

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

Semantic markup for Chinese Buddhist texts and its Application - A platform for Querying and Spatio-temporal Visualization of the "Biographies of Eminent Monks"

洪振洲

Jen-Jou Hung

法鼓佛教學院 圖書資訊館 助理教授

Assistant Professor

Library and Information Center

Dharma Drum Buddhist College

馬德偉

Marcus Bingenheimer

法鼓佛教學院 圖書資訊館 館長

Director

Library and Information Center

Dharma Drum Buddhist College

許智偉

Zhi-Wei Xu

法鼓佛教學院 圖書資訊館 組員

Staff

Library and Information Center

Dharma Drum Buddhist College

【摘要 Abstract】

《高僧傳》文獻的語意標記是在 2008 到 2010 年間，法鼓佛教學院於浩然基金會贊助下，所完成的數位文獻專案成果之一。此專案完成梁、唐、宋、明四本高僧傳，共一千三百多位的高僧傳記之內容語

email：洪振洲 jenjou.hung@ddbc.edu.tw；email：馬德偉 m.bingenheimer@gmail.com；

email：許智偉 marxlovebaseball@gmail.com

意標記。標記的內容除包括了文獻中所有的人物、地點、時間實體外，也包含三者交錯所形成的「關係點」。本專案完成的語意標記文獻，為目前所知最大量的以 TEI 標準完成之中文語意標記文庫。利用標記成果，我們得以使用特定的人物、時間或地點作為搜尋條件，來搜尋高僧傳的內容。並且透過線上介面，結合 GIS 的技術應用，讓使用者體會到《高僧傳》不一樣的面向。

The semantic markup of the biographies of eminent monks is part of a project conducted at Dharma Drum Buddhist College between 2008 and 2010 with support from the Haoran Foundation. It contributes to the field of Digital Humanities and electronic librarianship in several ways. To our knowledge it is the largest classical Chinese corpus with structural and semantic markup in TEI. The four biographical collections from the Liang, Tang, Song and Ming dynasty contain more than 1300 biographies of the most famous figures in Chinese Buddhist history. With markup distinguishing and referencing all personal and place names as well as all dates in these texts, it is possible to present the texts in much more user friendly format. In this paper we describe an approach to interface design based on semantic markup and inspired by GIS technology that maps the contents of the biographies in geographic space. The semantic markup allows us to search the dataset in ways that go far beyond available indexes. It is for the first time now possible to search e.g. for events at a certain period in a certain area. Another innovation is the concept of nexus point which describes basic events in the form "One or more person(s) at a certain time in a certain place". With these nexus points expressed in markup, we can display events in Chinese Buddhist history on a map and reference these visualizations directly back into the texts themselves.

[關鍵詞 Keyword]

佛教數位典藏 語意標記 地理資訊系統

Buddhist Digital Archive ; Semantic Markup ; Geographic Information System

壹、緒論

臺灣「數位典藏專案」的目的在於：「希望以眾多計畫產出的各類寶貴數位化資料為基礎，與其他主軸計畫共同進行教育、商業及學術等領域的加值應用，並配合協助進行國際交流（拓展臺灣數位典藏計畫，2008）。」而十年來，臺灣漢文佛典數位化的重心都在於建設「基礎」，也就是「數位化資料」。然而對於「加值應用」卻較少著墨。漢文佛典的數位化，例如佛光山在 1997 年完成《佛光大藏經》（佛光山電子大藏經，1997）的數位化，中華佛典電子協會（Chinese Buddhist Electronic Text Association, CBETA）更在 1998 年到 2010 年間完成《大正藏》、《卍續藏》及《嘉興藏》與其他藏外文獻，這些都提供了數量與品質均令人讚嘆的數位產出。雖然專案產出龐大，但可惜的是，目前佛典的內容數位化方式大都仍只停留在對於文獻形式、文獻背景資料，如經名、年代、作者、譯者等後設資料（Metadata¹）的建制。在

¹ 後設數據（Metadata），又稱元資料、中介資料，為描述數據的數據（data about data），主要是描述數據屬性（property）的資訊，用來支持如指示儲存位置、歷史資料、資源尋找、文件紀錄等功能。（維基百科，2006）。

全文數位化工作已經進行了十年的今日，藉由日新月異的資訊技術的協助，我們可以對於這些數位文本進行更精細的加工，以提供更多語意上的知識與產生更進一步應用。在本研究專案中，我們利用語意標記與相關資訊處理技術創建了一種跨資料庫、結合數位文獻、可視覺化的數位典藏應用。

《高僧傳》²是研究中國佛教歷史時一個不可或缺的重要參考來源，高僧傳收錄的傳記內容橫跨了東漢到明朝初年的重要高僧，合 1382 篇傳記。且更重要的是包含了魏晉南北朝與隋唐兩個中國佛教發展的關鍵時代。但是由於《高僧傳》的內容充斥著許多古代的人名、地名、時間，以及由這三者間交錯而成的關係訊息。因此，閱讀者除需要一定程度的佛學知識之外，也必須對於古代的地名與時間有相當程度的瞭解，否則將不易閱讀其內容。再加上中國古代文獻中，將對於人名的敬、暱、諱、貶包含於稱呼之中，使得許多佛學研究者，仍無法駕馭《高僧傳》中寶貴的知識內容。

為解決此一問題，我們利用取自中華佛典電子協會提供的《高僧傳》數位化資料並進一步進行語意的標記。其目的除提供「時間」、「人物」、「地點」與「關係點」的詳細參考資料之外，也旨在利用這些語意的標記資料，構建一數位資料庫系統，以提供進階的語意查詢功能。因此使用者得以在本系統中，詢問類似於「在《高僧傳》的紀錄當中，在唐代有哪些僧人曾拜訪過長安？」這樣的問題，並可快速得到解答。為此，本研究專案建立了《高僧傳》時空地理平台 (Bingenheimer, M., Hung, J. J. & Wiles, S., 2009)。我們除提供資料豐富的線上閱讀介面外，也成功結合地理資訊系統的技術，提供高僧傳內容的時空視覺化。此一平台的誕生讓使用者得以在閱讀《高僧傳》的當下，透過地理資訊系統的介面，一窺文本中所蘊含的時空資訊的分佈，藉以給讀者一個全新的角度來瞭解《高僧傳》的內容。我們確信，這一研究的成果將可以提供有關中國佛教史一嶄新的觀點，以及新的佛教歷史研究方法。

本篇論文之後續章節架構如下：第二章簡介本研究專案對於《高僧傳》文本內容的標記方式。第三章介紹利用此一標記完成之《高僧傳》數位文本，加上地理資訊技術所打造的《高僧傳》時空地理平台的功能與技術細節。第四章中，利用一些簡單實例，展現《高僧傳》時空地理平台的效果。在第五章中，則對本文進行總結，並簡述未來可能的研究方向。

貳、《高僧傳》標記處理

本研究使用的《高僧傳》數位底本，是取自中華電子佛典協會製作之 CBETA 光碟 (CBETA 電子佛典集成, 2010)。這些數位底本已經經過 CBETA 工作人員利用 TEI 標記語言³進行過相當程度標記作業。

² 這裡的所指的《高僧傳》包括了慧皎的《梁高僧傳》、道宣的《續高僧傳》、贊寧的《大明高僧傳》與如惺的《宋高僧傳》等四本傳記的集合。這四本並稱為《四朝高僧傳》，在本研究簡稱為《高僧傳》，其內容可將眾多佛教發展史中零星散亂的片段串連起來，乃為研究中國佛教歷史的重要著作。《高僧傳》是現今被保存下來最古老且完整的一部「僧傳」，也被譽為中國佛教高僧傳記的鼻祖。《高僧傳》不但在佛教發展史上佔有重要的地位，而且有大量佛教相關的地名、人名與時間收藏其中，非常適合更進一步進行語意的標記與時空資訊的視覺化運用。因此，本研究選擇《高僧傳》做為視覺化古籍研究的初探。

³ TEI 標記語言可以很完整地記錄原始文獻，並將其很精確地轉化到電子文獻上，讓我們對有興趣的目標進行編

TEI 標記語言可以讓我們對電子文本中有興趣的資料進行標記，並視專案與研究的需求，決定標記的目標與重點。在 CBETA 提供的電子文本中，所有的校勘記錄、行號、文本的經名、年代、作者、譯者的標記已經被妥善標記（CBETA 文獻資料—簡介學術版 XML 標記，1999），這也是我們目前能找到的最好的數位版本。因此本研究使用 CBETA 提供的數位版本，作為進一步標記的出發點，並將研究產出回饋給 CBETA 組織，以提供更多佛學工作者使用。

