

<國外電子檔案最新技術或制度>

veraPDF」(PDF/A 驗證工具簡介

各國在面臨電子檔案議題時，都深切了解長期保存的重要，須適當規劃長期保存政策，以確保電子檔案的真實性、完整性及可及性。在行政院頒布之「文書及檔案管理電腦化作業規範」附錄 8 內容中列舉許多適合長期保存的電子檔案格式，其中大家最熟悉、應用最廣泛的莫過於 PDF 格式。

PDF 格式有四種主要國際標準：PDF/A、PDF/X、PDF/H、PDF/E，而使用在檔案歸檔的標準即是 PDF/A(Archive)，這種格式標準能確保電子檔案經過數十年，依然能正常讀取及開啟，而且文件呈現格式與儲存時一樣，非常適用於電子公文檔案長期保存。

veraPDF 組織是由 PREFORMA(PREservation FORMAts for culture information/e-archives)計畫及歐盟委員會的 FP7-ICT 方案共同贊助。該組織主要目的是開發開放原碼(OpenSource)的認證器，並為解決建置長期保存檔案格式所面臨的問題，滿足博物館、檔案館、圖書館，文物機構等機關(構)對於長期保存數位資料進行完善管理的需求。

一、驗證模型與工具

veraPDF 驗證模型包含：

- (一) 物件類型的層級結構。
- (二) 可繼承屬性以及可連接其他物件類型的命名鏈結。

此檢查工具透過解析字串或本文的方式來檢查 PDF 文件。透過解析可輸出兩種報告，一種描述該 PDF 文件和其詮釋資料，一種描述該 PDF 文件和 PDF 特徵的相似度。報告格式為 HTML5。

(二) veraPDF 詮釋資料修正(Metadata Fixer)

此修正工具提供對 PDF 文件詮釋資料的修正功能，例如當文件與 PDF/A 格式不一致時移除 PDF/A 格式標籤，或是當損壞的 XMP(Extensible Metadata Platform)詮釋資料是唯一導致驗證失敗的原因時，修復此受損的 XMP 詮釋資料。此修正工具會產生一個修正後的 PDF 文件和修正報告，報告內容說明修正成功或失敗的資訊。

(三) veraPDF 政策檢查(Policy Checker)

veraPDF 政策檢查工具會藉由分析 PDF 文件及比對公司政策設定檔，判斷 PDF 文件是否符合公司政策並產生報告。此工具針對 PDF 文件內的幾乎所有的品質都可進行檢查(例如註解的用法)，不限定是否為 PDF/A 格式。

(四) veraPDF 報告(Reporter)

此報告工具將執行檢查、詮釋資料修正及政策檢查等 3 個所產生的報告轉換成可供下載的形式。

(五) veraPDF 殼層(Shell)

veraPDF 殼層負責管理 veraPDF 的其他元件並確認各活動序列間的互動。使用者藉由命令提示字元介面(CLI)、桌面圖形使用者介面及網頁圖形使用者介面與 veraPDF 殼層進行互動。

二、軟體使用介紹

2016/2/28 發布之 veraPDF 軟體使用介紹如下：

(一) 選擇要驗證的 PDF 檔案及要使用的驗證規則檔(如圖 2)，可依照要驗證的政策選定不同規則檔(如 PDF/A-1A、PDF/A-1B、PDF/A-2A 等)。

