

擴大我國電子檔案範圍之規劃 與實作(期末報告)

目 次

提要.....	2
第一章 緒論.....	8
第一節 計畫背景與動機.....	8
第二節 計畫目標.....	9
第三節 執行方法與步驟.....	10
第二章 文獻探討.....	12
第一節 電子郵件保存.....	12
第二節 網站保存.....	17
第三章 國內外現況.....	22
第一節 美國.....	23
第二節 英國.....	58
第三節 加拿大.....	81
第四節 澳洲.....	93
第五節 我國.....	109
第四章 作業指引與雛型系統.....	121
第一節 電子郵件保存.....	121
第二節 網站保存.....	135
第五章 擴大機關核心業務電子檔案之可行性分析.....	147
第一節 電子郵件.....	147
第二節 網站.....	161
第六章 結論與建議.....	175
第一節 結論.....	176
第二節 建議.....	178
附錄一 電子郵件保存系統安裝手冊.....	185
附錄二 電子郵件保存系統操作手冊.....	192
附錄三 網站保存系統安裝手冊.....	205
附錄四 網站保存系統操作手冊.....	229
參考資料.....	269

提要

一、 研究緣起

各級政府機關對於電子檔案的需求不斷攀升，使得電子檔案已在檔案事業發展上扮演舉足輕重的角色。然而，綜觀國外先進國家所採用的電子檔案範圍，大多皆涵蓋電子郵件與網站，甚至也有部分國家包含資料庫內容。因此，若與各先進國家相較，我國電子檔案的範圍仍有所侷限。為了因應時代所趨，並與國際間之作法接軌，本分項計畫將針對擴大我國電子檔案範圍的相關議題，進行解決方案之規劃與實作，以作為檔案管理局修訂相關法規之參考。

因此，為研擬適合我國之電子檔案範疇，並提出相關管理辦法，本分項計畫擬定之目標為：研析擴大機關核心業務電子檔案之可行性，如機關業務之電子郵件、網站、資料庫及機關重要檔案，訂定分期納管策略、範圍及推動期程等；以檔案管理局歷年網站及公務電子郵件保存系統實作，提出保存系統應蒐集範圍、收集週期、規劃時程及應用方式，並依系統實作結果，研提具體建議事項及相關配套措施，並提出保存系統所需硬體設備與系統軟體需求，俾利檔案管理局另案採購。

二、 執行方法與過程

本分項計畫依擬定之目標，藉由文獻探討、實地參訪、專家焦點座談等方法，規劃符合國際趨勢與我國需要之「電子檔案範圍」與相關配套措施。

（一）文獻資料彙整與探討

本分項計畫將針對相關議題廣泛收集資料，並進行彙整、研讀、與分析。資料來源主要包括：國內、外期刊、研究報告、考察報告、書籍、手冊、法規、網站、...等。

藉由文獻探討，預期彙整完成以下資料：本分項計畫同時彙整美國、英國、加拿大、澳洲與我國之相關資料，以瞭解並分析各國作法，作為研提「電子檔案範圍」與相關配套措施之參考。

（二）國內外機關參訪

本分項計畫依計畫目標，選擇挑選國內具有網站保存與處理電子郵件豐富經驗之機關與廠商，實地參訪，以彌補文獻資料的不足。

（三）管理規範與技術配套之研擬

本分項計畫擬配合我國政府機關的需要，針對將電子郵件與網站納入電子檔案範圍之相關議題，提出所需之管理規範與配套措施。

（四）系統實作

本分項計畫以檔案管理局之電子郵件系統與網站為對象，依本分項計畫所規劃之作法，進行系統實作，以了解其可行性。

（五）焦點座談

本分項計畫邀請相關專家、學者、資訊廠商代表、政府機關代表等，舉辦座談會，針對本分項計畫所規劃之內容，進行廣泛的意見交換，以作為修正的參考。

三、 主要建議事項

本分項計畫根據國內外文獻探討、機關參訪、焦點座談會與雛型系統建置經驗，提出以下建議：

（一）制定電子郵件保存詮釋資料

現有電子檔案詮釋資料中，增加屬於電子郵件相關之詮釋資料。

（二）避免採用公務信箱處理私人事務

處理私人事務時，請採用私人信箱，若處理公務時，請採用公務信箱，換言之，機關承辦人員應避免採用公務信箱處理私人事務。

（三）處理機密或敏感性電子郵件

各機關得選定可確保資訊安全的方式，如採用加密方式傳遞電子郵件。

（四）設定電子郵件標題名稱

承辦人員在命名電子郵件的標題時應審慎考量。或者，建議把電子郵件保存至檔案管理系統時，給予一個更適切的標題作為案由或案名，並且可保留電子郵件原本的標題。

（五）保存電子郵件的時機

承辦人員得採定期保存電子郵件的方式，或當電子郵件傳送或接收時，便同時進行保存。

（六）匯送電子郵件至檔案管理系統

建議將電子郵件也保存至檔案管理系統中。若機關已建置檔案管理系統，則建議在電子郵件系統中增加匯出電子郵件成為一筆電子檔案之功能，並透過檔案管理系統中電子檔案點收之功能，將電子郵件匯入檔案管理系統，使電子郵件進入檔案管理的生命週期，以利後續管理與維護；若機關尚未建置檔案管理系統時，則建議把電子郵件放到電子郵件系統之共用資料夾(shared public email folder)，或將電子郵件保存在共用的特定系統或硬碟中。

（七）選擇電子郵件保存格式

採用目前電子郵件格式中最主要的標準「.eml」，它也是現大多數電子郵件系統支援的格式。此外，因應科技變化，應定期

審視此保存格式是否適合。

(八) 決定網站保存優先順序

決定該保存哪些網站時，可採用 MoSCoW 分類方式以標明保存重要性，將網站分為以下四種，且因屬於 M 及 S 分類的清單可能壟長且混亂，因此 MoSCoW 方法可在這兩個分類再實行一次，以確認保存優先順序。

(九) 網站保存執行方式

機關若需進行網站保存時，可採用自製、外包、協作、聯盟等四種方式，並同時請先考量機關之成本、風險、優先權、可得資源及基礎設備等因素後。

(十) 網站保存工具

若機關決定自行進行網站的保存計畫，目前已有一些網站典藏的相關工具可供機關參考。分別為：

1、蒐集(Acquisition)工具

例如：Heritrix、HTTack。

2、存取(Access)工具

例如：Wayback、WERA、NutchWAX。

3、管理(Curator)工具

例如：Web CuratorTool、PANDORA Digital Archiving System、NetarchiveSuite。

本計畫所建置的雛型系統已整合了 Web Curarator Tool (含 Hertrix) 與 Wayback 等工具，運作順利。

(十一) 網站抓耙方式

機關可依需要選擇在伺服器端備份網站資料、瀏覽器端抓取網站畫面，或是使用抓耙軟體保存網站。

(十二) 選擇網站的保存格式

從「授權系統或伺服器端」著手保存網站是以 Mirror 方式完全備份原始網站檔案，此種方式保存下來的檔案格式會與檔案原來的格式完全相同。而從「瀏覽器端」的保存方式會以圖檔的方式將畫面完整的保存，檔案格式會以常見的 jpg、gif 及 png 等圖檔為主。若使用抓耙軟體保存網站，則建議能採用 WARC 國際標準，以符合國際間的發展趨勢。

第一章 緒論

第一節 計畫背景與動機

近年來，隨著資訊科技的蓬勃發展，我國積極發展檔案資訊化作業，各級政府機關對於電子檔案的需求也不斷攀升，使得電子檔案已在檔案事業發展上扮演舉足輕重的角色。

何謂「檔案」？何謂「電子檔案」？依「檔案法」與「檔案法施行細則」的規範，檔案是指各機關依照管理程序而歸檔管理之文字或非文字資料及其附件，包括各機關所持有或保管之文書、圖片、紀錄、照片、錄影（音）、微縮片、電腦處理資料等，可供聽、讀、閱覽或藉助科技得以閱覽或理解之文書或物品。依「文書及檔案管理電腦化作業規範」之定義，電子檔案是指指電腦可處理之文字或非文字資料，且符合檔案法及相關法令規定等。

然而，綜觀國外先進國家所採用的電子檔案範圍，大多皆涵蓋電子郵件與網站，甚至也有部分國家包含資料庫內容。因此，若與各先進國家相較，我國電子檔案的範圍仍有所侷限。為了因應時代所趨，並與國際間之作法接軌，本分項計畫將針對擴大我國電子檔案範圍的相關議題，進行解決方案之規劃與實作，以作為檔案管理局修訂相關法規之參考。

第二節 計畫目標

本分項計畫以擴大電子檔案範圍為主軸，針對管理與技術等不同面向之相關議題進行探討。預期達成之主要目標如下：

一、研析擴大機關核心業務電子檔案之可行性，如機關業務之電子郵件、網站、資料庫及機關重要檔案，訂定分期納管策略、範圍及推動期程等。

二、以檔案管理局歷年網站及公務電子郵件保存系統實作：

（一）提出保存系統應蒐集範圍、收集週期、規劃時程及應用方式，並依系統實作結果，研提具體建議事項及相關配套措施。

（二）提出保存系統所需硬體設備與系統軟體需求，俾利檔案管理局另案採購。

第三節 執行方法與步驟

本分項計畫依擬定之目標，藉由文獻彙整、實地參訪、系統實作、專家焦點座談等方法，規劃符合國際趨勢與我國需要之「電子檔案範圍」與相關配套措施。

一、文獻資料彙整與探討

本分項計畫針對相關議題廣泛收集資料，並進行彙整、研讀、與分析。

資料來源主要包括：國內、外期刊、研究報告、考察報告、書籍、手冊、法規、網站、...等。藉由文獻探討，預期彙整完成以下資料：

- (一) 電子檔案之定義；
- (二) 電子郵件歸檔與保存之作法；
- (三) 網站歸檔與保存之作法；
- (四) 資料庫內容歸檔與保存之作法。

二、國內外機關參訪

本分項計畫依計畫目標，選擇至少五家的國內機關或保存機構，及一到兩個其他國家，前往實地參訪，以弭補文獻資料的不足。實地參訪前，本分項計畫將先做好充分準備，除了熟悉國內機關或受訪國家相關機構與本分項計畫有關之背景資料外，也須事先準備訪視時參訪及詢問之重點與細節，並讓受訪單位了解參訪目的，以避免流於走馬看

花。參訪時則須就受訪單位有關電子檔案涵蓋範圍與配套措施，詳細詢問並完整收集相關補充資料。

三、管理規範與技術配套之研擬

本分項計畫配合我國政府機關的需要，針對將電子郵件、網站、資料庫等內容納入電子檔案範圍之相關議題，提出所需之管理規範與配套措施。

四、系統實作

本分項計畫以檔案管理局之電子郵件系統與網站為對象，依本分項計畫所規劃之作法，進行系統實作，以了解其可行性

五、焦點座談

本分項計畫邀請相關專家、學者、資訊廠商代表、政府機關代表等，舉辦座談會，針對本分項計畫所規劃之內容，進行廣泛的意見交換，以作為修正的參考。

第二章 文獻探討

第一節 電子郵件保存

電子郵件是一種傳送數位化訊息的方式，電子郵件保存的重要性已是許多公司企業與學者關注的議題。Anon 指出對於企業而言，保存電子郵件是相當重要的。因為，資訊是企業中相當重要的資產，電子郵件也是其中相當重要的一種資訊。Anon 發現 70% 的企業認為電子郵件可以增進資訊處理效率，這也是企業投資電子郵件與文件管理系統的主因。但是，該如何確保電子郵件、如何看待電子郵件保存系統所得之報酬與效益、如何有效管理電子檔案，使得這些資訊不僅可為作為企業效率提升的助力，並符合法律規範等議題，都是許多企業關注的焦點(Anon, 2008)。

Craine 針對 926 家公司的調查研究，發現 79% 的公司都知道法律規定要保存公司文件，但是只有 57% 的公司有正式制定相關政策來保存電子郵件。更令人驚訝的是，100% 的公司都把決定哪些電子郵件該刪除或保存的決定權交給終端使用者，這表示大部份的公司對於如何進行電子郵件保存，仍然感到相當困惑(Craine, 2007)。

有鑑於此，Wagner 提出電子郵件保存系統所應具備的需求。Wagner 指出大多數的公司，在發展有效率且有用的電子郵件保存與搜尋系統(Email Archiving and Discovery, EDA)時，必須要認清系統的需求，例如：法律規範的要求、電子化搜尋的需求、文件管理的需求等。因此，Wagner 提出了一個創

新的方法，讓所有公司所使用的企業內容管理系統(Enterprise Content Management, ECM)也可以成為一個專門處理電子郵件保存與搜尋(Email Archiving and Discovery, EDA)的系統。目前，電子郵件保存與搜尋(EAD)系統必須面對大量資訊吸收率、支援分散式全文檢索、允許法院的搜查的問題。同時，必須納入檔案管理與系統管理的考量，並且以最低成本的方式進行。Wagner 提出運用一個具象化的企業內容管理系統儲存界面，使得企業以最少的作業成本，在受管理的環境中，同時維護關聯式資料庫系統和全文檢索系統。(Wagner, 2008)

一般而言，微軟 Exchange Server 是許多公司用來節省電子郵件管理成本的方法。早期，使用電子郵件保存的公司發現，如果忽略某些步驟或要素，會導致管理電子郵件保存系統的困難，甚至增加相關成本。因此，Spurzem 指出倘若公司要建立電子郵件保存系統時，應考量的關鍵要素包含：伺服器、基礎建設、儲存、專業服務、支援服務五個構面(Spurzem, 2008)。

為協助公司順利建立適當的電子郵件保存系統，Taylor 認為已有許多公司企業深切瞭解保存相關的重要資訊的重要性，而且往往是公司中的資訊部門負責管理這些存在於公司中的重要資訊。其中，所謂重要資訊是指與核心業務、公司相關業務、管理相關之資訊，或是為符合某些規範所保存之資訊(Taylor, 2007)。

就電子郵件保存系統而言，Srinivasan 介紹電子郵件保存的技術，以及電子

郵件與其他溝通管道的演變。他認為現今，電子郵件已經被廣泛地用來作為溝通的管道，甚至被視為法律上的證據。每年許多國家的政府單位都會制定新的法律來規範電子郵件的使用與處理，例如美國政府就是每年制定相關標準的國家。此外，企業所採用的標準也將保存電子郵件這項功能視為必要的需求。因此，若公司企業無法符合政府所制定之規範，將必須支付罰金；若公司企業有任何電子郵件資料遺失，將容易造成法律訴訟案件的敗訴。為遵循電子郵件保存的規範，電子郵件保存技術必須要能夠配合這些需求。Srinivasan 指出在提供電子郵件保存系統時，所面臨的重要議題有：保存(retention)、管理(Supervision)、搜尋(Search and Discovery)、儲存管理(Search and Discovery)。電子郵件保存系統所面臨的挑戰則包含：郵件分類(Mail Classification)、管理技術(Surveillance Techniques)、非文字資料的解讀(Understanding Non-textual Data)、搜尋(Search and Discovery)(Srinivasan, 2008)。

在挑選電子郵件保存系統時，除了考量有效的管理方式，也必須滿足法律相關的規範，以及其他必須遵循的政策。如果公司企業不符合這些保存規範，可能必須支付罰款，或者受到股東的質疑、甚至降低顧客對公司的信心。電子郵件保存期間的長短則是和公司的行業與所在地點息息相關。例如：有些法律規定電子郵件必須要保存 7 到 15 年，有的甚至到 26 年。例如，製造業依據 Office of Fair Trading ruling 這個規定，必須要保存他的門電子郵件七年(Craine, 2007)。

因此，Craine 提出了 17 項電子郵件保存系統的需求，讓使用者依照這些需求對電子郵件軟體廠商與其產品進行評估。其中，17 種需求進行評估包含：相容性(compatibility)、儲存媒體(storage)、離線儲存媒體(off-line storage)、系統還原能力(restore)、保存狀況(archive conditions)、檢索(retrieval)、可擴充性(scalability)、支援性(supportability)、使用者安裝(client installation)、規畫能力(capacity planning)、搜尋(searching)、存取(Access pointers)、PST 檔案(PST files)、報告(reporting)、存取控制(access control)、安全性(security)、效能(performance) (Craine, 2007)

此外，仍有其他學者探討使用電子郵件保存系統時，相關之議題，包含：如何釐清電子郵件之間的關係(Perer, 2006)、電子郵件保存的安全機制(Li, 2006)。其中，Perer 探討當電子郵件保存後，該如何釐清電子郵件之間的關係，Perer 認為目前已經有許多使用者透過電子郵件進行溝通、訊息傳遞。大多數的使用者都只知道該如何管理自己的電子郵件。但是，就歷史和社會的角度而言，電子郵件是一個很重要的人工產物，藉由此電子郵件，可以瞭解個人與團體。為了要瞭解這些人類溝通的訊息，除了瞭解電子郵件本身之外，也需要瞭解前後文的內容。因此，Perer 提出一個取得這些必要性前後文資訊的新方法。這個方法是從電子郵件的標頭(header)來建立有意義的關聯，並且辨別電子郵件間彼此的關係。藉由具像化的技術，瀏覽電子郵件保存的人可以發掘到存原本隱藏在這些電子郵件中的重要資訊。同時，Perer 將這個方法應用到個人 15 年內的電子郵件，其中包含四萬五千筆訊息，超過四千條關聯性 (Perer, 2006)。

就電子郵件保存的安全機制而言，Li 認為電子郵件提供了一個簡單、易存取的方式來儲存個人或公司企業經年累月的資訊。一般而言，讓使用者自行設定密碼是最常用來保護電子郵件內容的作法，但是，這樣是不夠安全的一種作法，很可能會導致資料外洩。因此，為了保護電子郵件內容，Li 提出一個新的方法，透過簡單的機率模型計算出使用者寄送電子郵件的行為模式，這個行為模式是找出電子郵件寄件者和使用者電子郵件管理的關係。接著，搜集三個月內的電子郵件使用資料，進行實驗測試。結果發現採用一個不規則的偵測方式，是一種可行的電子郵件安全機制，而且這種方式不需要改變使用者存取行為(Li, 2006)

就電子郵件保存軟體而言，「Mail Safe」是一套電子郵件保存軟體，這套軟體會自動的蒐集所有接收或寄送的電子郵件，並且將這些資料保存在網路交換機之外。電子郵件保存有兩個元素，一是 SQL 伺服器，用來儲存所有分散在各地的電子郵件，另一個是分散的 Mailstore，是負責把電子郵件的附件壓縮、加密並儲存(Anon, 2007)。

Perkio 設計了一個查詢和分析電子郵件的系統，這個系統以四個面向為主，分別為：以內容為基礎的搜尋引擎、統計的標題模型、自動參照社交網路、時間序列的分析。從資訊檢索的觀點來看，這個系統提供瀏覽、近似全文檢索、和資料具象化的功能(Perkio, 2005)。

英國第四大的出版商「強斯頓新聞公司(Johnston Press)」，是擁有超過 175

個地方網站。考量法律規範和內部研究需求，該公司已經開始考慮如何保存公司中的各種文件，包含電子郵件及其附件。最後，強斯頓新聞公司選擇「GFI MailArchiver」這套軟體來管理微軟 Exchange Server 上的許多電子郵件。「GFI MailArchiver」是一套容易使用的電子郵件保存軟體，可以讓內部和外部的系統都被保存到單一的 SQL 資料庫，也可以讓系統管理者提供網頁搜尋界面，讓使用者輕鬆地透過單一管道就可以存取過去的電子郵件，同時，也讓所有使用者在管理電子郵件時，輕易的滿足相關法規的要求，例如，沙賓法案(Anon, 2005)

第二節 網站保存

網路已成為目前重要的資訊管道，網站則提供一般民眾與學術人士豐富的內容。但是，網站的迅速興起衰落使今日能看到的網站內容很可能隨時因伺服器關閉、網站改版等因素而無法再次被利用(余顯強，2008)。因此，重要網站的長久保存為時勢所趨，需制定相關策略確保珍貴文化的傳承。

首先，負責網站保存的機構是一個關鍵因素。雖然 Google 早開始採取全域搜尋保存政策，但是其目的不在於永久保存。因此，還是交由國家或長久性機構擔負責網站保存才是較有保障的方法(洪淑芬，2006)。網站也可被歸為電子資源的一部份，而負責長期保存電子館藏資源的機關，應在館藏建立與保存議

題上積極投入研究，並且致力於維護與促進適當的標準，甚至扮演主導性的開發工作，包括：研究技術、參與制訂標準、建置維護系統、建立書目及詮釋資料，引導電子館藏的發展方向(林巧敏，2007)。

然在發展過程中難免會面臨一些挑戰，例如：台灣國家圖書館目前在電子資料送存、網站保存、館藏數位化及數位保存等議題上，面臨缺乏統合協調的問題，國內學者建議館內應有統籌規劃的業務單位，並建立一合作的資源共享平台，利用群策群力，收錄自願性開放使用的電子資源，達到資訊交流的目的(林巧敏 陳雪華，2008)。Michael Day 亦認為協同合作(Collaboration)為蒐集與保存網站的關鍵成功因素(Day, 2003)。以現行運作情況而言，許多網站保存僅針對特定主題及領域保存，很難有單一機構能蒐集全域的網站。因此，不同機關推行網站保存時，彼此緊密合作或溝通是相當重要的。如果一開始可配合機構特長發展不同主題保存中心。最後，網站資源保存可以建立資源分享平台，不僅提供合作機構上傳資源的工具也兼收各界自願性開放的資源(林巧敏，2009)。

國內許多學者已萃取出美國、澳洲、加拿大等國網站資源保存計畫之精華，進而提出建議計畫約略步驟。首先，第一步應先釐清收錄原則並研究網站保存相關技術；其次，開始複製初步擇選的有限網站，並檢視複製過程與內容的正確性；最後，再將收錄的網站加以編目，上載至館藏目錄系統，同時建置一個網站匯集系統作為提供民眾檢索或瀏覽的窗口，並注意資訊內容的隱私權

問題(林巧敏 陳雪華，2008)。

網站保存應以規劃保存策略為主，並建立有助於永續發展的合作基礎。

Lyman 認為，保存策略與兩大領域息息相關——保存圖書專業及電腦科學，需參照兩方建議才有機會成型合適的保存策略(Lyman, 2002)。接著，應制定網站保存的範圍及方式。一般而言，對國家而言，具有重要性的資料，基於完整性與經濟效率，宜考慮採用混合式保存之模式。除了批次選擇收集特定類別網站(例如 edu、gov、org 等網域之網站)，由於重要人物的個人網站網域可能是 com，因此，也可針對特定的重要網站進行保存，作深入的蒐集，並選擇有關教育上的系列主題與重要時事之網站，加以備份保存及整理為主題群網站(洪淑芬，2006)。

為達到網站資源長久保存的目的，納入網站保存範圍應包含網站的相關描述，也就是詮釋資料及網站本身。保存系統必須有一完整規劃的詮釋資料格式用以描述與管理，並詳細紀錄網站的內涵與產生方式。而為了落實資源長久保存政策，從根本著手是最簡易的方式。首先應推廣將網站的詮釋資料納入保存範圍，以便網站被保存軟體抓下後的自動編目。機關也可利用此資訊進行檢核工作，掃描並檢核所設計的網站，確認網站設計是否符合標準理想(洪淑芬，2006)。而所有產出數位資源的機構，對於該機構之網站，不論在圖形利用、連結的有效性、「保存目的之詮釋資料」之完整性等方面，都應該仔細審視並修正，使得機關網站可以符合資源可及性原則。如此多管齊下的配合與努力，

才可能使國內未來的網路資源備份保存計畫順利執行(洪淑芬，2006)。

網站保存也有幾項困難急待解決。首先是網站數量與內容急遽成長；其次是網站的動態性質，許多網站、網頁、及領域迅速的更動、消失；再加上網站技術的快速演進，許多網站內容由動態資料庫產生，使網站的長期保存愈發困難(Day, 2003)。目前未能克服的技術困難約有下述二端：一、網站 Flash 以及 Javascript 撰寫的網站擷取問題，此為目前國外內網站擷取普遍存在的技術困難；二、網站的版本控制問題。經由網站版本的比對，可望節省相當大的儲存空間，然如何達到目標則尚無定論(賴重仁，2007)。

為因應網站保存所遇到的技術阻礙，已有許多專家學者嘗試提出網站保存方式，用以協助保存機構克服此過渡期。其中，Eysenbach 等學著提出了一個 Webcite 的構想，是一種引用網路參考資料的新標準。一旦「WebCite」網站資料，該篇被引用的網站將被保存至「www.webcitation.org」，可有效避免動態網站消失不見的問題。而 WebCite 在克服著作權的作法，只要網站著作權擁有者選擇退出，不接受被保存，則可使用「Robot」、「拒絕貯藏」、或「拒絕保存」等標籤，WebCite 都會尊重，並排除收納資料庫中(Eysenbach, G. and Trudel, M., 2005)。另外 Fitch 等學者也建議一種可紀錄網站版本差異的方異，為 pageVault，當網頁內容一產生或傳送至網站，便會追蹤並保存更動的內容，此方法是一個既有效率又有效果的方式(Fitch, Kent, 2003)。國內學者余顯強則指出了不具結構化特性的網頁語法，例如：HTML，未來的改變可能導致無法使

用現有保存之網頁資源。雖然透過現有軟硬體系統的保存、模擬、更新等方式，可解決資訊改版所造成無法使用的問題，但成本與時效卻是另外衍生的問題。因此，建議藉由主題地圖的方式，將非結構化特性的個別網頁，另外透過解構的技術，將內容的各種屬性分離出來，予以結構化重組，藉此達成長久保存之效(余顯強，民國 97 年)。

除了技術障礙之外，多數國家的法令環境也使網站保存面臨法律問題，最顯著的阻礙即為著作權。但是，法令爭議並非無法克服，只要謹慎挑選來源，結合有效的權利管理政策(rights management policy)，確保具爭議及潛在非法資料僅可被機密性的取用，且在抵觸法律及權利相關人時，移除相關網頁內容(Charlesworth, A., 2003)，即可有效在保存與著作保護中取得平衡。

台灣的網站保存的起步已晚美國許多，而美國 Internet Archive 對臺灣網路資料之備份是始於 1996 年。若能與 Internet Archive 建立合作關係，得到其過去備份保存之臺灣網路資源，則可更完整地建立國家的網路資源保存(洪淑芬，2006)。

第三章 國內外現況

各先進國家對於電子檔案範圍，大多皆涵蓋電子郵件與網站。因此，本章藉由彙整美國、英國、加拿大、澳洲以及我國對於保存電子郵件與網站之現況，瞭解並分析各國差異，作為研提我國電子郵件與網站保存作業指引之基礎。

為深入了解各國作法，進而研擬適合我國之作業指引，就電子郵件保存而言，本分項計畫將針對以下九個面向進行探討，分別為：法源依據/資料來源、電子郵件保存目的、電子郵件保存範圍、電子郵件管理政策、電子郵件保存系統、詮釋資料、人員權責、員工訓練、與檔案管理系統之關係；其次，透過以下七點作為比較各國網站保存之面項，分別為：法源依據/資料來源、網站保存範圍、網站保存方式、更新頻率、詮釋資料、蒐集/抓取(capture/harvest)方式、與檔案管理系統之關係。

第一節 美國

一、電子郵件保存

由於本分項計畫蒐集到與電子郵件保存相關之資料以北卡羅萊納州較為豐富。因此，本分項計畫彙整美國電子郵件保存內容時，以北卡羅萊納州(North Carolina，簡稱北卡)為主要介紹對象，詳述電子郵件在管理時應注意的相關事項，並以「美國國家檔案暨文件署(National Archives and Records Administration, NARA)」與美國其他州之資料為次要介紹對象，作為研擬本國電子郵件保存作業指引的參考。

(一) 法源依據/資料來源

電子資訊交換日漸普及，各國開始制訂相關法案，藉此闡明電子郵件或檔案規定的保存期限，並使企業不僅要遵循相關法規，也必須重視資料外洩所帶來的損失及應付的責任。其中，與電子郵件保存相關之法案包含：「醫療保險流通與責任法案(Health Insurance Portability and Accountability Act，簡稱 HIPAA)」及「沙賓法案(Sarbanes-Oxley Act，簡稱 SOX)」。

「醫療保險流通與責任法案(HIPAA)」是於西元 1996 年通過，目的是規範醫療工作人員須保障病人隱私，維護員工、客戶及股東資料的完整性，同時，簡化健康醫療的行政程序，以及增加健康照護體系的效率。「醫療保險流通與責任法案(HIPAA)」要求無論是電

子郵件或檔案都需保存至少 6 年。

「沙賓法案(SOX)」則是於西元 2002 年訂立，起因於美國安隆公司、世界通訊公司爆發財務欺詐事件破產，暴露出公司和證券監管問題。「沙賓法案(SOX)」要求上市企業有責任將所有的電子檔案以及電子郵件紀錄，備份保存至少 5 年。此法案企圖讓企業營運及財務資訊更加透明化，並且建立相關規範以確保財務資訊正確性及可靠性。

以下為美國與電子郵件保存相關之法源與資料來源：

1、資料來源：

- (1) 北卡羅萊納州政府檔案管(Government Records Branch of North Carolina)
- (2) 美國國家檔案暨文件署(National Archives and Records Administration, NARA)
- (3) 「阿拉斯加州(State of Alaska)」
- (4) 「西肯德基大學(Western Kentucky University)」
- (5) 「猶他州(Utah)」
- (6) 「明尼蘇達州 Minnesota)」
- (7) 「南卡羅萊納州(South Carolina)」

2、相關法規：

- (1) 「醫療保險流通與責任法案(Health Insurance Portability and Accountability Act , HIPAA)」
- (2) 「沙賓法案 (Sarbanes-Oxley Act , SOX)」
- (3) 「北卡羅萊納州-電子郵件保存及銷毀政策(E-mail as a Public Record in North Carolina: A Policy for Its Retention and Disposition)」
- (4) 「北卡羅萊納州法律彙編(North Carolina General Statutes)」
 - Executive Order No. 18- Email retention and archiving policy
 - §14-113.20. Identity Theft.
 - §14-454 (b). Accessing computers.
 - §121-2(8). Definitions
 - §132-1(a). "Public Records" defined.
 - §132-1.10. Social security numbers and other personal identifying information.
 - §132-1.2. Confidential information.
 - §132-2. Custodian designated.
 - §132-6. Inspection and examination of records.

(5) 「阿拉斯加州電子郵件保存政策(Email Retention Policy)」

- AS 40.21 ("Management & Preservation of Public Records")

- 4 AAC 59.005 ("Retention & Preservation of Electronic Records")

- AS 39.50.200(a)(9)

(6) 「美國聯邦規則法典(Code of Federal Regulations)」

- 36 CFR 1222.38

- 36 CFR 1234.24

(7) 「西肯德基大學網路與電子郵件使用政策(Western Kentucky University-Internet and Electronic Mail Acceptable Use Policy)」

(8) 「猶他州電子郵件指引(Utah-Email Guideline)」

(9) 「明尼蘇達州電子郵件管理(E-mail Management)」

- Official Records Act [Minnesota Statutes, Chapter 15.17]

- Records Management Act [Minnesota Statutes, Chapter 138.17]

- Minnesota Government Data Practices Act (MGDPA)

(10) 「南卡羅萊納州電子郵件管理(E-mail Management)」

- South Carolina Public Records Act [PRA] (Code of Laws of South Carolina, 1976, Section 30-1-10)
- South Carolina Uniform Electronic Transactions Act [UETA] (Code of Laws of South Carolina, 1976, Section 26-6-10 through 26-6-210)
- South Carolina Freedom of Information Act [FOIA] (Code of Laws of South Carolina, 1976, Section 30-4-10 through 30-4-165)

(二) 電子郵件保存目的

公務員已逐漸仰賴電子郵件作為完成政府相關業務的快速、便利之溝通工具，也因為這些郵件檔案具有證明性、保護個人或政府權益之功能、文件化業務決定過程、及歷史和文化的價值，因此一些郵件訊息需因應機關需求、檔案保存需要、及法律用途而被保存下來。

(三) 電子郵件保存範圍

根據「北卡羅萊納州 General Statutes (N.C.G.S.) §121-2(8) and § 132-1(a)」的定義，北卡羅萊納州機關及其子部門根據法令及法條於公務中所產生及接收的電子郵件，應被視為「公務檔案(Public

Records)」。

1、屬於公務檔案的範圍包含：

- (1) 政策與指令
- (2) 最初的草稿或報告及建議
- (3) 公務相關的回應及備註
- (4) 工作排程及分配
- (5) 會議議程及記錄
- (6) 業務過程相關文件
- (7) 判例訊息

2、非公務檔案的範圍包含：

- (1) 私人信件

根據「行政命令第十八號—電子郵件保存與保存政策 (Executive Order No.18-E-mail Retention and Archiving Policy)」，所有在郵件系統上的信件都可被大眾閱覽，公用帳號裡的私人信件不受隱私權保護。

- (2) 垃圾信件
- (3) 無關信件

與其它業務相關或完全無關之信件，需與私人信件一樣，在合理的時間內刪除。

（四）電子郵件管理政策

1、一般政策

（1）電子郵件的擁有權

凡使用政府資訊系統所產生的資料，含電子郵件，公務員不具有任何隱私權，除非有經過機關或政府特准。

（2）個人電子郵件帳號的使用

不鼓勵員工使用私人電子郵件帳號處理政府相關業務。若上述情形發生，則相關電子郵件需轉寄到政府提供之電子郵件帳號；無帳號之員工則需印出郵件訊息，後續遵循檔案保存及銷毀的排程管理。

公務員使用私人電子郵件帳號處理該州公共事務時，需確保所有公務檔案的保存符合「行政命令第十八號—電子郵件保存與保存政策(Executive Order No. 18 - E-mail Retention and Archiving Policy)」之規範。而主管級的公務員需將私人帳號中的公務電子郵寄轉寄到公家的檔案保存處。

2、管理公務電子郵件

(1) 電子郵件的保存(retention)

根據「行政命令第十八號—電子郵件保存與保存政策(Executive Order No.18- E-mail Retention and Archiving Policy)」，將電子郵件依照電子郵件之內容與功能，分為以下三種保存期間：

- 短期保存檔案：對機關亦無重大保存價值的溝通及宣告的訊息均屬之。
- 長期保存檔案：評估過其行政、財政、及法務價值後，判定其對機關有重大價值但不到永久保存的檔案均屬之
- 永久保存檔案：包含州政策、決議、程序、及重要業務的證據且具持續性的歷史價值之檔案均屬之。

(2) 電子郵件的鑑定與評估

北卡羅萊納州曾在「公務電子郵件管理指南(Managing Your Inbox E-mail as a Public Record)」中參考肯德基州所建議的電子郵件的責任歸屬定義，並指引使用者快速判斷並採取動作，詳細作法如圖 3- 1 所示。肯德基州規

範應保存之電子郵件，需依照規範之程序判斷該電子郵件之價值後，才將電子郵件保存，而其規範之程序如圖 3- 2 所示。

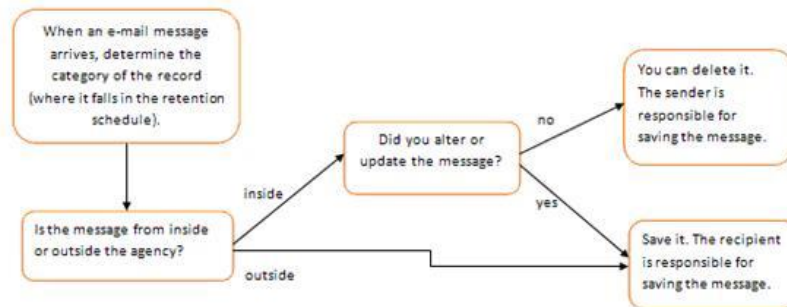


圖 3- 1 保存電子郵件的責任定義

Decision Sequence for Determining E-Mail Retention

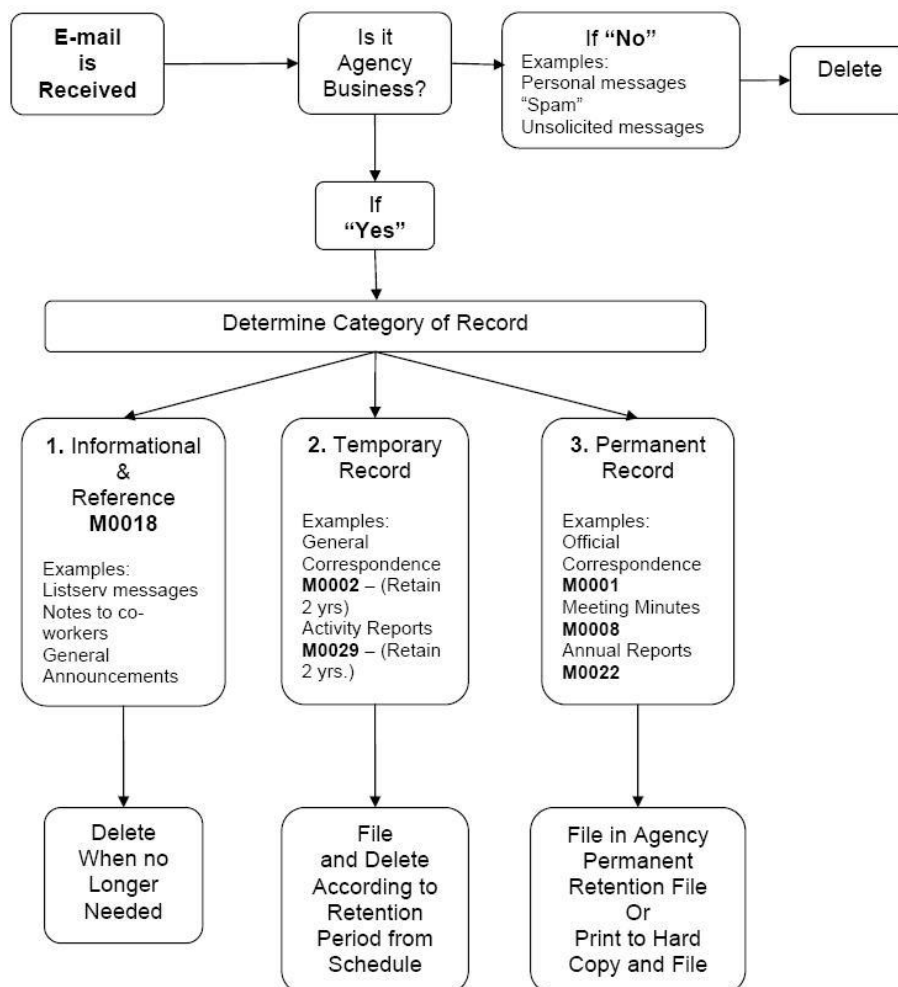


圖 3- 2 電子郵件保存判斷圖

3、電子郵件的存取(Access)

電子郵件的存取或借閱是有限制的，若符合以下條件之一者，才可以存取該封電子郵件：

- 此郵件的寄收者
- 郵件寄收者的主管
- 部門裡網路系統或電子郵件系統的管理者
- 透過法定程序，如 N.C.G.S. §132-2 and §132-6，的個別申請人

此外，若有人需要存取一份包含機密機密性資料之電子郵件時，檔案管理人員必須調整機密資料的內容。所謂機密性，是指涉及以下幾點內容者，均屬機密性檔案：

- 醫療保險轉移和責任法裡包含的資訊(與處理醫療保險資訊的可攜性及保護病患的隱私權問題有關)
- 家庭教育權利和私隱條約裡包含的資訊(保證在校學生的私隱權)
- 社會安全號碼與其它私人認證資訊
- G.S. § 132-1.2.中定義之機關交易秘密

（五）電子郵件保存系統

「北卡羅萊納州電子郵件保存及銷毀政策(E-mail as a Public Record in North Carolina: A Policy for Its Retention and Disposition)」並未對電子郵件保存系統提供相關的注意事項。因此，以下將以「美國國家檔案暨文件署(National Archives and Records Administration, NARA)」及猶他州所建議的電子郵件保存系統注意事項作為此議題之參考。

1、美國國家檔案暨文件署(National Archives and Records Administration, NARA)

美國國家檔案暨文件署(NARA)曾於 2008 年 5 月公告「電子郵件保存系統指引(Guidance concerning the use of e-mail archiving applications to store e-mail)」，並提出以下對於電子郵件保存系統之建議、注意事項、效益、爭議：

(1) 電子郵件保存系統相關建議：

- 「電子郵件保存系統」是指可將電子信件從信箱伺服器移至中央貯藏處並統一管理之保存軟體。
- 符合聯邦檔案法規定之電子郵件，均屬於保存範圍。
- 電子郵件保存系統應具備之功能包含：具分類功

能、簡易與即時的檢索、合宜的保存格式、可取用檢視、保留寄收資料、支援永久保存檔案的轉移。

- 電子郵件保存系統可提供自動或手動移轉出郵件伺服器中的電子郵件到電子郵件保存系統中，並將這些進入到保存系統中的電子郵件建立索引。

(2) 使用電子郵件保存系統之注意事項

- 系統有其限制，不一定能滿足各項法規要求
- 機關須負責處理電子郵件保存系統不足之處
- 若電子郵件保存系統是唯一存放聯邦檔案的方式，機關必須詳細定義系統功能、負責處理不足之處、制定相關政策以確保系統符合「36 CFR 1234.24.」規範

(3) 電子郵件保存系統的效益

- 提升儲存效率
- 加強搜尋能力
- 確保備份及災後恢復

(4) 電子郵件保存系統的爭議

- 群組、維護功能不足
- 難區分永久、暫存檔案；無保存屆期後處置方式
- 電子保存檔案與書面保存抵觸。

2、猶他州

(1) 檔案管理及其保存系統

電子郵件檔案需放置於某種檔案保存系統中。此保存系統可根據目的及保存排程分類檔案，並提供有效存取檔案之功能、保存重要資訊之功能、以及相關的安全機制，例如：可拒絕無授權取用及破壞的安全機制。此外，倘若檔案儲存於此系統後，便可自原先儲存電子郵件的地方移除該郵件，且任何原先電子郵件資料的備份，亦可被銷毀。

電子郵件保存系統可分為「電子郵件系統(E-mail System)」及「中央系統(Centralized Systems)」。

- 電子郵件系統

截至目前為止，大部份電子郵件管理系統仍仰賴使用者透過系統資料夾，或類似選項管理電子郵件。藉由資料夾管理郵件對於使用者而言，較為簡單、

易學。當具保存價值之郵件被寄收，此郵件可被立即移至合宜的資料。若電子郵件較不常使用時，可根據保存年限(Retention Schedule)刪除該電子郵件，或移轉至永久保存之格式。

但是，此種由使用者管理資料夾的簡易型的電子檔案管理系統，容易因員工的判斷誤差，而產生與保存政策相抵觸之情形。

- 中央系統

電子郵件保存也可以透過委外開發、採購的方式取得。這種中央系統提供較強的檔案管理功能，例如：不僅檔案創建者可以檢視、銷毀或移轉該電子郵件。同時，這種中央系統也提供監督的功能。這種系統透過集中管理的方式，不僅可以減輕複製品所引發的爭議，並有條理地回應許多使用者的需求。

(2) 管理政策

- 電子郵件管理系統的基本要素

根據相關政策、管理規則或法令，每個州機關被要求使用經同意的電子檔案管理系統，或建立一套自

己的政策是可符合上述系統的基礎標準。政策需描述系統如何使用及所包含的檔案。為確保所有重要的電子郵件可在管理系統中被存取，政策應要求所有州業務需在可連接已授權管理系統之電腦及設備上處理。

- 附件管理

電子郵件附件資料應採用電子郵件檔案系統做更有效之管理，使得使用者透過一般電子郵件系統搜尋到附件資料。

- 儲存空間(Storage)

為使州政府具有電子郵件保存之能力，應先建立以下類型的儲存空間：

類型 1. 具有自動保存及銷毀使用頻繁(至少 60 天內)

特性之電子郵件於中央儲存空間。

類型 2. 可供電子郵件使用頻繁的單一機關或多機關的儲存空間。

類型 3. 過渡型的儲存空間。

類型 4. 可保存「已移轉至州檔案管理局之電子郵

件」、「特殊類型電子郵件」、「可透過檔案管理系統存取之電子郵件」之儲存空間。

類型 5. 離職員工電子郵件的儲存空間。

類型 1、2 和 4 需高速高取得性的環境；而類型 3 及 5 只需低價硬碟即可。

- 電子郵件的保存及銷毀

以下保存年限不僅適用於電子郵件，亦適用於各類信件：

暫時性通信記錄：正本或複本可由機關保存直至使用需求終止，而後銷毀。

政策、方案通信記錄：正本可由機關暫存至使用結束，然後移轉至州檔管局；當使用完畢後，複本可予以刪除。

- 電子郵件的保存

猶他州要求在檔案的保存年限內，其全文可開放檢視。電子郵件的保存，即使保存年較限短之電子郵件，仍需以開放格式儲存電子郵件、電子郵件詮釋資料、電子郵件附件於於檔案保存系統。

（六）詮釋資料

「北卡公務電子郵件保存及銷毀政策(E-mail as a Public Record in North Carolina: A Policy for Its Retention and Disposition)」並未提及電子郵件保存時應包含的詮釋資料，因而此議題將參考肯德基州及明尼蘇達州的建議，作為詮釋資料的補充。

1、肯德基州

詮釋資料囊括電子郵件檔案的所在位置、活動軌跡及使用目的等資訊，包含：

- (1) 寄件者及收件者姓名
- (2) 郵件產生及寄送日期
- (3) 郵件產生的應用程式相關資訊
- (4) 郵件所行經的所有系統、電腦的歷史記錄

某些電子郵件軟體可支援詮釋資料直接在郵件裡顯示，然其它軟體將詮釋資料置於 Header，或存於「屬性」檔。

2、明尼蘇達州

所謂電子郵件檔案應包含以下幾項特性：

- (1) 完整性：電子郵件記錄需完整的文件化交易紀錄，必須

包含以下元素

- 收件者
- 寄件者
- 主旨
- 內文
- 寄出資料
- 寄出時間
- 完整的附件—附件不能僅有檔案名，附件檔須完整附加，因其亦為此檔案的一部份
- 群組成員名單—若電子郵件檔案僅註明群組名，收件者也可藉由群組名單辨識成員資訊。
- 電子郵件帳號與使用者名稱的對照表

(2) 正確性：含交易紀錄完整始末資料。

(3) 可存取性：區分可開放及禁止開放的資料。

(4) 易管理性：電子郵件應為全體職員日常業務的一部份。

為使職員執行並使用電子郵件檔案管理政策，程序必須直覺且直接。

(5) 安全性：因電子郵件易散佈特性，更須嚴謹的系統控管。

(七) 人員權責

北卡羅萊納州與阿拉斯加州分別對機關、職員的責任範圍定出了一些可參考的注意事項：

1、北卡羅萊納州：「行政命令第十八號—電子郵件保存與保存政策(Executive Order No.18-E-mail Retention and Archiving Policy)」

(1) 職員權責

- 職員以州政府電子郵件帳號所收發的電子郵件均須視為公務檔案，並遵循與紙本、實體檔案相同法令規範。
- 在州電子郵件系統中的文件一律視為可公開，無論對象為大眾或政府職員，且不受隱私權保護。
- 任何從行政部門電子郵件帳號外寄的信件均有適當語言可辨識此訊息傾向於公開檔案法亦或需對第三方隱匿。
- 職員不可永久性地刪除 24 小時內寄收的信件，除非收到的信件與政府行政業務無任何關連。

- 職員以私人帳號處理的政府業務信件需確保遵循

「行政命令第十八號—電子郵件保存與保存政策
(Executive Order No.18-E-mail Retention and
Archiving Policy)」及合適的檔案保存排程保存。

- 州政府電子郵件帳號禁止被用於政黨用途。

(2) 機關權責

- 機關需至少每日一次備份所有寄收資料。
- ITS (Information Technology Services) 將提供機關電子郵件保存系統。此系統裡的信件應保存 10 年。
若信件需保存差超過 10 年，ITS 會再與北卡羅萊納州文化資源部門〈簡稱 DCR〉商議。
- 北卡羅萊納州文化資源部門(DCR)需提供線上訓練手冊。
- 北卡羅萊納州文化資源部門(DCR)應隨機稽查機關，確保職員遵守保存銷毀排程。

2、阿拉斯加州：針對主管職位(Executive)與非主管職位(Non Executive)的員工有不同的電子郵件保存程序建議。

(1) 非主管職位

- 資料夾中的電子郵件

根據電子郵件保存政策，所有非主管職員的信件，除了已移至保存資料夾的信件除外，均將在 90 天後自動刪除。使用者需將電子郵件檔案放置於合適的 Enterprise Vault (簡稱 EV)保存資料夾。非主管職員應有五個電子郵件保存資料夾，分別為保存「1 年」、「3 年」、「5 年」、「7 年」、及「10 年」，如圖 3-3 所示。。雖然大多信件為短期，但小比例的需要長期資料夾的使用者可向他們的部門主管或機關檔案官員申請使用「25 年」、「50 年」、或「永久」的資料夾。

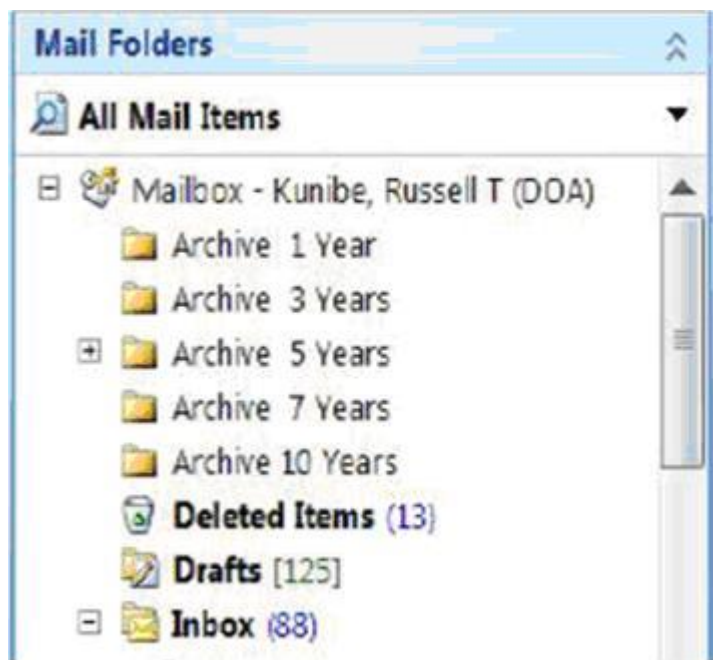


圖 3- 3 非主管職員的資料夾管理

- 刪除資料夾中的電子郵件

使用者在刪除電子郵件之後，該郵件會被移至刪除資料夾。所有在刪除資料夾中的電子郵件都將在七天之後被永久性的刪除，詳細規則如表 3- 1 所示。

- 寄件備份及使用者定義資料夾皆會在 90 天後自動刪除

表 3- 1 非主管職員的電子郵件政策指引

項目	規則
寄收的郵件	90 天內應保存具保存需求之郵件，並移至合適資料夾。若郵件同時符合多項保存要求，擇最久期限的保存項目執行。
未封存郵件	30 天後被封存，90 天後自動刪除。
封存資料夾選項	1 年、3 年、5 年、7 年、10 年（25 年、50 年、及永久）。
子資料夾	繼承父資料夾的保存。
已刪除項目	七天後清除
行事曆項目	五年後刪除

(2) 主管職位

- 資料夾中的電子郵件

根據電子郵件保存政策，所有主管職員的信件，除了已刪除的之外，需在 90 天後自動封存。使用者需將電子郵件檔案放置於合適的 Enterprise Vault〈後簡稱 EV〉保存資料夾。主管職員應有九個電子郵件保存資料夾，分別為保存「90 天」、「1 年」、

「3 年」、「5 年」、「7 年」、「10 年」、「25 年」、「50 年」、及「永久」。如圖 3-4 所示。



圖 3- 4 主管職員的資料夾管理

電子郵件年限分類準則與紙本同步。EV 裡的電子郵件在截止日前，不可被刪除。即使 Outlook 信箱裡的郵件捷徑被刪除，該郵件依然保存於 EV 中，並遵循著它的保存期限。然而，捷徑可被移至不同的保存資料夾，而其保存期限會因此被重置。

- 刪除資料夾中的電子郵件

使用者在刪除電子郵件之後，該郵件會被移至刪除資料夾。所有在刪除資料夾中的電子郵件都將在七天之後被永久性的刪除。詳細規則如表 3-2 所示。

- 行事曆項目的保存：期限為五年。

表 3- 2 主管職員的電子郵件政策指引

項目	規則
寄收的郵件	90 天內應保存具保存需求之郵件，並移至合適資料夾。若郵件同時符合多項保存要求，擇最久期限的保存項目執行。
未封存郵件	30 天後被封存，90 天後自動刪除。
封存資料夾選項	1 年、3 年、5 年、7 年、10 年（25 年、50 年、及永久）。
子資料夾	繼承父資料夾的保存。
已刪除項目	七天後清除
行事曆項目	五年後刪除

（八）員工訓練

「北卡公務電子郵件保存及銷毀政策(E-mail as a Public Record in North Carolina: A Policy for Its Retention and Disposition)」在此議題僅提供可諮詢之部門，並未對訓練的細目作出建議。因此，以明尼蘇達州與肯得基州對員工訓練的設計作為參考。

1、明尼蘇達州

明尼蘇達州提出合適的員工訓練應可培養員工回答以下問題的能力，例如：

- (1) 此電子郵件是否為正式記錄？此電子郵件為公務或私人的？
- (2) 此電子郵件是否有長期重要性？此電子郵件是否內含業務或營運資訊？
- (3) 根據明尼蘇達州政府資料實務法（Minnesota

Government Data Practices Act)，此電子郵件為可公開
或非公開記錄？

- (4) 當保存此電子郵件，我應該抓取哪些詮釋資料
(Metadata)？
- (5) 此電子郵件應屬於哪個檔案系列？
- (6) 我是否該完整的保存此電子郵件，包含其附件及群組列
清單？
- (7) 此電子郵件是否被申請過成為法定活動的證據？

2、肯德基州

肯德基州建議員工訓練範圍包含以下幾項議題：

(1) 正確的使用

職員的電子郵件帳號為處理公務使用，機關有權利監視
及瀏覽內容。

(2) 機密性議題

敏感及機密性資訊不應透過電子郵件傳遞，除非有合適
的協定(Protocols)能保護此訊息。

(3) 明確主旨

應使用簡潔且明確的主旨作為電子郵件的標題。

(4) 防毒概念

政府機關為駭客常攻擊的目標，職員應留意電子郵件可能夾帶的病毒，機關也需建立開啟附件時應注意的相關政策。

(5) 加密時機

加密可能與檔案長期保存策略衝突，郵件應在描述加密需求及闡述解密方法之後，才允許使用加密。

(九) 與檔案管理系統之關係

南卡羅萊納州認為的保存電子郵件檔案之方式中，有兩種較趨於集中化管理電子郵件及其它的電子檔案，分別為電子歸檔及EDMS；

1、電子歸檔(Electronic Filing)

電子郵件系統非檔案保存系統，重要信件若與其它相關資訊斷連，易導致資料流失。無歸檔系統也需將電子郵件放至命名清楚的資料夾中。最後再將此資料夾匯出至電子歸檔系統(Electronic Filing System)。

2、電子檔案管理系統(Electronic Document Management

System, EDMS)

可將電子郵件與其它電子檔案共同置於電子檔案管理系統一併管理。為維護電子系統，需採用標準流程。電子檔案管理系統會逐漸與內容管理系統(Content Management System)整合。

二、 網頁保存

(一) 法源依據/資料來源

1、 資料來源：

(1) 北卡羅萊納州

- Collection Procedures for State Government Web sites using Archive-It (2006)
- Standard for Automated Web Site Capture(2006)

(2) 堪薩斯州

- Guidelines for Managing Records on Kansas Government Agency Web Sites. January 2004

(3) 南卡羅萊納州

- Electronic Records Management Guidelines—Web Content Management. March 2005

2、相關法規：

(1) 卡羅萊納州

- General Statute 132.1 and 125.11.

