體制

(一) 取決於國家檔案管理結構形式

在 1960 年召開的第四屆國際檔案大會上，比利時國家檔案局曾透過對世界各國檔案事業管理實踐進行分析比對，提出檔案管理體制的兩種形式：集中式與分散式。

1、集中式

中央檔案機構對地方檔案機構進行監督與指導，且地方檔案館與中央檔案館或國家檔案館彼此相互構成一個完整的檔案館網絡體系。這種類型的代表國家為中國大陸及北歐國家等。

2、分散式

中央檔案館或國家檔案館只需保管自身檔案，並不過問地方檔案事務，地方檔案工作由地方政府自行指導及監督，並不隸屬於中央檔案機構，彼此之間沒有上下的從屬關係。這種類型的代表國家為美國、英國等。但是美國聯邦政府這一層級的檔案工作，具有以下特點：

- ❖ 所有聯邦政府機關的文件全部移交給國家檔案館 2 個總館及 12 個分館。
- ❖ 國家檔案館接收和保管各種載體檔案文件，有紙質檔案、影音檔案，也包含電子檔案。
- ❖ 國家檔案館對聯邦政府機關產生的檔案和文件，實行一體化管理。

(二) 取決於管理業務範圍

從檔案管理業務上，可以區分檔案管理體制為兩類：

1、文檔一體化管理模式

將文件管理納入檔案工作系統中，納入檔案管理體制，統一成一個有機整體，以便進行管理。這種模式從體制上說，就是賦予檔案行政管理機關指導文件管理之職，例如美國國家檔案及文件署下設文件管理處便是為了解決與日俱增的文件對於檔案庫房的壓力問題。

2、文件管理與檔案管理分別管理模式

將文件與檔案視為兩種不同性質的事物，分別納入兩個工作系統與不同的管理領域。由於文件生命週期理論強調，文件從始至終都是一個完整、動態、連續的生命運動過程。因此能否實行文件管理與檔案管理的一體化管理，主要在於對文件生命週期理論的認可程度，目前中國大陸可歸為此類模式。

二、檔案管理策略

世界各國的檔案管理策略多半依據自身的管理傳統、電子化政府策略以及文檔管理體制等，選擇適合自己國家的策略規劃。例如英國以架構政策框架為先導策略，明確訂定電子文件管理規範系體，強調目標管理與績效評估，以達成電子化政府及提供電子檔案線上整合服務功能。美國則以技術研發為首要，透過電子文件典藏策略計畫（Electronic Records Archives, ERA），期望成為電子檔案管理之領先典範，致力於研究電子文件長期保存的軟體系統，透過建立滿足文件功能需求的文件管理系統，得以實現電子文件之集中保管。澳大利亞則藉由檔案館的聯合行動，建立並推廣電子文件產生與管理的統一平台。

這些國家都似乎有著相同的經歷，從最初的自發研究分散探索到國家整體規劃，最後集中控制等過程。雖然採取的路徑不一定相同，但在策略面思考上，具有以下共同特徵：

(一) 高度重視電子文件管理

將電子檔案管理納入資訊化策略範圍，作為推進檔案電子化建設的重要內容。例如英國將電子文件管理納入負責資訊化政府工作的「e大使辦公室」（Office of the e-Envoy）職責範圍，同時制定及發布「電子文件管理系統計畫」（Electronic Records from Office Systems, EROS），由跨部門的資深管理人員組成工作小組，成員包含中央電腦及通訊局（The Central Computer and Telecommunications Agency, CCTA）等資訊技術人員，顯示其結合相關領域專家學者以及資訊技術人員等，共同制定電子文件管理策略。美國將電子文件管理列入其電子化政府行動方案，並納

入 ERA 計畫當中。澳大利亞則將電子文件管理作為政府績效檢查及審計的內容與依據。

(二) 以國家檔案館為集中控制中心

以國家檔案館為中心，形成電子文件管理和提供服務的網路體系，作為永久保管政府文件的專門機構。在早期電子文件的保管模式曾經經歷集中和分散相互博弈過程，目前國際的發展趨勢逐步傾向集中管理模式，集中管理具有永久保存價值的電子文件，成為許多國家目前的共識。集中控制型的國家的檔案館，集中管理政府各部門電子文件，並協調全國電子文件管理系統的跨平台互通，確實發揮其核心及樞紐的重要作用。例如美國國家檔案及文件署(National Archives and Records Administration, NARA)的 ERA 計畫，經過 6 年之久的基礎研究階段並耗資 3.08 億美元，規劃為期 10 年的系統研發階段，讓 NARA 得以接收並保存聯邦政府各部門所產生之各種類型、格式的電子檔案，為政府部門及公民提供便捷、有效的利用服務。ERA 採用開放式架構設計，其所開發之軟體系統將提供給各州使用，對於各州政府的電子文件管理將產生全面的示範效應。英國、澳大利亞、加拿大、新加坡、丹麥等國的國家檔案館也都已經或計畫接收政府部門所產生的電子文件。

(三) 以合作為基礎，大力推動基礎研究

合作是一項重要的管理基礎策略，各國的電子文件管理規劃都必須建構在相關理論和學術研究的基礎之上，因此結合其他政府機關、資訊工程專家、檔案學者、學術單位及私人企業等，提出合作計畫，對於基礎理論、標準研究以及國際合作等項目進行研究，其中重要堅實的理論基礎為紀錄連貫性模式(Records continuum model)和電子文件生命周期理論(the Life Cycle of Electronic Documents)，促使電子文件管理活動由機關層面轉向國家層面、從分散式管理轉向集中式管理模式、從環節式分離管理轉向無接縫式流程管理。英國、美國及澳大利亞的電子文件管理策略也都基於 ISO15489(檔案管理標準)、ISO14721(詮釋資料技術文件標準)、開放檔案信息系統參考模型(Open Archives Information System, OAIS)等國際標準制定。此外，一些跨國界的合作計畫，例如 The International Research on Permanent Authentic Records in Electronic Systems (InterPARES)，致力於研究電子系統中文件真實性、永久性之保障等專業技術問題，同時也提供國家層面之電子文件管理策略和標準參照框架。

(四) 重視檔案人才培育

由於對檔案學地位認識的不同，各國檔案教育所採取的形式也就各具特色。可分為以下兩種說明：

1、學院教育

把檔案學視為一門獨立的學科，擁有正規的檔案教育。綜觀世界各國，這一類型的教育有兩種形式：

(1) 正規檔案院校

如法國檔案學院、奧地利維也納檔案學院、俄羅斯國立人文大學歷史檔案學院、德國馬爾堡大學檔案學院以及中國人民大學檔案學院等，都是典型寓教育與檔案學研究於一體的教學研究機構。

(2) 檔案館附設的檔案學校

有些國家在檔案館附設了檔案學校，以檔案館的設施為基礎培養具備理論基礎又有實踐能力的檔案專門人才。如印度國家檔案館附設的檔案學院、德國普魯士國家機密檔案館附設的達萊姆歷史與檔案進修學院等。

2、培訓教育

這一類型的國家相對屬於少數，但在歐洲、北美洲和大洋洲都不乏其例；然而培訓教育的普及是世界性的，由於其實用性、針對性強，在有正規教育的國家中，每年也都會舉辦大量的培訓班。這一類型的教育也有兩種形式：

(1) 大學中設定檔案學專業課程

最具代表性的是美國、英國、加拿大和澳大利亞。美國的檔案教育通常是在歷史學、圖書館學碩士或哲學博士學位課程中，設定一系列檔案專業課程。英國也有少數大學開設檔案學課程，著名的如倫敦大學、利物浦大學等。加拿大開設檔案課程比較著名的大學有不列顛哥倫比亞大學、蒙特利爾大學和拉爾大學。其中不列顛哥倫比亞大學的電子文件管理研究在國際檔案界享有盛譽。澳大利亞開設檔案學課程的大學有新南威爾士大學、坎培拉大學、墨爾本大學、柯廷技術大學及蒙納士大學等。

(2) 國際檔案理事會和各國檔案工作者協(學)會舉辦的培訓班

國際檔案理事會(the International Council on Archives, ICA)

創辦過國際檔案講習班，並建立非洲、拉丁美地區和阿拉伯地區三個培訓中心，1979年於理事會內部設定了專業教育與培訓委員會。目前許多國家的檔案工作者協會，都將定期舉辦檔案培訓班視為主要任務之一。如英國、荷蘭、美國、加拿大及澳大利亞等國家的檔案工作者協會，每年都定期舉辦各種培訓班，除講授專業基礎知識外，還會依據工作需求和最新發展趨勢，及時加強培訓內容。

三、國外實施現況

美國、澳大利亞、歐洲等很多先進國家和地區的電子文件管理都不約而同地經歷了從機關層面自行研發到國家層面有組織規劃、從分散管理到集中管理的轉變過程，代表逐漸朝國際化、正規化和標準化的方向發展。由於篇幅所限，接下來將歸納說明美國、澳大利亞、加拿大、紐西蘭、荷蘭、英國及中國大陸等國家的電子文件管理發展狀況。

(一) 美國

美國的 ERA 計畫(2001~2011)，設計獨立的電子文件管理系統，集中管理各部門產生的電子文件。此項計畫的目標就是要將聯邦政府各部門所產生具有保存價值的電子文件，不論其具體的格式為何，都能夠通過系統間的無接縫移交到國家檔案館的 ERA 系統中進行保存，以提供政府和公民未來使用。

最近，美國國家檔案及文件署在官方網站上公布了其 2007—2017 年戰略規劃。該項規劃由法律依據、目前的情勢、要達成的目標、如何達到目標、如何知道已經成功、已經完成哪些成就以及附錄共七個部分組成。

(二) 澳大利亞

政府的一切活動都需要有真實、可靠的文件作為憑證，政府與公民之間的交流和溝通，通常需要借助電子文件來實現。澳大利亞政府將文件管理政策作為評價業務活動的一項重要指標，並將其納入政府的事務管理框架當中。

同時，澳大利亞針對電子政務活動的文件和資訊管理，提出且構建了 DIRKS 策略(Designing and Implementing Recordkeeping System, DIRKS)，採用事務活動分析的方法提出電子文件管理系統設計及實施的八個方法步驟，用以指導聯邦政府機構的電子文件管理活動。

此外澳大利亞是標準化管理的典型範例，它以記錄連貫性模式（Records Continuum Model）為基礎逐步建立出一套完善的電子文件管理制度規範系統，內容涵蓋電子文件管理所有主要環節，並體現了電子文件標準化管理的思維。這些制度標準包括其國家標準 ASISO 15489—2002、指導電子文件產生機構進行電子文件管理之聯邦政府工作手冊—文件管理系統的設計與實施（DIRKS）以及其他針對電子文件控制、檢索、保管、儲存及利用等環節制定的操作標準。以上這些制度標準對於統一協調管理聯邦政府各部門的電子文件管理發揮極重要作用。

（三）加拿大

加拿大國家圖書檔案館（Library and Archives Canada, LAC）是目前世界上唯一由圖書館和檔案館合而為一的國家級機構，於 2004 年 5 月 21 日經國家立法通過後創建。LAC 所保管的檔案文獻是加拿大和加拿大人歷史和經歷的唯一紀錄，被視為關於加拿大的資訊和知識源泉。

LAC 期望成為經得起各方面批評、能滿足各方面需求的成功學術團體，其下還設置有加拿大家譜中心、線上學習中心、資訊管理中心、會議中心、網路工作中心以及獨一無二的加拿大人肖像館。

加拿大國家圖書檔案館的檔案保存技術，在世界先進國家當中佔有一席之地，其對於殘缺文獻及蟲蛀檔案採用紙漿修復的保護技術措施，非常值得借鑒和學習；最近國家檔案館實施了一種稱作葉鑄（Leaf casting）的工藝來修復紙製文件。

另外，加拿大在電子文件管理研究議題上也有卓越貢獻，為解決電子文件長久真實性難題，加拿大大不列顛哥倫比亞大學和美國國防部合作的「保護電子文件完整性研究」（InterPARES），自 1993 年至今，已獲得顯著成就。

（四）紐西蘭

紐西蘭於 2005 年 4 月 14 日通過公共檔案法案（Public Records Bill），取代 1957 年通過之檔案法（Archives Act）以致力推動檔案管理現代化。

