

凡例

文本隨著時間的推移而有各種變化。從最基層的單字或詞彙，到句子、句群、段落，再到較大單元的章節、作品與文集，也許是刪減或添加了一個單字、片段或章節，也許是進行了順序的重新排列。

不只是文本本身，連它的書目資訊也是不穩定的。標題、創建日期或著作權的歸屬等，都有被遺忘、偽造、或者再現真相的可能。這些改變，或有意或無意，或大或小，都可見於文本的各類形式和製作中：如語言、圖像、音樂、口述、筆錄、印刷或數位化。文本的變化，起於作者的創造，終於徹底的被遺忘——遺忘了文本曾經存在過。所以，我們所看到的文本上的每一個事例，都不過是一張快照，是文本不斷實體化的長流中的一次定格。本書作為《中華佛寺志叢書》系列中的一冊，也不例外。

依照常規，此系列的所有寺志皆屬文集，意即它們囊括了來自不同作者、時代、體裁的短小文本。它們最初印刷於 16 世紀末到 20 世紀初之間，部份是木刻版印刷，部份是活字版印刷。在 2007 到 2011 年間，經過掃描並數位化，它們變成了 XML / TEI 的格式¹。編輯者添加了各種資訊：較顯著者如文本校勘、文字釐訂、新式標點、註解和編者序言等；較不顯著者如版面的設計和段落的區分等。在數位媒體上，這些編輯修改可以以 TEI 來記錄。TEI 是一種標記規範，表現為 XML 格式，它從 1991 年起，由 TEI 聯盟（<http://www.tei-c.org>）開發、維護。本書則是將我們的 TEI 版本轉檔成 PDF 後付印的。在我們加註的標記中，部分樣式適於付梓發行的，就在書中予以呈現。另一方面，同一個 TEI 版本，還能轉成

1 《中華佛寺志叢書》數位化的底本依據是《中國佛寺史志彙刊》，杜潔祥主編，台北：明文書局（亦名「丹青圖書出版公司」、「宗青圖書出版公司」）印行，1980-1994。

凡例

可在網路平臺上供人閱覽的 HTML 檔，也可做成可供下載的檔案格式。由於兩種格式間是採用程式的自動轉換，所以我們的版本看上去與傳統的古文本有些不同，但期望讀者能夠瞭解我們這個做法的優勢。

以下將詳細描述我們在文本中所添加的資訊。

1. 人名與地名

人名和地名是我們把佛寺志當作歷史資料來研究所遇到的最大障礙之一。很顯然，寺志中含有大量的人名：皇帝和方丈、文人、僧侶和俗眾。這些名字往往在沒有任何背景的情況下出現，我們常常無從得知其生卒年月和生平履歷。如今在數位媒體上，當某某名稱在文本中出現時，我們已能將之連結到有其人或其地背景資訊的資料庫，而順利地閱覽文本。

建立這樣的資料庫實非易事。要記錄一個人的生卒年月，就要精確重建其出生與死亡的日期；然而這些資料若不曾有過明確的記錄與留存時，我們常常只能用一個時段來代替一個單純的日期而進行標記。而這些訊息都必須註明來歷，提供讀者追溯本源的可能性，以方便學術上的使用。

歷史地名的辨識也自有其難處。首先，規範資料庫必須竭盡所能判斷一個地點的地理位置；但要做到這一點，必須考慮到地點可能遊移（相同的地名，不同的位置）、地名可能改變（不同的地名，相同的位置）等等因素。此外，資料來源也經常是指向寬泛的區域，而非單一的地點。在理想的狀態下，規範資料庫對此提供的座標定位，應為面狀而不該是點狀。然而，若是缺乏述及史實疆界的確切資訊，我們就無法證據確鑿地做到這一點。

自 2007 年以來，法鼓佛教學院一直在搭建規範資料庫，諸多佛教史料出現過的相關人物和地點資訊，在資料庫裡不斷擴充、改善。截至 2012 年 8 月，資料庫裡，人物資訊多達 22,000 條，地點

