

## 杜正民居士訪談紀錄（二）\*

卓遵宏、侯坤宏訪問，李美寬記錄\*\*



《大正藏》（杜正民提供）

### 陸、CBETA的承先啟後工程

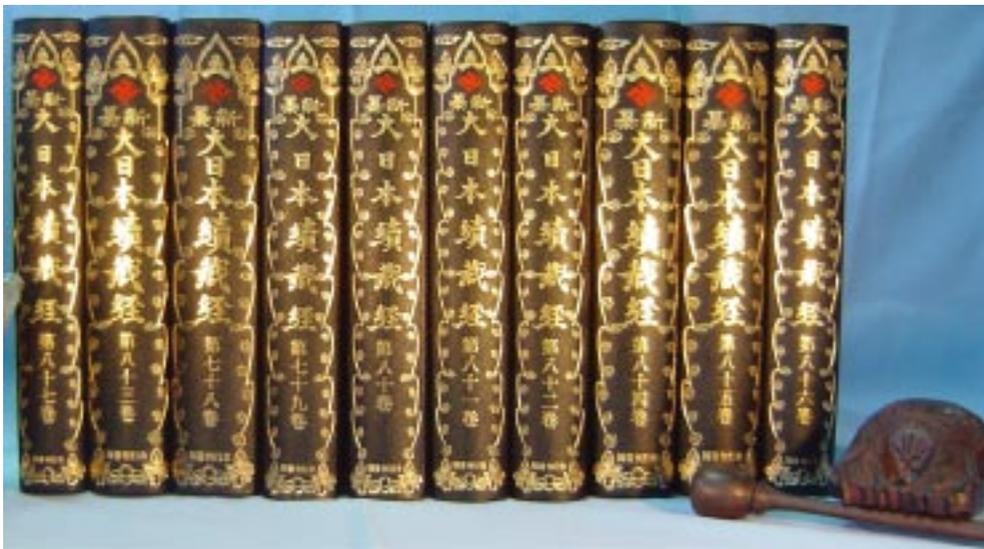
#### 一、卍字續藏的電子化

CBETA的發展是第一個做《大正藏》，再做續藏。現在做的是《卍字續藏》。《卍字續藏》的特色是它大部

分是中國祖師的著作，以及《大正藏》沒有收的印度專書部分，所以我們先把這部分補齊。我們常說《大正藏》國際上大家會搶著做，但是《卍字續藏》就不會有人搶著做，除了大陸以外，但是大陸目前的方向也不在這裡。《卍字續

\* 本文係卓遵宏及侯坤宏纂修於2005年7月29日上午9時30分至下午4時在中華佛教文化館訪問杜正民居士之訪談稿。杜正民先生，1953年出生於北投，曾留學於美國馬里蘭大學，回國後擔任英文老師多年。1987年，因年屆中華佛學研究所考試年限35歲，乃毅然決然放下高薪的教學工作，報考中華佛學研究所，畢業後留校擔任佛學英文老師。1994年，受臺大哲學系教授恆清法師聘請，為其建立臺大佛學研究中心，自此逐漸走入佛學數位化領域，開拓新一代佛學與資訊的研究。歷任中華電子佛典協會總幹事、中華佛學研究所網路資訊室主任；現任中華佛學研究所圖書資訊館館長及法鼓佛教研修學院副校長。

\*\* 卓遵宏、侯坤宏現任國史館修纂處纂修，記錄為中華佛學研究所李美寬小姐。



《續藏》（CBETA 大正藏與卍續藏所參照之底本）（杜正民提供）

藏》這部分就跟聖嚴法師有關係，這部分的經費都是聖嚴法師募來的。有一位新加坡的居士，他本來自己想做，後來他跟聖嚴法師打菩提禪三之後，提到他的願望，聖嚴法師就說我們正在做這個，所以就帶他到這邊來，惠敏法師幫他做介紹。他本身對電腦資訊也是很有概念的，他大部分的學術訓練都是在英國受的，所以對於標準做法都很清楚。他覺得我們會做得比他好，所以就決定贊助這個計畫。

這個計畫是從2002年9月9日開始，那天聖嚴法師宴請這對夫婦，我隨李志夫所長到農禪寺作陪，當天就在那邊簽約，開始長達五年的贊助計畫。因為我

們《大正藏》做了五年，所以續藏也應該是五年的時間。從那時候開始我們就分成兩條路線，一個是把《大正藏》的校勘繼續做完。校勘很複雜，每一條都要做標記、查對，以及版本考證等等，需要花的時間比較多。我們把人力分一些出來，譬如說輸入校對的人此時已沒事，我們就開始轉移到做續藏。我們的工作流程，首先是審核標準，佛典協會會先做一個定義，授權哪些小組討論哪些文獻，確定之後就會開始輸入。原則上大概有2到3種版本，有兩個方法，一個是人工輸入，一個是掃描辨識輸入，這樣兩個版本比對起來會比較精確一點，再往下就是校對。

《大正藏》的電子化首先要感謝蕭振國居士，前面提到他提供25冊的佛典文字檔。剛好臺灣也有一群人，那時候電腦還未普遍流行，他們很喜歡打字，就拿佛典來練習打字，所以一篇《心經》，可以在網路上找到許多不同的電子版本。這些對我們來講都是很好的比對，電腦比對跟人工不一樣，人工是你每次都要耗費心血的去比對，電腦則是我們寫好程式讓它直接比對，找出錯誤。錯誤會呈現不同的狀況，只要人去看一下，去選擇正確的就好，因此我們有發展一些工具出來。譬如一比對就有圖檔跳出來，馬上就可以看到原書是什麼字，就不需要再耗費時間去查，非常省事，在電腦上看到一個紅的一個黑的，就代表有錯了，再按一下紅的，圖檔就會跳出來，那個地方位置在哪裡，一看就可以做決定，真的有問題時再去查原書。有這樣的工具，就不需要3、400人來做大工程的輸入和比對了。所以在製作《大正藏》的時候，我們很感激網路上已經有很多現成的文字，我們在上面註明資料來源是誰提供的，這也許是非常有意義的工作。

在做續藏的時候這些來源沒有了，我們就需要做輸入。我們發給大陸打，價格很便宜。文字輸入完成以後，往

下就是要處理成為可以應用的文字，一方面是缺字問題，要一個字一個字的處理，去查對現在Unicode有沒有這個字？原碼是什麼？輸入的是什麼？甚至最後我們還要查對字的意義，這真的是一個很細緻的工作。所以做完以後我常說，我們幾乎可以做一本佛典的字書了。這也是我們下一期的工作，我們處理到一個程度以後，可以藉著這樣的工具，把佛典所有的用字全部跑過一遍。這樣我們就可以全部分期、分類，把它整理出來。

接下來我們要做成電子檔，其中有一個基本的概念就是基本檔。這個概念很重要，因為在早期做數位化的人比較沒有這種概念，通常是你改一個，我改一個，他改一個，結果大家的版本都不相同；同時又用不一樣的呈現，有的人用HTML檔，有的人用DOC檔。為了釐清這些差異，就要回到源頭基本檔（source file），從基本檔再產生不同檔，產生的不同檔若有問題，再回到基本檔。我們把基本檔定位為XML檔，因為XML檔本身可以產生多種版本，這也是為什麼CBETA推展得很好的原因，因為這部分我們早就選擇標準化。