Continuum 是一個完全由紐西蘭政府推動的文件保存計畫，這個計畫的目的是要提供最有效率的工具及服務給政府機構，以使其達到文件保存標準的最佳狀態；而且這個計畫也能夠協助各機關發展自己所需要的系統，以滿足業務及歸責需求，同時提昇

文件管理品質，因此政府機關重要的文件得以為現在及將來後世子孫保留下來。

Continuum 也是一個整合系統，企圖於紐西蘭國家檔案館（Archives New Zealand）與其他政府機關之間建立穩固、彼此合作及成熟的互利夥伴關係。

同時，Continuum 計畫也提供執行要點及最佳實踐給 2005 年所提出的「公共文書法案」（Public Records Act）。

(五) 荷蘭

歐盟委員會根據歐盟各個會員國的實際情況提出若干關於數位資訊管理的計畫，其中「數位歐洲」（e—Europe），被視為最具影響力，該計畫致力於為所有歐盟成員國建立統一的數位資訊交流平台，並保證電子文件資訊得以無障礙傳遞。配合此一計畫，荷蘭特別制定由國家圖書館負責的國家數位計畫。

從 InterPARES 計畫內容加以參考引用，荷蘭選擇了幾項特殊類型的電子檔案作為研究測試對象，如電子郵件、文本文件、Excel 資料表、資料庫等，研究如何對這些文件進行長期有效保存，並根據實驗結果來制定最佳的實驗策略。此項實驗同時比較了三種保管數位資訊的方法：轉置（Migration）、模擬（Emulation）和 XML，並探討這三種方法在管理不同類型電子文件時，其優劣及可行性。

此外，荷蘭國家圖書館所主導之 e-Depot 計畫，在數位保存領域上成為一個開創者，幾乎所有大型的科學期刊國際出版商都參與該計畫，目前已包含 1000 萬篇科學文章。

(六) 英國

英國提出「支持電子政務的電子文件管理政策框架」，將電子文件管理與電子政務緊密結合，在英國 95% 政府文件都以電子文件形式產生的情況下，電子文件管理將被納入電子政務建設的框架，成為另一種國家行動力。

國家檔案館發布一項新的網站策略，宣示在未來三年內將會如何進行一系列改造：

1、新的入口網站

國家檔案館的網站將被改造成單一入口網，藉此反映機關提供的各種功能，並滿足使用者不斷變化的需求。同時確保網路服務不中斷。

2、數位服務網路

提供資訊以及決策支援能力，透過網際網路將提供文件的網路轉變成為數位服務網路。這被認為對英國國家檔案館網站的未來具巨大影響，根據官方表示，期望使更多的資訊可以被搜尋並能夠重新獲得使用。

3、了解使用者

國家檔案局意識到需要尋求更好的方式來了解使用者及滿足他們的需要，同時不斷地將有限的資源經由服務遞送給其他政府機關及公民使用。同時建立一個雙向的交談機制，努力實現此一機制，促進交流，發展社交網絡服務（Social Networking Service, SNS）。

4、推動變革

提供資訊以及決策支援能力，為網路服務提供了一個決策支援框架。

(七) 中國大陸

中國大陸以綜合檔案館為主，在中央一級也設有專門檔案館；地方各級（省、自治區、直轄市）設有綜合檔案館，以及各縣(市)檔案館。全國各級國家檔案館分別由中央和地方各級檔案行政管理部門於業務上實行監督及指導。由於幅員廣大，層級複雜，因此較難一窺其政策全貌，

中國大陸最近正展開從上而下部署、指導全國檔案館，推動「圍繞電子政務建設，推進檔案信息化建設」，其內容主要在於推動電子政務建設的平衡發展、促進電子政務內外網建設、公共資料庫建設、檔案資訊化成果應用以及檔案資訊化人才建設等工作。

其他國際參與部分，2008/5 召開“數字信息長期維護與保存”中美論壇以及參加第 16 屆國際檔案大會等。

參、電子檔案長期保存相關技術之應用現況

隨著資訊科技的發展以及我國推動電子化政府的努力，近 10 年來無論在提升政府效率及服務品質方面，已有相當具體的成果，同時，各級政府機關原生數位 (digital born) 資料比例與數位化檔案應用需求不斷提高；此外，各機關廣泛使用非屬公文類型之資料庫執行業務、推陳出新的網路申辦服務等，更加顯著電子檔案對於未來檔案事業發展的影響。電子化帶來許多便民優點，但是在大量資料數位化/電子化或原生數位資料不斷產生後，自然面臨資訊安全、保存、典藏與備份等問題，這與過去傳統紙質文件的保存所面臨之問題，是截然不同。本章主要探討國外電子檔案長期保存相關技術問題，以作為我國後續發展規劃之參考依據。

一、相關技術

本節主要就技術角度來探討電子檔案保存的方法與策略。經由文獻探討，彙整國內外專家學者對於長期保存所提出之策略有：系統保存/技術保存、更新、轉置、模擬、標準化、封裝、複製或分散儲存、轉印成紙質文件或其他類比型式資料等。

以下所要探討的技術將依據本案建議徵求書所提需求，限縮於系統保存/技術保存、轉置、模擬、標準化、封裝 (含詮釋資料) 等。

(一) 過時必備 (System Preservation)

過時必備在文獻上也經常被稱為技術保存 (Technology Preservation)，是指將數位資訊以其原始的格式儲存，連同電腦硬體及系統的保存，當然可以依照實際情況需求，選擇只保存電腦硬體及其作業系統，或只保存應用軟體等。這項技術策略認為如同博物館一般，專門保存與維護完整之重要過時設備，同時包含能夠複製任何老舊過時的軟硬體系統；其所涉及之保存內容包括原始應用執行程序、作業系統軟體以及硬體平台，強調應保存電子檔案的原始運作環境，才能呈現電子檔案的真實面貌，進而實現電子檔案的長期保存。

由於這個策略必須持續將過時或即將被淘汰的電腦系統保存起來，以長遠的觀點而言，需要大量的金錢、維護人力及實體存放空間，對於許多財政拮据的政府檔案管理機構而言，被視為一項奢侈且不切實際的選擇，因此也大大地降低了被採用的可能性，所以較適合短期目標或者與其他策略搭配使用。

(二) 轉置 (Migration)

轉置是指將數位資訊從前一代的軟體或硬體設備定期轉換到新一代設備上，包含轉存（refresh）媒體以及轉移（migration）系統二種概念，是電子檔案長期保存的主要技術手段；適用於新技術的發展、採用及電子儲存媒體衰退等問題之上。電子檔案長期保存過程當中，採用轉置策略的因素有很多，其中最主要的因素有硬體設備更新汰換、電子檔案儲存格式標準過時、作廢及保存需求變化（例如，把 WordStar 文件轉移到 WordPerfect 格式，再依次轉移到 Word3.0、Word5.0、Word97、Word2000、Word2003 格式，如此隨著軟體版本的更新不斷進行轉移）等都可採用轉置技術解決。

以電腦科技的發展趨勢看來，轉置主要有二種情況，一是把電子資料從種類繁多的格式轉移到目前廣泛使用的幾種標準格式當中。二是將電子資料從穩定性低的媒體轉移到穩定性較高的媒體之上，或從對軟硬體環境相依性較高的格式轉移到對軟硬體環境依賴程度低的環境。

有效的轉置必須保證電子檔案的完整性（註 1）、真實性（註 2）、可及性（註 3）及不可否認性（註 4），否則將造成內容資訊的嚴重錯漏。因此在規劃執行上需要考慮多種因素：

1、預測新技術的出現

包括新的技術和媒體為何、功能為何、是否與現行系統相容。

2、觀察市場上的主流產品

目前市場上主流產品為何、新的技術是否有可能成為市場主流、是否與現行系統相容。

3、人力、時間、經費之消耗

需投入更多的人力、時間及經費，達到有效的轉置並符合經濟成本效益。

（三）模擬（Emulation）

模擬是指使用、開發模擬軟體，在今後未知的系統上模擬舊系統，使得電子檔案所依賴的原始操作程序能夠在非原技術平台上正常運作，並可確保資料不會遺失。實際上模擬就是延遲技術淘汰的一種方法，它並非如技術保存一般保存並維護舊有軟體系統與硬體設備，而是經由軟體工程師所撰寫之模擬系統，模仿過時的軟硬體平台的行為與相關操作系統，進而形成電子檔案的舊

有運作環境，達到電子檔案的原始格式及原始資料的保存。

模擬不僅能使已過時的系統所產生之電子資料以其原始面貌（原始格式、版面與內容等）展現出來，其功能性也隨之得到保障。然而模擬引起多數的討論與重視，在於其有助於保護、辨識那些與特殊軟硬體相依性極高而又無法在新、舊技術平台間進行轉置的電子資料檔案等，例如超文件（在文件當中的某些字、符號或短語具有熱連結（Hotlink）的作用，在顯示上其字體或顏色產生變化或標有下橫線，用以區別於一般的正文）、多媒體等複雜文件類型。

從純粹技術觀點而言，模擬自有其可行性，但是在實際的運作環境當中，它的相容性並不高，對於軟、硬體的技術人才要求非常高，而且模擬器自身的耐用性也必須得到一定程度的維護與不斷升級，因此，從長遠觀點來看，這未必是一個有效的方法，系統的開發者也可能最終會放棄對某些軟體的向下相容性。

（四）標準化（Standardization）

電子資料的長期保存所涉及的眾多問題當中，標準化試圖以某一種廣為大家接受的標準來進行電子資料的儲存、描述、組織與檢索，因此被視為緩解技術更新與電子資料讀取衝突之良方妙帖，此方法的重點在於資料的標準格式上，同時也須考慮與其他策略方法搭配運用以避免資料遺失，例如系統進行轉置時，需要把資料進行標準化以利轉置到新一代的設備上使用。

此外，選擇標準時也須考慮此標準格式之資料是否存在保存格式、遇到格式更新時，是否會因格式轉換而損及資料的完整性、是否能夠支援原始資料的各方面功能。由於標準同樣會隨著時間的過往而演變，因此定期的檢視目前標準的變化以及進行標準的更動，都是必要的措施。

SGML 是由國際標準組織（ISO）於西元 1986 年所正式公告之電子文件交換標準，其目的在於規範文件製作及交換的標準，以確保電子檔案能在不同的電腦系統上交換應用，也因此符合電子檔案長期保存的需求。

SGML 使用文件類別定義（Document Type Definition, DTD）來指定文件的結構。SGML 可以轉換成多種的輸出格式，包括純文字、HTML、LaTeX、PostScript 等格式。其中最重要的，它為所有的文件提供了一致的外觀，便於管理與維護。SGML 文件本身包含三個部份：一是 SGML 宣告（SGML Declaration），二是

文件類別定義 (Document Type Definition, DTD)，三是文件實例 (Document Instance, DI)。其中，SGML 宣告是用來告訴 SGML 剖析器 (SGML Parser) 在剖析該份文件時，所使用的字元集和文件特性；而文件類別定義是用來宣告該份文件文件結構與文法參數；至於文件實例，就是加上標記處理後的文件。

由於 SGML 的結構太過複雜，因此衍生了二個子集：

1、HTML

HTML 是從 SGML 語法簡化而來的子集，在 1991 年由 Tim Berners-Lee 建立了 HTML 語法，簡單易學，也使得今日網路上的資訊能夠快速傳播，互相共享。但是 HTML 語法太過於鬆散，不夠嚴謹，所以才有目前的 XML 出現。

2、XML

XML 是以 SGML 的格式精簡後制定出來的，制定者是全球資訊網協會 (World Wide Web Consortium, W3C)。XML 是 SGML 的一個子集合，誕生的目的是為了擴充網路的應用、彌補 HTML 的缺點，以及讓 SGML 也能容易地在網路上應用。隨著資訊的普及，資料量的劇增，與資料的複雜化，HTML 的語法格式則顯得過於簡單，無法支援一些專門格式，而 XML 與 HTML 正好可以相輔相成，解決這些問題。

所以 XML 並不是用來終結 HTML，取代 SGML，或是廢除舊的標準。XML 是一種用來定義其它語言的另一種語法系統，這正是 XML 功能強大的主因。

(五) 封裝 (Encapsulation)