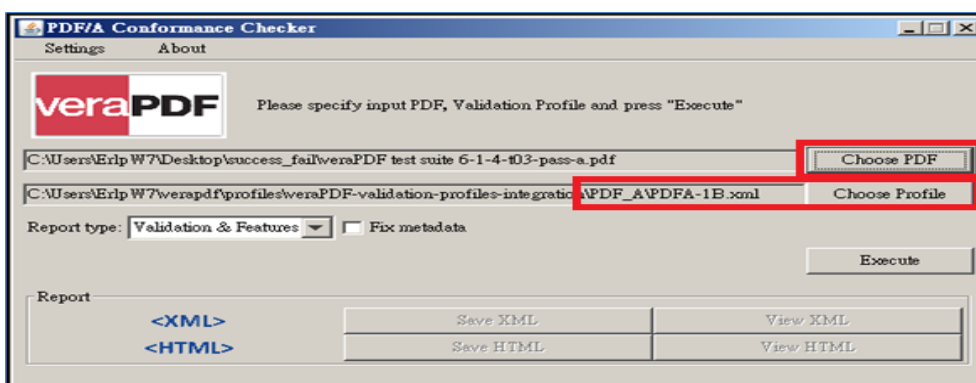


圖 2 veraPDF 軟體圖形化操作介面

(二) 也可依需求調整驗證規則檔的內容，如調整驗證的條件值。圖 3 為驗證規則內容。

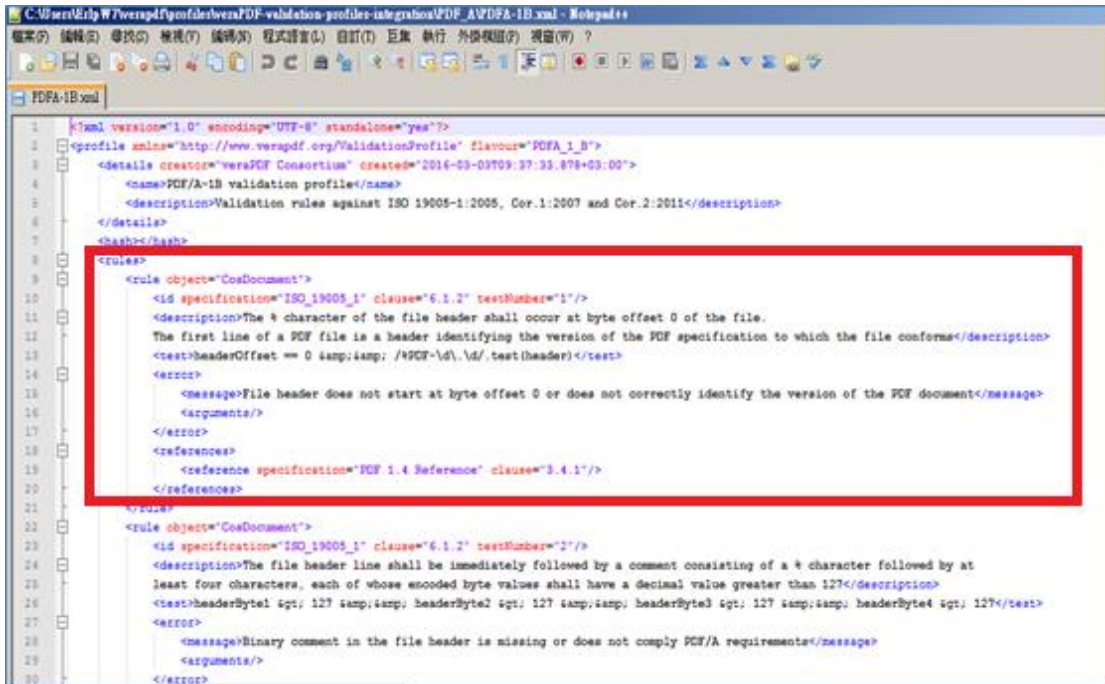


圖 3 規則 XML 檔內容

(三) 執行驗證後產生 XML 或 HTML 之報告(如圖 4 及圖 5)。

