

# 由資料庫到數位研究平台—— 談佛典文獻數位研究工具之發展與演變

From Database to Digital Research Platform: Development and Evolution  
of Buddhism Literary Digital Research Tool

洪振洲 (Hung Jen-jou) \*

編按：去（104）年漢學研究中心與中央研究院數位文化中心合辦「全球視野下的漢學新藍海國際研討會」，於11月19-20日在國家圖書館文教區國際會議廳舉行。會中邀請來自荷蘭、英國、德國、美國、韓國、捷克及臺灣本地學者就1. 當代漢學研究之特色與前瞻；2. 漢學數位資源之發展與趨勢；3. 漢學研究之國際合作與交流等子題，共發表13篇論文。有關此研討會之報導，請見本期「會議報導」專欄。

此次發表論文的內容，具有研究參考價值，《漢學研究通訊》特別選擇適合本刊性質之論文，分二期刊登。限於篇幅，於本期先行刊登洪振洲教授、徐約和（Joerg H. Huesemann）助理教授、金把路先生等3位主講人論文。

## 一、前言

佛教的傳播，在印度本地曾經是師徒口耳相傳教法，書寫貝葉紀錄。當佛教思想自東漢經由中亞流傳到中國之後，透過譯師的傳譯，古代高僧大德們陸續將印度佛教經、律、論三藏典籍翻譯為漢文佛典，並掀起了中國本土長達千年的佛經翻譯熱潮。這場巨型的文化活動，不僅完成了大量佛典的漢譯，也奠定佛教於中土廣為流傳的根基。後來中國僧人再依自身對於佛教的理解，為傳譯的佛典書寫經序、編撰目錄、傳抄流通，並加入中土高僧撰述的論著，包括：疏注、史傳、音義、

法數、語錄、儀制等，逐漸形成具中土特色且博大精深的佛學體系，而成「漢文佛教大藏經」，作為佛法的傳承。大藏經不但是歷代祖師大德、後世景仰者學習佛教義理的主要資訊來源，它的紀錄也反映著中國佛教歷史的發展，載錄了文明變遷的痕跡。

今日研究漢傳佛教者，無不以大藏經為主要研究文獻，但這些佛教文獻本身也產生了許多文獻研究上的謎團待解決。為尋找這些問題的合理解答，使用傳統研究方法的研究者，多半以自身能力搭配先賢的知識引導，以人工方式進行文獻內容的整理與比對，並由詞彙、語法、文體、內容意涵等各方面進行探索與討論。但因研

\* 作者為法鼓文理學院佛教學系專任副教授，並兼任圖書資訊館館長一職。

究過程皆需要大量人力來協助，且人工對於大量資料同時進行整理與比對實有進行上之困難，以致佛典文獻研究進展緩慢，大規模進行資料比對的成果實屬少見。

自 1990 年開始，建立「數位佛學資源」以輔助研究者或使信眾得以一窺佛典精要的想法，便開始陸續在各團體組織間萌芽。歷經多年的努力，在臺灣許多「數位佛學資源」陸續被建立起來，其成果不僅豐碩，也頗獲國際好評，為臺灣保存並累積了相當豐富的文化資財。這些已經完成的數位佛學資源，其主題各異，內容十分豐富，提供研究者許多研究上的方便，讓使用者省下大量的資料查閱時間。而這些數位資源的製作團隊，也在一次次的工具資源製作過程中，得以慢慢摸索出使用者的習慣和對於資料使用上的需求，製作出使用方式更便利、更能貼近研究者需求的研究資源與工具。

現有的各種數位佛學資源雖然都是具有豐富內容的數位資料庫，但惜因各資料庫間並無良好之整合，實難發揮相輔相成之效。且由於這些資料庫之設計，仍以輔助研究者閱讀與查找資料為主，因此雖然加快了資料搜尋的腳步，但卻未解決人力無法進行大量資料比對與分析的問題。為突破此困境，近年來引起各方注意的數位人文研究浪潮，就是嘗試以數位資訊技術的長處，輔助人文學者在資訊科技的幫助之下，進行大規模的資料分析與比對，藉以更進一步探索經典文獻的內涵，擺脫人文研究進展緩慢且多有見樹不見林之缺憾。

隨著全球性的數位人文研究浪潮，我們需要思考的是如何從跨領域的結合中，創造數位專案之間彼此的整合，藉由重新詮釋無形文化之深層涵意，將不同時期所建置的不同型式而彼此相關的「數位內容」，透過「知識管理」能力，建構成更有價值的整合型數位「知識資產」。為此，我們現正著手進行佛典文獻之數位整合研究平台的建置。此平台的主要設計理念，就是以輔助人文學者應用數位方法進行漢籍佛典文獻研究為主旨，並以將現有相關數位資源於平台上進行良好的整合，提供更全面的數位資料內容為目的。在提供良好整合之資料

後，再將數位研究所需之工具，建置於此平台之上，以引導研究者使用數位研究方法於研究問題之中。

為使讀者對於現有之佛典數位典藏專案與未來之可能發展有更全面的認識，本文後續將先從歷史沿革的角度，討論佛典文獻數位研究工具之發展與演變，再針對目前進行中的數位整合研究平台作詳細的介紹與解析。期能以此讓讀者快速掌握佛典數位資源的整體發展現況，並提供系統開發設計人員於進行類似系統開發時可參考之模板。

## 二、佛典文獻數位化工作之緣起

90 年代中後期，資訊技術有了長足之進步，帶動個人電腦的普及；同時間，網際網路的發明，不僅讓人們首次感受到數位資訊傳遞的便利性，也連帶引發佛教界人士開始討論關於將佛教典籍數位化，以便能在網路流通等相關議題。一些因緣較為具足的機構與個體擁護者，已經開始著手進行將某些特定的經典文字輸入於電腦系統的作業。其中一個以組織方式，有系統化的進行佛典數位化工作的要角，就是沈家禎居士。沈居士在西元 1994 年於紐約大莊嚴寺成立「佛教電腦資訊庫功德會」，以便持續進行佛典數位化工作。在該協會 1995 年的綜合報告第一號中就已提到，當時正在進行《大般若波羅蜜多經》、丁福保《佛學大辭典》等資料的輸入，且成果已達數百萬字。<sup>1</sup> 另外，約略同時間，一位臺商蕭鎮國居士，以獨立出資的方式，進行《大正新脩大藏經》（後續簡稱《大正藏》）的內容輸入作業，時至 1997 年初，蕭居士已經完成《大正藏》1-25 冊的輸入作業。<sup>2</sup>

這一階段的數位化工作，多半以基礎的資料輸入作業為主，其目的主要是為了滿足使用者利用電子媒體閱讀佛典資料的需求。因此完成的成果，多半為大量的純文字檔案，之後交由使用者自行運用。以功能面而言，自然是十分陽春，尚無法稱為一個完整的系統或是資料

1 參閱：佛教電腦資訊庫功德會，「綜合報告第一號」，<http://www.baus-eps.org/report/epsr001.htm>。（2015/9/17 上網）

2 蕭鎮國，〈由《大正藏》輸入經驗兼談古籍整理〉，《佛教圖書館館刊》24（2000），頁 92-102。

庫，然而在當時的環境與條件之下，已經算是前無古人的突破了。也因為有了上述初步嘗試的成功，才促使相關人士開始思考如何能夠製作出更大規模、更具組織系統，而且具有更多功能的數位文獻與工具。因此催生出了「中華電子佛典協會」（Chinese Buddhist Electronic Text Association，以下簡稱 CBETA）的成立，以及後續製作出讓整個佛學研究界改以數位資源為主要研究來源的「CBETA 電子佛典集成光碟」。<sup>3</sup>