## 二、資訊個性：佛法三法印

這兩年佛典資訊化為什麼在臺灣特別蓬勃發展，尤其在佛教界也成為相當重要的一部分，我想這應該跟資訊個性有關。依個人淺見，第一，我認為資訊個性跟佛法很相像，資訊是開放的，只要看得到就拿得到（what you see what you get），除非特別去設限保護；第二，把佛法的三法印用到資訊界也很符合。大部分做網路的人是無我的，因為自己做的東西要跟別人分享，換句話說，也就是全然開放的。提到無我，CBETA一開始做HTML版就是不設防，只要看得到就可以整個拷貝下來，因為我們的基本原則就是要跟別人分享。這用來做佛典很好，因為佛典本來就要推廣、跟別人分享，所以能夠跟資訊個性結合在一起。

第三，做資訊的時候我們可以體會無常，像剛開始做資訊的時候常常斷電、當機，這是表面的無常，而資訊持續發展且沒有固定的模式更是無常。我們常說資訊三個月，人間一年，這是以前謝教授常說的一句話，資訊三個月就是一個循環，每三個月就推出一個新產品。不管是軟體或硬體，都有它的週期存在，在成長過程中很快的發展。這樣的發展讓我們體會到無常，不是固定的法則在運作，而是自己要時時配合著

整個發展。對我而言，資訊是屬於緣起法，看起來漂亮的網頁，其實背後只有電腦語言，是我們加給它條件後所呈現出來的。我們要怎樣的圖像？怎樣的格式？做怎樣的呈現？這是程式語言的組合，讓這些組合呈現我們視覺所看到的假象。這點跟佛法很貼切，正如同緣起法，是各種因素的組合。當一個人開始做資訊，他可以很心平氣和的在環境裡工作，雖然不能達到涅槃寂靜的境界，至少也會有某種程度的安靜祥和。所以對我而言，這和三法印變相應的。

從這裡我就可以感受到為什麼資訊跟佛教界，不管是在臺灣或是國際上，可以很快的融合在一起，當然跟佛教的個性有關，因為佛教的包容性很強，可以跟新的文化、新的資料做結合，這是我的個人淺見。

看看日本，他們的資訊起步比臺灣早，但這兩年感覺上比臺灣慢了點。我覺得是日本的資訊個性不夠，因為他們比較嚴謹、保守，他們做的不輕易開放。資訊個性是要放出去才有回收，要跟別人分享。基本態度是分享才能夠得到大眾的回饋，這些回饋可能會有倍數甚至更多的成長，所以資訊在個性比較保守的民族可能就走不出去。歐洲的資訊也走得比較早，但他們的個性同樣也

比較保守，在資訊的發展比較沒有那麼快。當然，我們不能說比較不快就表示以後會落後，因為過了一段時間以後，他們很快就會發展起來。相對的，走美式路線就很快，因為他們常說網路是美國人的玩具，是美國人自己弄出來，自己在玩的一套東西，變成全世界得跟隨他們的標準。

早期臺灣很多東西都是跟著美國走，所以這方面我們就比較跟得上。在整個學術水準來講，臺灣也是藉此機會提升我們的國際水平，也水漲船高。以佛學研究來講，我們離國際水準還差得遠，但比較好的是在資訊方面大家同時起步，我們走得快了一點，所以水準就跟著提高起來。早期很多工具是日本開始研發的，不管我們做什麼研究，需要的參考工具都來自日本。目前臺灣除了CBETA以外，中華佛研所也做了很多學術研究資料，這些都是我們可以回饋給國際的。大陸有些佛教學者來臺灣訪問，他們就很感激臺灣，尤其中華佛學研究所把學報、出版刊物都放在網路上，提升大陸佛學研究一、二十年的功力。因為他們買不到書，即使有書也是很貴，難以取得。現在則是每一個佛學研究所都可以很快的在網路上找到他們所需要的資料，不光是電子佛典，甚至

連當代的學術研究都可以拿得到。這樣的資訊個性很適合臺灣這種海島個性，才能夠不斷地創造資料、提供資料，也能得到國際上的回饋。

### 三、資訊與佛學研究

瞭解資訊個性之後就要提到資訊素養，廣義的資訊素養包括研究的基本態度。早期大家都是紙型的研究方法，但是現在時代不一樣了，我們開始從紙型角度改為數位研究。數位研究大家是一起起步，雖然基本素養有差別，至少在起步的時候我們不落人後。現在反而是人家來瞭解我們怎麼做，在做的過程中如何去培養新一代的人才等等。所以以前是向國外取經，現在反過來是別人要向我們學習，或是彼此互相切磋。

早期我們做的都是紙型訓練，所以我們的資料、來源、研究方法都是做紙型的研究，即使是田野調查，到最後還是回到紙型的部分。我們可以看到下一代年輕人的資料是數位的，紙型慢慢會成為他們的輔助工具，數位才是他們基本的來源。所以他們的思考方式就會不一樣；他們會想到，如何在數位資源裡面做研究。載體不一樣，研究方法會跟著不同，成果也會跟著不一樣。這也是我們必須要去面對的事情！

我們的國家也開始警覺到這樣的問題，所以也開始去做數位研究典範這類的調查。我們是應該要及早做準備，因為在國外有一些課程，從小學就帶小孩子做資料庫，他們已經開始讓小孩子從小就去準備未來的挑戰，用數位的資料建構自己的資料庫，自己去應用等等。這些都是需要加入學習的課程裡面，我們都同樣在起步，對臺灣來說這是很好的機緣，這是所謂資訊素養的基本概念。

#### 四、資訊研究人才

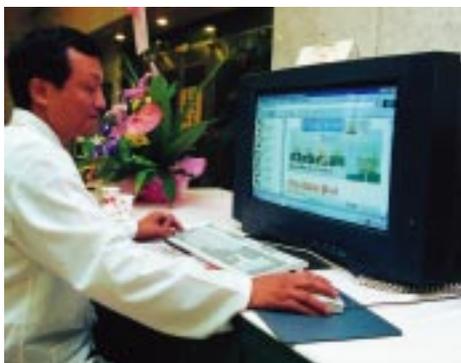
我覺得中華佛學研究所是走在世界的前端，因為我們開設佛學資訊學程，希望未來成為佛學資訊組或是佛學資訊系。這種把佛學跟資訊整合在一起的學程，據我所知，現在在世界上是唯一的。對我們來講，資訊不等於電腦，電腦只是工具，資訊則是一種思考方法，也是一種研究方法。在學界來講，資訊科系可以將資訊工程與生物科學結合，成為生物資訊學（Bioinformatics）。就人文來講，我們有圖書資訊，甚至有人文資訊。近十幾年來有個很大的思考方式，就是如何把人文學與資訊學結合在一起。同樣地在佛學來講也是，如何把佛學與資訊這樣的認知研究方法，重新

的加以整合。以佛學來講，就是在學習如何認知外界，我們有自己一套認識外界的方法，但是資訊傳介的媒體不同，整個的認知也會不同。

像我們常講的網頁設計，就很像我們所看的外相一般，表面都很漂亮，有一天佛陀來了，他說你看的都是假相，都是緣起法。我們看網頁的背後也是如此，看不到圖片，看不到漂亮的東西，看到的只是幾個程式語言，這些程式語言就組合成這樣的假相。就資訊來說，也是類似這樣的觀念，如何把人文的認知、學習和研究，結合在這裡面。

