網路環境下圖書館數位館藏的建置與評估

詹麗萍　所長
中興大學圖書資訊學研究所

壹、網路環境下圖書館數位館藏的建置

傳統圖書館的館藏發展，過去幾年正面臨經費縮減、書刊價格不斷上漲、典藏空間不足、文獻保存維護不易以及館藏利用缺乏效益等幾個困境。事實上，經常會被使用到的大約是20%的館藏，另外80%館藏使用率通常是非常的低。如果傳統圖書館的館藏發展面臨這樣的困境，那數位館藏是否爲一個解決方案呢？仔細審視所面臨的困境後，發現它可能已不是最好的解決方案了。數位館藏至今也有幾年的發展經驗，事實上它並不會比較經濟，也許還需要更多的支出。那麼數位館藏是否有效益呢？個人覺得也許至今仍有大部分的圖書館還不清楚，理由是用高價位引進了如此多的數位館藏，圖書館是否眞的了解讀者使用的程度，讀者是否覺得數位館藏比紙本館藏更方便，這一點其實也還不太確定，除非有確切的數據，來加以證明。

總而言之，圖書館發展數位館藏的原因，可歸納爲：資訊科技的進步、電子出版的盛行、網路資源的成長與讀者使用習慣的改變。如果讀者習慣在遠距，或任何一個讀者覺得方便的地點，在方便的時間內得到讀者想要的資訊，此時圖書館提供數位館藏才能發揮最大的效益。

一般而言，數位館藏建置的方式有下列幾種途徑：(1) 向出版商購買電子出版品；(2) 將圖書館實體館藏數位化；(3) 運用網路上現成免費的資源。其實目前已有很多圖書館館員，主動在網路上尋找適合提供給讀者利用的資訊，將它放在目錄裡面，以便讀者利用。

一、數位館藏類別

依照儲存媒體、使用模式等的不同，數位館藏可以分爲不同類型，目前國內圖書館所提供的包括光碟資料庫、線上資料庫、電子期刊、電子書…等。

1. 光碟資料庫

光碟資料庫可能是書目資料庫或全文資料庫，這類資料庫通常只能夠買到使用權，無法像買書或訂期刊一樣，紙本館藏資料可以持續的讓讀者永久使用，但數位館藏有版權上的限制，有時無法順利滿足讀者的需求。

2. 線上資料庫

一般線上資料庫均是買檢索權，部分是書目資料庫，另有些則是全文資料庫。

3. 電子期刊

國內圖書館非常熱衷於電子期刊館藏的發展，而且讀者非常喜愛全文的電子期刊，有部份在台灣設有 Mirror Site ，絕大部分則是連線到供應商的資料庫去取得期刊全文。

4 .電子書

目前國內淡江大學圖書館已提供 Net-Library電子書服務，逢甲大學正推動電子書聯盟，全稱是 Taiwan E-Book Net ，簡稱 TEBNet，目前約有二十幾個圖書館計劃參與。電子書的引進發展很快，這 必然成爲一個發展新趨勢。

5.「隨選視訊」(Video On Demand，VOD)

國內有幾所大學圖書館已發展隨選視訊系統，提供視訊資訊的使用服務。它牽涉到網路的頻寬、版權的問題，所以並不是那麼容易發展。如果圖書館有意想要建置「VOD」系統，首先要評估圖書館需要什麼內容、要如何取得內容以及克服技術上的問題。

6. 網路資源

圖書館可選取網路上值得參考的資源，加以編目整理，提供讀者利用。

二、選擇數位館藏的考慮事項

圖書館建置數位館藏之前必須訂定館藏政策，依照館藏政策執行。除了符合館藏政策外，選擇數位館藏時還有一些相關事項需特別注意。

1. 確認作者及出版者的信譽

紙本資料的選擇可使用選書的工具，例如：書評。而數位資源或資料庫，則可能需要試用並加以評估。

2. 檢查資料實體及內容

資料庫供應商通常會先提供測試版，圖書館可先行試用後再決定是否採購。

3. 確定有足夠的技術及設備

4. 檔案格式要與設備相容

5. 方便利用

有些資料庫廠商要求讀者安裝檢索的軟體，要考慮這種不方便的程度，是否爲讀者所接受，因爲它會影響到將來這些資料庫的使用及效益。

6. 考量價格

採購方式與買書不同，很多資料庫是不

賣的，僅能購買檢索權或使用權。

7. 了解合約的內容

注意合約的條文規定是否合理。

8. 購買「使用權」還是「擁有權」

有關這個議題的討論文獻相當多。圖書館曾經幾乎相信 Access可以取代 ownership，認爲資料在哪沒關係，只要能提供讀者使用就好。但有另一派學者覺得很憂心，考慮到以後永久典藏資料到底該由誰來負責？所以圖書館應該尋找一個平衡點，由圖書館互相合作建立豐富均衡的館藏。