本研究的進階標記目的在於識別文本中的人、時、地實體，並紀錄彼此間的互動關係。進階標記內容，可分為兩部份來描述，首先是人、時、地實體的標記：包含有基本的「時間」、「人物」、「地點」的辨認。而關係點的標記：關係點是由「時間」、「人物」、「地點」三個基本元素所構成，因此當完成基本的人時地之實體標記後，我們可以於其上進行關係點的標記。

一、人、時、地實體的標記

人名、地名、時間三者的標記，雖然處理準則類似，但在處理過程中，各種型態之標記有各自不同的挑戰。我們在此分述如下：

(一)人名實體標記

在本研究專案中，我們採用相當寬鬆的人名認定方式。因此，幾乎所有的人名實體，都在我們的標記範圍內。舉凡在文本中出現的所有的人物，包括神祇、鬼靈或虛構之神話人物等等，以及那些只知道姓氏但無法確定其全名者，譬如文本中出現的「藹法師」、「炬法師」等，看來像是暱稱或簡稱的稱呼方式，都被納入人名實體的標記範圍內。更重要的是，只要文本中出現指涉這個人的文字就標記，甚至代名詞也都標記。例如以玄奘法師為例，包括「唐三藏」、「奘」，甚至是文句中以「之」或是玄奘的別名等等都標記。這樣的作法雖然會增加工作負擔，但另一方面也確保了所有的文獻中的人名都被標記到。人名實體使用的標記規則如下：

<persName key="人名唯一碼">人名實體</persName>

其中<persName>是 TEI 中用來標記人名的標記符號，而 key 屬性內容則是用來識別此人名實體的唯一碼⁴。今舉一範例如下：

碼，以執行更高階的記錄與搜尋，TEI 標記語言因此非常適合本研究的需求。由於 TEI 是一種非常開放，且為國際泛用之標準。因此標記後的文本不但可在不同的平台上做文件處理，且具有國際交流以及資源交換的特性。更多有關 TEI 的細節請參照「TEI 使用指南——運用 TEI 處理中文文獻」（Burnard, Sperberg-McQueen & Bingenheimer, 2009）。

⁴ 為識別在文本中相同與不同之人名實體，我們給與每一個經過辨識出來的不同人名實體一組唯一碼。在文本當中，相同人物的不同別名，將使用相同之唯一碼。唯一碼的標示對程式自動化處理資料內容，有十分重大

<persName xml:id="A012516">釋圓震</persName>姓陳，中山人也

而上述之“A012516”就是釋圓震之唯一碼。若在文本中出現指涉多人的名稱，且該名稱為歷史上經常使用的稱呼，例如竹林七賢、彌陀三聖等，則使用人名群組⁵的方式標記。例如：

<persName xml:id="G000001">竹林七賢</persName>

經過人名實體的標記後的文本將使相關的搜尋功能變得非常強大，因為傳統的全文檢索功能下只能搜尋「玄奘」這兩個「字」，但經過標記的文本將讓使用者可以去搜尋這個「人」出現在哪些文本之中，超越了以往傳統搜尋只能對「字」進行檢索，這樣的搜尋方式將更貼近研究者的需求。

(二)地名實體標記

和人名標記方式一樣，在文章中出現的所有地名與地名的代名詞都要標記。但實際作業情形中，發現文本為釐清其指稱的地點，有時會在目標地名前包括一長串更大範圍的地名，例如「天都長安京兆府草堂寺」。在此情況下，依文意，此地名乃指「長安草堂寺」而言，則標以「長安草堂寺」之唯一碼標記該地名。這樣做的目的是希望不要讓使用者閱讀文本時產生混淆。地名實體使用的標記規則與人名實體使用標記相同，表示如下：

<placeName key="地名唯一碼">地名實體</placeName>

另外，有時文本會出現指涉多處地名的名詞，例如五嶽、安西四鎮、三韓等等，則使用地名群組的標記方式，例如：

<placeName xml:id="CN000000G01AA">五嶽</placeName>

不過古代地名距離今天年代已久遠，文本中常常會出現許多地名是完全無法考證其確切位置的，而當本研究團隊無法確定其正確位置，但仍知道其大概位置時，仍會標記出來。此時，我們將使用最小可能範圍所在地區之區碼，並使用該區域的行政座標建立的地名權威資料。例如龍丘寺在史書中完全找不

的幫助。此外，將該唯一碼與「法鼓佛教學院規範資料庫」（佛學規範資料庫，2008）使用之編碼相同，讀者可以藉由唯一碼於此資料庫中探尋更進一步的參考資料。而本專案所創建之時空地理平台，也引用了「法鼓佛教學院規範資料庫」的資料，對所有已辨識的實體，提供詳細的參考內容。

⁵ 人名群組為一群人名實體的集合總稱，在「法鼓佛教學院規範資料庫」中，人名群組的內容，實際上僅紀錄便是各相對應的人名實體的參考連結，實際上的資料是紀錄在各人名實體紀錄之中。利用此一彈性作法，資料庫內容將可允許使用者以人名群組或單一人名實體為條件來進行查詢。

到相關的資料，其寺廟也未保留至今，所以座標無全無法確定。但是文本中有指出是「衢州龍丘寺」，因此我們知道龍丘寺在「衢州」。對於這樣的地名實體，我們還是選擇將其標記，以「衢州」的行政中心，作為「龍丘寺」的座標，並注記為「估計值」。

(三)時間實體標記

時間實體的標記採用 TEI 中用來標記時間標準標籤<date>。但相較於人名與地名實體的標記作業，時間實體使用的之標記較為複雜，這是因為在文本中，所指稱的時間有時並無法明確指出該段時間的起訖點。因此，當文本中所指時間十分明確，且為單日時，我們使用以下標記：

```
<date when="0394-03-26" xml:id="d53766155376615">晉太元十九年甲午之歲二月八日</date>
```

其中 when 屬性的內容是採用推算格里式曆(Proleptic Gregorian Calendar⁶)來表示該日期，而 xml:id 內容則是該日期的唯一碼，用來讓程式自動化連結規範資料庫，以取得參考資料。而若是內容時間是指某一段可以確定的時間，例如文本指涉開元二年三月、貞元六年、天寶年間、唐玄宗當政年間、整個明朝等，則使用"from" 與 "to" 屬性的搭配，取代單一的 when 屬性。以貞元六年為例，"from" 與 "to" 屬性的值將會紀錄貞元六年第一日與最後一日之日期，如下所示：

```
<date from="0790-01-24" to="0791-02-11" xml:id="d55211905521573">貞元六年</date>
```

在某些情況中，文本所指稱的時間是很模糊的，例如貞觀年間唐玄奘在長安，因為我們無法確定玄奘是到底是在哪一年出現在長安，所以使用" notBefore " 與 " notAfter " 屬性的搭配，用來表示這種不確定性。範例如下：

```
<date notBefore="0627-01-26" notAfter="0649-07-13" xml:id="d54616575469861">貞觀年間  
</date>唐玄奘在長安
```

除上述規則的訂定之外，實際上進行標記作業時，經常遇到文本中並無直接表示時間，而是利用其他重要事件來表示時間。例如"唐武帝廢佛之時"，這時本團隊則會利用外部參考資料，找出唐武帝打壓佛教的時間為何，來進行標記。此外，我們也利用此一機會，進行文本內容的修訂。當本團隊發現文本內容的時間資料不是最精確的，且我們找到更精確的日期時，我們會標記更精確的資料，例如「釋處寂」於開元二十二年示寂（佛光大辭典，2008），但更精確的資料指出是開元二十二年正月示寂（CBETA 電子佛典集成，2010），在此我們會標記開元二十二年正月的時間碼。最後，當文本中所指時間是某個年號