(2) 堪薩斯州

- Government Records Preservation Act (K.S.A. 45-401 through 45-413)
- Public Records Act (K.S.A. 75-3501 through 75-3518)
- Open Records Act (K.S.A. 45-215 through 45-223)

(3) 南卡羅萊納州

- Public Records Act [PRA] (Code of Laws of South Carolina, 1976, Section 30-1-10 through 30-1-140)
- South Carolina Uniform Electronic Transactions Act [UETA] (Code of Laws of South Carolina, 1976, Section 26-6-10 through 26-6-210)
- South Carolina Freedom of Information Act [FOIA] (Code of Laws of South Carolina, 1976, Section 30-4-10 through 30-4-165)

(二) 網站保存範圍

以政府機關而言，北卡羅萊納州的文化資源部門負責鑑定網站是否具數位資產的抓取價值，並用以下準則判斷網站是否在保存範

圍中：

- 1、保存具正式州政府資料的公私立網站。例如：北卡美術館的網站，雖然架設於私人伺服器上，也是在網站保存範圍中。
- 2、不保存私人營利或非營利組織網站。例如：北卡動物協會。
- 3、保存含可協助政府策略的重要資料之私人網站。例如：北卡教育協會。
- 4、不保存重覆資料之網站。例如：北卡稅收部門及其別名。

網際網路檔案館(Internet Archive，簡稱 IA)則保存所有公開且可進入的網站，除了以下四種網站之外：

- 1、需密碼才能取用的網站
- 2、網站擁有人貼上 Robot exclusion 的標籤之網站
- 3、需人為輸入產生表格之頁面
- 4、存於安全伺服器的網站

(三) 網站保存方式

(四) 更新頻率

依據北卡羅萊納州的網站保存方式，所有定義在範圍內且可被抓取的網站將會被送至宏觀鑑定(Macro-Appraisal)網站，使用該評分系統去描述網站抓取標準文件。網站的分數與網站的價值相關於Archive-It的服務限制，也決定了抓取頻率。保存網站的種子列表會因應網站翻新及廢止而隨年調整。抓取頻率的調整也會仰賴保存檔案的大小及成長率。若某領域因技術限制而落於範圍外，該網站也不會被抓取，因此機關需要與北卡羅萊納州文化資源部門(DCR)合作確保網站有被保存。北卡羅萊納州文化資源部門(DCR)職員運用網際網路檔案館(Internet Archive，簡稱IA)所提供之報告瞭解網站抓耙器(crawler)不能抓取網站，再聯絡網站機關討論抓取及保存的替代方案。

1、網站需以 Macro-Appraisal Score Table 評分網站的七項面向，如表 3- 3 所示。

2、根據抓取頻率表得知網站的抓取周期，如表 3- 4 所示。

表 3- 3 鑑定積分表

	1 point – Low value	2 points – Medium value	3 points – High value
1: Size	Small Web site with few directories	Large Web site with multiple directories	Large Web site with multiple directories from multiple offices
2: Originality	Mainly copies of paper publications	Combination of original material and copies of paper material	Substantially original material not available in paper form
3: Frequency of update	Static Web site with time-insensitive material; information rarely changes	Web site updated on a regular basis; information current and fresh	Web site information changes frequently; timeliness and currency of information extremely important
4: Historical value	Information is ephemeral; little value beyond the present day	Some original information with historical value; partially documents decisions/trends of government and programs	Rich original information; statistics/facts document government decisions/trends and programs
5: Evidential value	Information never litigated and little or no legal risk	Potential for litigation	Substantial liability
6: Public interest	Low public interest; little press coverage	Medium public interest; some press coverage; potentially controversial information	High public interest; frequent press coverage; public interest groups monitor sites
7: Governmental interest	Little legislative, executive, or judicial interest	Some legislative, executive or judicial interest	Active and vigilant legislative, executive, or judicial interest

表 3- 4 抓取頻率表

Frequency of crawl, based on Macro-Appraisal Score (4/2806-6/15/06)	
7.0-7.99:	Crawled annually
8.0-10.99:	Crawled quarterly
11-21:	Crawled monthly
Frequency of crawl, based on Macro-Appraisal Score (6/15/06-)	
7.0-7.99:	Crawled annually
8.0-21:	Crawled quarterly

(五) 詮釋資料

南卡羅萊納州為確保檔案可被完整的呈現內容、背景、及架構，強調詮釋資料的重要性。多數機關選擇使用都柏林核心集 (Dublin Core Metadata Element Set)作為標準，包含以下元素，星號「*」代表若為長期保存或永久保存之檔案，則建議應包含之基本元素：

1、*標題：資源(Resource)創造者或出版者所給予的名字。

2、*主題：資源主題。

- 3、*描述：資源內容簡短文字性之描述。
- 4、*創建者：資源創建者的名字。
- 5、公開者：公開發佈資源者姓名。非指將資源貼到網站上的對象，而是負責發佈的單位，如機關部門。
- 6、貢獻者：創建者除外對此資源有重大貢獻者。
- 7、*日期：創建及公開日期。機關需定義日期格式，如
yyyy-mm-dd。
- 8、*來源類型：資源所屬種類，如會議記錄、新聞稿、或報告。
- 9、*格式：資源檔案格式。
- 10、*識別項目：資源獨有的文字列或數字，如 URL 或檔案名稱。
- 11、 相關資料：資源所參考的資料。
- 12、 來源：目前資源的取得來源。
- 13、 權利管理：關於版權所有的文字敘述。
- 14、 語系：資源使用到的語言。
- 15、 範圍：地區(如南卡)及時間範圍(如 2000-2001)。

(六) 蒐集/抓取(capture/harvest)方式

北卡羅萊納州與國會圖書館皆對網站蒐集方式，運用不同工具以進行抓取。

1、北卡羅萊納州

北卡羅萊納州使用網際網路檔案館(Internet Archive，簡稱IA)所開發的 Archive-It 作為永久保存網站的搜集、儲存、及使用媒介。然由於 Archive-It 的限制，有些落於保存範圍內的網站，也會被 Archive-It 排除在外，例如下述項目

(1) 因技術、程序、及經費限制，地方機關及公立大專院校之網站不保存。

(2) 需要使用者互動產生的資料內容不保存。

(3) 高度仰賴描述來呈現內容的網站不保存。

若機關選擇不用 Archive-It 自動保存，或因為 Javascript 的限制而無法被保存，DCR 建議機關可採用網站快照(snapshot)方式。DCR 建議快照可在主要版本變動時再實施，或至少一年一次。

2、國會圖書館

國會圖書館是盡可能完整的保存網站所有元件，以及網站

的詮釋資料。因此，國會圖書館(Library of Congress)使用下列工具進行網站的獲取及使用媒介。

(1) 獲取方式：主要工具為 Heritrix 網站抓耙器(crawler)。

(2) 存取工具：國會圖書館以 WERA 和 Wayback 為存取工具並用 NutchWAX 為搜尋引擎。

(七) 蒐集/抓取(capture/harvest)方式

北卡羅萊納州與國會圖書館皆對網站蒐集方式，運用不同工具以進行抓取。

1、北卡羅萊納州

北卡羅萊納州使用網際網路檔案館(Internet Archive，簡稱 IA)所開發的 Archive-It 作為永久保存網站的搜集、儲存、及使用媒介。然由於 Archive-It 的限制，有些落於保存範圍內的網站，也會被 Archive-It 排除在外，例如下述項目

(1) 因技術、程序、及經費限制，地方機關及公立大專院校之網站不保存。

(2) 需要使用者互動產生的資料內容不保存。

(3) 高度仰賴描述來呈現內容的網站不保存。

若機關選擇不用 Archive-It 自動保存，或因為 Javascript 的限制而無法被保存，DCR 建議機關可採用網站快照(snapshot)方式。DCR 建議快照可在主要版本變動時再實施，或至少一年一次。

2、國會圖書館

國會圖書館是盡可能完整的保存網站所有元件，以及網站的詮釋資料。因此，國會圖書館(Library of Congress)使用下列工具進行網站的獲取及使用媒介。

- (1) 獲取方式：主要工具為 Heritrix 網站抓耙器(crawler)。
- (2) 存取工具：國會圖書館以 WERA 和 Wayback 為存取工具並用 NutchWAX 為搜尋引擎。

(八) 與檔案管理系統之關係。

第二節 英國

一、電子郵件保存

本分項計畫以英國國家檔案管理局(The National Archive, 簡稱 TNA) (TNA)所規範之電子郵件保存相關資料為主要蒐集對象，並蒐集蘇格蘭的地政機關(Register of Scotland)與里茲大學(University of Leeds)所依循之電子郵件保存政策。

(一) 法源依據/資料來源

1、資料來源：

- (1) 英國國家檔案局(TNA)-Management, appraisal and preservation of electronic records
- (2) 英國國家檔案局(TNA)-Guidelines on developing a policy for managing email
- (3) 蘇格蘭的地政機關(Registers of Scotland)的 email policy
- (4) 里茲大學(University of Leeds)的電子郵件保存(Email Archiving)

2、相關法規：

- (1) 資料保護法(Data Protection Act)
- (2) 資訊公開法(Freedom of information Act)

(3) 隱私權(Regulation of Investigatory Privacy Act)

(二) 電子郵件保存目的

根據英國國家檔案局(TNA)所制訂的電子檔案管理、鑑定與保存(Management, appraisal and preservation of electronic records)，中規定，所謂電子檔案，包含：文字為基礎的文件(text-based documents)、文書處理的文件(word-processed documents)、電子郵件訊息(e-mail messages)、電子資料表(spreadsheets)、簡報(presentations)、掃描文件(scanned documents)。

因此，電子郵件可視為檔案，機關必須負責妥善保存與管理，並且確保電子郵件管理相關規範依循資料保護法資料保護法(Data Protection Act)、資訊公開法(Freedom of Information Act)、隱私權(Regulation of Investigatory Privacy Act)。

此外，英國國家檔案局(TNA)所制訂的電子郵件政策發展指引(Guidelines on developing a policy for managing email)建議各機關單位應制定電子郵件政策(email policy)，以免失去重要資料、或組織成員不當的使用電子郵件；該指引中並說明各機關制定電子郵件政策的目的是在於確保組織成員瞭解何為不適當的使用電子郵件，以及了解電子郵件與業務活動間之關聯性。同時，制定電子郵件政策可協助機關界定期對電子郵件管理之權責，且機關有責任去確保員工

不受到來自電子郵件的言語或文字上的污辱。

（三）電子郵件保存範圍

根據英國國家檔案局(TNA)所制訂的電子郵件政策發展指引
(Guidelines on developing a policy for managing email)，認為會成為檔案的電子郵件包含：討論、一群人之間傳遞的訊息、核准辦理的同意書、其他與履行業務相關的資訊，且所有員工都必須能辨認出電子郵件是否該視為檔案。

（四）電子郵件管理政策

1、英國國家檔案局

根據英國國家檔案局所制訂的電子郵件政策發展指引
(Guidelines on developing a policy for managing email)中建議各機關單位之電子郵件管理方式包含以下幾點：

(1) 信箱管理

由於公開信箱(public mailbox)與分享信箱(shared mailbox)會涉及許多參與同一工作之員工，因此較不易管理。若公開信箱(public mailbox)式開放給組織中所有員工使用，而分享信箱(shared mailbox)是用來處理特定工作相關的資料傳送，則建議公開信箱(public mailbox)

應該針對單一個目的來作使用，且在設立一個公開信箱(public mailbox)時，應該要釐清其設立原因、如何使用、使用者會是誰、這個信箱中的討論內容多久之後會被移除，也就是要讓使用者之到這個信箱可以使用多久，以及會移到哪裡去等問題。

此外，若公開信箱(public mailbox)是用來作員工間討論，並非用來管理討論的紀錄，那麼這些電子郵件可能會印成紙本，在電子檔案管理系統中管理，或者要放在一個與公開信箱不同的地方特別管理。在管理公開信箱時，應該要評估所有電子郵件是否要視為檔案與以保存，而非在信件往來的最後才決定。

(2) 附件管理

電子郵件管理指引時，應決定是否要將附件一併視為檔案。

(3) 電子郵件加密

如果要管理加密後的電子郵件會有很多問題，如果組織收到或發送加密的電子郵件，則必須考慮到未來如何存取此電子郵件。

(4) 與檔案管理系統整併

如果組織有電子檔案管理系統，那應該很容易可以說服使用者把電子郵件歸檔到電子檔案管理系統中，因為那是其他檔案都存放、管理的地方。若組織並沒有電子檔案管理系統，則建議可以把這些是為檔案的電子郵件放到共用資料夾(shared public email folder)或者把這個電子郵件以「.msg」的格式保存在共用的硬碟中。

2、英國里茲大學(University of Leeds)

根據英國里茲大學(University of Leeds)電子郵件保存(Email Archiving)作法中提到，該校保存政策(Archiving Policy)包含：

(1) 符合保存資格的項目(Items eligible for archiving)

使用者信箱中開啟過的電子郵件、任何使用者定義的資料夾、寄出的信件。

(2) 拒絕保存的項目(Items excluded from archiving)

未讀之信件、刪除之信件、行事曆、備忘錄、工作、公開資料夾。

(3) 自動保存政策(Automated archiving policy)

所有符合保存條件滿三個月者(從收到資料或是最後修改資料算起)，將會被自動的保存。

(4) 使用者導向之保存(User-driven archiving)

由使用者可以手動設定自動把檔案保存，而非等到系統自動去保存。

(5) 離線存取保存信件(Offline access to archived items)

使用者有安裝離線存取電子郵件信箱者需要事先提出申請，才可以離線存取保存之檔案。

(6) 保存信件(Retention of archived items)

保存的信件不一定會被儲存，直到該校實行機構電子郵件保存與刪除政策(institutional email retention and deletion policy)

(7) 已刪除檔案的復原(Recovery of deleted archived items)

從保存信件中永久刪除者，將會被保留 90 天。如果是不小心刪除者，可以重新儲存到保存的信件中。

(五) 電子郵件保存系統

根據英國里茲大學(University of Leeds)電子郵件保存(Email Archiving)作法中提到：

- 1、該校採用的是" Symantec's Enterprise Vault" 市面上的電子郵件保存系統(email archiving system)
- 2、電子郵件保存是藉由每晚掃描 Exchange database 以符合保存政策(Archiving policy)
- 3、這些資料會移到次要的儲存媒體上，並且在使用者的信箱中會被訊息捷徑(message shortcut)取代
- 4、如果要取得這些資料，只要在捷徑上雙擊便可檢索出以被保存(archive)之電子郵件。
- 5、使用者可搜尋已保存之訊息，也可以永久刪除這些訊息。
- 6、使用者可以透過手動選擇，把沒有被自動保存的資料加以保存(archive)

(六) 詮釋資料

英國國家檔案管理局參考美國針對檔案管理系統所規範之標準——「DoD5015.2」中所規範之電子郵件標頭(header)基本要素，而修訂適合英國的電子郵件保存詮釋資料。其中，英國發現「DoD5015.2」中所列之詮釋資料和英國國家檔案管理局所規定的要求相當一致，僅用語上有差異。因此，規範機關在保存電子郵件詮釋資料時，應涵蓋以下六項詮釋資料：

- 1、 寄件者全名
- 2、 主要地址的全名
- 3、 其他地址的全名
- 4、 資料傳送日期與時間
- 5、 資料接收日期與時間
- 6、 訊息主題

（七）人員權責

根據英國國家檔案局所制訂的電子郵件政策發展指引 (Guidelines on developing a policy for managing email)，認為管理電子郵件之公開信箱(public mailbox)與分享信箱(shared mailbox)時，相關人員權則應從部門與擁有者的角度思考。部門應負責控制與存取之責任，而擁有者應負起監看之責任。當兩者同時並行，才可確保電子郵件妥善保存。

（八）員工訓練

根據英國國家檔案局所制訂的電子郵件政策發展指引 (Guidelines on developing a policy for managing email)，應讓組織成員對電子郵件保存之認知包含以下幾點：

1、何時該使用電子郵件

使用電子郵件有許多優缺點，優點是可以快速的傳遞與交換資訊，但是，缺點是對收信者而言，可能會收到過多不必要的資訊。同時，為了要最大化電子郵件的優點，與最小化其缺點，組織應該要決定在傳遞或交換資訊時，是否有些狀況可以採用電子郵件以外的替代管道。例如，如果要減少電子郵件被傳送給組織中所有人，可能要檢討公司內部網路(intranet)的使用方式，並且採用其他辦公方式(procedure)來取代電子郵件的傳遞。

2、撰寫郵件時該注意的規矩

在撰寫電子郵件時，應該要注意信件內容的形式，是否屬於正式的商用書信，以免危害到公司。例如，電子郵件內容被誤解、或是造成處理失當。所以員工應該要瞭解，不論用電子或是紙本方式在傳遞訊息時，都應該遵守正式溝通的一些原則，例如文法、拼字。

3、遇到機密或敏感性資料時該如何處理

在傳遞電子郵件時，可能會涉及到敏感性資料的範圍主要有：與國家安全相關的資訊、與商業機密相關的資訊、個人隱私相關的資訊，但因各單位屬性不同，可視情況而定。

這些資料都應該要謹慎處理，以免送錯。因為組織中一定會有一些敏感的資料，因此必須要選定依些可確保各種資料溝通的方式。也因此，電子郵件應該要考慮採用加密的方式。

4、不要濫用電子郵件

電子郵件的濫用將會是組織的一大問題，而且很難察覺。有各種不同的電子郵件濫用範圍，像是過度使用私人電子郵件作出歧視性的講話，因此，組織有責任要讓所有員工知道，再傳遞電子郵件時，哪些評論是不被接受的。

5、處理大量電子郵件該注意：

有些員工必須處理大量的電子郵件，這些人可能會不易管理其電子郵件，但是許多電子郵件管理系統都有提供一些協助管理的工具，因此建議要給員工具體的管理建議還有有效整理電子郵件的作法。

6、命名

檔案在電子檔案管理系統（Electronic Records Management System, ERMS）中標題是檢索檔案重要指標，許多電子檔案管理系統都把電子郵件的標題直接用來當電子檔案管理

系統中的標題，但是這樣會無法清楚的描述電子郵件被視為檔案的原因。可是又不可以更改原本電子郵件的標題。因此建議把電子郵件存到電子檔案管理系統時給予電子郵件一個新的標題，而原本電子郵件的標題維持原狀，就像是加上一個標籤。

（九）與檔案管理系統之關係

據英國國家檔案管理局所制訂的電子郵件政策發展指引 (Guidelines on developing a policy for managing email)，電子郵件與電子檔案管理系統有以下幾點關係；

- 1、當電子郵件被視為檔案時，這個資料必須和其他檔案一起保存。換言之，這表示這些電子郵件應該要從個人的信箱中移到特定的檔案管理系統中。而電子郵件蒐集到檔案管理系統的時間點，就是這份檔案蒐集的時間點。
- 2、就電子郵件而言，蒐集(Capture)檔案是指找出電子郵件與其他檔案間的關聯性，其他檔案可以是在電子檔案管理系統中或是紙本。
- 3、如果組織有電子檔案管理系統，那應該很容易可以說服使用者把電子郵件歸檔到電子檔案管理系統中，因為那是其他檔案都存放、管理的地方。如果組織並沒有電子檔案管

理系統，則建議可以把這些是為檔案的電子郵件放到 shared public email folder 或者把這個電子郵以.msg 的格式保存在共用的硬碟中。

4、檔案在電子檔案管理系統中標題是檢索檔案重要指標，許多電子檔案管理系統都把電子郵件的標題直接用來當子檔案管理系統中的標題，但是這樣會無法清楚的描述電子郵件被視為檔案的原因。可是又不可以更改原本電子郵件的標題。因此建議把電子郵件存到 ERMS 時給予電子郵件一個新的標題，而原本電子郵件的標題維持原狀，就像是加上一個標籤。

二、 網站保存

本分項計畫網站保存之資料來源以英國網站保存聯合會（UK Web Archiving Consortium，簡稱 UKWAC）為主要蒐集對象，並且就英國其他政府單位所制訂之網站保存資料進行彙整，其中以英國運輸部公路局之資料較為豐富。

- 英國網站保存聯合會（UK Web Archiving Consortium，簡稱 UKWAC）由以下六個機構所組成：
- 大英圖書館（British Library）

- 威爾斯國家圖書館 (National Library of Wales)
- 蘇格蘭國家圖書館 (National Library of Scotland)
- 國家檔案局 (The National Archives)
- 聯合資訊系統委員會 (The Joint Information Systems Committee)
- 惠康圖書館 (The Wellcome Library)

英國網站保存聯合會(UKWAC)成立於西元 2003 年，並且於 2003 年開始與網際網路檔案館(Internet Archive，簡稱 IA)簽合約，獲得存取其資料之權力。網際網路檔案館(Internet Archive，簡稱 IA)提供自 1997 年以來的網站保存資料，意即英國網站保存聯合會所擁有之內容可以追溯到 1997 年。截至 2010 年四月，英國網站保存聯合會(UKWAC)所保存之網站已達 6,287 個，資料量已達 5.99TB，檔案數量多達 145,103,185。

(一) 法源依據/資料來源

1、資料來源：

- (1) 英國網站保存聯合(UK Web Archiving Consortium, UKWAC)
- (2) 英國資訊系統聯合委員會(Joint Information Systems Committee, JISC)-The Preservation of Web Resources Handbook
- (3) 英國運輸部公路局(Highways Agency)-Web Archive Policy

2、相關法規

(1) 「資訊公開法」(Freedom of Information, FOI Act)

(2) 「資料保護法」(Data Protection Act, DPA)

(3) 「環境資訊規章」(Environmental Information Regulations, EIR)

(4) 「法定送存圖書館法」(The Legal Deposit Libraries Act, 以下簡稱 Act 2003)

(二) 網站保存範圍

1、英國網站保存聯合會 (UK Web Archiving Consortium, UKWAC)

根據英國網站保存聯合會 (UK Web Archiving Consortium, UKWAC) 成員英國資訊系統聯合委員會(Joint Information Systems Committee, JISC)於 2008 年所出版的網站保存手冊(The Preservation of Web Resources Handbook, PoWR)中指出，凡屬於出版品、檔案、任何值得保存的網站內容會被視為藝術品/人工產物，這些珍貴的網站都應該要予以保存。反之，倘若這些資料已經在其他地方受到妥善管理或是保存價值很低者，便不該予以保存。因此，網站保存的範圍，可能包含：

(1) 一般的網站(HTML, CSS, JPG)

(2) 廣為流傳的格式(PDF, DOC, MP3, PPT)

(3) 以資料庫為主的資料 (database-driven blog, wikis, or data resources)

網站或系統已有專業人士在負責維護管理，因此不在保存範圍內者，可能包含：

(1) MIS

(2) 管理系統的資料

(3) 虛擬的學習環境

(4) 線上資產管理系統

至於如何判斷網站價值，網站保存手冊(The Preservation of Web Resources Handbook, PoWR)中提供以下幾個問題，以協助評估網站資源之價值：

(1) 員工是否需要這個資料(resources)來完成特定工作？

(2) 過去六個月內這個資料是否被存取過？

(3) 這個資料是否只有複製品，還是只有這種存取方式？

(4) 這個資料是否為組織網站刊物的一部份？

(5) 這個資料是否可以被再利用或是重新策劃？

(6) 這個資料是否可以協助稽核？

(7) 是否有法律因素以至於需要保存此資料？

(8) 就員工投入的時間與金錢而言，這個資料是否具有財務投資方面的重要意義？

(9) 這個資料是否具有歷史或文物價值？

經過上述評估後，由於需要保存之網站量仍然相當可觀，因此該手冊中建議在決定是否保存該網站時，應先瞭解該網站與其他網站服務的使用目的，還有數位內容的特性，然後考慮是否應該要保存整個網站，或者選擇性的保存部份網站、是否可依照優先順序分批予以保存，故其建議採用"莫斯科方法(MoSCoW method)來將需保存之資料進行分類。此外，該手冊認為保存方法不能一體適用到所有網站，因此建議選擇性保存，且不可能保存所有單一資源的所有版本。

英國網站保存聯合會(UK Web Archiving Consortium, UKWAC)根據「法定送存圖書館法」(The Legal Deposit Libraries Act，以下簡稱 Act 2003)中將送存義務擴充到「任何出版品」

(anywork)，而「任何出版品」則包括：傳統印刷型態之出版品(圖書、期刊、報紙、地圖、樂譜)與非印刷型態出版品

(CD-ROM、DVD、微縮資料、網路資料，但不包括電影與錄

音資料)。因此，解決保存各機關單位網站時，授權取得之問題。

2、英國運輸部公路局(Highways Agency)

根據英國運輸部公路局(Highways Agency)的網站保存政策(Web Archive Policy)，該機關僅保存該機關網站

(www.highways.gov.uk)內容，以下內容則不屬於其保存範圍：

(1) 外部網站(Externally hosted websites)

(2) 簡易資訊聚合(RSS Feeds)

(3) 資料庫(Databases)

(三) 網站保存方式

根據英國運輸部公路局(Highways Agency)的網站保存政策(Web Archive Policy)，該機關的網站保存方式共包含七個步驟：

1、內容稽核(Content Audit)

內容稽核是在判定：內容的數量、內容的類型、存在哪邊、在可以展示的網站中，網站內容要存多久。

2、風險分析(Business Risk Analysis)

風險分析是要確認哪些資料具有重要價值，因此必須要列入蒐

集範圍。例如：法律相關的內容或是對於其他機關來說很重要的資料，有些資料可能 TNA 有永久保存之需求。

3、選擇/鑑定(Selection/(Appraisal)

鑑定策略列出具有選擇、蒐集、保存之目的的內容範圍。

4、蒐集(Collection)

蒐集的方法有很多種，可依照依些主要的原則予以分類。例如：保存整個網站則稱為"snapshots" (快照)、蒐集特定部分、蒐集的頻率。

5、品質保證(Quality Assurance)

蒐集來的資料必須要確認這些資料是否採用恰當的蒐集方式，並且順利蒐集。

6、保存(Preservation)

數位保存是一個相當新而且不斷在改變的，必須要不段的確認是否可讓使用者持續的存取。

7、傳遞(Delivery)

當資訊被保存或是長期保存，必須要確立瀏覽內容的機制，這會與在各種溝通管道上瀏覽內容有關。

(四) 更新頻率

根據英國網站保存聯合會 (UK Web Archiving Consortium, UKWAC) 網站中指出，網站蒐集頻率一般是每六個月會去重新蒐集該網站，但是也會依照網站內容而改變蒐集頻率，例如有的網站即將關閉或該網站之內容不會改變，則此網站只會被保存一次。因此，網站蒐集頻率主要是取決於網站的內容，再視情況而予以調整。

(五) 詮釋資料

根據英國網站保存聯合會 (UK Web Archiving Consortium, UKWAC) 成員英國資訊系統聯合委員會(Joint Information Systems Committee, JISC)於 2008 年所出版的網站保存手冊(The Preservation of Web Resources Handbook, PoWR)中指出，網站保存之詮釋資料應有以下五類：

1、管理(Administrative)

例如：產生日期、產生部門

2、法律(Legal)

例如：著作權、數位版權

3、保存(Preservation)

例如：格式的詮釋資料、軟體的詮釋資料

4、技術 Technical

例如：格式、資料大小

5、結構 Structural

(六) 蒐集/抓取(capture/harvest)方式

英國網站保存聯合會在西元 2005 年至 2007 年間，採用澳大利亞國家圖書館 (National Library of Australia) 所提供 PANDAS 軟體來進行網站保存工作，2008 年時，大英圖書館從國際網際網路保存聯合會 (International Internet Preservation Consortium, IIPC) 引進新的網站保存工具 Web Curator Tool (簡稱 WCT)，並將原有資料轉換至新系統。

此外，根據英國網站保存聯合會成員英國資訊系統聯合委員會於 2008 年所出版的網站保存手冊(The Preservation of Web Resources Handbook, PoWR)中指出，依照抓取範圍來區分，可將蒐集或抓取網站的方式分為三種，包含：全部/區域收割(Bulk/domain harvesting)、依原則進行挑選(Criteria-based selection)、以事件為基礎的選擇方式(Event-based)。

1、全部/區域收割(Bulk/domain harvesting)：

(1) 抓取全部網站，並且/或者與其相關的範圍。此方法又稱為「非選擇性方法(unselective approach)」。

(2) 相對於耗費時間的選擇策略，這是一個省錢又快速的方法，且較不主觀的方法。但是，這種方法比較適合用在數位保存或者是想要再在某些特定範圍與界限內蒐集資料，而且若已經有大量資料，沒有必要去保存所有資料，況且產生大量可能沒有用的資料會需要額外的儲存空間。

2、依原則進行挑選(Criteria-based selection)：

(1) 此方法是指依照事先訂好的原則，挑選適合保存之檔案。此方法又稱之為「選擇性方法(selective approach)」。

(2) 事先訂好的原則例如：系所擁有的所有資料、某一類網頁資料(ex blogs)、與某一主題相關之資料(尤其和研究相關的)、影響學生的資料、網頁型式傳遞的補助計畫、可能會有遺失風險的資料、所有檔案、所有出版品、對於外部使用者很有用的資料。

(3) 這種方法並不能保證所有選到的資料可以符合組織對資訊方面的需求，所以如果先有一個試驗性的計畫，找到哪些可能資料可以保存，且其結果是可被呈現或快速傳遞的。

3、以事件為基礎的選擇方式(Event-based)：例如選舉或疾病

- (1) 任期結束(End of term)
- (2) 任期開始(Beginning of term)
- (3) 新的學年(New academic year)
- (4) 新官上任(Appointment of a new senior official)
- (5) 教授離職(Departure of a professor or senior academic)
- (6) 研究完成(Completion of a major piece of research)
- (7) 發表新章程(Publication of the new prospectus)
- (8) 採購對於網站內容有影響的新軟體
- (9) 更換名稱(Corporate or institutional rebranding)
- (10) 成立新部門(Formation of a new department)

若依照抓取方式來區分，可分為以下三種，包含：在授權的系統或伺服器內(within the authoring system or server)、)在瀏覽器內蒐集資料(Capturing at the browser)、透過網站抓耙器(crawler)蒐集資料(Harvesting content with a crawler)。

- 1、在授權的系統或伺服器上抓取(Capture within the authoring system or server)

- 優點：簡單、短期或中期內可完成
- 缺點：抓到的是原始資料，不是呈現的資料、可能太依賴授權架構或內容管理系統(Content Management System, CMS)、不適合外部存取

2、在瀏覽器上抓取(Capture at the browser)

- 優點：可以抓到你所看到的、對於控制良好的網站而這個作法相對簡單、以有現成工具
- 缺點：將網站內容是為出版品、失去網站的行為或其他特性

3、透過網站抓耙器(crawler)抓取內容(Harvest content with crawlers)

- 優點：最廣泛使用、解決一些存取的問題(link re-writing、嵌入式的外部內容)
- 缺點：會造成負責蒐集資料者更多的工作負擔，例如：負責處理此工作者必須學會如何使用網站抓耙器(crawler)、呈現抓取的資料也會遇到困難，通常發生在沒有抓到所有資料、或者抓了太多資料

(七) 與檔案管理系統之關係

目前尚無資料。

第三節 加拿大

一、電子郵件保存

本分項計畫以加拿大圖書館及檔案館(Library and Archives Canada, LAC)針對電子郵件保存之資料為主要蒐集對象外，並彙整分析亞伯達(Alberta)省檔案管理單位所制訂之電子郵件相關規範作為補充。

(一) 法源依據/資料來源

1、資料來源：

(1) Library and Archives Canada

(2) Government of Alberta

2、相關法規：Library and Archives Canada Government

(1) Access to Information and Privacy

(2) Library and Archives Canada Act

(3) Treasury Board's Management of Government Information Policy

(4) Government Security Policy

3、相關法規：Government of Alberta

(1) Alberta Evidence Act

(2) Electronic Transactions Act

(3) Freedom of Information and Protection of Privacy Act (FOIP)

(4) Government Emergency Planning Regulation

(5) Health Information Act

(6) Historical Resources Act

(7) Records Management Regulation

(二) 電子郵件保存目的

電子郵件廣泛用於管理決策、分配任務、提供進度報告、確認交易，以及許多其他與政府業務相關的用途。這種與政府業務有關的訊息須依政策及法律加以記錄並保存一段時間或永久保存。而除了少部份是以最短保存期限進行保存的資訊外，其他依據最長保存期限的資訊則必須在期滿之後，根據法律進行銷毀。有鑒於電子郵件已經成為政府機關中溝通的主要渠道，這些紀錄也必須謹慎且有效地加以管理，以確保這些電子紀錄被正確的分類及歸檔，確保這些電子檔案能在保存期限中被檢索並加以應用。

(三) 電子郵件保存範圍

政府處理一般業務過程中所產生、收集、接收，或傳送的電子郵件都是加拿大政府的檔案。這些訊息及其附件反映了政府的功能、業務活動，以及決策。為了政府整體紀錄的完整性，必須對電子郵件進行保存。

過渡性質的電子郵件則應該在完成作用後加以刪除，這些包含

轉寄郵件、垃圾郵件、休閒訊息、內部作業用版本、通知或參照用的文件、草案與註記等。以紙本進行保存維護的電子郵件也會被認為是過度性質的紀錄。換言之，若將電子郵件列印成紙本後，便不在保存範圍中，僅用於處理暫時性工作，這種類型之資料稱之為過度性質之電子郵件。

個人性質的電子郵件則不屬於加拿大政府的檔案，不屬於加拿大圖書館與檔案館法的適用範圍。例如像是利用電子郵件訂購午餐、員工的個人訊息像是興趣嗜好有關的電子郵件、業餘活動、公告、廣告等。這些電子郵件應在用途達成後加以刪除。但當機構收到根據 Access to Information or Privacy Acts 提出的正式請求，相關的紀錄就不容許使用者加以刪除。

（四）電子郵件管理政策

根據加拿大圖書館及檔案管理局(LAC)所制訂的電子郵件管理指引(Email Management Guidelines)中建議各機關單位之電子郵件管理必須達成以下幾點：

- 1、制訂管理電子郵件的計畫
- 2、收集業務相關的電子郵件
- 3、運用符合組織業務需求的分類系統組織業務相關的電子郵件

- 4、確保電子郵件採用適當的保存期程進行維護、保護，以及保存
- 5、遵守所有相關的政策及法令進行電子郵件的銷毀
- 6、定期審視電子郵件管理政策的有效性並視需要進行改進
- 7、機關內部的每個人都必須了解他在電子郵件方面的相關權責

(五) 電子郵件保存系統

目前尚無資料。

(六) 詮釋資料

目前尚無資料。

(七) 人員權責

加拿大圖書館及檔案管理局(LAC)所制訂的電子郵件管理指引(Email Management Guidelines)將機關內部人員分成五種角色：使用者、網路或系統管理人員、資訊管理專業人員、作業經理、執行長。

(八) 員工訓練

電子郵件的管理以及相關員工訓練是被設計用來確保員工都能養成足夠的知識與技巧，如此能夠更有效率以及以更適當的方式來使用電子郵件。

這裡首先說明一個對於政府員工來說最保險的策略，這個策略

的主要目的是避免組織得到過多不必要的電子資訊而產生的負擔：對於所有的文件，都把它當作是以後要給法官或是陪審團仔細閱讀的文件來描寫。

針對不同的人要設計不同的訓練課程，有些人要親自 1 對 1 教學，其他人則比較習慣在一間教室上課。有些人能夠透過閱讀以及聽講就能學習，其他人則需要親眼看到或是親自動手的實際教學。

以下是幾個訓練電子郵件管理新手的指引：

- 1、有效率以合適的方法來使用電子郵件系統
- 2、有禮貌且貼切的來使用電子郵件系統
- 3、建立並且使用密碼來保護電子郵件系統
- 4、在整個資訊流的生命週期裡面都確保電子郵件的安全性
- 5、對於無用的紀錄或是資訊，即刻刪除
- 6、根據政府的法令或規章來保存記錄
- 7、對於記錄或是資訊的訊息內容、詮釋資料與附加檔案，保存一定週期
- 8、根據政府的法令或規章來確保隱私權與機密性
- 9、對於已經達到保存期限的記錄或是資訊予以刪除

10、據政府的法令或規章來確保著作權

11、對電子郵件加以分類，並將之儲存到適合的儲存位置

12、固定一段時間就把相關的電子郵件檔案傳送給 LAC

13、確認誰該對於這份電子郵件的保存負責(寄件人、收件人，或是雙方?)

14、對於一般員工、IT 專業人員以及主管教導不同的管理電子郵件的權限

為了更有效的訓練員工，訓練課程應該要提供各種媒體介面的訓練資源以符合不同的學習風格。例如：有些人可以透過自己閱讀相關資料或是自己在線上學習就能得到不錯的訓練效果。有些人則習慣坐在教室裡面學習，其他還有些人需要在桌上寫下提示。

各個機關都應該分派電子郵件管理指標、政策與程序，如此可確保員工、系統承包商，和其他的使用者都能瞭解電子郵件管理指標。除此之外，任何錯誤的使用電子郵件系統行為，都將受到紀律嚴懲且終止政府員工雇用或是承包合約。

政府機構應該將電子郵件管理的學習活動(像是訓練課程或是演講的出席)當作紀錄保存下來。

（九）與檔案管理系統之關係

目前尚無資料。

二、 網站保存

本節資料依據加拿大圖書館及檔案管理局(LAC)針對網站保存之相關作法為主。

（一）法源依據/資料來源

1、 資料來源：

(1) 加拿大政府網站保存 (Government of Canada Web Archive)

(2) 亞伯達省網站標準(Government of Alberta Web Site Standards)

2、 相關法規：

(1) 加拿大圖書館及檔案館法 (Library and Archives of Canada Act)

(2) Treasury Board of Canada Secretariat. Common Look and Feel for the Internet

(3) Information Management and Privacy Branch, Alberta Government Services. Alberta Freedom of Information and Protection of Privacy Act (FOIP)

(4) Information Management and Privacy Branch, Alberta
Government Services. Alberta Freedom of Information and
Protection of Privacy: Guide to Developing Privacy Statements
for Government of Alberta Web Sites. May 2001.

(二) 網站保存範圍

依據加拿大圖書館及檔案館法所授與加拿大圖書館及檔案館
相關的權利來蒐集並保存對加拿大人有意義的公開存取網站，特別
是以下四種情況：

- 1、為了造福現在及後代而需要保存的加拿大文化遺產；
- 2、加拿大所服務的機構，其內容具有長期保存價值、應對大眾公
開的知識、或傳達加拿大作為自由民主社會在文化、社會，以
及經濟上的進步；
- 3、加拿大中與社群合作蒐集、保存及分享知識的組織
- 4、提供加拿大政府及其下組織相關記錄資料的組織

(三) 網站保存方式

加拿大圖書館及檔案管理局(LAC)對於所有加拿大政府機構的
網站進行網站保存。首先對這些網站執行網路爬蟲(crawl)，2004 年
到 2005 年間，已經進行了 3 次的網路爬蟲。這些網路爬蟲會對於
網站的公開存取介面執行"snapshot"並且保存起來這些"snapshot"。

值得注意的是：這些網路爬蟲的設定是---儘可能深度抓取網站，因此保存起來的 snapshot 能夠讓之後的使用者完整的存取各部門過去的網站，包含完整的內容以及內部與外部的連結。

然而，不是所有的網站資訊都能被網路爬蟲給抓取下來。舉例來說：像是架構在部門之間的內部網路(intranet)的網站，爬蟲因為無法穿透防火牆而無法抓取這種網站的資訊。此外遇到需要使用者輸入資訊的註冊點，爬蟲也無法工作。像是使用者需要付費才能觀看的網站資訊，或是需要使用者註冊才能瀏覽的網站。

(四) 更新頻率

人力與技術資源許可的情況下，加拿大圖書館與檔案館傾向每六個月抓取一次加拿大政府的網站。

(五) 詮釋資料

1、加拿大國庫委員會

加拿大國庫委員會(Treasury Board of Canada)從 CLF

6.3(Common Look and Feel for the Internet, Standard 6.3.)裡面挑

出 5 個主要的 metadata element 來描述加拿大政府的網站資源

(Web Resource)：

(1) 標題(Title)

(2) 作者(Creater)

(3) 日期(Date)

- dc.date.created：這份 web resource 首次發表在網路上的時間

- dc.date.modified：最近一次這份 web resource 修改的時間

(4) 語言(Language)

(5) 主題(Subject)

此外，加拿大國庫委員會從 TBITS 39.1(Treasury Board Information Management Standard) 裡面，採用 DC (Dublin Core) 的詮釋資料標準，在 16 個 Dublin Core 標準中提到的 metadata element 中，跟 CLF 相同 5 個 elements，為主要的。其他 11 個則是可選擇的。以下介紹 5 個可選擇的 DC element：

(1) 對象(Audience): 對於這份 web resource 有興趣或是感到有用的人

(2) 覆蓋範圍(Coverage): 這份 web resource 所涵蓋的範圍或是廣度

(3) 描述(Description)

(4) 檔案格式(Format)：The physical or digital manifestation(顯示) of the resource.

