



# 電子檔案保存工具之 102 年度成果

## Achievement of Preserving Electronic Archives & Records Suite (PEARS) Development in 2013

■ 高夢謙 Kao, Meng-Chien

電子檔案保存實驗室技術服務專員

Technical Service Executive, Electronic Records Preservation Laboratory

### 壹、前言

國家發展委員會檔案管理局以電子檔案保存為重心，永續保存政府電子檔案為宗旨，為協助解決國家及機關電子檔案面臨之保管、應用與安全等問題，成立「電子檔案保存實驗室」（以下簡稱實驗室），開發「電子檔案保存工具」（Preserving Electronic Archives & Records Suite，簡稱 PEARS），期望達到電子檔案可及性（accessibility），藉由法規制定一致性的線上簽核電子檔案格式，並限縮附件檔格式機制，配合法定保存年限，適時辦理電子檔案格式轉置，維持電子檔案及其管理系統的可用性。電子檔案格式轉置是目前國際間最常被使用的保存方法，以改變原始電子檔案格式結構的特性，使電子檔案可因應資訊科技環境的變化，仍能有效讀取檔案的內容。

本文將介紹電子檔案保存工具提供的主要功能及 102 年度精進部分，讓機關同仁能妥善運用保存技術與工具，以有效降低電子檔案風險並有效保存與管理電子檔案。

### 貳、電子檔案保存工具精進

#### 一、主要功能

文書及檔案管理電腦化作業規範明訂各機關應配合電子檔案之保存年限，評估保存維護成本，選擇適當儲存媒體，適時辦理電子檔案之轉置。電子檔案保存工具主要提供電子檔案格式轉置及品質驗證兩大功能，協助機關檔案管理人員執行轉置作業。

（一）檔案轉置功能：提供單一檔案及批次檔案的格式轉置功能（如下頁圖 1）。將電子檔案格式轉置為符合文書及檔案管理電腦化作業規範附錄 8 之電子檔案格式，並產生轉置報告及轉置實體檔案，以維持電子檔案可讀性。

1. 文件類型格式轉置（DOC、DOCX、WDL、POSTSCRIPT、XLS、PPT、EML）。
2. 影像類型格式轉置（TIFF、JPEG、GIF、LEH、PCD）。
3. 視訊類型格式轉置（MPEG-2、WMV）。



圖 1：電子檔案格式轉置項目

資料來源：電子檔案保存實驗室

4. 音訊類型格式轉置（MP3）。

5. 封裝檔外部檔案格式轉置。

（二）驗證功能：係整合國立清華大學所開發的品質驗證軟體，利用客觀的數學公式檢測檔案品質的差異，確保轉置後檔案與原始檔案內容的正確性，提供影像及視訊等 2 種類型的品質驗證。

## 二、102 年度功能精進

（一）檔案轉置功能整合匯入「檔案類型清查詳細清單（HTML5<sup>（註 1）</sup>）」

電子封裝檔工具箱之電子檔案技

術鑑定工具，可以協助檔案管理人員執行電子檔案清查及鑑定作業（如下頁圖 2）；如機關檔案管理系統未具備電子檔案技術鑑定功能，可將待鑑定電子封裝檔路徑清單及實體檔案匯出至本機電腦，透過電子封裝檔工具箱之電子檔案技術鑑定工具的介面，以進行技術鑑定作業，並產生技術鑑定報告。

檔案管理人員可透過電子封裝檔工具箱所產生的技術鑑定報告之附件 10「檔案類型清查詳細清單（HTML5）」，



查看是否有辦理電子檔案格式轉置需要。若機關檔案管理系統未提供格式轉置功能時，可透過電子檔案保存工具的介面控制，匯入前述清單，以同時進行多種電子檔案格式轉置，並產生轉置