三、數位館藏採購方式

數位館藏可以透過不同的管道採購，如：直接向出版商購買、請代理商代爲購買、或是以聯盟的方式購買。以下就代理商及聯盟方式進一步說明。

(一) 代理商模式

常見的代理商服務模式多半以提供的整合性服務爲主要訴求。相關服務包括：

1. 西文期刊全文資料庫的查詢及利用。
2. 電子期刊的傳遞系統，如圖書館訂購某種期刊，就提供其電子期刊使用權。
3. 提供文獻傳遞服務。
4. 提供期刊目次。
5. 圖書館線上公用目錄及網站服務：例如代理商 EBSCO，Swets， Silver Platter 等所提供的服務。

(二) 聯盟模式

早在1930~1960年國外就有Consortium的形成。所謂 Consortium是由一些圖書館結合形成聯盟，主要目的是跟政府爭取經費，以進行合作採購、合作編目等計劃。1980年代，則藉由自動化系統及加入大型的書目中心，進行書目資料共享及館際互借的活動。到了1990年代，有許多聯盟的產生目的是爲了電子資源共享，像國內的 CONCERT ，就是爲達到這樣的目的。

國內 CONCERT 所引進的國外電子資源，其訂購方式不外有兩種。一種是全國性的，就是National Academic Licensing，是由國科會科資中心主導，教育部提供部分補助。另有一種型態是由圖書館組成一個大的 Group ，然後再與資料庫的廠商談判， Group Purchasing 是由各館用自己的經費來購買，因爲每個圖書館讀者需求不太相同，所以其下還有好幾個子聯盟。聯盟的共同採購可結合成員集體力量，爭取合理的價格，由聯盟出面談判爭取最大的權益，另一方面又可節省人力物力，並提昇服務品質。 CONCERT 引進的方式：第一就是建置 Mirror Site，例如國內的科資中心、中研院、交大就有 Mirror Site。第二就是直接連線到國外廠商的主機。

四、數位館藏採購合約的簽訂

由於數位館藏使用模式不同於以往，因此在簽訂採購合約時，除了一般合約需注意的事項，針對數位館藏的利用還需注意或釐清下列事項。

(一) 界定使用者與使用地點

數位資源最大的好處就是可以多人同時使用，但是廠商爲了維護自己的利益，會限制使用者的人數及使用者的身份。如果以使用地點來判斷會比較省事，例如校園網路，可以從IP來判定，若是從使用者來判定，則略爲麻煩。如果系統設定爲教授、學生、職員才能使用，當一個陌生人走到這個校園裡面，用任何一部電腦，雖然圖書館知道此讀者並非合格使用者，但此讀者目前在規定的使用地點裡，圖書館該不該讓此讀者使用並爲其使用行爲負責？所以圖書館館員必須要謹慎，因爲已與廠商簽約，如果廠商發現有異常的使用狀況，圖書館可能必須負起相關的責任。

(二) 提供使用者遠端認證的方法

廠商應提供彈性的使用者認證方法。例如我目前的工作地點在台中，週末回到新竹的家，如果用使用地點來限制，那麼我在家就屬於不合格的使用者。如果可以用遠方認證，只要查明我是合格使用者就可以使用，那對使用者來說就方便多了。雖然要做一些控管的工作，可是對使用者來說是方便的，只要符合使用者的需求就值得圖書館去努力。