### （一）CBETA 電子佛典集成

CBETA 組織，乃是在 1998 年 2 月 15 日，由恆清法師號召有志於佛典數位化的人士，在取得「北美印順導師基金會」、「中華佛學研究所」及各界的全力支持之下，於法鼓山安和分院舉行籌備會議後成立。<sup>4</sup> 其原始之設立目的為：

- 收集所有的漢文佛典，以建立電子佛典集成。
- 研發佛典電子化技術，提昇佛典交流與應用。
- 利用電子媒體之特性，以利佛典保存與流通。
- 期望讓任何想要閱藏的人都有機會如願以償。

CBETA 的工作規劃，以五年為一期，時至今日，已經到了屬於第四期計畫的第 17 個年頭。在第一期計畫中（1998.02-2003.02），<sup>5</sup> CBETA 挑選了佛學研究界最常使用且最具權威的《大正藏》作為數位化的目標，並遠赴日本取得日本大藏出版株式會社的授權，確立 CBETA 的合法地位。同時間，與蕭鎮國居士合作，以其出資製作輸入的《大正藏》文字為基礎，進行文字編碼轉換、缺字處理、內容校對與數位標記處理，製作成高品質且可長久流傳的數位文獻。<sup>6</sup> 除了提供高品質的佛典數位文獻之外，CBETA 也在第一時期，就已確立將提供一個單機的閱覽程式（CBReader），<sup>7</sup> 輔助使用者閱讀資料藏

經。在此閱覽程式中，系統除提供經典目錄瀏覽、完整經文顯示，也具有全文搜索與產生符合論文引用格式的引用複製功能。在後續三期的工作之中，CBETA 持續進行更多佛典數位化內容的製作，截至目前為止，已經完成《卍新纂續藏經》、《嘉興大藏經》、「歷代藏經補輯」、「國家圖書館善本佛典」、「漢譯南傳大藏經」元亨寺版、《藏外佛教文獻》、《正史佛教資料類編》、《北朝佛教石刻拓片百品》等資料的數位化，內容總計約有 1 億 9 千萬字。

CBETA 能夠在經歷了 17 個年頭之後，依然毅立不搖，並躋身為佛學研究界目前主要的研究工具之一，其成功因素，筆者認為可以歸納為以下幾點：

#### 1. 時機得宜且資源充足

專案成立之初，佛教研究界莫不引頸期盼大藏經數位文獻工具的出現，而 CBETA 的成立，正為此需求找到曙光。CBETA 組織由各種不同專業的人士所組成，結合了佛學與資訊兩個極端專業的技術，使其產出結果能兼顧學術研究與一般流通需求。由於臺灣漢文環境是以正體字為主，在處理佛典古籍上有其方便性，而且臺灣佛教興盛，義工投入意願高，在專案初期需要大量人力協助之時，能提供 CBETA 許多人力上的支援，讓 CBETA 得以快速產生成果，首先達陣，打開名聲。

#### 2. 資料內容豐富

雖然學界多半以《大正藏》內容作為研究參考之圭臬，但《大正藏》並非是內容最完整的佛教文獻集合。在中國歷代集結的藏經之中，有許多重要內容，並沒有被收錄於《大正藏》之內，而 CBETA 在多年的製作過程中，陸續將這些資料補齊。以內容來說，已經遠超過《大正藏》紙本內容，並且也是其他佛典數位化專案目前尚無法匹敵的。

3 CBETA 電子佛典集成，由中華電子佛典協會（CBETA）所製作，最新版本為 2014 年所發表之內容。詳見：「CBETA 電子佛典集成」，<http://www.cbeta.org/cbreader/help/index.html>。

4 杜正民，〈漢文佛典電子化記事〉，《佛教圖書館館刊》24（2000），頁 60-83。

5 中華電子佛典協會，「CBETA 目標」，<http://www.cbeta.org/intro/mission.php>。（2015/9/17 上網）。

6 釋惠敏，〈大藏經電子化的實作〉，《佛教圖書館館刊》18/19（1999），頁 18-26。

7 中華電子佛典協會，「CBReader 使用說明」，[http://www.cbeta.org/cbreader/help/cbr\\_index.htm](http://www.cbeta.org/cbreader/help/cbr_index.htm)。（2015/9/17 上網）

### 3. 國際交流與時並進

CBETA 不但成員與作業皆十分國際化，多年來並積極參與國際交流，傾聽學者需求，觀察最新科技發展趨勢。因此 CBETA 電子佛典集成在內容與功能上，皆能與時並進，符合使用者需求。

### 4. 後端資料處理技術選擇得宜

專案在啟動之時，便決定除了單純記錄文字內容外，也利用標記語言 (XML)，並搭配數位文獻之國際處理標準 Text Encoding Initiative (TEI)，<sup>8</sup> 進行資料內容的結構記錄。因此，後端原始資料的記錄，不僅豐富，而且十分符合電腦處理需求，即便是使用於今日新系統的開發時，都還能夠妥善的運行。

### 5. 具專責組織，負責長久維護

CBETA 為一個獨立的專責組織，以非營利方式經營。該組織持續維護此工具，並擴充與更新內容，因此在軟體程式變化飛快的今日，仍能有一個可以平順運行的環境。同時間，此機構持續給人生氣蓬勃之觀感，也增加使用者對於使用此工具的意願與信心。

## (二) 其他早期佛典數位文獻

1990 年，印順文教基金會即開始進行印順法師的佛學著作數位化工作。<sup>9</sup> 歷經 8 年的努力，將印順法師佛學著作 42 冊，約七百萬字的內容全部數位化。其光碟所附之程式，也提供全文檢索、萬用字元檢索、結果再檢索等功能。目前該程式仍由印順導師基金會協助維護中，光碟程式內容已更新至 4.0，<sup>10</sup> 並且提供方便平板電腦閱讀的 iPad 線上版，與電子書 EPUB 格式下載。值得一提的是，導師光碟的營運模型與 CBETA 有些類似，也是由印順文教基金會多年來不斷的進行維護與更新，

而其後端資料格式也是使用 XML 為基礎的方式進行標記處理。因此，多年來，歷經多次資料更新與程式功能變革，底層資料格式仍能符合程式處理需求。

佛光山於 1997 年製作完成八大冊的《佛光大辭典》的數位版本，並以光碟的形式發行。《佛光大辭典》內含二萬二千六百五十餘條目，七百萬字的釋文，光碟內並附有一個獨立的運作程式，是全世界第一份中文現代佛學辭典軟體，<sup>11</sup> 而後續在網站上也提供此辭典內容的查詢服務。<sup>12</sup> 直至今日，《佛光大辭典》仍是佛學研究者愛用的研究工具之一。爾後，佛光山繼續啟動《佛光大藏經》的數位化計畫，並於 2002 年陸續完成內容的數位化。<sup>13</sup>

以上兩者僅是早期佛教數位化文獻的成果之二，因此綜觀此時期所完成的數位工具，可發現其建構目的多以提供研究上需要的重要文獻，或辭典工具書的數位內容為主要目標。而專案製作理念也是朝著將紙本資料，轉化數位形式之後，搭配獨立程式提供資料查詢與閱讀之用。在傳播媒體上，則習慣以光碟搭配單機程式作為主要的傳播媒介。

## 三、法鼓文理學院數位化專案

在承續 CBETA 的成功經驗之後，並同國家對於數位典藏計畫的重視，法鼓文理學院圖書資訊館之數位典藏組，成功爭取到來自各方的經費贊助，執行了許多佛典數位化專案。本團隊歷年來製作之佛學數位專案超過 20 個。<sup>14</sup> 在這些專案的製作過程中，除考慮數位化更多不同內容的古籍文獻之外，有些專案也改以建立單一特定主題為核心之跨文獻集合專題型研究資料庫為目的。

8 Text Encoding Initiative, 「TEI」, <http://www.tei-c.org/index.xml>。(2015/9/17 上網)

9 釋厚觀、釋廣淨, 〈《印順法師佛學著作集》光碟版簡介〉, 《佛教圖書館館刊》18/19 (1999), 頁 47-51。

10 《印順法師佛學著作集》光碟, 下載網址為: <http://goo.gl/P7Ww3C>。(2015/9/17 上網)

