



電子檔案 保存技術與工具介紹

檔案管理局

101年5月



電子檔案技術服務中心

大綱

- 電子檔案風險
- 電子檔案安全議題
- 線上簽核封裝檔
- 電子封裝檔工具箱
- 電子檔案保存工具及共用元件



電子檔案風險



電子檔案風險(1/3)

— 硬體過時

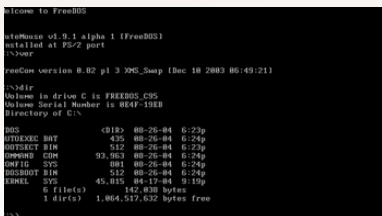


ENIAC, 1946



APPLE II, 1977

— 軟體過時



DOS, 1981



PEII, 1991

— 格式過時



WDL等格式

— 儲存媒體過時



電子檔案風險(2/3)

- 資料依附載體儲存，須透過機器才能閱讀。
- 易於修改、複製的特性，外洩或遭到篡改時不易察覺。
- 資訊科技軟硬體快速銳變，檔案格式時有變革，降低電子化檔案的可讀性。
- 電子檔案的內容或其它相關紀錄可能會遺失，以致於僅存的電子檔案沒有價值。
- 不同使用目的，靜動態格式眾多不一。



電子檔案風險(3/3)

- 如何確保電子檔案於法定保存年限仍可讀取？
 - 軟硬體過時、格式過時、儲存媒體過時。
- 如何確保電子檔案的完整性，證據性是否足夠？
 - 電子簽章、雜湊值機制，確認電子檔案是否遺失、遭破壞或竄改。
- 線上簽核電子公文與公文影像儲存的困難
 - 儲存媒體老化及劣化，無法以人工持續檢查，電子媒體損壞時，亦無法自行修復。
- 業務性電子資料保存困難
 - 部分電子資料必須配合應用軟體及系統才能查閱，大幅提高保存難度。
- 電子檔案保存面臨風險但缺乏管理機制
 - 電腦軟硬體環境變化、人為操作錯誤及資料外洩等風險。



電子檔案安全議題



電子檔案管理目標

- 真實性(authenticity)：確保網路中個體(Entity)的身分確實如他所表明的，或由網路所接收的資料確實為該傳送者(Sender)傳送。亦即可鑑別及確保電子檔案產生、蒐集及修改過程的合法性。
- 完整性(integrity)：確保網路中所傳輸之資訊與原來的資訊一致，不會遭到竄改或偽造。亦即在電子檔案管理流程中，確保儲存電子檔案的內容、詮釋資料及儲存結構之完整。
- 可及性(accessibility)：藉由電子檔案保存機制，配合法定保存年限，維持電子檔案及其管理系統的可用性。



電子檔案管理目標對應作為

- 真實性(authenticity)：運用電子憑證加簽。
- 完整性(integrity)：運用雜湊值計算。
- 可及性(accessibility)：制定一致性的線上簽核電子檔案格式，並限縮附件檔格式。