⁶ 格里式曆。所有曆法皆按格里曆（Gregorian calendar）向前推算。（ISO 8601:2004(E)，2004）。

的初年或是末年時，本團隊為了使標記更接近史實一致，因此我們定義「初年」泛指該年號的前四分之一的時段，「末年」則選擇後四分之一的時段來表示。

二、關係點的標記

除了人名、時間、地名之外，本團隊還會記錄每一篇文章中所發生的事件，其事件本研究稱之為「關係點」(nexus points)，關係點指的就是文本中有一段文字，其時間、人物、地點俱為說明同一標的。對於一個關係點，我們使用一個<linkGrp>來標記，並儲存於每一篇文章的最後面。關係點的標記可說是把前面的人名、地點與時間串連起來的關鍵，不但讓使用可以從「事件」的角度去搜尋，也可將其「事件」視覺化在介面中。

由於一個關係點，指的是在同一時間，同一地點，一人或多人之間所產生的互動。因此，一個關係點僅能包涵一個地名實體與一個時間實體，但相關人物可以包含一個以上。而一篇文章中可能有好幾個事件同時存在，例如《唐高僧傳》中《唐京師崇義寺釋慧顛傳》的內容中某一段標記如下：

```
<persName xml:id="A005757">釋慧顛。</persName>.....<date when="0573" xml:id="d54419125442276">十歲</date>出家。師事舅氏<placeName xml:id="CN0410502T01AA">光嚴寺</placeName><persName xml:id="A005758">明智法師</persName>。.....。
```

由於這事件是包含：「釋慧顛」與其師「明智法師」於釋慧顛十歲時，在光嚴寺的事件。因此我們在該文章之後，新增關係點標記如下：

```
<linkGrp>
  <link targets="#0535a07 #0535a08"/> (標明這事件在文本中行號的起始)
  <ptr target="#A005757" type="person"/> (釋慧顛)
  <ptr target="#A005758" type="person"/> (明智法師)
  <ptr target="#CN410502T01AA" type="place"/> (光嚴寺)
  <ptr target="#d54419125442276" type="time"/> (十歲)
</linkGrp>
```

在一個<linkGrp>標記中，我們會新增許多<ptr>的標記，每一個都指向與此事件相關的人時地實體。而<ptr>標記中的 target 屬性值，就是該參考項目的唯一碼。利用此一標記，我們可以成功紀錄文本中所發生的重要事件，目前本團隊在《高僧傳》中共記錄了六千多筆的關係點，讓使用者可以使用人名、時間、地點三者做交叉查詢。在關係點的標記中，時間的決定是很重要的。但很不幸的，在許多情形下文本的內容，並沒有清楚的指出一個事件的確切發生時間。對於這些事件，我們嘗試的提供一個可能發生的時間範圍，這個時間範圍可能來自文章前後的相關性，或經過其他相關文獻參考而得。但是對於那些

無論如何都找不到更清楚的時間範圍的事件，我們使用該篇文章主角的「預設時間（default time）」來表示。預設時間的主要用意，就是利用主角的出生時間到卒沒時間，作為相關於該作者的事件時間判斷之最寬鬆準則。而由於預設時間與傳記主角的生卒年有關，因此預設時間的決定，也受到該傳記主角是否有清楚的生卒年紀錄所影響。我們將主角的預設時間的決定方式，條列如下：

- (一) 若主角有確定的生卒年，則是取主角的十歲（生年+9）到卒年為他的活動時間範圍，並使用"from"與"to"屬性的搭配，例如釋圓震（705-791）貞元六年終，享齡八十六，預設時間就決定為為 714 年的一月一號到貞元六年，以求其最大範圍，標記為：**<persName xml:id="A012516"><date from="0714" to="0790" type="default" xml:id="d54934085521573"/>釋圓震</persName>**。
- (二) 若只確定卒年而不確定生年，則使用"notBefore"與"to"屬性的搭配，並假設其年壽為一百歲。
- (三) 只確定生年而不確定卒年，則使用"from"與"notAfter"屬性的搭配，並假設其年壽為一百歲。
- (四) 當生卒年都不確定，但確定歲數與知道在某個年號中卒時，則使用"notBefore"與"notAfter"屬性的搭配。例如某僧人齊永明年間卒，年七十二。齊永明共十一年（C.E. 483 - C.E.. 494），則確定其「卒年」最晚不會超過 494 年，最早不會超過 483 年。因此十歲年為 $483-72+9=420$ ，卒年為 494，因此預設時間 420 年，跟齊永明年間的最後一年。也就是**<date notBefore="0420" notAfter="0494" type="default" xml:id="d53860275412930"/>**。
- (五) 當生卒年與歲數都不確定，但知道在某個年號中卒時，也使用"notBefore"與"notAfter"屬性的搭配，例如某僧人齊永明年間卒，年不詳，則假設其僧人一百歲。齊永明共十一年（C.E. 483 - C.E.. 494），則確定其「卒年」最晚不會超過 494 年，最早不會超過 483 年。因此最寬鬆的十歲年為 $483-100+9=392$ ，卒年為 494。此時預設時間就設定為 392 年，到跟齊永明年的最後一年。也就是**<date notBefore="0392" notAfter="0494" type="default" xml:id="d53758005412930"/>**。
- (六) 當生卒年都不確定，但確定歲數與知道其在某個年號中生時，與當生卒年與歲數都不確定，但知道其在某個年號中生時，其處理方式與上面相同，也使用"notBefore"與"notAfter"屬性的搭配，並一上述原則求算求其最大範圍。
- (七) 當生卒年、歲數與在哪個年號中卒、生都不確定時，就利用其他相關事件來輔助判斷。例如某僧人在齊永明年間做了某一件事，標記者可假定其僧人是在五十歲時做了那件事（其假定歲數要依前後文與各方資料判斷），並假定其年壽為一百歲，所以其十歲年數假定為 $483-50+9=442$ ，卒年為 $494+50=544$ 。因此預設時間就取 442 年到 544 年。也就是**< date notBefore="0442" notAfter="0544" type="default" xml:id="d53940635431682"/>**。

參、《高僧傳》地理資訊系統

《高僧傳》地理資訊系統的主要設計目的，是將經過妥善標記的《高僧傳》數位文本中的時空資訊

由高僧傳文本中取出，並且將這些資訊與地理資訊系統（Geographic information system，以下簡稱 GIS）相結合，提供一個新的檢視《高僧傳》資料內容的角度，並希望藉由此一工具發掘出佛教歷史的新知識。而將文獻內容與地理資訊結合，也是一種前所未見的嘗試。為此，我們結合了許多網頁、XML、與 GIS 的相關技術，來打造這個線上地理資訊系統的平台。此系統的一個最大的特點就是，我們使用了最經濟解決方案的方式來發展這一個線上平台。也就是說，本專案所使用的所有發展技術，都是免費或是甚至是開放原始碼授權大眾使用的技術，也因此，我們相信這一個線上平台的架構與開發經驗，可以非常簡單被複製到其他類似的專案當中來使用以作為其他人文資訊專案的典範。最後，本系統使用 Google Earth 做為時空地圖的呈現方式，Google Earth 軟體（Google Earth，2010）把地理資訊、經緯度、航空照片、3D 照片與衛星照片整合在同一個介面上，不但是目前較穩定且普遍的虛擬地球軟體，重要的是可將搜尋結果轉成 KML 檔案供使用者保存，這樣的特性能夠讓使用者更容易存取與分享自己的研究發現。

一、系統架構

圖 1 是此系統的架構示意圖。根據圖 1 所示，本專案網站主要由 php 撰寫而成，而本平台的內容資料，都以 xml 的方式儲存在名為 eXist（eXist，2010）的 XML 原生資料庫之中。當前端介面需要 XML 資料時，php 程式將執行一段由 XQuery 寫成的腳本，以便由 eXist 資料庫中取回所需的 xml 文件，或文件片段。而由 eXist 取回的 XML 資料，系統將利用一段預先寫好的 XSLT 腳本，將 XML 內容轉換成 HTML 的格式輸出至使用者的瀏覽器當中。而同時間，php 腳本將會把 XML 資料中的事件資訊取出，產生相對應的 KML 資料，並交給 Google Earth 來進行視覺化處理。同時間回傳的網頁內容將引用法鼓佛教學院規範資料庫的網頁服務（web services），來提供線上閱讀《高僧傳》的使用者同時可以看到更豐富的參考資料。由於文本的視覺方面大致上只掌握在 XSLT 的腳本之中，進行一致的風格管理，當有需要修改呈現的頁面時只需要修改 XSLT 腳本即可，不需要對所有文本都進行修改，此可以節省修改的時間。下圖 1 為本系統的架構圖。