11 佛光山人文教基金會電腦大藏經, 〈電子佛典萬「碼」奔騰:《佛光大辭典》光碟片介紹〉, 《佛教圖書館館刊》14 (1998), 頁 55-56。

12 《佛光大辭典》電子資料庫, 網址: [https://www.fgs.org.tw/fgs\\_book/fgs\\_drser.aspx](https://www.fgs.org.tw/fgs_book/fgs_drser.aspx)。(2015/9/17 上網)

13 永本, 〈佛光山電子佛典的功能與應用〉, 《普門學報》27 (2005), 頁 351-360。

14 「法鼓文理學院數位典藏專案」, [http://lic.dila.edu.tw/digital\\_archives\\_projects](http://lic.dila.edu.tw/digital_archives_projects)。(2015/9/20 上網)

此外，我們也嘗試引進更多種可能的資訊技術，包含結合地理資訊系統和社會網絡系統來呈現資料內容，以及引入資料視覺化、數位分析等方式，讓原本以文字為主的文獻資料，透過新時代的資訊技術，呈現全新的風貌。雖然各專案間的製作內容與理念並非完全相同，但是我們確實可以觀察到，這些專案的製作理念之改變，皆有一些脈絡可尋。這些脈絡將可提供後續規劃新的文獻數位化專案時，作為重要的參考資料。我們將此流變整理為下列五個主題來說明，此五個主題為：數位文獻資料庫、數位參考工具、主題式文獻資料庫、文獻內容視覺化、佛典文獻之數位分析。

### （一）數位文獻資料庫

此類型的數位專案製作之目的，與前述佛學經典數位化剛起步時之概念相同，也就是以有系統的提供更多數位全文資料為目的。本團隊執行過不少此類型之專案，例如：「台灣佛教史料庫」<sup>15</sup> 專案，就是一個以蒐集臺灣早期（包含明清時期、日據時期與戰後時期）佛教文獻資料為主的數位資料庫。而「《臺灣佛教》期刊數位典藏」<sup>16</sup> 專案，則是將民國 36 年至民國 62 年間發行的《臺灣佛教》期刊進行全文數位化，並提供完整的圖文資料。因文章篇幅關係，筆者在此僅以「中國佛教寺廟志數位典藏」專案為例，來詳細說明此類型資料庫之特色。

中國佛教寺廟志數位典藏（以下簡稱佛寺志數位典藏），<sup>17</sup> 顧名思義，是以「佛寺志」作為主要的數位化對象來源。所謂的「佛寺志」，乃是指以某一特定佛寺或地域為敘述核心的史志類作品。大部分的佛寺志是以單一佛寺為主要的描述對象，但也不乏以一個區域，或一座山頭為範圍所編撰的佛寺志。有別於以朝廷為重心的官方正史，佛寺志的內容，主要記錄佛寺的歷史沿革、

附屬地域空間發展，以及佛寺發展情況等事。而曾於該寺院住錫之歷代祖師、文人雅士之軼事，也通常會成為該寺志的記錄對象。以文體而言，佛寺志的內容有別於一般的論述性文體或依年代條列式的記事體裁，而是採用將不同文類的文本（傳記、詩、散文、地圖、銘文、序跋、圖像等）匯集而成的作品，因此與其說是一個單一的創作，不如說是當時將所有各地蒐集而來的相關著作，統整編輯而成的合集作品。

這些佛寺志文本，在近代被收集並編輯成為兩部內容豐富的叢書。第一部為：1980 年至 1985 年由杜潔祥主編之《中國佛寺史志彙刊》，此部叢書共有 110 冊，前後共分三輯，分別由臺北明文書局（前兩輯）與丹青圖書（第三輯）發行。另一部是：2005 年由張智主編，杭州廣陵書社出版的《中國佛寺志叢刊》，內容共 130 冊。110 冊的《中國佛寺史志彙刊》，內容共包含 100 部寺志，而《中國佛寺志叢刊》的 130 冊內容中，收錄 197 部寺志。兩叢書所收錄的所有寺志當中，共有 60 部寺志是完全相同或僅有小部分差異。因此扣除重複的部分之後，我們可認為內容總共有 237 部具參考價值的寺志。<sup>18</sup> 本團隊自 2007 年開始進行之佛寺志數位典藏，其目的就是要將此兩套叢書所錄之 237 部寺志，進行高品質之全文數位化處理。

在數位資料處理方面來說，為使資料能夠長時間保存，本專案將上述兩套叢書的內容，逐一進行「圖檔掃描」、「全文打字輸入」、「內容標記」等作業。其中「內容標記」是指將每一部寺志，都當作一個標記單元，並將內文中的資訊，如：目錄、標題、表、詩偈、作者、印章、底本小字、註解、人名、地名、時間、原書頁碼行號等訊息，<sup>19</sup> 都利用符合 TEI 標準的 XML 技術，鉅細靡遺地記錄下來。有了這樣的詳細標記，若未來要針對文本進行進一步的自動化處理，將變得較為可行。如圖

15 「台灣佛教史料庫」，專案網址為：<http://buddhisticinformatics.dila.edu.tw/taiwanbuddhism/tb/>。（2015/9/20 上網）

16 「《臺灣佛教》期刊數位典藏」，專案網址為：[http://buddhisticinformatics.dila.edu.tw/taiwan\\_fojiao/](http://buddhisticinformatics.dila.edu.tw/taiwan_fojiao/)。（2015/9/20 上網）

17 「中國佛教寺廟志數位典藏」，專案網址為：<http://buddhisticinformatics.dila.edu.tw/fosizhi/>。（2015/9/20 上網）。

18 Marcus Bingenheimer（馬德偉），〈中國佛寺志初探及書目研究〉，《漢語佛學評論》2（2011），頁 377-408。

19 詳細標記方式可參見：洪振洲，〈中國佛教寺廟志數位典藏系統之建置〉，《法鼓佛學學報》12.1（2013），頁 145-157。

## 1 為佛寺志專案 XML 標記文件片段。



圖 1 佛寺志專案 XML 標記文件片段

由於專案處理之文獻數量龐大，且工作人員眾多，因此，為確保標記品質與一致性，我們針對文獻內容的標記規則，訂定了詳細的規範。而此一標記規則，不僅可以讓工作人員在進行標記工作時有所依循，也適用於新進人員之教育訓練，或新起類似專案在訂定標記規範時可有所參考。我們將專案內使用之文獻標記規則，記錄於法鼓文理學院的知識網，<sup>20</sup> 同時公開給全世界任何有興趣的人來閱讀。

在資料呈現的媒體方面，佛寺志數位典藏系統選擇以網頁作為主要的呈現媒體。系統所提供之閱讀介面如圖 2 所示，介面包含文件的樹狀結構導覽資訊、文章的主要內容，以及對應目前閱讀內容的圖檔。

