

## 目錄

壹、簡介 .....	- 2 -
貳、沿革說明 .....	- 3 -
參、PDF 介紹.....	- 4 -
肆、轉置品質驗證方法說明 .....	- 8 -
伍、轉置程序 .....	- 10 -
陸、PPT 格式轉置 PDF/A 格式 .....	- 11 -
柒、XLS 格式轉置 PDF/A 格式 .....	- 18 -
捌、EML 格式轉置 PDF/A 格式 .....	- 24 -
玖、MP3 格式轉置 WAV 格式 .....	- 30 -
壹拾、結論與問題檢討 .....	- 35 -

## 壹、簡介

近年來由於電子化政府之推動，機關電子檔案之類型與數量均呈現急遽成長，且資訊技術發展日新月異，電子檔案之格式與儲存媒體，幾乎是以3至5年為一個世代快速轉變。因此，如何確保被保存之電子檔案能持續被存取與利用，已成為一個非常重要之議題。目前已有許多長期保存方法與策略被廣泛地探討，但仍以轉置、封裝、模擬與系統保存等為主要及常見之保存策略，其中又以轉置為最常使用之保存策略。然而電子檔案每經過1次轉置都有可能造成資訊的流失，甚至可能是極為珍貴的資料。因此，本報告彙整本年度開發整合之轉置工具與驗證方式，提出檢討與解決方式。

## 貳、沿革說明

98 年度之轉置作業為影像檔案轉置(TIFF 格式轉置 JPEG 格式)、文字檔轉置(WDL 格式轉置 PDF 格式)、視訊檔案轉置(MPEG-2 格式轉置 H.264 格式)，相關轉置工具業已整合至電子檔案長期運作平台。

99 年度之轉置作業為文字檔轉置(DOC 格式轉置 PDF/A 格式及 DOC 格式轉置 ODT 格式)、視訊類型轉置(WMV 格式轉置 MPEG2 格式)及電子檔案封裝檔格式轉置。

100 年度之轉置作業為文字檔轉置(DOCX 格式轉置 PDF/A 格式及 POSTSCRIPT 格式轉置 PDF/A 格式)，並將相關轉置工具整合至電子檔案保存工具。

101 年度之轉置作業為文件類型轉置(PPT 格式轉置 PDF/A 格式、XLS 格式轉置 PDF/A 格式及 EML 格式轉置 PDF/A 格式)及聲音類型轉置(MP3 格式轉置 WAV 格式)，並將相關轉置工具整合至電子檔案保存工具。

## 參、PDF 介紹

### 一、特色

#### (一)PDF 為結構化物件

將 Office 文件直接轉為 PDF 檔，可保持其結構化物件的文字、向量圖形，得以全文檢索，相較於 TIFF 檔案格式保存的檔案，在被搜尋之前必須先經過文字辨識(OCR)，惟辨識結果將受圖形清晰度及辨識軟體辨別度影響，須人為的校正。

#### (二)PDF 檔案的大小

結構化物件的 PDF 檔通常比 TIFF 檔案來得小，也提供較好之呈現品質，在電子傳遞上較具優勢。

#### (三)PDF 檔案裡可以儲存各式的詮釋資料

標題、作者、建立日期、修改日期、主旨、關鍵字等詮釋資料得以保存，使得 PDF 檔案可自動化分類，完全無需人為介入。

#### (四)PDF 檔案之內容與裝置無關

PDF 檔案因其保存字型、顏色、位置等相關屬性資料，日後不會因新的輸出設備不支援舊版規格而產生呈現的問題。

PDF 檔案格式係國際上被廣泛接受的標準電子文件格式，許多網際網路應用的表單與文件亦以 PDF 檔案格式儲存，大多數的電腦也已經安裝免費 PDF 閱讀程式 Adobe Acrobat Reader。

近年來電子公文與附件的呈現、交換、保存上普遍地採用 PDF 檔案格式，「文書及檔案管理電腦化作業規範」附錄中已明訂 PDF 格式為電子檔案格式之一。

#### (五)PDF 的四種主要國際標準

1. PDF/A：A 是 Archive，是歸檔用之 PDF 標準，PDF/A 只允許唯讀的屬性，即使過了數十年以後，依舊被讀取，而且文件呈現格式與儲存時一模一樣，非常適用於電子檔案長期保存使用。
2. PDF/X：X 是 eXchange，用於交換用的 PDF 標準，屬印刷上運用之檔案交換格式。
3. PDF/H：H 是 Health，醫療用設備可讀取之 PDF 標準，屬目前醫療體系之標準檔案交換格式。
4. PDF/E：E 是 Engineering，工程工業上運用的 PDF 標準，屬工業上交換檔案時的標準格式。