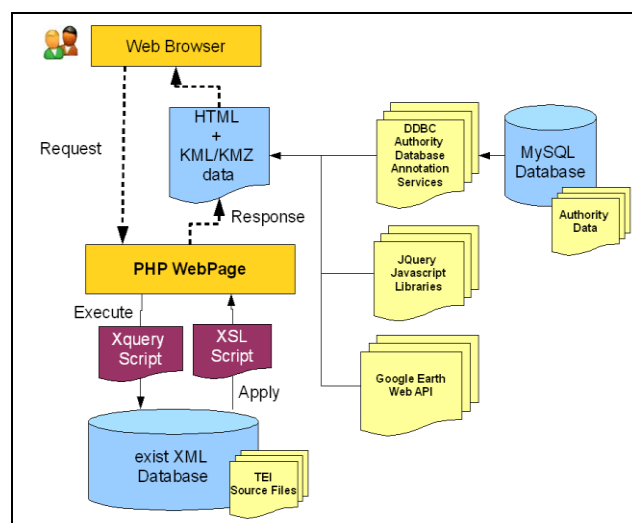


圖 1 高僧傳地理資訊系統架構示意圖

經由這樣的系統架構，使得法鼓佛教學院規範資料庫與《高僧傳》XML 文本、Google Earth 三者之間可以作迅速有效率的存取，下圖 2 是本系統所提供的五大服務：

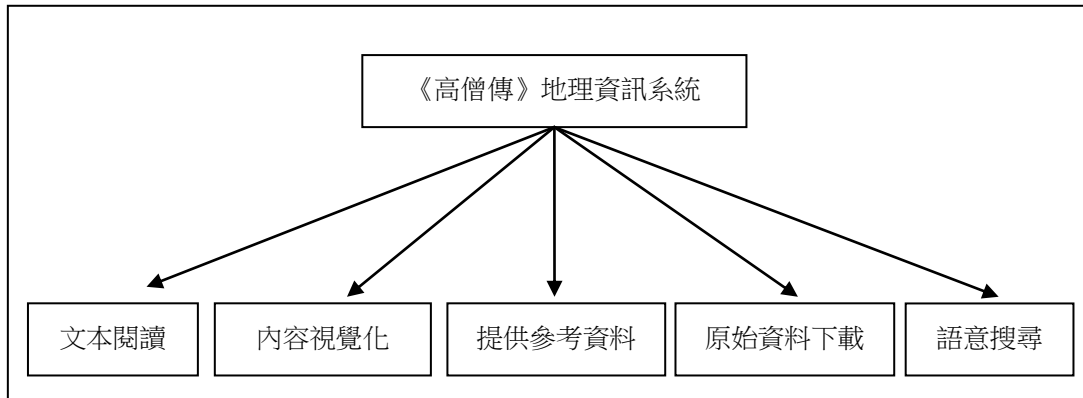


圖 2 《高僧傳》地理資訊系統的五大服務

下面將依序介紹本系統所提供的服務，包括文本的閱讀、視覺化，以及與法鼓佛教學院規範資料庫的連結，和語意搜尋的功能。

二、文本的閱讀、視覺化與提供參考資料

文本閱讀的選擇是利用下拉式選單讓使用者直接選擇想看的傳記，選單的結構是以「傳別」、「類別」、「人名」的階層式架構而成，除了可藉由三層結構來過濾資訊外，也可以節省網頁擺放文本篇名需要花費的空間。而這樣為節省使用者讀取網頁的時間，我們使用 Ajax 搭配 jQuery (jQuery, 2010) 的技術將使用者的查詢送至後端，有別於傳統送出查詢後網頁必須重新整理的缺點，使用 Ajax 可以提供使用者一個更即時、更方便的操作模式。下圖 3 是目前的介面首頁。



圖 3 《高僧傳》地理資訊系統操作介面首頁

上圖中，框框所指示的部份，就是讓使用者選擇傳記的下拉式清單。選擇後，畫面左方將隨之出現欲閱讀的傳記文本，而畫面右邊的 Google Earth 區塊則會將文本中所提到的「地名」與「關係點」呈现在虛擬的地球模型上。文本視窗中的時間、人名與地名實體，都以不同顏色標記出來。

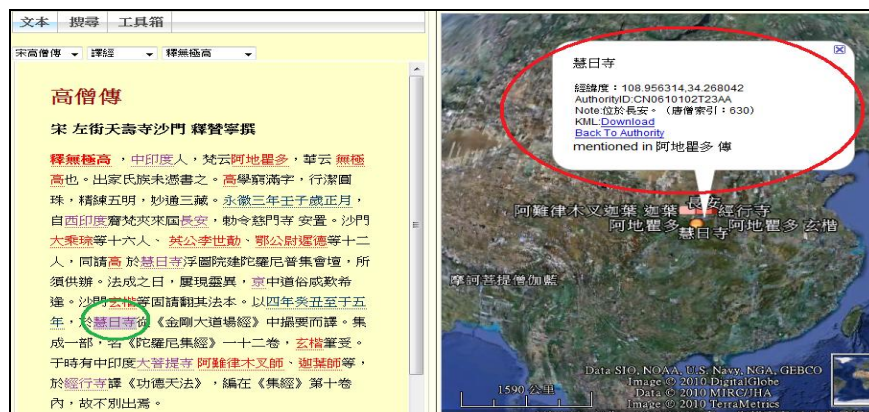


圖 4 地名的彈出視窗

由於本系統使用了法鼓佛教學院規範資料庫註解服務（洪振洲、李志賢，2009），因此，使用者可使用滑鼠點擊這些實體，系統將會連結到法鼓佛教學院規範資料庫並跳出介紹其人名、時間與地名的視窗，視窗內除了相關的重要資訊之外，還包括「Back to Authority」的連結，可以讓使用者獲得該實體更多的參考資訊。若是點擊地名的話，不只是會跳出介紹視窗，連帶的介面中的 Google Earth 也會把其地點標

示出來，如圖 4 所示。

如下圖 5 所示，而在人名的彈跳視窗中，包含有該人名的重要參考訊息，例如朝代、生卒年、籍貫與其他別名等。這些資料都是輔助讀者瞭解文本內容的重要參考。

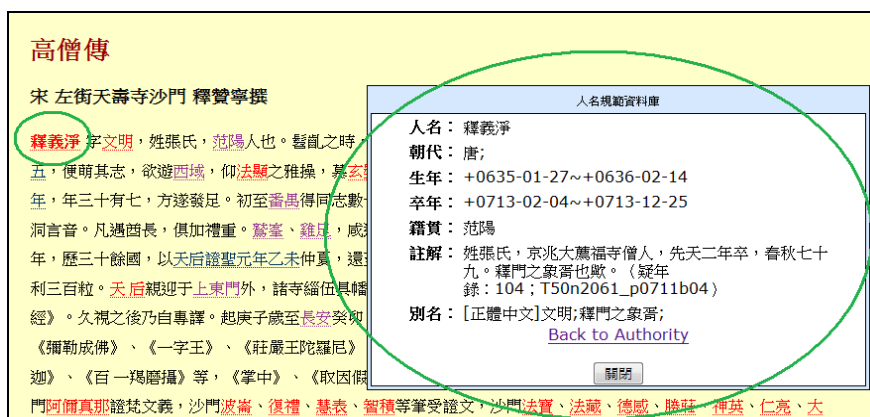


圖 5 人名的彈出視窗

下圖 6 是時間的彈跳視窗的一個範例。由圖中，我們可以清楚知道時間彈跳視窗所包涵內容，並非只有文本紀錄之時間相對應的西元時間而已，也包括中國該時間所有的存在的王朝的對應時間，以及日本曆法的時間。這些豐富的資料，將可以提供最完整的時間參考資訊。