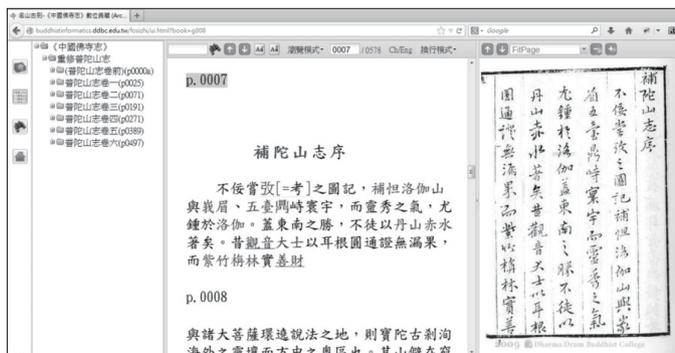


圖 2 佛寺志數位典藏專案線上閱讀介面

資料閱覽介面是以標準的 HTML 語法，搭配 Javascript 函式庫所製作，因此介面相容於大部分的主流

網頁瀏覽器，使用者毋須預先安裝任何外掛程式，即可使用此一多功能線上閱讀介面來閱覽全文。介面上也提供圖文對照模式，使用者可選擇只閱讀原文或圖檔，或是兩者一起瀏覽。在開發過程中，我們很快速地做出線上介面的雛型，並同時完成可將後端資料庫中的 TEI 標記文件即時轉換並呈現在網頁上的機制。因此資料標記人員在驗證的過程中，就可以立即使用線上介面檢視其最終成果，以便過濾許多可能的細節錯誤，或對於介面的使用上，立即提供調整的建議，使文獻內容與程式操作方便性，都能持續改進，算是一個十分成功的開發模型。

此外，佛寺志專案的另一個大突破，就是將專案所產生的符合 TEI 標準的 XML 文件資料，再度轉換成 PDF，並付印成實體版。相較於新式的數位媒體，目前傳統紙本圖書實際上還是具有長期保存和無障礙閱讀等優勢。目前專案已經完成一套十二部寺志的《中華佛寺志叢書》<sup>21</sup> 的出版作業。在這舊酒新瓶的紙本出版品中，我們附加了極具參考價值的資訊，諸如人名及地名索引之編製、中西曆時間轉換之呈現，還有文本之修補、新式標點之添加等。對於為數可觀的異體字，也適當的提供了通用字。這樣的成功嘗試，讓我們在實體書籍與數位資料的轉換上，開啟了一條雙向的應用路徑。

## (二) 數位參考工具

在進行文獻研究的過程中，除閱讀研究文獻外，不免也會需要查閱參考書籍尋找相關的內容。因此在數位文獻資料庫逐步被建立的過程中，許多專案開始將重點擺到相關的參考書籍上。在此方面，本團隊也進行過許多相關專案，例如：「佛學術語字辭典」數位化專案<sup>22</sup> 就完成了 12 部，內容含約 11 萬筆的佛學辭典資料之數位化。以下，筆者挑選兩個此類型的代表性專案「佛教藏經目錄數位資料庫」與「佛學規範資料庫」進行介紹。

20 詳情參閱：「名山古剎——《中國佛寺史志》數位典藏 工作手冊」，<http://wiki.dila.edu.tw/pages/名山古剎——《中國佛寺史志》數位典藏。>（2015/9/20 上網）

21 馬德偉總編輯，《中華佛寺志叢書》（臺北：新文豐出版公司，2013）。

22 「佛學術語字辭典專案」，專案網址為：<http://buddhistinformatics.dila.edu.tw/glossaries/>。（2015/9/19 上網）

### 1. 佛教藏經目錄數位資料庫

在研究佛教經典內容之前，查閱經典的背景資料是必備的功夫。大部分的藏經背景訊息可以透過經典前後序跋，或於大藏經目錄的記載而得知。但實際上，自宋代起，佛教文獻先後被集結為一部又一部的大藏經。由於各藏經的編輯時代、編輯理念與收錄重點皆不盡相同，因此歷代藏經所收錄的內容多有出入，且在經典順序的安排上也有所差異。因此，閱讀各藏經目錄，以理解經典的背景資料是研究佛教經典的基本工作之一。但因歷代藏經為數不少，加上同一經典在不同藏經內可能記錄為不同標題，或甚至有連翻譯者之記錄也不同的現象，因此使得目錄的比對與查閱成為了一個耗時的工作。「佛教藏經目錄數位資料庫」的設計目的，便是為了提供一個將歷代經錄整合比對之後的友善查找介面，<sup>23</sup> 如圖 3 所示。

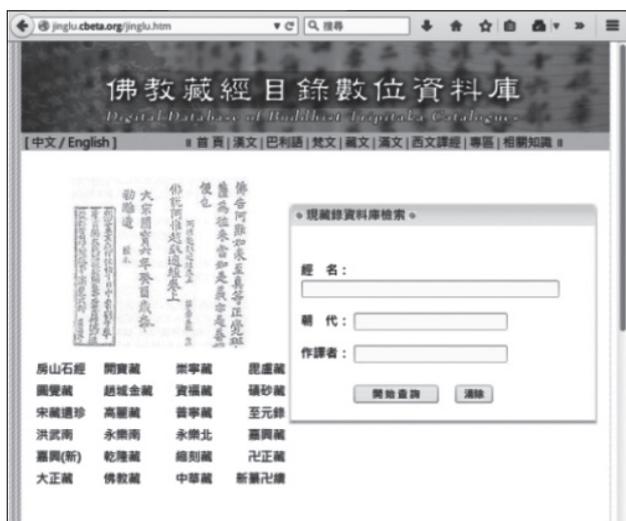


圖 3 佛教藏經目錄數位資料庫整合查詢介面

此資料庫目前針對漢文部分，將歷代藏經的資料經過比對與整合之後，提供一個整合的查詢介面。系統中所完成之資料建檔與連結的藏經目錄有：《房山石經》、《高麗藏》、《大正藏》、《嘉興藏》、《碓砂藏》、《乾

隆藏》等 22 部歷代藏經目錄資料庫的建置。此系統主要提供使用者以單一經典為出發點，查閱該經典的基本資料，與該經在歷代漢文大藏經內的收錄狀況，但也可查閱某一特定藏經的完整目錄內容。

在數位資料處理方面，由於此類資料並非照錄文獻內容，而是以提供整理後的結果為主，因此專案採用以表格形式為主體的關聯式資料模型來儲存經錄資料。資料被規劃成數個資料表格，<sup>24</sup> 而前端閱讀介面則是以 PHP 撰寫而成的網頁，是標準的網頁資料庫應用程式架構。

### 2. 佛學規範資料庫

佛學規範資料庫<sup>25</sup> 建立之目的，就是要提供佛學文獻中所提及的人名、地名與時間等名稱實體的基本背景資料。這些資料將不僅可以讓讀者作為了解某一實體的詳細資料之用，也可整合於各專案的程式介面之中，增加數位專案的內容豐富度，同時提高文獻內容的可讀性。

有別於其他將已成書的資料進行數位化的處理方式，在此資料庫中，除了時間資料部分是利用中、日、韓各國各朝代的曆書與 NASA 所提供之天文資料所彙整之外，人名與地名實體之參考資料，則是來自於其他專案製作的副產品。也就是說，當其他各專案在處理文獻的過程中，一旦遇到需要查找人名與地名之實體相關的背景資料時，各專案人員會把所查到的資料，記錄於規範資料庫中。一來，可以將這些資料作有系統的整理，讓使用者查詢閱讀；二來，可避免不同專案重覆查找相同資料的不必要工作。

我們將規範資料庫分成三大子資料庫，分別是：人名規範資料庫、地名規範資料庫與時間規範資料庫。在資料庫中的所有資料，皆被賦予一個「唯一碼」(Authority ID)，作為該識別對象的唯一識別記錄。人名資料庫的內容記載個別人物之別名、生卒年、朝代、籍貫、說明等欄位。而地名資料庫則為包括地名之別

23 「佛教藏經目錄數位資料庫」，<http://jinglu.cbeta.org/>。(2015/9/19 上網)

24 有關經錄資料庫之資料表格規劃的細節，可參考：「經錄資料庫程式與資料庫架構」，<http://jinglu.cbeta.org/tech/database.htm>。(2015/9/19 上網)

25 「佛學規範資料庫」，專案網址為：<http://authority.dila.edu.tw>。(2015/9/19 上網)