封裝的定義在於使用詮釋資料將被保存之數位資料以及相關資訊，例如文件說明等，包裹於一個封包之內，與其他的策略方法搭配運用，例如透過轉制或模擬等方式，提供使用者讀取與了解檔案內容。

詮釋資料為一種描述、說明與定義資源的結構性資訊，可用於描述任何層次的資源及資訊模式，如整體館藏或單一資源。保存性的詮釋資料 (Preservation Metadata) 是各種數位保存策略之基礎，模擬策略或轉置策略都與保存性詮釋資料關係密切，尤其模擬策略需要許多相關資訊的支援，當保存機構採取模擬策略時，應該在封裝保存階段就把未來可能需要的資訊完整記錄下來。其中有關詮釋資料項目，應該參考國內外機關對於保存性詮釋資料的發展方案，整理彙整出所需使用的詮釋資料項目。

目前最廣為人知的國際標準為都柏林核心集 (Dublin Core)，以簡易與彈性為原則，任何社群皆可採用都柏林核心集描述和檢索各式的資訊資源，以確保不同類型的描述性詮釋資料間的互通，但此類標準卻無法保證數位資源的長久取用。

二、國外應用現況

隨著電子資訊的激增、數位化環境的形成，現代人日益依賴電子資訊以及網路服務。電子檔案長期保存伴隨著其問題的複雜度也日益受到重視。在歐美、澳大利亞、紐西蘭等國家，有關電子檔案長期保存策略研究和應用已有長足進展，一些國家也早已開始應用性部署，同時累積了大量的知識與經驗，分別介紹如下：

(一) 電子檔案轉置發展

以各國主要電子檔案保存機構之網站為分析對象，從中瞭解該機構透過網站所傳遞之數位化服務與相關長期保存技術，各國以英國、美國、加拿大、澳洲等國家之國家檔案館或國家圖書館網站為主，將其電子檔案長期保存策略與轉置技術說明如下：

1、英國國家檔案局(The National Archives, TNA)

國家電子檔案資料集(National Digital Archive of Datasets, NDAD)包括：British Crime Survey、Agricultural & Horticultural Census、Primary Births、Judicial Statistics、Schools Census、Department of Transport、British Bats、Mining Reports、Heavy Goods Vehicles、Beer Duty、Public Health、North Sea GIS 等，其處理過程包括從其他組織取得資料集(Acquisition)：要確認其完整性與一致性，並安全地傳輸資料集與詮釋資料；存取資料(Accession)：檢核、建立目錄、資料轉換為標準化、載入資料與詮釋資料；長期保存(Preservation)：媒體轉置、穩定性檢核並備份；顯示(Presentation)：資料與詮釋資料管理、建立目錄、擷取與顯示。

此外，為因應快速增加的電子檔案數量，英國國家檔案局推動「Seamless flow Program」計畫，其程序涵蓋評估與挑選(Appraisal and selection)、詮釋資料與目錄(Metadata and cataloguing)、移轉至國家檔案局(Transfer to The National Archives)長期保存與維護(Preservation and maintenance)、技術監控(Technology watch)、傳送與顯示(Delivery and presentation)、管理與安全(Management and security)及職責改變與訓練(Business change and training)等，希望電子檔案從政府機關產生

後，經過檔案管理局「保存」到公開給民眾在網路「應用」，整個電子檔案生命週期管理能完全自動化，將使英國 TNA 從以紙本檔案為主，邁入以電子檔案為主的里程碑。

任何在辦公室電腦所產生的文件，都列入作業環境涵蓋範圍：包含文書處理軟體、電子郵件、電子試算表、簡報資料、圖形、影像與網頁資料。

根據英國國家檔案館之數位保存聯盟網站對轉置策略之敘述，轉置的特色在於簡單的程序被建立的很好、是目前大多數數位保存所優先選擇的保存策略，隨著技術的進步，轉置會變得更簡單。電子檔案長期保存策略採轉置作業，連帶保持所有原本的檔案格式，並發展可以自動轉置軟體。

TNA 發展 PRONOM 計畫之線上資訊系統，用以長期保存電子檔案，並提供給任何需要利用資訊的人。採用文件閱覽軟體(Viewer)技術瀏覽檔案，亦採取檔案轉置措施，即轉換檔案為新格式，以便在新技術環境可以存取，以確保檔案可利用，並採用模擬舊有技術環境以便檔案可以其原格式存取。其電子檔案長期保存策略如下：

(1) 消極性長期保存(Passive Preservation)

執行 Digital Object Archive(DOS)，提供安全性儲存(store)與存取控制(Security and access control)。DOS Management Console 請參閱圖 3-1。

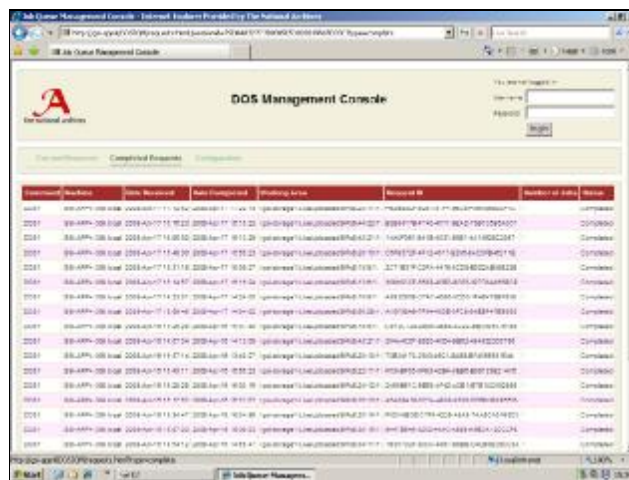


圖 3-1 DOS Management Console

A. 實體與系統安全

使用者與系統存取控制(Physical and system security, user and system access control)。

B. 電子檔案完整性管理(Integrity management)

C. 儲存管理(Storage management)

媒體選擇、管理與清除、備援與備份(Media selection, management and refreshment, redundancy and backup)。

D. 災害復原(Disaster recovery)

(2)積極性長期保存(Active Preservation)

A. 確保日後已儲存紀錄之持續可及性，並跨越技術改變

B. 積極性長期保存系統，納入 PRONOM

C. 服務導向架構

D. 所有服務均以 Web Services 展示(區域或遠端)

E. 使用 J2EE 開發

F. 採用現行新的 TNA 與第三者工具

G. 協調使用流程引擎(Workflow Engine)

英國國家檔案局積極性長期保存循環請參閱圖 3-2，Active Preservation Management Console 請參閱圖 3-3。

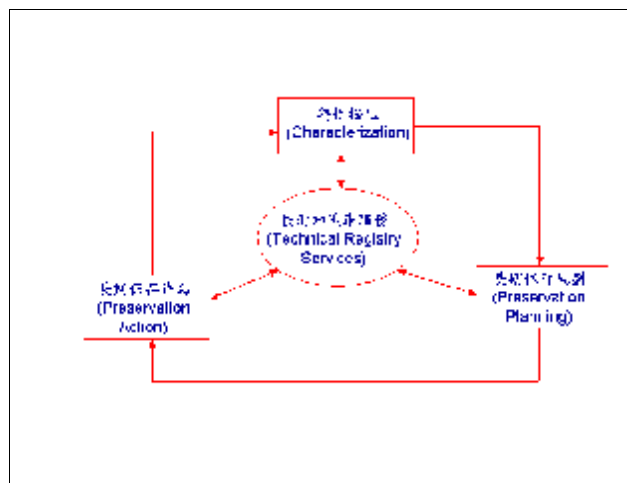


圖 3-2 英國國家檔案局積極性長期保存循環

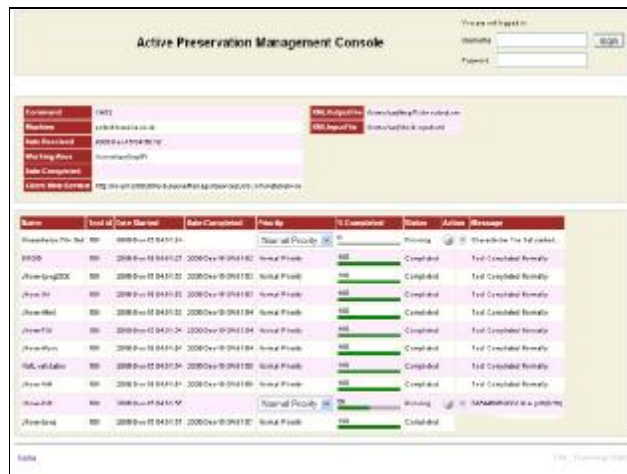


圖 3-3 Active Preservation Management Console

(3)TNA 電子檔案特性描述(Characterization)

A. 功能

衡量對電子檔案長期保存意義重要的屬性。

B. 識別(Identification)

格式化版本以 PUID(PRONOM Unique Identifiers)顯示，一種為 PRONOM 知識庫中之格式提供持久、唯一與清楚的識別機制。

C. 確認(Validation)

- ☆ 格式完好(Well-formed)：在語句構造上(syntactically)正確。
- ☆ 有效性(Valid)：格式完好且在語義上(semantically)正確。

D. 屬性擷取


- ☆ 技術屬性(Technical properties)：包括檔案格式(The format of the object)、電子影像壓縮演算法(The compression algorithm used in a digital image)、Office 文件之 Macro 特性(The nature of macros contained in an office document)。
- ☆ 固有屬性(Inherent properties)：表示電子檔案長期保存必要品質與經由類似轉置之長期保存作為。屬性包括影視剪輯之持續時間(The time duration of a video clip)、文書處理之本文內容(The textual content of a word-processed

document)。

E. 透過 PRONOM 自動運作

☆ PRONOM

技術資訊線上註冊，例如檔案格式特徵。PRONOM 可以下載全部資料，其欄位包括格式代號、格式名稱、格式版本、格式風險與延伸檔名。PRONOM 請參閱圖 3-4。



The screenshot shows the PRONOM website interface. The search results table is as follows:

PRONOM Unique Identifier	Format Name	Format Version	Risk Level
x-m23	Microsoft Document		doc
x-m2	Microsoft Word to Microsoft Document	2.0	doc
x-m23	Microsoft Word to Microsoft Document	X	doc
x-m271	Microsoft Word to doc, DOC Document	1.0	doc
x-m274	Microsoft Word to doc, DOC Document	4.0	doc
x-m275	Microsoft Word to doc, DOC Document	5.0	doc
x-m276	Microsoft Word to doc, DOC Document	5.5	doc
m23	Microsoft Word to Microsoft Document	1.0	doc
m23	Microsoft Word to Microsoft Document	1.0	doc

圖 3-4 The Technical Registry PRONOM

☆ DROID 2.0(Digital Record Object Identification)

由 TNA 發展之軟體工具，依據 PUID 執行自動批次識別檔案格式。DROID 是不附屬於平台之 Java 工具，在開放原始碼許可(Open Source License)情況下可自由下載，DROID 新版本將涵蓋資料庫，以儲存日誌資料。DROID 請參閱圖 3-5。

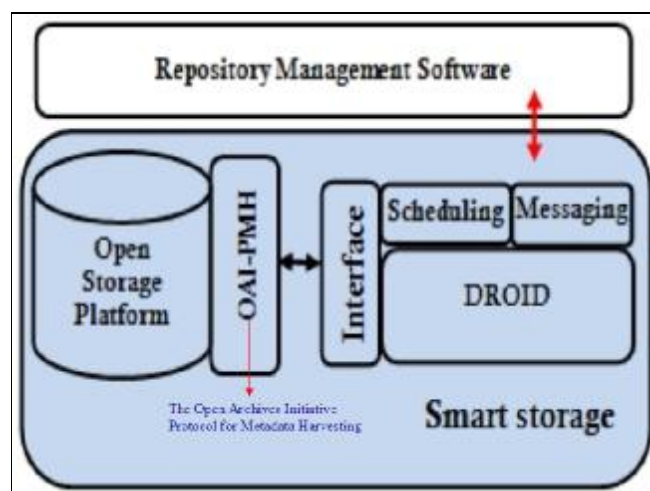


圖 3-5 識別工具 DROID

☆ JHOVE

確認與屬性擷取(Validation and property extraction)。由 JSTOR 與哈佛大學圖書館共同發展，提供系列電子檔案型態之自動識別、確認與特性描述(characterization)。在 GNU 較寬鬆公共許可(Lesser General Public License) 情況下其先期發行 Beta 版本可自由下載，JHOVE 請參閱圖 3-6。其功能如下：