報告及轉置實體檔案。檔案管理人員可視轉置情形，將轉置實體檔案匯入檔案管理系統，並更新相關檔案詮釋資料（如圖3）。

1. 選擇處理類型為「檔案類型清查詳細

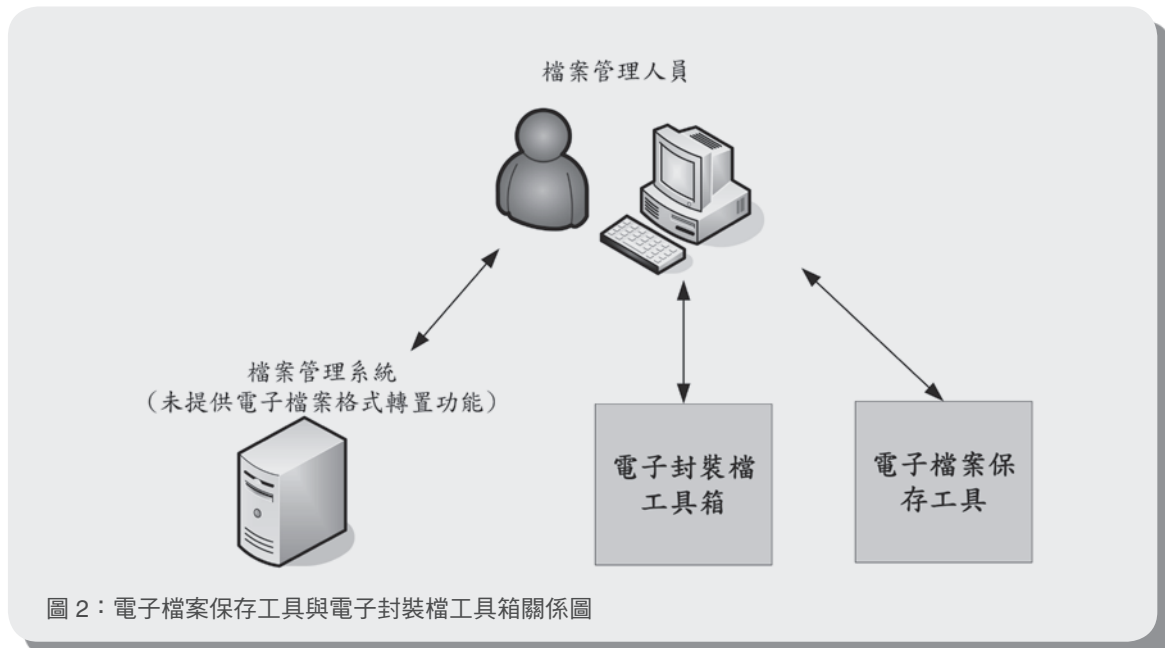


圖2：電子檔案保存工具與電子封裝檔工具箱關係圖

資料來源：電子檔案保存實驗室

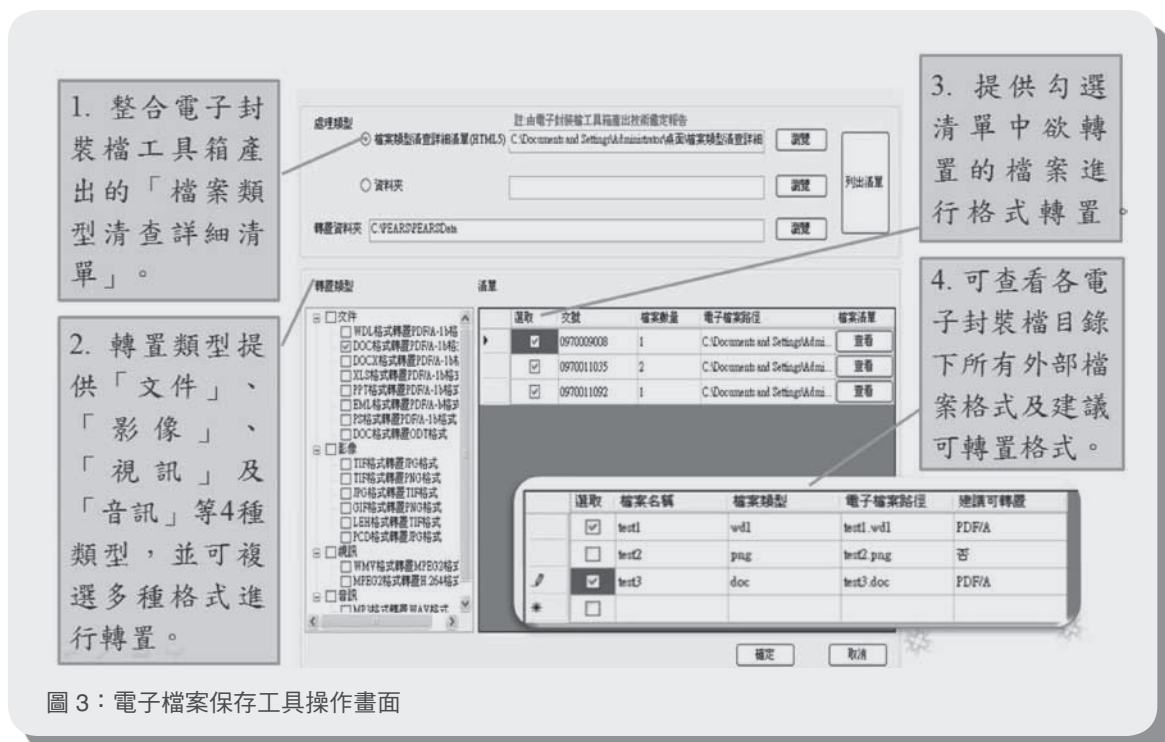


圖3：電子檔案保存工具操作畫面

資料來源：電子檔案保存實驗室



清單 (HTML5)」。

- 點選「列出清單」按鈕，系統將會讀取附件 10 檔案類型清查詳細清單的內容，並將內容呈現於清單中。