#### 五、佛學資訊學程

佛學資訊學程首重佛學的基礎，不僅僅是一般資訊的研究，所以學生要先學好佛學的基礎，然後再把新的研究方法加進來，加進來以後就可以開創一套新的領域，新的研究成果。這套佛學資訊學程在臺灣可說是一個新的嘗試。維博士以前在臺灣時也瞭解我們這套計畫，所以現在他在東京京都大學人文情報中心，也成立一個類似的學科。維博士曾在這邊授課，回到東京他也開類似的課，慢慢兩邊就會有類似的課程在進行，只是標的不太一樣。他們的標的比較屬於漢籍或是人文類的，我們的標的



第三屆中華國際佛學會議資訊展示（杜正民提供）

則是佛學的，但是兩者方法和進行的過程蠻類似的。

國際上有些學校瞭解到這個概念很重要。開始進行數位化的工作後，我比較常參與國際的活動，覺得自己這點很像臺商的個性，要把東西提到國際檯面，就要自己去開創，去告訴別人。早期大家對臺灣的東西並不是那麼瞭解，參與國際性的會議有個好處，就是很容易把東西推展出去。惠敏法師和我就是一直走向國際，從成立CBETA以後，我們一年就出國好幾次，主要是為了標準化的問題。也因此國際會議上認識維博士，並邀請他來臺灣，因為我們有共同的理念，製作佛典要標準化和國際化，與國際分享。至於走向國際化，是因為我們瞭解佛典語文有好幾種，不只是漢文，還有梵文、巴利文、藏文，還有當代文獻的英文，或是其他語文。

語文一定要國際化才能共同分享。也因為這個關係，所以我們開始往外走，參加的國際會議主要也是和佛學、資訊發展等幾個大的領域有關。幸運的是，我在1997年就成為EBTI在臺灣的代表，所以我們有機會向國際推展我們的資料。1998年CBETA成立後，我們就很積極的去推展電子佛典，最遠的走到蘇俄的莫斯科、聖彼得堡。他們看到臺灣所做的資料，也很有心做交流。我們曾跟大英圖書館簽約做敦煌資料，我們跟劍橋大學的關係也很好；我們跟美國柏克萊大學也一直維持很好的關係，因為有蘭卡斯特教授在那邊。我們在德國、英國設有據點，這也是我們走出去的成果。

由以上我們可以看出國際對臺灣的東西有所肯定，不只是因為我們做數位化，更是因為我們的水準，具有德國人的精緻、日本人的深究，還有臺灣人的勤勉。這樣的水準在整個漢文界來說是數一數二的，超越中研院早期做的二十五史。二十五史已經做得很好，只可惜做得太早，標準就沒有辦法國際化，後繼的人也沒有辦法去做轉換，這是我們覺得很可惜的地方。

說來很榮幸，我在2005年7月底接下國家數位典藏漢籍全文召集人的任務。這個計畫是中研院在執行的，他們想把

整個漢籍數位化，做出一套規範和標準，所以我們成立這樣的小組會議。有天主席在會議中提到：數位化雖然是中研院跑得最早，後來卻沒有繼續發展，這是很可惜的事；但是很高興今天有CBETA、中華佛學研究所一起來參與，把臺灣的數位化整個往上提升。

國家數位典藏計畫有一個很大的目標，就是把臺灣目前的所有研究用數位來呈現，把典藏、研究、教學整個提升上來。這裡面劃分很多小組，其中有一個就是漢籍全文部分，這部分臺灣和國外走的不一樣，尤其大陸。大陸這幾年成長得很快，早期是我們做什麼他們就拷貝過去，今天做的隔天他們的網站就有了；但反過來講，這幾年他們的速度也很快，因為他們人工便宜，很多人甚至不拿薪水，也把全文輸入一直做下去，所以現在大陸的網站，可以找到很多漢籍資料。以量來說，臺灣贏不了大陸，但是以質來說，臺灣的質非常好。在數位典藏會議中，主席也指出，連日本都要到臺灣來瞭解我們的數位化，因為臺灣有一個好處，就是我們走標準化，譬如說我們的數位走Mata Data，我們有做背後的管理。目前國際上有很多數位化只做表面的呈現，表面做得很好，但是背後的資訊沒做好，以長期保

存來講會有問題。臺灣這幾年走出數位化的特色，在做文獻、典籍的時候，掌握的很好，這和CBETA走的方向蠻貼近的。其實CBETA是同時跟著中研院的全文資料庫成長的，我們最早是參考中研院二十五史的輸入法，因為比較晚加入，所以又加入自己創新的做法，在過程中一直在很類似的國際單位裡面成長。這是因為謝清俊教授在帶的人文資訊也是一起往這樣的方向走，而且可以看到有大量的數位文獻出來，容易引起人家的注意而應用它。

## 柒、漢籍資料庫與佛典資訊化

### 一、漢籍資料庫的進展

在進行數位典藏之前，最早是中研院史語所在做漢籍資料庫，一開始做數位化就發現大陸的文獻一直出來。史語所黃寬重所長時期就開始考量臺灣是不是要跟大陸重覆做。當時大陸跟香港合作《全唐詩》、《十三經》、《四庫全書》，《古今圖書集成》則是大陸和臺灣都做。類似這樣的典籍臺灣做大陸也在做，是不是有需要？我們做出來的品質或特色在哪裡？都是他們一直在評估的。在評估的過程中他們就停頓下來，其實也不需要兩邊都做。大陸一直在走打字輸入到紙型呈現概念這條路，臺灣

則比較早走後面的技術工作。這是我們的特色。假如臺灣想在數位的領域裡再成長，應該要走後面的部分。我們雖然在量上比不上大陸，但是在知識的觀念方面能力很強，所以我們應該著眼於 Mata Data 的成長。我認為這是一個方向，也是我們在帶佛學資訊的學程時，一再告訴學生的：我們不要去做外相，因為外相的東西將來可以請得到這樣的技術人才。譬如說怎樣呈現漂亮的網頁，這種技術人員早有，但是後面的佛學內涵怎麼去掌握，才是我們的生命。這是我們在學程裡一再強調的，也是當初為什麼要請維博士來臺灣的原因。最近我們又請了一位我在網路上認識的老朋友，德國籍的馬德偉老師，他是我在網路上認識有特色的人之一。我把他請到臺灣來，他也留下來了。他很喜歡這裡，我們今年就聘請他加入，想借重他。第一，他是德國人，有相當的精緻度；第二，他有國際的敏感度，可以再把臺灣跟國際接軌在一起。

我不是說臺灣人不好，而是覺得本地人走出去的腳步還是差了一點。所以善用外面的聯繫，可以在國際上有所提升，這也是佛研所在這幾年做得不錯的地方。如何跟國際同步或跟國際接軌？維博士常說：你們常在講接軌國際，其

實你們已經早走在國際的前面了。他常常這樣鼓勵我們。就他來講，某些東西我們早在國際之上。他這麼說也有他的道理，所以這段時間，我們能順利的從資訊的素養走到資訊人才的培養過程，以面對未來的開展。