(三) 價格談判—降低價格

若價格差距很大，圖書館有需要和廠商協商。

(四) 訂定用途

所謂用途就是說圖書館能夠如何使用資料庫，有的資料庫限定讀者可以瀏覽但不能列印或下載，或其他種種限制。這些都要在合約裡訂定清楚，以免衍生糾紛。

(五) 訂定合理的使用權

圖書館提供書刊資料時，是用fair use的法律保障，在圖書館、在教學場合使用資料是被允許的，是合乎法律規定的。資料庫的合理使用範圍也應該清楚記載在合約內。

(六) 訂定館際互借數量

國外圖書館有集體訂購一種電子期刊的聯盟，它是可以館際互借的，但有數量限制，這些相關規定都應寫在合約內。

(七) 可備份或永久保存

若是將來圖書館經費減少，必須刪訂時，是否可將已訂購的資料備份或永久保存呢？過去圖書館都被動的接受廠商的規定，沒有去爭取，這樣到底可不可行，圖書館必須要試試看，否則圖書館保存資料的功能就沒有辦法充分發揮。

(八) 賠償條款

合約簽定後，如果有任何一方違法，要怎樣去處理？賠償條款是否訂的很清楚呢？例如美國德州大學（UT Austin）的一個網站，裡面有很多關於採購合約的文件，甚至還附有 checklist，列了很多重點，可以逐條檢查合約裡是否充分合理，或者面對這個條文圖書館應該要如何去做。

五、「館藏數位化」的作業流程

除了以採購方式取得數位化館藏外，自行建置數位化館藏亦是圖書館可以考慮的途徑之一，但在投入成本以及永續經營上必須有所規劃。

(一) 文獻的選擇

館藏數位化的第一步驟即是文獻的選擇。選擇哪些文獻值得數位化，或者是順序上哪些應該是要優先被考慮的。另外還要考慮到數位化的目的是爲了「利用」還是爲了「典藏」，例如數位化的目的是爲了「典藏」，因爲要把它轉換在另外一個儲存媒體上收藏，必須要注意它要比較符合原來的面貌。但如果目的是爲了「利用」，圖書館可將它放到網路上讓大家都能廣爲利用。另外，在網路上傳輸必須要考慮到速度、方便性，所以可能採檔案壓縮的方式來處理。

(二) 可行性研究

探討所選定的文獻是否適合數位化，是否實際可行？選定文獻的總量是多少？它的屬性是什麼？是否爲紙本？有無其他型態？比如說立體物件，若要把它數位化，就不像紙本單純掃描那樣的簡單，這就表示需要比較複雜的配備。另外要預估所需的人力、設備及所需耗費的時間。

(三) 確定系統架構

建立系統前必須確認檔案格式、儲存的方式、建立怎樣的 Database、前端做網頁、如何連到資料庫、整個硬體的架構、使用怎樣的軟體等等，這對一個小圖書館來說，其實是個非常沉重的負擔，所以利用國家級的計劃來做，應爲一好的建議。

(四) 建立數位化檔案

若是文字資料要轉成怎樣的檔案格式呢？若是非常精細、非常接近眞實的圖案，要選擇怎樣的檔案格式？如果一個檔案是要永久保存的可能要比較精細，如果是在網路上傳輸，可以壓縮成較小的檔案，或者轉成兩個或三個不同的檔案以便達到不同的目的。所以數位化檔案的數量非常的多，檔案的命名也要非常注意。

Metadata 其實有很多種，MARC 也是 Metadata 的一種。在十一月一日交通大學的研討會中，中正大學洪政欣教授就談到 Metadata 其實有很多缺點，以其參與過的蘭嶼計劃爲例，展示了一些相關圖片，其實那些圖片是有故事的，因爲做 Metadata

是以一張一張圖片來做描述，也就是說以 Metadata 根本不能夠提供 user 需要的充分資訊，讀者查詢的時候是查到一張一張的圖片，讀者不能得到故事性與完整性的資訊。Metadata 其實不只有洪教授所提到的這些缺點，可能做 Metadata 的人都很疑惑〝我該怎樣去呈現我所描述的東西〞因爲 Metadata 尚在發展中，不像 MARC 已經非常結構化且修訂了很多次。

(五) 建立資料庫和網頁

數位化計劃的團隊要有各種不同專長的人，至少要有內容方面的學者專家，要有懂得寫程式的人，美工的人，更要有懂得推廣的人。

貳、網路環境下圖書館數位館藏的建置

要有效提供資訊服務，需針對流程、業務、利用等不同項目進行了解與評估，做爲服務規劃的參考。傳統圖書館的評估已有具體的模式及方法，數位館藏是否有不同的評估準則？以下就傳統圖書館以及數位館藏的評估進行討論。