(5) 類型(Type)：這份 web resource 的內容的類型

因為大多數的搜尋引擎不能夠結構化的去閱讀，而且使用被加拿大政府採用的 DC(Dublin Core)詮釋資料元件(metadata elements)。除了主要的 DC 詮釋資料元件，最好再加上兩個 HTML 的 metatag。

(1) 描述(Description)：跟 DC 裡面的描述指不一樣的東西。這裡的描述是指對於這份 web resource 的一個簡短的簡述內容，就像某些搜尋引擎會將這個簡短的描述內容，顯示在搜尋結果。

(2) 關鍵字(Keyword)

2、亞伯達省

根據 Government of Alberta Web Site Standards。所有加拿大 Alberta 省的政府單位在進行網站保存的時候，有以下四個詮釋資料：

(1) 標題(Title)

(2) 描述(Description)

(3) 關鍵字(Keywords)

(4) 語言(Languages)

(六) 蒐集/抓取(capture/harvest)方式

加拿大圖書館與檔案館透過三種開放原始碼軟體來達成保存

加拿大的網站：

1、Heritrix v1.12.1，網路抓爬器。

2、NutchWAX v0.10.0，資料庫暨存放庫軟體

3、WAYBACK v0.8.0，網路呈現工具

第四節 澳洲

一、電子郵件保存

本章節主要以澳洲政府針對電子郵件保存之資料為主要蒐集對象。

（一）法源依據/資料來源

透過澳洲政府系統建立的電子郵件必須遵守法規 Archives Act 1983，同時也跟 Freedom of Information Act 1982、Privacy Act 1988 及 Evidence Act 1995 相關。

根據 Managing Electronic Messages as Records (May 1997)一文，澳洲政府正式將電子郵件視為記錄的一種。

澳洲法庭已將電子通訊看作一個合法的證據來源，因此，電子郵件也被視為一個合法的程序，例如：可以當作證據用來發送傳票。

1、資料來源：

(1) 澳洲檔案管理局

(2) 亞伯達省(Government of Alberta)

2、法源依據：

(1) Archives Act 1983

(2) Freedom of Information Act 1982

(3) Privacy Act 1988

(4) Evidence Act 1995

(二) 電子郵件保存目的

澳洲政府在辦理行政業務時，廣泛的使用到電子郵件。因此，這些電子郵件記錄了很多政府的行為，也因為這些電子郵件具有證明性、保護個人或政府權益之功能、文件化業務決定過程、及歷史和文化的價值，澳洲政府認為這些電子郵件是重要的資訊資產。這些與政府機關進行溝通之用的電子郵件，應妥善控管相關作業規範與資訊管理制度，才能保護所有民眾之利益。

為妥善保存電子郵件，澳洲政府根據「保護性安全手冊 (Australian Government Protective Security Manual)」制訂一套為電子郵件所定的標準化安全詮釋資料。

(三) 電子郵件保存範圍

電子郵件主要可分為公務檔案與非公務檔案兩類：

1、公務檔案的範圍包含：

(1) 政策與指令

(2) 最初的草稿或報告及建議

(3) 公務相關的回應及備註

(4) 工作排程及分配

(5) 會議議程及記錄

(6) 業務過程相關文件

(7) 判例訊息

2、非公務檔案的範圍應包含：

(1) 私人信件：與國家機構不相關檔案即為私人信件，即可被銷毀。

(2) 垃圾信件

(3) 無關信件：與其它業務相關或完全無關之信件，需與私人信件一樣，在合理的時間內刪除。

（四）電子郵件管理政策

1、電子郵件管理可依機關屬性不同適度調整

電子郵件必須透過機構的政策以及程序來管理，並且需要有訓練有素的員工來配合。機構之間可以擁有針對每個機構所需不同，而運用不同的政策來管理郵件。

2、電子郵件保存與銷毀

一封電子郵件銷毀與否以及如何銷毀都依照檔案法

(Archives Act)，以下共有三種處理方式，分別為：

(1) Record Authorities

(2) General Disposal Authorities

(3) Normal Administrative Practice

經過條件篩選後，相關人員可從三個方式中選擇出最好的

銷毀方式(Tools)。至於如何選擇何種保存或銷毀較為恰

當，則是根據圖 3-5 之流程進行判斷。

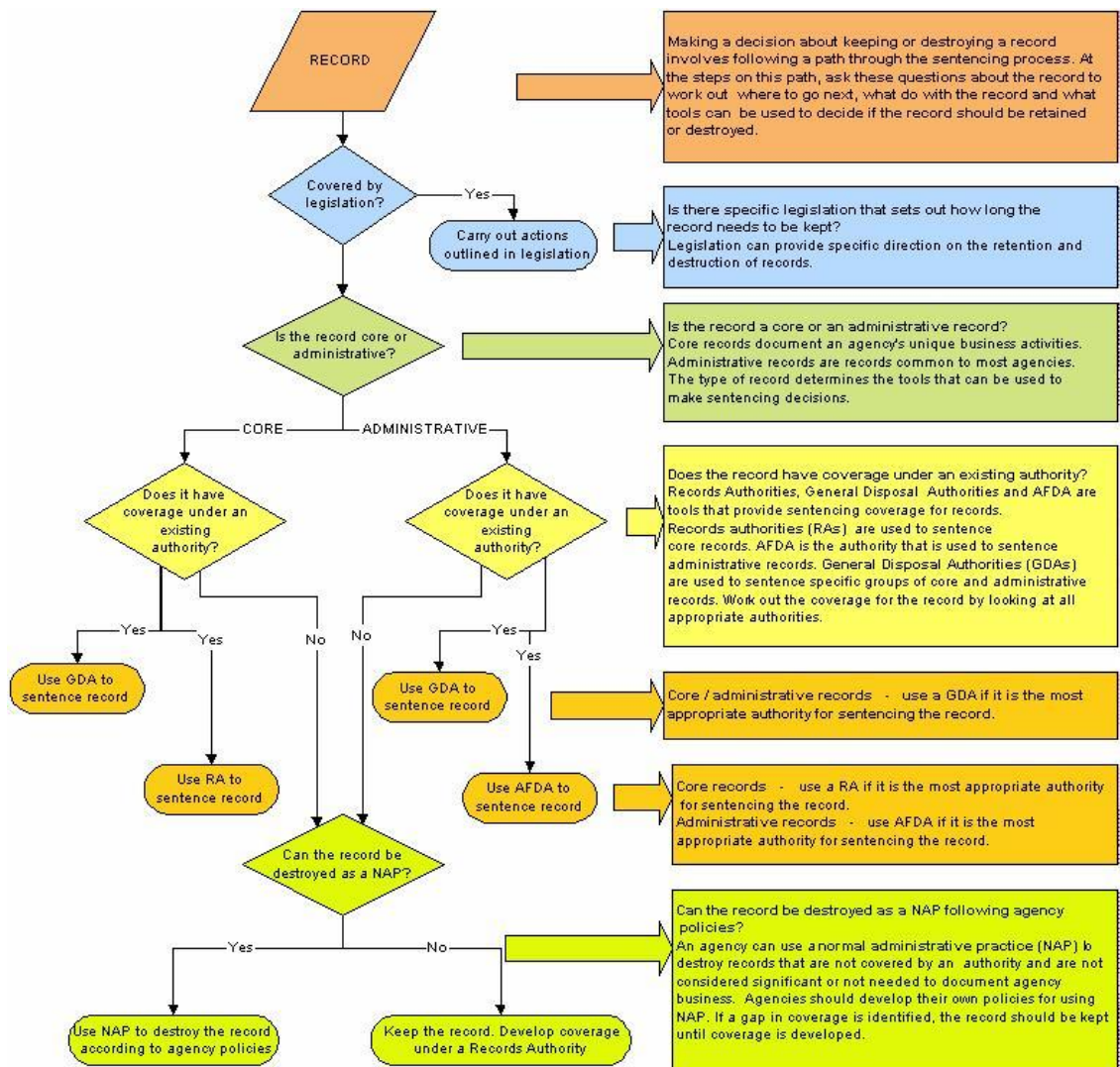


圖 3-5 保存與銷毀所選擇哪種方式之流程圖

3、電子郵件鑑定

為協助相關人員判斷哪些電子郵件與現在進行的業務有關，澳洲檔案管理局提供以下四個問題協助相關人員進行判斷：

- (1) 現在收發的郵件是否跟所負責的工作有關
- (2) 現在收發的郵件是否跟所負責的專案有關
- (3) 現在收發的郵件是否批准某個決定(決定導向)
- (4) 現在收發的郵件是否提供建議或是告知方向

(五) 電子郵件保存系統

目前尚無資料。

(六) 詮釋資料

澳洲對於電子郵件所使用的詮釋資料是採用 2005 年所定義的 Australia Government Email Metadata Standard(AGEMS)標準，有以下 9 個詮釋資料：

- 1、代理人(Agent)
- 2、權限管理(Right Management)
- 3、標題(Title)/主題(Subject)

4、關鍵字(Keyword)

5、關係(Relationship)

6、日期(Date)/時間(Time)

7、訊息編號(Message Identifier)

8、訊息順序(Message Precedence)

9、訊息重要性(Message Importance)

(七) 人員權責

目前尚無資料。

(八) 員工訓練

目前尚無資料。

(九) 與檔案管理系統之關係

如果電子郵件是跟機構現在正進行的業務有關，就必須被儲存在該機構的檔案管理系統內(依照電子檔案或是紙本方式)，而保存以及銷毀可以依照銷毀當局或是依照該機構的(NAP)政策。一旦電子郵件被當成電子檔案儲存到檔案管理系統前，必須符合真實性(Authentic)及可靠性(Reliable)，也就是說該電子郵件不管多久都是有用的(Usable)以及可讀取的(Access)，且電子郵件需受保護，且不

接受非授權的讀取(Access)、取代(Alternation)及刪除(Deletion)。

二、 網站保存

本節資料依據澳洲國家圖書館(National Library of Australia)針對網站保存之相關作法為主。

(一) 相關法規/資料來源

1、 資料來源

(1) 澳洲檔案管理局

2、 相關法規：

- Archives Act
- The Selection Committee for Australian Online Publications

(二) 網站保存範圍

1、 潘朵拉計畫(PANDORA)

潘朵拉計畫(PANDORA)發展的過程，係於 1996 年 1 月成立「澳洲線上出版品選擇委員會」(The Selection Committee for Australian Online Publications)發展線上出版品的選擇指引，同年 4 月圖書館的數位保存部門(Digital Archiving Section)開始與出版機構溝通授權，國家圖書館同時進行兩項作業：一是發展網頁保存的政策並開始嘗試保存；另一是發展下載及儲存的

技術。1998 年當技術及政策愈臻成熟，國家圖書館開始邀請其他圖書館參與合作，維多利亞省立圖書館是自願參加潘朵拉計畫的第一個合作機構，此後陸續有八個機構加入，總計十個合作參與機構，這些參與者分工合作蒐集不同重點的網路資源並建立目錄，由國家圖書館負責軟硬體的建置與維護。各館蒐集重點簡述如下：

- (1) 澳洲國家圖書館：內容具有澳洲象徵意義的網站與線上出版品。
- (2) 澳洲原住民及托勒斯海峽島民研究中心（Australian Institute of Aboriginal and Torres Strait Islander Studies）：有關澳洲原住民研究的網站與線上出版品。
- (3) 澳洲戰爭紀念館（Australian War Memorial）：有關澳洲軍事歷史的主題網站與線上出版品。
- (4) 國家影片及錄音資料館（National Film and Sound Archive）：有關澳洲電影、電視、廣播及聲音紀錄產業發展的網站及線上出版品。
- (5) 北領地區圖書館（Northern Territory Library）：內容有關北領地區的網站與線上出版品，特別是對於社會、政治、文化及科學發展上具重大意義者。

(6) 昆士蘭省立圖書館 (State Library of Queensland)：內容有關昆士蘭地區的網頁，以及昆士蘭省政府較具價值且以線上形式存在的政府出版品。

(7) 新南威爾斯省立圖書館 (State Library of New South Wales)：內容有關新南威爾斯省網頁與線上出版品。

(8) 南澳大利亞省立圖書館 (State Library of South Australia)：內容有關南澳大利亞的網頁與線上出版品。

(9) 維多利亞省立圖書館 (State Library of Victoria)：內容有關維多利亞省的網頁與線上出版品。

2、西澳大利亞省立圖書館 (State Library of Western Australia)

澳洲國家圖書館根據 2005 年修訂的「線上出版品徵集指引」明列對於線上出版品收錄的一般考量、優先收錄與明確不收錄的類型。對於線上出版品選擇的一般考量是：

(1) 有關澳洲的出版品應優先收藏，包括：內容介紹澳洲，有關澳洲的社會、政治、文化、教育、宗教、科學或經濟發展的內容，或是澳洲權威性作者的創作。

(2) 當同時存在紙本與線上形式，除非線上有更多的內容，否則因考量資訊技術的更迭，應優先收紙本；如果是線上與實體

的電子資源（光碟或磁碟）比較，因線上版長久保存的問題比較容易解決，國家圖書館應收錄線上版。

(3) 由國家支援經營的網站，具有影響國家決策的內容。

(4) 具有電子資源創新性的出版範例，可提供未來回顧技術里程碑的價值。

根據上述一般原則考量，羅列應予優先收錄的線上出版品有下列六大類：

(1) 澳洲聯邦及中央政府出版品（地方政府部分由地方政府優先保存）：對於政府機構網站定期擷取內容更新收錄，因而對線上政府出版品也儘可能完整收錄。

(2) 有關高等教育機構的出版品：大學院校等高等教育機構的學術出版品應建立機構保存，國家圖書館基於保存具學術價值文獻的考慮將這些機構保存網頁以及開放取用的資源加以蒐藏。

(3) 主題網頁，同時採取兩個作法搜集資源，一是以三年為週期，每年針對特定主題重點選取收集網路資源，每三年再輪回一次。第一年的重點主題是政治、醫藥、教育、技術與工業、科學；第二年的重點是運動與休閒、社會、環境與資源、

藝術、人文與媒體；第三年是法律、商業經濟、民眾與文化、歷史。另一個方式是不定時針對社會重大議題保存網路資源，如雪梨奧運。

同時，在此徵集指引內容中，也明列確定不收錄的線上資源，包括：

- (1) 為宣傳上載影像的網站攝影機（Web Camera）。
- (2) 論壇、聊天室或布告欄性質的網頁。
- (3) 遊戲網站。
- (4) 個人的文章和網頁。
- (5) 已有紙本發行的線上版新聞網頁。
- (6) 私人機構的紀錄。
- (7) 促銷和宣傳的網站。
- (8) 取材自其他網站內容而無自創性內容的網站。

（三）網站保存方式

網頁保存方式有四種作法：

- 1、選擇性保存（Selective Approach）：依需求選擇設定特定的擷取對象，如：加拿大、丹麥、日本及澳洲等國家均屬之。

2、定期擷取 (Periodic Harvesting)：將境內所有網域的資源定期擷取並定期更新，如：瑞典、法國、冰島、挪威等國家均採此方式。

3、主題蒐藏 (Thematic Collecting)：針對特定主題蒐集相關的網頁內容，如美國國會圖書館蒐集 2000 至 2002 年的選舉資訊網站。

4、寄存保存 (Deposit Collections)：針對出版機構自願送存的出版品建立網站保存，如荷蘭國家圖書館成立 e-Depot 收錄來自於出版者送存的電子資源。

(四) 更新頻率

澳洲政府採取的策略是：當網站改變的時候，就去抓取。首先，澳洲政府對於網站的改變，劃分為以下四種不同的情況：

1. 頻繁的(不到 3 個月就更新一次)
2. 不頻繁的(超過 3 個月才更新一次)
3. 定期的(根據某個規劃的時間表)
4. 非定期的(因為需要改變網站才改變)

澳洲政府會將網站更新分成以下四類：

1. 定期且頻繁的更新網站

2. 定期但是不頻繁的更新網站

3. 非定期但是頻繁的更新網站

4. 非定期且不頻繁的更新網站

(五) 詮釋資料

根據 AGLS (Australian Government Locator Service) and
RMSCA (Recordkeeping Metadata Standard for Commonwealth
Agencies)兩個標準。澳洲政府制訂了網站保存的詮釋資料：

1、作者或創造者(必要)

2、出版者(必要)

3、其他貢獻者

4、權限管理

5、標題(必要)

6、主題(必要)

7、描述

8、資料來源

9、語言

10、關係

- 11、 涵蓋範圍
- 12、 功能描述(必要)
- 13、 日期(必要)
- 14、 資料類型
- 15、 格式
- 16、 資料辨識碼(必要)
- 17、 可否存取
- 18、 聽眾
- 19、 命令(Mandate)
- 20、 抓耙檔案(Harvest file)

(六) 蒐集/抓取(capture/harvest)方式

澳洲政府目前使用 HTTrack 抓耙器。而目前正在使用的 PANDAS 系統，原本是為了滿足各式各樣的抓耙需求而設計的，而非大範圍的抓耙網站。目前，開發者正在重新設計 PANDAS 系統，未來將以 IIPC 的工具作為主要抓耙工具（例如：Heretrix、WERA）。

澳洲的潘朵拉計畫雖然是以選擇性擷取方式開始，但運作多年後，澳洲國家圖書館於 2005 年 7 月至 2006 年 9 月試行擷取網址.au

的澳洲網域所有的網頁資源，嘗試綜合選擇性與全網域定期擷取的作法。

可分為兩種：目標驅動(Object-driven)及事件趨動(Event-driven)。

1、目標驅動：

- 靜態網站，如：純 HTML 網站。
- 全部的 URI 都必須要能對應的其目標
- 對於每個目標要如何跟 URL 建立關連，要有詮釋資料說明
- 要保存完整的 URI、資料目標、資料目標的 MIME 類型、連結開始與結束的時間、此份資源的授權或發行與其他資料的可能關連。
- 優點是能減少保存的負擔，不用抓取整個網站
- 抓取所有東西，包含各種描述語言以及額外插件
- 最好能夠跟追蹤記錄一起應用

2、事件驅動：

- 動態網站，如：互動式網站。

- 要保存日期與時間、IP 位置或是網域名稱、瀏覽過的網頁、執行的動作、使用的網頁瀏覽器

(七) 與檔案管理系統之關係

目前尚無資料。

第五節 我國

一、電子郵件保存

因我國尚未將電子郵件納入檔案管理的範圍，各政府機關並未正視電子郵件的保存與管理機制，大多僅提供電子郵件的備份功能。為深入了解目前市場上有關電子郵件保存系統的發展現況，本分項計畫特別參訪國內發展電子郵件系統具代表性的廠商－網擎資訊(Openfind Information Technology, Inc.)。

網擎資訊於 1998 年成立至今，持續致力於電子郵件系統之發展，軟體產品包含：企業搜尋引擎、電子郵件系統、郵件防護系統、郵件歸檔管理與稽核加密系統，到線上服務之電子信箱服務與企業郵件代管。其中，以「MailBase 郵件備份管理系統」與電子郵件保存系統最為相似。

以下為參訪時所討論的重點。

(一) 電子郵件與法規

美國於 2002 年 7 月簽署「沙賓法案」(Sarbanes-Oxley Act of 2002)，規範企業有責任將所有的電子檔案以及電子郵件紀錄，備份保存至少 5 年。同時，根據美國 HIPPA 規定，無論是電子郵件或檔案，包括與公司往來合約、與醫療政策及流程相關的文件、與病人溝通的紀錄表格或是消費者抱怨等紀錄，都需保存至少 6

年，所有與病人相關文件最少需保留至病人過世後 2 年。日本也制訂「個人情報保護法」與「日本版沙賓法案」，以規範資料的保存，並明訂相關刑責與罰則。

（二）電子郵件成長量

根據法規中明定必須保存之電子郵件範圍，並參考網擎客戶資料量，網擎估計若開始保存電子郵件時，第一年約有 0.73TB，平均每年成長率為 1.1 倍。

（三）電子郵件生命週期

電子郵件存放之儲存媒體可依照存取速度分為在線(Online)、近線(Near-line)、離線(Off-line)。例如：存在硬碟中，可即時查詢者，屬於在線，存在磁帶等較慢速之儲存媒體者，屬於近線；存在光碟者，屬於離線。

（四）電子郵件調閱方式

依據組織階層，分別控管各層級人員對電子郵件之調閱權限。

（五）電子郵件標準

RFC 822 (MIME) (.eml) 是目前大多數電子郵件管理系統所依循之電子郵電格式。此外，過去曾有微軟 Exchange 的另一種格式，稱為：TNEF(.pst 或.msg)，但後來在 2009 年三四月時，已經將其

格式公開，因此現在市面上電子郵件格式可將微軟所採用之格式轉成 MIME 格式。

（六）市場現況

目前電子郵件伺服器(Mail Server)以 Exchange 和 Notes 為主，這兩套系統是在微軟的作業平台上運作，另外約有 10%的系統是在 Linux 系統上運作。

（七）網擎系統

採用 C 語言作為系統開發之程式語言，配合自定檔案格式。我國 23 個縣市中，約有 12 個是網擎的客戶。

（八）問題討論

1、機關電子郵件保存所會面臨之挑戰

（1）機關保存電子郵件之意願低

根據網擎提供各機關電子郵件管理服務之經驗，會有許多機關不願意接受電子郵件被備份起來，而且實際上若要求機關保存電子郵件，可能會面臨的問題是，各機關本來就沒有保存電子郵件的習慣。

（2）機關不瞭解保存電子郵件之目的

各機關對於電子郵件保存與備份的觀念並不清楚，多半會認為電子郵件保存是當意外事件發生時，復原電子郵件資料相關。例如：有些機關的使用者可能電腦損壞時，才會有需要從保存之電子郵件中，把資料還原回來以供使用。

2、與檔案管理系統整併之挑戰

(1) 詮釋資料標準不明確

若各機關要將電子郵件歸檔，便須將電子郵件匯入檔案管理系統當中，但是由於目前電子公文檔案之詮釋資料與電子郵件可提供之詮釋資料有差異，因此未來將需要訂定一詮釋資料標準，便可使電子郵件順利轉入檔案管理系統當中。

(2) 多元性電子郵件系統

一般而言，各家電子郵件系統平台不同，需要針對各種軟體進行客製化，因此難與電子檔案管理系統整合。此外，有的機關會把工作流程和電子郵件綁在一起，例如 Notes，更增添其複雜性。

3、電子郵件保存技術相關問題

(1) 電子郵件保存之資訊安全防護問題

由於保存電子郵件時，可能會需要檢查保存前後是否有遭竄改，以免未來電子郵件歸檔後，若需作為證據性造成法律上的困擾。因此，保存電子郵件時是否需要加密，仍待探討。

(2) 電子郵件編碼問題

電子郵件編碼方式不同，會影響開啟電子郵件後呈現之問題。因此，若要保存電子郵件時，需要探討如何解決各種不同編碼方式所引起的問題。

二、 網站保存

雖然網站已是我國政府機關提供資訊的重要管道之一，但目前各機關本身鮮少有系統化的進行網站保存，更未將它納入檔案管理的範疇。為了解國內針對網站保存的發展現況，本分項計畫參訪了國內進行網站保存兩所重要機構：台灣大學及國家圖書館。台灣大學於 2006 年建置台灣大學網站保存庫(NTU Web Archiving System)，可算是國內的先驅。此外，國家圖書館也於 2007 年建置台灣網站保存系統(Web Archive Taiwan)。因此，以下將分別介紹由台灣大學所開發之「台灣大學網站保存系統」以及由國家圖書館所建置之「臺灣網站保存系統」。

(一) 法源依據/資料來源

- 1、台灣大學－網站保存庫(NTU Web Archiving System)
- 2、國家圖書館－台灣網站保存系統(Web Archive Taiwan)

(二) 網站保存範圍

1、台灣大學網站保存庫

台灣大學網站保存庫評選網站的考量因素為：以與台灣大學相關者為主，進而擴大到台灣相關資料，並規畫富有地方性及時代價值的特色網站達到保存臺灣記錄的計畫目標。因此，該計畫以達成保存機構、保存臺灣和保存時代的三大目標，規畫收錄內容及保存範圍如下：

- (1) 台大相關網站(完整收藏)：臺大首頁、各行政學術單位、合作網站、計畫網站、個人網站等。
- (2) 政府機構網站(完整收藏)：中央各部會、各級地方政府所屬單位之合作網站及個別網站。
- (3) 教育學術網站：大專院校及學術單位之網站、合作網站(結盟網站)、計畫網站及個人網站等。
- (4) 藝文相關網站：文學、美術、古物、民間風俗、民間藝術等及相關保存單位(如博物館、美術館)網站。各教學資源及史

料檔案等相關網站入此。

(5) 族群相關網站：原住民、客家、臺灣各宗親會網站；婦女團體、少數民族、弱勢族群入此。

(6) 非營利社團網站：職業公會及工會、慈善團體、宗教團體(含臺灣傳統民間信仰及民俗團體)、其他非營利社團等。

2、國家圖書館台灣網站保存系統

依據〈圖書館法〉及其他相關法規的規定，國家圖書館具有徵集本國出版圖書資訊，並予以整理編目保存的使命。然網站資源成長快速，且內容良莠不一，考量永久保存及相關人力物力的投入，本系統對網站之收錄範圍進行審慎諮詢，將具時代及歷史文化意義之優質網站列為優先收錄對象，主要如下：

(1) 各級政府網站：中央及地方各級政府機關網站。

(2) 教育及學術機構網站：研究機構、社會教育機構及全國大專院校及所屬系所網站。

(3) 圖書館網站：各大專院校圖書館、專門圖書館、公共圖書館、圖書資訊學研究單位及圖書館團體網站。

(4) 國家圖書館網站：國家圖書館及附屬單位網站。

(5) 重大社會議題網站：以政治、經濟、文化等各領域社會重大

議題為導向，保存與該議題論述相關網站，目前本系統保存環境保護、教育改革、媒體改革、金融改革、國土規劃、全民反恐、公民投票及 2008 總統大選等議題網站，日後將視時空環境之變化，陸續保存其他議題網站。

(6) 台灣網站保存系統所能擷取的網站範圍包括：

- 一般靜態網頁
- 附加檔案
- 影音檔案

(7) 無法保存之網站：

- 需要輸入帳號密碼或資料庫性質等網站內容
- 部份網站若以目前擷取技術仍無法完整擷取者(如：flash、JavaScript)

(三) 網站保存方式

1、台灣大學網站保存庫

台灣大學網站保存系統前端功能主要包含：

- (1) 主題探索：展示部分重要主題供使用者瀏覽
- (2) 保存時代：展示部分重要事件供使用者瀏覽

(3) 保存縮影：以畫面輪播方式，展示部分重要主題給使用者瀏覽

(4) 時光迴廊：以畫面輪播方式，展示某網站在不同時期的變化

(5) 保存類目：將各網站依照台灣大學所制訂的類目歸納整理，供使用者瀏覽。

(6) 時光拉霸機：提供使用者依照時間順序瀏覽各種已保存之網站

(7) 地理瀏覽：提供使用者依照地理位置瀏覽各種已保存之網站

(8) 熱門網站：以跑馬燈方式，提供較熱門或最新保存之網站

(9) 關鍵字查詢：提供使用者輸入關鍵字檢所網站，或全文檢索

(10) 使用者推薦：使用者可以推薦值得保存之網站

2、國家圖書館台灣網站保存系統

國家圖書館台灣網站保存系統前端功能共包含八個部份，分別為：

(1) 保存查詢：提供使用者關鍵字、網址等查詢。

(2) 分類瀏覽：將所有保存之網站依照主題、重大議題、機構、地區等方式分類，供使用者瀏覽。

- (3) 保存網站：以輪播方式提供使用者最新與最熱門的網站。
- (4) 熱門查詢：提供使用者熱門查詢之網站。
- (5) 推薦議題：提供使用者瀏覽重大議題相關之網站。
- (6) 保存內容精選：以縮圖方式呈現保存之特殊或重要之網站。
- (7) 保存資訊：提供使用者本保存系統保存資訊之數據，包含以
保存之網站版本數、網站數、檔案數。
- (8) 保存檔案：統計出本保存系統保存網站時所保存之檔案類型
與數量
- (9) 使用者推薦：使用者可在網站中推薦值得保存之網站。

(四) 更新頻率

1、台灣大學網站保存庫

台灣大學網站保存庫一般是每年會重新蒐集該網站，有時也會依照網站內容而改變蒐集頻率，例如某些網站即將關閉、網站內容異動頻繁等。

2、國家圖書館台灣網站保存系統

國家圖書館每年會先決定可能蒐錄保存的網站清單，再輔以使用者的推薦網站。由於出版社相關之網站多屬私人，因此目前

未列入保存範圍。網站保存深度以可完整保存該網站為準，而抓取頻率則依照網站的屬性而選定，並可事先設定抓取排程。

（五）詮釋資料

1、台灣大學網站保存庫

台灣大學網站保存庫的詮釋資料是由負責處理網站保存的人員將資料著錄以便使用者查詢、索引。目前並未採用國際的詮釋資料標準，僅填寫必要的詮釋資料，例如：網站名稱、URL、出版者、關鍵字。

2、國家圖書館台灣網站保存系統

國家圖書館保存網站後，採用的詮釋資料以 Dublin core 為主，包含：題名、主網址、英文題名、出版者、主題、機構、地區、關鍵詞、簡述、語言、資源類型、開放範圍、保存起迄日期、保存版本數、保存檔案數、保存容量、瀏覽次數。此外，所有網站的詮釋資料，僅在抓取第一個版本時由人工方式填寫，往後若網站改版後所保存之網站詮釋資料均由系統自動產生。換言之，每一個網站只需由人工輸入一次詮釋資料。

（六）蒐集/抓取(capture/harvest)方式

1、台灣大學網站保存庫

台大圖書館自行開發一套程式讓館員產生系統自動抓取網站的清單；接著，將此清單交由 HTTrack 去抓取網站；最後，會將抓取下來的資料夾存入 Web site Repository。請參考圖 3-6。

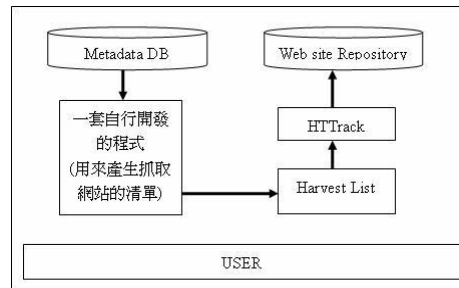


圖 3-6 網站保存系統

2、國家圖書館台灣網站保存系統

國家圖書館的網站保存則是由軟體廠商協助開發，該系統採用 JAVA 語言，SQL Server，Windows+ Linux 系統，並透過 HTTrack 抓取網站，因為它提供完整的 API。

（七）與檔案管理系統之關係

1、台灣大學網站保存庫

目前尚未列入館藏或與檔案管理系統整併，若未來國家圖書館將詮釋資料明訂後，才可能會有此打算。

2、國家圖書館台灣網站保存系統

由於網站保存在編目時仍有許多問題尚待討論，因此還未將其納入館藏或與檔案管理系統整併。

第四章 作業指引與雛型系統

根據國內外現況彙整與文獻探討，本分項計畫研提「電子郵件保存指引」、「網站保存指引」，並規劃與實作「電子郵件保存之雛型系統」、「網站保存之雛型系統」。因此，本章將分別說明電子郵件與網站保存相對應之作業指引、雛型系統、相關建議，藉此瞭解保存電子郵件與網站可能面臨的困難，以及解決之道。

第一節 電子郵件保存

一、 作業指引

（一）前言

電子郵件已逐漸成為機關人員處理公務的主要溝通工具之一，且由於機關人員利用電子郵件進行溝通的過程中會產生具有證明性、歷史性的價值，使得電子郵件在未來檔案事業發展將扮演舉足輕重的角色。目前，隨著電子郵件使用量日益漸增，若不及早重視電子郵件保存，將可能失去珍貴檔案資料。

（二）適用範圍

電子郵件保存所涵蓋範圍(對象)主要是針對機關處理公務或與公務相關，且具保存價值之電子郵件。因此，機關所提供之電子郵件信箱中，符合下列原則之電子郵件均屬之：

- 1、機關人員於處理公務中所產生與公務相關之電子郵件。
- 2、機關人員於處理公務中所接收與公務相關之電子郵件。
- 3、非屬私人性質工作之電子郵件。
- 4、機關為處理特定公務所提供信箱中之電子郵件。

此外，機關得視資訊安全或公務需求調整電子郵件保存範圍。

（三）相關法令

- 1、文書及檔案管理電腦化作業規範
- 2、個人資料保護法

（四）名詞定義

1、電子郵件

是指通過電子通訊系統進行書寫、發送和接收的信件。

2、保存

是指將蒐集來之檔案加以處理，加上詮釋資料的描述，以檔案的形式儲存。

（五）人員權責

1、機關人員

擁有機關所提供電子郵件信箱之人員應注意以下事項：

- (1) 以機關提供之電子郵件帳號所收發的電子郵件均可被視為公務檔案，並遵循與紙本、實體檔案相同法令規範。
- (2) 視為公務檔案之電子郵件均須匯入檔案管理系統歸檔管理。
- (3) 將公務相關之電子郵件歸檔前，機關人員應先行查檢擬歸檔之電子郵件是否辦畢及填列分類號與保存年限，未填列者，應由機關人員補填。
- (4) 屬於機關檔案保存範圍之寄收電子郵件，尚未匯出檔案管理系統前不應刪除。

此外，各單位承辦人員應先行檢查擬歸檔之電子郵件是否辦畢及填列分類號與保存年限。

2、檔案管理人員

- (1) 得負責著錄電子郵件匯入檔案管理系統時，所需之詮釋資料。
- (2) 若檔案管理人員發現檔案分類號錯誤者，應查明更正之。
- (3) 制定電子郵件保存規則及電子郵件命名原則。

3、資訊人員

4、負責安裝與維護電子郵件系統，並提供保存功能。

（六）處理程序

各機關進行電子郵件保存時，包含兩個階段：電子郵件蒐集與電子郵件保存，作業流程圖如圖 4- 1 所示。各階段作業內容說明如下：

1、電子郵件蒐集階段

(1) 承辦人員辦畢電子郵件時，應判斷該電子郵件是否與公務相關，如：會議通知、公務事項討論、其他與履行業務相關的資訊等，若屬於私人信件者，便不須予以保存。若此信件為處理特定公務所提供信箱中之電子郵件，如首長信箱，則須予以保存。

(2) 若電子郵件與公務相關，則判斷此電子郵件是否具有保存價值，如：處理相關業務時所需之參考資料、保存歷史軌跡所需之參考資料、法律佐證之參考資料等參考價值，若不具參考價值者，便不予以保存。

(3) 若具有保存價值者，則判斷此電子郵件是否與其他電子郵件具有關連，如：開會通知之往返信件、事項討論之通訊紀錄與討論內容等。

- (4) 若存在相關之電子郵件時，則依照併案方式辦理，並且註記相關電子郵件之關聯性，如多封電子郵件之共同主題、參與人事等，將所有相關電子郵件一併設定單一適當之保存年限；若無相關之電子郵件，則直接針對該電子郵件設定期適當之保存年限。

2、電子郵件保存階段

- (1) 承辦人員得採書面或電子方式產生電子郵件歸檔清單，並應記載下列事項：

- 文(編)號。
- 主旨或事由。
- 承辦單位及人員。
- 媒體類型及數量。
- 附件媒體類型及數量。
- 附件抽存續辦者，其名稱及應辦畢日期。

- (2) 承辦人員可透過電子郵件系統中的保存功能將電子郵件相關詮釋資料著錄，並且匯出至檔案管理系統。若以案件層級歸檔者，應著錄以下項目；

- 案由項：包括案由、並列案由及其他案由。
- 發（來）文者項：包括主要發文者、次要發文者、主要來文者、次要來文者及受文者。
- 文件形式項：包括文別及本別。
- 密等及保存年限項：包括機密等級、保密期限、解密條件及保存年限。
- 相關編號項：包括收文字號、發文字號、來文字號、檔號、電子或微縮編號及其他編號。
- 日期項：文件產生日期。
- 媒體型式項：包括紙本、底片、微縮片、幻燈片、照片、硬式磁碟、磁片、磁帶、光碟、錄音帶或錄影帶等檔案之媒體型式。
- 檔案外觀項：包括檔案之數量、單位、附件及外觀細節。
- 關聯項：包括可參照之有關案件及所屬案名。
- 若以案卷層級歸檔者，應著錄以下項目；
- 案名項：包括案名、並列案名及其他案名。
- 檔案產生者項：包括檔案產生及管有機關（構）、團體或

個人。

- 密等及保存年限項：包括機密等級及保存年限。
- 檔案應用項：包括應用限制、應用註記及複製限制。
- 相關編號項：包括檔號之年度號、分類號、案次號及其他編號。
- 日期項：案卷內文件起迄日期。
- 媒體型式項：包括紙本、底片、微縮片、幻燈片、照片、硬式磁碟、磁片、磁帶、光碟、錄音帶或錄影帶等檔案之媒體型式。
- 檔案外觀項：包括檔案之數量、單位及外觀細節。
- 關聯項：包括案情摘要及可參照之有關案卷。
- 主題項：包括檔案內容涉及之人、事、時、地、物等關鍵詞彙。
- 附註項：各著錄項目需補充或解釋之必要事項。

檔案管理人員應將待點收之電子郵件匯入檔案管理系統

中，若發現檔案分類號或相關詮釋資料有缺漏或錯誤者，應查明更正之。

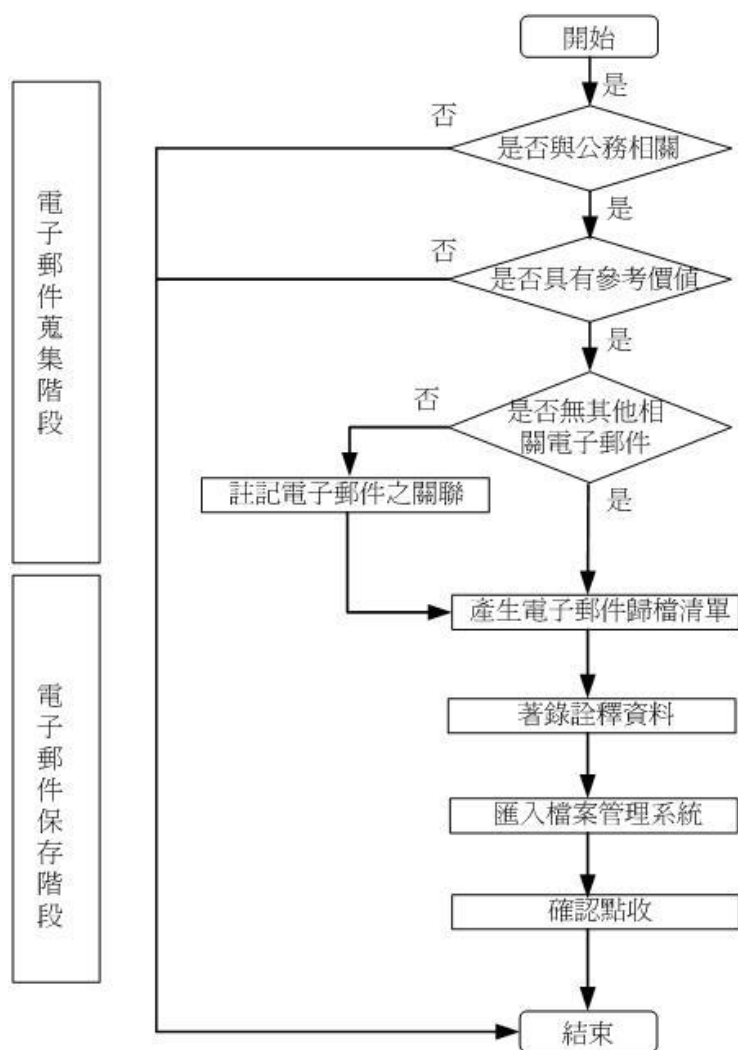


圖 4- 1 電子郵件保存作業流程

二、 雛型系統

(一) 系統示意圖

為驗證上述處理程序之可行性，本分項計畫建議在現有電子郵件系統中增加「保存」的功能，藉此將電子郵件轉換為電子檔案，進而匯入檔案管理系統中，以納入一般電子檔案管理的範疇，如圖 4- 2 所示。

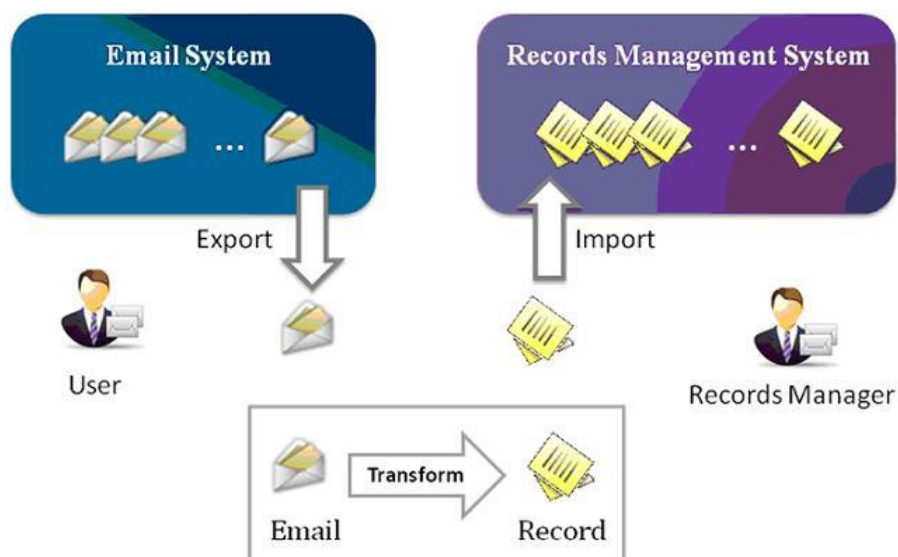


圖 4- 2 電子郵件保存示意圖

（二）系統實作

電子郵件保存系統安裝手冊，請參考附錄一；電子郵件保存系統操作手冊，請參考錄附二。

（三）詮釋資料

為將電子郵件轉換為電子檔案，首先，必須將電子郵件從電子郵件系統中匯出，且為使電子郵件可順利匯入檔案管理系統中，必須依循檔案管理系統中對於電子檔案所規範之詮釋資料。目前，檔案管理局所制定之文書及檔案管理電腦化作業規範中將電子檔案分為案件與案卷兩種層級，電子檔案若為案件型式之必要欄位如表 4- 1 所示；若為案卷型式之必要欄位如

表 4- 2 所示。然而，為了完整且詳實保存電子郵件所隱含之

資訊，本分項計畫建議增加部分詮釋資料欄位，包含：寄發郵件者、接受郵件者、電子郵件本文及其附件。為釐清電子郵件轉換為電子檔案前後，詮釋資料之對照，詳細說明請參考表 4- 3。

表 4- 1 電子檔案詮釋資料之必要欄位(案件)

資料欄位	欄位說明
系統名稱	代碼；指資料來源之系統名稱， 1:公文管理系統 2:檔案管理系統
功能	代碼；只 N 新增 M 修改 D 刪除
案由	文字，指扼要表達案件主旨或事由之文字。
文別	(公文類別,(函類別 令類別 其他類別)?)，現行檔案為必要欄位。 1 令 2 函 3 公告 4 開會通知單 5 簽 6 簽稿會核單 7 會銜公文會辦單 8 公文時效統計 9 呈 A 咨 Z 其他
本別	收發處理本別，現行檔案為必要欄位。 1 正本 2 副本 3 抄本 4 影本 5 譯本 A 稿本 B 草稿 C 定稿 D 底圖 E 藍圖
密等	代碼；1" "2 密 3 機密 4 極機密 5 絕對機密
保存年限	依機關檔案保存年限區分表所定年限，99 表示永久保存，?表示未判定，定期則以數字表示之
分類號	指公文依檔案分類表所載類目名稱之文、數字組合號碼
文件產生日期	年月日，發（來）文日期、簽署日期。
檔案類別	指原生檔案型式，代碼； 1:紙本檔案;2:電子檔案
數量	文字；總頁數、總件數、總卷數或總檔案數，不含附件，案卷層級為必要欄位。
計量單位	文字；如：頁、件、張、卷、幅、檔案或(其他)
承辦單位	單位名
承辦人員	姓名
承辦人員職稱	職稱

表 4- 2 電子檔案詮釋資料之必要欄位(案卷)

資料欄位	欄位說明
功能	代碼；只 N 新增 M 修改 D 刪除
案名	文字，扼要表達案卷內容之名稱。案卷之必要欄位，現行檔案以案件方式著錄時，亦為必要欄位。
檔案產生及管有機關	文字，描述檔案產生及管有機關、團體名稱、姓名
密等	代碼；1" "2 密 3 機密 4 極機密 5 絕對機密
保存年限	依機關檔案保存年限區分表所定年限，99 表示永久保存，?表示未

	判定，定期則以數字表示之
應用限制	Y 開放 N 不開放 R 限制開放
年度號	指該案卷(件)之年份號碼
分類號	指公文依檔案分類表所載類目名稱之文、數字組合號碼
案次號	指區分同類不同案次之號碼
案卷檔案起始日期	文字，著錄該案卷內最早之文件產生日期之年制，如聖祖康熙 23 年則為「聖祖康熙」，如為民國年則空白。
案卷檔案迄止日期	文字，著錄該案卷內最晚之文件產生日期之年制，如聖祖康熙 23 年則為「聖祖康熙」，如為民國年則空白。
數量	文字；總頁數、總件數、總卷數或總檔案數，不含附件，案卷層級為必要欄位。
計量單位	文字；如：頁、件、張、卷、幅、檔案或(其他)
分類號啟用日期	年月日，檔案分類表所載類目名稱之產生日期，如民國 88 年 01 月 15 日為 0880115

表 4- 3 電子郵件與電子檔案詮釋資料之對照表

電子檔案 詮釋資料	文書及檔案管理電腦化作業規範之欄位說明	電子郵件 詮釋資料
系統名稱	代碼；指資料來源之系統名稱， 1:公文管理系統 2: 檔案管理系統	新增一類 3 電子郵件系統
功能	代碼；只 N 新增 M 修改 D 刪除	N
案由	文字，指扼要表達案件主旨或事由之文字。	主旨
文別	(公文類別, (函類別 令類別 其他類別)?), 現行檔案為必要欄位。 1 令 2 函 3 公告 4 開會通知單 5 簽 6 簽稿會核單 7 會銜公文會辦單 8 公文時效統計 9 呈 A 咨 Z 其他	請承辦人填寫
本別	收發處理本別，現行檔案為必要欄位。 1 正本 2 副本 3 抄本 4 影本 5 譯本 A 稿本 B 草稿 C 定稿 D 底圖 E 藍圖	1
密等	代碼；1" "2 密 3 機密 4 極機密 5 絕對機密	請承辦人填寫
保存年限	依機關檔案保存年限區分表所定年限，99 表示永久保存，?表示未判定，定期則以數字表示之	請承辦人填寫
分類號	指公文依檔案分類表所載類目名稱之文、數字組合號碼	請承辦人填寫
文件產生日期	年月日，發（來）文日期、簽署日期。	寄件日期
檔案類別	指原生檔案型式，代碼； 1:紙本檔案;2:電子檔案	2
數量	文字；總頁數、總件數、總卷數或總檔案數，不含附件，案卷層級為必要欄位。	請承辦人填寫
計量單位	文字；如：頁、件、張、卷、幅、檔案或(其他)	請承辦人填寫

承辦單位	單位名	請承辦人填寫
承辦人員	姓名	請承辦人填寫
承辦人員職稱	職稱	請承辦人填寫
寄發郵件者		新增 收信者
接受郵件者		新增 寄信者
電子郵件本文 及其附件		新增 電子郵件本文

三、 相關建議

（一） 制定電子郵件保存詮釋資料

若將電子郵件保存至檔案管理系統時，需依照電子檔案之詮釋資料才可順利進入系統中，但電子檔案詮釋資料現有之欄位無法完整保存電子郵件內容，例如：寄發郵件者、接受郵件者、電子郵件本文及其附件等詮釋資料，因此建議在現有電子檔案詮釋資料中，增加屬於電子郵件相關之詮釋資料。

（二） 避免採用公務信箱處理私人事務

由於處理公務相關之電子郵件需要妥善保存，但許多承辦人除了使用公務信箱處理公務相關電子郵件，也會使用公務信箱處理私人事務，導致承辦人必須篩選出與公務相關之電子郵件，並且予以保存，如此一來，承辦人可能會誤判此電子郵件是否與公務相關，

以至於未保存公務相關之電子郵件。因此，建議處理私人事務時，請採用私人信箱，若處理公務時，請採用公務信箱，換言之，機關承辦人員應避免採用公務信箱處理私人事務。

（三）處理機密或敏感性電子郵件

在傳遞電子郵件時，可能會涉及到敏感性資料的範圍主要有：與國家安全相關的資訊、與商業機密相關的資訊、個人隱私相關的資訊。處理機密或敏感性資料時，應謹慎處理，以免誤送或遭竄改等狀況發生。因此，各機關得選定可確保資訊安全的方式，如採用加密方式傳遞電子郵件。

（四）設定電子郵件標題名稱

檔案管理系統中案由或案名是檢索檔案的重要指標，若將電子郵件的標題(subject)直接作為案由或案名，可能會無法清楚描述電子郵件被視為檔案之原因，因此建議承辦人員在命名電子郵件的標題時應審慎考量。或者，建議把電子郵件保存至檔案管理系統時，給予一個更適切的標題作為案由或案名，並且可保留電子郵件原本的標題。

（五）保存電子郵件的時機

機關處理公務相關之電子郵件來往頻繁，為確保與公務相關之

電子郵件均妥善保存，並避免遺漏重要公務相關之電子郵件。因此，建議承辦人員得採定期保存電子郵件的方式，或當電子郵件傳送或接收時，便同時進行保存。

（六）匯送電子郵件至檔案管理系統

由於處理公務相關之電子郵件也應視為電子檔案，若將此類電子郵件保存於電子郵件系統中，將不易依照電子檔案管理辦法妥善管理。因此，建議將電子郵件也保存至檔案管理系統中。若機關已建置檔案管理系統，則建議在電子郵件系統中增加匯出電子郵件成為一筆電子檔案之功能，並透過檔案管理系統中電子檔案點收之功能，將電子郵件匯入檔案管理系統，使電子郵件進入檔案管理的生命週期，以利後續管理與維護；若機關尚未建置檔案管理系統時，則建議把電子郵件放到電子郵件系統之共用資料夾(shared public email folder)，或將電子郵件保存在共用的特定系統或硬碟中。

（七）選擇電子郵件保存格式

由於不同電子郵件系統支援之電子郵件格式也不同，各機關保存電子郵件時，應採用何種格式保存電子郵件？為使各機關有可依循之保存格式，本計畫建議採用目前電子郵件格式中最主要的標準「.eml」，它也是現有大多數電子郵件系統支援的格式。此外，因應科技變化，應定期審視此保存格式是否適合。

第二節 網站保存

一、 作業指引

(一) 前言

網站已逐漸成為機關傳遞訊息或提供民眾服務的管道，且由於機關網站中所發布的資訊往往詳實的記載著機關發展的軌跡，這些珍貴資訊都具有歷史性的保存價值。因此，使得網站在未來檔案事業發展扮演舉足輕重的角色。目前，各機關紛紛建置網站，若不及早重視網站保存，將可能失去珍貴檔案資料。

(二) 適用範圍

網站保存涵蓋範圍(對象) 主要是針對機關處理公務或與公務相關，且具保存價值之網站。因此，機關所建置之網站，符合下列原則者均屬之：

- 1、機關用以處理公務相關資訊之網站。
- 2、機關用以提供民眾與公務相關資訊之網站。
- 3、與機關執掌或業務範疇相關之網站。
- 4、機關認為具有檔案保存價值之網站。

網站保存主要區分二種方式：集中式及分權式。兩種方式的主

要差異在於，集中式保存是由單一機關徵集特定範圍機關內，包含所屬及非所屬機關，與保存主題相關之檔案；而分權式由一主要機關收集所有所屬機關內與保存主題相關之檔案。因此機關需先決定網站保存採取的方式，進而定義相關機關的範圍。

（三）相關法令

1、個人資料保護法

（四）名詞定義

1、網站

是指網際網路上有特定網域，提供各種多元化資訊的網路服務。

（五）人員權責

1、檔案管理人員

(1) 負責挑選具有保存價值之網站。

(2) 負責取得網站保存授權。

(3) 得負責保存網站。

2、資訊人員：

(1) 負責安裝/建置或維護網站保存系統。

(2) 得負責保存網站。

（六）處理程序

各機關進行網站保存時，包含三個階段：網站挑選、網站授權取得、網站保存，作業流程圖如圖 4-3 所示。各階段作業內容說明如下：

1、網站選擇階段

集中式及分散式的網站保存皆包含三個階段：網站資料挑選、網站授權取得、網站資料保存，作業流程圖如圖 8-3 所示。兩種網站保存方式僅在「網站授權取得」階段有差異，因集中式會徵集到外部機關之檔案，需提出授權許可申請；而分散式主要徵集所屬機關檔案，僅需提出網站保存通知。集中式可能包含提出授權許可申請及網站保存通知兩種流程；而分散式只需提出保存通知即可。各階段作業內容說明如下：