- 由 JSTOR 與哈佛大學圖書館共同發展特性描述 (Characterization)工具。
- 識別(Identification)：經由特徵比對確定推定的格式。
- 確認(Validation)：格式完好(Well-formed)且具有效性 (Valid)。
- 提供 Module：AIFF、ASCII、Bytestream、GIF、HTML、JPEG、JPEG2000、PDF、TIFF、UTF-8、WAVE、XML 等格式辨識模組。
- 屬性擷取：格式、版本、作者、主題、產生時間、修改時間、字型、TrueType 字體、頁數。

```
RepresentationInformation: Zmhr5e3e9v6ocra1_emo3ator_4s_4_1ong-
term_preservation_strategy_for_digital_objects.pdf
ReportingMode: PDF-hul_Rai 1.7 (2008-02-20)
LastModified: 2008-08-20 14:09:34 CST
Size: 82274
Format: PDF
Version: 1.3
Status: Well-Formed and valid
SignatureMatches:
PDF-hul
MimeType: application/pdf
PDFMetadata:
Objects: 02
FreeObjects: 1
IncrementalUpdates: 0
DocumentCatalog:
PageLayout: SinglePage
PageMode: UseNone
Info:
Title: Modular emulation as a long-term preservation strategy for digital
objects
Author: Jeffrey van der Hoeven, Hilde van Wijnngaarden
Subject: Paper about the application of modular emulation in digital
preservation
Keywords: modular emulation, digital preservation
Creator: PDFCreator Version 0.8.0
Producer: AFPL Ghostscript 8.14
CreationDate: Tue Aug 09 14:45:49 CST 2006
ModDate: Tue Aug 09 14:45:49 CST 2006
ID: D:8b2442e9395bfcecb06246b991b21b725, 0x8b2442e9395bfcecb06f
246b991b21b725
Filter:
FilterPipeline: FlateDecode
Fonts:
Type1:
Font:
```

圖 3-6 確認工具 JHOVE 測試結果

☆ Apache Java POI(Poor Obfuscation Implementation) library
MS Office 文件之確認與屬性擷取，依據 MS OLE2 複
合文件格式(Compound Document format)，POI 專案採用純
Java 組成運用各種檔案格式之 API。可使用 Java 讀寫 MS
Excel、Word、PowerPoint 及 Visio 檔案。

☆ Java JAXP API

確認 XML(Validation XML)，使用各式不同 API 確認
(Validate)、語意分析(Parse)及轉換(transform)XML。

(4)TNA 長期保存規劃(Preservation Planning)

長期保存規劃架構是積極性長期保存系統決策核心，決定
了長期保存作為適用於何種檔案與適當的應用時機。長期保存
以風險為基礎，考慮格式風險(Format risk)與情境風險(Instance
risk)，前者來自於格式之一般屬性，由 PRONOM 依關鍵格式
屬性資訊推測與記錄，推測屬性包括可支援該格式之軟體工具
與格式的開放性。後者來自於檔案之特定屬性，該屬性例如在
Office 文件中 Macros 出現或應用在影像檔案之 Compress
Algorithm。長期保存規劃如下所示：

A. 風險評估(Risk Assessment)

- ☆ 識別作為之緊迫性，依儲存於知識庫中標準規範評估。
- ☆ 如檔案之長期保存風險分數高於設定之門檻數值，即自行

啟動產生長期保存規劃，隨即執行轉置程序。

B. 技術監視(Technology Watch)

- ☆ 監控會影響電子檔案持續可及性之技術的改變。
- ☆ 取得技術的改變並相對應修改 PRONOM 知識庫內容。
- ☆ 結果回饋已修正風險評估準則。

C. 影響評估(Impact Assessment)

- ☆ 監控進行中風險。
- ☆ 量化徵集的影響。

D. 產生轉置路徑(Migration Pathway)

- ☆ 經由分析知識庫以確認可能的轉置路徑
 - 確認最適當的轉置目標格式，即支援具最低風險之特定來源的屬性。
 - PRONOM 容納之資訊足以確認特定來源與目標格式之間可能的轉置方法組合。
- ☆ 測試並確認方法
 - 其結果可作為所喜轉置方法之選擇，並確認該方法適用於該特定的情境。

(5)TNA 長期保存作為(Preservation Action)

A. 轉換檔案以確保其持續可及性

B. 經由 PRONOM 自動運作的轉置工具

- ☆ 執行(Execution)：實施轉置作業，轉置工具自動轉換特定檔案並產生新表現形式(Manifestation)。
- ☆ 特性描述(Characterization)：採用標準特性描述程序描繪已轉置的檔案。
- ☆ 確認(Validation)：自動將特性描述結果與原始檔案進行比對，以確保有效性。

C. 轉置工具

- ☆ Stellent Transformation Server

提供內容管理系統, Stellent 在 2006 年 11 月併入 Oracle, 該產品即為 Oracle Universal Content Management。

☆ TNA validation tool

2、美國國家檔案及文件署(National Archives and Records Administration, NARA)

美國國家檔案及文件署於 2005 年 9 月 8 日授與 Lockheed Martin 3 億 8 百萬元的合約, 委任承辦永久檔案管理系統之建置工作, 此系統需提供由聯邦政府所產生之電子檔案的保存與管理, 總計約 135,000 人負責進行研究設計發展製造整合等工作。首都華盛頓管理局之前推行為期 6 年之電子檔案管理計畫 (Electronic Records Archives, ERA) 也將與此計畫共同合作。

ERA 於 2005 年開始進行系統發展, 2007 年首次使用電子檔案管理系統, 2008 年協助政府機關改善電子檔案管理績效, 2011 年完成全部系統開發與建置。ERA 時程表與 ERA 架構圖請分別參閱圖 3-7 與圖 3-8。

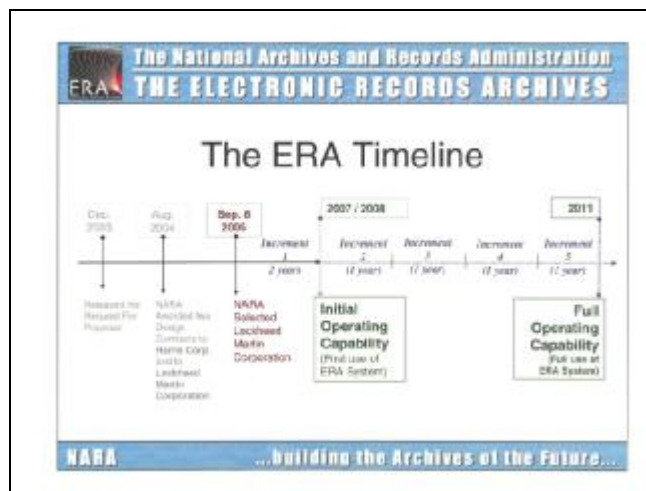


圖 3-7 The ERA 時程

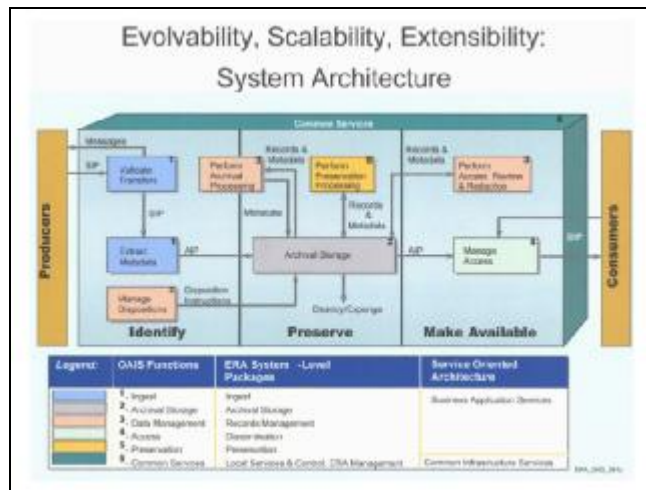


圖 3-8 ERA 架構圖

(1) 架構說明

A. 擷取(Ingest)

完成將電子檔案導入 ERA 的程序，包括：

(A) 實體電子檔案轉移至 ERA

系統必須檢核轉移功能整合性、正確性與完整性，並檢核轉移之電子檔案與配置條款(Disposition Agreement)一致，並按長期保存與存取規劃(Preservation and Access Plan)進行保存與服務。

(B) 驗證轉移的內容

轉移的實體必須能夠轉移所有支援的檔案格式與媒體上的電子檔案。

B. 檔案儲存庫(Archival Storage)

ERA 需要可靠的電子檔案構成之資料儲存庫，包括：

(A) 儲存管理服務(Storage Management Services)。

(B) 媒體管理(Media Management)確保所有儲存資料的實體完整性。

(C) 資產保存清單(Assets Inventory)維持資產與各類型電子檔案關聯。

(D) 系統資產之多元複製品，管理儲存多版本複製品，支援持續運作與資料復原、例行備份及儲存電子檔案長期保存與存取目的的不同格式版本。

(E) 媒體轉置(Media Migration)

- 耐用媒體。
- 媒體必須是市場上可行且受支援。
- 媒體必須能以自動方式操作。
- 媒體按開放式標準格式化(Formated)、標籤化(Labeled)與錄製(Recorded)。
- 媒體可被位元錯誤偵測(Bit Error Detection)與描述支援。
- 媒體必須能自我描述(Self-Describing)與非加密形式(Non-encrypted)。
- ERA 支援能將毀損媒體復原資料的自動維護與工具，能將儲存資料轉置至新媒體以強化效率與可靠性。

C. 檔案管理(Records Management)

系統提供 NARA 對所有型態資料生命週期管理的決策支援，包括支援檔案保存程序活動例如評估(Appraisal)、排程(Scheduling)與說明(Description)等，檔案管理程序包括：

(A) 管理配置條款(Managing Disposition Agreement)。

(B) 轉移實體保存(Physical Custody)與法定保存(Legal Custody)資料。

(C) 檔案保存說明(Archival Description of the Records)。

(D) 權限來源管理(Managing Authority Sources)。

(E) 管理檔案生命週期資料(Managing Records Lifecycle Data)。

(F) 使用範例檔案(Utilizing Sample Records)。

D. 長期保存規劃(Preservation Planning)

(A) ERA 提供適當的工具、技術與方法，俾長期保存規劃確定預防性(Preventive)、反應性(Reactive)與前瞻性(Proactive)的步驟，以保證電子檔案免於毀損，在任何時間產生、儲

存、移轉、管理或保存程序皆能產出資訊科技之真實性複製品。

- (B) 長期保存與存取規劃(Preservation and Access Plan)對在 ERA 長期保存電子檔案均相關連。
- (C) 長期保存程序包括產生(Generation)、註冊(Registration)與確認(Validation)模型與其他確認檔案屬性的控管。
- (D) 理想長期保存格式應能自我描述(Self-describing)，並按開放性(Open)與非專屬性(Nonproprietary)標準確認(Validating)。

E. 存取(Access)

- (A) 提供存取 ERA 資產需要搜尋能力以尋找檔案與系統內其他資產，ERA 並針對所建立檔案以群組(Groupings)與排序(Arrangements)方式存取。
- (B) 健全的搜尋能力，包括描述檔案生命週期資料與檔案內容搜尋能力。
- (C) 在顯示與輸出時，電子檔案仍需維持真實性，包含檔案特定屬性。此外，提供自動編寫與使用者編寫能力。

F. 使用者介面(User Interface)

ERA 安全包括確保資產避免不正當存取與破壞，並保證授權使用者持續可存取。

G. 安全性(Security)

- (A) ERA 安全包括確保資產避免不正當存取與破壞，並保證授權使用者持續可存取。
- (B) 安全方法應用至系統資源與服務、使用者及資訊資產等。
- (C) 對資產保護包括對特定權限資源的強化性權限存取，並由健全的稽核軌跡(Audit Trail)支援。