PDF/A19005-1:2005 係經 ISO(International Standard Organization)定義為電子檔案長期保存格式標準。19005-1:2005 標準 document management-Electronic document file format for long-term preservation-Part 1: Use of PDF 1.4 (PDF/A-1) (ISO 19005-1:2005)，因此，辦理電子檔案長期保存時，應符合 PDF/A 標準。

## (六)PDF/A 標準規範

1. 不能包含音訊或視訊資料：如 PDF 檔案中含有音訊或視訊資料，必須額外之影音解碼器才能開啟，將與影音解碼器相依而導致保存困難。
2. 不能執行 JavaScript 與其他夾帶附件的功能：夾帶附件的檔案使得 PDF 檔案需要額外程式處理附件檔；JavaScript 動態內容將使呈現行為控制不易。
3. 文字必須內嵌：避免不同電腦上文字編碼與字體所帶來之視覺上差異。
4. 顏色必須與設備無關（device-independent）：於任何顯示設備上都可以正確顯示原本顏色，必須符合 ICC(International Color Consortium)國際色彩協會之顏色內嵌規範，避免不同電腦色彩演繹所帶來之視覺差異。
5. 避免使用透明物件。
6. 不可以加密。
7. 必須包含詮釋資料（Metadata），以方便搜尋。

## (七)PDF/A 標準

1. PDF/A-1b：基本上完全符合 PDF/A 標準，同時也是符合標準之最小要求，達到本類要求屬 B 等級相容，可以確保以後文件開啟時以一模一樣之格式呈現。
2. PDF/A-1a：PDF/A-1a 為確保文件之邏輯架構，規定文字內容必須依讀取順序排列，以便萃取文字，對於日後在移動裝置上之讀取是相當重要的。由於多

數的移動裝置因受限於畫面大小，其文字必須重新排列。此標準除了要符合 PDF/A-1b 所有規範外，還需要加入 Tag(標籤)，以利各種裝置，如手持裝置或是閱讀輔助裝置等，閱讀 PDF/A 檔案。

3. PDF/A-2：此標準加入了 PDF1.5、1.6、1.7 中的新功能，具有相容性，所有符合 PDF/A-1 的檔案都符合 PDF/A-2，但符合 PDF/A-2 的檔案不一定符合 PDF/A-1。

## 肆、轉置品質驗證方法說明

電子檔案保存工具針對文字檔案類型格式轉置作業之品質驗證方式，係將文字類型(PPT 及 XLS 格式)之來源檔案及轉置後檔案(PDF/A-1b 格式)，分別轉置為 JPEG 檔，再將兩個 JPEG 檔案進行影像結構相似性指標(Structure Similarity, SSIM)及通用性影像品質指標(Universal Quality Index, UQI)，以檢測轉置前後的檔案品質。針對聲音檔類型格式轉置作業之品質驗證方式，係將轉置後檔案利用 JHOVE 格式檢測工具進行檢測，檢測是否為 WAV 聲音檔的格式。

### 一、通用性影像品質指標 UQI(Universal Quality Index)

通用性影像品質指標以相關係數降低、亮度改變及對比改變等因素來判定差異度。其最佳及最差的量化值介於 1 至 100 之間，若量化值 $\geq 60$ ，代表品質可接受，若量化值 $< 60$ ，則代表品質不可接受，因此可得到較具體的評估值。

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}$$

圖1. UQI 計算公式

### 二、結構相似性指標 SSIM (Structural Similarity)

結構相似性指標係用於測量兩幅影像間相似性之方法，係從圖像結構角度，以亮度、對比度及反映場景中物體結構的屬性因素，分別就其屬性之平均值、標準差及變異數作為衡量指標。其最佳及最差的量化值介於 1 至 100 之間，若 SSIM=100 表示與原始檔案完全一致，若 SSIM $\geq 60$  代表品質可以接受，若 SSIM $\leq 60$ ，代表品質不可以接受，肉眼即可察覺到明顯的畫面劣化，因此被判定為無實際觀賞價值。



### 三、JHOVE 格式介紹

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，功能摘述如下：

(一)識別(Identification)：經由特徵比對，推定電子檔案格式。

(二)確認 (Validation)：判斷電子檔案之格式完好 (Well-formed)且具有效性(Valid)。

(三)提供 Module：AIFF、ASCII、Bytestream、GIF、HTML、JPEG、JPEG2000、PDF、TIFF、UTF-8、WAVE、XML 等格式辨識模組。