名、群組、經緯度、現在行政區、朝代、註解等資料。地名資料庫也提供地名的十進位地理經緯度座標資訊 (Longitude, Latitude)，可用於 GoogleMap 或 GoogleEarth 等圖形化的地理資訊系統軟體上面，用以明確辨識出該地名的地理位置。透過座標資料的記錄，相關的佛學數位專案便可依此產生地理視覺化的效果，不再侷限於文字上的描述。而時間資料庫則以一日為一個基本單位，每一天被賦予一唯一的規範碼。系統中記錄每一個日期於中國、日本、韓國與西曆曆法的完整轉換對應與解釋。圖 4 為人名規範資料庫查找的範例。

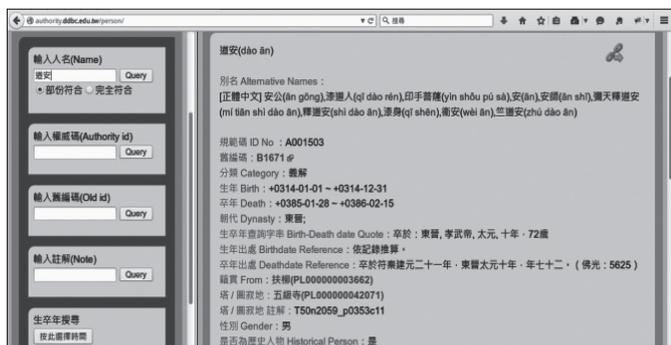


圖 4 人名規範資料庫查找的範例

如同上述經錄資料庫專案，此一規範資料庫專案之資料也是以關聯式資料模型來儲存，並搭配 PHP 撰寫而成的網頁介面提供查詢。此專案也設計了一組 WEB API，<sup>26</sup> 讓其他資料庫應用在所需的網頁介面中，輕鬆整合此一查詢服務。

### (三) 主題式文獻資料庫

在文獻數位化專案執行初期，專案多以數位化單一文獻集合為目標。但實際進行研究時，卻需要使用來自不同文獻集合的資料段落，相互參照比對，以推敲出問

題的答案。因此蒐集符合某主題的相關研究資料，利用文獻彼此間關係與實際研究使用的方式，將文獻作適當排列後所製成的主題式文獻資料庫，反而更能貼近研究者的使用需求。

製作主題式文獻資料庫，主要是要將散落各地而未整合的資料，整合於相同系統之中。其中有些資料可能已經被數位化，收錄在某個大型的數位文獻資料庫中；當然，有些部分也可能僅有紙本資料而仍待進行數位化程序。一旦把相關資料皆蒐集至同一系統，並加以適當的整合之後，將可以發揮原有資料數倍以上之功用，這就是主題式文獻資料庫的魅力。本團隊歷年來已完成多個主題式資料庫，例如「法華經數位資料庫」<sup>27</sup> 是以漢譯《法華經》為主軸，佐以其他各種版本的《法華經》，進行文獻內容之比對。內容包括梵文寫本、漢譯本、藏譯本及英譯本等，建構而成此套完整文獻比對的資料庫。而「《別譯雜阿含經》(T.100)的版本比對資料庫」<sup>28</sup> 是將《別譯雜阿含經》的 364 篇短經文與其對應的漢譯、巴利語平行經文、梵語殘篇經文與英譯內容並列。在下面段落中，筆者以「瑜伽師地論資料庫」<sup>29</sup> 做為此類型的代表性專案來詳細說明。

「瑜伽師地論資料庫」為法鼓文理學院利用行政院科技部之研究計畫補助，再加上後續其他經費挹注所製作之資料庫。其內容是以《瑜伽師地論》<sup>30</sup> 為核心，加上其綱要書、異譯本 (三經二論)、注釋書、梵文原典、藏譯本的電子資料庫。此系統以如圖 5 左方所示之選單，將上述之文獻電子全文，表列於一選單中，供使用者選擇閱讀。此外，為了加強資料之間的緊密程度，此專案以唐朝僧人遁倫所集撰的《瑜伽論記》中的科判資料，做成貫穿此資料庫所有文獻主要綱要。所謂的科判，乃是古德進行佛經內容解析時的一種方式。主要就是將經

26 「佛學規範資料庫資料服務」，網站：<http://authority.dila.edu.tw/docs/services/>。(2015/9/19 上網)

27 「法華經數位資料庫」，專案網址為：<http://sdp.chibs.edu.tw/>。(2015/9/20 上網)

28 「《別譯雜阿含經》(T.100)的版本比對資料庫」，專案網址為：<http://buddhisticinformatics.dila.edu.tw/BZA/>。(2015/9/20 上網)

29 「瑜伽師地論資料庫」，專案網址為：<http://ybh.chibs.edu.tw/>。(2015/9/20 上網)

30 《瑜伽師地論》舊名《十七地論》，簡稱《瑜伽論》，是佛教唯識宗根本大論，亦是玄奘菩薩西行取經求法之最大原因。約於公元前 300 年左右出現在古印度。漢傳譯本中最完善完整的是唐朝玄奘法師於貞觀二十一年 (647) 至二十二年 (648) 間，在東都弘福寺翻譯者，共一百卷。在此之前，另有部分譯本。

文拆解成多層次的小段落，之後再加以逐段解釋，並以一短語來代表該段落的內容。若將這些短語與其間的層次一併取出後，就形成文章的結構綱要了。利用此綱要，專案將各文獻拆解成小的語意段落，並加以辨識文獻彼此的對應段落，並將這些段落間的關係，記錄於數位化完成的文獻之中。因此，系統中的文獻，不再是單純的並排關係，而是形成一個文件間具有參考與比較的網狀結構。

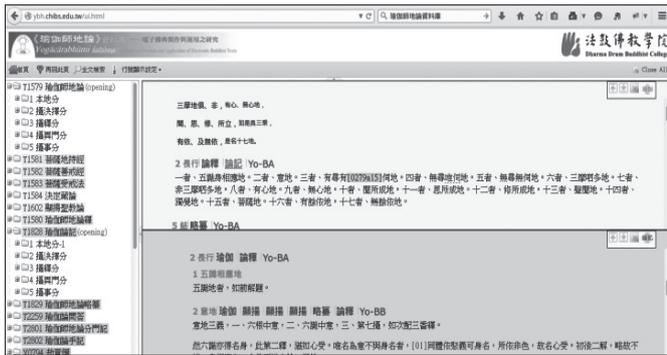


圖 5 瑜伽師地論資料庫介面示意

在閱讀介面上，為使資料間網狀關係能正確呈現，專案人員花費了不少心思在構思介面的前端呈現方式。如圖 5 所示，在全文閱讀介面中，在每個內文段落之後，皆會列出被識別為與此段落有關係的其他文獻參考段落連結。當使用者按下這些對應參考段落之連結後，系統將會開啟新的小分割視窗，並顯現其參考內容，以供讀者進行對讀之用。透過此系統，研究者將不須一直花費時間在多個文獻間尋找對應的參考段落，而可以快速取得完整的相關內容。如圖 5 所示，畫面右方切成上下兩個視窗，上方視窗為《瑜伽師地論》之主要內容，下方視窗是當使用者按下「2 長行」右方的「論記」（圖上標記為紅色）連結後，自動開啟的《瑜伽論記》中相對應的段落，以方便研究者進行內容對讀。

#### (四) 文獻內容視覺化

在大部分的數位文獻資料庫中，系統仍然選擇以類

似原書模樣的方式來呈現資料內容，但當文獻內容被數位化後，文字資料實際上已經被轉化為可運算的資料型態。若將其搭配上適當的資訊技術，這些傳統文獻資料，也可以有全新的面貌。在「佛教傳記文學地理資訊系統」和「佛教傳記文學社會網絡視覺化系統」兩個數位專案中，我們將佛教文獻資料與近年來快速發展的「地理資訊系統」及「社會網絡視覺化技術」相結合，提供讀者一個全新的資料閱讀體驗。

