

# 臺大佛學數位圖書館暨博物館現況與未來展望

高巧倫、翁藝芳

## 壹、緣起

台大佛學研究中心設立的宗旨，是希望凝聚國內從事佛學研究的專家和學者，藉由現代學術研究的方法，對佛學的特質、流變、潛在價值、現代意義等問題進行深入的研究。方法是廣泛地收集國內外佛學文獻資料，運用現代科技，建構系統化的佛學文獻資料庫，並透過網際網路，讓有興趣的人們來查詢；而主要的經費來源，都是由善心人士所捐贈。恆清法師在自己的研究室成立了佛學研究中心之後，建立了一個佛學資料庫的網站。從事這樣浩大的工程，需要運用很多的人力和財力，而中心的各項資源全靠十方的資助，恆清法師表示：「佛學資料庫的成立，除了要感謝十方的支持之外，還要特別感謝悟因法師」—由於悟因法師慷慨贊助，佛學資料庫才得以開辦，恆清法師至今仍念念不忘她的協助；另外杜正民老師、館員林慍如小姐在這段期間幫忙最多，是建構佛學資料庫的幕後大功臣。「只要進入佛學研究中心的網站，全部有關佛學的資料，都可以在這裡搜尋到」—這是恆清法師當初建立佛學資料庫的唯一目標。在進行建立資料庫期間，雖然工作很辛苦也很困難，但恆清法師都一一克服，她說，最終的目的是希望使用者上網時，中文、梵文、巴利文的一手資料以及二手資料都垂手可得。

（資料來源：台灣佛教數位博物館－佛教人物《恆清法師》  
<http://www.chibs.edu.tw/academic/projects/taiwan/formosa/people/2-heng-qing.html>）

## 貳、整體功能簡介

- ☞ 書目全文檢索系統：  
目前資料庫中有十二萬四千餘筆書目，以及近四千七百筆全文資料，囊括中英日三種語言，除提供全欄位查詢與進階欄位檢索介面外，另提供關鍵字列表，讓使用者能夠一目了然尋找所需之資料。
- ☞ 佛學語言教學：  
佛教經典原文除了漢文外，亦包含梵文、巴利文與藏文，因此本計畫設有基礎梵文、巴利文、藏文教學，以利使用者運用各種經典語文查閱原文資料。
- ☞ 數位化佛教經典：  
目前已提供學術界通用的《大正新脩大藏經》，以及乾隆大藏經、巴利文藏經等

線上閱覽，相對應的梵文、藏文佛經原典也將陸續上線，未來希望能提供各種不同版本與不同語言的藏經線上閱覽。

- ☞ 佛學工具提供：  
研究佛學資料時，除資料庫外，辭典更是不可或缺的工具，本計畫提供其他佛學相關的資料庫連結，以及各種不同語言的佛學辭典下載。
- ☞ 全球佛學院所網站連結：  
佛學院所乃最專門研究佛學之單位，全世界各地佛學院所各有特色，也囊括不同領域之佛學研究，我們盡所能將所有佛學院所連結整理於此，供使用者參考。
- ☞ 佛學網路資源：  
本計畫收集全球佛學網路資源，並分為「華文佛教網路資源」與「國際佛教網路資源」兩大類，收集之網站幾乎涵蓋全世界的佛學網路資源。

### 參、檢索引擎的演進

本計畫有一段時間，有兩個檢索引擎。一為吳昇教授帶領團隊開發的 GAIS 系統，一為台大資工所歐陽彥正老師率領的團隊所製作，這兩個檢索引擎在今年整合為一。整合之初，緣於我們需要一個效能良好的檢索核心，同時也要解決中英日梵文混合的問題，因此曾經考慮直接引進一般市售的產品。但經過考量以下因素：委外成本高，且導入所帶來的變化需花很多時間與精力投入，技術智識無法累積。因此，最後決定自行開發系統，除了不必負擔昂貴的軟體授權費之外，更能針對本計畫的需求進行細微調整。

檢索核心採用以 Java 為核心的全文檢索引擎 Lucene，目前 12 萬筆書目資料的檢索可以在一秒之內完成；毋需加掛任何字型檔即可顯示日文資料，相較一般市售產品毫不遜色。技術說明如下：