(四)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。

## 伍、轉置程序

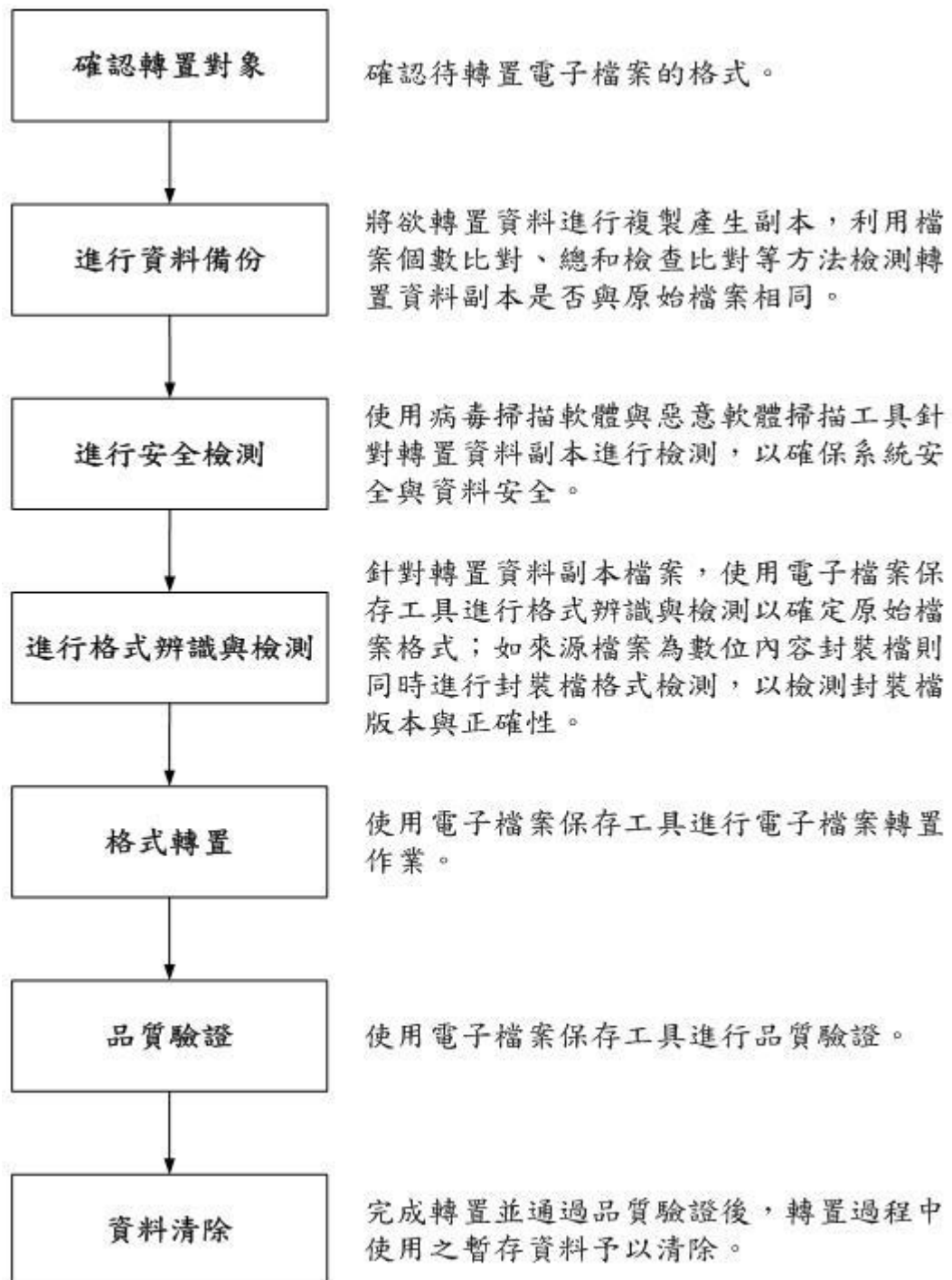


圖2. 轉置程序圖

## 陸、PPT 格式轉置 PDF/A 格式

### 一、轉置目的

Microsoft Office 目前被廣為使用之辦公室軟體，許多電子公文附件皆以 Microsoft Office 之檔案格式儲存(DOC、DOCX、XLS 及 PPT 等)，由於 Microsoft Office 檔案格式並未對外公開，在歷次 Office 改版過程中也遇到各版本呈現方式不一致之情況，因此該格式並不適合做為長期保存格式。

經本專案評估後，建議文字檔類型採用 PDF/A 作為電子檔案格式轉置之目的格式。

### 二、轉置測試環境

表1. 轉置測試環境表

項目	規格
中央處理器	Pentium (R)Dual-Core CPU P8700 @ 2.6GHz
硬碟	110GB
記憶體	3.0GB
作業系統	Microsoft Windows XP Professional Version 2002 Service Pack 3
顯示晶片	Intel G33/G31 Express ChiPSet Family