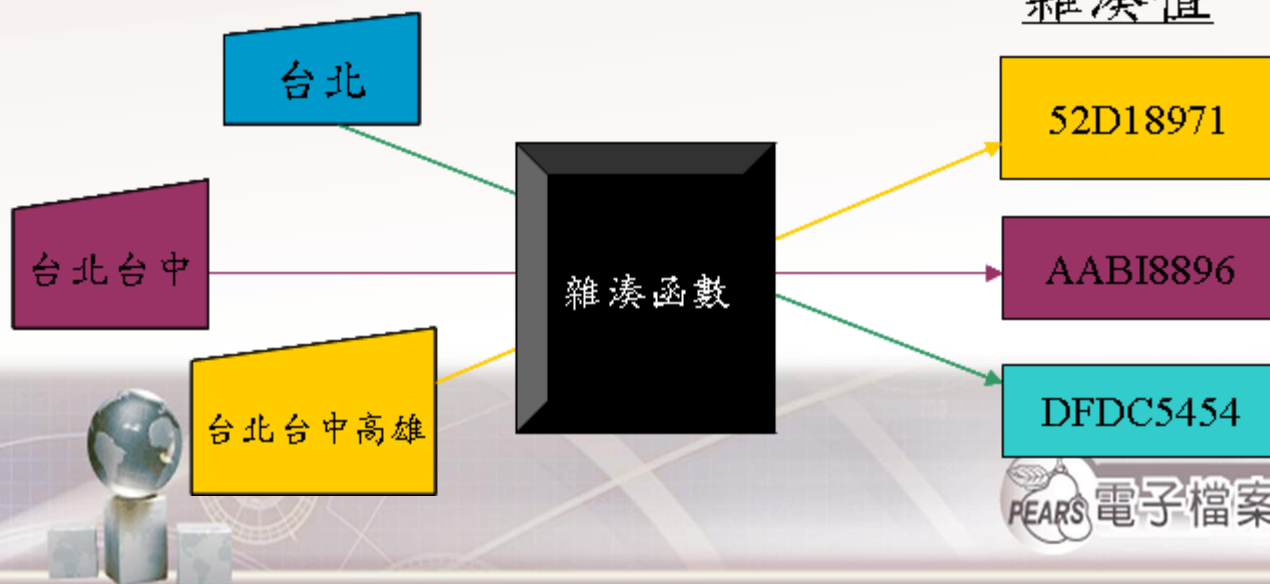


雜湊值(Hash Values)概念

- $Y = f(X)$ Y 為 f 函數的值
- 雜湊函數 (Hash Function)