H. 系統管理(Administration)

管理使用者註冊與帳號、使用者協助、管理使用者與 ERA 溝通、保持事件日誌(Event Log)並提供報表、系統測試、監控、參數調整與系統故障處理等。

(2) ERA 的主要功能

A. 確保電子檔案的長久保存

- (A) 把檔案備份、異地儲存，防止實體文件的損毀。
- (B) 使用非專屬介面系統，防止檔案被個別軟體提供商所控制。
- (C) 謹慎設計和開發，避免人為失誤。
- (D) 不同系統間的功能互補，防止軟體出現漏洞。

B. 確保文件的長期可存取

- (A) 客觀評價文件轉置、模擬等技術。目前尚沒有一種技術或者方法能解決電子檔案逐漸過時的難題。現在應該做的是把文件的原始檔案以詮釋資料的形式統一保存，以便將來更好地理解該文件。
- (B) 除了保存文件的原始內容，還可以把文件轉化為一些更好的衍生形式，這些形式應具有通用性、本文性和持久性。
- (C) 為避免技術上的過時，不能過於依賴於某一個獨立的技術平台或單一的供應商。不能把所有的文件都轉化為某一個技術標準格式。
- (D) 保留重要的詮釋資料，通常認為保存的詮釋資料越多越好，但事實上，最有價值的詮釋資料是那些無法從文件自身抽取且很可能會丟失的詮釋資料(例如，結構性詮釋資料)。
- (E) 保存那些文件中明確列出或蘊含在內容之中的重要的參考資料。數位化檔案經常會涉及到其他的數位對象，比如文件中包含的影像、項目表格或調用的其他組織的文件。
- (F) 盡可能多保存那些用於加工和處理原始文件的軟體和工作流程的相關資訊，最好的辦法是記錄文件處理的全部過程，一旦發現文件的保管形式不正確，便可以利用這些紀錄，分析哪些環節需要調整，從而得到文件最佳處理方案。

C. 確保電子文件的安全

文件資料很容易丟失，原因也很複雜，例如，惡意破壞、自然災害、病毒侵入、人員操作失誤、硬體的毀壞。因此，

必須考慮這些問題可能帶來的損失。首先，應該把系統設計與運轉規範和流程結合起來；其次，制定關於文件拷貝、完整性檢查、訪問控制和安全審核等方面的規範。對於那些相對重要的文件，應予以特別的關注，確保其安全。

D. 確保電子文件存儲介質可靠

- (A) 從經濟性、效用性和持久性的角度考慮，光碟系統的優勢已經超越了包括磁帶在內的其他媒體。
- (B) NARA 要仔細考慮不同的儲存方案，比較不同技術的成本、複雜性、功能和可靠性。
- (C) 媒體轉置包含維護電腦系統以便從現行儲存媒體將資料轉置至新媒體。

E. 確保系統的遠端存取

將來很可能只需要使用某種單一的、穩定的工具，就能存取 NARA 的資料儲存庫，檢索到所需的電子文件。

- (A) 各機關至少每 6 個月應測試儲存清理中或有意延長為永久性電子檔案。
- (B) 移轉檔案必須採 ASCII 或 EBCDIC 方式編碼，或資料文件與資料庫、文字文件、數位空間資料或其他種類電子檔案。
- (C) 徵集式永久物件保存(Collection-based persistent object preservation)架構由 U.S.National Partnership for Advanced Computational Infrastructure 所發展。其方法包括：
 - 確認被保存物件的特性。
 - 以明確且摘要之模式展現特性。
 - 應用該模式將物件轉換為適合長期保存之獨立性技術格式。
 - 採用多層次文件(Multi-Valent 或 Multi-layered Documents)技術擷取並保存 bitmapped 影像文件。
 - 資料庫與地理資訊系統(Geographic Information System, GIS)保存呈現方式採 XML 標準。

(3) 長期保存策略

- A. 維護固有軟硬體元件，確保檔案可再製能力。

- B. 隨著科技的演化將舊有檔案格式轉換為新的格式。
- C. 提供檔案產生時之環境，以供檔案重新產生之用。適時轉換檔案格式，減少對特定軟硬體依賴性，使檔案在各種技術轉換過程仍保有原貌。
- E. 保存的經驗大都限定在與簡單技術相關的格式(例如循序檔)。
- F. 不可能長期保存電子檔案，只可能維持再產生電子檔案的能力。
- G. 保存系統應建置為可能會取代硬體或軟體的組件，但對系統有最小的影響，對內容沒有影響。
- H. 檔案紀錄產生或從公眾的或個別的活動接受的資訊，由內容、全文與結構所組成，以提供活動的證據而不需顧慮類型或媒體。延伸涵蓋檔案的顯示與電子檔案惟一的特性，即超連結。
- I. 長期保存採 XML 顯示資料結構，此為單一資料進入長期保存之內容模型。
- J. 多層次文件(MVD)技術取得並維護文件的位元影像，XML 標準被考慮保存檔案的外觀。
- K. 展現這些方法包含再次決意一個多樣化的檔案之徵集。
- L. 美國國防部電子檔案保存標準(DoD5015.2-STD)分為必備功能需求、非必備功能與機密檔案管理功能等三類。
- M. 不可將檔案壓縮，除非經同意並提供解壓縮軟體。
- N. 應付快速增加的電子檔案之可能做法：採傳統方式處理、控制並轉置至新媒體或格式、提供新的儲存庫。
- O. 針對下列電子檔案要有新解決方案：涵蓋 E-Mail 系統、工作流程系統、Digital films、TV(HDTV)、Models 3D(VMR-Virtual Model Reality)、CAD(Computer Aided Designing)與 GIS(Geographical Information Systems)等資料。

(4)轉置策略

- A. 當現行格式可能不受支援，過時資料必須轉置為能符合未來技術的格式。

- B. 透過單一套裝軟體版本與某些標準化類型，涵蓋了系統轉置。
- C. 替代性格式在某些個案可能不存在。
- D. 標準化產品包含延伸性標準或以不同方法發展標準的特性。
- E. 每一個轉置帶來選擇性的風險，並與顯示資料的整合性相連結有困難度。
- F. 隨著格式數量的成長而複雜度遞增。
- G. 稱呼技術過時，而非檔案的長期保存。

(5)轉換策略(Transformation Strategies)

- A. 包含 NARA 參與協同研究的長期保存的方法。
- B. 不以原技術狀態保存資料。
- C. 需要長期保存需求的明確規範，該需求與內容、全文、結構、資料之保存及所屬的集合相關連。
- D. 集中在檔案保存的需求。

(6)選擇儲存媒體或媒體轉置時考慮要項

- A. 判定保存期限時檔案被授權的期限。
- B. 保留檔案時必要的維護工作。
- C. 檔案儲存與檢索的經費。
- D. 檔案的量。
- E. 檢索檔案的時間。
- F. 媒體的可攜式及資訊從一媒體轉換至另一媒體的能力。
- G. 媒體須符合聯邦資訊現行處理標準。
- H. 各機關應確保資訊不因技術的改變，或將轉置的儲存媒體匯入現有軟硬體時遺失，轉置至其他媒體前，各機關應確定經授權清理的電子檔案得在轉置後順利執行。
- I. 各機關至少每 6 個月應測試儲存清理中或有意延長為永久性電子檔案。
- J. 各機關保存永久性或清理中資料的磁帶應不逾 10 年。

K. 每一捲用於儲存永久性或清理中電子檔案磁帶的外部標籤均應有統一標示。

(7)美國電子公文檔案產生方式

A. 透過掃描方式，將原為紙本形式文件檔案，如信函與備忘錄等，轉換為電子檔案。

B. 各機關所使用之已採電子方式進行文件，可直接轉換為電子檔案。

3、加拿大

(1)加拿大圖書館與檔案館(Library and Archives Canada, LAC)

A. 檔案管理之功能

(A) 檔案分類計畫

(B) 檔案存取及安全機制

(C) 檔案資訊 (包含位置、描述及清理)

(D) 檔案報表

(E) 檔案之保留、負責及借閱

(F) 檔案之稽核機制

B. 長期保存策略

(A) 詮釋資料以都柏林碼(Dublin Core, 2001)為基本核心依據。加拿大主要詮釋資料描述作業，並不在機關中的紀錄管理中心，而是在完成清理確立保存價值後，由檔案館人員進行。

(B) 檔案學界認為業務處理過程，非紀錄文件表達的主題。加拿大不保存鑑別過程的數位簽章，合乎檔案科學的分析原則，也降低保存作業之困難度。

(C) 對於現行各類電子文件之加密技術，加拿大國家檔案館保存政策中，明訂不處理任何加密後之電子檔案。當加密作業會變更原有內容者，就會破壞原始檔案的完整性。

(D) 文件處理過程之數位簽章流程，對目前管理上技術過於複雜，加拿大直接採取不允管理的決定。

- (E) 加拿大檔案學界認為封裝技術是在形成檔案之後，對檔案加於變更形式，破壞文件主體而持反對看法。
- (F) 除公文外，政府所屬電子資料庫資訊，也更具有保存之歷史價值。如屬科學性資料之氣象局天氣地震資料庫、財稅機關納稅資料、民航局飛航管制系統等。
- (G) 電子檔案的保存是以資訊系統為對象，而非公文為主體。

C. 轉置策略

- (A) 檔案之保存年限管理、轉置及清理。
- (B) 每 2 年進行資料轉置之評估，需要注意到電腦硬體、軟體及儲存媒體之轉置與更新。
- (C) 資料轉置：每 2 年進行資料轉置之評估，需要注意到電腦硬體、軟體及儲存媒體之轉置與更新。
- (D) 電子檔案模式分析：本文(Text)、doc、PDF、wpd、圖表(Graphic)、影像/Images)、電子簽章(Electronic Signature)與電子戳記(Electronic Seal)。字碼採 ASCII 碼、EBCDIC 碼、UTF-8。

D. Electronic Archives Preservation System (EAPS)系統

- (A) 在電子檔案徵集時有自動轉置、檢驗與擷取功能，自動轉置減低人工介入並節省人力。
- (B) 當進行長期保存程序時，相關技術資訊擷取為詮釋資料，如此，確保原有與新的檔案格式細節均得以保存。
- (C) 轉置的自動化對極小的資料徵集或組織極為實用

E. 電子公文檔案之資訊安全

- (A) 存取控制(Access Control)。
- (B) 金鑰管理(Key Management)。
- (C) 稽核機制(Audit)。

(2)The International Research on Permanent Authentic Records in electronic Systems(InterPARES)計畫

目的是為電子檔案長期保存發展必要知識，並提供標準、

政策、策略與作為之計畫以確保該資料長期性與確實性。共分為三個階段：

- A. 第一階段，時程為 1999 年至 2001 年。主要著重在原生單位對其工作、使命與目標已不需要電子檔案真實性之長期保存，此階段產生多樣的結果，包括電子檔案挑選與長期保存相關確實性與方法之概念性需求。InterPARES1 團隊由很多專家學者組成，涵蓋法律、歷史、電腦工程、資訊工程與化學等，而專案主要的觀念與方法論是從古文書研究與檔案學研究而來。
- B. 第二階段，時程為 2002 年至 2006 年。專案的目的在發展與闡述能確保電子檔案長期保存精確性與可靠性，即存在於藝術、科學與政府活動中經由經驗、互動與動態資訊科技所產生的電子檔案。全世界藝術與科技專家、檔管人員、藝術家、科學家、工業專家與政府代表等共同合作，應付數位系統運作與不相容性、技術老化與媒體易脆性等所產生之挑戰，以保證社會數位紀錄的記憶得以在未來可以存取。
- C. 第三階段，時程為 2007 年至 2009 年。此專案將把 InterPARES 與其他研究至今努力所發展的電子檔案長期保存的理論與方法，為現存而將長期保存在組織內以有限的資源將其轉化為具體的行動計畫。處理過程中，詳細知識將被發展情況如下：一般性的理論與方法將如何被發展為小型或中型檔案與單元，且實際有效。什麼因素決定對每一檔案的內文的發展型態是適當的。執行這些操作需要什麼專業技能。

依此基礎，將為內部訓練課程、持續訓練專題、學術課程等對社會數位文件遺產之長期保存提供有競爭性的專業，發展模組，此外，經由保護電子資訊的正確性與真實性，以確保組織與機構的責任。

4、澳洲國家檔案局(National Archives of Australia, NAA)

