



佛教禪修的對話與交流

◎陳陶

2021年10月16至17日佛教學系主辦「對話與交流——佛教禪修研究工作坊」，於法鼓德貴學苑展開，因應疫情，同時開放線上參與，共有近百人參加。本著跨域與對話的精神，此工作坊規劃「禪修文獻與佛教傳統」、「禪修與現代社會」與「禪修與宗教經驗」等3項與佛教禪修相關的研究面向進行交流。

在禪修文獻與佛教傳統場次，莊國彬、廖本聖、見弘法師3位老師分別介紹印藏禪修文獻如《瑜伽經》、《俱舍論》、《修習次第》之禪修傳統；林佩瑩、鄧偉仁、廖肇亨、謝世維4位老師則介紹漢傳禪修文獻，如南北朝時所譯的《禪祕要法經》、《治禪病要法經》、唐代禪修釋疑的《修禪要訣》、明清禪修文獻

以及道教內丹文獻與佛教禪修之關聯。議題討論中，多位學者就不同佛教傳統之禪修文獻所呈現之禪修的異同等議題進行深入的交流。

禪修與現代社會場次中，張嘉如、溫宗堃、林建德、謝世維4位老師分別討論了禪修在當代美國的發展，馬哈西傳法的現代性，禪修與心識研究的局限及貢獻，當代西方不二論與禪宗禪修的發展等議題。在禪修與宗教經驗場次，蔡怡佳、李維倫、劉怡寧3位老師分別從心理學、現象學、社會學等視角分析佛教禪修經驗；本校2位博士生林悟石與智諦法師分別報告了Robert Sharf關於宗教經驗論述的難題以及禪修經驗與維特根斯坦所提出的私人語言之關係。

佛教學系主任鄧偉仁老師



▲ 2021年10月16至17日佛教學系舉辦佛教禪修研究工作坊，照片由法鼓山文教基金會提供。

在採訪中表示，他希望這次的工作坊可以透過這3個領域，提供禪修者一個跨佛教傳統與跨領域的交流與對話平台。一方面提升佛教禪修研究的深廣度，另一方面探索佛教禪修在現代社會推動中的多重意義、局限與挑戰。

2021圖書館週 以饗讀者

◎釋天佑

2021年11月29日至12月3日本校圖書資訊館舉辦第19屆圖書館週系列活動。首先登場的是10月份暖身活動——圖資館Logo設計比賽與藏書票設計比賽，經全校師生網路投票後，分別選出前3名作品，未來將實際運用在圖書館的海報、文創商品與圖書館周邊小物中。

11月29日開幕茶會由館長洪政洲老師致詞與頒獎，並備有摸彩活動，希望大家踴躍的參與。11月30日Fun電影活動，觀賞奧斯卡2021金獎影片「游牧人生」，描繪美國在金融風暴衝擊、工廠紛紛關閉後，發展出以車為家的新型態社群。影片結束後由心靈環保中心主任果光法師與法鼓山僧眾服務院常捷法師帶領觀眾分享心得、互動與交流，從中學習更

寬闊的見解。12月1日冬季五分鐘書評活動共有9位同學參賽，由潤迦法師介紹的「Love for Imperfect Things」票選為最有趣閱讀的一本書。惠敏校長也與現場大眾分享，書評活動是為了鼓勵學生分享讀書的樂趣，學習簡潔有力的表達。同日下午為中西參（餐）找碴比賽活動，現場學生參與踴躍，有的同學獨自作答、有的合作答題。然而，此活動的目的是為了讓同學們更了解圖書館的館藏與資訊的運用，並善用於課業及解行之中。

圖書資訊館於2002年正式啟用，每年都會如期舉辦圖書館週系列活動，讓全校師生與教職員共同歡慶參與，期待今年慶祝20週年的系列活動，再次共襄盛舉。

▼ 中西參（餐）找碴比賽得獎同學。

終身學習工作坊 再訪台中寶雲寺

◎顏美惠

2021年12月25日本校連續第二年造訪法鼓山台中分院寶雲寺，舉辦「終身學習菩薩行」工作坊，共有超過70位學員參與。首先由校史館主任辜琮瑜老師介紹大學院教育，由終身學習的脈絡梳理、深化學習至完整觀點的系統化規劃。學群長陳定銘老師則由心靈環保管理學的角度切入，從善良比聰明更難得延伸至社會企業，並引用《基業長青：高瞻遠矚企業的永續之道》一書，試算出每投入1美元的效益，經過多年竟可出現6.65倍的巨大差距，證明有理念的願景也能同時為企業帶來高額效益，推動社會企業絕非空中樓閣的空談。