不同文字輸入相同的雜湊函數必產生不同的函數值。

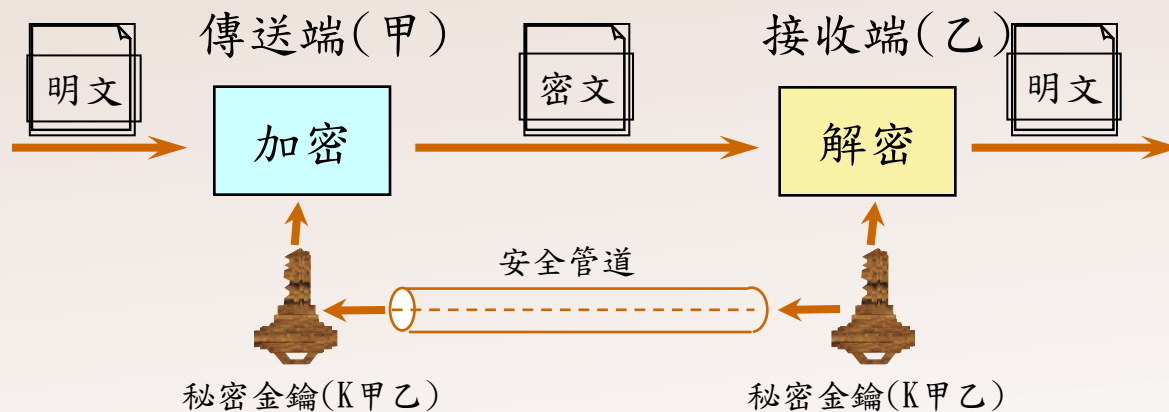
輸入



加密機制

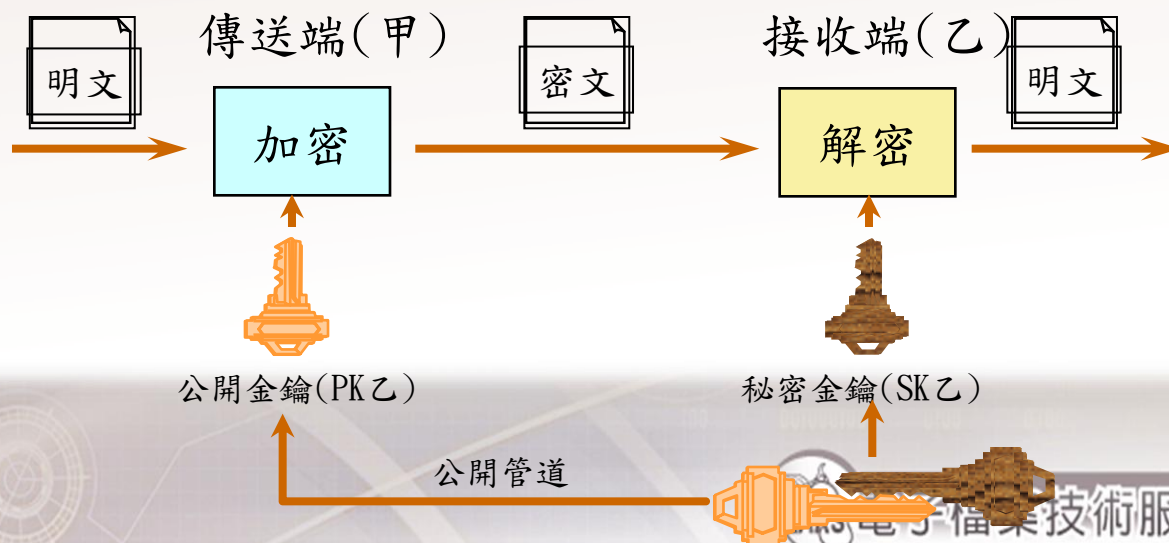
對稱性
加密法

DES
3-DES
AES



非對稱性
加密法

RSA
ECC
DL-Based

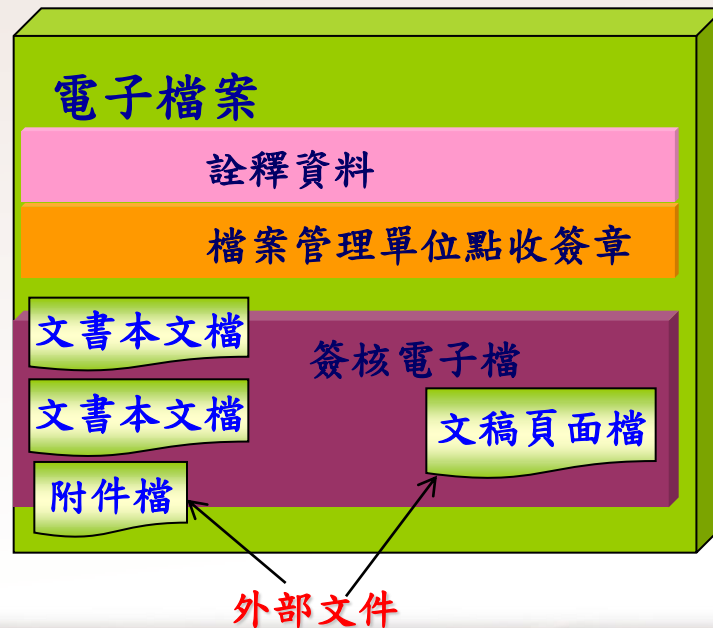


線上簽核封裝檔



線上簽核封裝檔(1/5)

- 包含文書本文檔、附件檔、簽核電子檔及文稿頁面檔。



線上簽核封裝檔(2/5)

■ 執行線上簽核檔案得包含以下內容：

□ 文書本文檔

- 內容以XML語法所組成，另稱DI(Document Instance)檔。
- 包括令、函、簽等公文本文，內容結構依文書及檔案管理電腦化作業規範，呈現格式依文書處理手冊規定。

內容結構

```
<?xml version="1.0" encoding="UTF-8"?>
<!--          99_2_utf8.dtd          函          2008.10.1  修改日期:2010.01.08
-->
<|ENTITY % 基本標籤 SYSTEM "99_basic_utf8.ent" >
%基本標籤;
<|ENTITY % 電子交換標籤 SYSTEM "99_exchange_utf8.ent" >
%電子交換標籤;
<|ELEMENT 函 (發文機關+,函類別,地址,聯絡方式+,受文者,發文日期,發文字號+,速別?,密等及解密條件或保密期限?,附件?,主旨,段落*,正本,副本?,署名*)>
<!-- /函 -->
```