(1)電子公文檔案數位資料庫

以經費 5.5 百萬澳幣建立一個能保存全維多利亞電子公文檔案數位資料庫：

- A. 提供一個線上傳送方式，可將機關電子公文檔案直接以線上方式傳送至 PROV。PROV 所推動的 VERS 是包含策略、標

準與建議的架構、可執行的計畫與解決電子公文檔案長期保存的方案。

- B. 提供一個數位檔案資料庫，永久儲存電子公文檔案。
- C. 提供一個管理資訊系統，有效管理電子公文檔案之詮釋資料，並保證其完整性。
- D. 提供線上存取電子公文檔案及一般實體檔案之入口點。
- E. 澳洲國家檔案局將轉置列為電子檔案長期保存機制，但避免重複轉置，將多樣性系列轉置為較小系列的標準格式。
- F. 降低風險策略。
- G. 開放原始碼鼓勵有興趣者提供回饋及改善建議。
- H. 未來應會有更佳長期保存格式，因此，從有限的標準化格式轉換為新格式會比從原格式容易。
- I. 保存所有版本舊格式。
- J. 軟硬體設備採易取得的工業標準產品。
- K. 以洋蔥形式層層包覆，鎖上的 Lock Signature Block 是將被標記物件與簽章區塊鎖起來，以確保不被惡意的竄改，並確保其安全性。
- L. 規定使用 PDF 或 TIFF 格式。

(2) 技術過時檔案之規劃(Planning)

機關需長期保存之電子檔案，必須為技術過時進行規劃，應確認資料能複製(Copied)、改版(Reformatted)、移轉(Converted)或轉置(Migrated)跨越至下一代電腦技術。

A. 產生電子檔案長期保存策略(Strategy)

- (A) 在組織內形成政策(Policies)、程序(Procedures)與指引(Guidelines)為發展策略提供正式的架構(Formal Framework)。
- (B) 提供員工手冊、參考資料與訓練，以確保長期保存策略會被正確地發展。

B. 電子檔案長期保存技術(Techniques)

(A) 轉置

- 將較舊或過時的軟硬體建構或上一代之電子檔案，固定轉置至現行建構或這一代，以便維持可及性。
- 將資料轉置至新的媒體時，同時提昇資料至較新的格式。
- 轉換(Conversion)：將電子檔案從其原資料格式轉移至標準且長期保存格式。轉換與「正規化」(Normalization)、「固定」(Stabilization)與「標準化」(Standardization)相關連。長期保存資料格式應是開放原始碼(Open Source)與非專屬性格式，以便提供可能長久性且較少限制。

(B) 模擬

原資料格式得以保存，可至能重現原「Look and Feel」之環境存取檔案資料。

(C) 封裝

封裝並不保存電子檔案，必須與轉置或模擬一起使用，以確保檔案的可及性(Accessibility)。

C. 發展電子檔案長期保存策略(Strategy)

(A) 挑選電子檔案長期保存實施方法(Approach)

- 發展的成本，包括為了持續長期保存處置的循環性成本(Cyclical Costs)。
- 所挑選實施方法的技術複雜度與機構支援長期實施方法的能力。
- 與現存軟硬體相容性。
- 對業務操作的影響。
- 為維持機關電子檔案長期保存整合性、可及性與功能性，實施方法應具有全然的效力與堅實。

(B) 確認電子檔案長期保存處置執行時機(Applied)

- 為擴大電子檔案長期保存前景，當檔案尚可存取時，長期保存技術即為可取與實際的執行時機。
- 機關更快提出長期保存事宜，決定並發展適切的長期保存實施方法，電子檔案可以成功被保存的可能性更高。

- 在電子檔案變成過時與不可存取前，機關被鼓勵進行電子檔案長期保存策略，並決定發展適當的長期保存技術。
- 在已升級 IT 環境，長期保存處置的執行主要是為確保現行的電子檔案，尤其是現存的業務核心檔案，可以從其原本的格式轉換為具功能性的新格式。

(C) 發展電子檔案長期保存策略之規劃(Planning)

- 維持電子檔案整合性與確實性執行不夠確實，轉置軟體與硬體系統將危及其有實證的價值。
- 機關對檔案格式實施長期保存處置，卻不確實評估處理程序，對電子檔案會有功能性、格式、架構與內容遺漏與限制情況，也可能對與檔案相關的詮釋資料有潛在性的遺漏。
- 在軟硬體成功升級後，電子檔案長期保存規劃將可使機關維持其功能性與整合性。
- 發展長期保存策略與挑選適當的實施方法，是機關內檔案與 IT 部門合作努力的結果。

(D) 發展電子檔案長期保存策略(Stratgy)

- 發展長期保存策略

發展長期保存策略包括確認需要長期保存資料、研究技術解決方案、測試建議之解決方案、備份確認要長期保存資料、執行長期保存處置(treatment)、稽核長期保存資料的整合、銷毀適當的來源資料與建立監督制度(Regimes)。

- 成功長期保存策略之需求

成功長期保存策略之需求必須慎重選擇與測試處理長期保存之軟硬體，採用非專屬且開放原始碼與詳述之資料格式，長期保存程序對機關中現行與非現行電子檔案均適用，長期保存當時即取得相關詮釋資料，用於轉置或轉換之複製(copying)或改版(Reformatting)必須在詮釋資料詳述，當檔案轉置、轉換、複製與改版，在來源檔案之副本被銷毀前，必須驗證並確認資料整合。

(3) 國家檔案局電子檔案長期保存實施方法(Approach)

- A. 轉換(Conversion)、封裝與模擬等技術的整合為電子檔案長期保存之實施方法。
- B. 採典藏(Archival)資料格式將電子檔案轉換或「正規化」，即 XML 標準機制。

(4) 澳洲開放原始碼工具

A. Xena V4.2(XML Electronic Normalizing for Archives)

Xena 請參閱圖 3-9。

- (A) Xena 是由 NAA 發展之開放原始碼軟體，對電子檔案長期保存有所助益。
- (B) Xena 以 Java 撰寫，可以跨系統執行，包括 Linux、Windows 及 OS X。
- (C) 標準採用：XML、ODF、PNG、JPEG、PDF、FLAC。
- (D) Xena 軟體執行下列 2 項重要工作有助於電子檔案長期保存：
 - 檢視電子檔案之檔案格式。
 - 將電子檔案轉換為開放格式以便長期保存。
- (E) Xena 功能性，中心模組加上「plug-in」模組可做到：
 - 檔案格式推測(File format guessing)。
 - 檔案正規化(File normalization)。
 - XML 封裝(XML encapsulation)。
 - 處理及資料確認(Process and data verification)。
 - 檔案查看(File viewing)。



圖 3-9 澳洲國家檔案局轉置工具 Xena 測試結果

B. DPR (Digital Preservation Recorder)

NAA 發展之程式組合之一，作為電子檔案長期保存系統之部分，並維護可受檢驗之電子紀錄活動過程的稽核軌跡 (Audit Trail)，將所有 NAA 處理過程紀錄儲存在電子資料庫中。DPR 的特性：

- (A) 採用 Java 執行環境以跨平台(Cross platform)。
- (B) 記錄 Metadata(Metadata recording)。
- (C) 使用者確認與描繪(User authentication and profiling)。
- (D) 多位使用者可同時執行 (Multiple concurrent users)。
- (E) 詳細的報表(Detailed reporting)。

5、中國大陸

(1) 國家檔案局

A. 電子文件長期保存格式應具有的特徵

- (A) 支持真實性
 - 支持數位浮水印、數位簽章與校正碼。
- (B) 格式透明
 - 格式開放，有公開發表的相應標準和技術規範。
 - 有與產品無關的技術專家組和標準化組支持該格式。

(C) 具可攜性元件

- 被多種作業系統和應用軟體支持。
- 支持多種儲存技術，或與儲存技術無關。
- 當用戶不能使用指定產品軟體時，可使用已有的介面讀取。
- 使用與設備無關的顏色規格實現準確列印和顯示，不必考慮軟硬體平台。

(D) 格式自我包含

- 不包含指定版本的格式之外的內嵌對象。
- 不包含外部對象的鏈接。

(E) 格式自我描述

用標準格式（通常是 XML）設置詮釋資料，描述對象的屬性特徵，滿足管理、保存與描述的需求。

(F) 固定顯示

- 維持固定的檔案頁面、章節、段落的邏輯組織結構，不因軟硬體平台和讀取器變化而不同。
- 以自然閱讀順序提供本文，以便文件能用基本本文編輯工具閱讀。

(G) 持續可解釋

不包含加密機制，也不包含加密選項。

(H) 持續可使用

- 支持無失真壓縮。
- 在壓縮機制中不使用分辨率的壓縮監控。

(I) 可轉換

- 支持其他格式轉換為長期保存格式。
- 支持過時的長期保存格式轉換為新的長期保存格式。

(J) 易儲存

- 格式具連貫性，資料結構合理，資料位元組數少。
- 支持在一個檔案中容納大量資料，不必把資料拆分成區塊。

- 具有整合能力，可以把幾個相關的資料對象整合到一個檔案中。

B. 電子檔案長期保存格式的選擇原則

(A) 滿足需求

電子檔案長期保存格式應具有長期保存格式的特性，滿足電子檔案長期保存的要求。

(B) 通用

- 符合相關的國際標準、國家標準或技術規範。
- 被廣泛應用和支持，採用的技術成熟、穩定，具有較強的生命力。

(C) 資訊安全

符合國家資訊安全政策，規避技術壁壘，維護資訊安全。

(D) 代價最小

選定的長期保存格式儘可能使電子檔案的管理成本最小化，這些成本包括使用、維護該格式軟體費用、相關格式轉換費用及人員培訓費用等。

(E) 限量

符合電子檔案格式的數量可能是無限增長的，但特定電子檔案保存處所接受的電子檔案長期保存格式的數量應是有限的。

(F) 可擴展

當某類電子檔案沒有合適的長期保存格式時，可採用工業標準格式作為長期保存格式。

(G) 評估

認定電子檔案長期保存格式，需要經過標準一致性評估。

C. 常見電子檔案格式

表 3-1 常見電子檔案格式

文件類型	格式	格式特徵
純文本文件	TXT格式	格式簡單透明、不含結構資訊和加密、具可攜性、能用基本文本編輯工具閱讀、資料位元組數少等。
	XML格式	遵循XML技術規範，格式開放、具可攜性、格式自我描述、不包含加密、易於轉換等。
格式化文本文件	UOF格式	遵循GB/T 20916-2007《中文辦公軟體文檔格式規範》，支持電子簽章、格式開放、具可攜性、格式自我描述、不包含加密等。
	RTF格式	格式開放、具可攜性、不包含加密、易於轉換等。
	WPS格式	支持電子簽章、可向其他本文格式轉換等。
版式文件	PDF/A格式	遵循ISO/DIS 19005《文件管理 電子檔案長期保存格式 第1部分 PDF1.4的使用(PDF/A-1)》，支持電子簽章、格式開放、具可攜性、格式自我包含、格式自我描述、固定顯示、不包含加密、可向其他本文格式轉換等。
	CEB格式	支持電子簽章、格式自我描述、固定顯示、可向其他本文格式轉換等。
	SEP格式	支持電子簽章、格式自我描述、固定顯示、可向其他本文格式轉換等。
圖像文件	TIFF格式	支持無失真壓縮、具可攜性、易於轉換、整合能力強等。
	JPEG-2000格式	遵循ISO 15444-1:2004《信息技術 JPEG 2000圖像編碼系統 核心編碼系統》，格式透明、支持無失真壓縮、具可攜性、易於轉換等。
	JPEG格式	遵循相關標準規範，格式透明、具可攜性、易於轉換等。
	GIF格式	支持無失真壓縮、格式透明、具可攜性、易於轉換等。
	PNG格式	支持無失真壓縮、格式透明、易於轉換等。
	DjVu格式	格式透明、格式具連貫性、具有整合能力、資料位元組數少等。