(1) 檔案管理人員選擇網站內容時，可依照下列原則進行挑選：

- 具有證據性之網站。
- 具有再利用價值之網站。
- 具有歷史價值之網站。
- 特殊事件相關之網站，且該事件與機關執掌或業務範圍相關。

- 僅存在網站中之網站，且具有保存價值者。
- 經專家推薦之網站，且具有保存價值者。

(2) 挑選出待保存之網站後，檔案管理人員需界定網站保存範圍，至少包含以下項目：

- 網站抓耙深度
- 網站抓耙頻率

(3) 界定網站保存範圍後，得針對以下項目評估網站保存之可行性：

- 欲抓取此網站的多少內容

抓耙前先瀏覽網站確認是否僅需此網站的網頁內容，或包含外部連結頁面，以確立抓耙之範圍。

- 此網站連接多少伺服器及主機

多數提供大量多媒體及 PDF 檔案的網站包含外部連結網頁，應確認抓耙工具有能力抓取所有需要的 URL 種子(seed URLs)。

- 此網站是否被 robot 保護

檢視此網站的 robots.txt 檔案確認是否限制抓耙。

- 抓耙工具的限制

此網站是否包含工具無法抓取的部分，如資料庫及 FTP 裡的資料。

2、網站授權取得階段

(1) 檔案管理人員得確認網站的所屬單位。若非所屬機關，即採用集中式網站保存方式時，則寄出授權許可申請書，請接續下一個步驟；若單位為所屬機關，即採用集中或分散式網站保存方式時，則發出網站保存通知，請跳過下一個步驟，直接進入網站保存階段。

(2) 網站所屬機關同意「網站保存授權許可」，則進入網站保存階段；若網站所屬機關無法同意「網站保存授權許可」，則停止保存其網站，請進入網站保存階段。

3、網站保存階段

(1) 檔案管理人員或資訊人員得設定網站保存對象，包含：網站名稱、種子 URL(抓耙工具開始之處)、範圍、時間、排程選項、描述資料及網站權益資料等資訊。

(2) 檔案管理人員或資訊人員可透過網站保存系統抓耙網站。

- (3) 當網站保存系統抓耙資料完成後，檔案管理人員或資訊人員需檢視網站內容，確保網站內容之完整性、真實性、可及性。此外，檔案管理人員或資訊人員也可藉由與該網站前次的抓耙內容相互比對，確認是否以相同或相異的設定方式進行。
- (4) 若經檢視後，發現網站正確或如預期時，則由檔案管理人員或資訊人員透過網站保存系統將此網站妥善保存。若經檢視後，發現網站內容並不完整或不如預期，則由檔案管理人員或資訊人員透過網站保存系統重新抓耙網站。

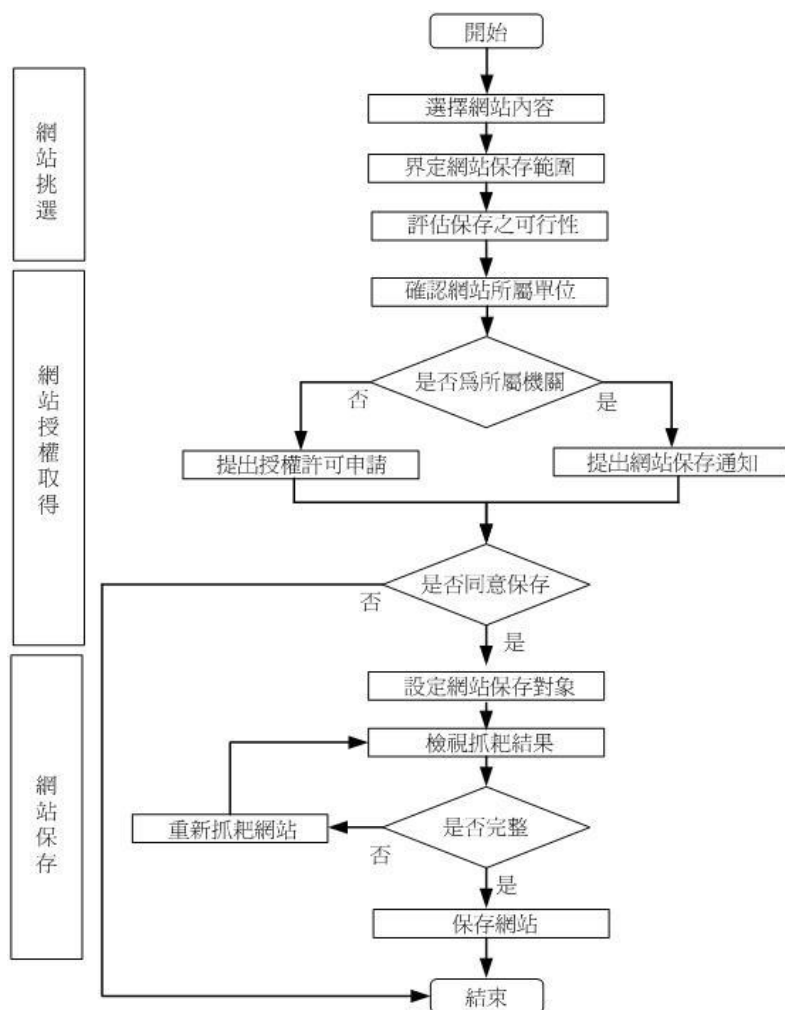


圖 4- 3 網站保存作業流程

二、離型系統

(一) 系統示意圖

為驗證上述處理程序之可行性，本分項計畫建議將網站匯出後，透過著錄電子檔案所應具有之詮釋資料，將網站轉換為電子檔案，進而匯入檔案管理系統中，以納入一般電子檔案管理的範疇，如圖 4-4 所示。

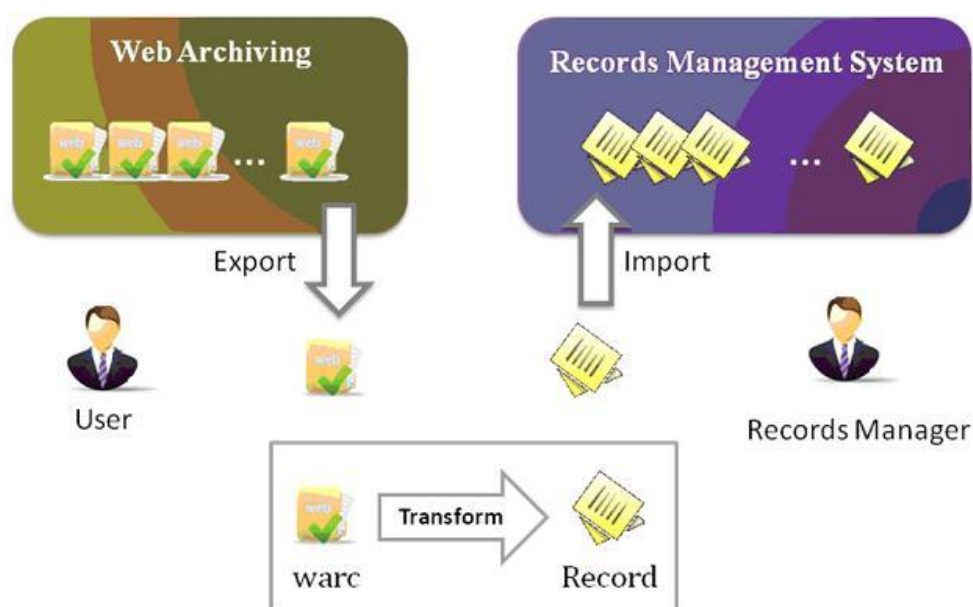


圖 4-4 網站保存示意圖

其中，網站保存系統，如圖 4-5 所示。管理者(Administrator)是指負責執行網站典藏工作的人員，而使用者則是指瀏覽網站者。換言之，管理者可以透過管理工具(Curator Tools)設定系統自動抓取與呈現網站，而使用者則是透過存取工具(Access and finding aids)瀏覽有興趣的網站。就系統運作流程而言，首先，管理工具(Curator Tools)會依據管理者的設定，透過蒐集工具(Acquisition)去蒐集網

站，接著，將蒐集回來的網站存成(Snapshot)網站典藏格式(WARC)，最後，將此網站典藏資料交由存取工具(Access and finding aids)負責呈現給使用者。

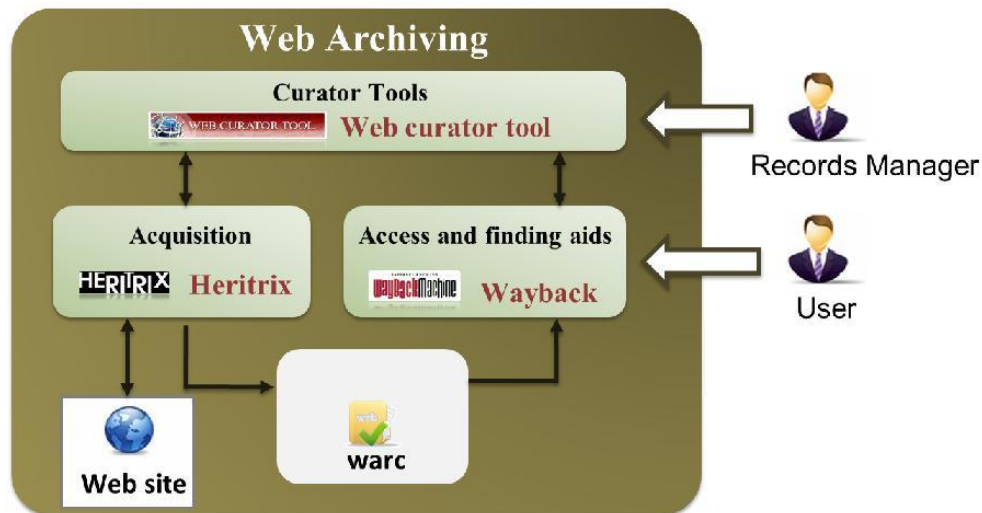


圖 4- 5 網站保存示意圖

(二) 系統實作

網站保存系統安裝手冊，請參考附錄三；網站保存系統操作手冊，請參考錄附四。

三、相關建議

(一) 決定保存優先順序

建立優先順序是成功保存網站資料的重要基礎，因為，保存「每個」網站幾乎是不可能。因此在決定何者該保存，何者不該保存時，可採用 MoSCoW 分類方式以標明保存重要性，將網站分為以下四

種，且因屬於 M 及 S 分類的清單可能壟長且混亂，因此 MoSCoW 方法可在這兩個分類再實行一次，以確認保存優先順序：

- 1、Must(M)：必須典藏之網站。
- 2、Should(Should)：應該盡可能典藏之網站。
- 3、C(Could)：可考慮典藏之網站。
- 4、W(Won't)：不需典藏之網站。

(二) 網站保存執行方式

機關若需進行網站保存時，可採用以下四種方式，並同時請先考量機關之成本、風險、優先權、可得資源及基礎設備等因素後，可採行以下方式：

1、自製 (In-house)

主要由機構負責資源規劃、管理與實行。此方式提供了機構較多的彈性與掌控，並易於評估所需資源及必須能力。

2、外包 (Contracted-out)

所有或部分的工作由合約商負責。若單一合約商不能擁有全部的技能與資源去完成整個計畫，機構需注意合約所包含的服務。比方說，由合約商負責網站的主機及存放空間，而機構保

留管理及挑選等程序。

3、 協作 (Collaborative)

一個以上的機構共同合作，共享技術與資源以達到相同目標。

4、 聯盟 (Consortium)

主要由組織中的聯合單位負責，旨在營造一互助合作的環境，提供會員持續性的政策建立及非會員可參與網站典藏的機制。

(三) 網站保存工具

若機關決定自行進行網站的保存計畫，目前已有一些網站典藏的相關工具可供機關參考。

1、 蒐集(Acquisition)工具

例如：Heritrix、HTTack。

2、 存取(Access)工具

例如：Wayback、WERA、NutchWAX。

3、 管理(Curator)工具

例如：Web CuratorTool、PANDORA Digital Archiving System、NetarchiveSuite。

本計畫所建置的雛型系統已整合了 Web Curarator Tool (含 Hertrix)與 Wayback 等工具，運作順利。

(四) 網站抓耙方式

機關可依需要選擇在伺服器端備份網站資料、瀏覽器端抓取網站畫面，或是使用抓耙軟體保存網站。

1、授權系統或伺服器端

從「授權系統或伺服器端」進行抓耙網站。採用此方式將有機會完整抓耙網站所有內容。但是，越來越多網頁是為配合使用者需求而動態產生，如瀏覽器的樣式、螢幕大小，桌上型電腦或智慧型電腦。因此，產生該保存何種版本之網站的議題。

2、瀏覽器端從「瀏覽器端」著手抓耙網站。採取此方式可採用快照(snapshot)軟體進行抓耙，但僅能抓耙到網站中無法變動的內容。

3、使用抓耙軟體

透過「使用抓耙軟體」保存網站。採取此方式或許能解決上述兩種方式之問題，但有遺失外部資源的疑慮，如資料庫產生之內容就無法囊括。

(五) 選擇網站的保存格式

從「授權系統或伺服器端」著手保存網站是以 Mirror 方式完全備份原始網站檔案，此種方式保存下來的檔案格式會與檔案原來的格式完全相同。而從「瀏覽器端」的保存方式會以圖檔的方式將畫面完整的保存，檔案格式會以常見的 jpg、gif 及 png 等圖檔為主。若使用抓耙軟體保存網站，則建議能採用 WARC 國際標準，以符合國際間的發展趨勢。

第五章 擴大機關核心業務電子檔案之可行性分析

本章將以前一章節中所言研擬之「作業指引與雛型系統」為基礎，分別就電子郵件與網站兩方面，針對經濟、技術、時程、資源與組織等五個構面，進行可行性分析，以供檔案管理局參考。

第一節 電子郵件

一、經濟可行性

探討電子郵件的經濟可行性時，主要包括成本與效益兩方面的分析。前者涵蓋將電子郵件納入電子檔案並保存之軟體成本；後者則涵蓋保存後可帶來的有形收益與無形成效。

（一）成本分析

依本分項計畫於第四章所規劃之電子郵件保存作法是，若要將電子郵件納入機關核心業務電子檔案時，建議由承辦人負責將電子郵件匯出，再交由檔案管理人員負責將匯出之電子郵件匯入檔案管理系統，藉由此過程將電子郵件轉換為電子檔案，進入檔案管理生命週期。因此，需要電子郵件系統提供匯出電子郵件之系統功能，並且需要檔案管理系統提供匯入此電子郵件之系統功能。根據電子郵件系統廠商與檔案管理系統廠商之評估，若需提此系統功能所需

之成本如表 5- 1 所示。

表 5- 1 軟體設備之規格與估價

項目	名稱	單價	數量	總價
1.	電子郵件系統(具備可匯出電子郵件功能)	\$40,000	1	\$40,000
2.	檔案管理系統(具電子郵件典藏功能)	\$100,000	1	\$100,000
總價				

(二) 效益分析

1、保存更多具證據、歷史、文化等價值的電子檔案

電子郵件已是大多數公務人員用以作為處理公務的溝通管道，因此，若將電子郵件納入電子檔案範疇中，將可保存更多具有證據性、歷史性、文化價值等之電子檔案。

2、提升機關處理公務效率

當機關承辦人在處理相關工作時，若將電子郵件納入保存範疇，也可提供承辦人員更多值得參考之檔案資料，進而提升機關處理公務之效率。

3、提供民眾更多電子檔案內容

機關處理公務的過程中，難免會有需要與民眾溝通的機會，目前已有許多機關也開始透過電子郵件與民眾進行溝通。因此，若將電子郵件納入電子檔案範疇，當民眾在查詢機關電子檔案時，機關也能提供更多具有豐富資訊的電子檔案內容，以滿足

民眾對檔案資訊的需求。

4、符合國際趨勢

目前國際間已有許多國家開始探討電子郵件保存相關作法與配套措施，例如：美國、英國、加拿大。因此，若儘早將電子郵件納入電子檔案範疇，便可及早符合國際趨勢，進而與國際接軌。

根據上述成本與效益分析，若機關考慮將電子郵件納入電子檔案範疇時，所需的成本為修改電子郵件系統與檔案管理系統之經費。實際上，機關可能會面臨缺少相關經費的困難。因此，若檔案管理局可提供部分補助，或機關於軟體系統定期更新時，再一併更新系統功能，如此一來，就經濟可行性而言，便可將電子郵件納入電子檔案範疇。

二、技術可行性

依本分項計畫於第四章所規劃之電子郵件保存指引，「電子郵件保存」所涉及的技術範圍主要在於軟體系統部分，相關人員有資訊人員、檔案管理人員、承辦人、電子郵件系統廠商、以及檔案管理系統廠商，以下則依照此三種角色分別介紹其所需具備之相關技術：

（一）資訊人員

1、系統維護

資訊人員須充分了解各種設備的使用狀況，並依規範定期維護保養檔案管理系統與電子郵件系統，以提供妥善的電子郵件保存環境。

2、系統管理

資訊人員須充分了解檔案管理系統與電子郵件系統環境，尤其應善加管理網路與資料庫，並維護資訊安全，以發揮電子郵件保存的最大成效。

（二）檔案管理人員

檔案管理人員則須充分了解電子郵件匯入檔案管理系統之功能與操作方法，以妥善保管電子郵件中重要檔案資訊。

（三）承辦人

承辦人員須充分瞭解匯出電子郵件之相關功能與操作方法，以妥善保管電子郵件中重要檔案資訊。

（四）電子郵件系統廠商：電子郵件典藏功能增修

電子郵件軟體廠商必須具有增修電子郵件保存功能之專業能力，尤其應依照電子檔案詮釋資料標準，將電子郵件匯出成為符合電子檔案格式，並且提供承辦人著錄相關詮釋資料之功能。

（五）檔案管理系統廠商：檔案管理系統匯入功能增修

檔案管理系統廠商必須具有增修電子郵件保存功能之專業能力，尤其應將符合電子檔案之電子郵件匯入系統中，成為待點收之電子檔案，並且將此電子郵件納入一般電子檔案管理範疇與作業流程中。

因此，根據以上技術可行性分析，若機關考慮將電子郵件納入電子檔案範疇時，可能面臨的困難在於電子郵件系統廠商與檔案管理系統廠商無法更新或修正系統功能。但是，根據本分項計畫實作電子郵件保存雛型系統後，發現目前相關軟體廠商應該都具備此系統修正能力。

三、時程可行性

電子郵件保存可從「機關」與「檔案管理局」兩者不同之角度探討，其中，機關是實際操作與落實電子郵件保存者，而檔案管理局則是負責推動各機關將電子郵件納入保存之重要角色。因此，以下將從此兩種角度探討電子郵件納入保存範圍之時程可行性。

（一）機關

根據依本分項計畫於第四章所規劃之電子郵件典藏，電子郵件保存作業可分為「電子郵件蒐集」與「電子郵件保存」兩階段，

若機關欲將電子郵件納入電子檔案範疇，則主要包含以下四個工作項目，所需建置時程約四個月，如表 5-2 所示：

1、更新電子郵件系統功能

當機關要開始進行電子郵件保存前，首先須與電子郵件系統廠商溝通，商討更新電子郵件系統功能，使機關中使用之電子郵件系統具備匯出電子郵件之功能，且此匯出之電子檔案必須依循電子檔案詮釋資料相關標準。

2、更新檔案管理系統功能

除更新電子郵件系統功能外，機關也須同時更新檔案管理系統功能，使檔案管理系統可以匯入從電子郵件匯出之檔案。

3、電子郵件保存方式與概念之教育訓練

當軟體系統廠商完成功能更新後，機關內承辦人員可能尚未了解電子郵件保存之概念與相關操作方式。因此，機關在正式進行電子郵件保存前，應先針對機關內所有相關人員，包含：承辦人員、檔案管理人員、資訊人員等，進行電子郵件保存方式與概念之教育訓練或宣導。

4、進行電子郵件保存

待上述工作完成後，機關便可依循電子郵件保存指引中所建議

之操作步驟進行電子郵件保存。電子郵件保存作業流程中，首先進行電子郵件蒐集階段，此階段是各承辦人定期或不定期將具有檔案價值之電子郵件匯出；接著，再經由檔案管理人員將此電子檔案匯入檔案管理系統中。

工作項目	工作時程			
	1 st 月	2 nd 月	3 rd 月	4 th 月
1.更新電子郵件系統功能				
2.更新檔案管理系統功能				
3.電子郵件典藏方式與概念之教育訓練				
4.進行電子郵件典藏				

表 5- 2 機關進行電子郵件保存之時程表

(二) 檔案管理局

若檔案管理局欲推行電子郵件保存作業，則主要包含以下工作項目，所需建置時程約 X，如表 5- 3 所示：

1、規劃電子郵件保存辦法與相關配套措施(相關辦法與指引)

檔案管理局推動電子郵件保存前，應先規劃電子郵件保存辦法與相關配套措施，以利後續各機關實際操作時可參考或依循之用。

2、召開專家座談會

當電子郵件保存辦法與相關配套措施規畫完畢後，應召集各機關與相關專家，從實務面探討此保存辦法與作業流程之可行

性，以及相關配套措施是否適當，並依據此會議討論結果修正電子郵件保存作業指引與配套措施。

3、進行電子郵件典藏試辦作業

為確保此電子郵件保存作業指引之可行性，以及瞭解並發現各機關實務面的困難。檔案管理局應就電子郵件保存作業進行試辦，實際以少數機關為試作對象。

4、修正電子郵件管理辦法與相關配套措施

經過電子郵件保存試辦作業後，檔案管理局應依據此試辦作業進行檢討改進，探討試辦過程中發現之困難與解決方法，進而提出更適合我國之電子郵件保存作業指引。

5、確立電子郵件保存辦法與相關配套措施

為推動我國電子郵件保存作業，檔案管理局應公告此作業指引，並確立此指引與相關配套措施之內容，以供各機關參考遵循。

6、機關檔案管理人員教育訓練

為順利推動此電子郵件保存作業，檔案管理局應在確立保存作業流程與相關配套措施後，主動對機關進行教育訓練與宣導工作，使各機關了解電子郵件保存之目的與重要性，並瞭解如何

進行電子郵件保存。

7、實行電子郵件保存作業

當上述工作項目完成後，檔案管理局便可正式實行電子郵件保存作業，以免各機關流失寶貴的電子檔案。

工作時程 工作項目	1 st 月	2 nd 月	3 rd 月	4 th 月	5 th 月	6 th 月	7 th 月	8 th 月	9 th 月
1.規劃電子郵件保存辦法與相關配套措施(相關辦法與指引)									
2.召開專家座談會									
3.進行電子郵件保存試辦作業									
4.修正電子郵件管理辦法與相關配套措施									
5.確立電子郵件管理辦法與相關配套措施									
6.機關檔案管理人員教育訓練									
7.實行電子郵件保存作業									

表 5- 3 檔案管理局推行電子郵件保存之時程表

因此，根據以上時程可行性分析，若機關考慮將電子郵件納入電子檔案範疇時，可能最耗費時間的工作在於機關內電子郵件保存方式與概念之教育訓練。因此，建議針對各承辦人員進行宣導時，注意各承辦人所提之疑慮與困擾，並妥善處理。此外，若檔案管理局考慮將電子郵件納入電子檔案範疇時，可能導致時程延宕的主要工作在於相關配套措施推行不易，因此，建議藉由教育訓練推廣電子郵件保存概念與做法，以免推行電子郵件保存時窒礙難行。

四、資源可行性

根據依本分項計畫於第四章所規劃之電子郵件保存作業指引，電子郵件保存作業所需資源包含以下四項：

（一）軟體設備

若機關打算將電子郵件納入電子檔案範疇時，則需要之軟體廠商將電子郵件系統、檔案管理系統功能修正。

（二）財務資源

若機關打算將電子郵件納入電子檔案範疇時，則需要相關經費，以支付電子郵件系統與檔案系統修正之費用。

（三）人力資源

若機關打算將電子郵件納入電子檔案範疇時，內部相關人員必須有時間處理額外工作，例如：系統安裝、檔案點收等工作。

（四）外部資源

若機關打算將電子郵件納入電子檔案範疇時，現有電子郵件系統廠商、檔案管理系統廠商是否可提供相關軟體功能。

因此，根據以上資源可行性分析，若機關考慮將電子郵件納入電子檔案範疇時，軟體資源、財務資源、人力資源、外部資源似乎取得並不困難。

五、 組織可行性

就組織可行性而言，電子郵件保存可分為「機關」與「檔案管理局」兩個面向來探討。其中，機關之組織可行性是指機關內部實際落實電子郵件保存時，可能面臨的抗拒心理；而就檔案管理局而言，則是從推動電子郵件典藏到各機關之角度來思考。因此，以下將從此兩種角度探討電子郵件納入保存範圍之組織可行性。

（一）機關

若機關欲推行電子郵件保存作業，則可能包含以下困難：

1、 檔案管理人員之態度

依據本分項計畫規畫之電子郵件保存作業指引，機關檔案管理人員須負責點收電子郵件，使其進入檔案管理生命週期。因此，當機關要落實電子郵件保存作業時，檔案管理人員可能會覺得未來工作量會遽增，例如：帶點收之電子檔案增加、須保存之電子檔案增加等。

2、 承辦人員之態度

依據本分項計畫研擬之電子郵件保存作業指引，機關承辦人員須定期或不定期將與處理公務相關之電子郵件匯出，以供檔案管理人員點收。因此，當機關要落實電子郵件保存作業時，承

辦人員可能會因為不知如何選擇具有檔案價值之電子郵件，或因為需處理之電子郵件數量龐大，而產生抗拒心理。

3、資訊人員之態度

依據本分項計畫規畫之電子郵件保存作業指引，資訊人員須負責更新電子郵件保存系統與檔案管理系統功能，使承辦人員與檔案管理人員可以順利使用電子郵件保存所需之功能。但是上述工作內容，原本就是機關資訊人員須負責處理之工作內容。因此，當機關要落實電子郵件保存作業時，資訊人員不會產生抗拒心理。

（二）檔案管理局

若檔案管理局欲推行電子郵件保存作業，則可能包含以下困難：

1、各機關內部首長是否願意配合辦理

若檔案管理局欲推動電子郵件保存作業，對於各機關內部首長與檔案管理主管而言，會是一項新的挑戰，而且倘若配合此政策是否對於機關執行公務有正向影響，以及配合所需付出之代價為何，會是機關首長主要考量。

2、對於現行法規之影響

若檔案管理局欲推動電子郵件保存作業，各機關檔案管理人員首先可能會關心此新政策對於現行法規之影響。例如：是否因而修正既有法規，導致現行作業必須調整。因此，檔案管理局須評估將電子郵件納入保存範圍可能影響之現行法規。

3、電子郵件保存指引與相關配套措施

若檔案管理局欲推動電子郵件保存作業，各機關在落實此作業時，將參考檔案管理局所制定之電子郵件保存指引以及相關配套措施。因此，電子郵件保存作業指引中各項作業流程、人員權責、相關配套措施等，將會影響機關是否願意配合處理之態度。

因此，根據以上組織可行性分析，若機關考慮將電子郵件納入電子檔案範疇時，可能面臨組織內檔案管理人員與承辦人員的抗拒心理。因此，若要順利落實電子郵件保存，則須事先宣導電子郵件典藏之目的與方法，以消弭檔案管理人員與承辦人員心中的疑慮。

若檔案管理局考慮將電子郵件納入電子檔案保存範疇時，則可能面臨的困難或挑戰有：機關首長的態度、現行法規修正所造成的影響、落實電子郵件保存作業指引的困難。因此，若要順利推動電子郵件保存，則須使機關首長瞭解電子郵件典藏之目的與效益、減少修正現行法規之內容、使電子郵件保存作業與機關檔案管理人員作業流程相仿。

六、 小結

以上針對經濟、技術、時程、資源及組織等五個夠面來分析電子郵件保存可行性。從經濟可行性的角度來看，電子郵件保存所需軟體成本是影響經濟可行性高低的關鍵，但若機關於軟體系統定期更新時，請軟體廠商一併更新功能，便可提升經濟可行性。同時，獲得保存電子郵件之效益，例如：保存更多具證據、歷史、文化價值之電子檔案、提升公務機關處理公務效率、提供民眾更多電子檔案內容、符合國際趨勢；就技術可行性而言，經本分項計畫實作雛型系統後，發現目前相關軟體廠商都具備發展電子郵件保存系統之能力，因此技術可行性相當高；時程可行性中，電子郵件保存方式與概念之教育宣導會是最關鍵的工作，若機關或檔案管理局可在教育宣導時，詳細的說明電子郵件保存之目的與效益、並解說保存方式，便可提升時程可行性；資源可行性方面，經本分項計畫實作雛型系統後發現，軟體資源、財務資源、人力資源與外部資源的取得都不算困難，因此資源可行性也頗高；組織可行性中，檔案管理人員與資訊人員的態度，及可否得到相關首長的支持是主要的困難，因此機關需事先針對電子郵件保存進行宣導並安排教育訓練，而檔案管理局則需使機關首長瞭解電子郵件保存之目的與效益。

第二節 網站

一、經濟可行性

探討網站保存的可行性分析，主要包括成本與效益兩方面的分析。前者包含建置網站保存環境所需軟硬體設備建置成本與每年的維運成本；後者則涵蓋網站保存實行之後可帶來的有形收益與無形成效。

（一）成本分析

1、建置成本

依本分項計畫於第四章所規劃之網站保存作法，檔案管理人員需運用網站抓耙工具蒐集列入保存範圍的網站，並建置查詢平台供機關內外部人士檢閱及瀏覽。因此，在硬體設備部份，可延用原電子檔案儲放環境，然儲放空間的硬碟容量則根據各機關需求而有差異。參考本分項計畫參訪台灣大學網站典藏系統的資料，台大平均每 3-4 個月增加約 1T 的資料量，因此本分項計畫概估一般機關一年的需求量約為 4T；而軟體方面，則需增設網站抓耙工具及查詢平台。具體而言，實行網站保存時，需建置的硬體、軟體設備之規格與經費如表 5-4 及表 5-5 所示，因此，建置成本之總價如表 5-6 所示。

表 5- 4 硬體設備之規格與估價

項目	名稱	參考規格	單價	數量	總價
1.	硬碟	3.5 吋 2TB 外接硬碟*2 顆	3988	2	\$7, 976
總價					\$7, 976

表 5- 5 軟體設備之規格與估價

項目	名稱	單價	數量	總價
1.	Linux Ubuntu 10.04 i386	\$0	1	\$0
2.	Postgresql 8.4	\$0	1	\$0
3.	Tomcat 5.5.30	\$0	1	\$0
4.	OpenJDK 1.6.0	\$0	1	\$0
5.	Web Curator Tool (WCT) 1.5	\$0	1	\$0
6.	Wayback Machine	\$0	1	\$0
總價				0

表 5- 6 網站典藏設備建置成本

成本項目	經費
硬體設備	\$7, 976
軟體設備	\$0
合計	\$7, 976

2、維運成本

網站典藏的軟硬體設施建置完成之後，每年所需的維運成本包括：連線費用及硬碟。連線費用及硬碟成本都可因各機關的需求差異而變動，而概估的經費如表 5- 7 示。

表 5- 7 網站典藏之維運成本

成本項目	經費	說明
連線費用	\$100, 000	一年之網路連線費用
硬碟	\$7, 976	每年需新增的儲存容量概估 4T，然須依據各機關每年保存的網站檔案大小而調整。
合計	\$107, 976	

（二）效益分析

網站典藏計畫開始執行之後，預期可獲得以下效益：

1、保存更多具證據、歷史、文化等價值的網站資料

網站可作為機關活動的歷史記錄，或代表機關文化的軌跡。若有妥善保存網站資料，便可提供民眾瞭解機關的活動事蹟或某事件的發展歷程，或者作為日後供學者研究的參考。

2、提升機關處理公務效率

歷史網頁資訊可能記載著某個事件的處理程序及成果，日後機關人員在接觸類似公務時，即可參考前人作法，有效提升決策速度。

3、資源重覆運用，節省成本

網站資料中所包含的圖表，都是前人花時間收集資料、精心設計出來的成果，若能善加運用既有的資源，則可達節省時間及人力成本之效。

根據上述成本與效益分析，若機關考慮將網站內容納入電子檔案範疇時，其所將要負擔的建置成本其實是相對低廉的。在硬體方面幾乎無需增購任何設備，而軟體方面機關也可以採用免費軟體，以節省經費。維運成本包含的人力及維護費用都可透過活用機關現有資源來達成，因此在經濟

方面相當可行。

二、技術可行性

依本分項計畫於第四章所規劃之網站保存，「網站保存」所涉及的技術範圍主要在於軟體系統部分，相關人員有資訊人員及檔案管理人員，以下則依照此二種角色分別介紹其所需具備之相關技術：

（一）檔案管理人員

1、挑選網站

檔案管理人員須有挑選具保存價值網站之能力，並根據價值差異定義合宜的保存年限、抓取頻率、抓取深度等。

2、系統操作

檔案管理人員須充份了解網站保存相關軟體的操作，例如：WCT 與 Wayback Machine 的功能與操作方法，詳細內容請參考附錄四。

（二）資訊人員

1、系統安裝

資訊人員在安裝系統時，須熟悉網站保存相關軟體安裝方法，

詳細內容請參考附錄三。

2、系統維護

資訊人員須充分瞭解各種設備的使用狀況，並依規範定期維護保養，及評估設備過期的可能性，以提昇保存設備的可用性。

3、系統管理

資訊人員須充分瞭解網站保存環境，尤其應善加管理網路與資料庫，並確保網站保存之資訊安全。

根據以上技術可行性分析，若機關考慮將電子郵件納入電子檔案範疇時，檔案管理人員可能面臨的困難在於無法精確的挑選具保存價值之網站，若篩選機制寬鬆，保存目標可能過於繁雜；若篩選機制嚴謹，也會有遺珠之憾。因此檔案管理人員須經過足夠的辨識訓練，才能精準網站典藏目標。對資訊人員而言，最大的挑戰應在於系統環境的架設，因軟體特性，安裝時需特別注意平台、語系等要求，本分項計畫也於附錄三、四中詳細說明網站保存系統安裝手冊與系統操作手冊，以協助資訊人員減少阻礙。整體而言，技術可行性在相關專業人員充足的訓練之下，可提升技術可行性。

三、時程可行性

網站保存可從「機關」與「檔案管理局」兩者不同之角度探討，其中，機關是實際操作與落實網站典藏者，而檔案管理局則是負責推動各機關將網站納入保存之重要角色。因此，以下將從此兩種角度探討網站納入保存範圍之時程可行性。

（一）機關

根據本分項計畫於第四章所規劃之網站保存，若機關欲將網站內容納入電子檔案範疇，其保存作業可分為「網站挑選」、「保存通知」與「保存執行」三階段，所需建置時程約 X，如表 5-8 所示：

1、網站挑選

當機關要開始進行網站保存前，檔案管理人員須在此階段決定挑選的途徑及範圍。

2、保存通知

檔案管理人員需在確認保存目標之後，依照機關屬性，寄出通知或許可申請，並協助機關排除困難，達成同意保存網站的協議。

3、保存執行

待完成前二項工作之後，檔案管理人員開始進行網站抓耙工具

的操作，並重覆檢核以確保抓耙成果的正确性。最後，需確認查詢工具裡顯示的網站資料能依據檢閱者的權限差異，存取到合宜的內容。

工作時程 工作項目	1 st 月	2 nd 月	3 rd 月
1、網站挑選			
2、典藏通知			
3、典藏執行			

表 5- 8 機關進行網站保存之時程表

(二) 檔案管理局

若檔案管理局欲推行網站保存作業，則主要包含以下工作項目，所需建置時程約 9 個月，如表 5- 9 所示：

1、規劃網站保存辦法與相關配套措施(相關辦法與指引)

檔案管理局推動網站保存前，應先規劃網站保存辦法與相關配套措施，以利後續各機關實際操作時可參考或依循之用。

2、召開專家座談會

當網站保存辦法與相關配套措施規畫完畢後，應召集各機關與相關專家，從實務面探討此典藏辦法與作業流程之可行性，以及相關配套措施是否適當，並依據此會議討論結果修正網站典藏作業指引與配套措施。

3、進行網站保存試辦作業

為確保此網站保存作業指引之可行性，以及瞭解並發現各機關實務面的困難。檔案管理局應就網站保存作業進行試辦，實際以少數機關為試作對象。

4、修正網站管理辦法與相關配套措施

經過網站保存試辦作業後，檔案管理局應依據此試辦作業進行檢討改進，探討試辦過程中發現之困難與解決方法，進而提出更適合我國之網站保存作業指引。

5、確立網站保存辦法與相關配套措施

為推動我國網站保存作業，檔案管理局應公告此作業指引，並確立此指引與相關配套措施之內容，以供各機關參考遵循。

6、機關檔案管理人員教育訓練

為順利推動此網站保存作業，檔案管理局應在確立保存作業流程與相關配套措施後，主動對機關進行教育訓練與宣導工作，使各機關了解網站保存之目的與重要性，並瞭解如何進行網站保存。

7、實行網站保存作業

當上述工作項目完成後，檔案管理局便可正式實行網站保存作業，以免各機關流失寶貴的電子檔案。

工作時程 工作項目	1 st 月	2 nd 月	3 rd 月	4 th 月	5 th 月	6 th 月	7 th 月	8 th 月	9 th 月
1、規劃網站保存管理辦法與相關配套措施									
2、召開專家座談會									
3、進行網站保存試辦作業									
4、修正網站管理辦法與相關配套措施									
5、確立網站管理辦法與相關配套措施									
6、機關檔案管理人員教育訓練									
7、實行網站保存作業									

表 5- 9 檔案管理局推行網站保存作業之時程表

根據以上時程可行性分析，若機關考慮將網站納入電子檔案範疇時，檔案管理人員可能會遇到的困難會發生在典藏通知階段，尤其保存的目標是由外部機關所擁有。外部機關可能因不瞭解保存的影響及重要性，不願配合保存，因此，檔案管理人員就應進行充份的溝通，使其了解保存的目的與效益，降低對方的抗拒，以順利完成網站保存，提升時程可行性。

此外，若檔案管理局考慮將網站檔案納入電子檔案範疇時，可能導致時程延宕的主要工作在於相關配套措施推行不易，因此，建議藉由教育訓練推廣網站保存概念與做法，以免推行網站保存時窒礙難行。

四、資源可行性

根據依本分項計畫於第四章所規劃之網站保存，網站保存作業所需資源包含以下四項：

（一）軟體設備

若機關打算將網站內容納入電子檔案範疇，則須取得所有需安裝的軟體工具。若機關選擇採用本分項計畫實作之雛型系統作為網站抓耙工具，則可參考附錄三、網站保存系統安裝手冊，並至相關網站免費下載所需軟體。

（二）財務資源

若機關打算將網站內容納入電子檔案範疇，財務將會是影響最小的阻礙。由於軟體工具多數免費取得，硬體資源又可以既有設備為主，因此不會增加財務上的負擔。

（三）人力資源

若機關打算將網站內容納入電子檔案範疇，內部相關人員的工作量勢必有所影響。對檔案管理人員而言，需多處理網站內容的挑選事務及申請同意等程序，且也需熟悉網站抓耙工具的操作；而資訊人員則要增加網站抓耙工具安裝、設定及維護等技能。機關須確保相關人員有多餘的時間可負擔增加的工作內容。

（四）外部資源

為提升資料完整性，在針對某個主題進行網站保存時，可能需向外部機關或一般民眾取得資料。

根據以上資源可行性分析，若機關考慮將網站納入電子檔案範疇時，主要的困難會出現在人力資源的負荷能力，機關應根據內部工作量安排確認是否增加人手。

五、組織可行性

就組織可行性而言，網站保存可分為「機關」與「檔案管理局」兩個面向來探討。其中，機關之組織可行性是指機關內部實際落實網站典藏時，可能面臨的抗拒心理；而就檔案管理局而言，則是從推動網站保存到各機關之角度來思考。因此，以下將從此兩種角度探討網站納入保存範圍之組織可行性。

（一）機關

若機關欲推行網站保存作業，則可能包含以下困難：

1、檔案管理人員之態度

依據本分項計畫規畫之網站保存作業指引，機關檔案管理人員須負責挑選、鑑定並操作網站保存的相關系統。檔案管理人員的未來工作量會增加，相對也將影響工作時數。然而只要有合理的人力、工時來處理此任務，檔案管理人員的抗拒心理應該不高。

2、資訊人員之態度

依據本分項計畫規畫之網站保存作業指引，資訊人員須負責網站保存工作的環境設定及後續維護。在網站保存計畫初期，資訊人員可能要花較長時間在確認系統可正常運作，只要渡過這段建設時間，之後僅需定期維護。只要有合理的人力工時來處理此任務，資訊人員的抗拒心理應該不高。

（二）檔案管理局

若檔案管理局欲推行網站保存作業，則可能面臨以下困難：

1、各機關內部首長是否願意配合辦理

若檔案管理局欲推動網站保存作業，對於各機關內部首長與檔案管理主管而言，會是一項新的挑戰，而且倘若配合此政策是否對於機關執行公務有正向影響，以及配合所需付出之代價為何，會是機關首長主要考量。

2、對於現行法規之影響

若檔案管理局欲推動網站保存作業，各機關檔案管理人員首先可能會關心此新政策對於現行法規之影響。例如：是否因而修正既有法規，導致現行作業必須調整。因此，檔案管理局須評估將網站納入保存範圍可能影響之現行法規。

3、網站典藏指引與相關配套措施

若檔案管理局欲推動網站保存作業，各機關在落實此作業時，

將參考檔案管理局所制定之網站保存指引以及相關配套措施。

因此，網站保存作業指引中各項作業流程、人員權責、相關配

套措施等，將會影響機關是否願意配合處理之態度。

根據以上組織可行性分析，若機關考慮將網站納入電子檔案範疇時，可能面臨組織內檔案管理人員與資訊人員的抗拒心理。因此，若要順利落實網站典藏，則須事先宣導網站典藏之目的與方法，以消彌檔案管理人員與資訊人員心中的疑慮。

若檔案管理局考慮將網站納入電子檔案保存範疇時，則可能面臨的困難或挑戰有：機關首長的態度、現行法規修正所造成的影響、落實網站保存作業指引的困難。因此，若要順利推動網站保存，則須使機關首長瞭解網站保存之目的與效益、減少修正現行法規之內容、使網站保存作業與機關檔案管理人員作業流程相仿。

六、小結

網站保存的可行性分析針對經濟、技術、時程、資源及組織等五構面進行。綜觀而言，經濟可行性因網站保存的軟硬體設備均無需額外費用，

因此相當可行；技術可行性可能會因為檔案管理人員的辨識技巧不足，或資訊人員無法完成環境的架設而遭遇阻礙，然而都可透過足夠的教育訓練克服，技術可行性仍高；時程可行性方面，檔案管理人員與網站擁有人的溝通則會成為是否可行之重點，因此檔案管理局需透過教育訓練使機關了解網站保存的目的，提升機關間的配合度；資源可行性可能會受到機關目前人力及工作量的影響，須視情況增加人手以利網站保存之進行；組織可行性中，檔案管理人員與資訊人員的態度，及可否得到相關首長的支持，則會影響網站保存計畫的推行，應透過法規及完善的流程規畫協助計畫進行。

第六章 結論與建議

由於資訊科技蓬勃發展，許多先進國家已擴大電子檔案所涵蓋之範圍，如電子郵件、網站皆已納入電子檔案之範疇。同時，電子檔案定義擴大後所延伸出的相關議題已備受各國檔案管理單位之重視，而我國國家檔案管理局亦積極規劃電子檔案定義擴大之各項議題與配套措施。為確保電子檔案定義擴大得以順利推行，本分項計畫針對擴大我國電子檔案範圍的相關議題，進行解決方案之規劃與實作，以作為檔案管理局修訂相關法規之參考。

目前國內具有電子郵件與網站保存相關經驗之機關或軟體廠商相當有限，因此本分項計畫實地參訪我國長期致力於電子郵件系統發展之軟體廠商（網擎資訊）以及具有網站保存經驗之台灣大學與國家圖書館。同時，為瞭解國外機構在相關議題之發展現況，本分項計畫於民國 99 年 7 月 5-18 日前往英國，參訪機構包含 The British Library、Digital Preservation Coalition、Joint Information Systems Committee、University of London Computer Centre、Cambridge University Archive、Oxford University Archive、Hanzo Archive。

參訪後，本分項計畫邀集相關專家、學者、機關、資訊廠商等，舉辦專家座談會，針對本分項計畫研擬之電子郵件與網站保存作法，進行意見交換。接著，為確保電子郵件與網站保存之可行性，本分項計畫委託二一零零科技與網擎資訊共同實作電子郵件保存雛型系統，並且自行實作網站保存系統。最後，

本分項計畫根據國內外參考文獻、國內機關訪談、國外實地參訪、專家座談會以及實作電子郵件與網站保存雛型系統，彙整待克服之問題，並提出解決之道。

第一節 結論

電子郵件與網站保存將是未來時勢所趨，然而根據本分項計畫所進行之機關參訪，我國對於保存電子郵件與網站之可行性仍有許多疑慮。因此本分項計畫針對經濟、技術、時程、資源與組織等五個構面，進行可行性分析，並且實作電子郵件保存雛型系統與網站保存雛型系統。

就電子郵件保存而言，電子郵件保存所需軟體成本是影響經濟可行性高低的關鍵，但若機關於軟體系統定期更新時，請軟體廠商一併更新功能，便可提升經濟可行性。同時，獲得保存電子郵件之效益，例如：保存更多具證據、歷史、文化價值之電子檔案、提升公務機關處理公務效率、提供民眾更多電子檔案內容、符合國際趨勢；就技術可行性而言，經本分項計畫實作雛型系統後，發現目前相關軟體廠商都具備發展電子郵件保存系統之能力，因此技術可行性相當高，詳細系統安裝與操作說明，請參考附錄一與附錄二；時程可行性中，電子郵件保存方式與概念之教育宣導會是最關鍵的工作，若機關或檔案管理局可在教育宣導時，詳細的說明電子郵件保存之目的與效益、並解說保存方式，便可提升時程可行性；資源可行性方面，經本分項計畫實作雛型系統後發現，軟體資源、財務資源、人力資源與外部資源的取得都不算困難，因此資源可行性也頗高；組織可行性中，檔案管理人員與資訊人員的態度，及可否得到相關

首長的支持是主要的困難，因此機關需事先針對電子郵件保存進行宣導並安排教育訓練，而檔案管理局則需使機關首長瞭解電子郵件保存之目的與效益。

網站保存所需之軟硬體設備主要是儲存空間與系統維護之成本，各機關可依照本身的需求自行調整，因此從經濟面來看相當可行；技術可行性之關鍵則在於檔案管理人員是否能選擇適當之保存對象，以及資訊人員建置系統之能力，但這些困難都可透過足夠的教育訓練克服，且本分項計畫也針對網站保存系統進行實作，發現技術方面仍為可行，詳細系統安裝與操作內容請參考附錄三與附錄四；時程可行性方面，檔案管理人員與網站擁有者的溝通則會成為是否可行之重點，因此檔案管理局需透過教育訓練使機關了解網站保存的目的，提升機關間的配合度，但若機關僅針對內部相關網站進行保存，則不需要與網站擁有者溝通，便可提升時程之可行性；資源可行性可能會受到機關目前人力及工作量的影響，須視情況增加人手以利網站保存之進行；組織可行性中，檔案管理人員與資訊人員的態度，及可否得到相關首長的支持，則會影響網站保存計畫的推行，應透過法規及完善的流程規畫協助計畫進行。

第二節 建議

本分項計畫針對電子郵件保存與網站保存提出以下相關建議：

一、電子郵件保存

（一）制定電子郵件保存詮釋資料

若將電子郵件保存至檔案管理系統時，需依照電子檔案之詮釋資料才可順利進入系統中，但電子檔案詮釋資料現有之欄位無法完整保存電子郵件內容，例如：寄發郵件者、接受郵件者、電子郵件本文及其附件等詮釋資料，因此建議在現有電子檔案詮釋資料中，增加屬於電子郵件相關之詮釋資料。

（二）避免採用公務信箱處理私人事務

由於處理公務相關之電子郵件需要妥善保存，但許多承辦人除了使用公務信箱處理公務相關電子郵件，也會使用公務信箱處理私人事務，導致承辦人必須篩選出與公務相關之電子郵件，並且予以保存，如此一來，承辦人可能會誤判此電子郵件是否與公務相關，以至於未保存公務相關之電子郵件。因此，建議處理私人事務時，請採用私人信箱，若處理公務時，請採用公務信箱，換言之，機關承辦人員應避免採用公務信箱處理私人事務。

（三）處理機密或敏感性電子郵件

在傳遞電子郵件時，可能會涉及到敏感性資料的範圍主要有：與國家安全相關的資訊、與商業機密相關的資訊、個人隱私相關的資訊。處理機密或敏感性資料時，應謹慎處理，以免誤送或遭竄改等狀況發生。因此，各機關得選定可確保資訊安全的方式，如採用加密方式傳遞電子郵件。

（四）設定電子郵件標題名稱

檔案管理系統中案由或案名是檢索檔案的重要指標，若將電子郵件的標題(subject)直接作為案由或案名，可能會無法清楚描述電子郵件被視為檔案之原因，因此建議承辦人員在命名電子郵件的標題時應審慎考量。或者，建議把電子郵件保存至檔案管理系統時，給予一個更適切的標題作為案由或案名，並且可保留電子郵件原本的標題。

（五）保存電子郵件的時機

機關處理公務相關之電子郵件來往頻繁，為確保與公務相關之電子郵件均妥善保存，並避免遺漏重要公務相關之電子郵件。因此，建議承辦人員得採定期保存電子郵件的方式，或當電子郵件傳送或接收時，便同時進行保存。

（六）匯送電子郵件至檔案管理系統

由於處理公務相關之電子郵件也應視為電子檔案，若將此類電子郵件保存於電子郵件系統中，將不易依照電子檔案管理辦法妥善管理。因此，建議將電子郵件也保存至檔案管理系統中。若機關已建置檔案管理系統，則建議在電子郵件系統中增加匯出電子郵件成為一筆電子檔案之功能，並透過檔案管理系統中電子檔案點收之功能，將電子郵件匯入檔案管理系統，使電子郵件進入檔案管理的生命週期，以利後續管理與維護；若機關尚未建置檔案管理系統時，則建議把電子郵件放到電子郵件系統之共用資料夾(shared public email folder)，或將電子郵件保存在共用的特定系統或硬碟中。

(七) 選擇電子郵件保存格式

由於不同電子郵件系統支援之電子郵件格式也不同，各機關保存電子郵件時，應採用何種格式保存電子郵件？為使各機關有可依循之保存格式，本計畫建議採用目前電子郵件格式中最主要的標準「.eml」，它也是現有大多數電子郵件系統支援的格式。此外，因應科技變化，應定期審視此保存格式是否適合。