## 二、資訊的超連結

依我個人的淺見，研究方法的不同，成果也不同，我們要趕快培養這樣的人才。目前很多人都使用電子佛典的光碟，其實使用者都用表面，沒有人真正用到背後。表面就是指文字的檢索、排列，大家就覺得用得很好。其實以後的發揮不在表面，而是我們已做了很多的標記。那些標記就是把知識隱藏在裡層，也就是把排版資訊弄好，再來就要把內容資訊開始加進來。我們把這幾樣用程式語文放在後面，未來的學者可以很快的把電腦中的知識轉移到這裡面。講淺一點，就是以人名、時代、朝代做分類。

現代網路有個功能叫做Hyperlink，就是超連結，這也是資訊的特性。超連結把平面的東西變成立體化，變成多層化，我可以從這個連結到那個，從上一個連結到下一個。若解剖網路，就是這麼一個功能，其他都只是花俏而已。這

個基本功能是人類知識的一大突破，從這個平面連接到下一個平面。所以整個Hyperlink的功能，可以幫我們把知識變成多層化。我們現在帶領學生，就是在教他們後面這一層，怎麼去應用知識的建構。建構很像古人在書上做眉批、點句逗。現在則是把這種方法放到電腦的第二層，也就是所有的眉批、句逗、考證等知識，都是放在第二層，從這裡再開展出去。國際上已經有人向CBETA取經，重做他們的資料庫。在國際上有幾個單位對我們的資料非常關心，像美國印第安那大學的約翰馬克瑞教授（Pro. John McRae）就對電子佛典很有興趣。他做得比我們早，還派了學生來跟我們學習XML及Mark Up，他一直想學，但好像切不到重點，後來乾脆請了他的博士生來臺灣瞭解我們做的背後資訊，這些東西才是未來的生命。大陸也有人看到，也一直很想來參與這個工作，目前的困難是大陸沒有跟我們作直接的交流，所以還沒有進行。

我們的態度是很開放的，有人願意來我們都很高興，這是需要大家一起來參與的。我一直很喜歡做佛教的數位化，因為佛教無我，強調分享。我們今天做完就希望明天能夠上網，所以我們做的包括XML都放在網路上，這就是

我常說的「網路個性」。有些學術單位，他們的網路個性沒有這麼開放。其實如果不開放出去，會有一段時間比較辛苦。談到私心當然一定會有，我們不能完全否認，但是我們總希望能夠大家一起來看看可以再做什麼。我個人覺得電子佛典的未來不是在臺灣，而是全世界。佛典資料那麼多，不可能臺灣一個小單位全部做完。臺灣的佛教學者大約有50位左右，但是在日本則有500位，如果有他們的投入，結果就會不一樣。每一句經文、文獻都需要去做更好更精緻的版本。因為我們的定位是Beta版，Beta之後必定會有更好更精緻的版本，所以要開放出來大家一起做。

這種合作模式很適合網路，這個時代是大家可以在網路上一起做這些工作的，我們往後要培訓第二代的學者如何使用後面這一層資料。所以我把它定位為「佛教數位文獻學」。這是很新的名詞，我們做的是古文獻，卻把它轉為最現代的數位化。怎樣把傳統文獻找出數位生命，也就是這個時代新的生命、新的選擇。我目前也開了類似的課程，就是佛教數位文獻學，帶學生走向古代文獻。研究古籍，不管是校勘、版本，都很有意思。在這一轉型的時代，如何把技術和訓練的學科轉移到數位化來，

是我們要持續深入考慮，也是佛學資訊學程要做的事情。

在佛學資訊裡不再只是古文獻的研究方法，而是要把新的現代化的思考方式和技能加進來，也就是將傳統和現代資訊相結合，如此就可以在國際上交流。做得比較早的像牛津人文資訊的部門，特別強化人文資訊的功能。我們也是把人文資訊裡的佛學和資訊加以發揮。

CBETA第二層的功能，牽涉到新的技能和新的研究方法。傳統訓練出來的人，就不知如何使用第二層的功能，還只是在做檢索，做編輯，永遠留在第一層，可能他們覺得已經夠了，符合他們的傳統訓練，而且非常的好。現在找個資料只要花3秒鐘，以前可能要花3個月。但是新一代的學者所重視的不再只是質的問題，還要有個基本的量，那個量可以從我們現有的資訊及數位資料裡面去建構。從這些量他才可以走出他的質，這個質就是所謂的標記、分析等等。這個時代我們有個很新的標記語文叫XML。CBETA以前做的是XTML，後來跟隨全世界的潮流走入XML，這是一種標記語言，把文獻做成標記，可以溝通電腦和人之間的知識，這是非常有彈性的。所以現在不管在做國防資料、

公文資料或是可以交換的資料，都會走這個系統，它有它獨特的方式。像有些符號不在文獻裡面，而是告訴電腦說：這是給你讀的，你不要給別人看。所以當你看到這種符號時就知道是給電腦解讀的。我們的工作就是把很多的資訊放在這裡面，或是藏在裡面，將來要有人進去把它挖掘出來，簡單的說，就是校勘。校勘以前是放在第三欄或第四欄的下面，方便比對，現在則把校勘都放在這些符號裡面，你在讀文章的時候，看到的都是文章，但只要點一下，資料會顯示這是某一朝代的版本，不同的程式就會出來；從這裡再去恢復或是創造新的資料，這是資訊時代一個蠻好的功能。

我們要往這方面去發揮，需要訓練，需要一步一步來。先要帶學生去做Tag（標記符號），把基本知識先做起來；第二步要帶領學生學習如何應用Tag。問題是目前沒有很好的工具，我們有程式、有功能，但是沒有很好的工具，因為商業軟體主要還是在做表面的應用。要做表面的呈現工具有很多，但是要做基礎的工作卻沒有工具可用，所以我們期待在未來發展更好的工具，沒有的話就只能自己去做。所以我們還有一段路要走，但是可以看得出來這個做

法對未來人類的影響會很大。

實際去解剖電腦，我認為它只在做一件事情，就是把資訊的關係建立起來。一般我們要靠人腦來建立關係，我們的知識是零散的、有限的，我們對這些零散做了研究，去想辦法把這些關係連結起來。現在則把建立關係的工作交由電腦來做，由電腦把人、事、時、地、物建立起來，這是最基本的，而且速度也很快。

### 三、中華佛教文化館與藏經的因緣

譬如現在國科會國家數位典藏的計畫中有一個計畫，是在做藏經的經錄。我們必須回頭來提中華佛教文化館，最早在1955年蔡運辰老先生就在舊的文化館推動大藏經的影印，他的動機就跟CBETA很像，大家都有一個心願，想做新一代的《大藏經》。那時候蔡運辰老先生以及東初老人，還有早期臺灣佛教界的大德，發起倡印臺灣的第一套《大正藏》，除此之外，他們也很想加以重編，做出一套屬於自己的《中華藏》，就像現在的CBETA面臨新的數位技術一樣。他們有新的排版技術，所以想把以前的雕刻本以及拓本再重印一次。重印前最重要的工作就是整理經錄，在這過

程中，產生了22種童璋先生所編輯的經錄，以及漢文《大藏經》的研究。這些都是他們研究過程中所產生的成果。早期的文化館還保有蔡運辰老先生的手稿，他年紀大到寫字會顫抖，但是你可以看出他對經錄的投注和關心，他到80幾歲了還肯做這樣的事，令我們非常的敬佩。