此外，社會企業與創新學程主任葉玲玲老師也分享幾個成功的社會企業，如甘樂文創、活水影響力投資等，並闡述自己在任教期間不斷考驗慈悲與智慧，與學生互



▲ 2021年12月25日舉辦終身學習菩薩行工作坊，共有超過70位學員參與。

動的小故事。黃信勳老師則以自身的專業背景，分析近百年來屢創新高的二氧化碳濃度所帶來的問題，同時也回應現場學員的提問，即便人類可留下污染的地球而遷移至其他星球，在倫理的角度下也有不合適的原因與論點。下午，僧伽大學副院長常寬法師引導全體學員進行禪修放鬆的體驗，其中有學員分享，一直以為自己是一

個靜不下來的人，這是第一次感受到放鬆安定的感覺。

穿插於教師群專業短講間的幾次小組討論，氣氛非常熱烈，一掃上午陰冷天氣帶來的陌生。許多學員不斷提出關於進入文理學院「都在學些什麼？」以及關心「畢業的條件」等問題，帶來許多笑語。最後，本工作坊在窗外雲開天見晴，大家收穫滿行囊，順利圓滿。



校長 Tea Time
惠敏法師

「佛典戲劇編導演」課程緣起

2021年4月8日法鼓文理學院校慶時，敝人預告明年是15週年校慶，將由前台北藝術大學戲劇學系汪其楣教授帶領本校「佛典戲劇編導演：生命教育之舞台實踐」課程的學生表演《悠悠鹿鳴》的佛教本生故事。

原先，我是電郵請問汪教授是否可邀請當初演出的「拈花微笑劇團或其他組合」再來表演。汪教授回覆：「何必召集外面的演員來為法鼓作戲，何不就到貴校去帶一個戲劇製

作，參與者就是佛學系的學生，經由劇本策劃，表演基礎，到排練製作的方法步驟，上台親自體驗，現身說法。……」

汪老師構想：「這門戲劇課應該是一學年。……經過一個學期的研習與實作課程，期末呈現就已見雛形。寒假期間有機會充實必須外求配合的製作，如服裝道具配樂等。明年下學期開

學，就是更多的排練、彩排，4月8日校慶上演，也可聯繫安排到附近中國國小去輕簡巡迴幾場，如此劇場經驗更紮實完整！」

我回覆：「的確是很有意義、且有創意的課程與表演，當然敝人也非常佩服您的智慧與勇氣（面對如此大的挑戰：以「表演素人」為對象），而且選課學生不知是否可以達到表演人

數？」汪教授答：「素人可塑性較高，流露本色時，比匠人更吸引人。課程也包括肢體和聲音口白的訓練，感情與連結的創造練習活動。要把課程設計得吸引多一點學生來上課，培養各種人才……」。

於是，經過汪教授、學校同仁、同學的一起努力，開課成功，我也托大家的福，參加這門課，於《人生》雜誌（2022年2月）刊登拙文「佛典戲劇編導演」角色體驗，歡迎大家指教。



學術
動態

我的學思及專業服務歷程

曾漢珍／社會企業與創新碩士學位學程副教授暨總務長

成大建築的教育，培養了我務實的专业取向，臺大城鄉所多元批判的跨領域學習，讓解決問題的思維更靈活。博士論文以學校建築為研究專攻，期間師大教研所黃政傑教授（前臺南大學校長）無私義務的指導，讓我真正進入教育學跨領域學習與實踐。

歷經南亞工專及中國工商專校建築科，再到中原大學室內設計系任教。1999年到國立臺北大學籌備處擔任建築顧問，2000年到臺北大學服務並任教於民俗藝術與文化資產研究所。兼任行政工作讓我從學校

建築研究者轉換為新校園建設的實踐者，必須面對更真實的挑戰。在臺北大學前後擔任三位校長的總務長，經常要處理校園建設理論與實務的衝突，挑戰很多成長更多，我認為行政大樓是三峽校區結合理念與實務最傑出的建設成果。果燦法師搭起了我與法鼓山的因緣，擔任建築顧問後於2015年接受惠敏校長之邀，來本校任教並擔任總務長迄今。

1998年執行政治大學委託之政大附中校園規劃，完成《國立政治大學附屬高級中學校園整體規劃研究》，作為新校園

設計依據。1999年完成《台北縣文化藝術長期發展計畫》，從空間角度思索文化藝術的發展策略。擔任台北縣、新北市都市設計審議委員近二十年的實戰經驗，培養了都市設計與建築間整合的實務能力。