DI檔

呈現格式

行政院研究發展考核委員會 函

地址：10051台北市中正區廣南路1段2之2
號6樓
承辦人：黃書儀
電話：02-23419066#822
電子信箱：syihuang@rdcc.gov.tw

發文日期：中華民國98年3月30日
發文字號：會研字第0982460308號
速別：普通件
密等及解密條件或保密期限：普通
附件：活動說明

主旨：本會謹定於98年4月17日假台南市政府6樓會議室辦理「機關電子公布欄推廣經驗分享座談會」，請派員參加，請 查照。

說明：

- 一、本會自96年8月起推廣機關電子公布欄，為持續推廣與擴大機關電子公布欄使用效益，爰辦理「機關電子公布欄推廣經驗分享座談會」。
- 二、本會邀請使用成效良好之縣市政府，分享其電子公布欄推廣經驗與應用方式(如前置作業、相關張貼規則訂定、政策宣導做法等)，活動議程及說明如附件。

正本：宜蘭縣政府、花蓮縣政府、南投縣政府、屏東縣政府、苗栗縣政府、桃園縣政府、高雄市政府、高雄縣政府、雲林縣政府、新竹市政府、新竹縣政府、嘉義市政府、嘉義縣政府、彰化縣政府、金門縣政府、福建省連江縣政府、臺中市政府、臺中縣政府、臺北市政府、臺北縣政府、臺東縣政府、臺南市政府、臺南縣政府、澎湖縣政府、基隆市政府

副本：

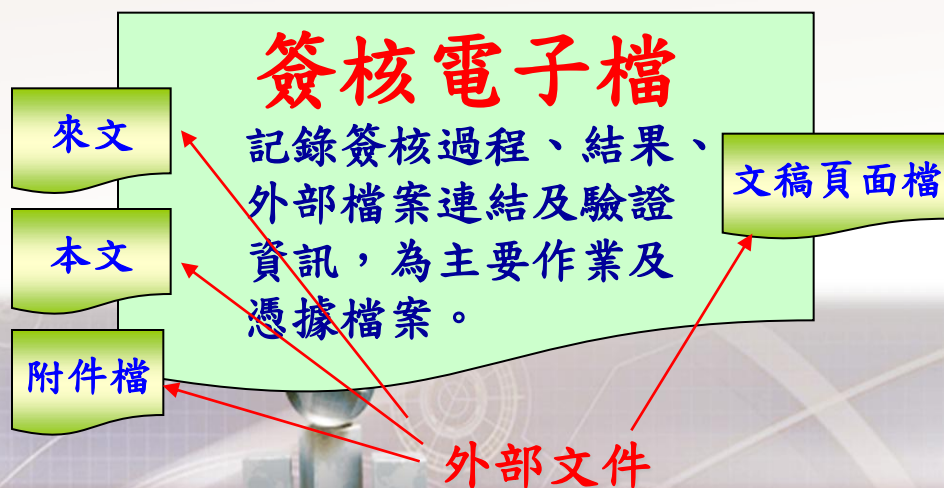
線上簽核封裝檔(3/5)

□ 附件檔

- 附件檔類型含文字檔、靜態圖形檔、工程圖檔、動畫檔、聲音檔、動態影像檔等電子或電子影音檔案。

□ 簽核電子檔

- 線上簽核文字化記錄工具，內容以XML語法所組成，用以記錄各簽核點簽核結果及過程，為主要作業及憑據檔案，另稱SI(Sign Instance)檔。



SI檔

```
<?xml version="1.0" encoding="utf-8"?>
<!-- 99_sign_utf8.ent 線上簽核(簽核電子檔) 2010.1.08 -->
<!ELEMENT 線上簽核流程 (簽核流程+)>
<!ATTLIST 線上簽核流程
  Id ID #REQUIRED
>
<!ELEMENT 簽核流程 (分會點|流程點)*>
<!ATTLIST 簽核流程
  Id ID #REQUIRED
  異動別 CDATA #REQUIRED
>
```


線上簽核封裝檔(4/5)

□ 文稿頁面檔

- 為線上簽核之視覺化記錄工具，因考量多元化之輔助工具技術產生及為免文字誤會，以影像方式記錄各簽核點人員作業完成後於清稿前之稿面，輔助重現文稿之原文原意(包含各簽核點簽核意見)。

文稿頁面檔

視覺化影像檔 (Tiff/Jpeg..)