臺灣第一套《大正藏》是在文化館倡印，而因緣巧合的，CBETA又搬回到中華佛教文化館來。原來我們並沒有把這兩者串連起來，等搬回來以後，才想到一開始倡印的地點就在這裡。像現在用的書架，是東老的時代就拿來擺《大藏經》的，我一直捨不得丟掉。當時影印《大藏經》到一個程度，他們就開始編藏經目錄，我們現在也是一樣。從今年開始，我們在國家數位典藏計畫裡，也開始做藏經目錄，多少也會突破古籍的格局，因為我們可以打通整個經錄。我們把古代的經錄跟後來的藏經目錄分為古經錄及現藏，現藏是指現在還存有的藏經的目錄。像道安和尚他們純粹只是在編目錄，我們試著把這兩個地方接通起來，因為一個有實體的藏經，一個完全是收集的資料，他的方式完全不一樣。後期比較像用圖書館的管理方式去做一個藏經閣的簡錄。我們就是往這方



佛教藏經目錄數位資料庫（杜正民提供）

向去做，這樣做下來就有古今結合的感覺。如何把資料文獻用現代的資訊去呈現，讓它變成入門的工具，可以瞭解一部經最早是在什麼時代，後來有那幾種不同的版本或譯本，甚至在什麼階段產生不同的章節被摘錄出去等等這些關聯，這些資料都可以建立起來，而這就是未來可以做的部分。

#### 四、製作藏經經錄與展望

國家數位典藏藏經的部分，在2005年4月計畫才被批准，我們也從4月份開始做了一些基本規劃，從7月份開始進行，到目前為止我們已有很大的收穫，就是我們在做經錄的時候，已經發現CBETA下一個階段可以做的事情。在做經錄的時候，是從現有的藏經開始去做。《大正藏》之後發掘出來的有

《趙城金藏》，還有《房山石經》。這兩個版本是《大正藏》之後大陸才發掘出來的，應該比《高麗藏》還早，因為《大正藏》的底本主要是以《高麗藏》為主。我們在做經錄的時候，一經排比就發現CBETA下一期的工程，可以從最早的文獻開始繼續補齊。我們最早就是希望把整個藏經補齊，叫「集成」，全名叫「電子佛典集成」。除了《大正藏》、《續藏》，我們要把一開始沒有的這兩種藏經補齊，這是相當龐大的工程。

我對未來資訊有很大的期望，我們這個時代只要做一個有標記的版本就好，未來的標記會用校勘呈現，因為還有另外一個超連結的功能，可以連結到圖檔。所以只要一個電子版，其他的都用圖檔，做學術研究時只要點一下，不

論《房山石經》的圖檔、《金藏》的圖檔，或是《磧砂藏》的圖檔，都可以一個個叫出來，學者就可以做比對，比對完就可以得到結論。這是現代資源共享的標準，最好的範例。

有關圖檔的大小，現在都不是問題。以前剛開始做時，1GB就叫做主機，還不到十年的工夫，現在的電腦則是100GB起跳，所以硬碟容量不是問題。早期還有頻寬的問題，現在頻寬也都解決了。一個從頻寬解決，一個從技術解決，我們內部也做過嘗試，因為工作上需要用到。講到比對就可以把圖叫出來比對，國際上玩圖檔的，他們的檔案所占的空間比我們還大量，因為很多做視覺藝術的，每天處理的就是大量的圖，比較之下，我們做文獻的檔就比較小，所以檔的大小不論現在或未來都不會是問題。

### 五、藏經成果與未來走向

我們從最原始的經文一路比對下來，發現CBETA下一步可以把《房山石經》補齊，這是我們能力可以做得到的。如果要全部重打實際上會有困難，但假如只是補齊，以學者來說只差幾十部而已，我們可以做得到，而且有經錄再比一次，就可以很清楚知道差哪幾

部。《金藏》也有同樣的問題，它也是一個早期的底本，所以我們也可以把它恢復再做。其實《金藏》是大陸《中華藏》的底本，只要有人做我們就可以做連結，只要願意合作大家都可以做連結，就算不是我們去連結，而是他們連結過來也可以，我們既不收錢也沒有限制，所以只要願意，幾乎都可以運用。假如他們不做，我們可以幫忙做，只要他們願意開放。

以長遠來看，我覺得這個工作蠻有意思的，做出一個版本後，只要按一下，就可以恢復各種版本。這些功能我們都已經有了，如果想看宋代版本，就點一下宋，宋代的版本就會出來，元、明、清也都可以點一下就出來，這也是為什麼我們要花兩年的時間去做校勘的原因，就是要把裡面的東西全部一個一個建立起來。這不只是做參照資料，而是在做恢復的工作，恢復到原來的版本，所以將來想看《房山石經》的所有經文，只要按一下，石經的部分就都會出來。當然，以目前來看，並不見得那麼完整，但是未來一定可以達到，我們這代沒有做完，下一代會有人做完。我常講反正我們這一代是不可能什麼都做完了，那麼做不完的就留給下一代來做，藏經在流傳過程裡，也是經過幾百



2007年4月法鼓佛教研修學院成立，校長惠敏法師致贈副校長聘書。（杜正民提供）

年的時間才形成的，所以我們也要往長遠的方向看。

### 捌、檢討與感言

#### 一、中華佛學研究所的支持

我個人很感激李所長具有寬廣的視野，開放讓我們來做，才能讓臺灣不管是CBETA或是中華佛學研究所跟國際同步，在這個領域上能有所貢獻，跟大家做交流。這是較具有未來性的方向，如今中華佛學研究所要改為法鼓佛教研修學院，佛學資訊學程就會改成組。再過

一、二年後，假如運作順利，就會朝改成一系所的方向，假如這樣的路線可以一直走下去，則臺灣佛教在這方面的前程是可以預見的。

我們走得比較先進，看得比較遠，各種不同方向的學習角度，可以做不同於以往的切入，這個方向在國際人文界也很重要。我在2005年8月底到英國參加IABS，這是佛教學者的國際會議，該會議主題很廣，從原始佛教，甚至從原始佛教和婆羅門教之間的關係，一直談到當代的佛教。我們在2000年加入之後，

就決定把電子佛典當做會議的一個主題來討論，到現在已經是連續第三屆，由此可看出大家對電子佛典的重視。而且不光是漢籍，各種語文的佛典都涵蓋在內。在9月份美國維吉尼亞有個很好的小會議，是以人文資訊為主，大家來探討如何因應這個時代，來製作一套新的教學思考方式，從資訊、人文、工具應用的角度，來探討一個新的方向。所以可以看到國際間一直往這方向去做，一個是應用，一個是因應。很幸運地，臺灣還可以跟上這樣的腳步，一起來參與。

我們這幾年參與的佛教會議能夠談的東西很少，最早是EBTI，是因蘭卡斯特教授的推動。但是後來CBETA做出來以後，基本目標也達成，所以EBTI的會議也告一段落。之後像EKAI就開始從EBTI

走出去，除了佛教，還從時間、空間跟語言的角度，來探討人文的數位資料。我們開始從這個大方向去探討，參與的團體也比較多，除了原來的佛教團體以外，還有國際上從事各種角度諮詢的人員來參與。所以我們稱EKAI為電子文化地圖，他們談的角度就是時間、空間、語言。從這三個角度來切入，未來在讀書或看資料時，都可以在一個文化地圖上呈現，可以把地圖裡面的資訊建構起來。譬如中古地圖，可以從隋唐一直建構下來，不同層面但是在同一個地方，同一個地方裡面可能有不同的人、不同的東西、不同的食物，可以一層一層的向前發展。現在圖書資訊系也開始往這方向去考慮，就看如何去管理這些資料。