二、網站保存

(一) 決定保存優先順序

建立優先順序是成功保存網站資料的重要基礎，因為，保存「每

個」網站幾乎是不可能。因此在決定何者該保存，何者不該保存時，可採用 MoSCoW 分類方式以標明保存重要性，將網站分為以下四種，且因屬於 M 及 S 分類的清單可能壟長且混亂，因此 MoSCoW 方法可在這兩個分類再實行一次，以確認保存優先順序：

- 1、Must(M)：必須典藏之網站。
- 2、Should(Should)：應該盡可能典藏之網站。
- 3、C(Could)：可考慮典藏之網站。
- 4、W(Won't)：不需典藏之網站。

(二) 網站保存執行方式

機關若需進行網站保存時，可採用以下四種方式，並同時請先考量機關之成本、風險、優先權、可得資源及基礎設備等因素後，可採行以下方式：

1、自製 (In-house)

主要由機構負責資源規劃、管理與實行。此方式提供了機關較多的彈性與掌控，並易於評估所需資源及必須能力。

2、外包 (Contracted-out)

所有或部分的工作由合約商負責。若單一合約商不能擁有全部

的技能與資源去完成整個計畫，機構需注意合約所包含的服務。比方說，由合約商負責網站的主機及存放空間，而機構保留管理及挑選等程序。

3、協作 (Collaborative)

一個以上的機構共同合作，共享技術與資源以達到相同目標。

4、聯盟 (Consortium)

主要由組織中的聯合單位負責，旨在營造一互助合作的環境，提供會員持續性的政策建立及非會員可參與網站典藏的機制。

(三) 網站保存工具

若機關決定自行進行網站的保存計畫，目前已有一些網站典藏的相關工具可供機關參考。

1、蒐集(Acquisition)工具

例如：Heritrix、HTTack。

2、存取(Access)工具

例如：Wayback、WERA、NutchWAX。

3、管理(Curator)工具

例如：Web CuratorTool、PANDORA Digital Archiving System、

NetarchiveSuite。

本計畫所建置的雛型系統已整合了 Web Curarator Tool (含 Hertrix)與 Wayback 等工具，運作順利。

(四) 網站抓耙方式

機關可依需要選擇在伺服器端備份網站資料、瀏覽器端抓取網站畫面，或是使用抓耙軟體保存網站。

1、授權系統或伺服器端

從「授權系統或伺服器端」進行抓耙網站。採用此方式將有機會完整抓耙網站所有內容。但是，越來越多網頁是為配合使用者需求而動態產生，如瀏覽器的樣式、螢幕大小，桌上型電腦或智慧型電腦。因此，產生該保存何種版本之網站的議題。

2、瀏覽器端從「瀏覽器端」著手抓耙網站。採取此方式可採用快照(snapshot)軟體進行抓耙，但僅能抓耙到網站中無法變動的內容。

3、使用抓耙軟體

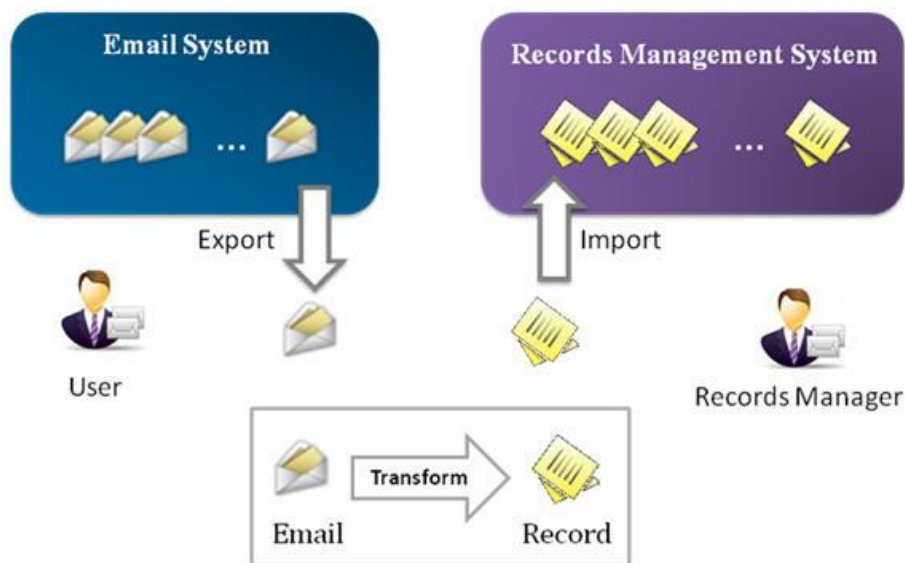
透過「使用抓耙軟體」保存網站。採取此方式或許能解決上述兩種方式之問題，但有遺失外部資源的疑慮，如資料庫產生之內容就無法囊括。

（五）選擇網站的保存格式

從「授權系統或伺服器端」著手保存網站是以 Mirror 方式完全備份原始網站檔案，此種方式保存下來的檔案格式會與檔案原來的格式完全相同。而從「瀏覽器端」的保存方式會以圖檔的方式將畫面完整的保存，檔案格式會以常見的 jpg、gif 及 png 等圖檔為主。若使用抓耙軟體保存網站，則建議能採用 WARC 國際標準，以符合國際間的發展趨勢。

附錄一 電子郵件保存系統安裝手冊

本分項計畫研擬在現有電子郵件系統中增加「保存」的功能，藉此將電子郵件轉換為電子檔案，進而匯入檔案管理系統中，以納入一般電子檔案管理的範疇，附圖 1- 1 為此電子郵件保存系統架構圖。



附圖 1- 1 電子郵件保存工具架構圖

由於使用者大多採用收信軟體或網頁瀏覽器作為收發信件之工具，因此本分項計畫分別針對此兩種方式進行系統實作，一種是使用 Outlook 擴充套件進行電子郵件保存，另一種則是使用 WebMail 進行電子郵件保存。

以下首先介紹 Outlook 擴充套件的安裝流程：

一、軟體清單

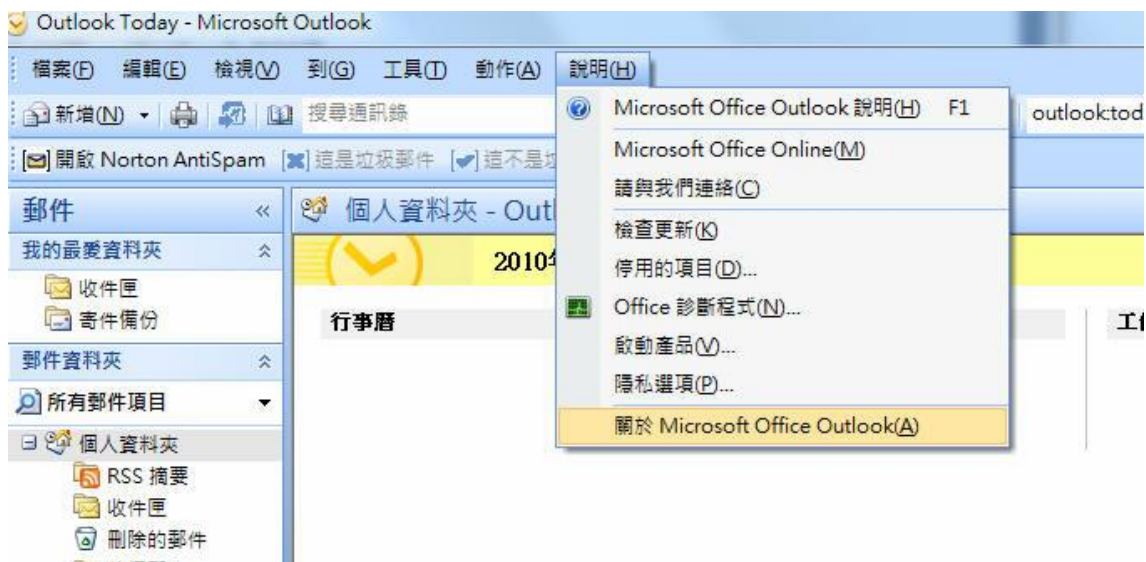
(一) Microsoft Office Outlook 2007

(二) ExportMSG

(三) outlookEMLandMSGconverter

二、安裝 ExportMSG 擴充套件

- (一) 首先必須確定 Outlook 的版本符合 Microsoft Office Outlook 2007 (12.0.6504.500) SP2 MSO (12.0.6425.1000)，如附圖 1-2 所示，點擊「關於 Microsoft Office Outlook(A)」的按鈕即可以確認 Outlook 的版本。



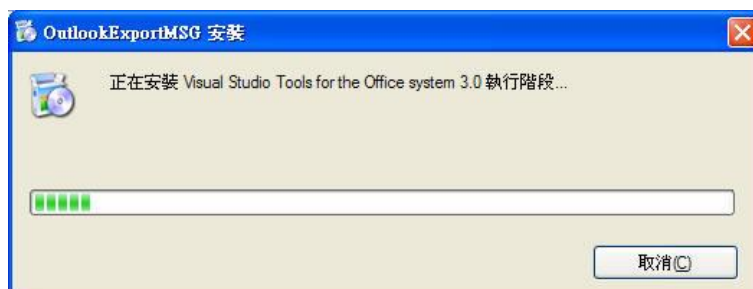
附圖 1- 2 Outlook 版本確認選項

- (二) 從取得之 ExportMSG 資料夾中，執行裡面的 setup.exe 進行安裝，並依照安裝精靈提示完成安裝步驟。如附圖 1- 3 中，點擊「接受」按鈕接受相關授權條款即可繼續安裝步驟。若是安裝環境有缺少 .NET Framework 3.5 或其他軟體，點擊「接受」按鈕這個步驟可能會需要重複數次。



附圖 1- 3 ExportMSG 安裝畫面

(三) 接下來請稍待片刻，等候相關軟體的安裝過程結束，安裝進度畫面如附圖 1- 4 所示。



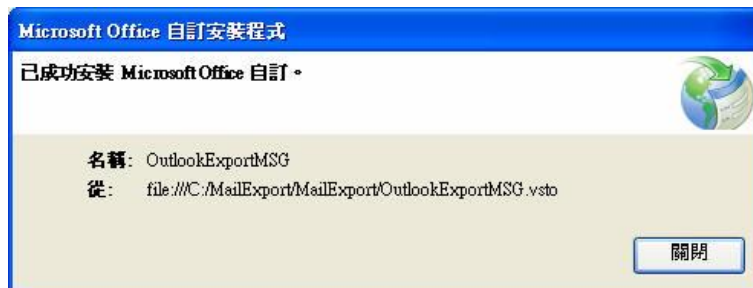
附圖 1- 4 ExportMSG 相關軟體安裝進度

(四) 相關軟體安裝完成後會出現 Office 自訂的安裝確認畫面，如附圖 1- 5 所示，點擊「安裝」按鈕即可繼續安裝步驟。



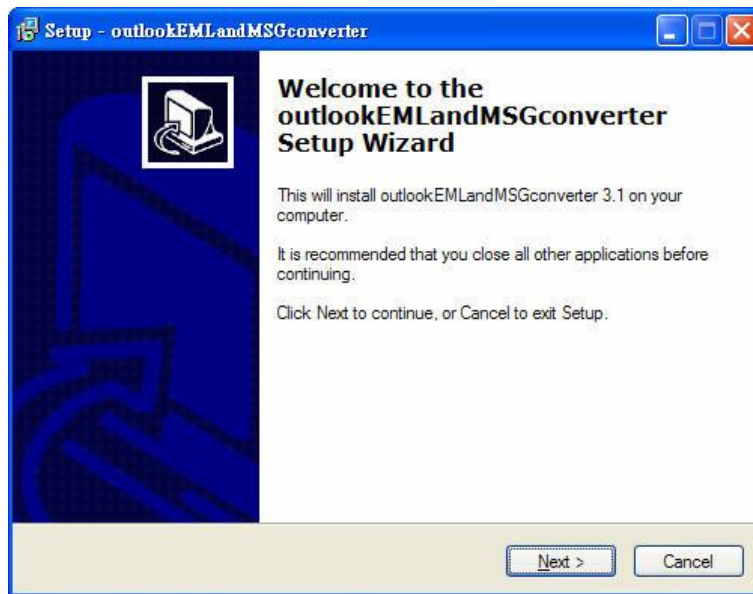
附圖 1- 5 ExportMSG Office 自訂安裝程式確認畫面

(五) 安裝完成後，點擊「關閉」按鈕即可繼續後續步驟，如附圖 1- 6 所示



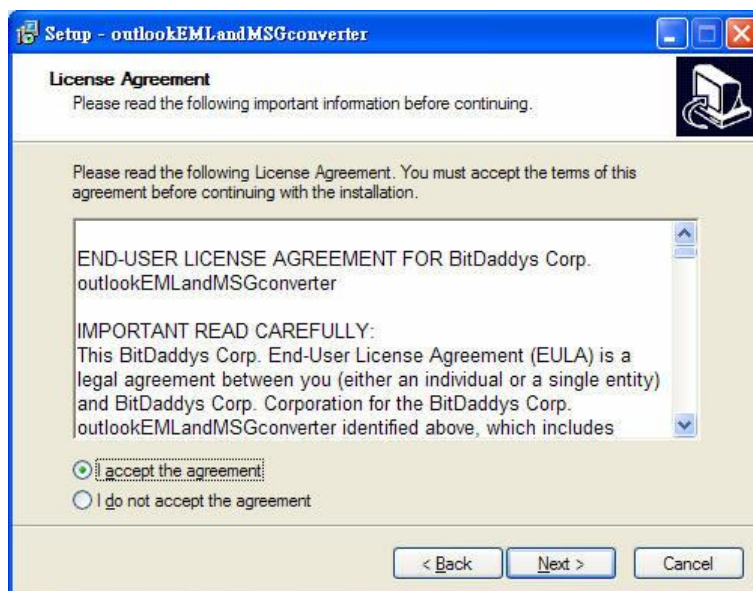
附圖 1- 6 ExportMSG Office 自訂安裝完成畫面

(六) 從取得之 outlookEMLandMSGconverter 資料夾中，執行裡面的 outlookEMLandMSGconverter-trial-setup.exe 進行安裝，並依照安裝精靈提示完成安裝步驟。如附圖 1- 7 中點擊「Next>」按鈕，即可繼續安裝步驟。



附圖 1- 7 OutlookEMLandMSGconverter 安裝畫面

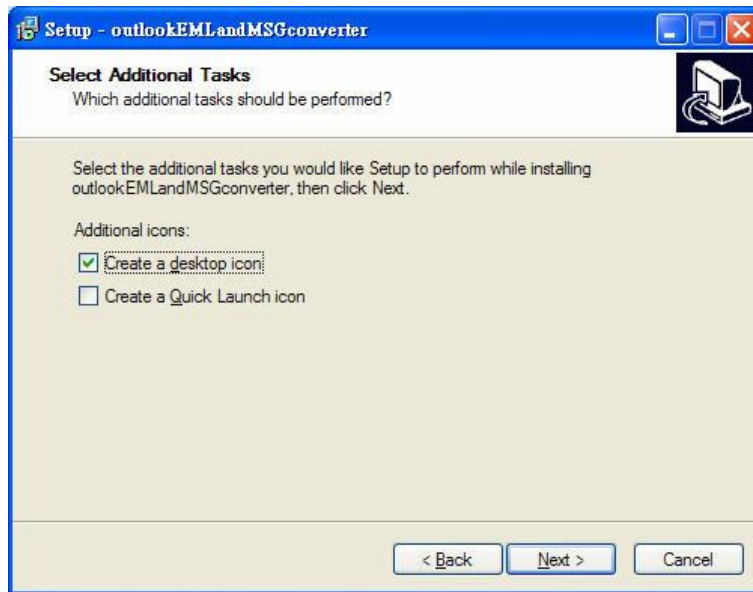
(七) 接下來需要接受 outlookEMLandMSGconverter 的授權條款，如附圖 1- 8 所示，勾選「I accept the agreement」之後再點擊「Next>」按鈕，即可繼續安裝步驟。



附圖 1- 8 OutlookEMLandMSGconverter 授權條款畫面

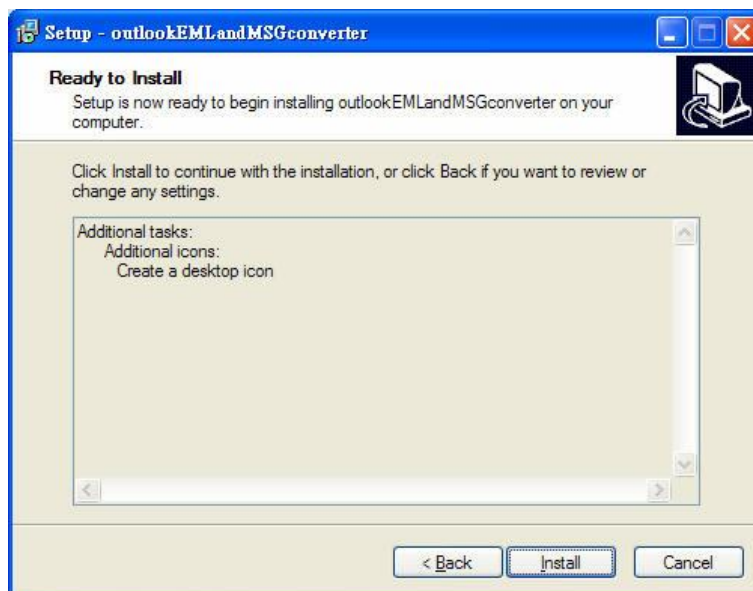
(八) 接受授權條款之後，會進入建立捷徑的畫面，如附圖 1- 9 所示，可視需要自行決定是否要建立其他的捷徑，勾選完成後點擊

「Next>」按鈕，即可繼續安裝步驟。



附圖 1- 9 OutlookEMLandMSGconverter 建立捷徑畫面

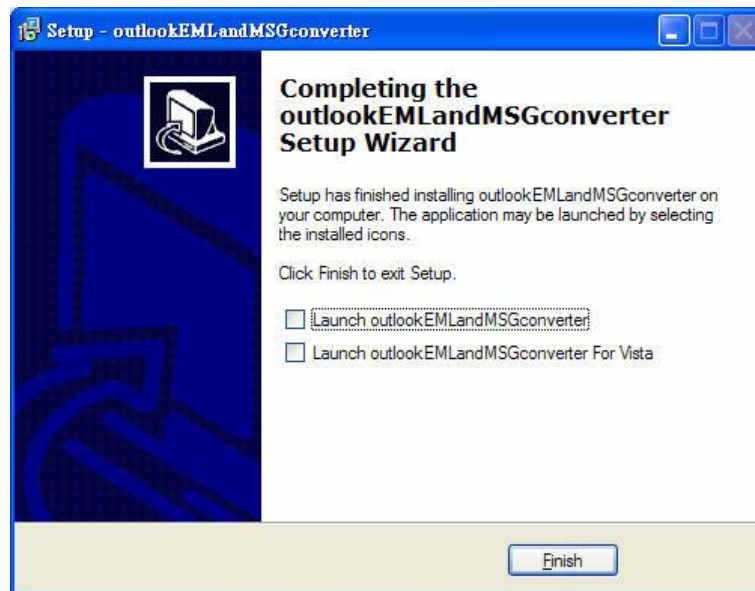
(九) 接下來會確認先前安裝的選項，如附圖 1- 10 所示，如果確認沒有問題，點擊「Install」即可進行軟體的安裝。



附圖 1- 10 OutlookEMLandMSGconverter 安裝確認畫面

(十) 安裝完成後，如附圖 1- 11 所示，點擊「Finish」後即完成安裝手續。隨後可使用 Outlook 安裝好的 MSGExport 轉出電子郵件的

msg 檔案及詮釋資料檔，並透過 outlookEMLandMSGconverter 將轉出的 msg 檔案轉換成 eml 檔，以供檔管系統後續的歸檔點收作業使用。



附圖 1- 11 OutlookEMLandMSGconverter 安裝完成畫面

附錄二 電子郵件保存系統操作手冊

電子郵件保存試作系統操作流程大體上可以分成兩個部分，第一個部份是從電子郵件軟體中匯出電子郵件及相關詮釋資料的步驟，第二部分則是將詮釋資料及相關的電子郵件匯入到檔管系統的步驟。

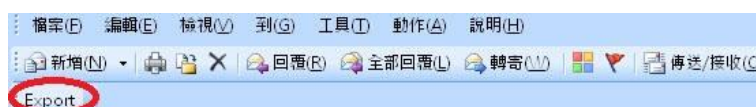
而在將電子郵件及其詮釋資料自電子郵件軟體匯出的第一個部分中，本計畫邀請網擎資訊依照常見之電子郵件軟體類型，實作了兩種電子郵件軟體的匯出工具，第一種是使用 Outlook 擴充套件進行電子郵件的匯出；第二種則是使用 WebMail 進行電子郵件的匯出。

在由檔管系統將詮釋資料暨相關的電子郵件匯入的第二個部分中，本計畫邀請了二一零零科技協助開發電子郵件匯入檔管系統的套件。

以下我們將介紹上述實作的操作步驟：

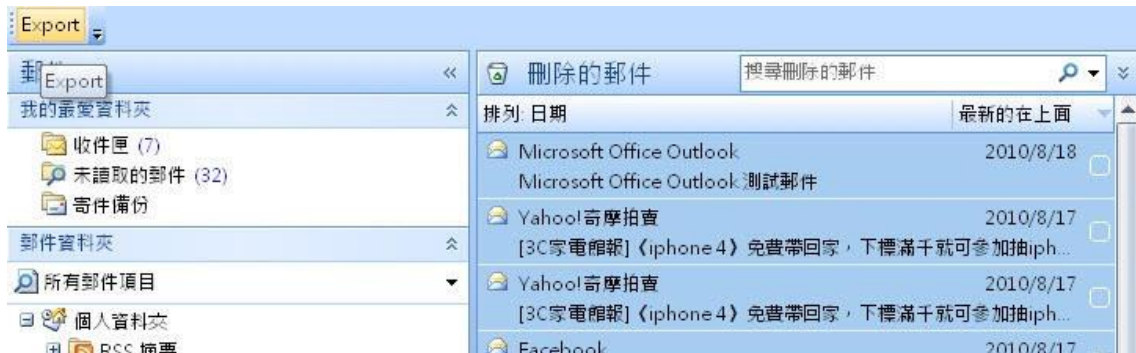
一、Outlook 擴充套件電子郵件保存操作範例

(一) 首先要開啟 Outlook，必須確定 Outlook 已按照附錄一的步驟完成相關擴充套件的安裝，在 Outlook 左上會出現 Export 的按鈕，如附圖 2- 1 所示。若無該按鈕，請按照附錄一的步驟完成相關擴充套件之安裝。



附圖 2- 1 Outlook 擴充套件按鈕畫面

(二) 確認 Outlook 中有 Export 按鈕之後，可以選擇需要匯出的電子郵件之後，點擊 Export 按鈕。如附圖 2- 2 所示。



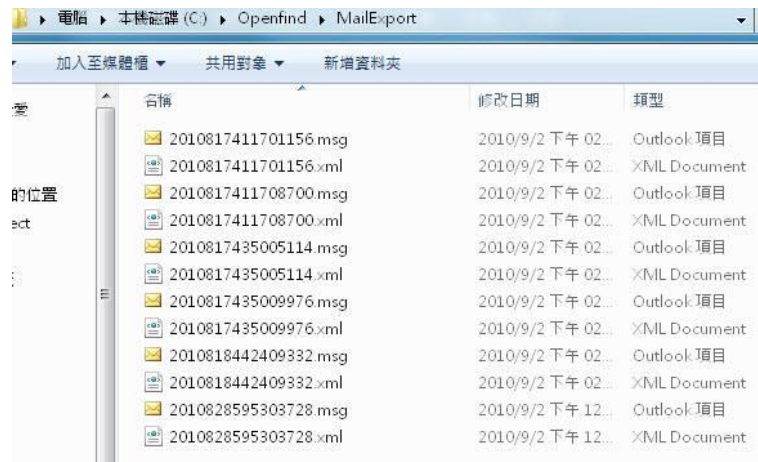
附圖 2- 2 Outlook 選取郵件畫面

(三) 按下「Export」按鈕之後，會跳出電子檔案需要的必要資料填寫視窗。如附圖 2- 3 所示。依照各欄位顯示之資料輸入該電子檔案所需之承辦人及文件資料之後，按「確認」的按鈕。

附圖 2- 3 電子郵件檔案詮釋資料填寫視窗

(四) 接著，Outlook 擴充套件會自動替你將電子郵件以 msg 的電子郵件格式匯出到 C:\Openfind\MailExport\資料夾裡面，並且將先

前輸入的資料以及其他相關資料匯入到 xml 檔案中，如附圖 2- 4 所示。但為了將這些檔案提供給後續的檔管系統匯入歸檔管理，這些 msg 的檔案格式必須轉換成 eml 格式才能作後續的匯入。

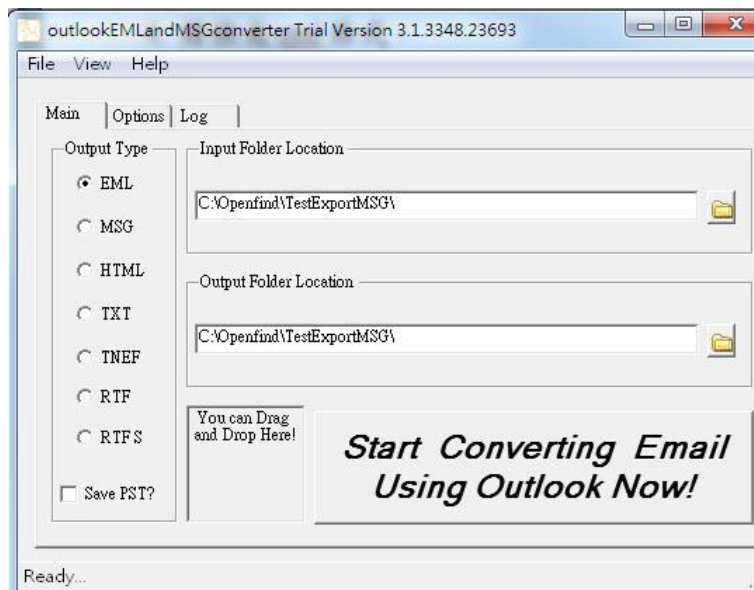


附圖 2- 4 匯出之電子郵件檔案

(五) 透過先前在附錄一安裝過的 outlookEMLandMSGconverter 工

具，可以將 msg 之電子郵件格式轉換成 eml 電子郵件格式。該工具的畫面如附圖 2- 5 所示，輸出類型(Output Type)選擇 EML，輸入(Input)及輸出資料夾位置(Output Folder Location)都填入預設的 Outlook 電子郵件匯出資料夾

C:\Openfind\MailExport\資料夾之後，點擊開始轉換(Start Converting Email Using Outlook Now!)的按鈕，即可將 msg 檔案轉換成 eml 格式。



附圖 2- 5 OutlookEMLandMSGconverter 工具畫面

(六) 如果有需要，可以直接開啟已存檔之 msg 或 eml 檔案，確認其內容是否正確。

二、WebMail 電子郵件保存操作範例

(一) 首先登入 WebMail 的系統，於欄位內輸入正確的帳號及密碼的資料後，點登入的按鈕進入電子郵件系統，如附圖 2- 6 所示。



附圖 2- 6 Webmail 登入畫面

(二) 根據需要進行電子郵件的搜尋，以尋找出需要匯出保存的電子郵件。如附圖 2- 7 所示。



附圖 2- 7 Webmail 電子郵件搜尋畫面

(三) 依據搜尋結果，如附圖 2- 8 所示，在此畫面中勾選要匯出的電子郵件，並點擊「工具選單」按鈕。



附圖 2- 8 Webmail 電子郵件搜尋結果

(四) 隨後會出現工具選單的下拉式列表，如附圖 2- 9 所示，點擊「批次匯出」按鈕以進行相關的匯出設定。



附圖 2- 9 Webmail 工具選單選項畫面

(五) 在檢索結果的畫面中，會出現匯出檔案格式的選項，如附圖 2- 10 所示，選擇 EML 格式後，點擊「匯出」按鈕。



附圖 2- 10 Webmail 匯出檔案格式選項畫面

(六) 按下「匯出」按鈕之後，會跳出電子檔案需要的必要資料填寫視窗。如附圖 2-11 所示。依照各欄位顯示之資料輸入該電子檔案所需之承辦人及文件資料之後，按「確認送出」的按鈕。

新增郵件詮釋資料

承辦人姓名	<input type="text"/>	* 姓名
承辦人單位	<input type="text"/>	* 單位
承辦人職稱	<input type="text"/>	* 職稱
文別	<input type="text" value="Z"/>	* 必要欄位：1令、2函、3公告、4開會通知單、5簽、6簽稿會核單、7會銜公文會辦單、8公文時效統計、9呈、A咨、Z其他
密等	<input type="text" value="1"/>	* 代碼：1普通、2密、3機密、4極機密、5絕對機密
保存年限	<input type="text" value="1"/>	* 依機關檔案保存年限區分表所定年限，99 表示永久保存，?表示未判定，定期則以數字表示之
分類號	<input type="text"/>	* 指公文依檔案分類表所載類目名稱之文、數字組合號碼

附圖 2-11 Webmail 電子郵件檔案詮釋資料填寫畫面

(七) 接下來 Webmail 會自動替你將電子郵件以 eml 的電子郵件格式匯出，並且將先前輸入的資料以及其他相關資料匯入到 xml 檔案中，以供後續的檔管系統匯入歸檔管理，如附圖 2-12 所示。如果需要下載相關檔案，可以點擊「前往工作紀錄列表」按鈕。



附圖 2-12 Webmail 電子郵件批次匯出畫面

(八) 在工作紀錄列表畫面中，在匯出工作最右側的執行欄位中，有提

供檔案下載的按鈕，如附圖 2- 13 所示。可以點擊「 」按鈕以進行下載。



附圖 2- 13 Webmail 工作紀錄列表畫面

(九) 按下「 」按鈕之後，會跳出儲存檔案的視窗，如附圖 2- 14 所示，可依需要存放到指定的目錄，以供檔管系統匯入使用。



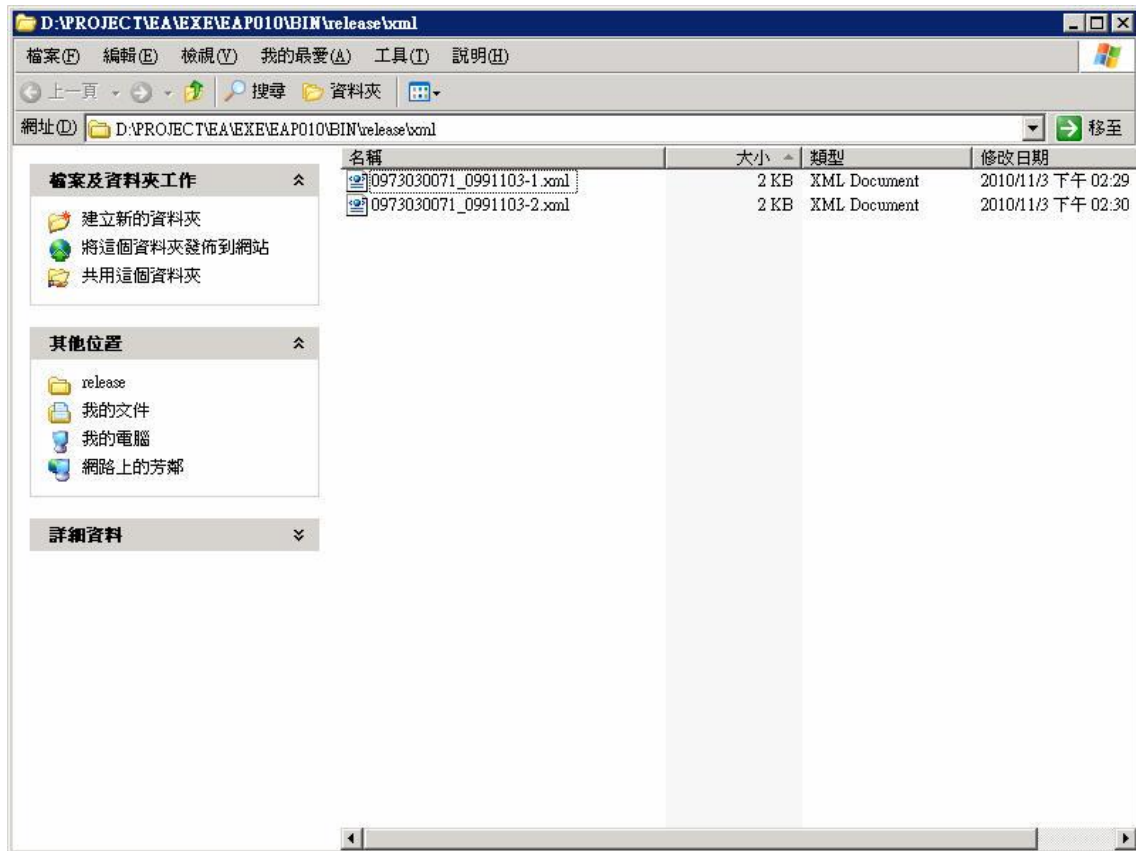
附圖 2- 14 Webmail 檔案下載畫面

三、檔管系統匯入操作範例

(一) 首先需要將欲匯入的電子郵件詮釋資料檔案搬移至指定的 xml

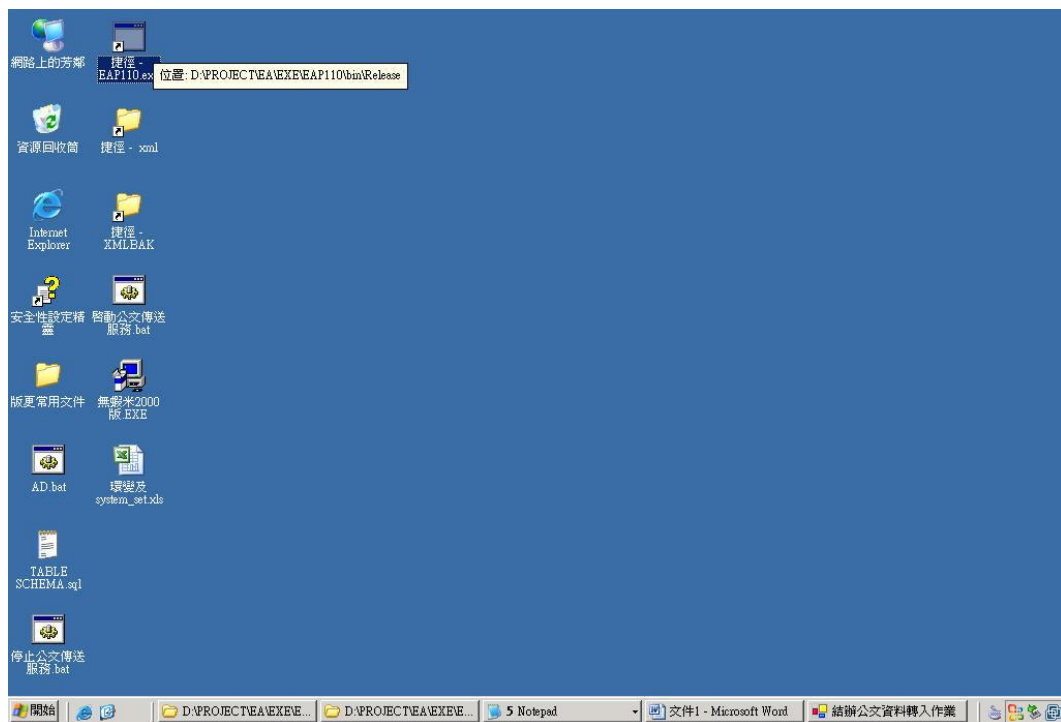
目錄，如附圖 2- 15 所示。目前匯入的詮釋資料受限於二一零零系統限制，須依照機關檔案作業手冊附件二之格式，轉成歸檔以及結案的兩個 xml 檔案，才能符合流程進而被系統接受處理，且

匯出之電子郵件 eml 檔案無法匯入到系統中。日後因應法規，此部分可能會有所更動。



附圖 2- 15 電子郵件匯入資料夾

(二) 接著執行 EAP110.EXE 這支程式，如附圖 2- 16 所示，執行後會出現結辦公文資料轉入作業的視窗，但是執行完成後才會顯示畫面，如附圖 2- 17 所示。顯示畫面後會自動關閉。

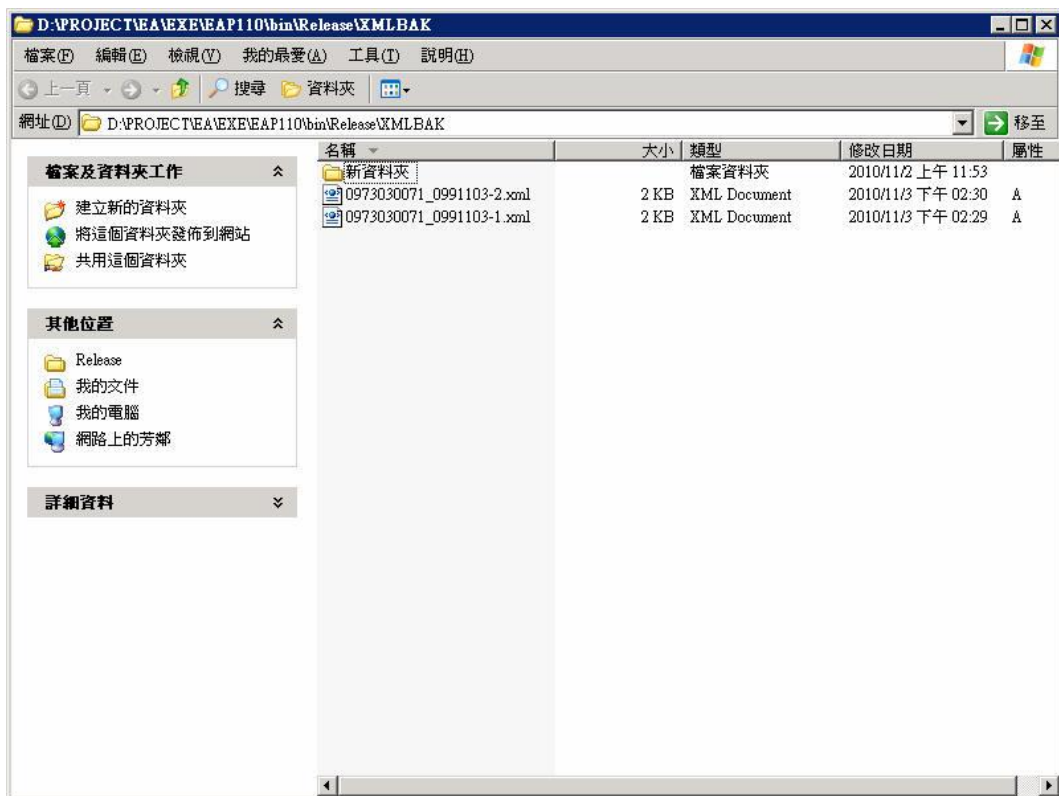


附圖 2- 16 結辦公文資料轉入作業程式



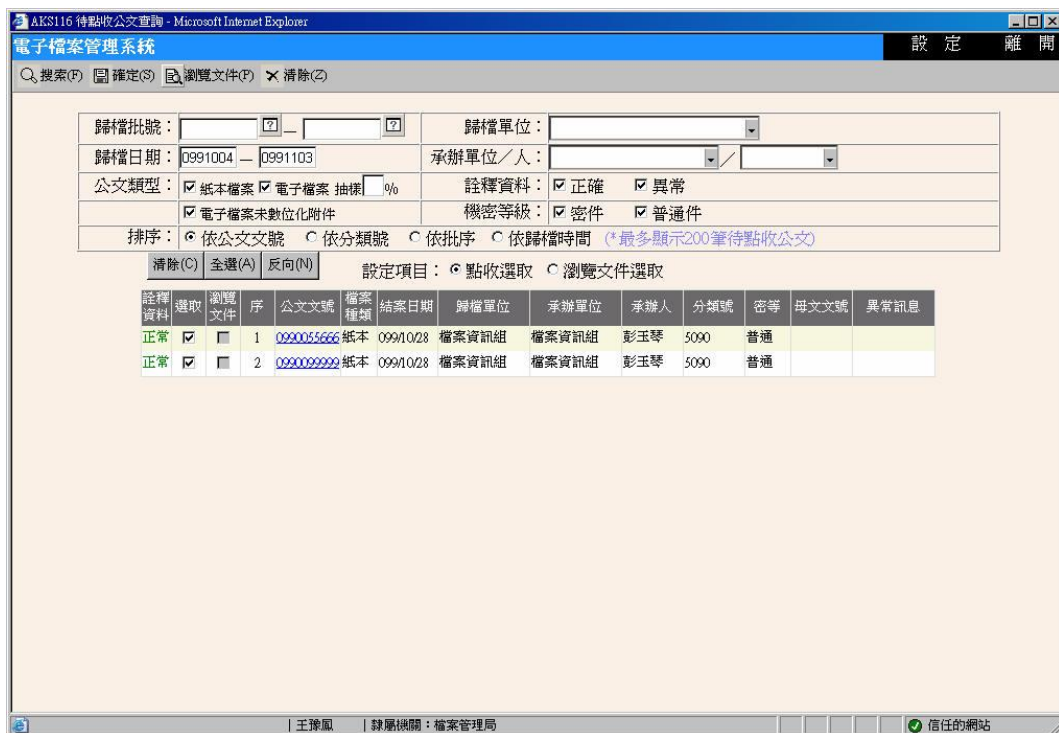
附圖 2- 17 結辦公文資料轉入作業畫面

(三) 此外可以從匯入備份的 XMLBAK 資料夾，觀察電子郵件詮釋資料檔案搬移的情況，來了解哪些檔案已被系統匯入至檔管系統之中，如附圖 2- 18 所示。



附圖 2- 18 電子郵件匯入備份資料夾

(四) 在完成匯入之後，可以從檔管系統中的待點收公文查詢功能中，查詢到已經匯入的資料。如附圖 2- 19 所示。



附圖 2- 19 檔案管理待點收公文查詢功能

(五) 接下來可以對待點收的電子郵件進行點收作業，如附圖 2- 20 所示。

附圖 2- 20 檔案管理系統檔案點收及退文作業

(六) 點收完成後會顯示點收完成的訊息，如附圖 2- 21 所示。

AKT116 點收及退文作業 - Microsoft Internet Explorer

電子檔案管理系統 線上說明 設定 離開

點收(R) 清除(Z) 待點收公文查詢(Q) 點收清單(B) 退文清單(P)

公文文號: 點收 ☒ 退文 註記: 更改註記

基礎維護(H) 確定(S) 取消(Q) 備註:

退文至(單位): ☐ 承辦 ☒ 歸檔 歸檔日期:

☐ 附件抽存 附件預計歸檔日期:

清除(C) 全選(A) 反向(N)

選擇	取消	註記	序	公文文號	檔案 種類	歸檔單位	承辦單位	作業 別	退文 類型	點收 / 退文註記	附件 抽存	附件預計 歸檔日期	實際歸檔 日期	備註	異常訊息
點收作業成功															

您已成功點收公文 1 份(點收批號: 0991103002), 退文 0 份。

新增 | 王豫鳳 | 登錄機關: 檔案管理局 | 信任的網站

附圖 2- 21 檔案管理系統點收結果畫面

(七) 隨後可以透過查詢觀看該歸檔資料詳細內容, 如附圖 2- 22 所示。

AKM330 檔案目錄維護作業 - Microsoft Internet Explorer

電子檔案管理系統 線上說明 設定 離開

儲存(S) 刪除(D) 取消(Z)

公文文號: 0990099999 待編目: 收創文日期: 參照文號: 併件 ☐ 併案查詢

版本別 年度號 分類號 案次號 卷次號 目次號

檔號: 03 5090 編卷 保存年限: 1

類目名稱: 其他事項 案名:

案由: League of Legends Newsletter #16 - August 2010 其他案由

主要來文者:

來文日期: 來文字號: 字第 號 >>

主要發文者: 檔案管理局 來發受文明細

發文日期: 0991025 發文字號: 檔資 0990099999 號 >> 檔資字 0990099999

受文者: 僑務委員會 辦畢日期: 0991025 辦畢時間: 文件產生日期: 0991025

庫房: 檔案庫房 本文儲存位置: 應用限制: 開放

密等: 普通 密件另存流水號: 應解密日期:

解密條件: 解密日期:

文別: 函 本別: 正本 檔案種類: 紙本檔案

承辦單位: 檔案資訊組 承辦人: 彭玉琴

數量: 6 頁 媒體型式: 紙本 微縮編號:

保存狀況: 良好 ☐ 毀損無法修復

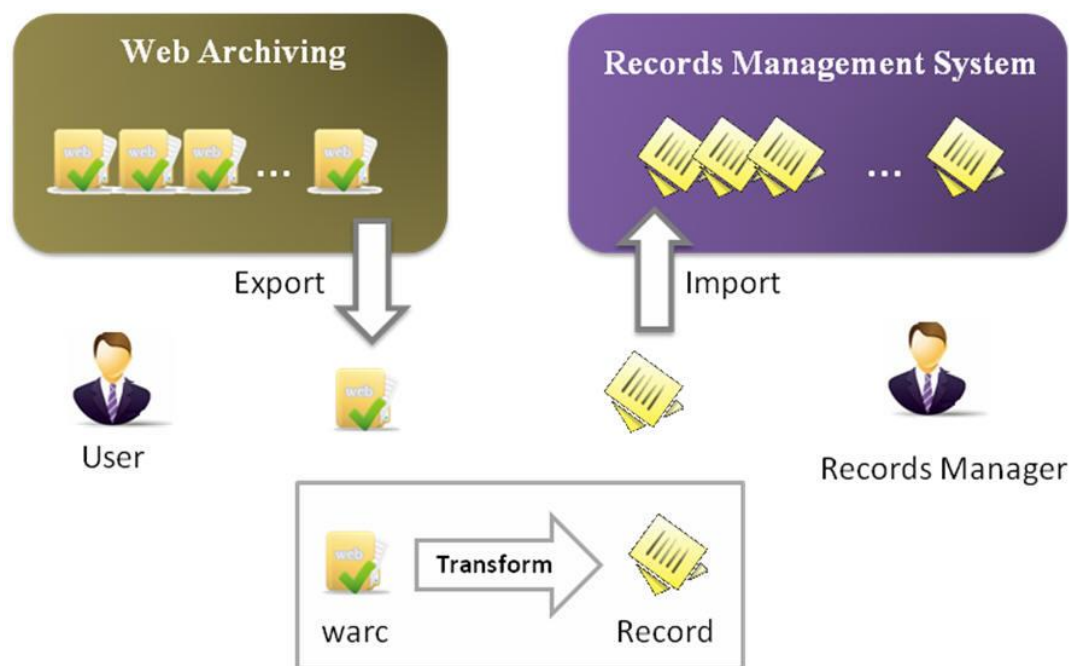
主題項 附註項 附件明細 電子檔資訊

修改 | 王豫鳳 | 登錄機關: 檔案管理局 | 信任的網站

附圖 2- 22 檔案管理系統檔案詳細資料畫面

附錄三 網站保存系統安裝手冊

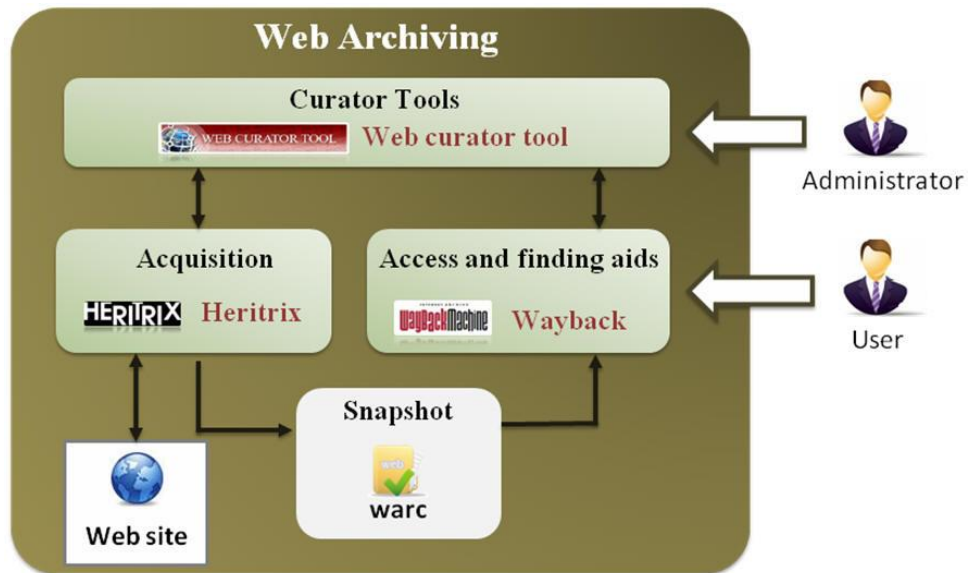
本分項計畫建議之網站保存系統，如附圖 3- 1 所示，將網站匯出後，透過著錄電子檔案所應具有之詮釋資料，將網站轉換為電子檔案，進而匯入檔案管理系統中，以納入一般電子檔案管理的範疇。



附圖 3- 1 網站保存示意圖

本手冊將依據附圖 3- 2 所示架構，提供網站保存系統之建置步驟。其中，管理者(Administrator)是指負責執行網站典藏工作的人員，而使用者則是指瀏覽網站資料者。換言之，管理者可以透過管理工具(Curator Tools)設定系統自動抓取與呈現網站資料，而使用者則是透過存取工具(Access and finding aids)瀏覽有興趣的網站資料。

以下針對 WCT 及 Wayback 兩套網站資料保存的相關工具提供詳細之安裝步驟，以減少檔案管理人員在安裝系統的過程中所遭遇到的阻礙或困難。



附圖 3- 2 網站保存工具架構圖

以下將介紹相關軟體的安裝流程：

一、 軟體清單

(一) Mandriva Linux 2010 Spring Free

(二) Tomcat 6.0.29

(三) MySQL 5.1.42

(四) Java Open SDK 1.6.0

(五) Web Curator Tool 1.5 GA

(六) Wayback 1.4.2

二、安裝 Mandriva Linux 作業系統

(一) 請至 Mandriva Linux 的官方網站下載 Mandriva Linux 2010 Spring Free 的 DVD 映像檔，並且將該映像檔燒錄成 DVD 光碟片。可於光碟上註記 DVD 光碟片所燒錄之內容，以便後續使用時易於辨識。