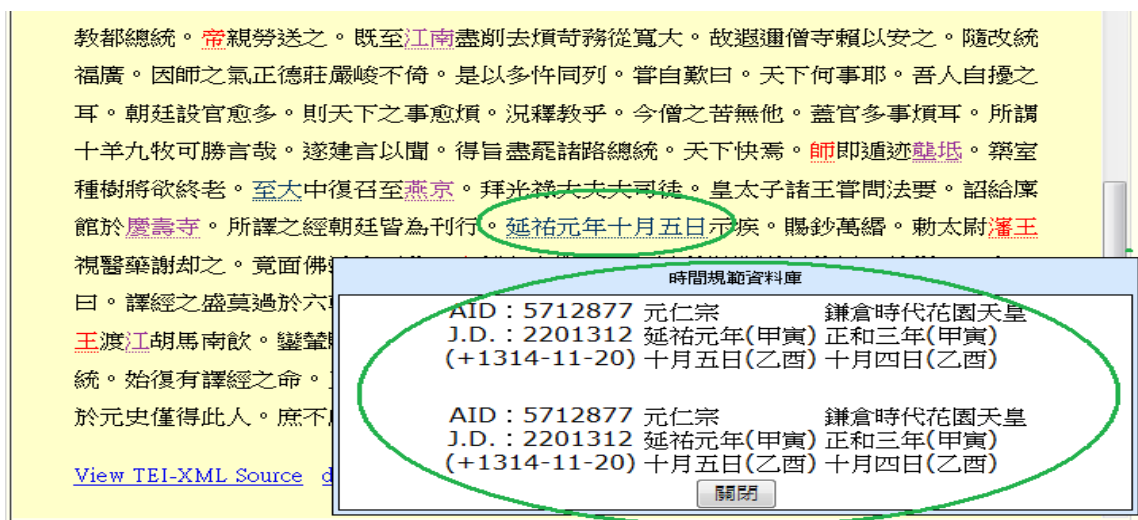


圖 6 時間的彈出視窗

在畫面右邊的 Google Earth 區塊中，把文本辨別出來的關係點，將被表示成一個個色彩相異的方塊，並且放至於 Google Erath 介面中相對應的地理位置。而當有多個事件同時發生於同一地點之上時，這些彩色的方塊將會被安置於一個格狀的版面配置之中，方塊之間並不會互相重疊。因此使用者將可以一眼看清在該地點上發生的事件總數。下圖 7 是一個實際的例子，由該圖中，我們可以清楚看到關係點集中

於三個地點。更準確的說，一個事件發生於最北邊點，八個事件發生於中間的地點，而三個事件發生於最南點。

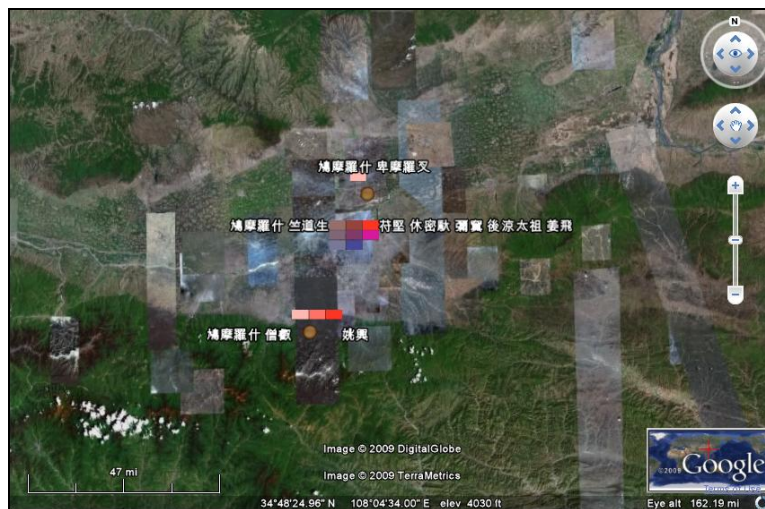


圖 7 高僧傳地理資訊視覺化範例

而當展現於地圖中的關係點被進一步點擊開來時，Google Earth 中將出現一個對話視窗，並將相關於該關係點的詳細資料，以及回到文本所在的連結，如圖 8 所示。



圖 8 高僧傳地理資訊系統地點擊後範例

關係點除表現在 Google Earth 之上外，系統也會在文本介面最左邊會整理出本篇所有的關係點。如下圖 9 所示。若使用者以滑鼠點擊這些關係點紀錄後，除了在 Google Earth 相關的地點會被標記出來外，

文本的介面會反白來告知使用者此事件是寫在文本的何處。並且為提供使用者更多互動的可能性，我們在每一篇的文本的最末，不但提供使用者下載標記後的 XML 原始檔案，也提供 KML 格式⁷的檔案下載，使用者只要安裝 Google Earth 即可在自己的電腦上觀看該篇文章視覺化後的結果。

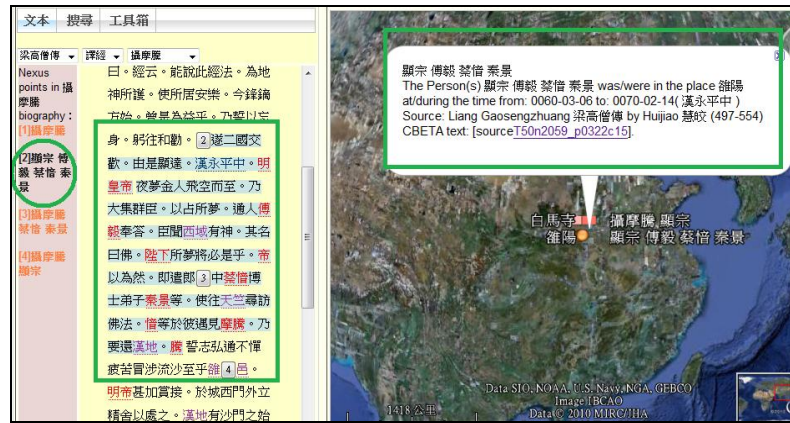


圖 9 關係點與文本之互動

三、語意搜尋

數位化文獻資料庫最大的貢獻，就是提供使用者便利的檢索功能。但隨著資訊技術的發展，傳統的檢索功能，如：全文檢索、目錄階層搜尋、經目查詢、行號搜尋，已經不是一個令人驚豔的新玩意。實際上，這些傳統的檢索功能也只是一種簡單的文字比對功能，也因此通常無法滿足研究者對於研究的需求。在傳統的全文搜尋方式之中，為了回答「玄奘在貞觀年間時在長安做了什麼事？與什麼人見面？」或其他類似問題，使用者必須精確的知道所需的關鍵字為何，並且在一次又一次的查詢之中，慢慢的修正關鍵字，以期望達到最好的結果。但實際上，全文檢索還是無法克服語意上的問題，例如玄奘有許多不同別名、唐貞觀年間的表示方式有許多種可能、長安城中大大小小的地名不可記數。這些問題都會造成傳統全文檢索準確度低落。

相較於傳統的搜尋功能，本研究建立的搜尋功能是以文本中標記出來的關係點為基礎，讓使用者得以進行以人名、地名、時間為條件，進行交叉檢索。由於關係點是經過人工解讀，並已經抽出相關人實地資訊的知識結構，因此以關係點為基礎的語意搜尋與全文檢索等傳統的搜尋方式比起來，可說是更直觀也更貼近研究者的使用方式。並且，更重要的是，這樣的搜尋結果往往可以提供使用者沒有注意到的新知識。

(一) 語意搜尋介面

⁷ KML 為 Keyhole Markup Language 之縮寫，是一種以 XML 為基礎，而設計來表達地理資訊資料的標記語言，目前已經被大多數的 GIS 軟體所支援 (KML, 2010)。

系統的搜尋介面以導引使用者設定所需搜索條件為主要設計方向。下圖 10 為系統畫面截圖。介面中包含「搜尋目標」、「搜尋條件」與「輸出模式」等三大設定區塊，使用者可以根據需求進行條件設定。而無論使用者做了何種選項的決定，系統會在搜尋介面的最下方，以有意義的句子，讓使用者清楚目前的查詢條件設定。



圖 10 高僧傳地理資訊系統語意搜尋介面

目前系統提供了兩種不同的搜尋目標：「關係點搜尋」與「籍貫統計圖」。在關係點搜尋中，系統按照訂定的搜尋條件，將符合條件的關係點，展現於 Google Earth 的展示平台中。而籍貫統計圖，則是讓使用者設定有興趣的年代與區域，將出生在該時空條件下的高僧⁸，根據其籍貫加以統計。

(二)關係點搜尋

在關係點搜尋下，搜尋條件中可以選擇傳別、人名、地名與時間進行交叉搜尋。由於高僧傳的重要事件，都已經被處理成以人名、地名與時間實體為主的資料結構。並且每一個實體也都有一個唯一編號。因此，在查詢過程中，我們並不針對內容進行全文比對，而是將使用者輸入之搜尋條件，轉換為人名、地名與時間實體的編號，利用編號的比對來進行搜尋，這樣的作法將可以成功解決中國文獻中對於人、時、地都有太多不同別名稱呼的問題。為成功轉換使用者輸入條件成為系統使用之唯一編號，我們在查詢介面中，完成了人名、地名與時間實體的選擇器。