#### 1. 佛教傳記文學地理資訊系統

「佛教傳記文學地理資訊系統」數位專案，<sup>31</sup>主要是將《大正藏》所收錄的多部高僧傳經典文獻，<sup>32</sup>進行全文數位化與標記作業，內容合計 1,382 篇傳記。這些高僧傳記的內容，皆提到許多人物、地點、時間，以及由這三者交錯而成的關係訊息。為使讀者能更容易明白傳記內容，專案不僅數位化文字資料，也識別文本中的人、時、地實體，並增加基本參考資訊。由於文本中的地點名稱都已經被辨別出來，地點的座標位置也已經過查找，所以我們得以結合地理資訊系統，讓使用者能夠透過地理分布的角度，一窺文本中所蘊含的時空資訊，給讀者一個全新的視野來瞭解高僧傳的內容。圖 6 展示的就是將釋玄奘傳的傳記中所提及的地點，在地理資訊系統上呈現的結果。我們可以看見，玄奘傳內容所提及的範圍，確實涵蓋至中亞與印度。



圖 6 佛教傳記文學地理資訊系統介面示意——以釋玄奘傳為例

31 「佛教傳記文學地理資訊系統」，專案網址為：<http://buddhisticinformatics.dila.edu.tw/biographies/gis/interface/>。（2015/9/20 上網）

32 內容包含有：《補續高僧傳》、《出三藏記集》、《比丘尼傳》、《梁高僧傳》、《唐高僧傳》、《宋高僧傳》、《名僧傳抄》與《明高僧傳》。

在數位資料處理方面，由於專案所需使用的各僧傳內容，實際上已經收錄於 CBETA 電子佛典集成之中，因此本專案使用 CBETA 已經完成的 XML 標記檔為底本，再於其上加入所辨識出的人、時、地實體之識別標記，<sup>33</sup> 大大節省了基本資料輸入的工作時間。

在程式介面設計方面，主要核心功能為如圖 6 所示之介面，讓使用者選擇單一文獻進行閱讀，並將文獻中所辨識出的事件點，即時展現於地理資訊系統之中。此專案除了成功結合地理資訊是其一大亮點之外，系統也提供如圖 7 般的語意搜尋功能，讓使用者得以使用人名、地名、時間為條件，進行交叉檢索。與透過全文檢索的傳統搜尋模式比起來，這是更直觀，也是更貼近研究者的使用方式。透過此系統，將可以回答傳統人工處理模型難以回答之問題，例如：「梁武帝從稱帝到開始篤信佛法之間和誰去過哪裡？」與「玄奘和誰何時在長安」之類的語意問題。

圖 7 佛教傳記文學地理資訊系統語意搜尋介面示意

## 2. 佛教傳記文學社會網絡系統

上述專案將高僧傳記資料用於地理資訊的視覺化呈

現，這僅是資料的一種可能呈現。實際上，運用同一份資料，經過不同方式的處理後，也可呈現完全不同的視覺化效果。在佛教傳記文學社會網絡系統，我們特別關注在高僧傳記資料中，人與人之間的關係。因此在處理上，我們略去地點的因素，統計不同人物存在於同樣事件內的次數。然後，我們便可以利用社會網絡系統的視覺化方式來呈現人物之間的關係，以及他們互動的緊密程度。圖 8 為釋玄奘傳記中的人物關係以社會網絡系統的視覺化的樣貌，其中人物以方框表示，而人物間的關係就以線條表示。某一人物出現於越多關係點中，其方框範圍將會較大，以彰顯其重要性。當兩位人物同時出現於關係點的次數越多，兩者中間的連線將會加粗，表示期間有密切往來的現象。

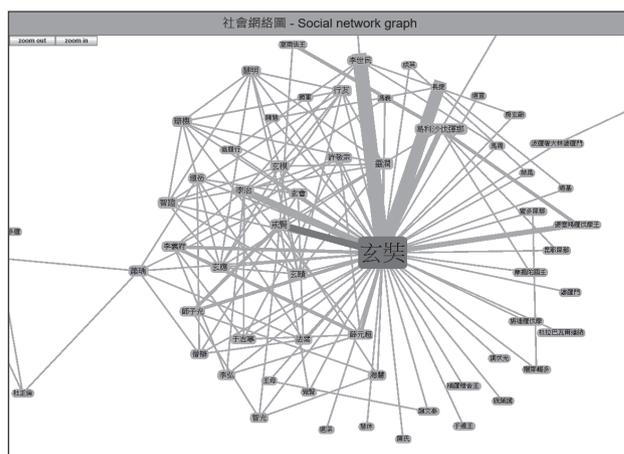


圖 8 佛教傳記文學社會網絡系統示意——釋玄奘傳人物關係

當然，電腦的強項是大量且快速處理資料的能力。因此進行跨經典的資料統計，找出在大範圍文本中，人物之間交往的狀況（例如：整部《唐高僧傳》中所有人物間的關係視覺化），才是此系統的魅力所在。透過資料的建置與資料視覺化技術之應用，將使大規模文獻內容的探勘工作，變得不再遙不可及。

## （五）佛典文獻內容之數位分析

數位資料除了提供讀者便利的閱讀功能之外，更大

33 詳細標記規則，可參閱：洪振洲、馬德偉、許智偉，〈漢文佛典的語意標記與應用：《高僧傳》文獻的時空資訊視覺化和語意搜尋〉《圖書與資訊學刊》2.3（2010），頁 1-24。

的好處是可以搭配電腦與數理統計的演算，以找出文獻研究問題的新研究契機。這類工作，實際上就是利用已經數位化完成的文獻內容，更進一步地進行文字的分析與統計來達到目的。

在「漢譯佛典 Ngram 量化統計資訊平台」<sup>34</sup>的專案中，我們將 CBETA 電子佛典所收錄的文獻內容，與經錄資料相結合，打造出一個簡單易用的文字分析工具，以便協助研究佛典語言的學者，透過統計方法來分析各語言單位在文章中出現的頻率和分布，以進行後續的分析與推論；並以巨觀的角度，找出傳統語言學研究方式不易發覺之語言特徵。在此工具中，研究者可以在輸入關鍵詞彙後，就快速查得這些詞彙在不同年代的發展趨勢。如圖 9，就是使用具有類似取代關係的關鍵字——「泥洹」和「涅槃」這兩個常用於佛經翻譯的替代詞彙——所執行出來的結果。

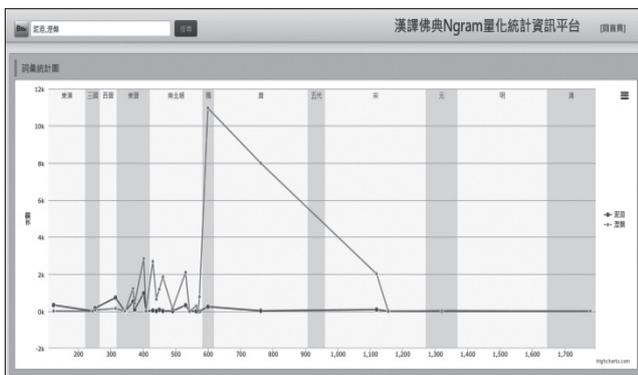


圖 9 「泥洹」、「涅槃」於各年代佛經內部出現情形

當使用者的游標移至畫面上的各點時，都可以即時看到所有關鍵字於相對年代的統計數字。此外，若使用者繼續點擊統計圖上的點時，將可以閱覽到（如圖 10 所示）關鍵字於各經出現的次數與出現的位置等細節，供使用者進行細部了解。