## 三、檔案轉置結果

## (一)測試檔 1

表2. 測試檔 1 資料表

電子檔案基本資料		
檔案名稱		轉置.ppt
檔案個數		1
掃毒工具名稱/版本		OfficeScan/6.3.1
病毒掃描結果		一切正常沒有病毒
轉置花費時間		31 秒
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	PPT
Size 大小	123KB	99KB
品質驗證	SSIM	100
	UQI	100

## (二)測試檔 2

表3. 測試檔 2 資料表

電子檔案基本資料		
檔案名稱		1010716 考試院保訓會高階班參訪簡報(組長)_1010711.ppt
檔案個數		1
掃毒工具名稱/版本		OfficeScan/6.3.1
病毒掃描結果		一切正常沒有病毒
轉置花費時間		39 秒
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	PPT
Size 大小	421KB	758KB
品質驗證	SSIM	100
	UOI	100

## (三)測試檔 3

表4. 測試檔 3 資料表

電子檔案基本資料			
檔案名稱		儲存媒體轉置實務.ppt	
檔案個數		1	
掃毒工具名稱/版本		OfficeScan/6.3.1	
病毒掃描結果		一切正常沒有病毒	
轉置花費時間		192 秒	
比較			
		<u>轉置後檔案</u>	原始檔案
電子檔案格式		PDF/A	PPT
Size 大小		4226KB	3390KB
品質驗證	SSIM	100	
	UQI	100	

## (四)測試檔 4

表5. 測試檔 4 資料表

電子檔案基本資料		
檔案名稱	960907-5.ppt	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	200 秒	
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	PPT
Size 大小	258KB	598KB
品質驗證	SSIM	100
	UOI	100

## (五)測試檔 5

表6. 測試檔 5 資料表

電子檔案基本資料		
檔案名稱		國立教育廣播電台參訪簡報 1000614.ppt
檔案個數		1
掃毒工具名稱/版本		OfficeScan/6.3.1
病毒掃描結果		一切正常沒有病毒
轉置花費時間		89 秒
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	PPT
Size 大小	417KB	416KB
品質驗證	SSIM	80
	UQI	100

## (六)轉置前後檔案容量與品質比較

1. 轉置前後檔案容量比較：轉置前後檔案容量如下表，測試結果顯示，轉置後檔案 PDF/A 格式之平均容量比原始檔案 PPT 格式大。

表7. PPT 格式轉置前後檔案容量比較表

序號	檔案名稱	轉置後檔案	原始檔案
1.	轉置.ppt	123KB	99KB
2.	1010716 考試院保訓會高階班 參訪簡報(組長)_1010711.ppt	421KB	758KB
3.	儲存媒體轉置實務.ppt	4226KB	3390KB
4.	960907-5.ppt	258KB	598KB
5.	國立教育廣播電台參訪簡報 1000614.ppt	417KB	416KB
平均		1089KB	1052.2KB

2. 轉置前後檔案品質驗證：利用 PDFCreator 工具將測試檔案 PPT 格式及轉置後檔案 PDF/A 格式，分別轉置為 JPEG 格式，再將 2 個 JPEG 格式檔案進行 SSIM 及 UQI 之品質驗證，轉置前後檔案品質驗證比較如下表。SSIM 值及 UQI 值大於 60 代表可接受。其中「國立教育廣播電台參訪簡報 1000614.ppt」的 SSIM 未達 100 分原因在於品質驗證需將測試檔案 PPT 格式及轉置後檔案 PDF/A 格式皆轉置為 JPEG 圖檔，其轉置為 JPEG 圖檔會產生一些誤差，例如：圖檔顏色的深淺或位置偏移等，導致品質驗證無法達到 100 分，但以肉眼的方式檢視 PPT 格式與 PDF/A 格式內容部分並無差異。因此測試 5 個測試檔計算平均值，SSIM 平均值約 96，UQI 平均值為 100，由此可知 SSIM 值及 UQI 值皆已符合標準。

表8. PPT 格式轉置前後檔案品質比較表

序號	檔案名稱	SSIM	UQI
1.	轉置.ppt	100	100
2.	1010716 考試院保訓會高階班 參訪簡報(組長)_1010711.ppt	100	100
3.	儲存媒體轉置實務.ppt	100	100
4.	960907-5.ppt	100	100
5.	國立教育廣播電台參訪簡報 1000614.ppt	80	100
平均		96	100

#### 四、檔案轉置結果

若 PPT 格式檔案簡報頁中只有圖片內容，此簡報頁面轉置時，工具將無法辨識其版面為直式或橫式，轉置後會轉成直式頁面，如圖 3 所示。針對此問題，經測試於轉置簡報頁面加上文字，轉置後的版面則會正常呈現，如圖 4 所示，因此修改工具，於轉置時系統自動加註文字「由 PPT 轉置」來解決此問題。