呈現文稿頁面
輔助原貌重現

理。

2樓研討室

二、本案焦點會議業於98年8月20日日假本局3樓會議室召開，計邀請吳委員宗成等5位委員出席，並請本局檔案徵集組及檔案典藏組列席，審查本案第2期交付文件。

三、本案如奉 核後，擬將會議紀錄以電子郵件發送與會人員。

敬請

核示

訂

承辦單位
檔案資訊組

分辦師 林玉美 0911 1025

科長 李殷 0911 1133

決行

可

檔案資訊組 張文熙 0911 1141

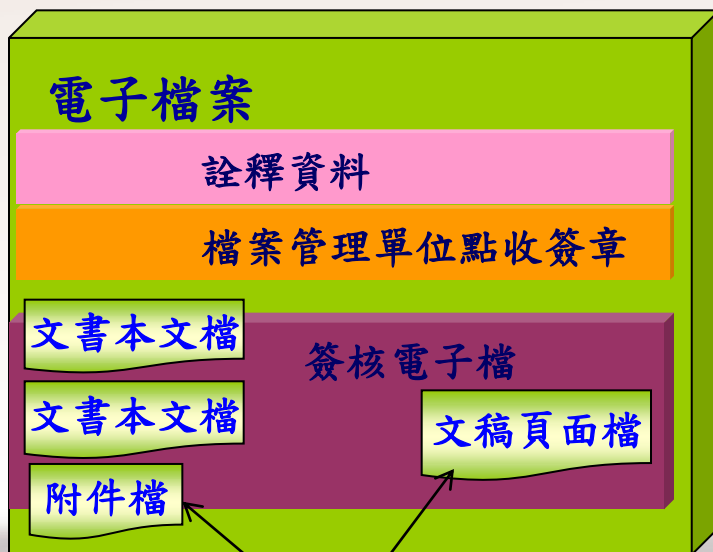
線



線上簽核封裝檔(5/5)

□ 電子檔案封裝檔

- 公文線上簽核電子檔案：內容以XML語法組成，以簽核電子檔為基礎，歸檔點收時以機關憑證加簽，年度封裝詮釋資料。
- 紙本檔案數位化電子影音檔案：以機關憑證加簽。



```
<電子封裝檔>+
<封裝檔內容 Id="Wrap">+
  <封裝檔資訊>電子檔案</封裝檔資訊>+
  <電子檔案>+
    <檔案管理單位點收簽章>+
    <Signature Id="CheckSignGCA">+
    <SignedInfo>+
      <CanonicalizationMethod Algorithm="http://www.w3.org/TR/2001/REC-
        xslt-1-4#-20010315"/>+
      <SignatureMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/>+
      <Reference URI="#FlowInfo">+
      <DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/>+
      <DigestValue>ADKDS+WuPEobYGrY7abDU85CMs=</DigestValue>+
      </Reference>+
      <Reference URI="#Sol">+
      <DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/>+
      <DigestValue>aCBtv+TVs91GsFdtEhx04ajH+4=</DigestValue>+
      </Reference>+
    </SignedInfo>+
    <SignatureValue>c2l24xp/4Ei2ufbBdC17SQD3aDDXcIdSI213Cn+nmvqFZLmWp
      1vRrAT3sGJzL0VxWzJiV7DabdVQ
      91u0jW4pFEC969rDe85LZpJBrg/4D+HGjHIQ7jZIDfci2W58QPEmM09/PAAbNysD
      ezOm2KjGhHIY YX2HGkbTVmAvuCz8UQM=+
    </SignatureValue>+
    <KeyInfo xmlns:ds="http://www.w3.org/2000/09/xmldsig#">+
      <X509Data xmlns="http://www.w3.org/2000/09/xmldsig#">+
        <X509Certificate>MIID7zCCAtegAwIBAgIQU6ux4wdmfUf7Ad+tpxD1DANBgk
          qhkiG9w0BAQUFADBKMqCwQCQYDQVQGEWUvzESMBAGA1UECgwJ6KG
          M5pS6ZmlMScwJQYDVOQLDB7mLLJupzmuKzoqabmhpHorYnurgHnkfbkuK31
```