座標已破 55,000 筆。²舉例來說，在網路上，讀者能點擊鏈接到資料庫中相關的條目；然而在印刷紙本上，我們只能編排一個不那麼方便的索引，附在書後，藉以達到近似的功能。這是印刷版本不能避免的弱點。

我們的實體版佛寺志編錄了人名及地名索引，以使讀者能獲得相關的基本資料。這套叢書的索引是由程式自動產生的，展現的是 2013 年初的資料庫狀態。在數位媒體上，可在多部寺志之間搜索字串；而在實體版索引的條目下，我們將書中屬於某條目的所有字串的頁碼，彙集一處，以便查找，希望藉此多少可以彌補在搜尋方面的功能。索引的條目依拼音字母順序編排，所依據的是聯合國授權的漢語拼音的羅馬拼音符號系統。³相較於以部首或筆劃編排的索引，這種方法的優點在於可以更為迅速地查找條目。然而任何以發音為基準的自動化編排，都要面對破音字的問題。有些漢字有兩個以上的讀音，在遇到人名、地名時，這種情況更為常見。從地理和歷史角度註釋條目，是 DDBC 地名規範資料庫所著重的，但不包括讀音。同樣，我們的實體版佛寺志索引只是為了條目的編排而採用了拼音，因此不應將之視為一份地名、人名讀音的標準提供。索引中的排序，所依據的不過是每個條目第一個字最常見的讀音而已。

在文本上的人名和地名，我們都分別標註了不同的底線，這類經過標註的名稱，大多會出現在索引中。然而有些人名、地名，當我們無法確認其身份、地點時，我們會在 TEI 文件中標註以「未知的人名」、「未知的地名」，屬於這類的詞條，便不會在索引中出現。原則上，我們也為人物的尊稱或職稱標註了虛線，方便讀者來閱讀文本。

2 地名規範資料庫內含多量台北中央研究院地理資訊科學研究專題中心所提供的條目。

3 ISO 7098:1982 Revised 1991.

2. 時間

佛寺志中含有大量的時間。正如您所料，這些時間都是使用中國傳統曆法的天干地支來紀年、紀日的。書中的年份通常是冠以帝號與（或）年號。然而，傳統紀年法有許多缺點。很少有讀者能記得住所有的年號，要是不參考對照表，往往很難判斷不同文章裡的兩起事件孰先孰後。所以，在古代文獻的現代版本之中，傳統曆法有需要對應、轉換到一個國際標準。為了達到這一目的，法鼓佛教學院創建了一個時間規範資料庫，將東亞地區各種曆法，對應、轉換到一種通用的標準——即是已被廣泛應用於天文學的「儒略日」（Julian Day Number）。⁴根據「儒略日」，就能輕而易舉地在不同的曆法間進行時間轉換。在我們的這一實體版寺志中，所有的日期與年代（朝代除外），都標上了其在「外推格里曆法」（Proleptic Gregorian calendar）裡的相應日期。如此，便有助於了解事件相應的時間點，並在時間軸上找到定位。

傳統的紀年法往往讓人很難明白兩起不同事件之間相隔有多久。例如，有篇文章先是提及一樁孝宗隆興元年七月間發生的事件，然後再提到另一件發生於寧宗嘉定間的事。⁵或許根據文字的出現順序，又或許讀者能背得出宋代的年號，那便可以推定第一樁事是發生在第二樁事之前。然而，我們大多數人並非宋代史專家，實難斷言後一事件是在前一事件的十年、五十年或者一百年之後發生的。若將各事件對應、轉換至通用標準（宋孝宗隆興元年七月＝1163.8.9-1163.9.6；宋寧宗嘉定間＝1208.1.26-1224.9.23），便能幫助我們理解文本所說的事件彼此相隔的時間長短，也使我們更容易去比較不同文本中的時間記錄。

此外，原書中的日期常常被含糊地以某年代之「初」或「末」來記錄（如天寶初年或萬曆末年）。凡此情形，我們皆以該年號的

4 參考 The XXIIIrd International Astronomical Union General Assembly: Resolution B1 "On the Use of Julian Dates".