(二) 在電腦開機後，若原先已安裝有作業系統，請進行相關資料之備份保存。

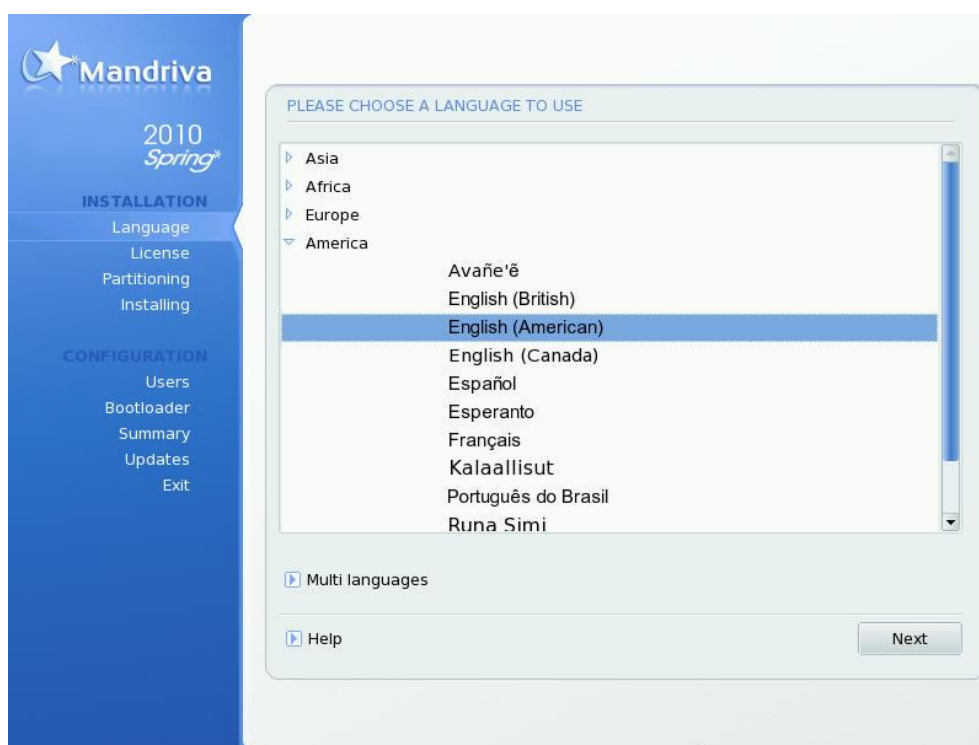
(三) 在開機畫面中進入 BIOS 畫面，設定從光碟機開機，並將光碟放入光碟機中，儲存並離開 BIOS 之後即可進入安裝畫面。

(四) 如附圖 3- 3 所示，請使用上下鍵選擇第二項 "Install Mandriva Linux 2010 Spring"，確認選取之後按下 Enter 鍵以進行 Mandriva Linux 的安裝與設定。



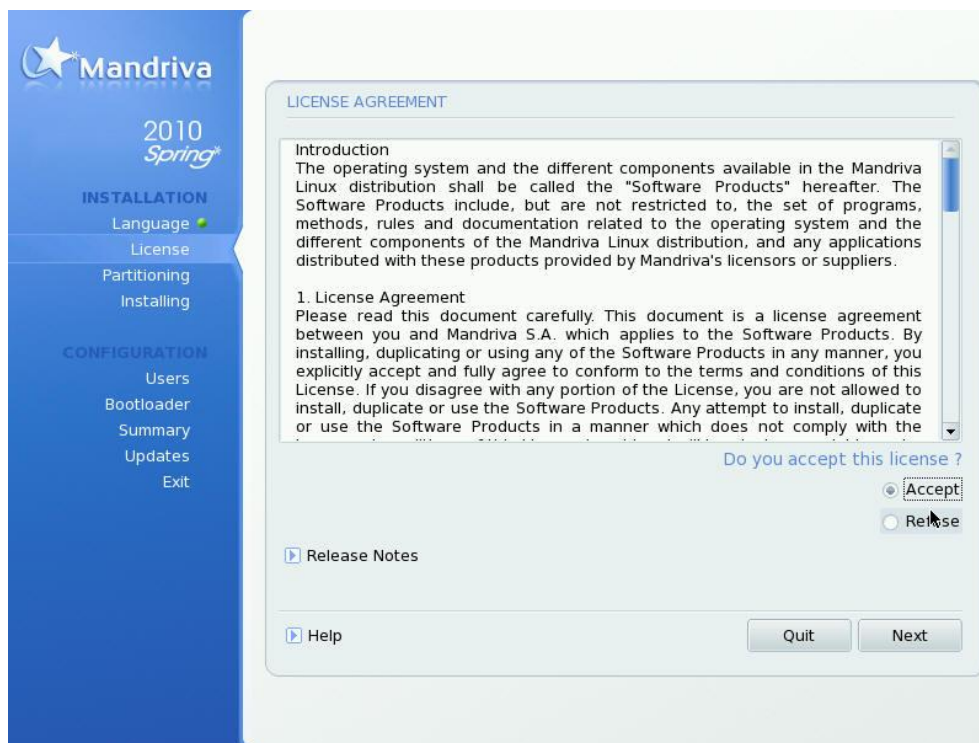
附圖 3- 3 Mandriva Linux 起始安裝畫面

(五) 由於 Web Curator Tool 在執行的過程中，會使用英語系國家的時間格式，若是系統預設語系不是英文的情況下，會導致 Web Curator Tool 無法正常地進行排程抓取網站的相關作業。為了避免此問題，需如附圖 3- 4 所示之系統預設語系設定畫面，選擇英文(English)作為預設的系統語系。完成後點擊下一步 (Next)進行下一個安裝步驟。



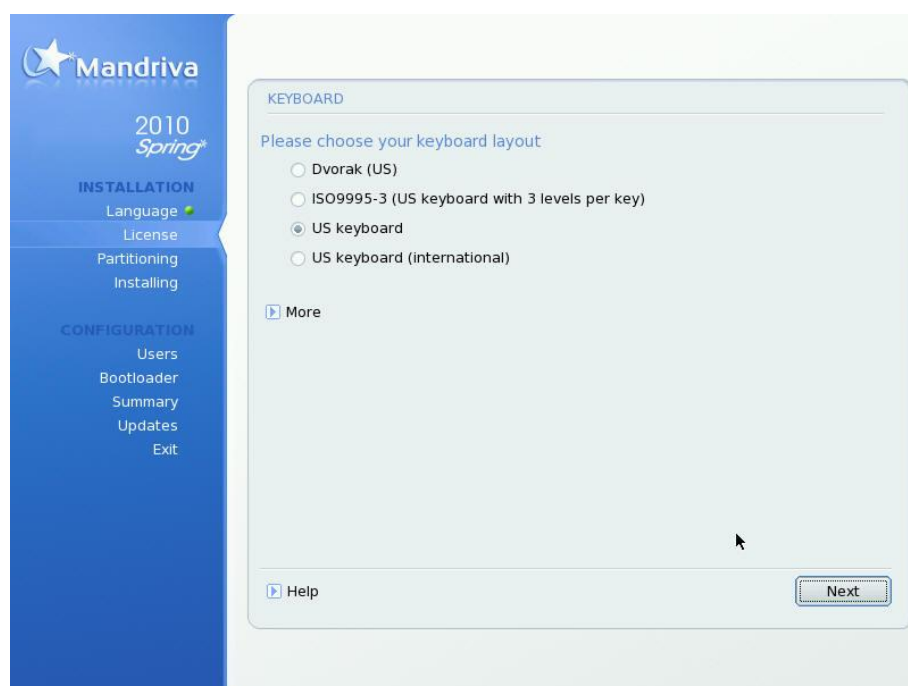
附圖 3- 4 系統預設語系設定畫面

(六) 附圖 3- 5 所示為 Mandriva Linux 之授權條款，若對於條款內容沒有無法接受之處，即可選擇同意(Accept)並點擊下一步(Next)的按鈕進行後續的安裝步驟。



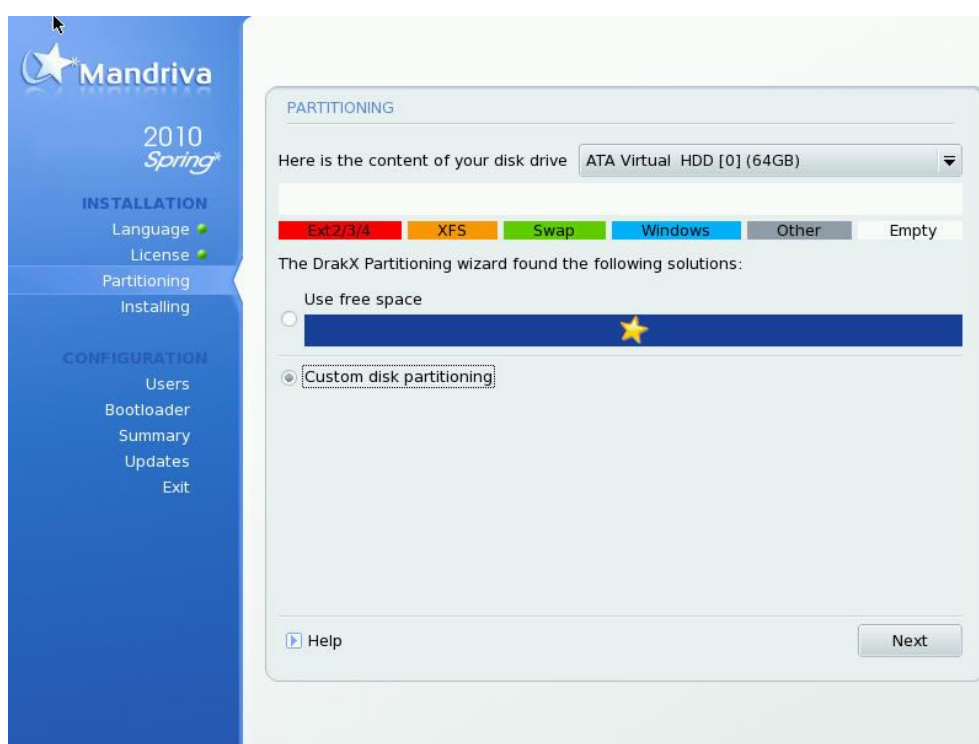
附圖 3- 5 授權條款畫面

(七) 接下來的步驟是選擇鍵盤的類型，如附圖 3- 6 所示選擇一般美式鍵盤(US keyboard)即可。完成後點擊下一步(Next)進行下一個安裝步驟。



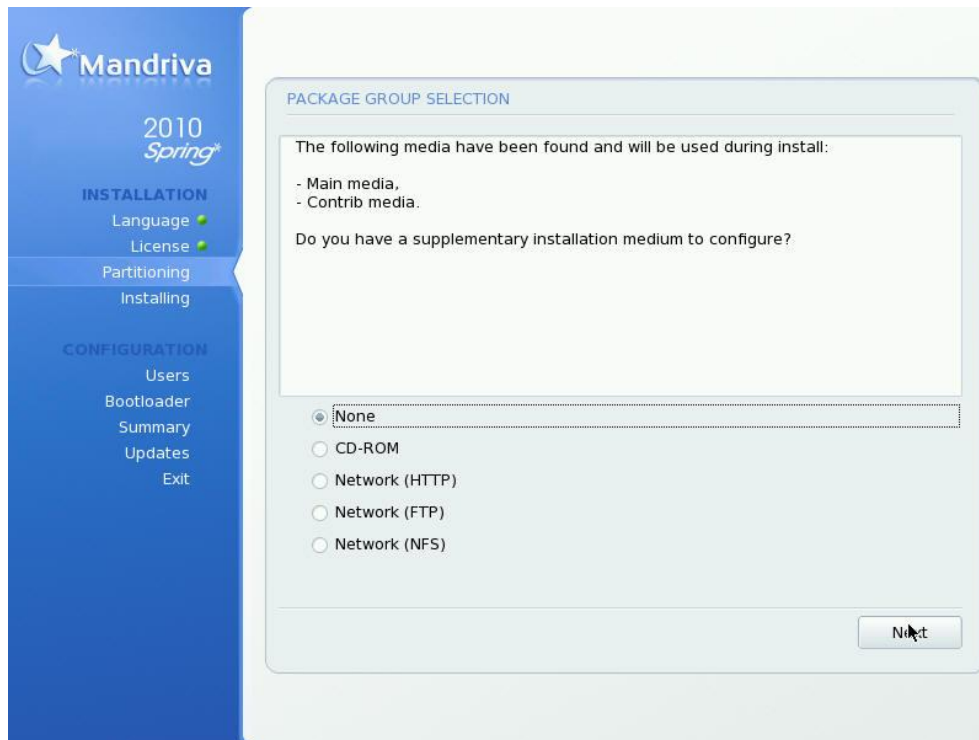
附圖 3- 6 鍵盤類型選擇畫面

(八) 接下來的步驟是設定磁碟分配的方式，若有自行設定之需求，可如附圖 3-7 所示，選擇自訂磁碟分配(Custom disk partitioning)的選項，將磁碟分配的方式根據需求來進行客製化的設定。一般建議 Swap 之空間設定為隨機存取記憶體(RAM)容量的兩倍大小，剩餘空間可全部分給根目錄。完成後點擊下一步(Next)進行下一個安裝步驟。



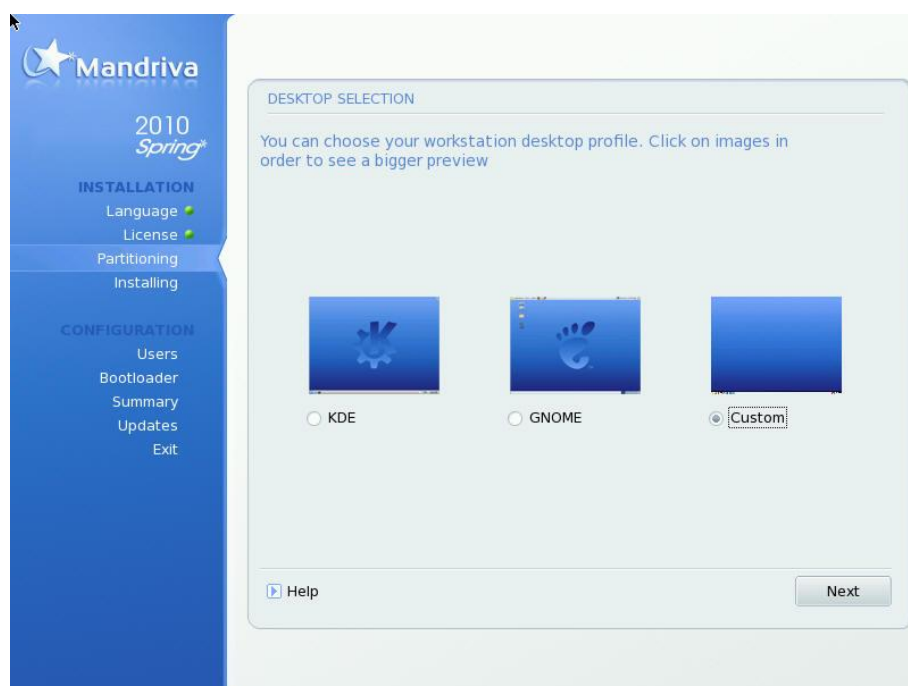
附圖 3- 7 磁碟空間分配畫面

(九) 接下來的步驟是設定安裝檔案的來源及方式，通常安裝來源為燒錄出來之 DVD 光碟片。因此，可如附圖 3- 8 所示，選擇沒有其他安裝來源(None)。完成後點擊下一步(Next)進行下一個安裝步驟。



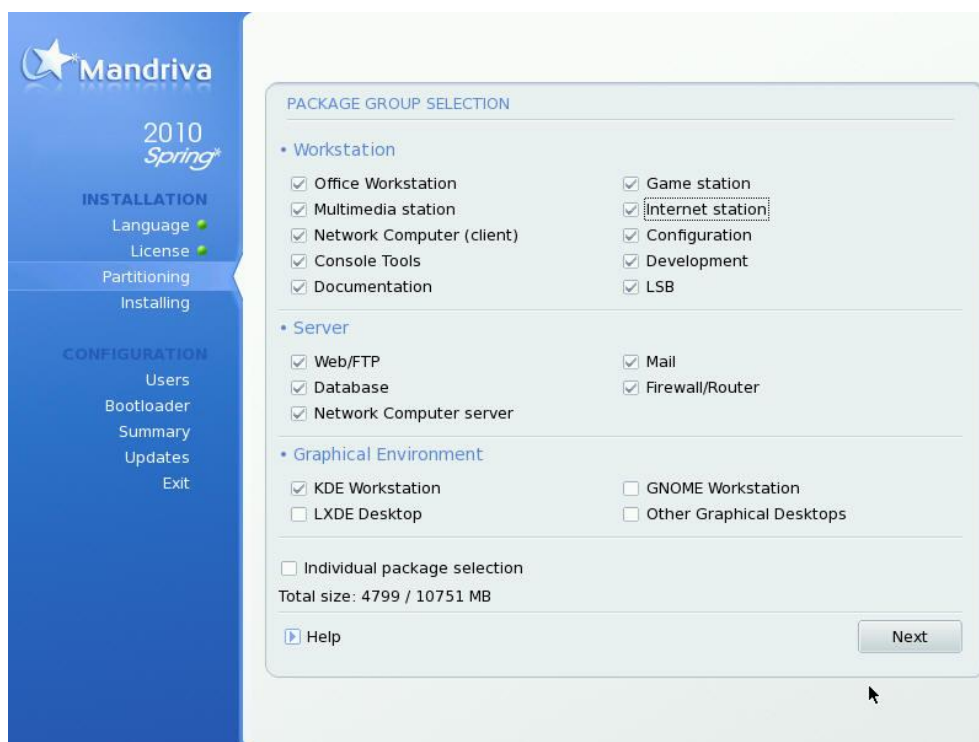
附圖 3- 8 安裝來源選擇畫面

(十) 接下來的步驟是設定桌面的類型，如附圖 3- 9 所示，請選擇自訂安裝桌面(Custom)。完成後點擊下一步(Next)進行下一個安裝步驟。



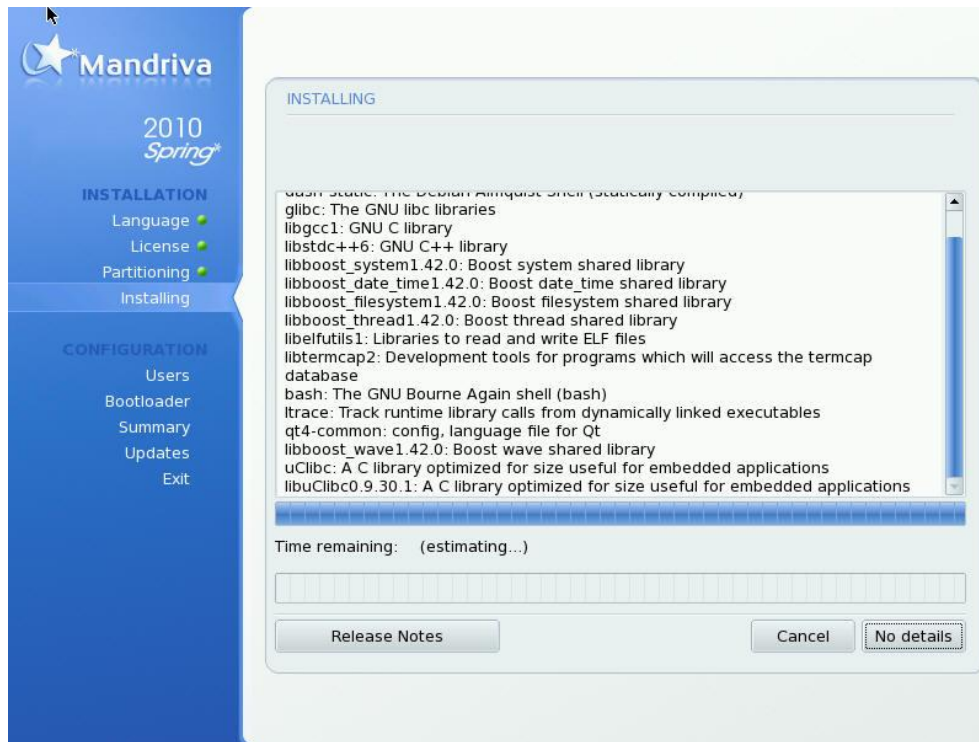
附圖 3- 9 桌面選擇畫面

(十一) 接下來的步驟是設定套件安裝的種類，請依附圖 3- 10 所示，勾選 KDE 工作站(KDE Workstation)以及其他畫面上已勾選之套件選項。完成後點擊下一步(Next)進行下一個安裝步驟。



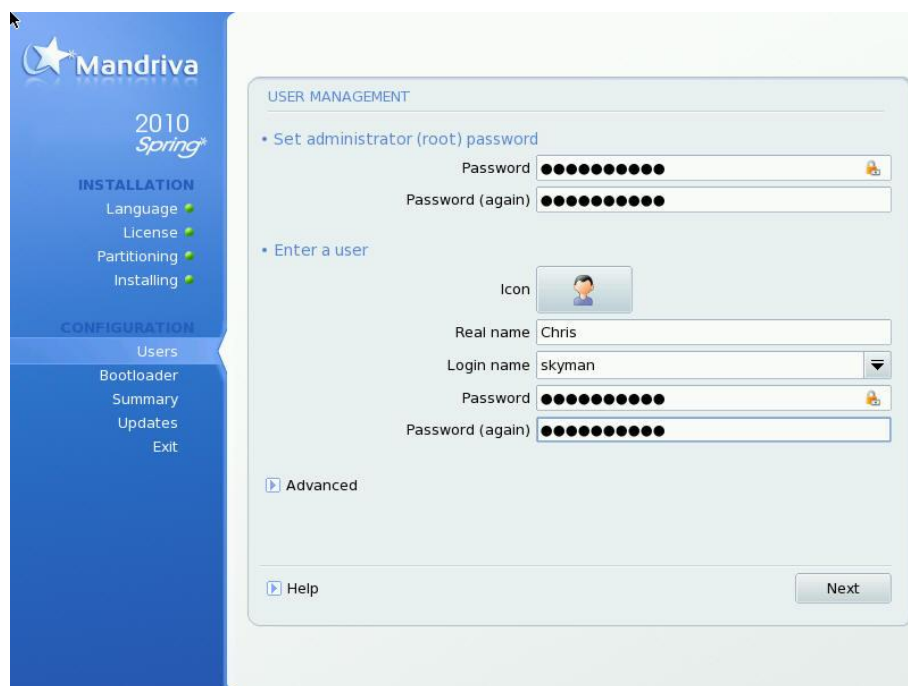
附圖 3- 10 套件安裝選擇畫面

(十二) 選擇套件的下一步(Next)之後，隨即會進入安裝畫面，如附圖 3- 11 所示。安裝所需時間依據電腦設備之間的差異以及套件選擇之多寡而有所差異，安裝程式估計之所需時間(Time remaining)及進度會顯示在畫面上以供參考，請耐心等待。



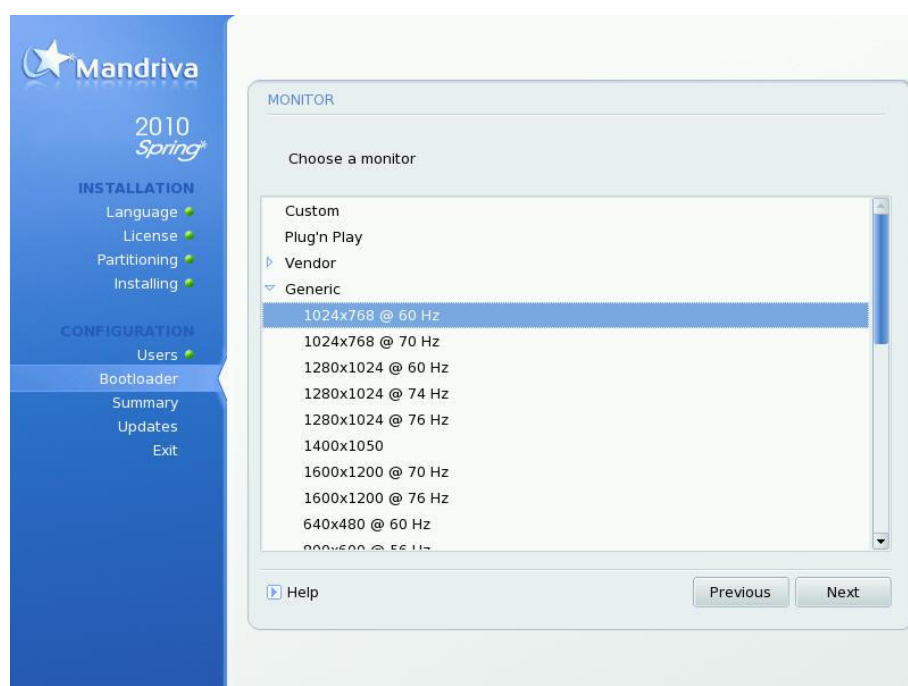
附圖 3- 11 系統安裝畫面

(十三) 系統安裝完成之後，接下來的步驟需要設定管理者密碼(Set administrator [root] password)，並且新增一個預設的使用者帳號，如附圖 3- 12 所示。密碼及帳號可以照個人喜好或實際資訊加以設定，但請妥善保管或牢記帳號登入名稱(Login name)及對應之密碼>Password)，以免日後無法進入系統進行操作。完成後點擊下一步(Next)進行下一個安裝步驟。



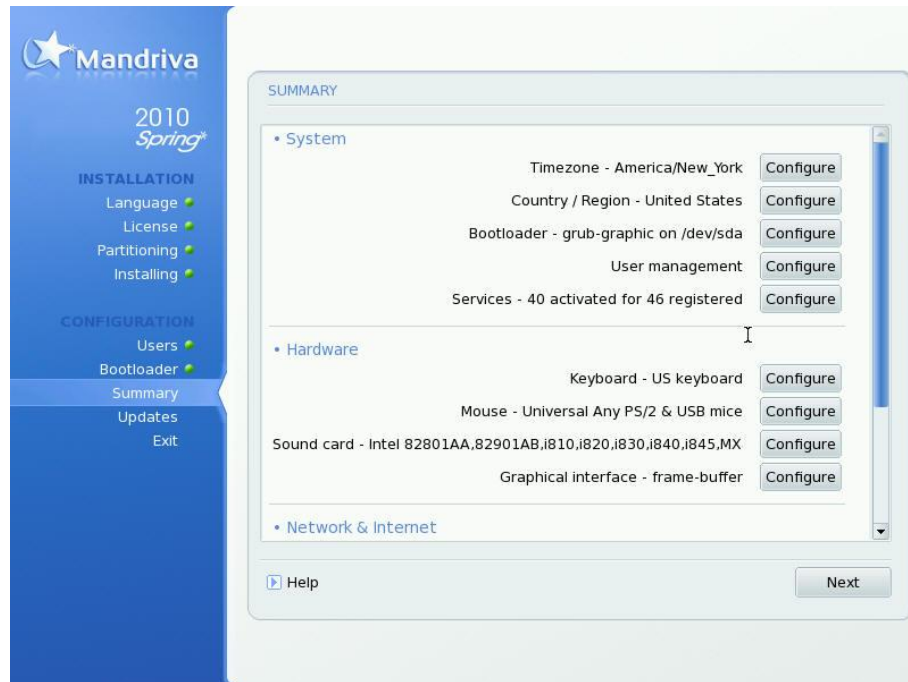
附圖 3- 12 使用者資訊設定

(十四) 接下來的步驟是調整螢幕解析度的種類及更新頻率，如附圖 3- 13 所示，可根據喜好選擇預設的解析度以及更新頻率。完成後點擊下一步(Next)進行下一個安裝步驟。



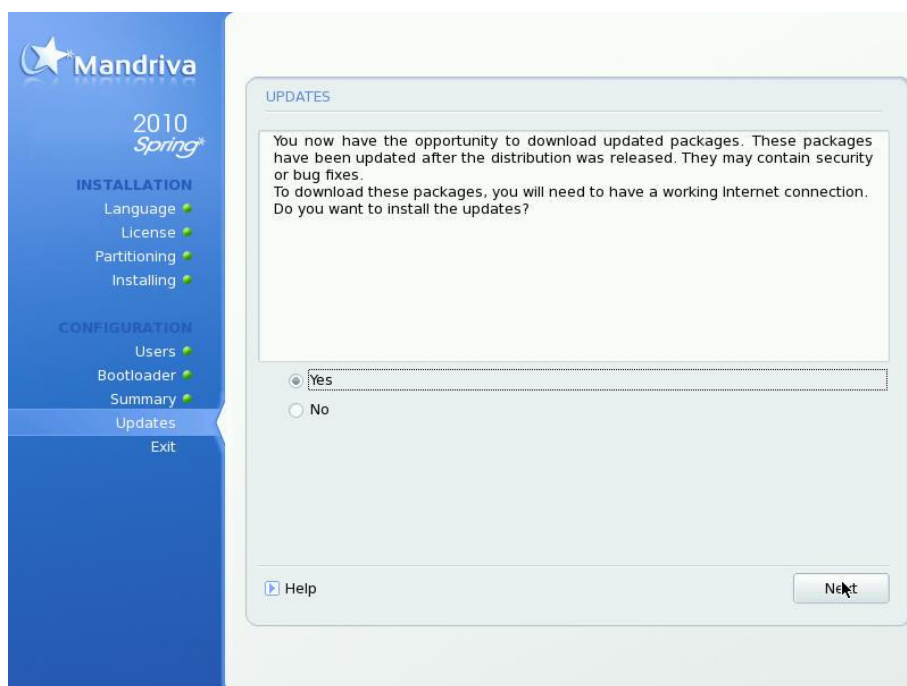
附圖 3- 13 螢幕參數設定畫面

(十五) 安裝步驟到此已經完成，如果有需要設定網路，如附圖 3- 14 所示，可以在此預先點擊設定(Configure)加以調整。完成後點擊下一步(Next)進行下一個安裝步驟。



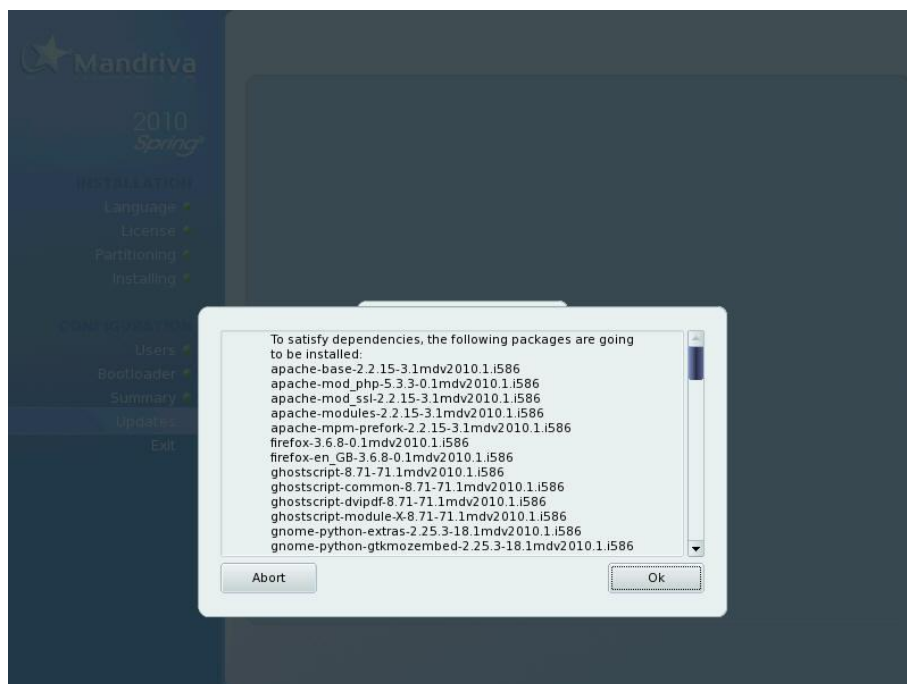
附圖 3- 14 系統設定畫面

(十六) 上一步驟中的網路相關設定若已設定完成，則在此可直接更新系統，如附圖 3- 15 所示，請選擇 Yes；如暫時無網路，則選 No，完成後點擊下一步(Next)。系統在重新開機後，即完成安裝。



附圖 3- 15 系統更新畫面

(十七) 上一步驟中如選擇要進行套件更新，即如附圖 3- 16 所示，點擊 Ok 鈕後即開始更新系統，系統將會在更新完成之後自動重新啟動，重開機後即完成安裝。



附圖 3- 16 安裝套件更新畫面

三、 安裝 Tomcat

(一) 首先需要下載 Tomcat 6.0.29 的壓縮檔，將該壓縮檔解壓縮之後，把解壓縮後的資料移動到自訂的資料夾之下，如附圖 3- 17 所示，此例是將解壓縮後的資料移動到/home 的資料夾下，並解壓縮完之後建立一個名為/home/tomcat 的連結。

```
[root@WCT~]# wget \
http://ftp.nsysu.edu.tw/Apache/tomcat/tomcat-6/v6.0.29/bin/apache-tomcat-6.0.29.tar.
gz
[root@WCT~]# tar zxvf apache-tomcat-6.0.29.tar.gz
[root@WCT~]# mv apache-tomcat-6.0.29 /home/
[root@WCT~]# ln -s /home/apache-tomcat-6.0.29 /home/tomcat
```

附圖 3- 17 Tomcat 下載及解壓縮指令

四、 安裝 MySQL

(一) 請將燒錄的 Mandriva Linux DVD 光碟片放入光碟機內，並藉由 Mandriva Linux 本身提供的套件安裝程式 urpmi 來安裝 MySQL。安裝完成後，可以透過 chkconfig 指令來設定在開機的過程中啟動 MySQL，接著即可新增 Web Curator Tool 所需要的資料庫，同時建立且授權該資料庫給特定使用者。注意，password 是使用者自訂給 usr_wct 使用者的密碼，而 newpassword 是給 MySQL 使用者 root 帳號所使用的密碼，如附圖 3- 18 所示。


```

[root@WCT~]# urpmi MySQL
[root@WCT~]# /etc/inint.d/mysql start
[root@WCT~]# chkconfig -add mysqld
[root@WCT~]# mysql -u root
mysql> create DATABASE DB_WCT;
mysql> \u DB_WCT
mysql> create user usr_wct@localhost identified by 'password';
mysql> grant all on DB_WCT.* to usr_wct@localhost;
mysql> quit
[root@WCT~]# mysqladmin -u root password newpassword

```

附圖 3- 18 MySQL 安裝指令

五、 安裝 Web Curator Tool

(一) 下載 Web Curator Tool 1.5 GA 壓縮檔並解壓縮後，將解開後 etc 資料夾中的 jta.jar 檔案複製到 Tomcat 的 lib 資料夾，再將 war 資料夾中所有檔案放到 Tomcat 的 webapps 資料夾，如附圖 3- 19 所示。

```

[root@WCT~]# wget \
http://downloads.sourceforge.net/project/webcurator/webcurator/1.5GA/wct-package-
bin-v1.5.tar.gz?r=http%3A%2F%2Fsourceforge.net%2Fprojects%2Fwebcurator%2Ff
iles%2F&ts=1283161542&use_mirror=ncu
[root@WCT~]# tar zxvf wct-package-bin-v1.5.tar.gz
[root@WCT~]# mv wct-package-bin-v1.5/etc/jta.jar /home/tomcat/lib
[root@WCT~]# mv wct-package-bin-v1.5/war/* /home/tomcat/webapps/

```

附圖 3- 19 Web Curator Tool 安裝指令

(二) 複製 jmxremote.password.template 並另存成 jmxremote.password 檔案，將該檔案權限改為只有管理者帳號(root)可以讀寫，並修改 Tomcat 下的 bin 資料夾中的 catalina.sh 檔案，在檔案裡面加上一行敘述，如附圖 3- 20 所示。

```
[root@WCT ~]# cp /usr/lib/jvm/java-1.6.0-openjdk-1.6.0.0/ \
/jre/lib/management/jmxremote.password.template
/usr/lib/jvm/java-1.6.0-openjdk/jre/lib/management/jmxremote.password
[root@WCT ~]# chmod 600 \
/usr/lib/jvm/java-1.6.0-openjdk-1.6.0.0/jre/lib/management/jmxremote.password
[root@WCT ~]# echo \
"JAVA_OPTS=-Dcom.sun.management.jmxremote.port=9004">>
/home/tomcat/bin/catalina.sh
```

附圖 3- 20 Java 相關設定指令

(三) 下載 mysql-connector-java 的壓縮檔並解壓縮後，將 mysql-connector-java-5.1.13-bin.jar 移動到 Tomcat 的 lib 資料夾下，如附圖 3- 21 所示。

```
[root@WCT ~]# wget \
http://www.mysql.com/get/Downloads/Connector-J/mysql-connector-java-5.1.13.tar.g
z/from/http://mysql.cdpa.nsysu.edu.tw/
[root@WCT ~]# tar zxvf mysql-connector-java-5.1.13.tar.gz
[root@WCT ~]# mv \
mysql-connector-java-5.1.13/mysql-connector-java-5.1.13-bin.jar /home/tomcat/lib/
```

附圖 3- 21 MySQL-JDBC 安裝指令

(四) 去除 wct-schema-1_5_0-mysql.sql 註解並產生新的檔案後，一一匯入到 MySQL 資料庫中，注意在留白區域部分，請輸入 MySQL 的使用者 root 帳號密碼，方可匯入成功，如附圖 3- 22 所示。

```
[root@WCT ~]# sed "--/d" sql/mysql/wct-schema-1_5_0-mysql.sql > \
sql/mysql/wct-schema-1_5_0-mysql-new.sql
[root@WCT ~]# mysql -u root -p DB_WCT < \
sql/mysql/wct-schema-1_5_0-mysql-new.sql
Enter password:
[root@WCT ~]# mysql -u root -p DB_WCT < \
sql/mysql/wct-schema-grants-1_5_0-mysql.sql
Enter password:
[root@WCT ~]# mysql -u root -p DB_WCT < \
sql/mysql/wct-mysql-bootstrap.sql
Enter password:
[root@WCT ~]# mysql -u root -p DB_WCT < \
sql/mysql/wct-indexes-1_5_0-mysql.sql
Enter password:
```

附圖 3- 22 資料庫匯入指令

(五) 完成上述步驟後，請先啟動 Tomcat，等候 Tomcat 自動解壓縮並掛載 war 檔案，這段時間大約要等候一分鐘左右，之後再關閉 Tomcat，如附圖 3- 23 所示。

```
[root@WCT ~]# /home/tomcat/bin/startup.sh
[root@WCT ~]# /home/tomcat/bin/shutdown.sh
```

附圖 3- 23 Tomcat 啟動關閉指令

(六) 如附圖 3- 24 所示，請編輯 wct-core.properties 檔案，修改 hibernate.dialect 等號右邊的敘述。

```
檔案名字：wct-core.properties↵  
檔案位置：/home/tomcat/webapps/wct/WEB-INF/classes/↵  
修改內容：↵  
hibernate.dialect=org.hibernate.dialect.MySQLDialect↵
```

附圖 3- 24 wct-core.properties 檔案位置暨修改部分

(七) 如附圖 3- 25 所示，請編輯 context.xml 檔案，修改下列參數等號右邊的敘述。同時請注意 MYSQL_PASSWORD-FOR-use_wct 及 YOUR_IP 為使用者自訂的資料。

```
檔案名字：context.xml↵  
檔案位置：/home/tomcat/webapps/wct/META-INF/↵  
修改內容：↵  
password="MYSQL_PASSWORD-FOR-usr_wct"↵  
driverClassName="com.mysql.jdbc.Driver"↵  
validationQuery="select 1+1"↵  
username="usr_wct"↵  
url="jdbc:mysql://YOUR_IP:3306/DB_WCT"↵
```

附圖 3- 25 context.xml 檔案位置暨修改部分

(八) 如附圖 3- 26 所示，將修改完後的 context.xml 檔案複製一份，放到 Tomcat 下的/conf/Catalina/localhost 資料夾內，並且更名為 wct.xml

```
[root@WCT ~]# cp -f /home/tomcat/webapps/wct/META-INF/context.xml \  
/home/tomcat/conf/Catalina/localhost/wct.xml↵
```

附圖 3- 26 context.xml 另存指令

(九) 如附圖 3- 27 所示，請編輯 wct-das.properties，修改下列參數等號右邊的敘述。同時請注意 YOUR_IP 為使用者自訂的資料。

```
檔案名字：wct-das.properties↵
檔案位置：/home/tomcat/webapps/wct-store/WEB-INF/classes/↵
修改內容：↵
wctCoreWsEndpoint.host=YOUR_IP↵
wctCoreWsEndpoint.port=8080↵
arcDigitalAssetStoreService.baseDir=/tmp/arcstore↵
```

附圖 3- 27 wct-das.properties 檔案位置暨修改部分

(十) 如附圖 3- 28 所示，請編輯 wct-agent.properties 檔案，修改下列參數等號右邊的敘述。同時請注意 YOUR_IP 為使用者自訂的資料。

```
檔案名字：wct-agent.properties↵
檔案位置：/home/tomcat/webapps/wct-harvest-agent/WEB-INF/classes/↵
修改內容：↵
harvestAgent.baseHarvestDirectory=/wct/harvest-agent↵
harvestAgent.host=YOUR_IP↵
harvestAgent.port=8080↵
harvestCoordinatorNotifier.host=YOUR_IP↵
harvestCoordinatorNotifier.port=8080↵
digitalAssetStore.host=YOUR_IP↵
digitalAssetStore.port=8080↵
heartbeatTrigger.startDelay=10000↵
heartbeatTrigger.repeatInterval=30000↵
```

附圖 3- 28 wct-agent.properties 檔案位置暨修改部分

(十一) 如附圖 3- 29 所示，請編輯 Tomcat 下 webapps/wct 資料夾中的 server-config.wsdd 檔案，修改下列參數等號右邊的敘述。

檔案名字：server-config.wsdd↵
檔案位置：/home/tomcat/webapps/wct/WEB-INF/↵
修改內容：↵
<parameter name="attachments.Directory" value="/tmp/attach"/>↵

附圖 3- 29 wct 資料夾中 server-config.wsdd 修改部分

(十二) 如附圖 3- 30 所示，請編輯 Tomcat 下 webapps/wct-store 資料夾中的 server-config.wsdd 檔案，修改下列參數等號右邊的敘述。

檔案名字：server-config.wsdd↵
檔案位置：/home/tomcat/webapps/wct-store/WEB-INF/↵
修改內容：↵
<parameter name="attachments.Directory" value="/tmp/attach"/>↵

附圖 3- 30 wct-store 資料夾中 server-config.wsdd 檔案修改部分

(十三) 如附圖 3- 31 所示，請編輯 Tomcat 下 webapps/wct-harvest-agent 資料夾中的 server-config.wsdd，修改下列參數等號右邊的敘述。

檔案名字：server-config.wsdd↵
檔案位置：/home/tomcat/webapps/wct-harvest-agent/WEB-INF/↵
修改內容：↵
<parameter name="attachments.Directory" value="/tmp/attach"/>↵

附圖 3- 31 wct-harvest-agent 中 server-config.wsdd 檔案修改部分

六、安裝 Wayback

(一) 下載 Wayback-1.4.2 的壓縮檔並將該壓縮檔解壓縮之後，把 wayback-webapp-1.4.2.war 檔案移動到 Tomcat 的 webapps 下並且重新命名為 wayback.war。接著，將解壓縮後的 wayback 下的 bin 的檔案放到 /usr/local/bin 下，同時也將 wayback 下的 lib 的 jar 檔案放到 Tomcat 下的 lib 資料夾內，如附圖 3- 32 所示。

```
[root@WCT ~]# wget \
http://downloads.sourceforge.net/project/archive-access/wayback/1.4.2/wayback-1.4.2
.tar.gz?r=http%3A%2F%2Fsourceforge.net%2Fprojects%2Farchive-access%2Ffiles%
2F&ts=1282729379&mirror=ncu
[root@WCT ~]# tar zxvf wayback-1.4.2.tar.gz
[root@WCT ~]# mv wayback/lib/wayback-webapp-1.4.2.war \
/home/tomcat/webapps
[root@WCT ~]# mv /home/tomcat/webapps/wayback-webapp \
/home/tomcat/webapps/wayback
[root@WCT ~]# mv wayback/bin/* /usr/local/bin
[root@WCT ~]# mv wayback/lib/*.jar /home/tomcat/lib
```

附圖 3- 32 Wayback 安裝指令

(二) 完成上面步驟後，請先啟動 Tomcat，等候 Tomcat 自動解壓縮並掛載 war 檔案，這段時間大約要等候一分鐘左右，如附圖 3- 33 所示。

(三) 下載 WCT 的 Wayback 範例程式並且解壓縮後，將 CustomBanner.jsp 放到 Tomcat 下的 wayback 內的 WEB-INF/replay 資料夾下，再將 ArchivalUrlReplay.xml 放到 WEB-INF 資料夾下，如附圖 3- 33 所示。

```
[root@WCT ~]# wget \
http://downloads.sourceforge.net/project/webcurator/webcurator/1.5GA/WCT_1_5_Wayback_Sample.zip?r=http%3A%2F%2Fsourceforge.net%2Fprojects%2Fwebcurator%2Ffiles%2F&ts=1282730141&mirror=ncu
[root@WCT ~]# unzip WCT_1_5_Wayback_Sample.zip
[root@WCT ~]# mv CustomBanner.jsp \
/home/tomcat/webapps/wayback/WEB-INF/replay
[root@WCT ~]# mv ArchivalUrlReplay.xml \
/home/tomcat/webapps/wayback/WEB-INF
```

附圖 3- 33 Wayback 範例程式安裝指令

(四) 如附圖 3- 34 所示，請編輯 wct-core.properties 檔案，修改下列參數等號右邊的敘述。注意，IP 及 port 為使用者自訂的資料。

```
檔案名字：wct-core.properties
檔案位置：/home/tomcat/webapps/wct/WEB-INF/classes
修改內容：
qualityReviewToolController.enableBrowseTool=true
qualityReviewToolController.enableAccessTool=true
harvestResourceUrlMapper.urlMap=http://<IP>:<port>/wayback/wayback/{$ArcHarvestResource.FileDate}/{${HarvestResource.Name}}
```

附圖 3- 34 Wayback 下 wct-core.properties 檔案位置暨修改部分

(五) 如附圖 3- 35 所示，請編輯 wct-das.properties 檔案，修改下列參數等號右邊的敘述。

檔案名字：wct-das.properties
檔案位置：/home/tomcat/webapps/wct-store/WEB-INF/classes
修改內容：
waybackIndexer.enabled=true
waybackIndexer.waybackInputFolder=/tmp/wayback/store
waybackIndexer.waybackMergedFolder=/tmp/wayback/index-data/merged
waybackIndexer.waybackFailedFolder=/tmp/wayback/index-data/failed

附圖 3- 35 Wayback 下 wct-das.properties 檔案位置暨修改資訊

(六) 如附圖 3- 36 所示，在 /tmp 內建立 wayback store 資料夾。

```
[root@WCT~]# mkdir -p /tmp/wayback/store
```

附圖 3- 36 wayback/store 資料夾建立指令

(七) 附圖 3- 37 所示，請編輯 ArchivalUrlReplay.xml 檔案，修改下列參數等號右邊的敘述，注意，請將 <!-- 及 --> 註解拿掉。

檔案名字：ArchivalUrlReplay.xml
檔案位置：/home/tomcat/webapps/wayback/WEB-INF/
修改內容：
<!--
<value>/WEB-INF/replay/CustomBanner.jsp</value>
-->
<property name="renderer" ref="archivalserverhtmlreplayrenderer"/>

附圖 3- 37 ArchivalUrlReplay.xml 檔案位置暨修改資訊

(八) 如附圖 3- 38 所示，請編輯 BDBCcollection.xml 檔案，修改下列參數等號右邊的敘述。

```
檔案名字：BDBCcollection.xml↵
檔案位置：/home/tomcat/webapps/wayback/WEB-INF/↵
修改內容：↵
<property name="prefix" value="/tmp/wayback/store/" />
```

附圖 3- 38 BDBCcollection.xml 檔案位置暨修改資訊

(九) 完成上述步驟後，Web Curator Tool 及 Wayback 已經安裝完成。如附圖 3- 39 所示，啟動 Tomcat 後，即可使用系統。

```
[root@WCT~]# /home/tomcat/bin/startup.sh
```

附圖 3- 39 Tomcat 啟動指令

(十) 將 FilePost.war 檔案移動到 Tomcat 的 webapps 資料夾下，並在 tmp 資料夾中建立 rms 資料夾，以供上傳的網站檔案或匯出的 XML 檔案暫存使用，如附圖 3- 40 所示。

```
[root@localhost wct]# cp Downloads/FilePost.war /home/tomcat/webapps/
[root@localhost wct]# mkdir /tmp/rms
```

附圖 3- 40 雛形系統安裝指令

(一) 在 mysql 中建立 rms 使用者以及 rms 資料庫、Record 資料表，並賦與使用者權限，如附圖 3- 41 所示。


```

[root@localhost wct]# mysql -u root -p
Enter password: 
mysql> create user rms@localhost;
mysql> create database rms;
mysql> create table rms.Record (ID bigint not null auto_increment unique,
-> sysname varchar(5),
-> functiontype varchar(5),
-> recordtitle varchar(50),
-> recordtype varchar(5),
-> copytype varchar(5),
-> secretlv varchar(5),
-> revyear varchar(5),
-> classsn varchar(100),
-> gendate varchar(7),
-> filetype varchar(5),
-> filenum varchar(5),
-> fileunit varchar(5),
-> username varchar(30),
-> userdep varchar(50),
-> usertitle varchar(50),
-> sitelink varchar(100),
-> sitepath varchar(100),
-> sitecollector varchar(30),
-> sitetitle varchar(50),
-> sitesubject varchar(50),
-> sitetype varchar(50),
-> siteformat varchar(20),
-> sitelang varchar(20),
-> sitedesc varchar(255));
mysql> grant USAGE, SELECT, INSERT, UPDATE on rms.* to rms@localhost;

```

附圖 3- 41 雛形系統資料庫設定指令

附錄四 網站保存系統操作手冊

Web Curator Tool (WCT)是一套 Open Source 的網站保存管理工具、目前英國與紐西蘭也採用這套軟體作為網站保存的工具，並且發展出各自的分支。由於各國國情不同、對於網站保存相關法律的差異等諸多因素，WCT 在各個方面都有一定程度的支援及預防，也導致其功能有許多檢查審核的部分需要處理，以下會先概略的說明 WCT 各部諸元功能，接下來以檔管局網站為例，實作網站抓取的過程，最後會展示一般民眾從外部瀏覽保存下來的網站的畫面。

一、 Web Curator Tool 介面概要說明

(一) 當軟體安裝完成後，請在網址列內打入

`http://127.0.0.1:8080/wct`，即會出現登入畫面，如附圖 4- 1 所示。

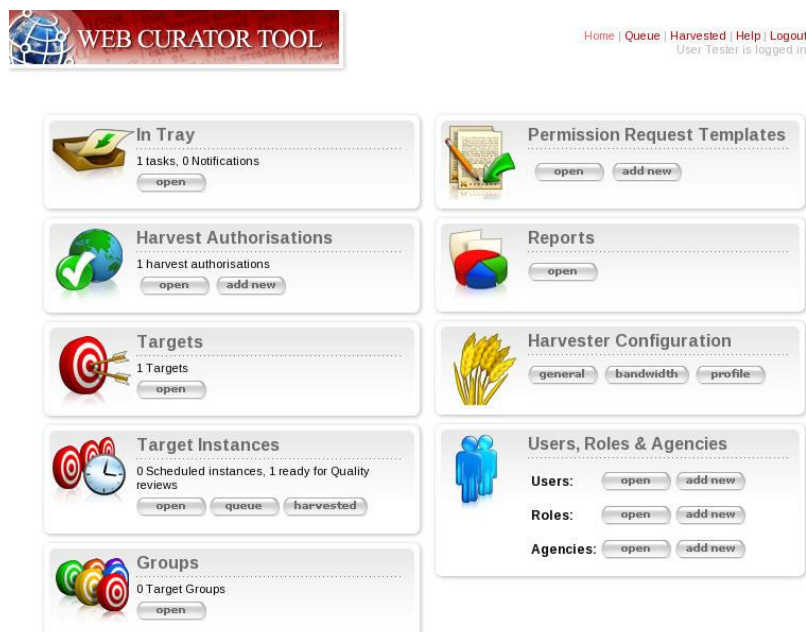
- username：使用者帳號，系統預設登入帳號為 bootstrap
- password：密碼，系統預設登入帳號的密碼為 password
- login：登入按鈕



附圖 4- 1 WCT 登入畫面

(二) 登入之後，會出現 WCT 的首頁，如附圖 4- 2 所示。

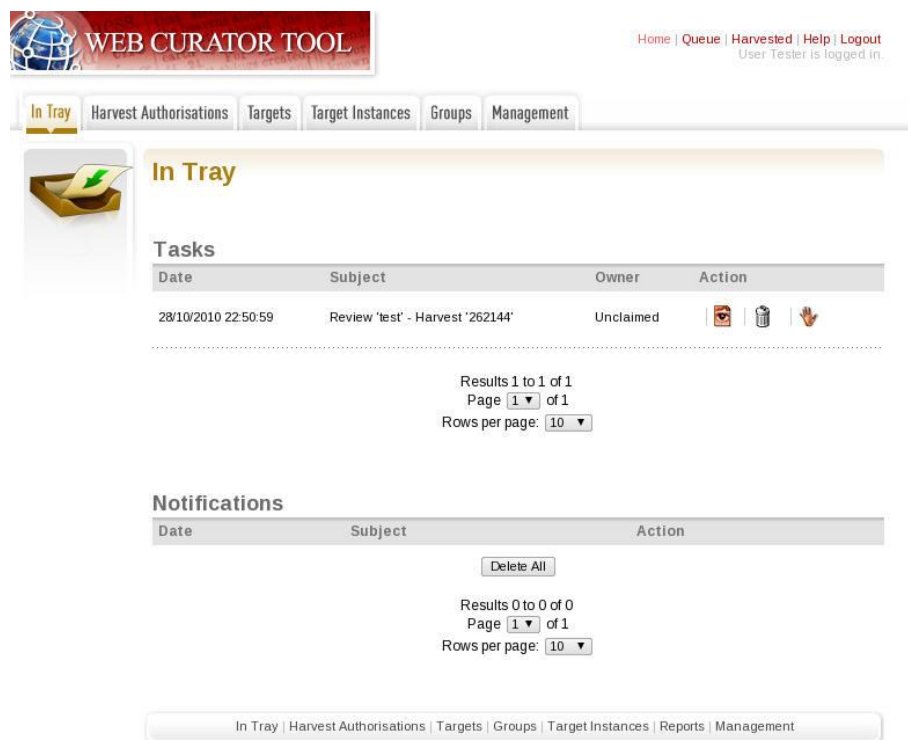
- In Tray：使用者的待處理事項以及通知事項
- Harvest Authorisations：抓取授權的相關設定
- Targets：目標網站的資料，包含抓耙器的相關設定
- Target Instances：目標實體的相關設定
- Groups：群組的相關設定
- Permission Request Templates：申請許可信的範本
- Harvester Configuration：抓取程式的相關設定
- User, Roles & Agencies：機關、角色、使用者的相關設定



附圖 4- 2 WCT 首頁畫面

(三) 在 In Tray 的部分中，點擊開啟(Open)按鈕，便可以觀看待處理事項以及通知事項，如附圖 4- 3 所示。

- Tasks：待處理事項，圖中的事項為抓取完成需要審核的網站。
- Notifications：通知事項。



附圖 4- 3 WCT In Tray 畫面

(四) 在抓取授權(Harvest Authorisations)的部分中，如附圖 4- 4 所示，可以新增及查詢欲抓取網站之授權設定。

- Search：搜尋的介面，可以透過序號(ID)、名稱(Name)、授權單位(Authorising)等參數進行搜尋。
- Results：搜尋的結果，根據需求可以對授權項目進行檢視(View)、編輯(Edit)、複製(Copy)以及產生信件(Generate letter/email)等動作

(Action)。

- create new：建立新的抓取授權

WEB CURATOR TOOL

Home | Queue | Harvested | Help | Logout
User Tester is logged in.