參與國際會議，分享佛典電子化的經驗。（杜正民提供）

參加電子佛典這類的國際性會議，大多是我跟惠敏法師去參加。如果是漢籍資料，在資訊方面，中研院參與的人很多，他們在這方面的腳步跟得很緊，事實上也走得很快。我個人覺得謝清俊教授對於帶動臺灣人文資訊這部分很投入，他也是早期國家數位典藏的發起人之一，他雖然在退休之後就交給下一代接手，不過他依然很熱心、積極的參與帶動。

對我們來說，在做一個工程的時候，要把過程，還有一些改變，以及為什麼要做這些改變等因素記錄清楚，後來的人就可以避免重蹈覆轍，會比較容易往下開展。當初《高麗藏》轉成印刷版的時候，可能也沒有人想到後來還會有個CBETA把印刷版轉成數位版。而在轉換的過程中，我們會面臨同樣或不一樣的問題，需要去瞭解。我手上這本古寫經，是我去（2004）年去京都開會時買的，它開啟了我的視野。我看到佛典的發展過程，就像CBETA的歷程一般，一開始很確定採用的底本就是《大正藏》，並且原版呈現，也不去更正；過了二百年之後，日本人就開始知道不對了，就開始有所謂的校勘本，用紅字在旁邊寫出錯誤的地方，然後字體也開始變化，開始走出日本人味道的字體。最

早就是原始版本用什麼字體他們就照用，慢慢地在字體上就有變化，到第三代以後就應用到生活裡，和生活做連結，譬如把經文繡在扇子上等等。

我覺得做電子佛典，好像在走相同的歷史過程。從一字不漏照抄，到開始做校勘校對，到我們開始推廣到生活化，如何做出PDA版，可以隨身帶著走。以前是一張光碟帶著走，現在是一張小小的SD卡就可以帶著走。這些都是跟生活結合，或是做成跟生活有關的網頁資料。在跟生活結合之後，日本就慢慢走出他們自己的佛教出來了，不過那已經是第四、第五代之後的事了。

## 二、走出新的資訊生命與無限的成長

我相信電子佛典往後會走出新的資訊生命來，或許不在我們這一代。也許我們現在無法預期，就像日本人剛開始接受新的佛學，他們也無法預期自己會走出自己的日本佛教一樣。可能在整個歷史發展的過程裡，也是這樣的狀況。所以在那樣的展覽會場裡，我真的很受感動。

從資料庫開始，資訊就是有生命的，因為它並非一出來就是史料。紙型資料除非有再版，否則就是定型了。資

料庫或網路的資訊則是隨時可以更新、改良的。所以對我來說，它有資訊生命，而且這種生命是無窮的，只要有人願意去維護，讓它成長，它就可以無限擴展。

早期臺灣做資料庫的缺點，就是做完就丟在那裡，不再做維護或發展，非常可惜。資訊不加以維護就沒有生命。假如資料庫是大家共用共享，而且基本上提供的單位是開放的，那麼這個資料庫應該是屬於大家的。就像我們的資料庫，我們定位在不屬於CBETA擁有，也不屬於某一單位，而是屬於大家的，所以大家一發現錯誤就會回饋回來。本來我們很天真地想要做一個機制，讓大家直接在上面修改，後來發現很難管理，所以目前還沒有開放到這種程度，但是至少開放到學者可以進來做某種程度的訓練。將來我們希望每個領域都有人去管理它，讓Beta版做得更好。這都是資訊時代可以做的創新，也是資訊的生命。資訊有它的生命循環週期，但也有可超越循環周期的發展。這是兩個不太一樣的開展。

在網路上我們有CVS的功能。CVS就是只要有登入密碼和權限，就可以進入網頁，在修訂版上做更正。不過，修訂以後不會影響到原始版本，我們可

以看到幾年幾月有什麼新的版本，後面的人還可以做修訂。這都是現在網路上的新概念，亦即「網路是所有智慧的集合」，要如何讓所有的智慧結合在一起，就需要有個機制，讓它具有活力，也就是我可以進來學一點，你也可以進來做補充。

### 三、駭客入侵

當然，開放性也有它的問題，譬如我們曾經碰到駭客進來破壞。最早在臺大的時候，某晚有一個從國外進來的駭客進來取走所有的資料，順便把檔案清得乾乾淨淨。那時候我們沒有防火牆的概念，臺灣也沒有做得那麼嚴謹，幸好我們有備份，發生問題時還可以恢復過來，雖然還是有損失，不過至少不是完全消失。因為難免還是有駭客存在，所以就需要有補救的機制。目前CBETA有防火牆，但是我們有個想法，資料不是只有在這裡，而是屬於全世界。所以我們設有好幾個點，這些點我們都是同時放出去，不會因為這個站不見就沒有，這是共享的好處。

### 四、傳承與感言

中華佛教文化館是非常重要的傳承點，把早期影印的《大正藏》跟現在



杜正民居士（Aming Tu）

的數位藏經做了結合，這樣的因緣很殊勝。我在做電子佛典時有個感覺，人是因緣在運作，像我常用英文名Aming Tu來稱呼自己。我常說其實Aming Tu實際上是虛擬的，是大家的組合，並不是有這麼一個實體，我只是這個實體的代表或一起工作的人。今天的訪談事實上也不是談我個人，而是這個時代有這群人共同做了這麼一件事，少掉一個人或少掉因緣就沒有辦法完成。很感激一開始有恆清法師的緣起，後來有北美印順導師基金會的贊助，後來又有聖嚴法師及中華佛學研究所的參與，使整個運作一直很順利。最重要的是後來有一群無名英雄包括CBETA同仁的投入，很多未曾謀面的人的參與以及很多在外層的贊助者。當我靜下來，覺得因緣好像在我身上環繞，這個環節身體的某個部分……，那個感覺很清楚，就個體

而言，是所有因緣的組合體，然後藉著某個身體來運作。非常感激這麼多的因緣，讓我們剛好在這個時代可以參與。我們常講「人身難得」，其實「時代難逢」更重要。

大部分參與這件事情的同仁都投入將近十年的時間，一旦計畫停止，大家就沒有工作，原本有很好的工作的，現在什麼都沒有。這都是因為大家有使命感，要完成佛典的數位化；否則在薪資不高，也沒有好的福利待遇，更沒有任何保障下，誰肯同甘共苦的去克服困難，完成工作？

剛開始的時候很辛苦，沒有經過這個歷程的人是無法體會的。我們常會擔心，做完今天不知道明天的經費在哪裡？又不知道可不可以做得下去？能不能完成？……等等的壓力一直存在。慢慢地，我們做了幾年以後就開始看到前景，整個情況就變好了。剛開始辛苦是因為：第一，這條路沒有人走過，所以我們沒有任何前例可以依循，沒有任何前人的智慧傳承，所以看到的每一個問題都是在創新、解決，雖然很有挑戰性，但是壓力也很大；第二，是財務的問題，中華電子佛典前兩個階段的經費都是從國外來的，臺灣贊助的很少。我們雖然也開放網路樂捐，不過收到的金