針對人名搜尋，系統還建立了「自動產生搜尋建議」功能，讓使用者更容易選擇到所需要查詢的人物，如下圖 11。

⁸ 此處所指的高僧，乃指在四部高僧傳中，有獨立傳記的僧人。

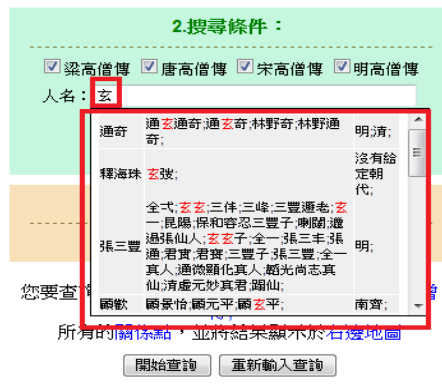


圖 11 人名搜尋欄位的「自動產生搜尋建議」

選擇時間的介面中有三個選項，首先可讓使用者選擇要查詢單一的時間還是一個範圍，第二，可讓使用者選擇要查詢中國年、日本年還是西元年。最後，日曆選單可讓使用者輸入更詳細的資料，例如朝代、帝號、年號、年、月、日等等。如下圖。

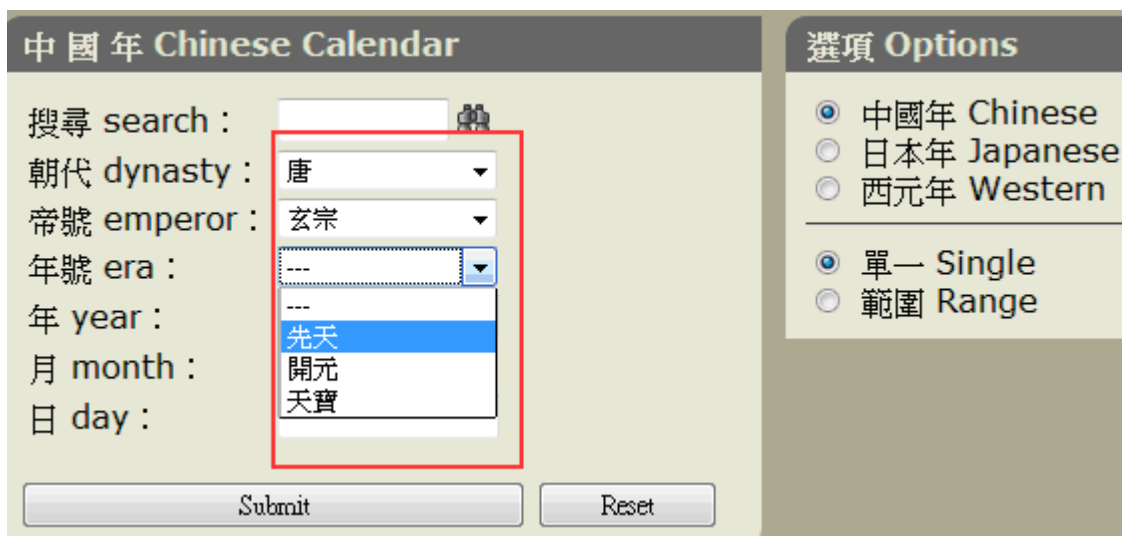


圖 12 時間條件選擇介面

而在選擇地點的介面中，分為區域選擇（Area Select）與單點選擇（Location Select）。當使用者按下區域選擇之後，將可以自由自在的在地圖上畫出所需的查詢範圍。如下圖藍色方塊所示。

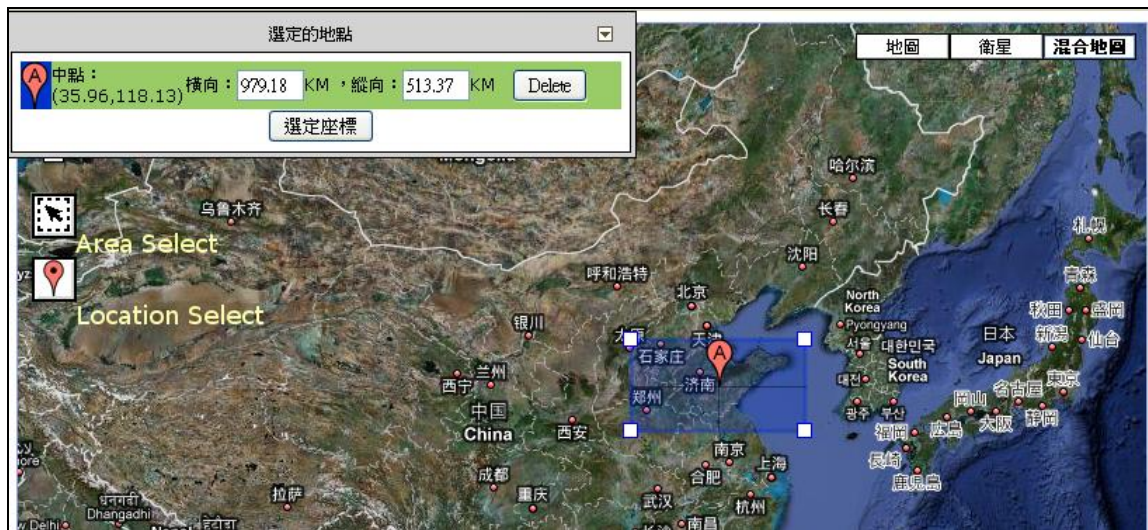


圖 13 地點搜尋條件選擇介面

而系統會為其選擇的地點自動展現區域的中點與橫向、縱向的公里數。使用者可以自由更改橫向與縱向的距離，來定義自己需要的搜尋範圍。而當使用者按下單點選擇功能時，本系統會結合法鼓佛教地名規範資料庫的資料內容，提供使用者歷史地名位置的搜尋的對話框，如下圖 14 所示。此一對話框中一樣會有「自動產生搜尋建議」功能，協助使用者找到所需的地點。

搜尋地名: 長安		搜尋	選擇勾選的地點	關閉
地名	別名			
乾明寺	長安寺;			
中寺	長安中寺;	沒有給定朝代		
長安五重寺				
長安八水		沒有給定朝代		
長安		西漢		
長安		東漢		

圖 14 地名搜尋欄位的「自動產生搜尋建議」

利用兩者的結合，利用本系統，學者可以簡單定義出「以唐朝的長安為中心，東西南北皆五百公里的範圍」。而相反的，若該查詢的地點是一個寺廟，且使用者只想查詢該寺廟發生了什麼事情，則可以把橫向與縱向的座標各改為一公里（或更窄），就可以只查詢在該寺廟所發生的相關事件。此外，為使更多專案可以使用本專案所開發之查詢介面，節省開發時間，我們將上述人名、地名與時間條件的輸入元件皆已獨立成可分離之網頁元件，以方便其他專案使用 (Hung, J. J., Bingenheimer, M. & Wiles, S., 2010)。

(三) 籍貫統計圖

籍貫統計圖目前就是統計高僧傳所有高僧們的出生地（籍貫），並以 Google Earth 的介面來視覺化。此功能的操作介面可以讓使用者分別選擇傳別與時間範圍，如下圖。

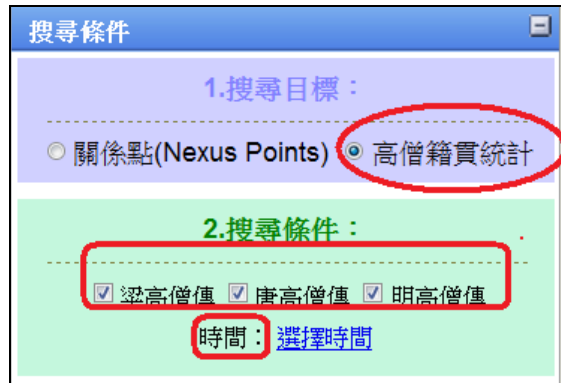


圖 15 籍貫統計的介面

在此一功能中，我們也善用了 Google Earth 優秀的 3D 處理能力，讓高僧的籍貫以立體長條圖的方式視覺化在 Google Earth 上。下圖是以梁高僧傳作為搜尋範圍的高僧籍貫統計圖。



