CBETA電子佛典集成

CBETA Chinese Electronic Tripitaka Collection

eBook

ZW09na053

談當前數字圖書館的幾個問題

財團法人 佛教電子佛典基金會

目次

- 編輯說明
- 章節目次
  - 編者按
  - 1 概念淺說
    - 1.1 什麼是數字圖書館
    - 1.2 數字化的方式
  - 2 弊端分析
    - 2.1 檢索問題
    - 2.2 標準問題
    - 2.3 版權問題
    - 2.4 使用問題
      - （1）速度
      - （2）費用
      - （3）習慣
      - （4）健康
  - 參考文獻
- 卷目次
- 贊助資訊

編輯說明

- 本電子書以「CBETA 電子佛典集成 Version 2025.R3」為資料來源。
- 漢字呈現以 Unicode 3.0 為基礎，不在此範圍的字則採用組字式表達。
- 梵文悉曇字及蘭札字均採用羅馬轉寫字，如無轉寫字則提供字型圖檔。
- CBETA 對底本所做的修訂用字以紅色字元表示（黑白色系版本則無顏色差別）。
- 若有發現任何問題，歡迎來函 service@cbeta.org 回報。
- 版權所有，歡迎自由流通，但禁止營利使用。

談當前數字圖書館的幾個問題

張伶　張寧

編者按：

數字化的浪潮正在衝擊全世界，包括佛教界。數字圖書館的出現正在逐步改變人們獲取信息的思維方式和行為方式，是圖書館必然的發展方向。佛教典籍的數字化進程也已經開始，漢文大藏經由此進入第四個時期——數字化時期。但目前這一進程目前剛剛起步，尚存在諸多問題需要探討和解決，如檢索、標準、版權、使用、重復勞動等。本文對當前數字化圖書館的若干問題進行了深入地思考，切中時弊。使我們對現階段全國數字圖書館的狀況，有一個較為客觀的、符合實際的認識。相信對我們的佛教典籍數字化的工作，也一定大有補益。

1 概念淺說

1.1 什麼是數字圖書館

目前對數字圖書館的一般理解是：不需要館舍、虛擬的、以數字轉換方式貯存大量文獻信息資源、方便地為公眾提供數字化方式檢索、閱覽文獻資料的圖書館。它體現為基於因特網分佈式的信息存儲與管理，支援信息的普遍存取，集成化、個性化信息檢索與服務這三個特徵。只要有電腦，任何人都可以極為便利地使用分佈在全球各地的圖書館中龐大的信息資源體系。

客觀地說，在全球範圍內至今還沒有一個大家公認的、標準的定義和理論闡述，可謂眾說紛紜。其實踐狀況，在歐洲及美、日等發達地區和國家也還是處於研究開發階段，並不比我們的現狀先進多少。

1.2 數字化的方式

目前有圖像掃描、錄校排、全息數字化三種方式，見下表。

方式	優點	缺點
圖像掃描	生產成本低	無檢索功能
	可保留原始版面	佔用空間大

		顯示效果差
		不能從版面上摘錄文字
錄校排	可實現文本摘錄和全文檢索	原版面信息無法完整保留
	佔用空間小	生產成本高
		誤差率高
		不能規模化生產
全息數化	可完整保留原版面的全部信息	佔用空間比錄校排大（但比圖像掃描小）
	檢索手段廣泛	

從表中可以看出，根據不同的使用需求三種方式各有長短，從製作技術和使用效果來看，前兩種方式明顯遜色於第三種，由於技術的成熟與普及的原因，實際上目前在國內最廣泛使用的還是第一種，即初級階段的方式。真正的數字圖書館除文字外還應包括有聲影像，現在採用的光碟刻錄只是初級形式，在概念的認識上不應過於把問題簡單化。

現在國際上較流行的是正文、圖像、視頻、音頻四種方式，對傳統紙質書刊的數字轉換主要是採用正文和圖像方式，國內亦大體如此，且一般認為正文方式將是發展方向。

2 弊端分析

從目前已有的實踐看，無論國外還是國內，數字化圖書館都面臨著以下幾個問題的困擾。

2.1 檢索問題

從數字圖書館的發展來說，應該有一種統一的、標準的分類方法，讓讀者通過電腦網絡查詢、檢索信息，快速準確地查到所需的數字化信息內容。但是目前的信息檢索方式還處在開發摸索階段，面對網上海量的信息源，用流覽方式根本無法滿足需要，所以檢索的效率和速度是數字圖書館面臨的最大障礙。

具體而言，面對不同的文獻載體形式，可以選用不同的數字轉換方式，而隨著轉換方式的不同，其檢索狀況也不相同。這就產生了一個問題，那就是數字化的目的是什麼，是僅僅為了節省藏書空

間，還是為了更便捷地使用？恐怕是應以後者為主。如前表所述，對古籍善本，採用圖像掃描方式可以很好地保留文獻原貌，但無法檢索。而對一般文獻採用錄校排的文本方式就便於檢索，但又無法保留文獻的完整原貌。