殯葬領域方面分別擔任臺北市殯葬諮詢委員、新北市殯葬設施審議委員，以及內政部地方政府殯葬管理業務績效評核委員，走遍全國各地有代表性之殯葬設施，尤其殯儀館與火化場，前後近二十年，對生命教育有更深的體悟。

九次參與海峽兩岸大學的校

園學術研討會，並提出服務學校之學校建築論文，切磋與交流校園建設議題。在法鼓文理學院服務期間發表（2015.10）〈聖嚴法師及其大學校園空間實踐初探：以法鼓文理學院為例〉，《第十五屆海峽兩岸大學的校園學術研討會》、（2019.6）〈聖嚴法師法鼓山世界佛教教育園區之大學校園空間實踐初探〉，《佛法與社會科學國際研討會》。2022

年3月將於《台灣教育研究期刊》發表學術論文〈佛教團體興辦大學之校園建設課題探討〉。

2022年2月1日我將屆齡退休，目前擔任行政院客委會客庄369專案的空間總顧問，可深入客家鄉鎮服務鄉親。希望多參與社會服務關懷弱勢，走入地方社區將是未來的生活重心。一切因緣合和，我將帶著全新的心情開創生命的因緣。

校譯新作 關於「定」的考察

廖本聖／佛教學系副教授暨語言與翻譯中心主任

野川博之博士中譯，而由我較譯的《世親《阿毘達磨俱舍論》及稱友《明瞭義疏》新譯：〈闡明等至品第八〉》（高雄：彌勒講堂，2021）已於今年八月出版。本書譯自櫻部建、小谷信千代及本庄良文日譯的《俱舍論的原典研究 智品·定品》其中的〈定品〉部分。

〈闡明等至品第八〉主要分為三個章節：〈一、成為功德所依的各種三昧〉、〈二、以三昧為所依的各種功德〉及〈三、結語〉。

簡單來說，定（等至）分為「有心定」與「無心定」兩大類。無心定有無想（*asamjñin*）與滅盡（*nirodha*）二定，因為屬於不相應行，所以在〈分別根品第二〉當中解釋。而〈闡明

等至品第八〉主要考察有心定，有心定又分為十七定，即八根本（*maula*）定、八近分（*sāmantaka*）定與一個中間（*antara*）定。

八近分定中，唯有初靜慮近分定可被稱為「未至定」，因為只有無漏的初靜慮近分定能斷除所有三界的煩惱，所以此定又被稱為「無所不能」

（*anāgāmya, mi lcogs pa med pa*）。

此外，也僅有初靜慮才有中間定，因為此定是由尋、伺二心所的有無而區分出來的，亦即中間定是「無尋唯伺」，而初靜慮根本定（含）以下皆為「有尋有伺」，第二靜慮近分定（含）以上皆為「無尋無伺」。



▲廖本聖老師專長為藏文文法、菩提道次第教法、格魯派三大寺六類基礎教科書等。

與定相關 梵文術語	漢譯		藏譯	定地				非定地 欲界
	意譯	音譯		有心定		無心定		
				色界定	無色界定	色界 無想定	無色界 滅盡定	
<i>samāpatti</i>	等至	三摩鉢底	<i>snnyoms par 'jug pa/ snnyoms 'jug</i>	✓	✓	✓	✓	
<i>samāhita</i>	等引	三摩呬多	<i>mnyam par gzhag pa/mnyam gzhag</i>	✓	✓	✓	✓	
<i>samādhi</i>	等持	三摩地，三昧	<i>ting nge 'dzin</i>	✓	✓			✓
<i>dhyāna</i>	靜慮	禪那，禪	<i>bsam gtan</i>	✓				
<i>śamatha</i>	止	奢摩他	<i>zhi gnas</i>	✓	✓			
<i>vipaśyanā</i>	觀	毘鉢舍那	<i>lhag mthong</i>	✓	✓			



▲曾漢珍老師於本月榮退，1月17日學校特舉辦茶會歡送。

本書還提到許多與「定」相關的術語，茲列表說明如下：

1. *samādhi* 跨有心定與非定地，但不適用於無心定，因為 *samādhi* 是心所。
2. *samāpatti* 與 *samāhita* 包含有心定與無心定。
3. *dhyāna* 僅限於色界的有心定。唯有淨（*śuddhaka*）與無漏（*anāsrava*）的根本靜