文件類型	格式	格式特徵
聲音文件	WAV 格式	支持浮水印技術、支持無失真或其他公開的壓縮演算法、易於轉換等。
	MP3 格式	遵循 GB/T 17191-1997《信息技術 具有 1.5Mbit/s 數據傳輸率的數字存儲媒體運動圖像及其伴音的編碼》，壓縮演算法公開、格式具連貫性、資料位元組數少、易於轉換等。
	WMA 格式	內置版權保護技術、格式具連貫性、資料位元組數少、易於轉換等。
	OGG Vorbis 格式	格式透明、格式具連貫性、資料位元組數少、易於轉換等。
視訊文件	AVI 格式	支持浮水印技術、支持無失真或其他公開的壓縮演算法、易於轉換等。
	MOV 格式	格式具連貫性、易於轉換等。
	MPEG 格式	遵循 GB/T 17191-1997《信息技術 具有 1.5Mbit/s 數據傳輸率的數字存儲媒體運動圖像及其伴音的編碼》或 GB/T 17975-2000《信息技術 運動圖像及其伴音信息的通用編碼》或 ISO/IEC 14496《信息技術 音視頻對象編碼》，壓縮演算法公開、具可攜性、易於轉換等。
注 1：表 1 所列的格式並非表示其通過了與本標準的一致性認證。		
注 2：表 1 所列僅為部分常見格式，更多格式參見互聯網媒體類型（IMT）。		

D. 電子檔案管理原則

- (A) 全程管理原則：對電子檔案從產生到永久保存或銷毀的整個生命週期進行全程管理。
- (B) 前端控制原則：在檔案形成階段或部分管理功能盡量在此階段實現。
- (C) 真實性保障原則：即保證檔案內容、結構和背景資訊經過傳輸、轉置等處理後依然與形成時的原始狀態一致。
- (D) 完整性保障原則：從制定合理的歸檔制度以及建立詮釋資料模型兩個方面著手，使每一份電子檔案的內容、結構和背景資訊沒有缺漏與損毀。

- (E) 可讀性保障原則：經過儲存、傳輸、壓縮、加密、媒體轉換與轉置等處理後能夠可以讀取和理解。

E. 有效性保證

- (A) 歸檔電子檔案的形成單位和檔案保管部門每年均應對電子檔案的讀取、處理設備的更新情況進行一次檢查登記。對磁性媒體每滿 2 年、光碟每滿 4 年進行一次抽樣檢驗，抽樣率不低於 40%，如發現問題應及時採取恢復措施。
- (C) 對磁性媒體上的歸檔電子檔案，每 4 年應轉存一次。原媒體同時保留時間不少於 4 年。
- (D) 各種不同類型的電子檔案，其檔案格式均應盡量採用 XML、PTF 與 TXT 格式；掃描圖片檔案採用 JPEG 與 TIFF 格式；視頻檔案採用 MPEG 與 AVI 格式；音頻檔案採用 MP3 與 WAV 等格式。

F. 電子檔案相應設備過時的對策

- (A) 再生性技術保護(媒體轉換技術)：將技術過時的媒體上的資訊適時地轉移到紙質或縮微影片上。
- (B) 建立電腦技術檔案館：收集過時的電腦軟硬體，提供利用，讀取那些過時技術記錄在媒體上的資訊內容。
- (C) 模擬：製造一種運作過時軟硬體之軟體。在這個軟體中對某一硬體與軟體的模擬，使得原來的軟硬體在形式上仍然可用。
- (D) 更新：在原來的技術環境下定時重寫電子檔案，為檔案館管理電子檔案的一項基本方法。
- (E) 轉置：隨技術變化適時改變電子檔案格式的處理過程，使資訊在將來可被存取。轉置要求電腦既可讀出舊格式也可將它寫在新格式上。

G. 電子檔案管理核心技術

- (A) 詮釋資料技術：描述電子檔案內容、結構和背景的资料，並加於封裝。
- (B) 長期可存取技術：模擬、轉置與媒體轉換技術。

(C) 系統安全及資訊安全技術：系統備份、防火牆、實體隔離與虛擬專用網等技術，保證電子檔案系統正常運作。

(2) 中國科學院

A. 系統架構

中國科學院長期保存系統架構請參閱圖 3-10 與圖 3-11。

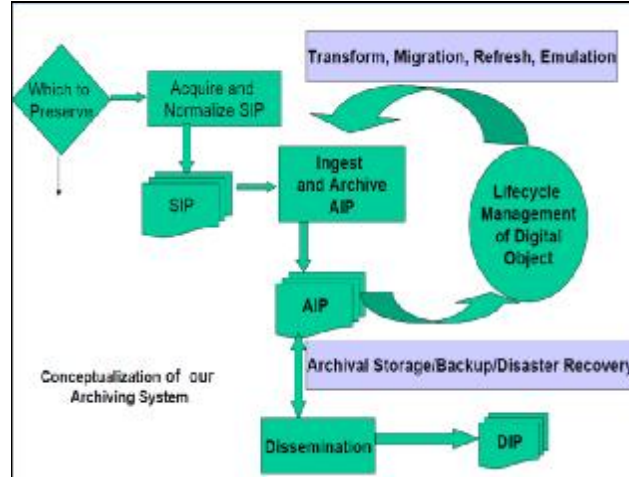


圖 3-10 中國大陸中國科學院長期保存系統架構圖一

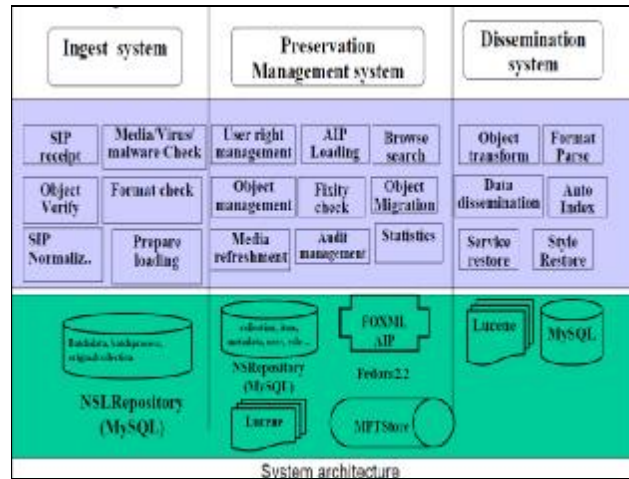


圖 3-11 中國大陸中國科學院長期保存系統架構圖二

B. 架構說明

(A) 以 OAIS 為基礎。

- (B) 採用 Fedora 作為基礎核心(Using Fedora as the fundamental core)。
- (C) 採用 MySQL 資料庫管理系統發展輸入系統與管理電子檔案(Using MySQL to develop ingest system and to manage the digital objects)。
- (D) 採用 Lucene 索引軟體發展電子檔案全文搜尋功能(Using Lucene to develop fulltext search function of the digital objects)。
- (E) 採用 MPTStore 儲存並搜尋資源描述架構(Resource Description Framework, RDF)三部分(Using MPTStore to store and search RDF triples)。

C. 擷取功能(Ingest Functions)

- (A) 提交資訊軟體(Submission Information Package, SIP)接收與註冊(SIP receipt and registration)。
- (B) 媒體/病毒/勒索軟體檢核(Media/Virus/Malware Check)。檔案確認(Object Verification(Object format, SIP format, Object number))。
- (D) SIP 正規化(SIP Normalization (PDF, FOXML))。
- (E) 備妥載入(Prepare for loading)。

D. 長期保存管理功能(Preservation Management Functions)

- (A) 檔案管理功能(Object Management Functions)
 - 電子檔案載入(Digital objects loading)。
 - 電子檔案瀏覽/搜尋(Browse/Search of digital objects)。
 - 電子檔案編輯/清除(Edit/Purge of digital objects)。
- (B) 長期保存管理功能(Preservation Management Functions)
 - 穩定性檢核(Fixity Check)。
 - 稽核管理(Audit Management)。
 - 統計(Statistics)。
- (C) 生命週期管理功能(Lifecycle Management Functions)
 - 檔案轉置(Object Migration)。

- 媒體清除(Media Refreshment)。

(二) 保存性詮釋資料 (Preservation Metadata)

依照美國康乃爾大學及密西根大學所發展之「美國的形成」計畫(Making of America II 1998) 對於數位保存詮釋資料的格式與功能分為三類，定義如下(註5)：

1、描述性詮釋資料 (Descriptive Metadata)

用於資源發掘與識別，可包含題名、作者、摘要與關鍵字等。例如機讀編目格式(MARC; MACHine Readable cataloging)、檔案編碼描述格式(EAD; Encoded Archival Description)、都柏林核心集(Dublin Core)。

2、結構性詮釋資料 (Structural Metadata)

用於協助系統呈現數位物件的資訊。

3、管理性詮釋資料 (Administrative Metadata)

提供資源的管理資訊，如建立日期與方式、檔案類型與其它技術資訊、使用權限等。

在國外，保存性詮釋資料(Preservation Metadata)已成為檔案界與圖書館社群相當普遍的研究與發展。保存性詮釋資料主要協助描述與定位資訊，提供保存完整與確實的資訊，較知名的相關發展如由國際共同合作的「電子記錄真確性永久保存研究計畫(InterPARES; International Research on Permanent Authentic Records in Electronic Systems project)」，目標是提出具備永久保存真確性的電子檔案之理論與方法。

在圖書館社群方面，英國大學研究圖書館聯盟數位化檔案庫方案(CEDARS; Consortium of University Research Libraries[CURL] Exemplars in Digital Archives)、澳洲國家圖書館(NLA; National Library of Australia)等計畫，則藉由美國線上電腦圖書館中心(OCLC; Online Computer Library Center)及美國研究圖書館組織(RLG; Research Libraries Group)之保存性詮釋資料工作小組(OCLC/RLG Working Group on Preservation Metadata)，進行保存性詮釋資料之比較與分析(註6)。以上這些計畫多根據或參照國際標準組織(ISO; International Standards Organization)所制定之開放式檔案資訊系統(OAIS; Open Archival information System)參考模式。

相關說明如下：

1、開放式檔案資訊系統 (OAIS ; Open Archival information System)

開放式檔案資訊系統，是國際標準組織 (ISO) 所屬美國國家太空總署 (NASA ; National Aeronautics and Space Administration) 的太空資訊系統諮詢委員會 (CCSDS ; Consultative Committee for Space Data Systems)，於西元 1999 年 5 月所發展的一套標準模式，是一個參考模式而非實作指南，依其標準提供系統間詮釋資料互通的架構，促使不同類型的詮釋資料得以交換與再使用。此模式可確立典藏應具備的相關資料 (註 7)，故許多研究計畫或方案均採用此模式發展其保存性詮釋資料之屬性。

OAIS Model 當中資訊物件 (Information Object) 由四種類別 (classes) 所組成：內容資訊 (Content Information)、保存描述資訊 (Preservation Description Information)、包裝資訊 (Package Information) 以及描述性資訊 (Description Information)。其中包存描述資訊 (Preservation Description Information) 建構與保存檔案相關活動的詮釋資料，並針對在長久保存的時間過程中，面對管理檔案所需的必要內容。但在內容資訊 (Content Information, Information Package 中另一個部分，與保存描述資訊在相同位階上) 中會記錄被保存檔案的靜態屬性 (例如保存時檔案的版本)；而保存描述資訊中除一般靜態屬性外，還會包括隨時間變動的狀態 (從檔案第一次被建立開始，被轉製為電子格式，到在儲存區中被保存)。

OAIS Model 對保存描述資訊提出四種分類：

- (1) 參考性資訊 (Reference)：描述識別系統，及識別子的指定機制，用以讓內外部的使用者能夠明確識別其內容資訊。
- (2) 情境資訊 (Context)：紀錄內容資訊與環境的關聯，包括內容資訊的建立的原因，及與其他內容資訊的關聯及原因。
- (3) 出處資訊 (Provenance)：紀錄內容資訊的歷史軌跡，包括出處，隨時間的內容變動過程，及對該內容所有的管理資訊。
- (4) 固定性資訊 (Fixity)：提供資料完整性及正確性的檢核與驗證，保證不會有未紀錄的資料異動發生。

2、美國研究圖書館組織詮釋資料典藏議題工作小組 (RLG Working Group on Preservation Issues of Metadata)