在同樣以「用」為主的前提下，檢索就成為使用的難點。從理論上說圖像掃描後進行OCR識別，再進行人工標引就可以達到檢索目的，但實際上人工標引所涉及的費用和人力，無疑將提高數字化的成本，而錄校排文本方式因為可以全文檢索就佔據了優勢，將來圖像掃描方式會漸次被其取代，至於全息數字化方式，因其技術要視IT業技術的發展而逐步完善，故目前尚不成熟，還未達到廣泛應用的階段。

2.2 標準問題

從概念上說，首先如果數字化標準的格式不統一，就不是圖書館的數字化。一個或若干個圖書館搞不成數字化，它應該而且必須是一大群圖書館或信息提供單位的聯手運作，所有參與者都按照統一的格式、統一的標準進行數字化，其產品才能便於用戶查閱使用。其次，數字圖書館是一項投資巨大的系統工程，有了標準不僅可以降低資源數字化過程中的研發成本及數字圖書館的建設成本，而且還能加強質量控制和提高生產效率，便於推廣。

但是，至今國家尚無一個權威性的、統一的可供具體操作的標準。利益驅動使得一些圖書館，甚至與圖書館根本不搭界的企業都紛紛參與其中，造成彼此資源庫和技術上的互不相容及浪費，在現有的數字化文獻中就已經出現了人力、物力上的浪費現象，比如同一本書，已經被某一圖書館數字化了，那麼其他圖書館就沒有必要因自身的需要將其再次數字化，由於無法協調造成了同一本書被不同的圖書館重複數字化，等於是若干次無意義的重複勞動，形成大量的初級層次的重複建設，這種不計成本的「虧」我們在歷史上是吃得太多了，但就是無法解決。再者，由於各圖書館的資源建設側重點不同，使用戶在眾多的數字化文獻面前難割取捨，無所適從，致使部分用戶可能支付兩倍或兩倍以上的費用，去選擇兩家或兩家以上的數字化產品。這也涉及相容問題，而相容問題說到底就是標準問題，如果說統一標準可以避免重複建設，那麼，規劃、制定這個標準就是首當其衝不可回避的問題。嚴格地說，這個標準還應當是國際性標準，參與者共同執行全球數字化的統一標準，做到世界範圍內的資源分享，避免不必要的重複建設。

2.3 版權問題

目前，國內外的數字圖書館都面臨著版權問題的困擾。著作權者抱怨被侵權,未經本人同意就將其作品數字化,嚴重損害了他的利益；出版商指責網路數字技術造成了大量的、無法統計的變相複印本，直接影響了他們的經濟效益等等。但如果獲取信息的主動權完全控制在版權擁有者手中，誰的出價高，就給誰數字化，勢必形成有錢才能使用信息的情況，而這又與知識的創造與傳播職能相悖。

保護作者、出版社的權益是極其重要和不可忽視的問題，但事物都是兩方面的，若對其過分地保護則必然會影響信息的傳播，實為兩難之事，對此比較一致的看法是著作人授權並有償使用其作品，這不失為一條可行之路。然而這還不是問題的全部，文獻數字化的目的是便於為公眾服務，即使上述問題能夠得到比較妥善的解決，接下來就是要從技術上解決數字化文獻本身不被盜版的問題，比如圖書館花費很大力量將文獻數字化為純文本供讀者使用，在網上傳播時很可能被人隨意盜取，全部轉過來據為己有，這就是數字圖書館自身的版權保護技術尚未解決的問題。所以說在保護不被盜版的前提下更好地提供服務，既是法律問題又是技術問題，其實質就是利益的重新整合。

現在出版界正在與法律界探討關於文獻數字化的版權問題，努力尋找兩全之策，使版權的法律不斷完善，既有效地保護著作者的知識產權，又便於為讀者提供服務，保護他們作為使用者的利益。客觀地說，版權作品的數字化，需要完善的法律去維護和保護，而要做到這一點並非易事，因為在那些數字化先行一步的發達國家，也是版權糾紛筆墨官司不斷，依然存在著數字版權立法不完善，版權缺乏必要的法律保護等問題，無法從根本上解決數字化圖書館中的這些矛盾。

2.4 使用問題

（1）速度

首先，在目前國內因特網帶寬的現狀下，文獻數字化的方式基本上是按頁掃描成像的圖形檔，沒有做OCR識別，故比相同內容的純文字檔案大數十倍以上，使得查閱速度相當慢，面對網上的這些海量無序且紛雜的信息源，讀者無所適從，很難快速查到所需要的信息。其次，現階段（包括今後相當長的一個時期）國內的多數個人用戶是以家庭普通電話線的方式上網的，這種電話線是單回路入