5 ZFSH 9, p.177.

前 25% 或後 25% 為一個時段，來標註其日期。在正文中的所有時間都有標記，括號中都提供了相對應的「外推格里曆法」日期（朝代除外）。朝代的年份請參閱凡例後面我們所附錄的〈中國朝代年表〉。

3. 強化文本

編輯，就是往文本中添加資訊的一個過程，藉此使讀者能夠對作品有更深刻的認識。除了上述提到的人名、地名、時間方面的資訊，我們還增加了如下四種訊息：

3.1 章節／文本標題：部分佛寺志有時收錄了沒有標題的內文。為了能較有效率地檢索到這些文章（例如從目錄中），我們為它們添加了簡要的描述性標題。在實體印刷版本裡，這些添加的標題都標上了角括號[]。

3.2 註腳：為了指出並解釋文中的問題，我們添加了註腳。但由於一部寺志收錄的文本可能來自不同的時期，而且體裁繁多，以本專案的規模，不可能加以全面性的註解，所以我們把註解的數量維持在一個較低的限度。

3.3 填補佚失的文字或段落：處理寺志木刻本或手稿抄本時，不時會發現一個字甚至一個段落缺漏，或者是難以辨識。我們儘可能地借助其他版本，嘗試填補。經填補的文字（都標上了雙角括號[[]]）或段落，全都清楚地在註腳中註明。

3.4 標點：標點最終成了本專案裡最費人力與時間的部份。我們為佛寺志提供了新式標點，這可能是這個版本在未來最易引起爭議之處。古文句讀是一門艱深的技藝，它無情地揭露了一個人的領悟力的高低。相對於新式標點符號的提供，本專案的作業規模顯得更為龐大也更具野心。我們試圖標點的文本，出自不同作者、不同

凡例

時代、不同體裁。雖然古文就某種意義而言是有標點的，⁶但它並沒有一個公認的標點符號標準。這主要是因為古文缺乏全面、公認的描述性文法，因而其句法成分常有歧義。古梵文曾有波你尼（Pāṇini）和波顛闍利（Patañjali）明文描述成套規則，但古代中文的詞法和句法從未以同樣的方式標準化過。

就算擁有這樣的文法，我們也要記得，標點符號並非語言本身的一部份，而是一種習慣、共識，它往往接受政治或文化機構的規範。要選擇什麼標點符號，應標點到什麼程度，這是一種主觀任意的決定。例如句號「。」和分號「；」之間的區別，常常只是詮釋之不同而已。在現代中文的標點符號中，逗點通常被分有兩種類型，用於列舉的「頓號」和用於區隔子句的「逗號」，而歐洲語系文本的標點符號，則未做這樣的區分。

簡言之，任何一種書面語言都不存在著「理所當然」的標點方式，最起碼古代中文是如此。我們試圖標上新式標點，盼能輔助二十一世紀初的中文讀者來閱讀文本。我們已逐冊審慎檢查我們新增的標點，不過，畢竟完成這項工作的人手陣容龐大，加上時時會碰到艱深難解的文本，所以成品的質量難免不齊。學者專家們也許會不認同某些標點，或者多多少少會發現一些錯誤，我們必當虛心承教。然而我們的標點符號相當有效地詮釋了文本、排除了句法上的歧義，希望這一點成就堪以彌補其不足之處。

4. 特異字處理

如同其他書寫系統所用的符號一樣，中文的字形寫法也不固定。在手抄本裡，這種差異分外明顯，同一個符號從來沒有一成不變的。木刻版則介乎手抄本與活字印刷版之間，它和活字印刷一樣，傾向於規範字形的書寫系統，期達成一種統一性。但是在製作木刻版的過程中，若是直接取用手寫稿，則木刻版仍會存留書法所