電子封裝檔工具箱



電子封裝檔工具箱

■ 單機版工具

□ 下載網址：<http://erlp.archives.gov.tw>。

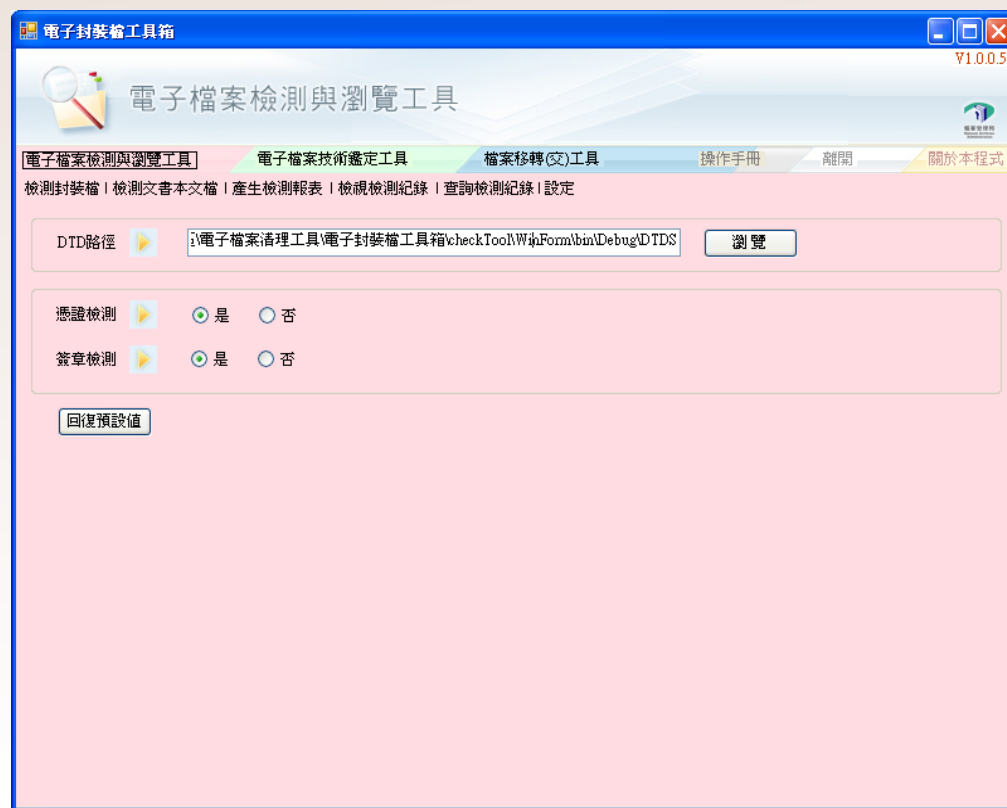
□ 工具架構

- 電子檔案檢測與瀏覽。
- 電子檔案技術鑑定。
- 檔案移轉(交)。



電子檔案檢測與瀏覽

- 檢測封裝檔
 - 格式、資料、外部檔案、憑證、簽章、其他。
- 檢測文書本文檔
- 產生檢測報表、紀錄
- 設定



電子檔案技術鑑定

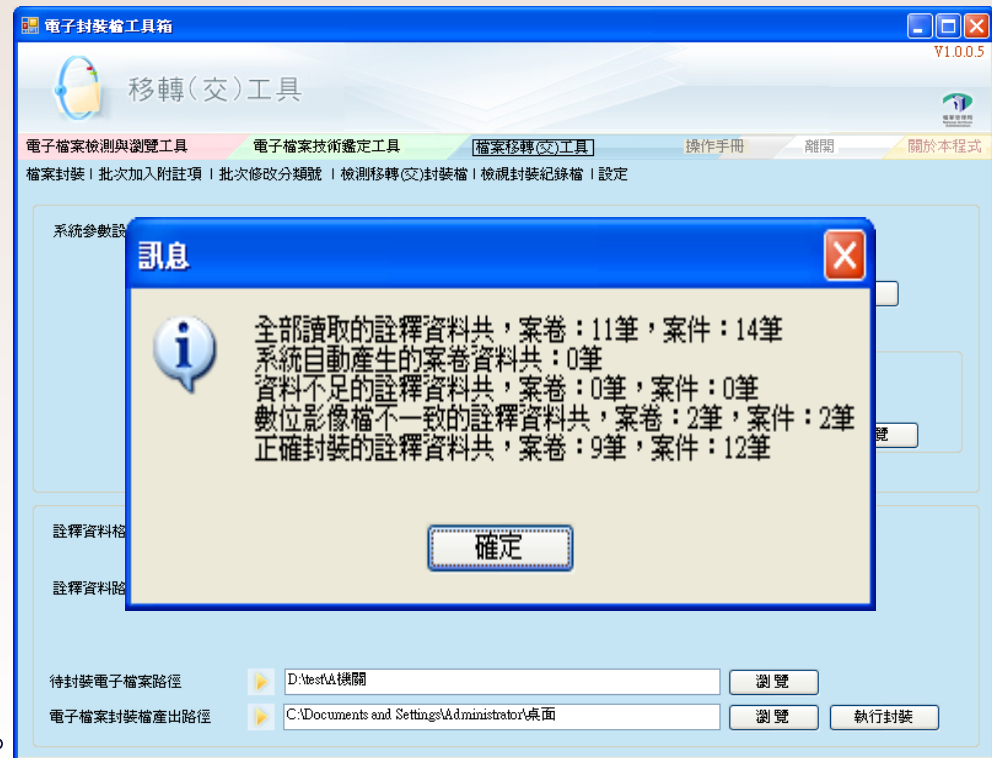
- 執行技術鑑定作業
 - 輸入待鑑定的電子檔案封裝檔路徑。
 - 輸入技術鑑定報告產出路徑。
 - 執行檢測。
 - » 自動產生清理批號。
 - 產生技術鑑定報告。

目次

五、 → 線上簽核封裝檔版本統計.....	15
六、 → 檔案清查歷程紀錄.....	16
第陸章 → 結論與建議.....	17
附件 1：電子檔案數量統計.....	18
附件 2：電子檔案有效性檢查檢測結果.....	19
附件 3：抽樣測試結果.....	20
附件 4：電子檔案格式及數量統計表.....	21
附件 5：電子檔案安全強度及數量統計表.....	22
附件 6：線上簽核封裝檔與保存年限分析.....	23