戶，銅線直徑細，信號傳輸能力差，噪音干擾大，使網路信息的錯誤率大為上升並被反復傳遞，這就自然延長了查閱的速度（時間），同時也就意味著讀者要多支付網費或電話費，這肯定會影響他們使用這種從根本上說就不是完全意義上的數字圖書館的積極性。

（2）費用

文獻數字化傳播會給讀者帶來極大的方便，這是毫無疑問的，但即使是以發展的眼光看，真正物有所值並能用得起的人尚屬少數。從現有情況看，那種想像中足不出戶就可得到所需文獻的人更是鳳毛麟角。「到」圖書館去查數字化文獻，這本身就是對數字圖書館的一種諷刺，也違背了數字圖書館的服務宗旨。即使如此，去查信息的人也不多，不論是高校的教授或科研機構的研究人員，不到萬不得已時是不會用這種方式去獲取信息的。道理很簡單，那就是每查詢或下載一個條目都是要支付費用的。一次、幾次尚可，如長期下去，高昂的費用會令人卻步，難以持久訪問。縱然有課題費，如若有其他方式可以解決問題，誰也不會出此下策。

昂貴的費用影響了數字圖書館的發展，使用它的人公費多，自費少，限制了讀者人數，這就是理想與現實之間不可否認的巨大差距。

（3）習慣

從多數人的實際情況來說，基本閱讀方式還是比較喜歡紙質的文獻載體形式，這不僅是幾千年傳承下來的方式與習慣，而且還是一種文化氛圍的體現和反映。試想一冊古籍線裝書在手翻閱研讀，與在螢幕上閱讀，是兩種完全不同的感受，後者的「書卷氣」勢必蕩然無存。傳統文化的氣息是數字化體現不出來的，如紙質、印刷、版別等。這並不是說要食古不化、抱殘守缺，更不是說數字圖書館與傳統圖書館二者之間此消彼長，而是說要尊重客觀現實，對傳統文化中仍有生命力的事物不能一概否定，應該是彼此共存、互為補充、合力發展。

（4）健康

人一生中只有一對眼睛，既不能再生也不能移植。不論是陰極射線顯像管（CRT），還是液晶顯示器（LCD）,長時間地坐在它們

面前閱讀，螢幕對人體尤其是眼睛都有一定程度的傷害。時間長了，不僅僅是視覺疲勞的問題，而且還會導致視力下降、身體乏力，直接影響人體健康。一般說近距離直視螢幕二十分鐘後，眼睛就會有發乾、發澀、視物不清的感覺，所以喜歡大篇幅從螢幕上閱讀文獻的人較以往逐年減少。當然這只是經驗之談。據觀察統計，多數讀者還是習慣于把那些自己需要反復閱讀的文獻章節打印出來後帶走，極少有讀者，也可以說幾乎沒有人為了獲取同一信息而反復多次在收費網上進出。

在時下一片數字圖書館熱的輿論中，筆者的這些「不合諧」音符並不是要標新立異反潮流，而是想通過現象，認清弊端，找出縮小差距的辦法，使真正的數字圖書館進入千家萬戶，使人們用得好（方便）、用得起，名副其實地達到其理想的目標。以往那種聽風是雨，不顧自己的具體實際情況，盲目地一哄而上爭拔頭籌，各自為政，造成諸多後遺症的教訓，不應在數字圖書館的實施過程中再次發生。

參考文獻

1. 聞生、倪光南：《質疑中國數字圖書館》，載《北京青年報》，2002年4月30日，第6、7版。
2. 孫坦、辛希孟、齊柳明、吳小京：《數字圖書館讓普通人學富五車》，載《光明日報》，2002年8月14日，c1版。
3. 孫承鑒、劉剛：《中國數字圖書館建設的起步與發展》，《國家圖書館學刊》，2000年第三期。

聯絡我們

財團法人
佛教電子佛典基金會（CBETA 基金會）
CBETA Comprehensive Buddhist Electronic Text Archive Foundation

service@cbeta.org
(02) 2383-2182 轉分機 12
官網 https://cbeta.org/
CBETA Online https://cbetaonline.dila.edu.tw/
立即贊助 https://cbeta.org/quick-donate/

贊助資訊

CBETA 長年來陸續將歷代重要佛典數位化為「CBETA 電子佛典集成」，不僅收錄了大量的珍貴文獻，更致力於校勘考證用字及增添新式標點，提升佛典的文獻品質及易讀性。「CBETA 電子佛典集成」是大眾廣為運用的公共資源，此成果也獲得國際學界的重視及肯定。

對於一個從事佛法志業的非營利團隊來說，能夠長期埋首理想、踏實耕耘，實屬不易。為了提供更優質的服務，CBETA 基金會擬增聘專業人力，擴大組織規模，懇請大家慷慨解囊、熱情贊助，讓未來有更多更好的電子佛典。

贊助捐款

「水滴雖微，漸盈大器，凡福充滿，從纖纖積。」
—— CBETA 2024.R2, T04, no. 210, p. 565a5-6