6 在中國古代，許多讀者在首次閱讀他們的書籍時會做句讀，即用單純的「圈」或「點」符號對接連不分的整串字進行斷句。

具有的一些彈性。有時在木刻版留存書法，是出於一種敬意。以佛寺志的情況為例，當中的序、跋往往以作者的書法原貌付梓。即使在非關書法的時候，相較於活字印刷，木刻版依舊使用了較多的特異字。多數時候這是出於審美的考量。而這些特質，在後代印本裡通常是被悄悄地抹去了的。在中國，佛寺志的新編版本甚至廣泛地採用了簡體字，因而與文本的本來面目相去更遠。我們曾試著取其中道，然而正如標點符號的處境一樣，某些讀者可能不會同意這樣的策略。我們的大方針是忠於原著，另一方面則採取下述做法，希望給予當代讀者一些協助。

數位化版本唯一適合的字元集是「標準萬國碼」（Unicode Standard）。目前，每個以長久保存、廣泛流傳為目標的重要數位計畫，都必須在標準萬國碼所限制的可編碼字元的範圍內作業。

於是遇到特異字時，首發的問題必定會是：這特異字是否為標準萬國碼的編碼字之一？若是，則我們的編輯方式會因以下三種情況而有不同：

a) 特異字本身清晰易辨時，則不需加以處理；比如常見的「德」字，在文本中時而寫成「徳」，不難辨識。不過中日韓漢字字元集所有的兼容異體字（標準萬國碼的編號 F900-FAD9 以及 2F800-2FA1D），如今都已經規格化了。⁷例如：

益（U+FA17）→ 益（U+76CA）；

福（U+FA1B）→ 福（U+798F）；

僧（U+FA31）→ 僧（U+50E7）。⁸

7 可擴展標記語言（XML）1.0（第五版）W3C 建議 2008 年 11 月 26 日
 (<http://www.w3.org/TR/2008/REC-xml-20081126/>) 2.2 : Ch. 2.2 : “Document authors are encouraged to avoid “compatibility characters”, as defined in section 2.3 of [Unicode].”

8 更多相關細節，請參閱網頁上的專案工作記錄：
http://wiki.ddbc.edu.tw/pages/佛寺志專案_特字處理

b) 在字形上，當代讀者已難以辨識的異體字，例如「𠂔」（即「去」字）、「灑」（即「法」字），諸如此類，我們以各部寺志為單位，蒐集、整理成一篇對照表，名為〈異體字與通用字對照表〉，附在每一部寺志之後，供讀者參攷。例如「𠂔」字不易辨識，〈對照表〉上就可查到「𠂔→去」這樣的資料；「灑」字不易辨識，〈對照表〉上就可查到「灑→法」這樣的資料。當然，辨識的難易度並沒有客觀的標準，所以〈對照表〉的功能也是有限的。

c) 第三種情況較為罕見：當一個普通字被用作另一字符的意義時，這類情況極可能使當代讀者身陷五里霧中。例如「懸」字在木刻版中有時被用作「懸」的意思，此時我們就在註腳裡提供「懸」字給讀者做參考。

此外，當某一特異字沒有標準萬國碼編碼，而我們還是可以確認它的通用字時⁹，在文本中就會直接採用其通用字來表現。只有在極少的情況下，有些特異字，我們無法指認其所屬的通用字為何；或者是因為避諱的關係而形成的異體字，我們無法使用其通用字，此時唯有以 SVG 造字來如實呈現之¹⁰。

為人名、地名、時間補充資訊，為文本加上註解與標點，為特異字提供通用字。我們一步一腳印，細心地經營，只望此版本的內容、品質，比諸過去已問世的各種印刷本，更完整、更可信，也更便於閱讀。

本版所據之底本是《中國佛寺史志彙刊》，為了保持前後版本之間的關聯、提供讀者以溯源查索之便利，我們將底本的頁碼，隨文本之先後，逐一標示在本版之頁側。既不妨礙文本之閱覽，而有查索參訪之方便。

9 這方面的作業，部份的參考是依據中華民國教育部國語推行委員會編纂的網路版異體字字典 (<http://dict.variants.moe.edu.tw/>，2012年8月)。

10 避諱字造字實例參見《寒山寺志》底本頁 210。