圖3. ppt 檔頁面為橫式，轉置後會轉成直式頁面

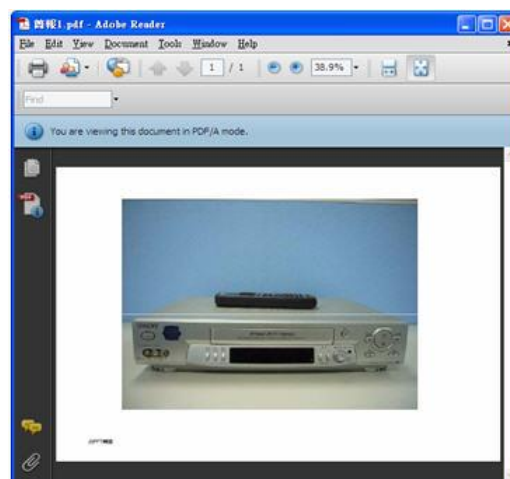


圖4. ppt 檔頁面為橫式，轉置後會轉成直式頁面

## 柒、XLS 格式轉置 PDF/A 格式

### 一、轉置目的

Microsoft Office 目前被廣為使用之辦公室軟體，許多電子公文附件皆以 Microsoft Office 之檔案格式儲存(DOC、DOCX、XLS 及 PPT 等)，由於 Microsoft Office 檔案格式並未對外公開，在歷次 Office 改版過程中也遇到各版本呈現方式不一致之情況，因此該格式並不適合做為長期保存格式。

經本專案評估後，建議文字檔類型採用 PDF/A 作為電子檔案格式轉置之目的格式。

### 二、轉置測試環境

表9. 轉置測試環境表

項目	規格
中央處理器	Pentium (R)Dual-Core CPU P8700 @ 2.6GHz
硬碟	110GB
記憶體	3.0GB
作業系統	Microsoft Windows XP Professional Version 2002 Service Pack 3
顯示晶片	Intel G33/G31 Express ChiPSet Family

## 三、檔案轉置結果

## (一)測試檔 1

表10. 測試檔 1 資料表

電子檔案基本資料		
檔案名稱		Magnetic.xls
檔案個數		1
掃毒工具名稱/版本		OfficeScan/6.3.1
病毒掃描結果		一切正常沒有病毒
轉置花費時間		42 秒
備註		內容含表格及圖表
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	XLS
Size 大小	206KB	563KB
品質驗證	SSIM	90
	UQI	90

## (二)測試檔 2

表11. 測試檔 2 資料表

電子檔案基本資料		
檔案名稱	內政部消防署磁片清單.xls	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	26 秒	
備註	內容含表格	
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	XLS
Size 大小	40KB	39KB
品質驗證	SSIM	80
	UOI	100

## (三)測試檔 3

表12. 測試檔 3 資料表

電子檔案基本資料		
檔案名稱		各廠商連絡相關資訊.xls
檔案個數		1
掃毒工具名稱/版本		OfficeScan/6.3.1
病毒掃描結果		一切正常沒有病毒
轉置花費時間		25 秒
備註		內容含表格
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	XLS
Size 大小	82KB	18KB
品質驗證	SSIM	100
	UQI	100

## (四)測試檔 4

表13. 測試檔 4 資料表

電子檔案基本資料		
檔案名稱	展示設備清單.xls	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	162 秒	
備註	內容含表格	
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	XLS
Size 大小	147KB	187KB
品質驗證	SSIM	100
	UOI	100

## (五)測試檔 5

表14. 測試檔 5 資料表

電子檔案基本資料		
檔案名稱	檔管局典藏組(正)轉置驗證表.xls	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	14 秒	
備註	內容含表格	
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	XLS
Size 大小	58KB	47KB
品質驗證	SSIM	80
	UQI	80

## (六)轉置前後檔案容量與品質比較

1. 轉置前後檔案容量比較：轉置前後檔案容量如下表，  
測試結果顯示，轉置後檔案 PDF/A 格式之平均容量比  
原始檔案 XLS 格式小。

表15. XLS 格式轉置前後檔案容量比較表

序號	檔案名稱	轉置後檔案	原始檔案
1. →	Magnetic.xls	206KB	563KB
2. →	內政部消防署磁片清單.xls	40KB	39KB
3. →	各廠商連絡相關資訊.xls	82KB	18KB
4. →	展示設備清單.xls	147KB	187KB
5. →	檔管局典藏組(正)轉置驗證表.xls	58KB	47KB
平均		106.6KB	170.8KB