實際上本專案並未增加額外的資料處理功夫，只是將原有的 CBETA 資料搜尋功能加上經典目錄資料與資料視覺化的技術，便可以達到目的，因此在製作上不需

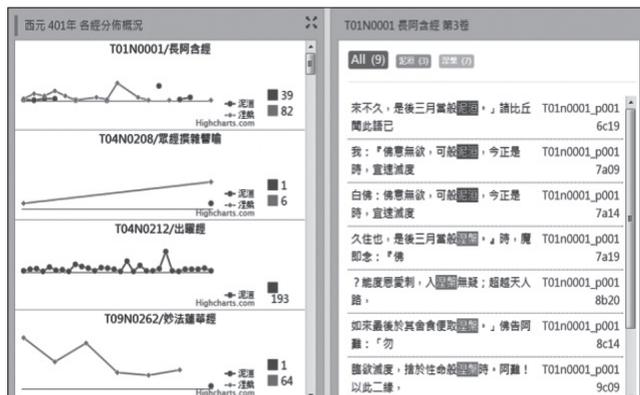


圖 10 「如是我聞」、「聞如是」於各年代佛經內部出現情形

要耗費大量開發時間。這讓我們體認到，以目前現有的數位資料，搭配一些資訊技術與巧思，我們便可以創造出許多前所未見的工具與新知識。

#### 四、佛典文獻之數位整合研究平台

佛典數位文獻工具，在經過十數年的發展期之後，已經有了許多內容豐富的「數位佛學資源」。只不過各資料庫雖然都具備豐富的資料與完善的搜尋功能，但因彼此之間缺乏整合，故而難以發揮相輔相成之效。再者，上述提及之數位量化分析趨勢，也是目前亟待發展的領域，需要更多更完整的工具來搭配。因此，本團隊於 2013 年開始，在國科會（科技部）專案贊助之下，<sup>35</sup> 開始進行建立適合「人文研究的數位研究環境」的開發計畫。其終極的目標，是將數位內容、研究需求與資訊技術，整合成一個能輔助進行以佛典為主之漢籍文獻的學習與研究之數位研究平台。

##### （一）數位研究平台之架構概述

平台的主要設計理念，就是想要顛覆以往獨尊資料閱讀體驗的設計方式，改以建立一個能讓使用者在此研究平台上，同時進行資料閱讀、文獻內容改善與進行量

34 「漢譯佛典 Ngram 量化統計資訊平台」，專案網址為：<http://dev.dila.edu.tw/BuddhaNgramViewer>，網頁取得日期。（2015/9/21 上網）

35 科技部 102 年數位人文研究專題計畫，「漢籍文獻數位研究環境：以佛典研究平台的建置為例——法相妙用：漢籍佛典數位研究學習平台之建置」，計畫編號：102-2420-H-655 -004 -MY2。

化分析的環境。我們將上述需求與使用者行為做統整考量之後，規劃出系統架構如圖 11。

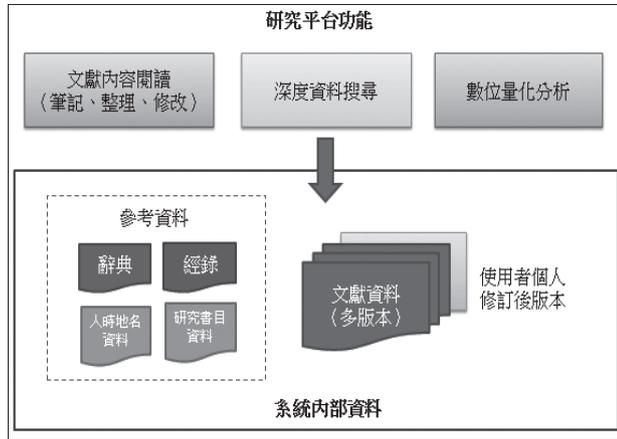


圖 11 數位研究平台之功能與資料架構概述

以功能面來說，系統將會提供「文獻內容閱讀」、「深度資料搜尋」、「數位量化分析」等三個大功能面。「文獻內容閱讀」功能，如字面所述，系統將提供使用者高品質、高正確性，且功能豐富的資料閱讀體驗。並加上相關的參考資料，讓使用者在閱讀資料時，可以方便取得其他相關內容。此外，讀者在閱讀資料時，難免會開始想要有「增加補充資料」、「標示重點內容」、「加入個人書籤」，甚至「進行底本資料的修正」等資料整理的需求。因此，此系統有別於以往的單一版本資料設計，進而考量資料將會有多個不同的版本並存於資料庫中之需求，如此有望成功解決古籍資料的版本問題，以及兼容使用者想要創造自我新版本的願望。

而「深度資料搜尋」是有別於一般的全文資料搜尋設計，此功能是以提供搜尋結果之相互比對，以找出特殊樣型的設計。傳統「資料搜尋」是為找尋某一特定內容在文獻中的出現位置，因此在介面的設計上，多半以單一條件、單一對象為搜尋的情境。但實際上，若介面能提供使用者進行「同一對象的不同條件搜尋結果」之比對，或「同一條件在不同對象的不同使用情形」，都可能引發研究者的研究契機。因此，我們除了在「文獻內容閱讀」介面提供基本的資料搜尋功能之外，也想打造一個獨立的「深度資料搜尋」介面，讓使用者不只能得到基本搜尋結果，更能以相互比對的方式，找出研究

的契機。

「量化數位分析」則是以量化統計分析為主要角度，預計提供文字分析常用的「文字雲」、「字詞統計」、「共用詞頻分析」和「前後綴詞分析」，讓使用者能簡易操作量化文字分析工具，找出隱藏於文字後的樣型，以全新角度回答研究者之文獻問題。

建立此一平台絕非易事，經過了兩年的努力，我們目前僅完成「文獻內容閱讀」功能中屬於資料閱讀的部份。儘管如此，系統中卻已經預載有完整的 CBETA 數位佛典集成之文獻內容，並提供整合良好之「佛教藏經目錄數位資料庫」、「佛學術語字辭典專案」、「佛學規範資料庫」、「漢譯佛典 Ngram 量化統計資訊平台」與「台大佛學數位圖書館暨博物館」的查詢功能。對於過往其他藏經系統來說，此閱讀介面已經解決了許多過往的問題。於本系統而言，此閱讀介面是系統的核心部分，在此部分完成了良好的基礎之後，後續功能將可順利向上繼續構築。

## (二)「文獻內容閱讀」介面介紹

### 1. 閱讀文獻內容

在經典閱讀介面之中，使用者可於圖 12 中標示 (1) 的區域內看到目前所在的經典標題與卷號。最左邊的大區塊為該經典的內容目錄，可以詳細瀏覽該「分」、「會」、「小經」或其他細分結構。讀者也可以藉由圖 12 中所標示的區域 (2) 的「卷」分頁，將內容目錄區塊改變成為該經典的卷次目錄，以便瀏覽該經的不同卷次。畫面中間為主要閱讀區，內容為所選的佛典文獻之文字。經典的閱讀內容中，保留了校勘資料（以中括號標示）。當按下文獻中以中括號標示的號碼後，會彈跳出含有校勘資料的文字框（如區域 (3)）。介面右上角（區域 (4)）的工具列將提供使用者選擇是否顯示文章行號、是否顯示校勘資訊，以及將字型放大、縮小等功能。

當使用者進行資料搜尋後，搜尋結果將列於畫面右邊的區塊（區域 (5)）之中。系統不只進行全文搜尋，同時也整合各個相關數位資源工具的查詢。目前系統將同時進行「全文檢索」、「查字典」、「Ngram Viewer」、「規範資料庫」和「台大佛學數位圖書館暨

博物館網頁」等五種查詢。在全文檢索的結果中，將會列出目標查詢詞句出現的文章與位置。在「查字典」頁籤內，會列出「佛學術語字辭典」的字典查詢服務所查到的相關詞條解釋。