圖 16 梁高僧傳籍貫統計

肆、實例應用與討論

由於《高僧傳》對於佛教史研究之重要地位，因此許多先賢大德，已經出版過許多針對《高僧傳》內容的統計研究。這些研究可以很概括的告訴我們一些重要事實，例如：「在梁高僧傳中，有 27 位正傳僧人之游錫地在長安」（嚴耕望，2005）。但礙於統計標的的選擇以及技術上的限制，這些統計文獻的資料都有其限制。首先在人名統計方面，傳統的研究只以正傳之主角為限，但這樣將會遺漏其它非正傳但與佛教發展有重要關係的人物，且也無法看到該正傳僧人與其他僧人之間的關係。在地名統計方面，傳

統的研究在地點上只能選擇容易定義的高僧標目地、最後住錫地、出生地、遊錫地等等，但其實該區域不一定是該僧人弘法最久與影響最大的之區域。並且，其地點的統計都是用大範圍的區域歸納之，例如學者嚴耕望在其著作中（嚴耕望，2005）將標目地為「上明寺」的竺僧輔與曇徽、「長沙寺」的曇翼與法遇等都歸類到「荊州」之下。這種忽略了小地點的統計方式，雖然仍然能提供一定程度的統計資訊，但是對於小地點與佛教發展之間的關係的資訊，便付之闕如。最後，時間統計方面，大部分相關研究在分析東晉南北朝高僧之地理分佈時，其只取梁高僧傳中東晉南北朝時期的僧人，而忽略其他三部高僧傳中，對於該時期事件的描述。這樣使用僧傳做為朝代統計的區分，將無法看到特定歷史事件與佛教發展之間的關係，並且對某些於後期的僧傳中紀錄的早期僧人事蹟，將可能被忽略。而本系統不僅僅可以像先前的研究一樣進行大範圍的統計，更進一步的是，可以讓使用者將《高僧傳》中所有的地點、人物與時間進行交叉統計。例如梁武帝與佛教發展之間關係密切，我們可以這樣問：「在《高僧傳》中，梁武帝在從天監元年（C.E. 502）稱帝之後到普通元年（C.E. 520）開始篤信佛法之間，曾經在哪裡見過哪些高僧？」也就是利用人名與時間的交叉搜尋，結果如下表 1：

表 1 梁武帝從稱帝到開始篤信佛法之間 (502-520) 和誰去過哪裡？

關係人	地點	時間	傳別	篇名
梁武帝；僧法	華光殿	Not Before：C.E. 502 Not After：C.E. 520	唐	僧伽婆羅傳
梁武帝；智藏	建康	Not Before：C.E. 502 Not After：C.E. 520	唐	釋智藏傳
梁武帝；釋慧超	南潤寺	Not Before：C.E. 502 Not After：C.E. 520	唐	釋慧超傳
梁武帝；僧伽婆羅；寶唱；袁曇允	正觀寺	Not Before：C.E. 502 Not After：C.E. 516	梁	求那毘地傳
梁武帝；釋明徹；僧旻	建康	Not Before：C.E. 502 Not After：C.E. 506	唐	釋明徹傳
梁武帝；寶誌；法雲	華光殿	C.E. 506	梁	釋保誌傳
梁武帝；僧伽婆羅；寶唱；釋慧超； 僧智；法雲；袁曇允	梁壽光殿	C.E. 506	唐	僧伽婆羅傳
梁武帝；僧伽婆羅；寶唱；釋慧超； 僧智；法雲；袁曇允	華林園	C.E. 506	唐	僧伽婆羅傳
梁武帝；僧伽婆羅；寶唱；釋慧超； 僧智；法雲；袁曇允	正觀寺	C.E. 506	唐	僧伽婆羅傳
梁武帝；僧伽婆羅；寶唱；釋慧超；	占雲館	C.E. 506	唐	僧伽婆羅傳

僧智；法雲；袁曇允				
梁武帝；僧伽婆羅；寶唱；釋慧超； 僧智；法雲；袁曇允	臺城	C.E. 506	唐	僧伽婆羅傳
梁武帝；僧旻	建康	C.E. 507	唐	釋僧旻傳
梁武帝；僧旻	定林上寺	C.E. 508	唐	釋僧旻傳
梁武帝；寶亮	建康	C.E. 509	梁	釋寶亮傳
梁武帝；寶誌；吳慶	華光殿	C.E. 514	梁	釋保誌傳
梁武帝；釋明徹	華林園	Not Before : C.E. 515 Not After : C.E. 520	唐	釋明徹傳
梁武帝；慧約	等覺殿	C.E. 519	唐	釋慧約傳

從以上的查詢結果，我們可以更清楚的知道梁武帝的佛法理念，可能受到哪些僧人的影響。由於梁武帝並非高僧，在僧傳中並沒有獨立傳記，要取得有關梁武帝的資料是需要十分大量的整理功夫。如上表所示，這些資料來自於 9 個不同的僧傳，而本系統可以在瞬間回答。另外也可利用本系統進行人名與地名的交叉搜尋，例如玄奘是佛教史上重要的譯經僧人，而長安是唐朝翻譯佛經的重鎮，我們可以這樣問：「在《高僧傳》中，玄奘在何時在長安，並且當時與誰見過面？」，其結果如下表 2 所示。

表 2 玄奘和誰何時在長安？

關係人	地點	時間	傳別	篇名
玄奘；釋宗哲；釋慧沼	長安	Not Before : C.E. 559 Not After : C.E. 656	宋	釋宗哲傳
釋明濬；玄奘；于志寧	西京	Not Before : C.E. 595 Not After : C.E. 686	唐	釋法沖傳
玄奘；釋道因	大慈恩寺	Not Before : C.E. 596 Not After : C.E. 658	宋	釋道因傳
釋道宣；玄奘	西明寺	Not Before : C.E. 605 Not After : C.E. 667	宋	釋道宣傳
玄奘；釋長捷；釋道基	大莊嚴寺	C.E. 617	唐	釋玄奘傳
釋義褒；玄奘	長安	Not Before : C.E. 619 Not After : C.E. 663	唐	釋義褒傳
玄奘；釋道岳；釋法常僧辯；釋玄會； 蕭瑀	西京	C.E. 624	唐	釋玄奘傳

釋慧淨；玄奘；房玄齡	長安	C.E. 628	唐	釋慧淨傳
釋窺基；玄奘；釋圓測	長安	Not Before : C.E. 641 Not After : C.E. 683	宋	釋滿月傳
釋窺基；玄奘；神昉；釋嘉尚；釋普光	長安	Not Before : C.E. 641 Not After : C.E. 683	宋	釋滿月傳
釋慧沼；玄奘；釋窺基；勝莊；大願； 塵外；武平一；盧藏用；陸景初；崔湜； 釋義淨	長安	Not Before : C.E. 641 Not After : C.E. 732	宋	釋慧沼傳
釋彥棕；玄奘	長安	Not Before : C.E. 644 Not After : C.E. 650	宋	釋彥棕傳
玄奘	西京	C.E. 645	唐	釋玄奘傳
玄奘	弘福寺	C.E. 645	唐	釋玄奘傳
釋懷素；玄奘	京兆府	C.E. 645	宋	釋懷素傳
玄奘；慧明；靈闍；行友玄曠；智證； 釋辯機；玄模；玄應	弘福寺	C.E. 645	唐	釋玄奘傳
玄奘；李世民	西京	C.E. 646	唐	釋玄奘傳
釋窺基；玄奘	長安	C.E. 648	宋	釋窺基傳
釋窺基；玄奘	廣福寺	C.E. 648	宋	釋窺基傳
玄奘；李世民	西京	C.E. 648	唐	釋玄奘傳
釋法藏；玄奘	長安	Not Before : C.E. 652 Not After : C.E. 713	宋	釋法藏傳
那提；玄奘	大慈恩寺	C.E. 655	唐	那提傳
玄奘；釋窺基	大慈恩寺	C.E. 656	宋	釋窺基傳
玄奘；李弘；薛元超；李義府	大慈恩寺	C.E. 656	唐	釋玄奘傳
釋法寶；玄奘	長安	Not Before : C.E. 662 Not After : C.E. 753	宋	釋法寶傳
那提；玄奘	大慈恩寺	C.E. 663	唐	那提傳
玄奘；馮義；唐高宗	大明宮	C.E. 663	唐	釋玄奘傳