2. 轉置前後檔案品質驗證：利用 PDFCreator 工具將測試檔案 XLS 格式及轉置後檔案 PDF/A 格式，分別轉置為 JPEG 格式，再將 2 個 JPEG 格式檔案進行 SSIM 及 UQI 之品質驗證，轉置前後檔案品質驗證比較如下表。SSIM 值及 UQI 值大於 60 代表可接受。其中「Magnetic.xls」、「內政部消防署磁片清單.xls」及「檔管局典藏組(正)轉置驗證表.xls」的 SSIM 及 UQI 未達 100 分原因在於品質驗證需將測試檔案 XLS 格式及轉置後檔案 PDF/A 格式皆轉置為 JPEG 圖檔，其轉置為 JPEG 圖檔會產生一些誤差，例如：圖檔顏色的深淺或位置偏移等，導致品質驗證無法達到 100 分，但以肉眼的方式檢視 XLS 格式與 PDF/A 格式內容部分並無差異。因此測試 5 個測試檔計算平均值，SSIM 平均值約 90，UQI 平均值為 94，由此可知 SSIM 值及 UQI 值皆已符合標準。

表16. XLS 格式轉置前後檔案品質比較表

序號	檔案名稱	SSIM	UQI
1. →	Magnetic.xls	90	90
2. →	內政部消防署磁片清單.xls	80	100
3. →	各廠商連絡相關資訊.xls	100	100
4. →	展示設備清單.xls	100	100
5. →	檔管局典藏組(正)轉置驗證表.xls	80	80
平均		90	94

#### 四、檔案轉置結果

XLS 檔若包含多個工作表，只會轉置某一工作表的內容。針對此問題，系統已可支援多個工作表轉置，檔案較大時需設定每個檔案的最大轉置時間，以避免轉置失敗，而導致系統無法正常的運作。

## 捌、EML 格式轉置 PDF/A 格式

### 一、轉置目的

EML(Election Markup Language)格式是微軟公司在 Outlook 中所使用的一種遵循 RFC822 及其後續擴展的文件格式，並成為各類電子郵件軟體的通用格式。

目前各機關的通訊方式大部分以電子郵件為主，若電子郵件的內容具有價值存在的話，應作保存以利後續查閱，但瀏覽 EML 檔須透過相關的瀏覽軟體才能檢視。經本專案評估後，建議 EML 檔可採用 PDF/A 作為電子檔案格式轉置之目的格式。

### 二、轉置測試環境

表17. 轉置測試環境表

項目	規格
中央處理器	Pentium (R)Dual-Core CPU P8700 @ 2.6GHz
硬碟	110GB
記憶體	3.0GB
作業系統	Microsoft Windows XP Professional Version 2002 Service Pack 3
顯示晶片	Intel G33/G31 Express ChiPSet Family



## 三、檔案轉置結果

## (一)測試檔 1

表18. 測試檔 1 資料表

電子檔案基本資料		
檔案名稱	FW_lab_diagram.eml	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	2 秒	
附件	無	
比較		
	<u>轉置後檔案</u>	原始檔案
電子檔案格式	PDF/A	EML
Size 大小	64KB	306KB
品質驗證	SSIM	100
	UQI	100

## (二)測試檔 2

表19. 測試檔 2 資料表

電子檔案基本資料		
檔案名稱	test.eml	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	1 秒	
附件	悠遊卡領用紀錄.doc	
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	EML
Size 大小	50KB	62KB
品質驗證	SSIM	100
	UQI	100

## (三)測試檔 3

表20. 測試檔 3 資料表

電子檔案基本資料	
檔案名稱	台灣高等法院花蓮分院報名表.eml
檔案個數	1
掃毒工具名稱/版本	OfficeScan/6.3.1
病毒掃描結果	一切正常沒有病毒
轉置花費時間	1 秒
附件	無
比較	
	<div>轉置後檔案</div> <div>原始檔案</div>
電子檔案格式	PDF/A
Size 大小	52KB
品質驗證	SSIM
	UQI

## (四)測試檔 4

表21. 測試檔 4 資料表

電子檔案基本資料	
檔案名稱	彰化縣政府封裝檔.eml
檔案個數	1
掃毒工具名稱/版本	OfficeScan/6.3.1
病毒掃描結果	一切正常沒有病毒
轉置花費時間	2 秒
附件	無
比較	
	<div>轉置後檔案</div> <div>原始檔案</div>
電子檔案格式	PDF/A
Size 大小	37KB
品質驗證	SSIM
	UOI

## (五)測試檔 5

表22. 測試檔 5 資料表

電子檔案基本資料		
檔案名稱	異常問題.eml	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	1 秒	
附件	無	
比較		
	轉置後檔案	原始檔案
電子檔案格式	PDF/A	EML
Size 大小	62KB	13KB
品質驗證	SSIM	100
	UOI	100