- 可點選檔案清單欄位的「查看」按鈕，查看各電子封裝檔所記載的外部檔案，於建議可轉置欄位顯示可轉置格式。
- 轉置類型主要分「文件」、「影像」、「視訊」及「音訊」等 4 種類型，可複選多種格式同時轉置，以批次辦理電子檔案格式轉置。

## (二) 擴充「PCD 格式轉置 JPEG 格式」及「LEH 格式轉置 TIFF 格式」

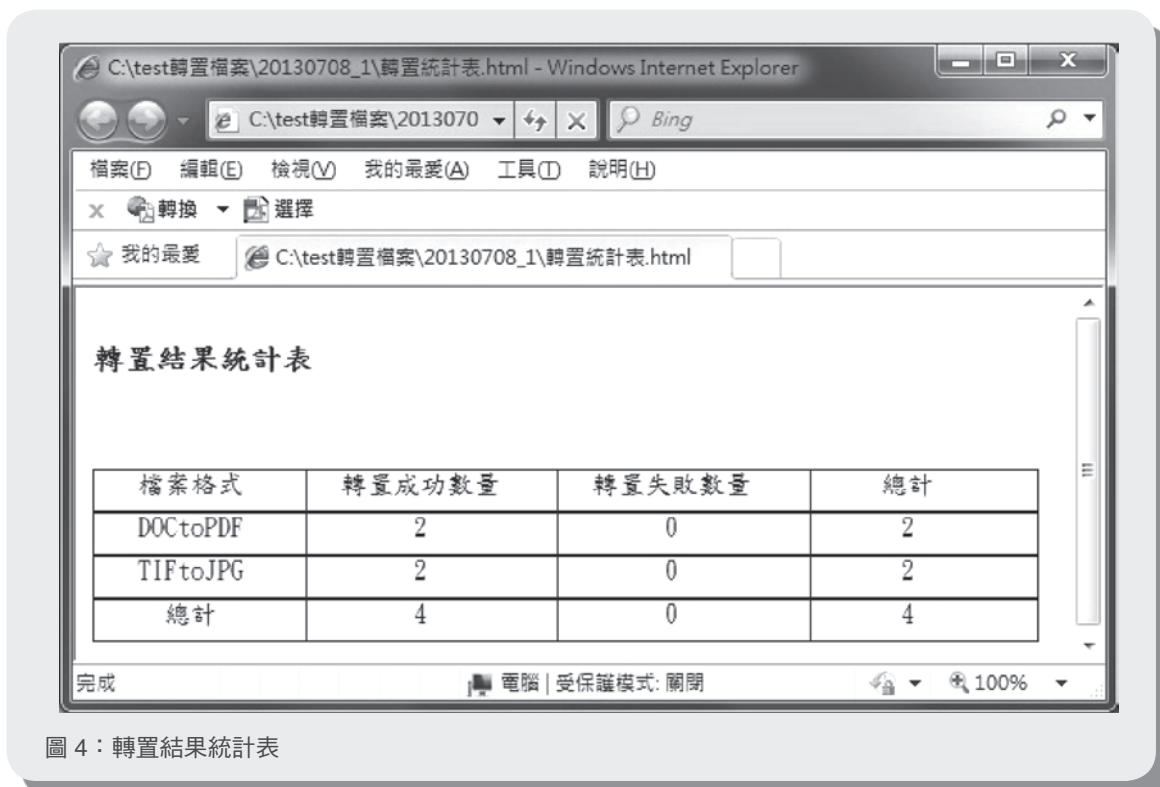
PCD (Photo CD) 為柯達 (Kodak) 公司一種獨家自製的相片光碟格式，無法使用一般瀏覽器開啟，必須使用支援