由於搜尋過程中是採用地理座標的對應查詢，因此在上表的結果中，不僅只有「長安」的結果，也會將玄奘在長安一些寺廟的活動都表列出來，讓學者得到更精細且豐富之答案。此外，這個跨傳搜尋的功能更可以讓我們針對某一地點，自動產生資料豐富的年表資料。例如，我們將查詢條件鎖定於建於南朝宋元徽年中（C.E. 473- C.E. 477）的寶林寺，我們將得到如表五所示的資料。在四本高僧傳中都曾今有

僧人造訪過，而此一資料已儼然是該寺的年譜雛型。

表 5 寶林寺的關係點

關係人	地點	時間	傳別	篇名
慧基	寶林寺	Not Before : C.E. 473 Not After : C.E. 477	梁	釋慧基傳
釋智瓌；慧成	寶林寺	C.E. 586	唐	釋智瓌傳
釋惠能	寶林寺	Not Before : C.E. 674 Not After : C.E. 676	宋	釋惠能傳
釋澄觀；洪霈	寶林寺	C.E. 731	宋	釋澄觀傳
釋希圓	寶林寺	Not Before : C.E. 804 Not After : C.E. 895	宋	釋希圓傳
釋了宣；善榮	寶林寺	Not Before : C.E. 1089 Not After : C.E. 1180	明	釋了宣傳
春谷	寶林寺	Not Before : C.E. 1299 Not After : C.E. 1371	明	釋大同傳
日益	寶林寺	Not Before : C.E. 1299 Not After : C.E. 1371	明	釋大同傳
釋大同	寶林寺	Not Before : C.E. 1335 not After : C.E. 1341	明	釋大同傳

最後，透過 Google Earth 的 GIS 視覺化功能，與語意搜尋功能的搭配，我們可以讓使用者從 GIS 的角度，進行較為宏觀的研究，以便探索佛教發展的過程，例如使用者可以問：「唐朝跟南宋佛教關係點分布的差異為何？」而下面兩張圖分別表示了唐朝（C.E.618 ~ C.E.907）以及南宋⁹（C.E.1127 ~ C.E.1279）的關係點分布圖，我們可以很清楚看到在圖 17(a)中，唐朝佛教重要關係點多半集中於河北到四川一帶，並因為譯經風氣盛行，因此有許多發生在西域及印度的事件。而南宋時，由於政治紛擾的關係，佛教的發展集中東南，而且規模也大為減少。從查詢結果中可以很明顯的觀察出此現象。

以上的查詢中，本系統不但將所有的「關係點」視覺化在 Google Earth 中，並且所有的關係點都可以回溯至文本之中，讓使用者了解其文本出處，不但可以驗證前人的研究成果，也可以當作學者研究成果的佐證。當然。我們更希望讓學者看到已往看不到的研究面向，進而發現出新的研究領域。

⁹ 由於宋高僧傳於宋朝早期便已編撰，因此對於北宋高僧紀錄甚少。而後續明高僧傳，其紀錄內容又以南宋、元、明之高僧活動為主。為避免資料內容誤導，本例中，僅引用南宋時期資料作為比較對象。



圖 17 唐朝與南宋之高僧傳文本關係點分布比較：(a)唐朝 (C.E.618-907) (b)南宋 (C.E.1127-1279)

伍、結論

本研究使用 TEI 標記語言，將高僧傳的內容進行「人名」、「地名」、「時間」與「事件」的語意內容標記，進而產出高品質的《高僧傳》數位文本。此外，本研究進一步探索數位文本應用加值的可能性，因此建立了《高僧傳》時空地理平台，藉由 Google Earth 讓使用者可以利用 GIS 工具看到視覺化後的《高僧傳》，提供嶄新的文獻觀察角度。該平台除提供時空資料視覺化之外，也將《高僧傳》數位文本與法鼓佛教學院規範資料庫連結，提供豐富的參考資料，讓使用者在閱讀時，能更容易的掌握文本的內容。此一時空平台也創造出更符合使用者需求的語意搜尋，協助使用者進行研究。由於我們已經利用 TEI 語言記錄了《高僧傳》中的事件，利用這樣的資料，在未來可以應用更多的資訊，展現《高僧傳》的重要內容，例如利用網路視覺化工具，將僧人的社會網絡 (Wasserman & Faust, 1994) 關係視覺化，或利用時間軸視覺工具，來自動產生中國佛教史的重要事件年表，都是可能的進行方向。

致謝

本研究為浩然基金會所贊助之漢傳佛教高僧傳之時空資訊系統研究計畫之部份成果，本研究團隊特此致謝。

參考文獻

- 佛學規範資料庫 (2008) 上網日期：2010 年 5 月 20 日，檢自：<http://authority.ddbc.edu.tw>
- 佛光山電子大藏經 (1997)。星雲大師略傳。上網日期：2010 年 5 月 20 日，檢自：<http://etext.fgs.org.tw/>
- 佛光大辭典 (2008)。名相：處寂。上網日期：2010 年 5 月 14 日，檢自：
http://www.fgs.org.tw/fgs_book/fgs_drser.aspx

- 法鼓佛教學院數位專案首頁 (2010) 上網日期：2010 年 5 月 20 日，檢自：
http://www.ddbc.edu.tw/digital_archives/projects.html
- 拓展臺灣數位典藏計畫 (2008)。計畫簡介。上網日期：2010 年 6 月 15 日，檢自：
<http://content.ndap.org.tw/index/?cat=6>
- 洪振洲、李志賢 (2009)。法鼓佛教學院權威資料註解服務。2009 電腦與網路科技在教育上的應用研討會
論文集。新竹：中華大學資訊管理學系。
- 維基百科 (2006)。元數據。上網日期：2010 年 6 月 15 日，檢自：<http://zh.wikipedia.org/zh-tw/Metadata>
- 嚴耕望 (2005)。魏晉南北朝佛教地理稿。中央研究院歷史語言研究所專刊之 105 (頁 33-69)。台北市：
中央研究院歷史學研究所。
- Bingenheimer, M., Hung, J. J. & Wiles, S. (2009) *Markup meets GIS - Visualizing the 'Biographies of Eminent
Buddhist Monks'*. Proceedings Information Visualization IV 2009 (pp. 550-554). Washington: IEEE Computer
Society.
- Burnard, L., Sperberg-McQueen, C. M., & Bingenheimer, M. (2009)。TEI使用指南——運用TEI處理中文文獻
(Chinese TEI - A guide to using TEI with Chinese texts) (謝筱琳、黃韋寧、吳恬綾、戎錫琴、劉純佑
譯)。台北市：數位典藏與數位學習國家型科技計畫。
- CBETA 電子佛典集成-CBETA Chinese Electronic Tripitaka Collection (2010)。台北市：中華電子佛典協會。
- CBETA 文獻資料—簡介學術版 XML 標記 (1999) 上網日期：2010 年 5 月 20 日，檢自：
<http://www.cbeta.org/data/cbeta/xmlsample.htm>
- eXist (2010) 上網日期：2010 年 5 月 20 日，檢自：<http://exist.sourceforge.net/>
- Google Earth (2010) 上網日期：2010年5月20日，檢自：<http://earth.google.com/intl/zh-TW/>
- Hung, J. J., Bingenheimer, M. & Wiles, S. (2010) *Digital Texts and GIS: The interrogation and coordinated
visualization of Classical Chinese texts, Proceeding of 2010 International Conference on Computer Design and
Applications (ICCCA 2010) vol. 1(pp.453-458)*, Hebei: Qinhuangdao.
- ISO 8601:2004(E) (2004)。Data elements and interchange formats-Information interchange-Representation of dates and
times。上網日期：2010 年 6 月 15 日，檢自：http://dotat.at/tmp/ISO_8601-2004_E.pdf
- jQuery (2010) 上網日期：2010 年 5 月 20 日，檢自：<http://jquery.com/>
- KML (2010) 上網日期：2010 年 5 月 20 日，檢自：<http://code.google.com/intl/zh-TW/apis/kml/documentation/>
- Wasserman, S. & Faust, K. (1994) *Social Network Analysis: Methods and Applications*, Cambridge: Cambridge
University Press.