## (六)轉置前後檔案容量與品質比較

1. 轉置前後檔案容量比較：轉置前後檔案容量如下表，測試結果顯示，轉置後檔案 PDF/A 格式之平均容量比原始檔案 EML 格式小。

表23. EML 格式轉置前後檔案容量比較表

序號	檔案名稱	轉置後檔案	原始檔案
1. →	FW_lab_diagram.eml	64KB	306KB
2. →	test.eml	50KB	62KB
3. →	台灣高等法院花蓮分院報名表.eml	52KB	21KB
4. →	彰化縣政府封裝檔.eml	37KB	3214KB
5. →	異常問題.eml	62KB	13KB
平均		53KB	723.2KB

2. 轉置前後檔案品質驗證：利用 PDFCreator 工具將測試檔案 EML 格式及轉置後檔案 PDF/A 格式，分別轉置為 JPEG 格式，再將 2 個 JPEG 格式檔案進行 SSIM 及 UQI 之品質驗證，轉置前後檔案品質驗證比較如下表。SSIM 值及 UQI 值大於 60 代表可接受。測試 5 個測試檔計算平均值，SSIM 平均值約 100，UQI 平均值為 100，由此可知 SSIM 值及 UQI 值皆符合標準。

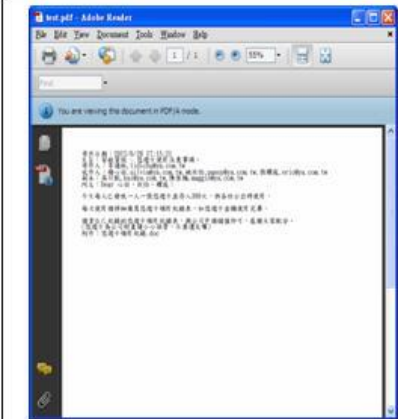

表24. EML 格式轉置前後檔案品質比較表

序號	檔案名稱	SSIM	UQI
1. →	FW_lab_diagram.eml	100	100
2. →	test.eml	100	100
3. →	台灣高等法院花蓮分院報名表.eml	100	100
4. →	彰化縣政府封裝檔.eml	100	100
5. →	異常問題.eml	100	100
平均		100	100

#### 四、檔案轉置結果

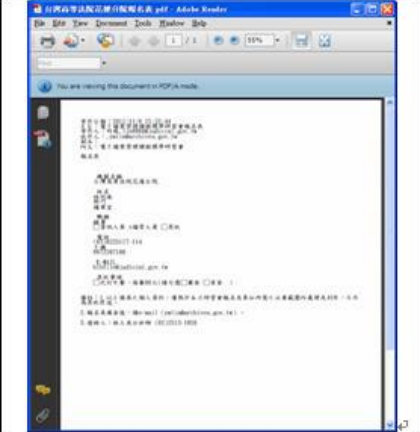

(一) 若 EML 信件內容文字有顏色，轉置結果會呈現黑色。經修正系統後，轉置後檔案的內容文字均可正常呈現。

表25. EML 格式轉置修正前後結果比較表(內容文字有顏色表)

修正前	修正後
	

(二)若 EML 信件內容有表格，轉置結果無法呈現表格。經修正系統，轉置後檔案可正常呈現。

表26. EML 格式轉置修正前後結果比較表(內容有表格)

修正前	修正後
	

(三) 由於附件格式眾多、病毒、加密或壓縮等因素，使得附件檔的轉置有困難，目前 EML 可轉置本文內容，附件檔部分則可轉置 PPT、XLS、DOC 及 DOCX 格式。

## 玖、MP3 格式轉置 WAV 格式

### 一、轉置目的

WAV (Waveform audio format) 是微軟與 IBM 開發的聲音編碼格式，在 Windows 平台的所廣泛應用。WAV 音頻格式無經過壓縮，音質不會出現失真的情況。由於 WAV 格式製作容易、可使用任何音源、且硬體需求少，所以建議聲音檔案採用 WAV 格式。

### 二、轉置測試環境

表27. 轉置測試環境表

項目	規格
中央處理器	Pentium (R)Dual-Core CPU P8700 @ 2.6GHz
硬碟	110GB
記憶體	3.0GB
作業系統	Microsoft Windows XP Professional Version 2002 Service Pack 3
顯示晶片	Intel G33/G31 Express ChiPSet Family



## 三、檔案轉置結果

## (一)測試檔 1

表28. 測試檔 1 資料表

電子檔案基本資料		
檔案名稱	Love Me.mp3	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	8 秒	
比較		
	轉置後檔案	原始檔案
電子檔案格式	WAV	MP3
Size 大小	32,999KB	4,491KB
格式驗證	正確	