額都不多，有時候收到一塊或兩塊錢的美金，扣掉手續費後，只剩下幾分而已，但是我們很高興有人來參與，無形中提供讓人做功德的機會。前兩階段的經費讓我們穩定下來安心的做，搬到中華佛教文化館後，這個場地才讓同仁穩定下來，這是後來的發展。在早期因為財力不固定，人事也比較不固定，比較擔心會有所變動，另外在技術方面也有許多問題等待克服。

我記得第一次五年計畫做完時辦了一場成果發表會，在會場上我哭出來了，我不是故意要哭的，就是話講著講著，那份感動就出來了。我想到這五年我是怎麼過來的，那些回憶好像一下子在身體再流轉一次，所以很自然的眼淚就掉出來了，因為我太感激有這樣的機會去參與。從第二期開始我比較有穩定感，我們在技術上已經走了這麼一段路，知道怎麼前進了。

目前（2005年）我們正在做第二期的第三年，前五年做《大正藏》，後五年是做《卍字續藏》。假如有第三期，我們就會做《集成》，把《房山石經》、《磧砂藏》一起補上來。我們這個單位也不會去化緣，雖然常有人建議我們募款，但我覺得不適合大肆去化緣，夠用就好了，這是我個人很大的缺點。

還好我們的主任委員惠敏法師很支持，並且把這個團隊帶得很好，讓我們有一個大的方向可以掌握。第二期計畫將會在2007年的3月份舉辦成果發表會，所以我們只剩下一年多的時間，相當緊迫。接下來還得再傷腦筋下一期的經費從哪裡來。財務的來源是我們很大的重點，沒有錢就很難留得住這樣難得的團隊。

### 五、中華佛學研究所的貢獻

從慧日講堂的時代開始，中華佛學研究所就是主要贊助單位，主要成員有惠敏法師、維博士、我，還有周邦信，周是屬於程式人員，我們4人就是協辦單位的主力。李志夫所長也非常支持，每次的活動他都會過來幫忙，或者擔任活動的發起人。中華佛學研究所一直是一股很重要的力量，但是我們盡量不要讓它顯得太強勢，因為實際上這是屬於大家共同的事業，是屬於佛教的事業，所以也不能讓大家認為只有中華佛學研究所在做，雖然大部分的事情其實都是佛研所在做。北美印順導師基金會非常好，他們不干涉我們，讓我們放心放手的去做，所以我們非常感激仁俊長老。但就整個過程來說，佛研所確實是最大的後援，假如CBETA沒有這樣的學術單位做後援，可能很難走入今天的國



中華佛學研究所圖書資訊館的珍貴線裝書（杜正民提供）

際化，假如沒有中華佛學研究所的學術性和國際性，我們是走不出去的。從得到國際上的認同，到普遍廣泛地應用，這段時間很長，一旦得到肯定就會很順利，進度很快。一是我們東西好，二是在國際上也得到肯定。

後來剛好慧日講堂有人事上的變動，同時中華佛研所正巧又搬到法鼓山，所以我們就跟聖嚴法師建議，把CBETA搬到中華佛教文化館。那時候法鼓山圖書館還沒有落成，有人認為圖書館尚未建好，就把書留下人先上去，不過，我則建議把佛研所的圖書先搬上去，我的想法是，就戰時來講圖書館是軍糧單位，人上去軍糧必須跟著上去。我們搬進來那天剛好是耶誕節，所有圖書在搬上去的第一週就開放，然後等圖書館整個落成以後，我們再搬第二次。

佛研所搬家之後沒多久，我們再把CBETA從慧日講堂搬到中華文化佛教館。在短時間之內連續搬家，所以我打趣說我們可以去開個搬家公司了。搬圖書館是找義工幫忙，CBETA搬家則是請卡車載物品，我們自行整理，同仁們非常好，事前就把辦公室規劃得很仔細，像隔板都找同樣的顏色。搬回來之後聖嚴法師也一直很支持我們，因為他也很認同這樣的事業，我們就很安穩的在這

裡發展下來。另外一點，就是法鼓山的信眾和團體，也慢慢接受這個事業跟法鼓山的因緣深厚，彼此也能夠結合在一起，未來佛學和資訊方面也將會有很好的發展。

## 玖、餘論與展望

### 一、經費收支概況

北美印順導師基金會每年贊助我們20萬美金，大約為新臺幣六百萬元，但是對我們來說仍然不夠，因為當時CBETA有15位成員，每位成員即使只拿3萬元月薪，每年最起碼也要有五百多萬的人事費，更別提其他設備開銷以及年終獎金了。記得有一年在尾牙的時候我幾乎掉了眼淚，因為不知道錢在哪裡？有同仁問我年終獎金在哪裡時我根本無法回答。

我非常感激北美印順導師基金會，因為他們後來也碰上財務的困難，但是仁老認為答應人家的事一定要做到，所以他們寧可自己在財務上有問題，還是讓我們繼續運作。只是錢進來的很晚，而且又沒有跟我們說為什麼錢那麼晚進來，所以我只能夠在這邊乾著急，不過後來總算是有驚無險，有一年就在過年的前一天錢才進來。

以我們資訊業一開始的運作來講，在那個時代應該是蠻省的。大概有90%左右是花在人事費，就一個團體來講很特殊，我們在設備方面很省，因為大部分的經費都是用在產出，也就是同仁的經費以及打字輸入的費用，這都是產出資料很重要的部分，這或許不一定是很健全的做法，但因為沒有錢，所以我們只好把重點放在產出的管道上，其他的能省就省。

後面這幾年比較好一點，因為人員減少，人事費用壓力比較輕一點，人員的減少是因為整個制度穩定了，可以用資訊取代，不一定要使用人力。另外我們學到的經驗是，可以委外的就盡量委外，因為委外可以省掉很多成本。因為他們是以量來計算，我們就可以知道實質可以回來多少，對我們來說出去多少錢就要回來多少字，這樣我們就能夠管理。在辦公室的管理成效反而不如委外好，所以各有優缺點。後來我們做了一些人事的調整，也讓經費稍微順暢一點。這幾年聖嚴法師的支持，讓我有安全感，錢沒進來沒關係，因為師父那邊的錢可以讓我們周轉。

第二期費用比照前期，也是逐年給。表面上看起來好像錢不少，但是做下去以後才知道根本不夠。像國家型的

計畫動輒要上億才能做下去，所以惠敏法師在評審國家型計畫時，常講我們CBETA很厲害，這麼一點點錢就可以做這麼大的事業，別人要了那麼多錢卻只做出很小的成果。所以我們在這方面的確是做到成本控制。

很多人說做網路好像是無底洞，怎麼也填不滿，而且還看不到，往往一開始什麼都看不到，我一開始也是擔憂這個問題，深怕如果沒做好，整個錢就泡湯了。

## 二、圖書資訊館的開展

早期在恆清法師臺大佛學研究中心做網站資料庫時，除了蒐集目錄和全文之外，也做一個全面的佛學網站，類似佛學網路資料的入口，所以我會關心整個華文界佛學的網站資料，後來沒有再接觸，所以瞭解的比較少。以我個人現在粗淺的看法，臺灣需要再一次翻轉，不要再停留在資料庫的階段，或只是做網路資料的階段，而是要去做下一個階段，即知識的產生與管理，這也是佛研所圖書資訊館現在開展的方向。圖書資訊館已經超越早期我在臺大做的資料庫，那時候我們純粹是在建構資料。圖書資訊館所做的服務除了資料以外，還應該提供作者一個很好的管道，如何從