PCD 格式的軟體才能檢視；LEH 檔案為一種自製編碼加密過的格式，無法用一般瀏覽器開啟，須做解碼的動作。因此，配合國家及機關檔案轉置需求，新增此 2 種影像類型檔案格式轉置功能，並使用 JHOVE<sup>(註 2)</sup> 格式檢測工具進行轉置後檔案的檢測，以確保格式的正确性。

## (三) 增修轉置報告

考量使用者與程式可讀性將轉置報告修改為 HTML5 格式，並將內容增修為「轉置結果統計表」及「轉置檔案詳細清單」。

- 轉置結果統計表提供檔案轉置成功及失敗的統計數量 (如圖 4)。
- 轉置檔案詳細清單列出所有轉置檔案詳細清單資料 (如下頁圖 5)。



資料來源：電子檔案保存實驗室





序號	檔案路徑	檔案名稱	檔案大小	轉置前檔案格式	電子檔案路徑	轉置後檔案名稱	轉置後檔案大小	轉置後檔案格式	轉置後電子檔案路徑	轉置狀態
097000000	C:\102400001\102400001\111\1005-10050-0001-0004\007000000-097000000.doc	007000000.doc	21K	doc	C:\102400001\102400001\111\1005-10050-0001-0004\007000000-097000000.pdf	007000000.pdf	61K	pdf	C:\test\轉置檔案\101307000_1\007000000.pdf	成功
097001005	C:\102400001\102400001\111\1005-10050-0001-0003\007001005-097001005.tif	007001005-1.tif	21K	tif	C:\102400001\102400001\111\1005-10050-0001-0003\007001005-097001005-2.tif	007001005-2.tif	38K	tif	C:\test\轉置檔案\101307000_1\007001005-097001005-2.tif	成功
097001005	C:\102400001\102400001\111\1005-10050-0001-0003\007001005-097001005.tif	007001005-2.tif	21K	tif	C:\102400001\102400001\111\1005-10050-0001-0003\007001005-097001005-2.tif	007001005-2.tif	38K	tif	C:\test\轉置檔案\101307000_1\007001005-097001005-2.tif	成功
097001002	C:\102400001\102400001\111\1005-10050-0001-0004\007001002-097001002.doc	007001002.doc	21K	doc	C:\102400001\102400001\111\1005-10050-0001-0004\007001002-097001002.pdf	007001002.pdf	61K	pdf	C:\test\轉置檔案\101307000_1\007001002.pdf	成功

圖 5：轉置檔案詳細清單

資料來源：電子檔案保存實驗室

## 參、結語

實驗室期望透過電子檔案保存工具之提供，協助機關檔案管理人員執行電子檔案格式轉置作業，以解決過時技術或檔案無法保存及應用的問題，降低電子化檔案的風險。

102 年度電子檔案保存工具精進的方向，主要結合了 HTML5 的格式，以 HTML5 網頁文件格式匯入及開啟轉置的資訊，相較過去使用 Office 文件格式，不但快速便捷許多，更提高了電子檔案的可用（讀）性。而電子檔案保存工具也會隨著科技的轉變，蒐集與彙整各機關意見與問題回饋，做為精進研究的方向，來擴充格式的轉置，以確保國家或機關保管之電子檔案能符合現在與未來典藏保存及應用需求。

為了擴大提供對機關服務的範圍與品質，將研究成果推廣至各機關及受理各機關委託，並持續研發及實驗電子檔案保存技術，訂定相關作業管理及營運機制，據以發展相關資訊系統及標準作業程序，同時提供機關相關技術諮詢服務，如有轉置或保存相關需求與問題，竭誠歡迎您與實驗室聯絡，將提供最佳服務與技術諮詢，協助您解決問題（網址：<http://erlp.archives.gov.tw>，電話：（02）2513-6099）。

## 註釋

註 1：由國際網路國際標準組織 W3C 所制定，用於規範網頁文件標準的超文件標示語言（Hyper Text Markup Language，簡稱 HTML），目前已經發展到第五版（HTML5）。

註 2：JHOVE 是由 JSTOR 與哈佛大學圖書館共同發展電子檔案型態描述（Characterization）工具，經由特徵比對，推定電子檔案格式，判斷電子檔案之格式完好且具有效性。