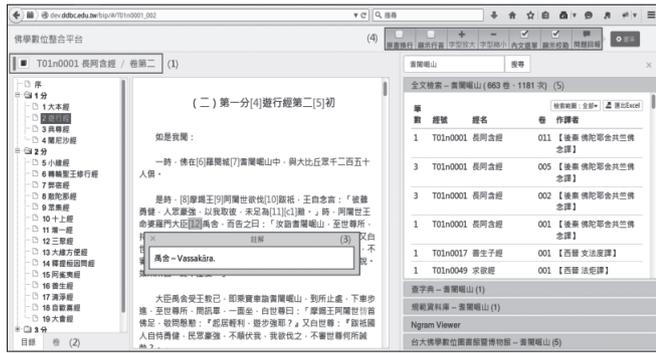
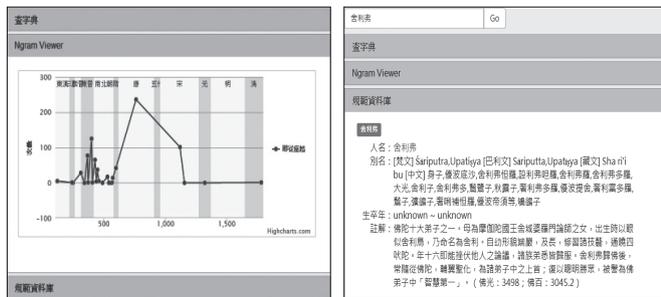


圖 12 經典閱讀介面功能展示

此外，若選擇「Ngram Viewer」頁籤，系統將會向「漢譯佛典 Ngram 量化統計資訊平台」的字詞統計查詢服務進行要求，並即時將統計資料傳回與進行統計圖製作，如圖 13(a) 所示。如此能讓使用者在閱讀過程中，馬上可以知道該查詢字詞在不同年代的經典文字中的使用次數。此外，系統也將同時查詢在「佛學規範資料庫專案」的結果，如圖 13(b) 所示，讓使用者可以立即為文獻中的人名與地名，找出相對的參考資料。



(a) 利用查詢漢譯佛典 Ngram 量化統計資訊平台查找字詞於不同年代的出現次數趨勢圖 (b) 利用規範資料庫查詢功能，尋找人名與地名的參考資料

圖 13 整合平台參考資料查找功能展示

2. 切換文字排列模式

在文獻內容的呈現上，系統提供了四種不同的調整

方式，而這四種方式主要是以兩個因素的調整來決定。這兩個因素就是：「原書換行」與「顯示行首」的兩個選項（如圖 12 區域（4）中的按鈕所示）。「原書換行」是指系統的文字排列，將完全模擬紙本文獻資料的樣貌。也就是，文字換行與換頁處的所在將與原書完全相同。但在數位媒體的顯示中，因為畫面的寬度與書本不同，強迫依照原書排列方式進行文字的排列並非是一個絕對完善的選擇。不過，依原書換行的文字自然有利於與原書對應，因此系統提供此按鈕讓使用者選擇是否文字要依原書的排列方式來換行。而另一個選擇就是「顯示行首」的功能，因為在 CBETA 電子佛典集成中，經典的每一行都有一個固定的行號，透過「顯示行首」的功能，可以讓系統開關這些行號資訊的顯示功能。因此兩相影響之下，便產生如圖 14 的四種不同的顯示模式，使用者可以自行選擇較好的方式進行閱讀。

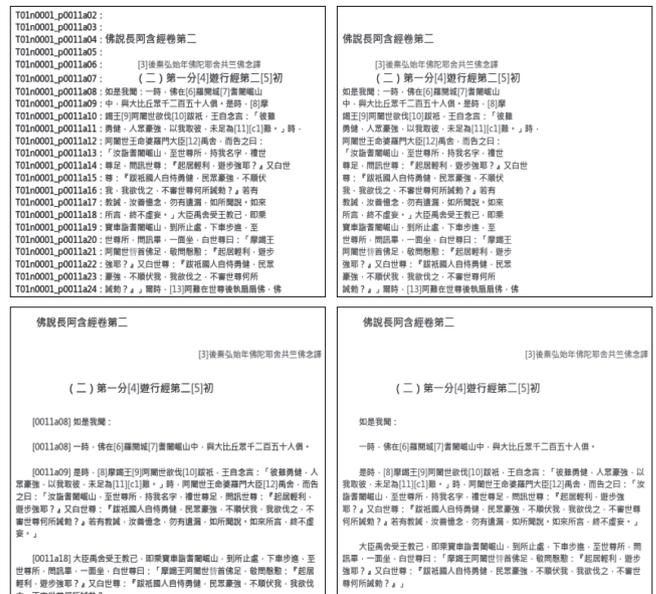


圖 14 四種不同的閱讀模式讓使用者自行切換

3. 經典內容選擇

在閱讀系統中，也提供良好的經典選擇方式。當讀者按下圖 12 區域（1）的書本圖示按鈕時，將會進入經典選擇畫面。我們在系統上提供了四種不同的經典選擇方式，包括有「依據部類」、「依據冊別」、「經目查詢」與「Goto 查詢」等四種方式。「依據部類」的選擇方式，就是利用 CBETA 所提供的整合經典目錄，引導

讀者找到所需經典資料的方式。在本界面中，我們採用類似檔案總管的方式來製作經典選擇器。右方主要選擇區域（如圖 15 中區域 3 所示）將提供選項之下的可能選擇項目，讓使用者進行選擇。而使用者選擇的項目，將記錄於上方的歷程區之中（如圖 15 中區域 2 所示）。



圖 15 經典選擇功能



圖 16 以「金剛」為關鍵字進行經目搜尋的結果

「經目查詢」是另一個系統的便利功能，當讀者僅記得經典編號，或經典題名或細節結構的部分名稱時，可以利用這個功能快速找到所需查閱的經典段落。圖 16 為以「金剛」為關鍵字進行經目搜尋的結果，系統除可以列出經名內有「金剛」的部分，若「金剛」二字出現在經典的小結構內，也會一併列出，並以紅色標記，方便讀者明瞭。

## 五、結語與平台發展

本文主要以「佛典文獻數位研究工具之發展」為題，簡述近二十年來，佛典數位文獻應用工具的發展脈絡。我們將這些發展工具的研發理念，依發展類型之先後順序，分為五大主題：數位文獻資料庫、數位參考工具、主題式文獻資料庫、文獻內容視覺化與佛典文獻之數位分析。文中除逐一介紹佛典文獻數位研究工具之各發展類型特性之外，也針對各類型相關專案舉例說明其數位資料模型設計重點與介面技術。

此外，考量到現有的佛典數位文獻工具，在經過十數年的發展期之後，已經有了許多內容豐富的「數位佛學資源」，但因為彼此間缺乏整合，難以發揮相輔相成之效，再加上以數位量化方式進行文獻內容分析也是一大趨勢，亟需更多更完整的工具來搭配，因此，本團隊進行「佛典文獻之數位整合研究平台」的開發計畫。其終極的目標，是將數位內容、研究需求與資訊技術，整合成一個能輔助進行以佛典為主之漢籍文獻的學習與研究之數位研究平台。

此整合平台在實作上，以「文獻內容閱讀」、「深度資料搜尋」、「數位量化分析」等三個大功能面為主要的功能設計。目前「文獻內容閱讀」之功能已接近完成，系統中將有完整的 CBETA 數位佛典集成之文獻內容，並提供整合良好之「佛教藏經目錄數位資料庫」、「佛學術語字辭典專案」、「佛學規範資料庫」、「Buddha Ngram Viewer」與「台大佛學數位圖書館暨博物館」的查詢功能，期望能提供完整的數位研究平台，同時也提供研究者所需之資料。

在後續的計畫執行過程中，將首先討論多版本資料在系統中的呈現與搜尋功能，並實做多版本的資料並存模式。然後開始考慮如何讓使用者於系統中建立自己的版本，並於前端介面上提供「資料修改」、「加註筆記」與「書籤管理」等功能。而同時也將以目前之資料平台為基礎，向上繼續建構「深度資料搜尋」、「數位量化分析」兩大功能面，以真實引導研究者使用數位研究方法於研究問題之中為最大目標。