這些資料去產生新的知識出來，就是我們在思考的問題。例如我們規劃一個寫作平台，可以讓作者很快在網站裡面寫出一篇文章，譬如說要做書目或是引文（citation）、註腳（footnote），都是基本的訓練工夫，但是很耗費時間以及精神，這些應該在資料庫就能產生出來，當我選擇以後，就能產生我們要的格式，這就是標準。做了基本的標準化以後，譬如寫論文可能要用MLA的寫作格式，或是APA的寫作格式，或是其他的格式，你可以做選擇，像我們現在建構的有《大正藏》和《卍字續藏》的部分經書全文，還有早期臺大和目前佛研所正在建構的當代文獻的全文，我們有兩套基本寫作的文獻，一個是原典，一個是當代文獻。我們寫文章只要就現有的這兩種就已經相當足夠，因為我們有現成的citation，這些東西可以很快的引用在文章裡，所以我常對學生笑說，把這些弄完以後再把and、or、not加上，你的一篇文章就出來了，再加上你自己的幾個見解，根據某一段經文，現代那些人做了什麼研究？譬如印順導師怎麼說？還有其他什麼學者怎麼說？然後我的看法是什麼？很快的你的一篇文章就可以出來了。這就是寫作平台的運用構想。



與圖書資訊館同仁合影（2006年）（杜正民提供）

另外就是徵引書目的部分，以前的書目下面我們都會寫上徵引的資料來源，其實引用的資料裡暗藏一條學術的脈絡，這一代的人都引用印順導師的論著，印順導師引用太虛大師的著作，或是引用他那時代其他的學者，我們這代教出來的學生又引用我們的資料。我們建構出這樣的脈絡，在現代的citation的技術裡面可以建構起來。我們現在做的是分析臺灣這幾年的學術趨勢，你可以看得出來這段時間大部分的人都用哪些佛典？因為我們可以追溯回去所引用

的佛典部分，包括引用哪些人的資料？臺灣佛教學術研究有很特殊的一點，學者都不太想用別人的，大部分都要寫自己的，所以引用的不像其他人文科學那麼多。這些東西可以產生引用的脈絡出來，這是屬於citation的等級。

### 三、從資料庫走向知識庫

往下我們思索的是應該如何從現有的資料庫產生知識出來，也就是要把關聯建立起來。我們現在花很多時間回到詞彙，把所有佛教的名相聯結起來。這

些事國際上已經有人做，並且整理了30萬以上的佛學名相。這些名相要如何重組？如何建立彼此的關係？便可以在這上面很快的找出彼此的關聯性。譬如上網路書店買書，他會告訴你還有其他相關書籍你可能有興趣，其實這就是背後的知識，告訴你這個key word跟其他key word會有關聯，或是這個關鍵字跟哪些佛典有關。這些我們都需要去做底層的分析，到最後必須回到處理字或是語意的問題。所以我們現在在建構的是佛學的語意網，這些都只是初步工作，但是我們已經看到如何走向知識面的部分，而不再只是資料面的部分。

我們之前利用臺大資料庫來作測試，發現用人的智能去想到的關聯，沒有辦法像電腦一樣，電腦會用統計分析把人沒有想到的資料串連起來，會把一些很奇怪的關鍵字幫你抓在一起。譬如你喜歡《華嚴經》的某一個章節，電腦可以找出這個章節裡面有哪些關鍵字出現的最多，以及它們之間的關聯性，除此之外，電腦還可以把不是《華嚴經》的其他經文的相關部分都結合在一起，然後再找出相關的當代的文獻，從古到今都可以幫你串出一條路線出來。這是臺灣未來可以走的路——從資料庫走到知識庫的路線。目前臺灣往這方面走的

比較少，大陸更不可能，因為他們目前還處於做資料庫的基本階段。早期資訊界有一個從資料到資料庫——從資料庫到知識——從知識到知識庫的概念，最後看看能不能開展出智慧來。這樣的學問在資訊界已經成為一個學門，就是知識工程，已經超越早期資料的建構，去作知識分析。我也講不出它未來的功能會在這裡，只知道它會超越我們過去的資料搜尋檢索，把裡面隱藏的知識呈現出來。所謂知識管理最基本的概念，是資料的輸入、整理、彙整之後再重組資料，再來就是資料的重現。未來這幾個步驟可能靠這樣的工具，而產生新的智能出來。這部分我們已經和臺大資訊工程學系合作開設一些課程，目前有歐陽彥正老師及黃乾綱老師開設相關的知識管理課程。這是比較有挑戰性、比較新的部分，期望未來研修學院一開始就往這方面大力推展。

這對未來的研究與寫作會是不一樣的呈現，跟古登堡一樣，我們做資訊業的常說：「我們現在做的就是資訊的革命，未來變得怎麼樣？我們目前還不能去預期。」但是就像日本的古寫經，可以看到它的發展步驟。古登堡的印刷術改變整個人類的思想，而資訊革命的力量比印刷術更強，它的傳播、儲存量、

再生的力量，足以超越紙型印刷的時代。1994年從網頁（Homepage）出來到現在，短短十年的發展，就已經有這麼大的影響力，我們預期在短短的幾十年之內，那種革命應該會很大。對我而言，走在這條路上可以感受到它的壓力。做資訊的人會有怕趕不上的壓力。沒有參與，生活在外面就不會感受到這種壓力；一旦投入，你便需要面對新的技術出現，新的標準以及功能出來。目前以資訊發展來說仍在嬰兒期，還不是成熟期，就像佛教剛在漢末傳入中國的時候，因為尚未成形，每樣東西都是新的挑戰，再過一段時間等大家都準備好以後，會在佛教界裡面產生不一樣的成果。

#### 四、資訊的革命

《聯合報》曾經刊出，90%的電腦使用者都使用微軟的作業系統，換句話說，比爾蓋茲控制了世界上90%電腦使用人口的思考方式，因為我們都要跟著他的作業系統走，整個思考方式慢慢會被他定型，這的確是一個問題。但是我們帶佛學資訊學程的人，不是帶他們用工具，而是帶領他們創造新的東西，所以就比較沒有這樣的限制。我一直認為CBETA不是在提供檢索、介面，而是

提供應用資料給人家，也就是後面的XML，只是現在還沒有拿來用而已。目前的學者還是用檢索編輯，但下一代會懂得應用XML。這是我們在第二層提供的，也是期待未來的人去做的部分。這股力量就潛藏在裡面，需要有人趕快去使用它。目前使用的人很少，幾乎不懂得去用，這是因為大家被限制在基礎訓練也就是紙型訓練上。紙型訓練很簡單，你只要給我一些資料，我把這些資料排列整合成一篇新的文章就好了。現在我們把XML版本提供在網路上，或是在以前的光碟裡面。以前出兩片光碟時，很多人都不知道如何使用XML的版本，所以現在只出一張光碟，另外把XML版放到網路上，提供大家來下載。那就是我們所謂的原始檔（Source file），從原始檔就可以看到裡面藏有很多的Tag，Tag就可以把你的知識、別人的知識都放在裡面，跟經文原版產生新的知識出來。