一、傳統圖書館的評估要素

目前對圖書館的評估可以從圖書館的使用、資訊流通狀況、資訊檢索機制等項目進行分析。

(一) 圖書館的使用

圖書館的使用是我們評估圖書館營運成效的一個要素。

(二) 館藏使用的狀況

圖書館的流通情形如何？每天大概有多少書被借出？那些在館內使用的資料，例如裝訂期刊，是否被充分使用？假如圖書館花了很多精力去建置館藏，可是讀者卻不喜歡利用，這也許表示圖書館選擇書刊的方法有問題，使得讀者無法從館藏中獲取所需的資訊。

(三) 館藏的檢索(OPAC)

讀者不管是使用哪個檢索點檢索，查到資料的比例是多少？從這裡可以看出讀者對於圖書館的館藏是否有興趣，如果圖書館的館藏質量不錯，讀者每次都能得到需要的資料，這樣讀者一遇到問題，就會想到這個圖書館來試試。

(四) 圖書館系統的使用

圖書館可能不只一個系統，所以需要了解整個系統經常來查詢的使用者是誰？讀者查獲資料的成功率或失敗率如何？以便日後做爲系統改進的依據。

二、數位館藏的評估要素

數位館藏的評估可以從三方面來看。

(一) 資訊的可及性 (accessibility)

資訊的可及性指的是讀者是否可以很方便使用。如果圖書館做很多的限制，讀

者使用困難，相對的就會減少使用。根據國外調查分析，發現很多讀者因爲麻煩所以中途就放棄，所以這是 accessibility 不足、不可親近，如果這是造成讀者半途而廢或沮喪的原因，那麼圖書館就必須要檢討了。

(二) 資訊的可用性 (usability)

資訊的可用性強調的是提供的資訊好不好用。讀者login之後，到每下一個指令及取得所需的文件爲止，是否可以很方便的得到他所需要的資訊。例如圖書館有許多全文資料庫，像 PDF文件就需download一個 Adobe 軟體才能使用，讀者可能覺得很麻煩或不曉得該如何操作，就這樣放棄了，結果最後不能達到讀者的目的，如果這個失敗率很高，那就表示有問題。資訊的可用性，有人認爲最簡單的方法就是看它的取得成功率。

(三) 資訊的使用 (use)

資訊的使用則是了解讀者各式各樣的使用狀況。例如想知道讀者 login 之後都做些什麼？只是瀏覽？列印或下載呢？還是選擇將全文資料 e-mail 回他的信箱？讀者的期望、使用與不使用的理由等等，這些評估數據都可以做爲改進的參考。

以上提到的這些評估要素，部分是屬於計量的方法，但有些目的是不能達到的。例如要知道使用者心裡期待什麼？除非有辦法把它化成可觀察的東西，又如讀者想要什麼？是不是可以從讀者 login 以後的行爲去猜，這是比較困難的，不過也並不是不可能得著。再來則以電子商務所使用的 E-metrics 來說明，它適不適合全盤拿來做爲數位館藏的評估仍待探討，但至少是一個可以參考的依據。

三、數位館藏評估所需要的數據

數位館藏的評估可從量化及質性兩方面著手，在量化方面，數位館藏評估的基本數據包括館藏量、使用量、使用時間。質性方面則包括爲何使用以及如何使用。

(一) 使用什麼？

讀者利用哪些數位館藏？圖書館訂了很多電子期刊，也購置了很多資料庫，哪些資料庫是常常被使用的？哪些資料庫是從來沒有人使用或用的人很少？哪些電子期刊沒有人使用？這些都是圖書館必須要得到的數據。

(二) 誰在使用?