附件 7：歷程紀錄.....	24
附件 8：自訂格式說明.....	25
第伍章 → 技術鑑定結果.....	11
一、 → 電子檔案封裝檔有效性檢查.....	11
二、 → 抽樣檢測.....	12
三、 → 電子檔案格式及版本統計.....	12
四、 → 安全強度統計.....	13

檔案移轉(交)

- 檢測移轉(交)封裝檔
- 批次修改分類號
- 批次加入附註項
 - 檔案徵集註。
 - 典藏歷史註。
- 檔案封裝
 - 原生檔案為紙本檔案。
 - 未依規定格式封裝加簽。



實機操作



電子檔案保存工具及 共用元件



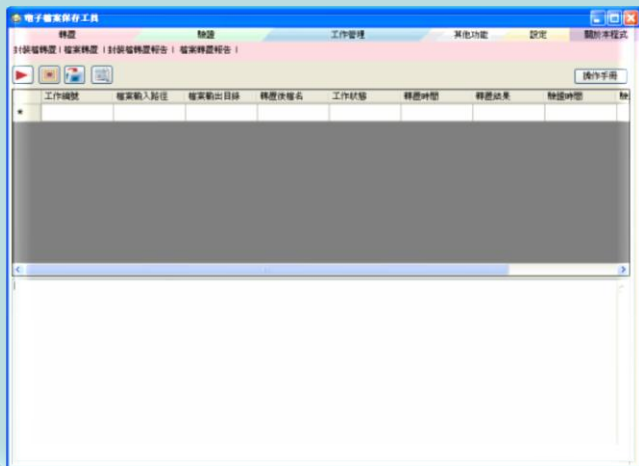
電子檔案保存工具介紹(1/3)

- 單機版之工具，區分簡易版及完整版
 - 下載網址：<http://erlp.archives.gov.tw>。
 - 簡易版功能
 - 電子檔案轉置及驗證。
 - 完整版功能
 - 電子檔案轉置及驗證。
 - 影像檔與圖片檔修補。
 - OCR 辨識及條碼辨識。



電子檔案保存工具介紹(2/3)

電子檔案保存工具



Open
Sources
軟體

- PDFCreator
- FFmpeg

學術學校
研究計劃

- 影像及視訊品質
驗證軟體。
- 影像及視訊修補
軟體。



電子檔案保存工具介紹(3/3)