In Tray | **Harvest Authorisations** | Targets | Target Instances | Groups | Management

Harvest Authorisations

Search

ID: Name: Authorising Agent: Order No: Agency: Show Disabled:
 ☐

URL Pattern: Permissions File Reference: Permissions Status:
 ☐ Pending ☐ Requested
☐ Approved ☐ Rejected

create new

Id	Name	Auth Agent	Order No	Status	Action
131072	test	local		Approved	

Results 1 to 1 of 1
Page 1 of 1
Rows per page: 10

In Tray | Harvest Authorisations | Targets | Groups | Target Instances | Reports | Management

附圖 4- 4 WCT 抓取授權畫面

(五) 在首頁中抓取授權的部分點擊新增(add new)按鈕或者是在抓取授權的頁面中點選新增(create new)的按鈕，都會進入到新增抓取授權的一般設定畫面，如附圖 4- 5 所示。新增抓取授權的頁面有四個頁籤，分別是一般設定(General)、網址樣式(URL Patterns)、授權機關(Authorising Agencies)、權限(Permissions)四個部分，以下先將說明一般設定部分的欄位資料。

- Agency：使用者所在的機關
- Title：欲抓取網站的專案名稱
- Description：欲抓取網站的描述說明

- Order No：單號
- Published：發佈與否
- Enable：啟用與否
- Annotations：額外的註記，會依據時間條列式紀錄下來，並可在之後

對先前註記過的資料進行修改或刪除。

附圖 4- 5 WCT 抓取授權一般設定畫面

(六) 點擊網址樣式(URL Patterns)的頁籤即可切換至網址的設定畫面，如附圖 4- 6 所示。網址可以藉由使用萬用字元(*)的方式來變更抓取的範圍。

- New URL Pattern：欲抓取網站的網址樣式
- add：新增網址，之後可以在網址樣式列表中對已新增的網址樣式進行

刪除的動作



附圖 4- 6 WCT 抓取授權網址樣式畫面

(七) 點擊授權機關(Authorising Agencies) 的頁籤即可切換至授權機關的設定畫面，如附圖 4- 7 所示。

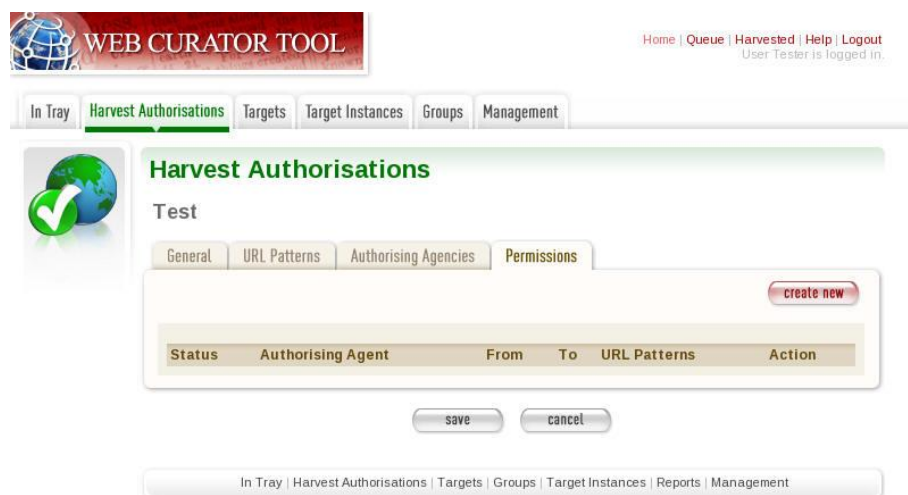
- search：搜尋曾經有創建過的機關資料，可直接從頁面中匯入
- create new：建立新的機關資料。姓名(Name)、聯絡人(Contact)為必填欄位，電話(Phone)、電子郵件地址(Email)、住址(Address)三項至少必須選填一項。



附圖 4- 7 WCT 抓取授權授權機關畫面

(八) 點擊權限(Permissions)的頁籤即可切換至抓取權限的設定畫面，
如附圖 4- 8 所示。

• create new：新增抓取權限的相關設定，新增的設定會以條列式的方式
列在下方，之後可以進行檢視、編輯、刪除及傳送的動作



附圖 4- 8 WCT 抓取授權權限畫面

(九) 在目標(Targets)的部分中，如附圖 4- 9 所示，可以新增及查詢欲
抓取網站之抓爬器等設定。

- Search：搜尋的介面，可以透過序號(ID)、名稱(Name)、網址(Seed)等參數進行搜尋。
- Results：搜尋的結果，根據需求可以對目標進行檢視(View)、編輯(Edit)、複製(Copy)等動作(Action)。
- create new：建立新的目標



WEB CURATOR TOOL

Home | Queue | Harvested | Help | Logout
User Tester is logged in.

In Tray | Harvest Authorisations | **Targets** | Target Instances | Groups | Management

Targets


Search

ID: Name: Seed: Agency: User:

Member of: State: ☐ Pending ☐ Reinstated ☐ Nominated ☐ Rejected
☐ Approved ☐ Cancelled ☐ Completed

search

Results [create new](#)

Id	Name	Agency	Owner	Status	Seeds	Action
32768	test	Test	Te Ster	Approved	http://localhost:8080/	  

Results 1 to 1 of 1
Page 1 of 1
Rows per page: 10

In Tray | Harvest Authorisations | Targets | Groups | Target Instances | Reports | Management

附圖 4- 9 WCT 抓取授權目標畫面

(十) 在目標的頁面中點選新增(create new)的按鈕，會進入到新增目標的一般設定畫面，如附圖 4- 10 所示。新增目標的頁面有八個頁籤，分別是一般設定(General)、網址(Seeds)、抓爬器設定(Profile)、排程(Schedule)、註記(Annotations)、網站說明(Description)、群組(Groups)、存取限制(Access)八個部分，以下先將說明一般設定部分的欄位資料。

- Name：目標的專案名稱

- Description：目標的說明
- Reference Number：參照編號
- Run on Approval：在被許可的狀況下是否執行
- Owner：任務的擁有者
- Status：任務的狀態

附圖 4- 10 WCT 目標一般設定畫面

(十一) 點擊網址(Seeds)的頁籤即可切換至網址抓取及權限設定的畫面，如附圖 4- 11 所示。

- Seed：網址樣式
- Authorisation：是否自動搜尋已經有相關授權資料

- link：將網址加入抓取的列表中，如果無法自動取得抓取授權的話，仍需要在 Havrest Auth 的欄位中點選新增(add)按鈕增加所需的授權。
- import：藉由檔案匯入所需資料



附圖 4- 11 WCT 目標網址設定畫面

(十二) 點擊抓爬器設定(Profile)的頁籤即可切換至抓爬器相關的設定畫面，如附圖 4- 12 所示。

- Base Profile：採用何種預先設置好的抓爬器設定
- Profile Overrides：僅在本次目標中更動的抓爬器設定，需要在啟用更動(Enable Override)的選項打勾才會真正生效
- Robot Honouring Policy：選擇是否遵從 robot.txt 的限制進行網站的保存
- Maximum Hours：該目標專案最多跑幾個小時

- Maximum Kilobytes：該目標專案最多抓幾 KB 的資料
- Maximum Documents：該目標專案最多抓多少個檔案
- Maximum Path Depth：該目標專案探索的最大深度
- Maximum Hops：該目標專案探索的最大節點數
- Exclude Filters：過濾欲排除的網站連結
- Force Accept Filter：過濾欲強制抓取的網站連結
- Excluded MIME Types：過濾欲排除之檔案類型
- Credentials Override：登入用的帳號資料


WEB CURATOR TOOL

[Home](#) | [Queue](#) | [Harvested](#) | [Help](#) | [Logout](#)
 User Tester is logged in.

[In Tray](#) | [Harvest Authorisations](#) | **[Targets](#)** | [Target Instances](#) | [Groups](#) | [Management](#)



Targets

Test

[General](#) | [Seeds](#) | **[Profile](#)** | [Schedule](#) | [Annotations](#) | [Description](#) | [Groups](#) | [Access](#)

Base Profile
 Base Profile: [Default - Test](#)

Profile Overrides

Profile Element	Override Value	Enable Override
Robot Honouring Policy	classic	<input type="checkbox"/>
Maximum Hours	<input type="text" value="0"/>	<input type="checkbox"/>
Maximum Kilobytes	<input type="text" value="0"/>	<input type="checkbox"/>
Maximum Documents	<input type="text"/>	<input type="checkbox"/>
Maximum Path Depth	<input type="text"/>	<input type="checkbox"/>
Maximum Hops	<input type="text"/>	<input type="checkbox"/>
Exclude Filters	<input type="text"/>	<input type="checkbox"/>
Force Accept Filters	<input type="text"/>	<input type="checkbox"/>
Excluded MIME Types	<input type="text"/>	<input type="checkbox"/>

Credentials Override
☐ Click to override the security credentials

- [Add Basic Credentials](#)
- [Add Form Credentials](#)

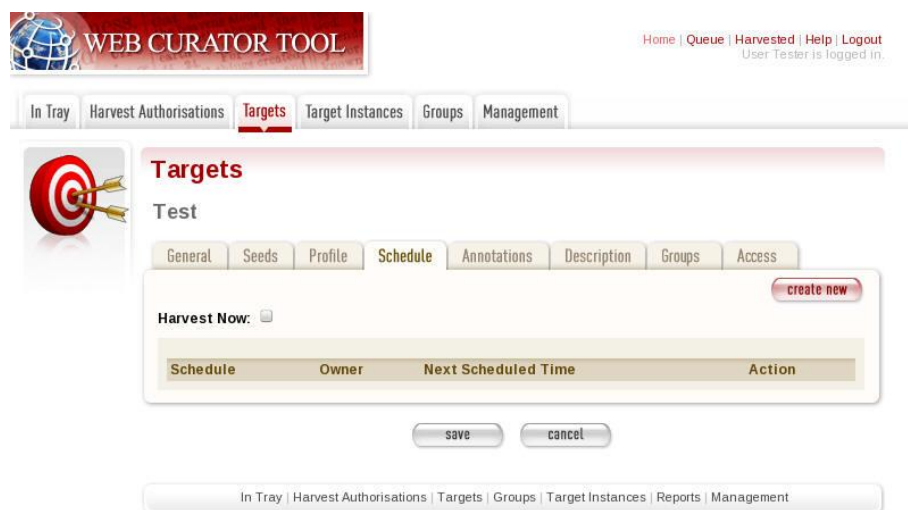
 No credentials provided in the overrides
 Profile Note:

[In Tray](#) | [Harvest Authorisations](#) | [Targets](#) | [Groups](#) | [Target Instances](#) | [Reports](#) | [Management](#)

附圖 4- 12 WCT 目標抓爬器設定畫面

(十三) 點擊排程(Schedule)的頁籤即可切換至網址抓取排程設定的畫面，如附圖 4- 13 所示。

- Harvest Now：不進行排程，目標設定完畢後直接進行抓取
- create new：建立新的排程



附圖 4- 13 WCT 目標排程設定畫面

(十四) 點擊註記(Annotations)的頁籤即可切換至註記資料的畫面，如附圖 4- 14 所示。

- Selection：這個目標選擇網站上的一些說明資料，如選擇的日期、種類、抓取的類型等等
- Annotations：註記說明的文字

WEB CURATOR TOOL

Home | Queue | Harvested | Help | Logout
User Tester is logged in.

In Tray | Harvest Authorisations | **Targets** | Target Instances | Groups | Management

Targets

Test

General | Seeds | Profile | Schedule | **Annotations** | Description | Groups | Access

Selection

Selection
Date:
Selection Type:
Selection Note:
Evaluation Note:
Harvest Type:

Annotations

Date	User	Notes	Action
There are no annotations available.			

In Tray | Harvest Authorisations | Targets | Groups | Target Instances | Reports | Management

附圖 4- 14 WCT 目標註記畫面

(十五) 點擊網站說明(Description)的頁籤即可切換至網站內容的都柏林核心詮釋資料編輯畫面，如附圖 4- 15 所示。WCT 並沒有使用到這些資料，日後移轉到其他檔案館或典藏機構時可能會需要這些資料。


WEB CURATOR TOOL

[Home](#) | [Queue](#) | [Harvested](#) | [Help](#) | [Logout](#)
 User Tester is logged in.

[In Tray](#) | [Harvest Authorisations](#) | **[Targets](#)** | [Target Instances](#) | [Groups](#) | [Management](#)



Targets

Test


[General](#) | [Seeds](#) | [Profile](#) | [Schedule](#) | [Annotations](#) | **[Description](#)** | [Groups](#) | [Access](#)

Title:
 Identifier:
 Description:
 Subject:
 Creator:
 Publisher:
 Contributor:
 Type:
 Format:
 Source:
 Language:
 Relation:
 Coverage:
 ISSN:
 ISBN:

[In Tray](#) | [Harvest Authorisations](#) | [Targets](#) | [Groups](#) | [Target Instances](#) | [Reports](#) | [Management](#)

附圖 4- 15 WCT 目標網站說明畫面

(十六) 點擊群組(Groups)的頁籤即可切換至群組的設定畫面，如附圖 4-16 所示。如果有需要可以將同類型的網站放置在相同的群組之中。


WEB CURATOR TOOL

[Home](#) | [Queue](#) | [Harvested](#) | [Help](#) | [Logout](#)
 User Tester is logged in.

[In Tray](#) | [Harvest Authorisations](#) | **[Targets](#)** | [Target Instances](#) | [Groups](#) | [Management](#)



Targets

Test

[General](#) | [Seeds](#) | [Profile](#) | [Schedule](#) | [Annotations](#) | [Description](#) | **[Groups](#)** | [Access](#)

add

Name	Action
This target does not belong to any groups.	

[In Tray](#) | [Harvest Authorisations](#) | [Targets](#) | [Groups](#) | [Target Instances](#) | [Reports](#) | [Management](#)

附圖 4- 16 WCT 目標網站群組設定畫面

(十七) 點擊存取限制(Access)的頁籤即可切換至群組的設定畫面，如附圖 4- 17 所示。

- Display Target：是否顯示目標
- Access Zone：可存取目標網站的限定範圍
- Target Introductory Display Note：目標顯示的說明



附圖 4- 17 WCT 目標存取限制畫面

(十八) 在首頁中目標實體的部分點擊開啟(open)按鈕，便會進入到查詢目標實體的搜尋畫面，如附圖 4- 18 所示。

- Search：搜尋的介面，可以透過序號(ID)、名稱(Name)、開始日期(From)等參數進行搜尋。

- Results：搜尋的結果，根據需求可以對目標實體進行檢視(View)、編輯(Edit)、立即抓取(Harvest Now)、刪除>Delete)等動作(Action)。

WEB CURATOR TOOL

Home | Queue | Harvested | Help | Logout
User Tester is logged in.

In Tray | Harvest Authorisations | Targets | **Target Instances** | Groups | Management




Target Instances

Search

ID: From: 03/11/2010 00:00:00 To: 03/02/2011 23:59:59 Agency: Test Owner: Te Ster

Name: Flagged Only: State: ☐ Scheduled ☐ Queued ☐ Running ☐ Paused ☐ Harvester ☐ Aborted ☐ Endorsed ☐ Rejected ☐ Archived ☐ Archiving

Results

Id	Name	Harvest Date	State	Owner	Run Time	Data Downloaded	Action
425984	test	04/11/2010 22:33:09	Scheduled	T. Ster			  

Results 1 to 1 of 1

附圖 4- 18 WCT 目標實體查詢畫面

(十九) 在目標實體的頁面點擊編輯(Edit)按鈕，便會進入到編輯目標實體的詳細資料畫面，如附圖 4- 19 所示。目標實體的頁面有七個頁籤，分別是一般設定(General)、抓爬器設定(Profile)、抓取狀態(Harvest State)、記錄檔(Logs)、抓取結果(Harvest Results)、註記(Annotations)、顯示(Display)七個部分。

WEB CURATOR TOOL

Home | Queue | Harvested | Help | Logout
User Tester is logged in.

In Tray | Harvest Authorisations | Targets | **Target Instances** | Groups | Management

Target Instances

test (425984)

General | Profile | Harvest State | Logs | Harvest Results | Annotations | Display

Id: 425984
Target Name: test
Schedule: 04/11/2010 22:33:09
Priority: Normal
Owner: Te Ster
Agency: Test
State: Scheduled
Bandwidth Percentage:
Flagged: ☐

save cancel

In Tray | Harvest Authorisations | Targets | Groups | Target Instances | Reports | Management

附圖 4- 19 WCT 目標實體設定畫面

(二十) 在首頁中群組的部分點擊開啟(open)按鈕，便會進入到查詢及新增群組的畫面，如附圖 4- 20 所示。

- Search：搜尋的介面，可以透過序號(ID)、名稱(Name)等參數進行搜尋。
- Results：搜尋的結果，根據需求可以對目標實體進行檢視(View)、編輯(Edit)、複製(Copy)、刪除>Delete)等動作(Action)。
- create new：建立新的群組

WEB CURATOR TOOL

Home | Queue | Harvested | Help | Logout
User Tester is logged in.

In Tray | Harvest Authorisations | Targets | Target Instances | **Groups** | Management

Groups

Search

ID: Name: Agency: Owner: Member of:

Type:

search reset

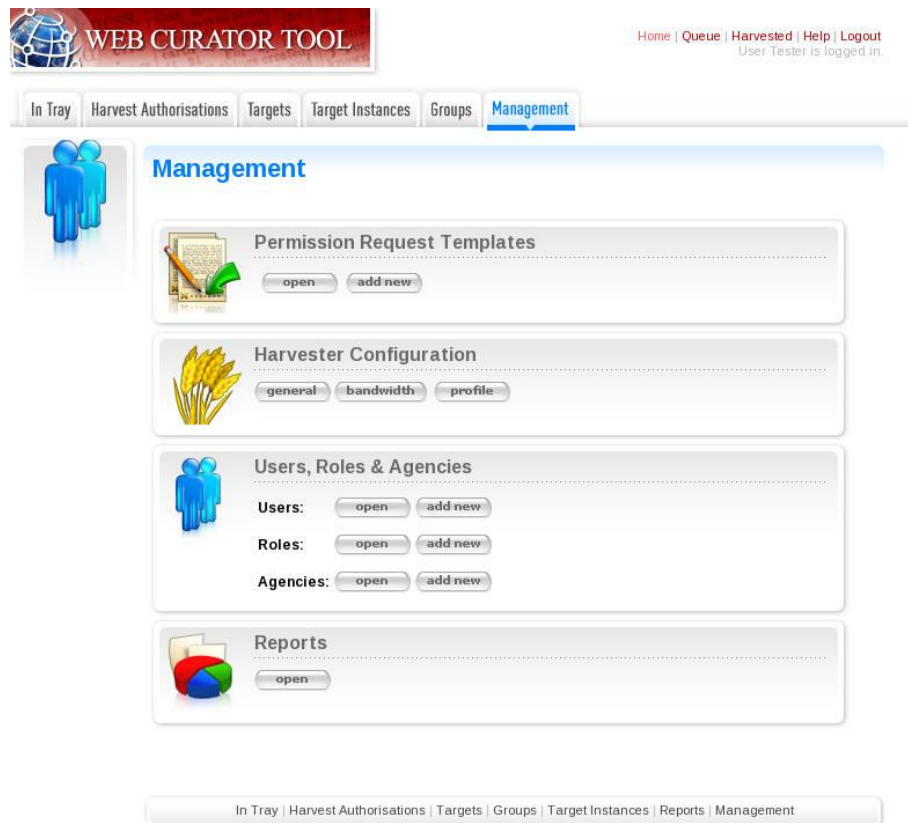
Results create new

Id	Name	Type	Agency	Owner	Status	Action
Results 0 to 0 of 0						
Page 1 of 1						
Rows per page: 10						

In Tray | Harvest Authorisations | Targets | Groups | Target Instances | Reports | Management

附圖 4- 20 WCT 群組設定畫面

(二十一) 從管理頁籤的頁面中可以看到跟首頁一樣的權限申請範本 (Permission Request Templates)、抓爬器(Harvester Configuration)、使用者、角色、機關(User, Roles & Agencies)、報告(Reports)等設定或資料查詢頁面，如附圖 4- 21 所示。

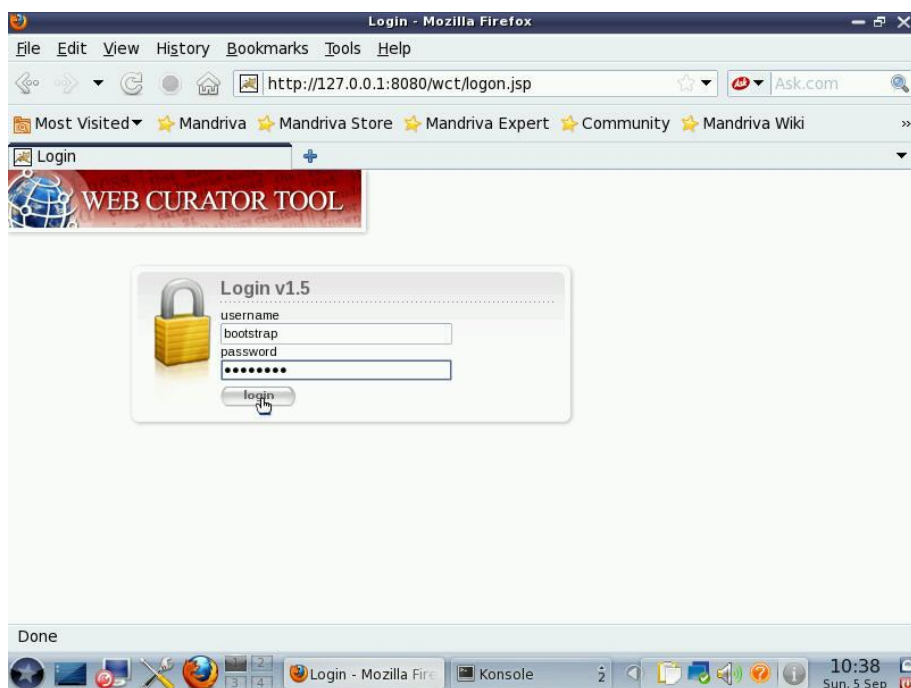


附圖 4- 21 WCT 管理畫面

二、 使用 Web Curator Tool 抓取檔管局網站

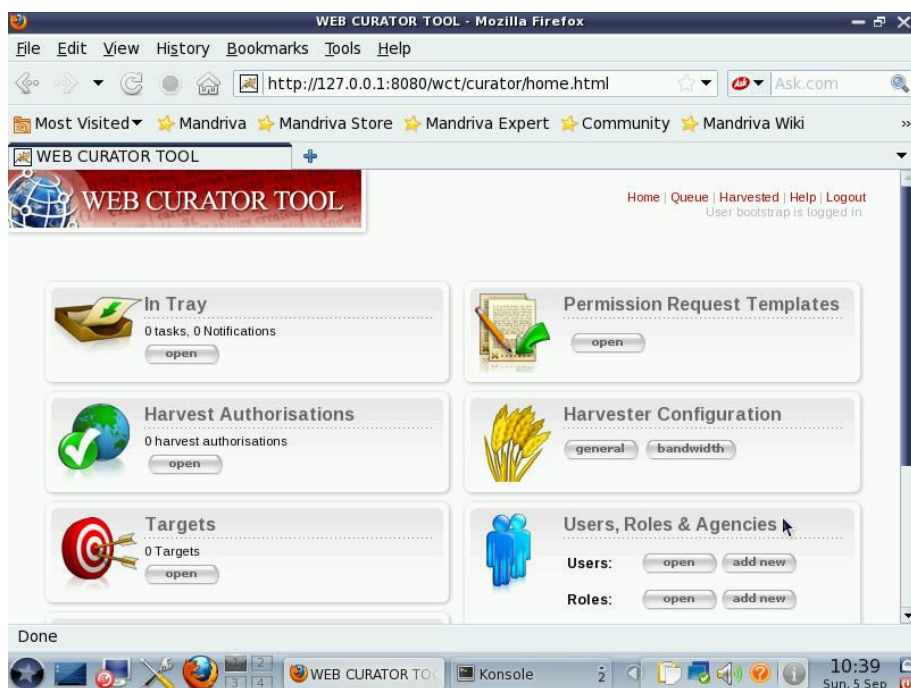
(一) 當軟體安裝完成後，請在網址列內打入

http://127.0.0.1:8080/wct ，即會出現登入畫面，第一次使用請利用預設帳號密碼登入，帳號請用 bootstrap，密碼請用 password，如附圖 4- 22 所示。



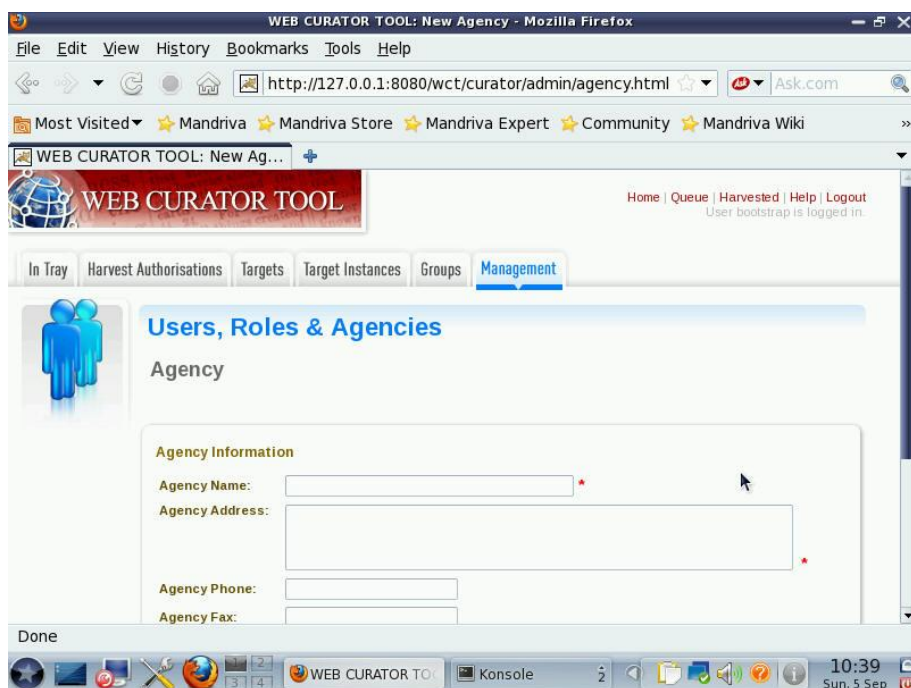
附圖 4- 22 WCT 登入畫面

(二) 使用預設帳號登入後，請在 Users, Roles & Agencies 欄位新增新的使用者、角色、單位，如附圖 4- 23 所示。



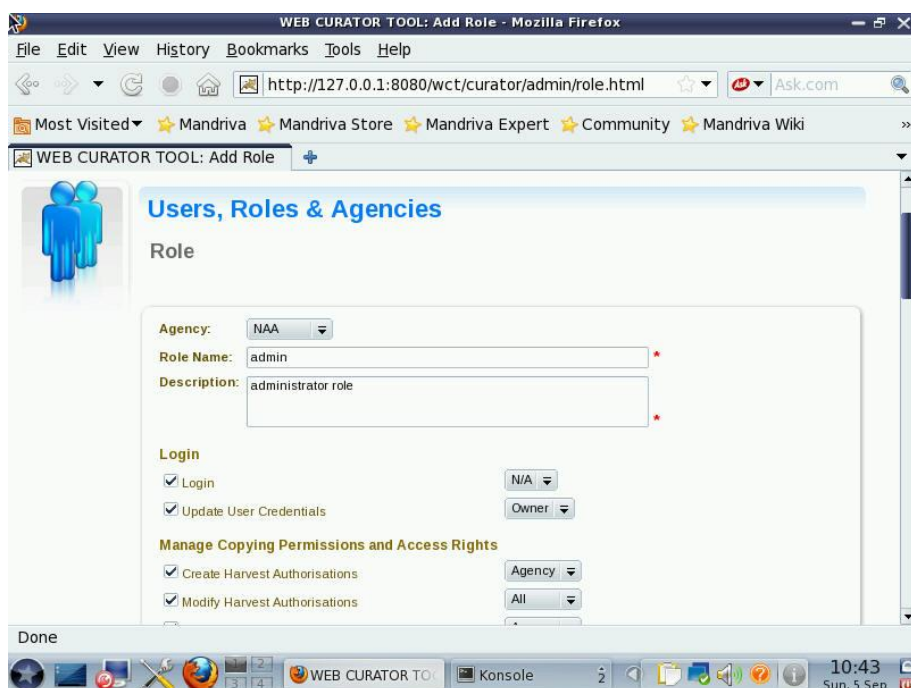
附圖 4- 23 WCT 首頁畫面

(三) 如附圖 4- 24 所示，請先建立新的單位。



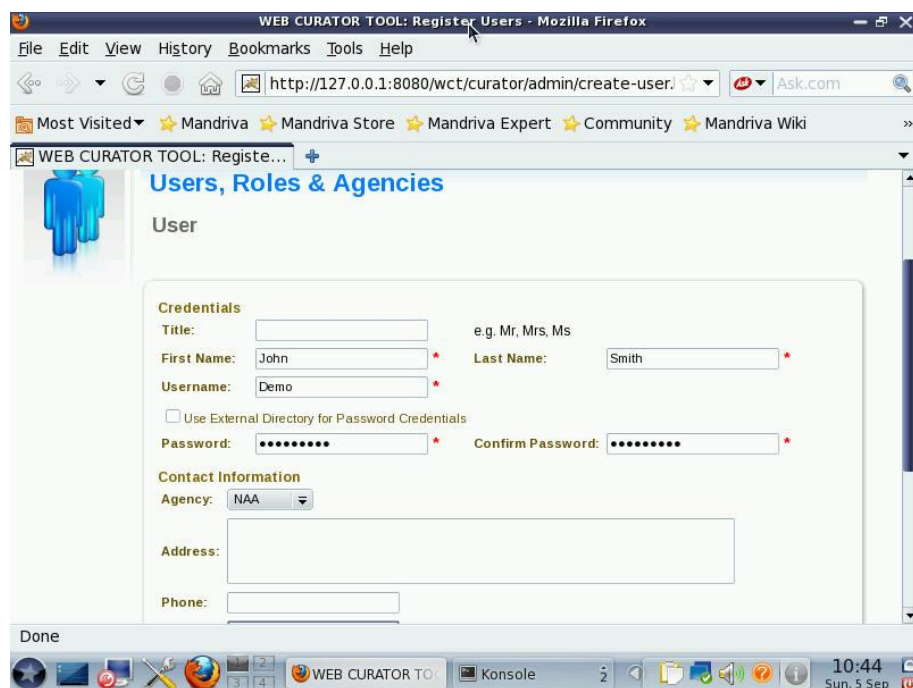
附圖 4- 24 新增單位畫面

(四) 如附圖 4- 25 所示，請建立新的角色，並且賦予所有權限。



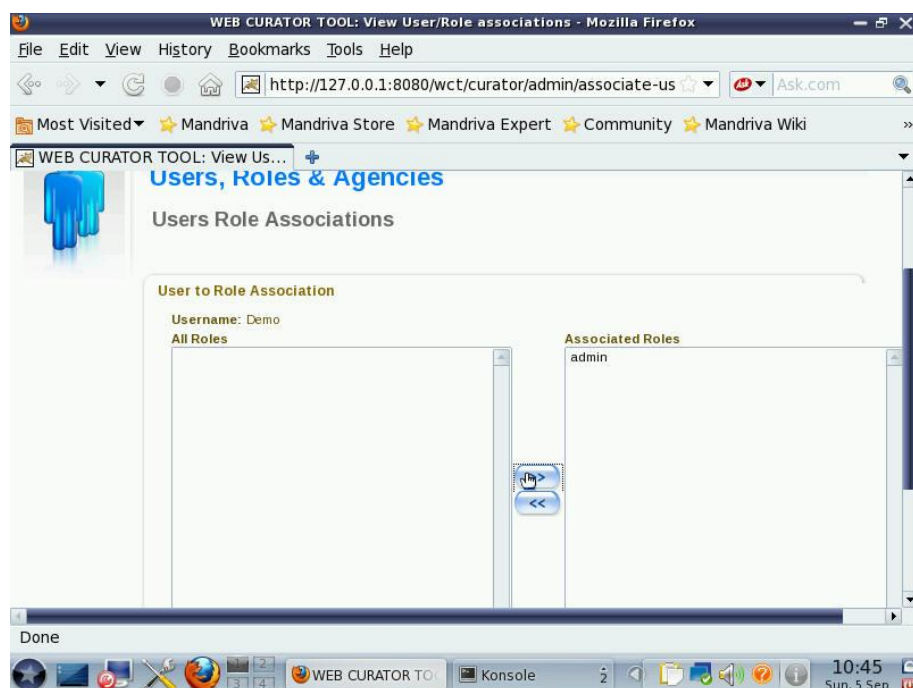
附圖 4- 25 新增角色畫面

(五) 如附圖 4- 26 所示，請建立新的使用者。



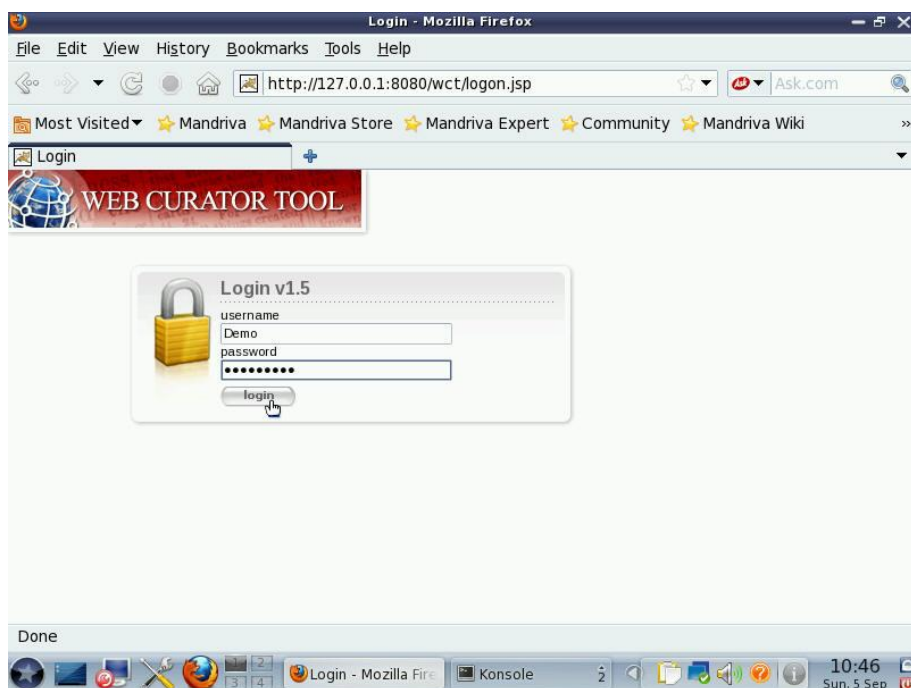
附圖 4- 26 新增使用者畫面

(六) 如附圖 4- 27 所示，請將建立完成的角色指派給新的使用者。



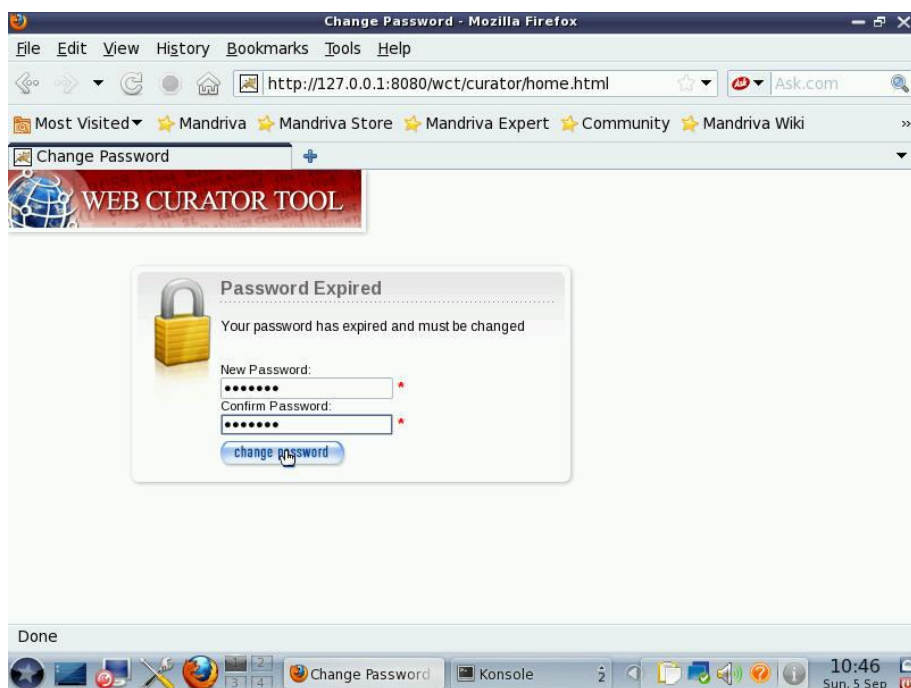
附圖 4- 27 角色指派畫面

(七) 如附圖 4- 28 所示，利用新建立的 User 登入 WCT。



附圖 4- 28 WCT 登入畫面

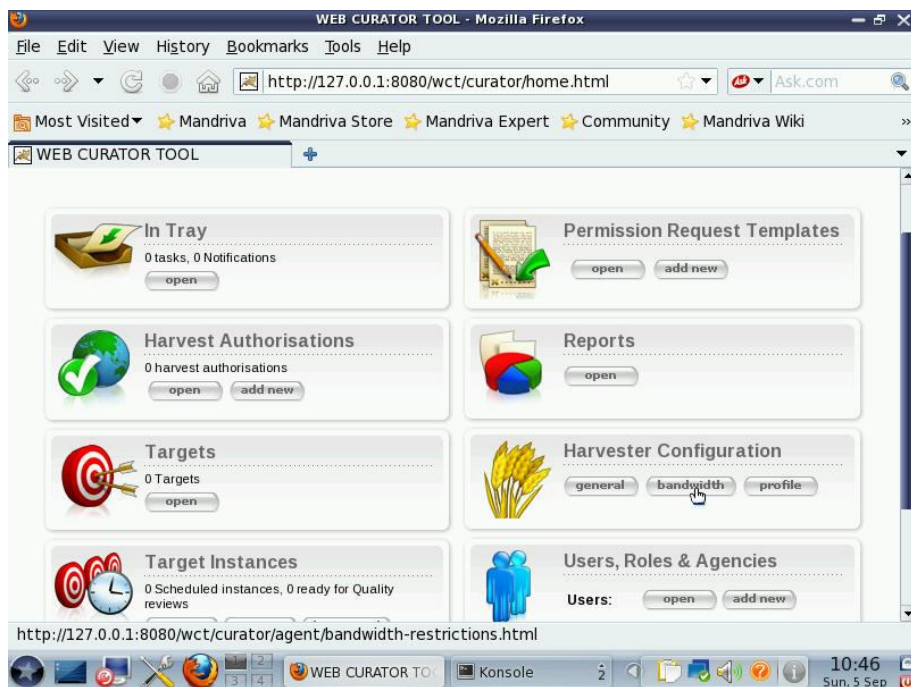
(八) 如附圖 4- 29 所示，使用新建立之帳號第一次登入後，必須修改密碼方可繼續操作。



附圖 4- 29 使用者密碼修改畫面

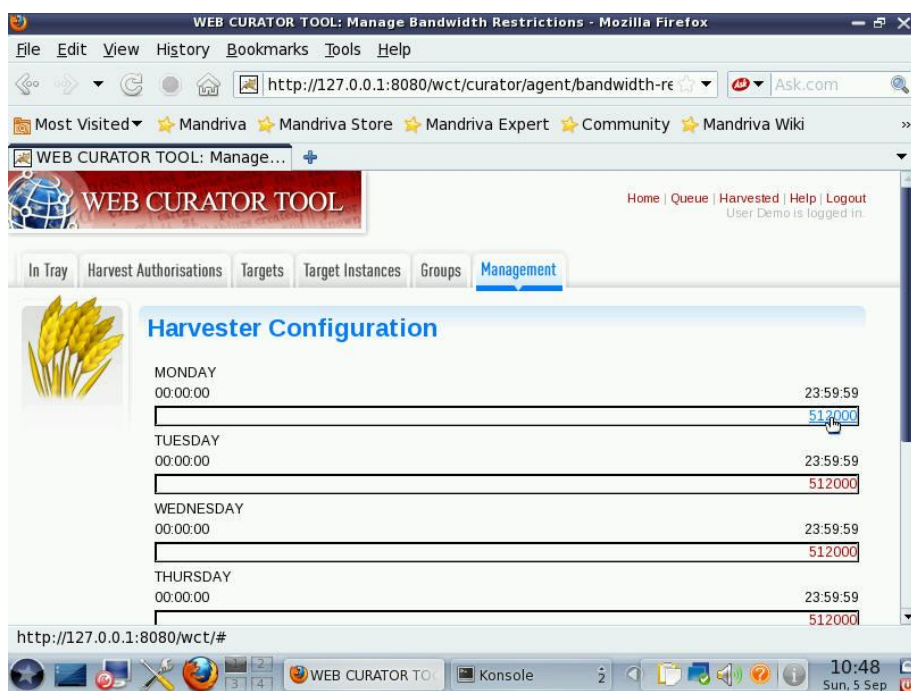
(九) 如附圖 4- 30 所示，操作 WCT 之前，必須在抓取設定(Harvest