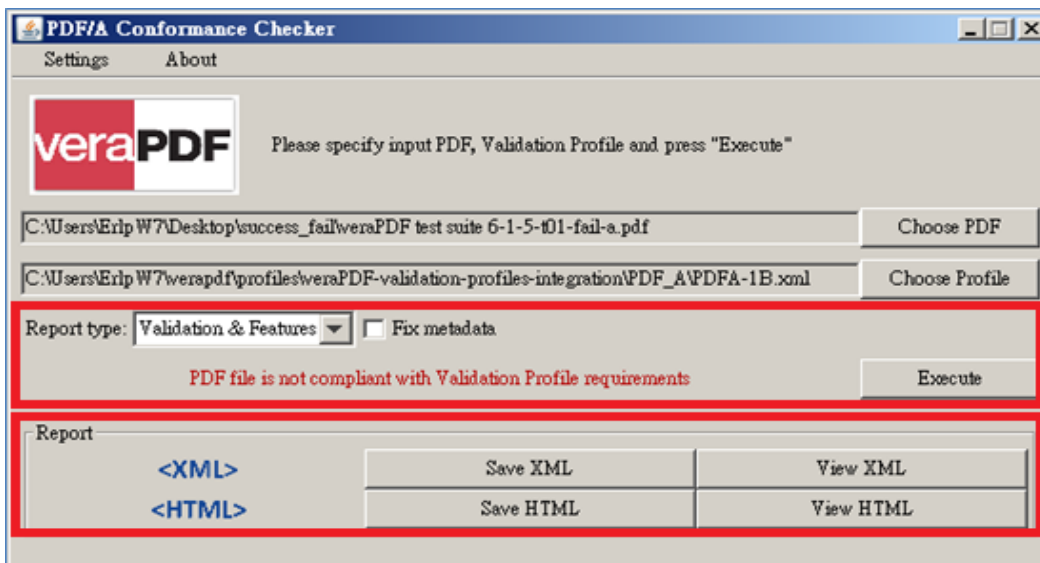


圖 4 veraPDF 驗證及報告產製



圖 5 veraPDF 驗證產製之 HTML 報告

(四) 可於驗證時勾選 Fix metadata，執行結果會先產製新的修正檔案(如圖 6)，再將該修正檔案進行驗證，即可發現經修正原檔案詮釋資料後產製的新檔案符合驗證規則(如圖 7)。