#### 甲、支援中日文字形

檢索引擎為 Java-based，內部處理字元使用的方式是 Unicode，資料則以 Unicode 格式儲存於關連式資料庫中。

#### 乙、檢索速度提升

- i. 利用 Inverted Files 來加快系統搜尋時之速度：  
現有文件處理方法，不外乎：Full-text Scanning、Inversion 和 Signature File。其中 Full-Text Scanning 所耗費之額外儲存空間最少，Inversion 速度最快，Signature 介於兩者之間。檢索核心 Lucene 採用 Inversion 的方式處理文件。所謂的 Inverted Files 即為一個排序過的特徵值集合，每個特徵值都有一個指標連結到包含此特徵值之文件。
- ii. 以 Java-based 網頁技術重寫原本 Perl-based 網頁，提高網站處理各式檢索的效率。

#### 丙、介面

i. 優化過程

有鑒於資料庫最重要的就是檢索介面，以及檢索後所呈現的結果，我們研究了許多海內外知名的資料庫檢索引擎，如台灣的國家圖書館、日本的 INBUDS、美國的 Harvard University Library...等，其中十分感謝佛研中心老師提供許多寶貴的意見。最後我們擷取各家精華，並以佛學研究為主要考量，配合我們現有的欄位，寫成現在的這個檢索介面。

測試期間，透過學者們使用，獲得許多珍貴的意見。隨著數位科技的進步與使用者需求的改變，我們將不會停止改良此檢索引擎，也歡迎各方先進提供寶貴的意見，讓我們有提升的空間。

## 肆、未來努力目標

### 甲、網站管理優化

i. 管理介面網路化與自動化

我們將管理介面建立在網路上，因為沒有地點與時間的限制，如此即可多方大量招募志工，解決目前人力不足的問題。我們已建立之網路介面為：書目上傳機制、書目審核機制、全文上傳機制、全文審核機制，未來將增加期刊更新機制等。設定管理權限，上傳者與審核者可從任何地方作資料上傳與審核。

ii. 審核中心的建立

佛學為一門精確與影響重大的學問，需要專精專門佛學各領域的學者共同把關，審核者必須由審核中心評估認可，授與審核書目全文的權力，審核者亦須對審核中心負責，同樣的有任何問題也可以向審核中心回報，若無法做出明確指示，審核中心可進行線上會議討論。審核中心將有定期會議，必要時可召開臨時會議。

iii. 資料庫規格正規化

參考全世界知名大學及資料庫，訂定資料庫的正規模式，並將過去的大量資料作正規化處理，方便學者使用資料庫軟體如 Endnote 下載管理。另外，將學術與非學術書目分類，方便學者搜尋真正想要的資料，減少不必要的時間浪費。

### 乙、接下來欲進行之專案

i. 成為佛學數位研究專案平台

目前台灣有許多佛學院所，也有很多相關科系或相關單位製作許多數位化專案，但往往因學生畢業或是經費不足，專案結束後就不了了之，甚至成為網路孤兒無人維護，甚為可惜。我們希望能利用臺灣大學豐富的網路資源，成為一佛學數位專案平台，歡迎各專案的加入，專案結束後若無人接手，將維持網站的基本運作，日後若有人對此專案有興趣，可接手繼續發展，在現有的基礎上繼續努力，

可直接傳承而免除從頭開始的人力浪費。

若要達成此目標，需要一套完整的專案製作規範，如何規範可使後人容易接手，但又不致造成專案的束縛，需要各領域專門人員詳細討論，前置作業雖較繁瑣，但將使後來的佛學數位計畫得以長久生存，並提高其應用價值。

## ii. 多語言版本藏經收錄

我們目前已收錄了大正藏、乾隆大藏經、巴利文藏經，未來希望能收錄更多不同藏經版本，並進行語言比對，目前遇到以下問題：

### 1. 來源取得

多數藏經仍未數位化，或是有版權限制。藏經數位化部份希望有更多專案能加入執行；至於版權限制，則需與更多單位洽談，希望能共同為佛學研究而努力，發揮資源共享的精神。

### 2. 語言障礙

目前網路上雖有很多翻譯成不同語言的經文，但由於工作人員對經文不了解，或是對語言不熟悉，即使取得這些資源，由於無法確保這些經文的正確無誤，便無法使用，因此急待專業人才的參與。

### 3. 校對問題

即使能夠使用該語言，但大量的經典需要大量人力支援校對，目前尚無此人力資源，故仍有無法確保經文正確無誤的困難。仍仰賴各大佛學院所或佛學中心，能提出經典數位化計畫，以確保經典數位化保存的正確性與完整性。

## iii. 跨國交流

不同的文化背景會發展出不同的學術成就，我們希望能接洽如 Stanford, Harvard, Berkeley 等具有佛學研究中心的海外院校，進行專案的交流，除充實本計畫的豐富性外，也希望能從這樣的交流中衝擊出一些新的想法，對佛學數位化領域能更有貢獻。