慮才有「靜慮支」或「禪支」（*dhyānāṅga*），四靜慮根本定總共有十八靜慮支。

4. *śamatha* 與 *vipaśyanā* 涵攝一切有心定。

除了 *samādhi* 未必是定之外，其餘五個必然是「定」，而這五個中，也只有 *dhyāna* 才可以稱為「禪定」。

學生會自製頻道—轉心轉世界

以PODCAST述說校園故事



◎學生會

轉心轉世界——讓你、我聽見法鼓文理學院，紀錄校園的生活點滴與學術交流，希望聽眾在其中挖掘心靈環保的寶藏，轉個心念，心讓世界大不同。

2020年開始至今，全球面對疫情的影響，舉辦活動受到限制，甚至面臨停辦的難題。學生會籌辦Podcast——轉心轉世界預錄型廣播節目，在線上與同學連結，拉近校園師生間的感情，傳遞學校資訊，讓更多人認識法鼓文理學院，發現學校的美好。特別感謝生命碩二的劉澧枋同學，為轉心轉世界製作開場音樂，非常有校園特色，讓人耳目一新。

目前進行的節目單元，分別為：「校園會客室」、「法鼓

文理學報」，以及「在山上的每一天」等3個內容。第一個單元「校園會客室」，每集帶大家認識一位法鼓文理學院的朋友，對象包括教職員生、義工等，分享生命故事、個人工作與生活經驗。首集由學生會李育彰會長與黃夙慧副會長上陣，錄製「跨世代跨領域服務因緣」，分享這段精彩的學生會選舉過程，並記錄任職的初心。

第2單元「法鼓文理學報」則以課業學習為主題，邀請同學來分享學習心得，讓我們的作業更有趣，研究更投入；也規劃邀請老師分享研究成果，有更多元的展現。首集中訪問佛學系博士班林悟石同學，侃侃而談自學士班到博士班在校

多年的體驗與學習歷程，以及對學術研究投入的熱情。

第3單元「在山上的每一天」以紀錄學校生活為主，讓DILA校園的日常生活、心情點滴等，都能在此抒發，分享關懷與祝福。首集造訪法鼓悅音社的聲樂課，本學期悅音社邀請法鼓合唱團創始人李俊賢老師與著名女高音賴珣好老師蒞臨指導聲樂課。在訪談中李俊賢老師提到認識自己聲音的重要性，音樂活動更能提昇人的品質。

我們期許用轉心轉世界，寫下法鼓文理學院的故事，一步一腳印地探索校園人文之美。邀請有興趣的同學一起參與，期待下一位受訪者和下一個美好的故事。



▲本月「校園會客室」單元由學生會會長李育彰（右）與前媒體人夏慧麟同學（左）擔任主持人，特別來賓是大家長惠敏校長（中）。

校友園地

博班心得：轉換思考與分享、發表

◎張雅雯

2021年2月初領到博士畢業證書，曾於同年4月《法鼓雜誌》（<https://www.ddm.org.tw/maze/pdf/202104vol376.pdf>）表達畢業感言與對惠敏校長、果鏡法師等師長之謝忱，今天藉此園地想與同學們分享求學心得與近況。

學習過程中的助力，包括親切引導的師長、藏書豐富的圖書資訊館等是大家耳熟能詳的。此外，我想特別提兩件事：一是論文與行門的交互運作，二是積極參與研討會與論文投稿。有關學術與實踐雙軌並進的好處，除了理論與實踐的交互增上外，當論文撰寫遭遇撞牆期、瓶頸時，我會先跳脫開，轉成設計企業營、青年營等生活禪修課程的教案，如此的場域轉換可以鬆綁被既有思考框架卡緊的腦袋，帶來新的思考路徑。

至於第二點，我對有興趣

的題目會先試著在研討會上發表，聆聽與談學者、老師的評論，然後再修正寫成正式論文投稿期刊，這樣一來，一個論題至少得到了三、四位專家的修正建議，幫助很大。例如：中華佛學研究所主辦的「漢傳佛教學者青年論壇國際研討會」，佛光大學的「近世東亞佛教的文獻和研究國際學術研討會」，聖基會的「漢傳佛教與聖嚴思想國際學術研討會」，這三者都提供了博士生、青年學者發表的機會。除可以聽取前輩的講評建議，與不同研究領域的學友交流意見，也都是很好的學習機會；畢竟，一個人的眼光、視角總有侷限，透過不同專家的觀點，可以協助覺察自己研究上的盲點，我會仔細觀察自己為什麼沒想到專家學者所提面向的原因，並藉此逐一校正自己研究過程上的缺失。