Configuration)的群組點擊中頻寬(bandwidth)項目，修改頻寬預設為零的設定。



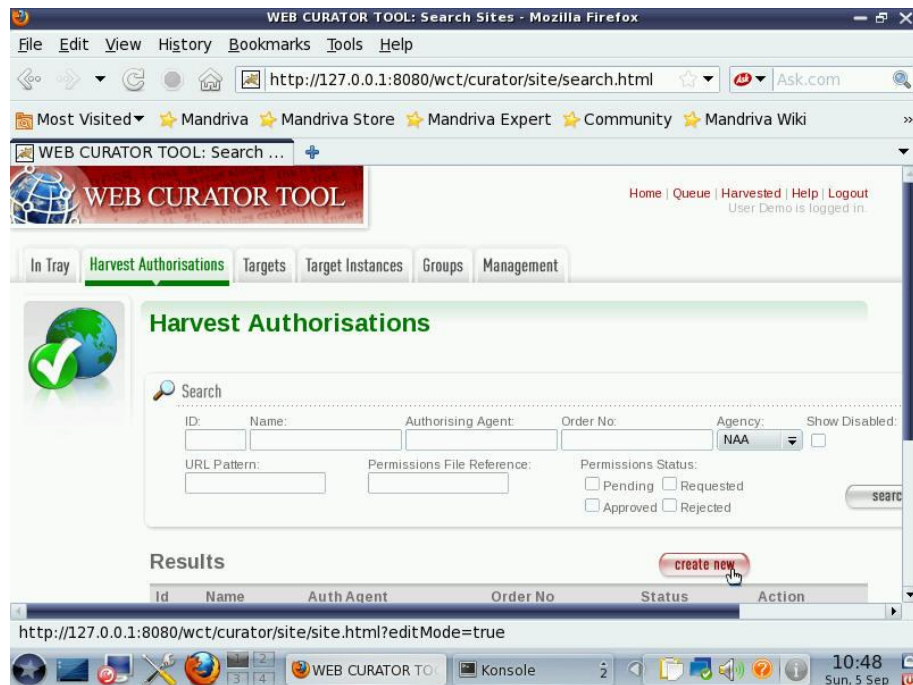
附圖 4- 30 WCT 首頁畫面

(十) 如附圖 4- 31 所示，此例頻寬設定為每天 512000 K/B。



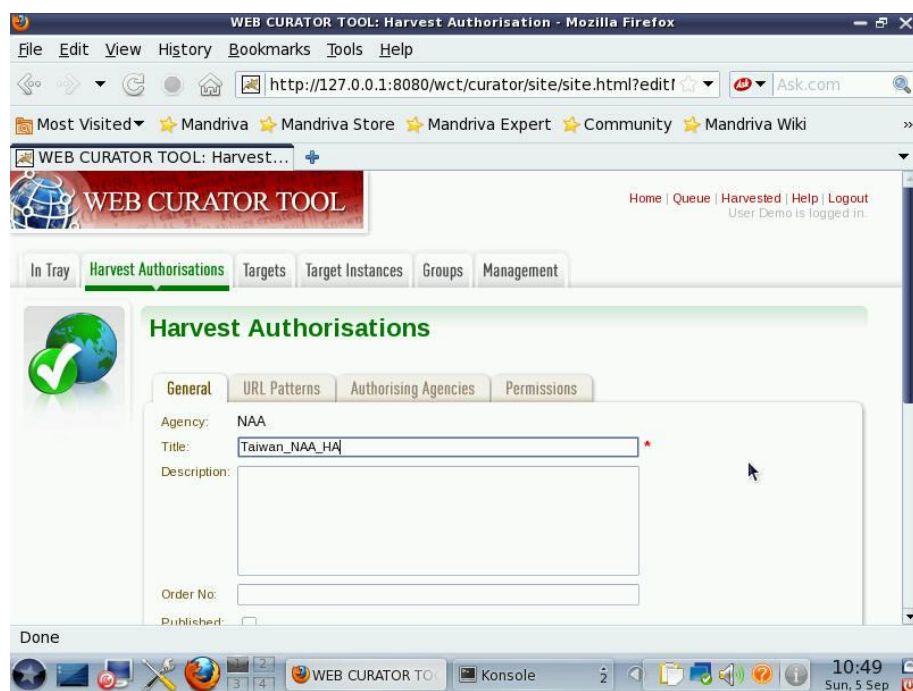
附圖 4- 31 頻寬設定修改畫面

(十一) 如附圖 4- 32 所示，頻寬設定完後，接著請建立新的抓取授權
(Harvest Authorization)。



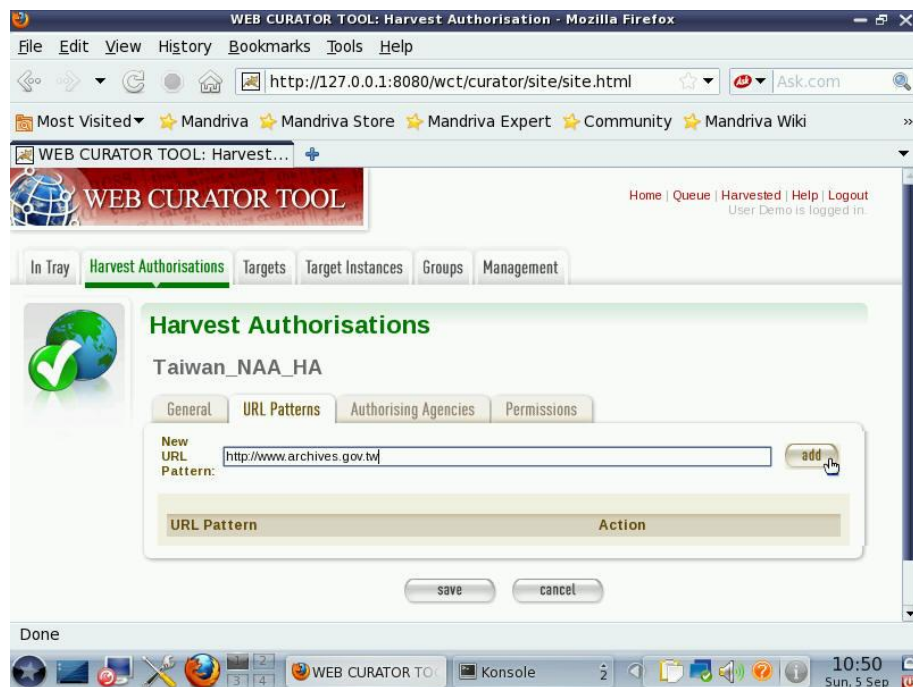
附圖 4- 32 抓取授權新增畫面

(十二) 如附圖 4- 33 所示，填入抓取授權的一般設定資料。



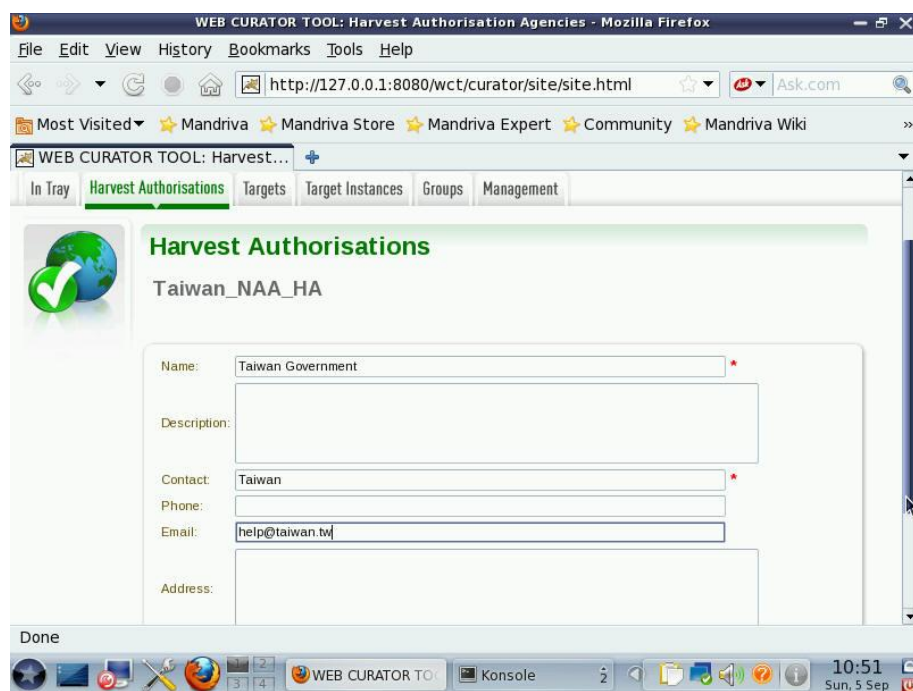
附圖 4- 33 抓取授權一般設定畫面

(十三) 如附圖 4- 34 所示，在新增抓取授權時，請加入檔管局網站位置
http://www.archives.gov.tw。



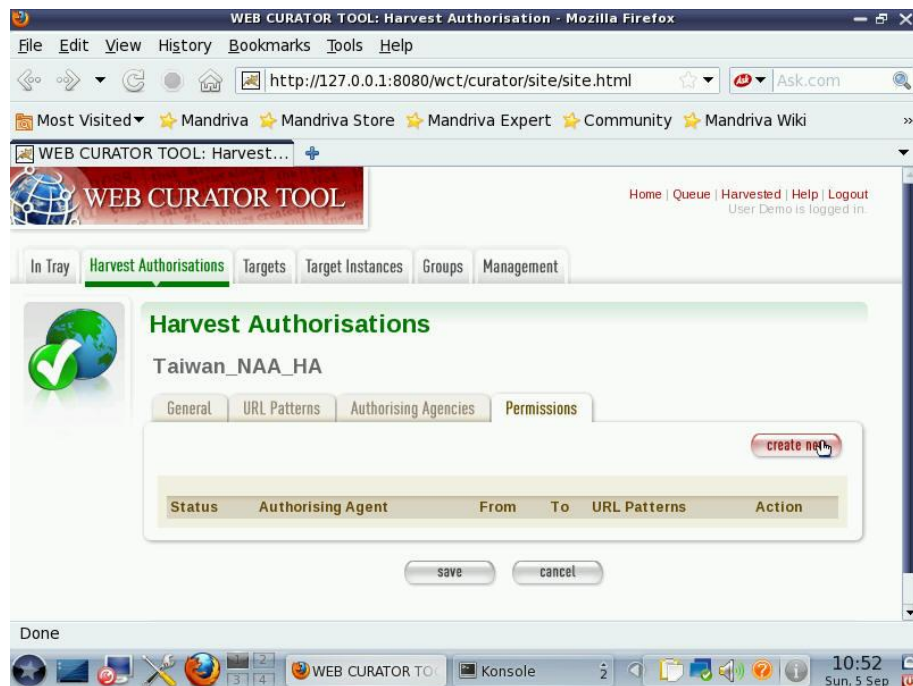
附圖 4- 34 抓取網址設定畫面

(十四) 如附圖 4- 35 所示，請填入抓取授權的相關部門聯絡人資料。



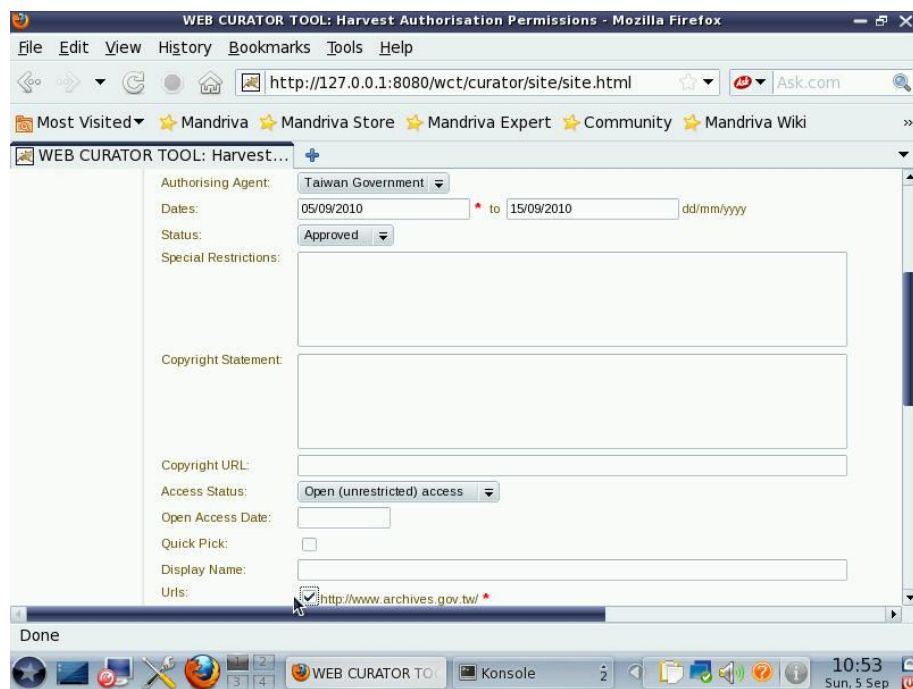
附圖 4- 35 抓取授權連絡人資料填寫畫面

(十五) 如附圖 4- 36 所示，請建立該抓取授權的相關權限。



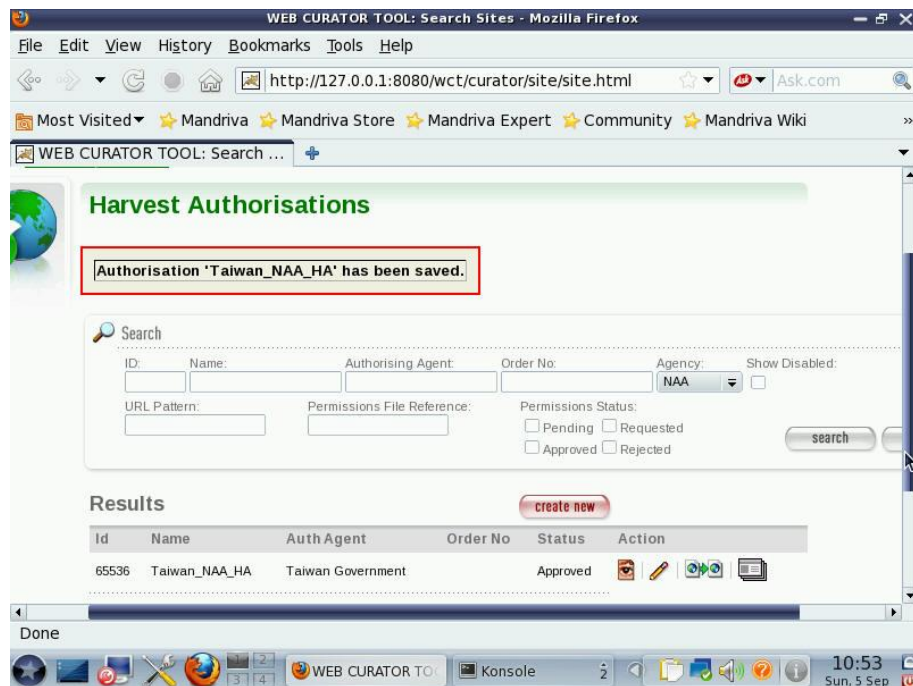
附圖 4- 36 抓取授權權限建立畫面

(十六) 如附圖 4- 37 所示，授權內容包含填入授權日期(Dates)以及授權網址，並且將狀態(Status)改成批准(Approved)。



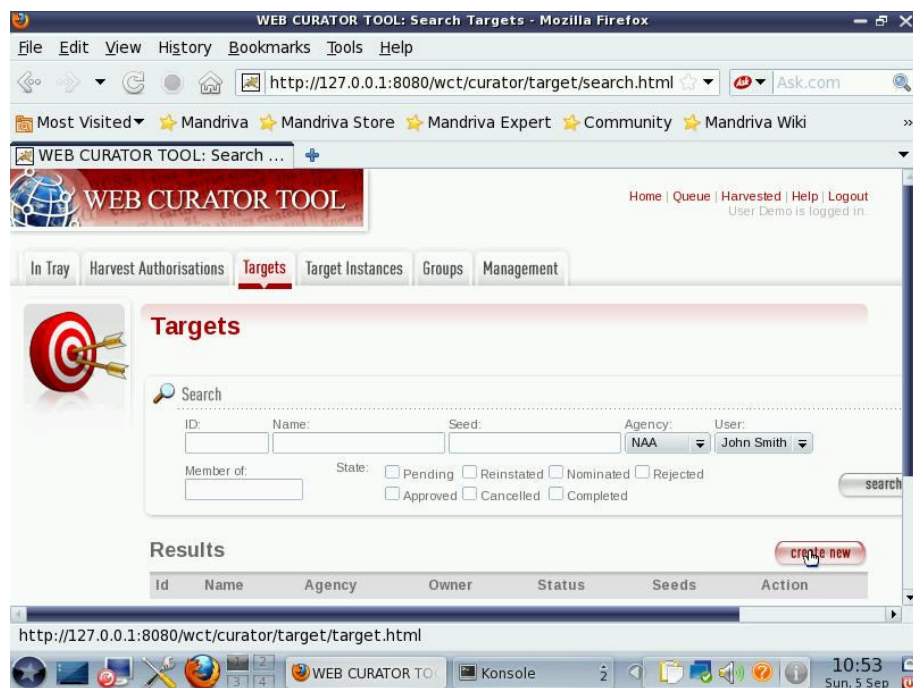
附圖 4- 37 抓取授權權限設定畫面

(十七) 如附圖 4- 38 所示，完成相關權限設定後，抓取授權即建立完成。



附圖 4- 38 抓取授權新增完成畫面

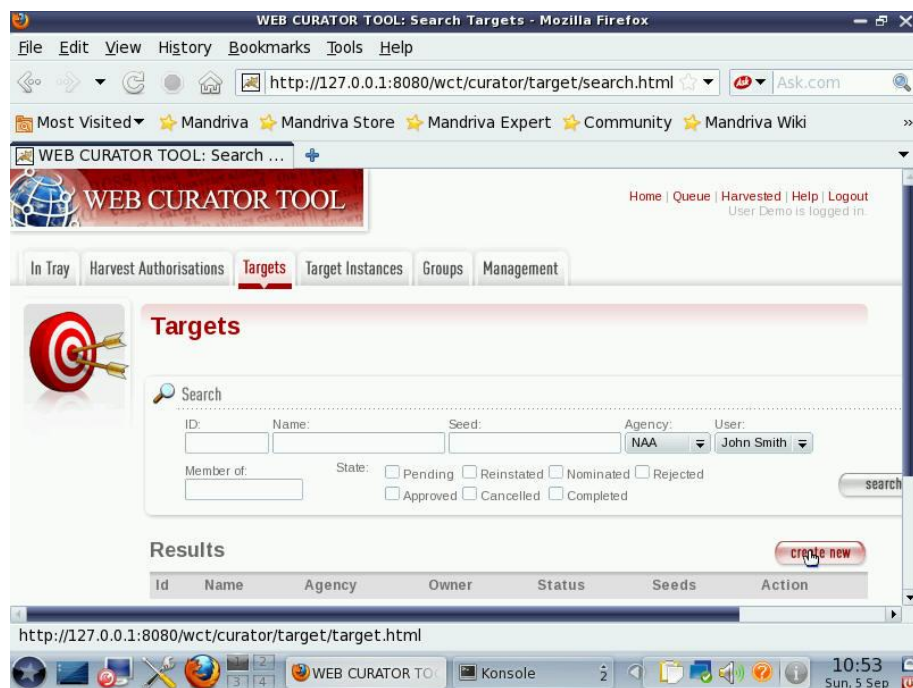
(十八) 如附圖 4- 39 所示，新建抓取授權後，請建立新的目標 (Target)。



附圖 4- 39 目標新增畫面

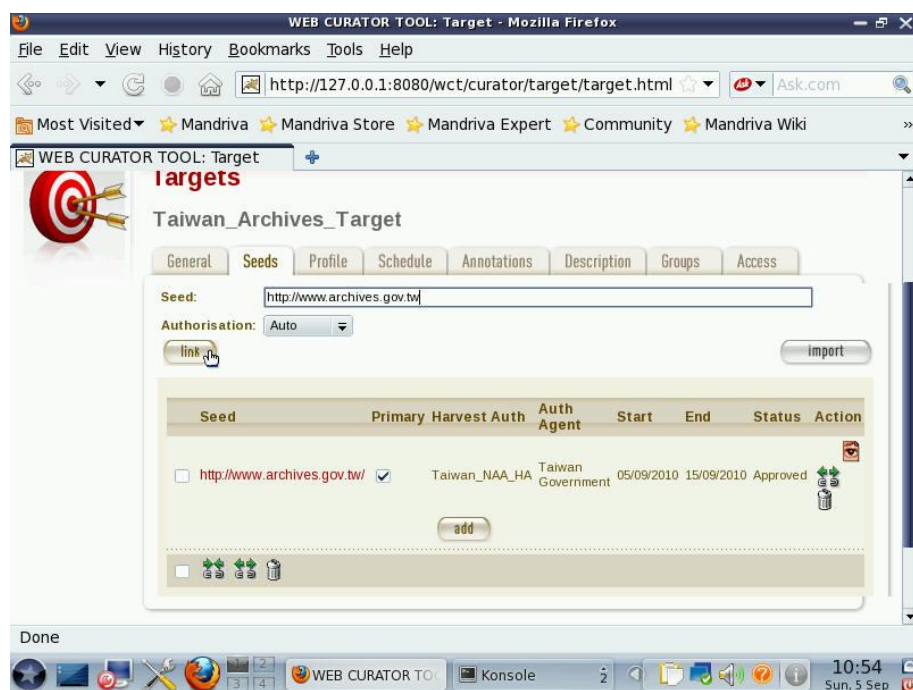
(十九) 如附圖 4- 40 所示，請填入目標的基本資料，並將狀態(Status)

改成批准(Approved)。



附圖 4- 40 目標資料設定畫面

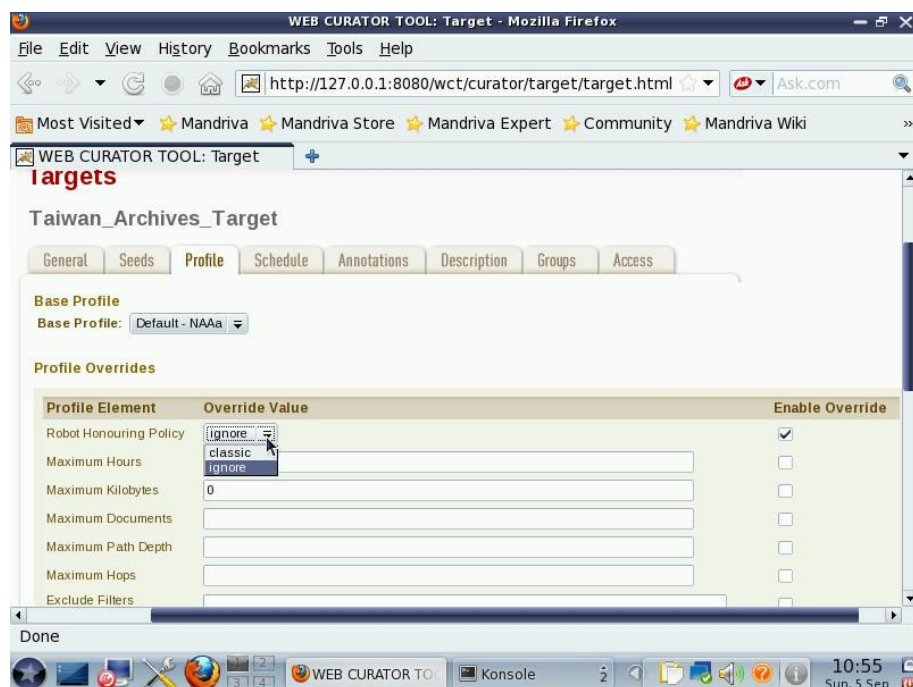
(二十) 如附圖 4- 41 所示，請將檔管局網址與該目標做一連結。



附圖 4- 41 網址鏈結畫面

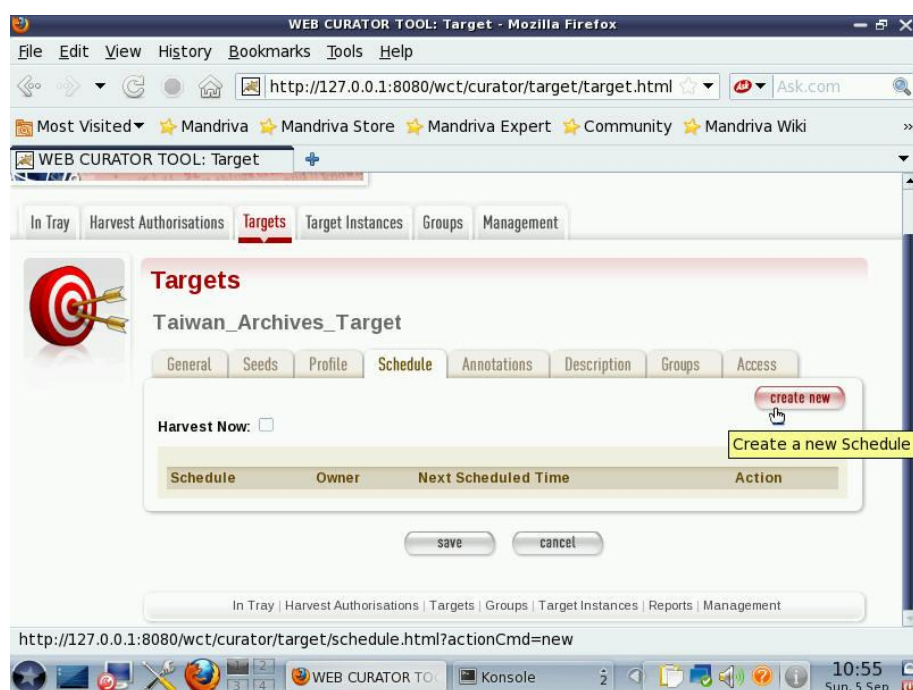
(二十一) 如附圖 4- 42 所示，請在 Profile 內設定 Robot Policy 為忽

略(ignore)，並且將勾取右邊的覆寫設定，因為檔管局網站目前尚未開放公開抓取資料，所以必須在此做設定。



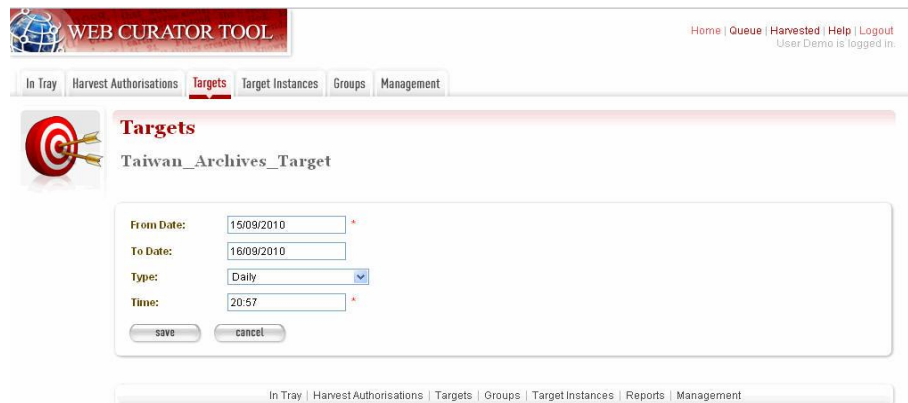
附圖 4- 42 Robot Policy 設定畫面

(二十二) 如附圖 4- 43 所示，將其目標設定抓取排程。



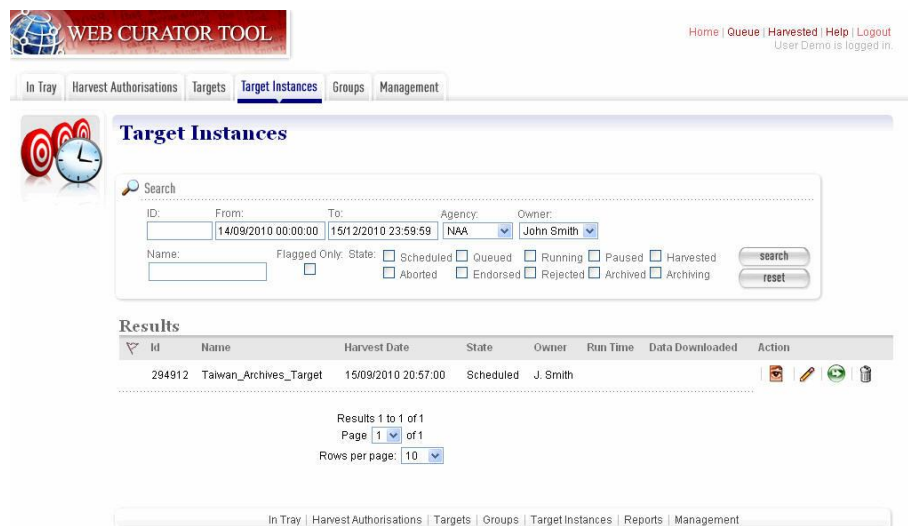
附圖 4- 43 抓取排程新建畫面

(二十三) 如附圖 4- 44 所示，抓取時間可以自訂，本例設定從 2010/09/15 晚上 8:57 開始抓取，直到 2010/09/16 晚上 8:57 分止，每天抓一次。



附圖 4- 44 自訂排程時間畫面

(二十四) 如附圖 4- 45 所示，在目標實體(Target Instance)區塊可看到已經排程的抓取工作，如要直接抓取，可按取綠色按鈕直接抓取。



附圖 4- 45 目標實體畫面

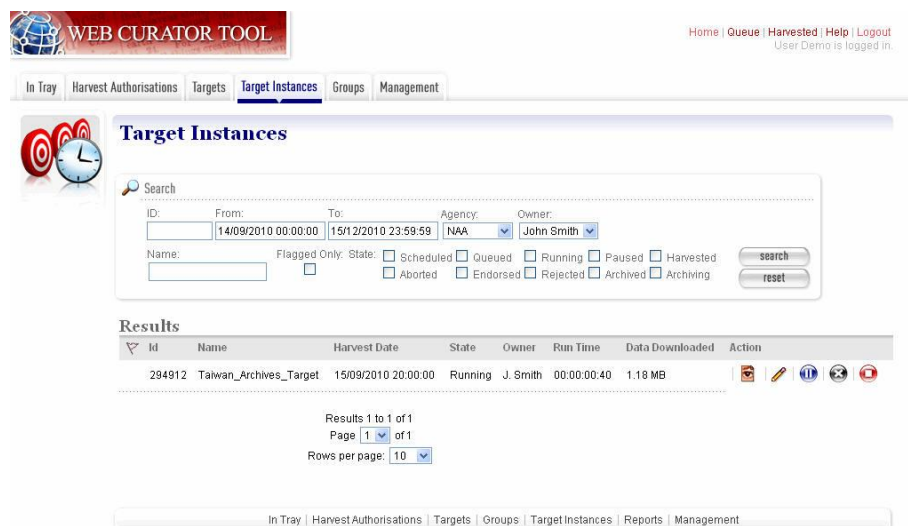
(二十五) 如附圖 4- 46 所示，按下綠色抓取按鈕後，會出現代理人(agent)分配的按鈕，本例預設為最多同時抓取兩個網站，目前有零個正在抓取，按下

紅色的配置(Allocate)鍵後，系統會馬上執行抓取工作。



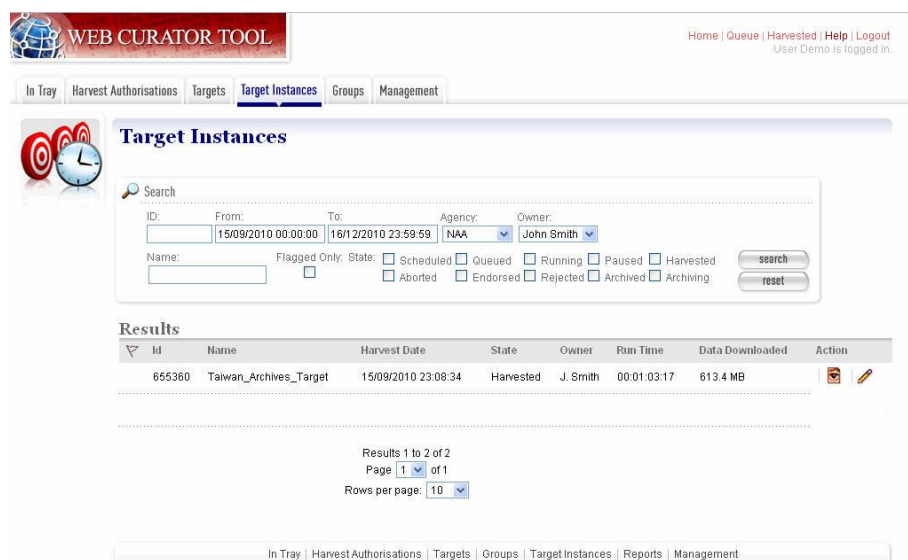
附圖 4- 46 抓取代理人配置畫面

(二十六) 上一步按下紅色配置(Allocate)按鈕後，系統馬上進行抓取工作，目前抓取了 1.18 MB 大小的內容，如附圖 4- 46 所示。



附圖 4- 47 目標實體抓取狀態畫面

(二十七) 如附圖 4- 48 所示，在狀態列顯示網站已經抓取完成(Harvested)後，在動作(Action)區塊選擇編輯，即可觀看抓取結果。



附圖 4- 48 目標實體抓取完成畫面

(二十八) 如附圖 4- 49 所示，請選擇抓取結果(Harvest Results)進到觀看步驟。



附圖 4- 49 目標實體抓取完成選項畫面

(二十九) 如附圖 4- 50 所示，直接點取審核抓取結果(Review This Harvest)即可觀看抓取的網站內容。



附圖 4- 50 目標實體品質審核工具畫面

(三十) 如附圖 4- 51 所示，此網站是利用 WCT 抓取檔管局網站後所存下的內容。



附圖 4- 51 抓取內容範例畫面

三、 利用 Wayback 搜尋及顯示 WCT 抓取之內容

(一) 如附圖 4- 52 所示，等待 WCT 抓取完網站後，開啟 Wayback 服務網址並且輸入檔管局網址 <http://www.archives.gov.tw>，接著按下 Take Me Back 按鈕。



附圖 4- 52 Wayback 搜尋畫面

(二) 如附圖 4- 53 所示，輸入檔管局網址後，出現 2010/09/15 有抓取結果，直接點取該連結，即可觀看抓取內容。



附圖 4- 53 Wayback 搜尋結果畫面

(三) 如附圖 4- 54 所示，此網站是利用 Wayback 所支援的 Arc 檔案，顯示檔管局網站的抓取內容。可以繼續選擇 Archived/Live Comparison，來觀看以抓取的內容以及線上的內容比較。



附圖 4- 54 Wayback 瀏覽保存網站畫面

(四) 如附圖 4- 55 所示，Wayback 提供已抓取網站內容以及線上網站內容比較，左邊為利用 WCT 抓取檔管局網站的內容，右邊則是現在檔管局網站線上的內容。



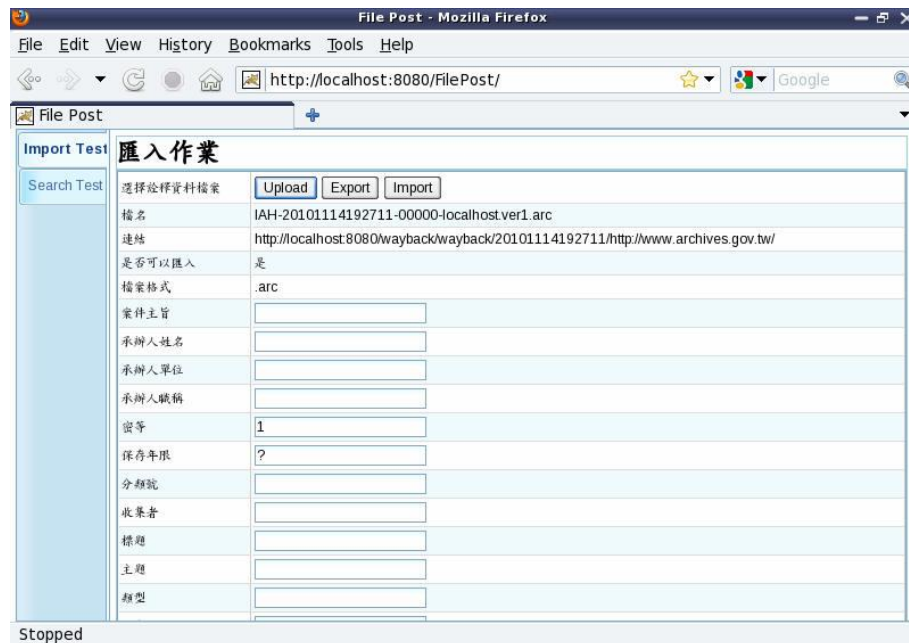
附圖 4- 55 Wayback 保存網站與現況比較畫面

四、 利用雛形系統匯出及模擬匯入 WCT 抓取之內容

(一) 在瀏覽器中開啟 <http://localhost:8080/FilePost/> 網址，即會看到匯入功能的主畫面，左側的頁籤提供了匯入(Import)功能以及模擬匯入之後，對已匯入資料進行搜尋(Search)的功能，如附圖 4- 56 所示。

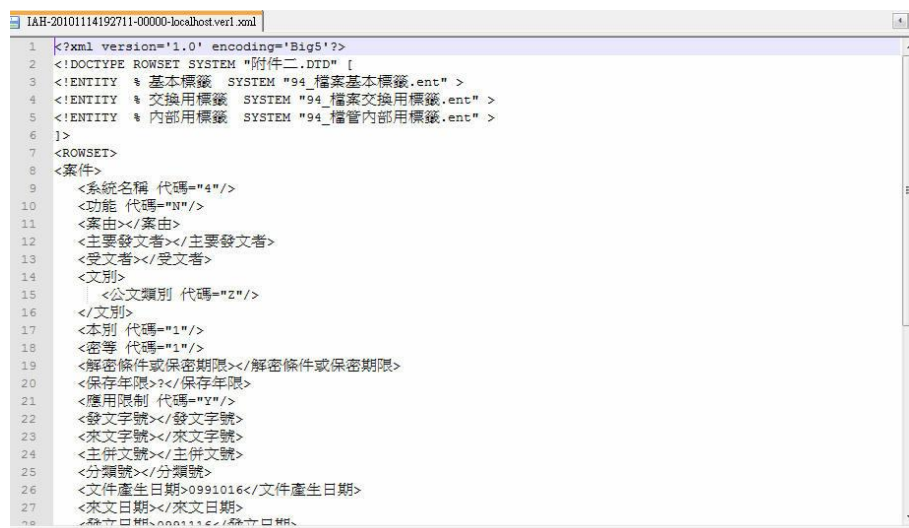
附圖 4- 56 雛形系統畫面

(二) 在匯入作業的畫面中點擊上傳(Upload)按鈕，若是上傳的檔案為符合網站保存格式的檔案類型的話，則會自動帶入部分資訊，其餘則需要使用者進行填寫，如附圖 4- 57 所示。



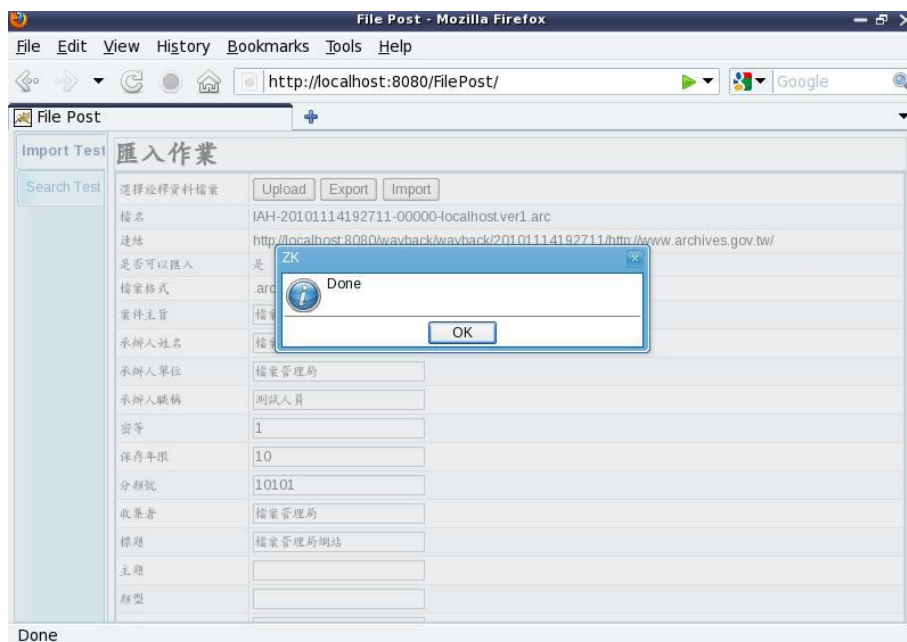
附圖 4- 57 雛形系統上傳檔案完成畫面

(三) 在完成上傳檔案之後，點擊匯出(Export)按鈕可以匯出符合附件二的詮釋資料檔案，下載之後的檔案內容如錯誤！找不到參照來源。所示。



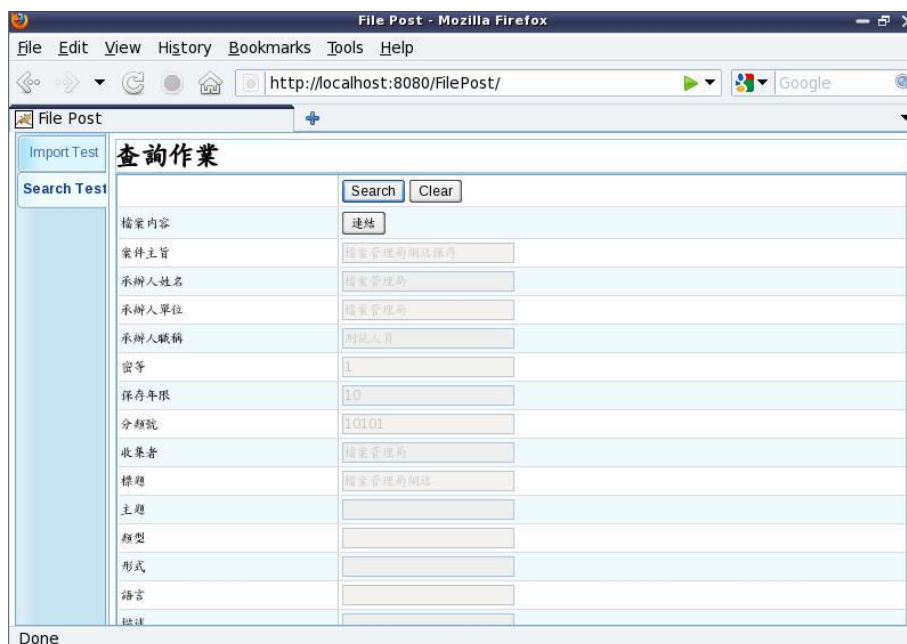
附圖 4- 58 雛形系統匯出之詮釋資料檔案內容

(四) 檔案上傳完畢且相關資料輸入完成之後，可以點擊匯入(Import)的按鈕，會將相關資料匯入到資料庫中，藉此模擬檔案管理系統的匯入功能，成功會出現 Ok 的視窗，如附圖 4- 59 所示。



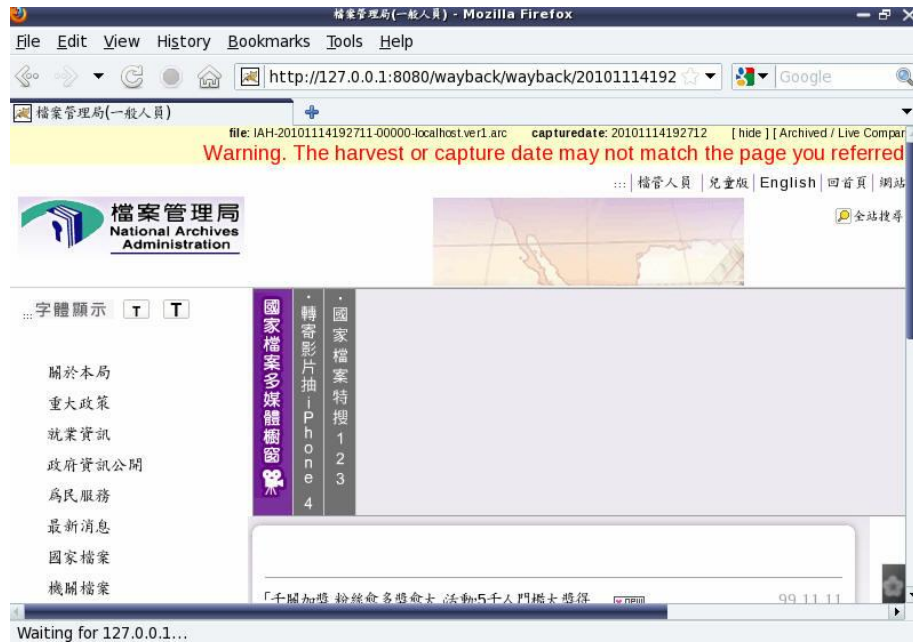
附圖 4- 59 離形系統匯入成功畫面

(五) 在搜尋功能中，可以輸入案件主旨並點擊搜尋(Search)按鈕進行查詢，如附圖 4- 60 所示。



附圖 4- 60 雛形系統搜尋功能畫面

(六) 搜尋成功完成後，會顯示搜尋的結果以及提供檔案內容的連結按鈕，點擊該按鈕即可觀看匯入的網站內容，如附圖 4- 61 所示。



附圖 4- 61 雛形系統網站顯示畫

參考資料

四、 中文參考資料

1. 林巧敏（2009），「網站資源保存計畫之調查分析」，圖書與資訊學刊，68 期，49-65 頁。
2. 洪淑芬（2006），「網路資源備份保存技術與應用探析」，國家圖書館館刊，95 年第 2 期，75-121 頁。
3. 林巧敏 陳雪華〈2008〉「國家圖書館電子館藏發展政策之研究」，國家圖書館館刊，95 年第 2 期，47-49 頁
4. 賴重仁〈2007〉，「國家圖書館網站保存先導系統」，國家圖書館館訊，97 年第 1 期，1-4 頁
5. 王麗蕉（2007）。網路資源長期保存：以多元層次描述模式建構之探討。教育資料與圖書館學，44(4)，455-471。
6. 余顯強（2008）。主題地圖與網頁資源保存之探討。教育資料與圖書館學，46(1)，55-80。
7. 林巧敏（2007），「澳洲國家圖書館數位館藏發展概述」，國家圖書館館刊，96 年第 1 期，185-187 頁。

二、 英文參考資料

8. Brown, A., "Archiving Websites: a practical guide for information management professional", 2006.
9. Brügger, N., "Archiving Websites. General Considerations and Strategies. Aarhus: The Centre for Internet Research", The Centre for Internet Research, 2008.
10. Charlesworth, A., "A study of legal issues related to the preservation of Internet resources in the UK, EU, USA and Australia", 2003, available at:
http://library.wellcome.ac.uk/projects/archiving_legal.pdf
.
11. Day, M., "Collecting and preserving the World Wide Web: a feasibility study undertaken for the JISC and Wellcome Trust", 2003,
http://library.wellcome.ac.uk/projects/archiving_feasibility.pdf
12. Day, M., "Preserving the Fabric of Our Lives: A Survey of Web Preservation Initiatives". Research and Advanced

Technology for Digital Libraries: Proceedings of the 7th European Conference, 2003, pp.

461-472. <http://www.ukoln.ac.uk/metadata/presentations/ecdl2003-day/day-paper.pdf>.

13. Eysenbach, G. and Trudel, M., "Going, going, still there: using the WebCite service to permanently archive cited web pages". Journal of Medical Internet Research, Vol.7, No.5, 2005.

14. Masanès, J., "Towards Continuous Web Archiving", D-Lib Magazine, Vol. 8, No. 12, 2002.

15. "How to Evaluate and Choose an E-mail Archiving Solution", available at:

<http://whitepapers.techrepublic.com.com/abstract.aspx?docid=315472>.

16. "Best Practices in Email Archiving", available at:

<http://whitepapers.techrepublic.com.com/abstract.aspx?docid=238444>.

17. Flynn, N., "Email Retention and Archiving: Manage Electronic Records, Minimize Workplace Risks and Maximize Compliance", 2008, available at:
http://www.usdatavault.com/library/email_retention.pdf.
18. "10 Best Practices for Archiving", 2007, available at:
http://searchcio-midmarket.techtarget.com/tip/0,289483,sid183_gci1159997,00.html
19. "Legal Considerations for E-mail Archiving", 2007, available at:
<http://whitepapers.techrepublic.com.com/abstract.aspx?docid=391734>.
20. "Top Five Strategies for Getting an Email Archiving Project off the Ground", 2008, available at:
<http://whitepapers.techrepublic.com.com/abstract.aspx?docid=927881>
21. Fitch, K. "Web site archiving – an approach to recording every materially different response produced by a website",

2003, available at:

<http://ausweb.scu.edu.au/aw03/papers/fitch/>.

22. Lyman, P., "Archiving the World Wide Web". Building a National Strategy for Preservation: Issues in Digital Media Archiving., 2002, available at:

<http://www.clir.org/pubs/reports/publ06/web.html>.

23. Wagner, F. ; Krebs, K. ; Mega, C., et al. "Towards the design of a scalable email archiving and discovery solution".

Advances in Databases and Information Systems. 12th East European Conference, 2008, pp.305-320.

24. Anon, "Keeping your assets tangible [email archiving system]". Business Info, Vol. 91, No. 91, 2008.

25. Srinivasan, A. ; Baone, G., "Classification challenges in email archiving", Rough Sets and Current Trends in Computing. Proceedings 6th International Conference, 2008, pp.508-5198.

26. Taylor, S., "Beware the pitfalls! [email archiving]", Document Manager, Vol. 15, issue 3, 2007, pp.15-17.

27. Spurzem, B., "Email archiving implementation – five costly mistakes to avoid", Software World, Vol. 39, issue 4, 2008, pp. 3–6.
28. Craine, K., "The strategic importance of email archive management", Records Management Bulletin, No. 136, 2007, pp. 12–20.
29. Perer, A; Shneiderman, B; Oard, DW, "Using rhythms of relationships to understand e-mail archives", Journal of the American Society for Information Science and Technology, Vol. 57, issue 14, 2006, pp. 1936–1948.
30. Yiru Li; Somayaji, A., " Securing email archives through user modeling", Proceedings. 21st Annual Computer Security Applications Conference, 2006.
31. Anon, "MailSafe [email indexing and archiving software]", Document Manager, Vol. 15, issue 1, 2007.
32. Perkio, J. ; Tuulos, V. ; Buntine, W., et al., " Multi-faceted information retrieval system for large scale email archives",

Proceedings. The 2005 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 557-64, 2005.

33. Anon, "Mail, ordered [Johnston Press email archiving]", Document Manager, Vol. 13, No. 4, pp.12-13, 2005.

三、 網站

34. 檔案管理局《Taiwan - National Archives Administration》
http://www.archives.gov.tw/Chinese_archival/Default.aspx

35. 加拿大圖書及檔案館《Canada - Library and Archives Canada》,
<http://www.collectionscanada.ca/index-e.html>

36. 美國國家檔案暨文件署《USA- National Archives and Records Administration》, <http://www.archives.gov>

37. 北卡羅萊納州政府檔案館《Government Records Branch of North Carolina》, <http://www.records.ncdcr.gov/default.htm>

38. 英國國家檔案局《United Kingdom- The National Archives》
web site, available at:

<http://www.nationalarchives.gov.uk/default.htm>

39. 台灣大學網站保存庫(NTU Web Archiving System),

<http://webarchive.lib.ntu.edu.tw/>

40. 台灣網站保存系統(Web Archive Taiwan),

<http://webarchive.ncl.edu.tw/nclwaFrontNew/>

41. Australia's Web Archive (PANDORA),

<http://pandora.nla.gov.au/about.html>

42. Austrian On-Line Archive,

<http://www.ifs.tuwien.ac.at/~aola/>

43. Government of Canada Web Archive,

<http://www.collectionscanada.gc.ca/webarchives/index-e.html>

1

44. Internet Archive,

<http://www.archive.org/index.php>

45. The Library of Congress Web Archives (Minerva),

<http://lcweb2.loc.gov/diglib/lcwa/html/lcwa-home.html>

46. Council of State Archivists

<http://www.statearchivists.org/arc/index.htm>

47. UK Web Archive, 2005

<http://www.webarchive.org.uk/ukwa/>

48. The U.S. National Archives and Records Administration

<http://www.archives.gov/>