## (二)測試檔 2

表29. 測試檔 2 資料表

電子檔案基本資料		
檔案名稱	Jessie J - Price Tag ft. <u>B.o.B.</u> .mp3	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	9 秒	
比較		
	<u>轉置後檔案</u>	原始檔案
電子檔案格式	WAV	MP3
Size 大小	37,958KB	3,031KB
格式驗證	正確	

## (三)測試檔 3

表30. 測試檔 3 資料表

電子檔案基本資料		
檔案名稱	心醉心碎.mp3	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	15 秒	
比較		
	轉置後檔案	原始檔案
電子檔案格式	WAV	MP3
Size 大小	49,883KB	11,319KB
格式驗證	正確	

## (四)測試檔 4

表31. 測試檔 4 資料表

電子檔案基本資料		
檔案名稱	Bii My Love.mp3	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	11 秒	
比較		
	轉置後檔案	原始檔案
電子檔案格式	WAV	MP3
Size 大小	43,376KB	6,887KB
格式驗證	正確	



## (五)測試檔 5

表32. 測試檔 5 資料表

電子檔案基本資料		
檔案名稱	陳勢安 - 天后.mp3	
檔案個數	1	
掃毒工具名稱/版本	OfficeScan/6.3.1	
病毒掃描結果	一切正常沒有病毒	
轉置花費時間	12 秒	
比較		
	轉置後檔案	原始檔案
電子檔案格式	WAV	MP3
Size 大小	49,352KB	6,716KB
格式驗證	正確	

## (六)轉置前後檔案容量與品質比較

1. 轉置前後檔案容量比較：轉置前後檔案容量如下表，測試結果顯示，轉置後檔案 WAV 格式之平均容量比原始檔案 MP3 格式大。

表33. MP3 格式轉置前後檔案容量比較表

序號	檔案名稱	轉置後檔案	原始檔案
1. →	Love Me.mp3	32,999KB	4,491KB
2. →	Jessie J - Price Tag ft. B.o.B..mp3	37,958KB	3,031KB
3. →	心醉心碎.mp3	49,883KB	11,319KB
4. →	Bii My Love.mp3	43,376KB	6,887KB
5. →	陳勢安 - 天后.mp3	49,352KB	6,716KB
平均		42713.6KB	6488.8KB

2. 轉置前後檔案品質驗證：利用 JHOVE 格式檢測工具來檢測轉置後檔案是否為 WAV 格式。由下表可知 5 個測試檔的檢測結果皆為正確。

表34. MP3 格式轉置後檔案格式驗證表

序號	檔案名稱	格式驗證
1. →	Love Me.mp3	正確
2. →	Jessie J - Price Tag ft. B.o.B..mp3	正確
3. →	心醉心碎.mp3	正確
4. →	Bii My Love.mp3	正確
5. →	陳勢安 - 天后.mp3	正確

#### 四、檔案轉置結果

目前驗證的方式是採用 JHOVE 檢測工具來檢測檔案的格式是否正確，以確保符合 WAV 的格式。但若要檢測原始檔案及轉置檔案的品質差異的話，目前並沒有適合的工具可進行檢測。

## 壹拾、結論與問題檢討

本專案 101 年度已於電子檔案保存工具擴充相關格式轉置技術，包含 PPT 格式、XLS 格式、EML 格式及 MP3 格式。以下整理於轉置過程中主要遭遇問題。

### 一、PPT 格式轉置 PDF/A 格式

若 PPT 格式檔案簡報頁中只有圖片內容，此簡報頁面轉置時，版面將無法辨識直式或橫式。此問題的解決方法為轉置時系統自動加註文字「由 PPT 轉置」，轉置後的版面則可正常的呈現。

### 二、XLS 格式轉置 PDF/A 格式

若 XLS 檔若包含多個工作表，只會轉置某一工作表的內容。針對此問題，系統已可支援多個工作表轉置，檔案較大時需設定每個檔案的最大轉置時間，以避免轉置失敗。

### 三、EML 格式轉置 PDF/A 格式

- 1.若 EML 信件內容文字有顏色，轉置結果會呈現黑色。此問題的解決方法為採用 PDFCreator 作轉置，則可正常呈現文字顏色。
- 2.若 EML 信件內容有表格，轉置結果無法呈現表格。此問題的解決方法為採用 PDFCreator 作轉置，則可正常呈現表格。

3.附件格式眾多、病毒、加密或壓縮等因素，使得附件檔的轉置有困難。針對 EML 的附件檔部分，目前可轉置 PPT、XLS、DOC 及 DOCX 格式。

#### 四、MP3 格式轉置 WAV 格式

目前驗證的方式是採用 JHOVE 檢測工具來檢測檔案的格式是否正確，以確保符合 WAV 的格式。但若要檢測原始檔案及轉置檔案的品質差異的話，目前並沒有適合的工具可進行檢測。