□ 單筆及批次轉置

- 文件類型(DOC、DOCX、WDL、PS)。
- 圖片類型(TIFF、JPEG)。
- 視訊類型(WMV、MPEG-2)。

□ 品質驗證

- 評估方法(PSNR、SSIM...)。
- 評估值1~100分，門檻值=>大於60分。
- 驗證方式(抽驗1頁或隨機抽驗或驗證全部)。

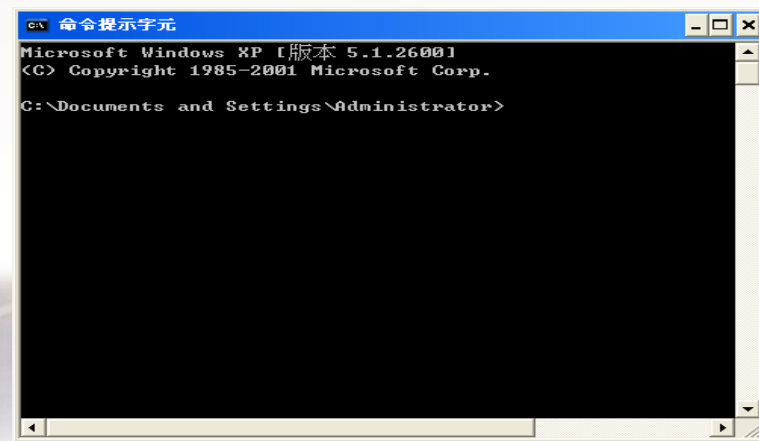
□ 免費使用

□ 提供技術諮詢



共用元件介紹(1/3)

- 功能
 - WDL格式轉置PDF格式或PDF/A格式。
 - DOC/DOCX格式轉置PDF格式或PDF/A格式。
- 使用方式
 - 使用命令提示字元，輸入相關參數，即可執行共用元件。
- 使用時機
 - 批次轉置。
 - 可整合公文檔管系統。



共用元件介紹(2/3)

- 共用元件參數說明

- wdl2pdf [options] source [destination]

參數	說明
source	設定原始檔案或目錄的路徑
destination	輸出轉置檔案或目錄的路徑
-l	Log檔路徑
-f	設定轉置格式 ● PDF/A (預設) ● PDF
-t	逾時(單位：秒，預設：15秒)
-v	轉置過程顯示於命令提示字元中
-h	查看參數說明

共用元件介紹(3/3)

- 範例

範例說明	說明
Source為「檔案」	C:\wdl2pdf.exe -l C:\1.log -t 60 C:\1.wdl C:\1.pdf
Source為「目錄」	C:\wdl2pdf.exe -l D:\轉置檔案\1.log -f pdf D:\原始檔案 D:\轉置檔案
查看參數說明	C:\wdl2pdf.exe -h



實機展示



簡 報 完 畢
恭 請 指 教



格式轉置

- 格式轉置：將某一電子檔案格式轉置為另一種檔案格式。

- 格式轉置種類

- 文件類型
- 圖片類型
- 視訊類型

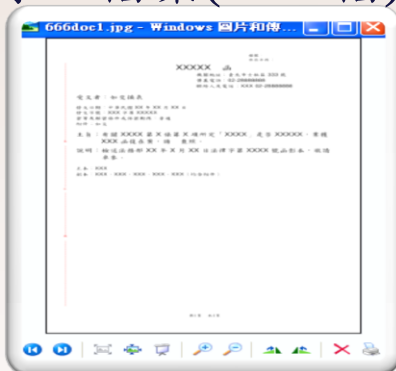
- WDL格式轉置為PDF/A格式
- DOC格式轉置為PDF/A格式
- DOCX格式轉置為PDF/A格式
- DOC格式轉置為ODT格式
- PostScript格式轉置為PDF/A格式



驗證

- 驗證目的：確保轉置後檔案品質。

原始檔案(TIFF檔)



轉置檔案(JPEG檔)



影像品質
評估方法

評估值 $\geq 60 \rightarrow$ 品質好
評估值 $< 60 \rightarrow$ 品質差

評估值
(介於1~100分數)



電子檔案技術服務中心