如果從 log file 來看，只能知道讀者從那個IP 進來，卻不知道是誰，如需要這樣的資料，就必須想辦法取得這個數據。

(三) 如何使用？

可以從 log file 來了解讀者的動作。

(四) 什麼時候使用？

(五) 爲什麼使用？

這項數據似乎比較難了解，也許要配合其他的方法來評估。

四、評估工具

E-metrics 是電子商務評估網站的一個工具。它是用科學方法來檢驗網站通行數 (traffic data) 的工具，測量一個網站成功程度的計量法。大家都知道 Bibliometrics 是書目計量法，E-metrics 還不知道怎麼翻譯。電子商務測量網站成效，基本上都是從量上面來看，目的是要了解網站使用者的行爲，還要確定網站是不是吸引適當的使用者來訪，也要評定網站經營策略是否得當， E-metric 已被使用一段時間，有很多相關的文獻可以參考。

E-metrics 實施的步驟有下列幾項：(1) 記錄網站的通行數據；(2) 利用web-server 的 log file 分析數據；(3) 爲網站行銷策略作改善建議。它的功能是要幫助網路企業界得到充分的資訊，藉以了解如何在競爭激烈的商場贏得勝利。

(一) E-metrics 在電子商務的運用

1. 分析使用者的資料並建檔，提供個人化的服務。
2. 了解使用者消費過程，以發展業務計畫。
3. 發展即時行銷的策略，提高銷售率。
4. 評估網站的內容，據以增修改善。
5. 加強網站的設計。
6. 減少服務成本。

透過數據所得到的分析結果，可以將許多服務納入營運系統中，減少很多人工的服務，如宣傳、廣告、客戶服務等。

(二) E-metrics 分析項目之解讀與運用

通用的E-metrics非常多，可以觀察多達上千個項目。有一家專門提供 E-metrics 電子商務分析服務的 NetGenesis 公司就在2000年完成一份研究報告，建議許多重要的計量方法，下面挑選一些比較容易了解的來參考。

1. 客戶造訪的頻率 (Frequency)
   知道客戶多久造訪一次。
2. 客戶最近造訪的時間 (Recencyness)
   知道客戶是否造訪過後很久都不再來，如果這個客戶最近還常來，那就比較不是問題。
3. 網站造訪人數與潛在客戶百分比(Reach)
   鄉鎮圖書館可能根本就不清楚潛在的讀者有多少，所以七、八萬人口只有五百人來使用也不在乎，而電子商務就蠻在乎的，他如果知道潛在客戶蠻多的，可是網站造訪或購買的人很少，就表示經營策略有問題。
4. 造訪者表示購買興趣(Acquisition)
   網上購物通常會叫你將想要購買的東西放到一個 shopping cart，除非你一次只買一個東西。如果你要買好幾種，當然

是查到就放到 cart 去，那一個 cart 就好像你去大賣場買東西的推車一樣，東西先放進去表示你有明顯的購買興趣。

5. 可能的客戶轉爲眞正的客戶(Conversion)
採取購買行動的客戶轉爲眞正的客戶。

6. 客戶的持續度(Retention)
知道客戶是不是持續造訪。

7. 忠誠度(Loyalty)
知道客戶是否持續購買。

8. 客戶停止購買行爲(Abandonment)
當客戶停止購買行爲發生時，表示是一個警訊。

9. 客戶耗損率(Attrition)
耗損率指現有客戶中停止購買行爲者的百分比。

10. 客戶的動搖率(Churn)
動搖率指耗損的客戶除以客戶的總數。客戶最好是穩定地成長，如果有流失情形，也表示是一個警訊。

事實上還有更多的E-metrics，目前只列舉十個通用的，而且是我們比較能了解的，它還可看出更多複雜的現象，並取得 data 做分析，做爲網站營運策略改進的一些數據。E-metrics在數位圖書館的應用，國外有好幾個計劃，最有名的是 ARL （American Research Libraries）E-metrics的計劃，這是一份相當有內容的報告，分成三個階段，現在第二階段的報告產生了，第三個階段今年年底會完成，也就是表示 ARL 會發展出一套 E-metrics 的方法，來評估數位館藏的使用。美國大概有二十幾個 Research Libraries 參與這樣的計劃，研究的成果可以提供參與研究的圖書館了解數位館藏的使用狀況，並做爲未來決策的依據。這個模式一旦做出來，其他的圖書館也可以拿來參考。另有幾個計劃，有的在英國（U. K. Examines Vendor Usage）、有的是整個歐洲的計劃（European Commission EQUINOX Project）。又如：The Digital Library Federation（DLF) Initiative Statistics， DLF 是美國數位圖書館的計劃，現在已是第二期，第一期有六個大學，第二期有多所的大學參與。