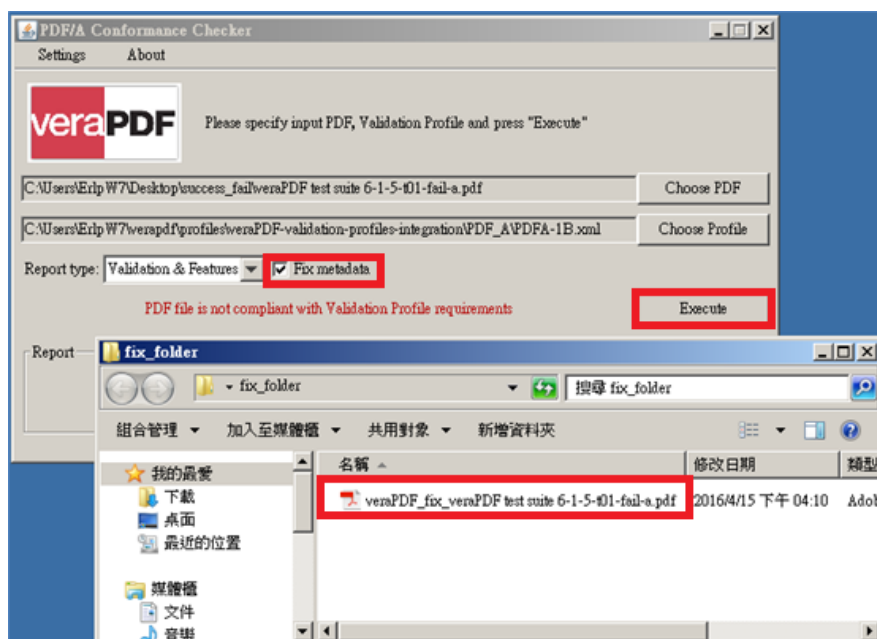


圖 6 veraPDF 修正詮釋資料會先產製新的 PDF/A 檔

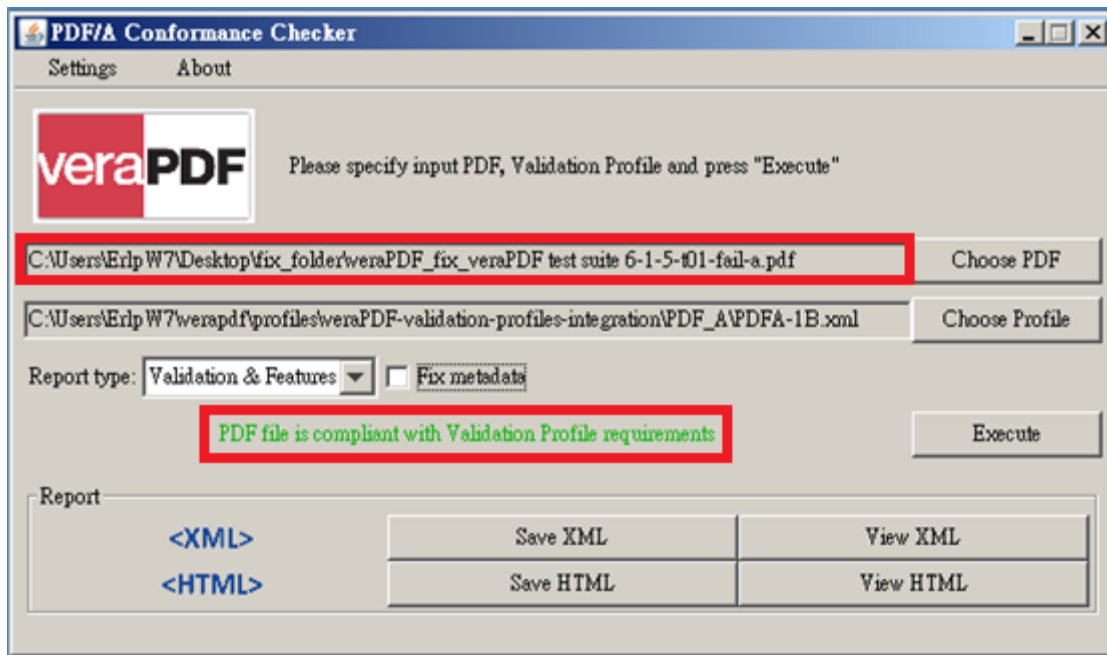


圖 7 修正後之 PDF/A 檔案符合驗證規則

veraPDF 軟體可驗證與 PDF/A 相似級別(如 PDF/A-1A、PDF/A-1B)之可攜式文件格式，其中 PDF/A-1A 表示該文件須符合 ISO 19005-1 所有標準，文件結構和語義特徵完全符合；PDF/A-1B 則只須符合 ISO 19005-1 最低標準，文件所呈現的視覺外觀可長期保存即可。

因此各機關(構)可訂定其長期保存政策，並藉由調整 veraPDF 驗證規則來驗證 PDF 文件是否符合機關(構)之需求。

三、參考資料來源

- (一) <http://verapdf.org/home/>
- (二) http://www.digitalmeetsculture.net/wp-content/uploads/2016/03/pfo_flyer_vera_exe.pdf
- (三) <http://files.dnb.de/nestor/veranstaltungen/Praktikertag2015/veraPDF-Cologne.pdf>
- (四) <http://www.digitalmeetsculture.net/tag/verapdf/>
- (五) <http://www.adobe.com/devnet/xmp.HTML>
- (六) <http://wiki.archives.gov.tw/index.php/100-151?id=2057>
- (七) <http://wiki.archives.gov.tw/index.php/component/content/article?id=1814>
- (八) <http://www.digitalpreservation.gov/formats/fdd/fdd000252.shtml>
- (九) <http://www.digitalpreservation.gov/formats/fdd/fdd000251.shtml>