早於西元 1997 年 5 月，美國研究圖書館組織 (RLG ; Research

Library Group) 即成立詮釋資料保存議題工作小組。西元 1998 年 5 月，工作報告推薦下列保存性詮釋資料項目，詳列如下：

- (1) Date
- (2) Transcriber
- (3) Producer
- (4) Capture Device
- (5) Capture Details
- (6) Change History
- (7) Validation Key
- (8) Encryption
- (9) Watermark
- (10) Resolution
- (11) Compression
- (12) Source
- (13) Color
- (14) Color Management
- (15) Color Bar/Grayscale Bar
- (16) Control Targets

數位資訊長久取用與保存的重要性，在國外受到普遍重視，且有許多相關研究計畫不斷在進行中，而保存性詮釋資料即為其重要成果之一，目前國外仍存在一些執行上的討論：

- ❖ 對於詮釋資料概念的發展甚於實際應用，且少有對於發展保存性詮釋資料需具備之知能說明。
- ❖ 數位典藏在經濟與社會成本上的不確定，故應考量建立詮釋資料的成本效益問題。
- ❖ 詮釋資料的建立問題，包括詮釋資料的建立工具、典藏機構是否了解詮釋資料的價值以及貢獻等。根據美國研究圖書館組織 (RLG) 的報告指出，典藏數位物件所遭遇的問題，大部份來自於未做好描述來源的工作。

以上的議題對於國內的發展也同樣重要，對於詮釋資料功能的認知與應用，不應只侷限於資源的描述，而是要進一步擴充詮釋資料的規格與標準，以符合現實環境需求。

肆、結論與建議

電子檔案的長期保存技術有許多國內外學者提出不同的技術分類方法，綜合各家所言，我們可以大致歸類（Waugh, et al., 2000）為實體保存（Physical Preservation）、功能保存（Functional Preservation）以及組織保存（Organizational Preservation）。實體保存主要關注於儲存媒體的議題，可採更新（Refresh）方式為之。功能保存主要關注於保有原來的軟體環境，可配合執行的策略有轉置、模擬、標準化以及詮釋資料等。組織保存則關注於是否有足夠資訊以利組織使用該資訊從事其組織活動，可經由封裝策略予以實現。

有關於對檔案的定義，國內與國外差距甚大，我國的電子檔案定義為完成線上簽核之電子公文，但是經由本案對於機關公文製作與管理之調查發現，目前國內線上簽核尚未普及，也因此形成對於電子檔案的認知不一，因實務上存在許多電子文件（例如公務使用之簡報檔案）存在，卻無法予以妥善處理。至於國外的部分，國外所指之「electronic records」即等同於我國的傳統檔案，必須經由選擇之後，認為具有保存價值者，才能晉升為「records」，相較之下，我國電子檔案管理要點的規定要項如線上簽核，大部分屬於公文製作而非「records」形成之要件。因此，除整體制度難以採用國外之外，由於整個精神與內涵均與西方先進國家規劃不同，較難引用其研究成果及方法。

本案經由需求分析與方案評估，對電子檔案長期保存策略，電子檔案轉置必須以封裝、詮釋資料與檔案格式為基礎，並與過時必備及模擬相互搭配，以達成電子檔案長期保存的目標，因此，規劃能整合以轉置加上封裝為主，模擬與過時必備為輔之運作模式。轉置、模擬與過時必備整合為「電子檔案長期保存實驗室」（簡稱實驗室）。提出實驗室有關短期（至 98 年底）、中期（至 100 年底）、長期（100 年以後）策略之建議如下：

一、短期（98 年）建議

（一）電子檔案轉置

1、訂定作業要點附件九檔案格式增刪準則

附件九電子檔案格式則對長期保存具關鍵性，持續評估新興或國際標準之電子檔案類型及格式，據以研擬我國電子檔案長期保存檔案格式增刪準則，並提出我國電子檔案格式修訂建議。

2、建置我國電子檔案長期保存運作雛型平台

按長期保存規劃(Preservation Planning)、擷取系統(Ingest System)、長期保存管理系統(Preservation Management System)與傳播系統(Dissemination System)、格式知識庫、長期保存電子檔案、相關工具等之間關聯進行分析、評估與設計。此外，並分析國內外電子檔案長期保存相關工具，如檔案格式判讀、轉置、安全管控等工具，評估驗測 貴局技術移轉所得的電子檔案格式轉置工具，研究納入電子檔案管理作業程序之可行性，規劃與試作我國電子檔案長期保存運作雛型平台，並提出試作檢討。

刪除: 技術

刪除: 雛型

刪除: 。

刪除: 現有

刪除: 實

刪除: 技術

刪除: 雛型

3、執行格式知識庫分析、評估與設計

內容涵蓋檔案格式等資料，其中格式風險應併同轉置風險一起評估，規劃之格式知識庫必須與整體轉置架構平台整合，提供作業所需相關資訊，以彌補詮釋資料之不足。另知識庫必須提供簡便的維護與回饋機制，俾內容足以達到即時更新情況。

4、進行轉置相關工具蒐集、測試與評估

從市場上取得可行工具，不足之處再自行修改或發展。分析評估現有轉置相關工具，包括格式辨識與確認工具、各類型檔案轉置工具、封裝檔 NAA Viewer、中文處理軟體與技術移轉軟體等，提出我國電子檔案長期保存平台之工具導入模式，選用具開放性、功能性、可攜性與普及性工具。

5、擬定電子檔案長期保存轉置作業規範

轉置運作機制相關包括電子檔案格式清單、詮釋資料、技術知識庫、作業程序與各類轉置相關工具等，均應長期保存管理，軟體工具亦應考慮其轉置議題，以維持其長期可用性，順利達成轉置工作，並維持轉置後電子檔案的應有品質。

(二) 電子檔案模擬

建議現階段局暫不投入模擬策略的發展及建置，模擬策略目前國外仍在研究階段，等待未來成熟的成果，可減少檔案管理局在發展建置上 Try and Error 的花費，且目前國內對模擬策略仍未有立即性的需求，未來時機成熟仍然可在適當的時機投入，現階段建議將有限的資源移做其它更有效益的用途，雖暫不投入發展及建置，但仍需持續關注國際間模擬策略及技術的發展情況，歐盟的 Planets 及荷蘭的國家圖書館對於模擬策略一直有持續的研

究及成果的發表，值得重視。

(三) 過時必備

1、建立機關軟硬體資訊蒐集分享的知識庫

由於目前蒐集資料分析的結果，尚無迫切需要馬上收入實驗室的需求，所以實驗室現階段的任務，重心應著重在前兩項，輔助推廣電子檔案長期保存的概念到機關，並建立起資源分享的知識庫，唯有蒐集到的資訊夠可靠與深入，才能確實找出機關需要實驗室服務的實際需求，避免購買設備、人力、時間等的浪費。

2、以檔案管理局為範例，研擬電子檔案長期保存的策略

要找出解決電子檔案長期保存的策略，尋找一個適合的機關來作為範例配合研究，會更能貼近現實的狀況。檔案管理局本身擁有一些舊的公文管理系統與檔案格式，加上本身就對長期保存有深入的認識並負責制訂相關法規，以其為範例合作研究，對電子檔案長期保存的策略擬訂，必定會達到雙贏的效果。

3、彙整分析教育訓練過程中，機關所提供的資訊

近期檔案管理局舉辦了多場教育訓練，在經過專家學者的講解後，機關人員已對電子檔案長期保存有了更深的認識，課後提供的資訊也就會更有價值，可彙整回收的問卷與上課提問的內容，分析機關有無電子檔案長期保存的問題。

4、建置電子檔案專家黃頁雛型系統

建議利用檔案線上百科系統建立專家黃頁雛型系，以彙整電子檔案各領域的專家資訊，如電子檔案保存策略、電子檔案保存技術、檔案鑑定、檔案管理、檔案典藏、檔案格式儲存應用修復轉置、儲存媒體保存應用修復轉置等。

5、研提電子檔案長期保存實驗室建議徵求書

在電子檔案長期保存的解決策略中，「過時必備軟硬體典藏實驗室」與「轉置模擬實驗室」的關係密不可分，與檔案管理局專家們討論後傾向將兩實驗室合併，成立「電子檔案長期保存實驗室」，實驗室組織結構下還是分成轉置模擬與過時必備等分項做相關的研究，預計在民國 98 年底前提出實驗室建議徵求書，載明試作研究所需軟硬體設備及相關技術規格，以利辦理後續採購招標事宜。

二、中期（99~100 年）建議

（一）電子檔案轉置

1、建構電子檔案長期保存轉置平台

尋找有意願機關進行驗測，內容涵蓋電子檔案長期保存轉置平台、格式知識庫與轉置相關工具，並與封裝、模擬、過時必備等長期保存策略整合，以搭配電子檔案技術服務中心的運作。

2、擬定電子檔案長期保存轉置作業環境之管理

轉置環境包括硬體如主機、伺服器、網路設備，軟體包括系統軟體如作業系統與網路軟體，均應進行管理，以維持其正常運作，並應留意作業環境的安全性，如病毒與駭客入侵等安全議題。

3、擬定法規中規範可接受之轉置風險

電子檔案長期保存轉置策略有其風險性，評估風險因素，透過監控風險與降低風險，並在法規中加以規範，以降低對長期保存的衝擊。

4、擬定或修訂轉置相關之封裝與詮釋資料格式

參考國際電子檔案長期保存相關標準，配合轉置所需資訊進行封裝與詮釋資料格式研擬或修訂，作為後續轉置的依循。

（二）電子檔案模擬

建議局內在此時發展模擬策略知識庫與模擬相關工具試用、發展及整合，模擬策略知識庫的發展需要多方面的經驗、測試，建議可尋求國內學術界與國際相關研究單位的合作，以國內的機關做為實際的例子進行相關研究及發展，並配合模擬相關工具進行多方面的整合評估，以期能發展適合我國國情的模擬策略知識庫及模擬相關工具。

（三）過時必備

解決服務機關實際問題：找到機關的需求，開始購買相關軟體硬體提供服務。由於機關遇到問題通常都會先找廠商解決，廠商無法解決才會找到實驗室。如何找到廠商無法解決的問題，是實驗室研究機關需求時該注意的事項。

三、長期（100 年以後）建議

(一) 電子檔案轉置

1、建構以服務導向架構(SOA)為基礎之架構平台

葡萄牙國家檔案局建構 SOA 架構與相關技術值得持續關注，為提供各機關較佳之轉置服務，建立以 SOA 架構為基礎平台有其必要性。

2、建構以網路服務為基礎之遠端服務平台

網路服務是必然趨勢，未來平台必須網路驗證、確認與安全整合，提供遠端轉置服務機制。

(二) 電子檔案模擬

建議局內在此時發展整合國家檔案電子檔案長期保存系統的模擬策略運作平台，包含模擬器的發展，可以參考國外如英國 PRONOM 電子檔案知識系統、Planets 的 GRATE 模擬器遠端執行系統與荷蘭的 Dioscuri 模擬器，建立以電子檔案知識庫為核心的模擬策略運作平台，如此便能確保模擬策略執行並降低模擬策略的風險。

(三) 過時必備

與廠商和其他相關實驗室建立合作關係：很多過時的設備，可能都要價不斐，使用率卻很低，譬如大盤的膠卷放映機，由於膠卷放置長時間很容易脆化，機關就算有放映機可能也不敢播放，必須委託專業廠商作處理。實驗室並不需要買進相關處理設備與訓練相關人力，可與廠商建立合作關係，相關實驗室亦然。

伍、註釋

- 註 1 機關電子檔案管理作業要點第四條第一項規定，當檔案隨著軟體更新或是系統更新時或是任何時空環境變遷仍保有檔案的完整資訊（檔案管理局 2006）
- 註 2 機關電子檔案管理作業要點第四條第二項規定，意指檔案不可遭非正常程序的竄改，以確保檔案資訊之真實，以利後續確認課責（檔案管理局 2006）。
- 註 3 機關電子檔案管理作業要點第四條第三項規定，當時空環境科技進步時，檔案仍可繼續保有使用的方便性（檔案管理局 2006）。
- 註 4 檔案簽核者均有電子簽章與相關證明其個人之簽署，系統應具備相關檢證機制，讓簽署者具備不可否認性，以建立課責。
- 註 5 <http://www.ukoln.ac.uk/metadata/presentations/rlg-npo/warwick.html>
- 註 6 <http://www.nla.gov.au/padi/topics/32.html>
- 註 7 http://www.oclc.org/research/projects/pmwg/pm_framework.pdf