五、數位館藏評估的困難

數位館藏評估在執行上有許多困難，以下就可能面臨的議題逐一討論。

(一) 缺乏標準的定義與計量法

電子商務爲什麼能夠發展出 E-metrics 供大家通用？雖然不是每一家電子商務廠商都用同樣的metrics，但是在上千種的E-metrics中你可以選擇你要評估的部分。而圖書館要發展一套標準的定義和計量法比較困難，因爲每個圖書館的差異性蠻大的，數位資訊的價格差異也很大，不過並不是不可能，也許要調查很多圖書館的意見

之後才能夠發展出來。

(二) 無法取得充分確實的數據

主要問題在很多圖書館是連結到廠商的資料庫去檢索的。如果自己沒有主機，那麼就要請廠商提供，雖然廠商都會提供月報表，但是有多少圖書館能夠判讀那些資料？國外圖書館發展 E-metrics 的經驗認爲，最大的問題在於沒有辦法從廠商那裡得到充分的數據，或是他提供的數據不標準，或是他提供資料的方法不是圖書館想要的。圖書館本身若有問題，那麼廠商勢必也不能配合提供數據，如此便產生更大的問題，這是所有參與 E-metrics 計劃的圖書館一致提出來的難題。

(三) 不易解釋數據

有了 data，可是怎麼去解釋它？電子商務廠商可能會想知道 user 是從哪一個 search engine 連到這裡來，他可能會很樂意去那個入口網站或搜尋引擎刊登廣告。user 從哪裡來、做什麼，這個比較容易解釋，可是要知道user 做那件事的動機，就沒有辦法從 data去解釋了。例如很多人來這個網站只稍作瀏覽便下站，可能從數據來看是失敗的使用，也有可能是他不經意連線進來，路過而已，進來之後沒什麼興趣就下站了，這時候就不太能從數據去解釋他的行爲。

(四) 不易做出結論

因爲很難解釋，所以不易做出結論，我們要清楚原因，才能做爲改善的依據。

(五) 技術日新月異

新的技術不斷推陳出新。當我們引進一個資料庫或做一個數位化計劃的時候，所用的技術在當時可能是新的，可是很快地它又落伍了，讀者可能覺得它缺少一樣什麼功能，並不是它不好，而是它的技術落伍了，這會造成評估工作進行的困難，是否要把技術列爲評估的要項？這一點讓很多做 E-metrics 計劃的圖書館認爲很難進行一致性的評估。

(六) 牽涉到個人隱私

Log file 可以顯示 user 進來的 IP 或身分，圖書館是否可以用這個資料來做分析？且須注意這是否會牽涉到個人隱私？

參、結論

數位館藏在今天的資訊服務機制中，是一個不可或缺的角色。如何評估、建立與利用數位館藏仍舊有許多議題等待我們進一步的討論。從永續經營的角度分析，數位館藏建置有其基本議題。

一、資訊組織問題

metadata 是不是一個最好的資訊組織的方法，它有什麼樣的問題，這是我們要探討的。

二、資訊保存維護問題

數位資訊的保存，事實上是有很多的問題，例如過去建立的一些檔案格式現在還能用嗎？還能讀取嗎？圖書館如果要長久的保存這些數位資訊，讓讀者可以長久使用，必須花很大的心血。

三、資訊的使用權問題

包括合理使用等問題。

四、經濟問題

此議題牽涉的層面可能比較廣，以前圖書館談收費或免費，現在我們已意識到數位資訊的成本是否合理的問題。廠商到底是用什麼樣的計價模式，我們可能要想辦法知道，就像過去建立傳統館藏，要熟悉出版界動態及出版的情形一樣重要，但我們對數位資訊的價格了解很少，所以未來圖書館要有經濟學方面的專家，要有法律方面的專家。

五、推廣教育問題

假如圖書館經過評估之後，了解database本身的設計及品質沒有問題，但仍然有很多user不知道該怎麼用，這就是推廣教育的問題。

CONCERT網址：

http://www.stic.gov.tw/fdb/consortium/serv.html

本文爲「知識經濟時代之圖書館服務系列〈一〉研討會」演講記錄，由鄒承平先生記錄，並經主講者寓目同意刊登。