有關近況分享：我於2021年3月開始任職中華佛學研究所博士後研究，同年5、6、9月分別於靈隱寺、佛光大學、聖基會舉辦的線上研討會以及「獨坐大雄峰——禪宗思想與寺院文化研習營」提出發表，主要就三峰派祖師尚未有人研究的文獻進行分析，以及對清初禪宗思想史進行梳理。此外，並以「溝通·禪」為介面，著手設計適合現代人的線上應用課程。個人心得分享總屬有限，僅推薦本校第一位博士李周淵（現為廣東中山大學博士後研究），所編「佛研青年」已累積六十多位佛學青年的研究心得，很值得閱讀。（https://mp.weixin.qq.com/mp/appmsgalbum?__biz=MzI4ODU3OTgwMg==&action=getalbum&album_id=1377585059678797824&scene=173#wechat_redirect）

訪學心得

文本與田野的相遇

◎馬浪吉拉雅山

從未造訪韓國這片土地，卻在走入法鼓文理學院的修習中，開啟了這趟殊勝因緣的旅程。

東國大學是韓國佛教最大宗派——曹溪宗所創辦的現代大學，毋庸置疑其館藏大量漢、韓、英版本的佛教各類書籍，對於梳理（漢傳）佛教於韓國思想史上的傳承及開展脈絡幫助很大；加上學校本身是綜合型大學，可供閱覽的跨領域文本資料更拓及至世界哲學與宗教系統。有趣地，文本間交互理解的因緣，看見了佛教學自身與其他知識系統在現象分析下的關係，更是此次交流訪學的意外收穫。

觀音禪林（Kwan Um School of Zen）及淨土會（Jungto Society）是在韓國訪學期間參與共修的兩個當代佛教團體。觀音禪林被視為全球最大的禪宗團體，一方面依循參話禪的傳統，藉著韓國佛教結夏安居及結冬安居各三個月的閉關共修，體現頓修頓悟的禪法修習；另一方面，採用英語弘法的方便因緣，有為數不少的非韓裔（歐美人士為主）信眾親近道場，進而成為法師與禪師，跨越文化傳承法脈。相對地，淨土會為法輪法師（Pomnyun Sunim）領導的現代佛教團體，積



▲鬧中取靜的東國大學佛堂，是筆者平日禮佛及休憩的最佳去處。

極投入韓國海內外的社會救助，在不同宗派、文化間進行交流；疫情期間，為避免大眾的身心失衡，除了提供全球信眾線上佛法開示及共修外，也利用APP禪修與拜懺。

在韓國的日常生活，對我來說就如同一個大型的修行道場，參與並觀察校園中師長、同儕間的利行、同事；或在公共場合，如餐廳、商店及公車上的布施與愛語。此外，也透過哲學咖啡館模式的團體互動，與共同生活在首爾本地的知識份子，如創業家、牧師、大學生等旅韓外籍人士一同對話與交流。

這趟韓國訪學之旅，如何詮釋文本與田野的相互關係？巧妙地，回應了讀萬卷書與行萬里路之間，解、行不二法門的想像。

111學年度單獨招生入學



佛教學系

博士班 4名 (佛教組3名、佛教與社會實踐組1名)
碩士班 13名 (佛教組10名、佛寶組3名)
學士班 14名 (加原住民考生組1名)

人文社會學群

生命教育碩士學位學程 22名
社會企業與創新碩士學位學程 14名
(社企組、社造組、環境組)

法鼓文理學院
DILA Dharma Drum Institute of Liberal Arts



4 Finding the Word Boundaries

Chinese Buddhist Scripture Word Segmentation Project

©Wang Yuchun & Huang Shuling

The present age is the era of big data and artificial intelligence, and the key to the information to be excavated in big data resources in Chinese will be word segmentation. Why is this the case? In the Natural Language Processing in Chinese, the word (詞) is the smallest independent element. However, Chinese writing is in units of characters (字), and characters are unlike words in English where there is a clear distinction between words and between word segments, and this often causes ambiguity. Where native speakers are concerned, detecting the segmentation of the words seems to be intuitive and easy, but in fact, it is not as simple as one might imagine, for instance, "second-hand vehicle" (二手機車), aside from the "second-hand / vehicle", can also be broken down as "two / mobile telephone / vehicle". If we do not have a way to process Chinese words correctly, to determine the segmentation of the words, the incorrect word segments significantly affect subsequent calculation processing, such as information extraction, sentiment analysis, or machine translation. Therefore, in terms of Chinese language processing, the importance of word segmentation is difficult to overstate, and we need to correctly distinguish between the segments of the words formed by Chinese characters.

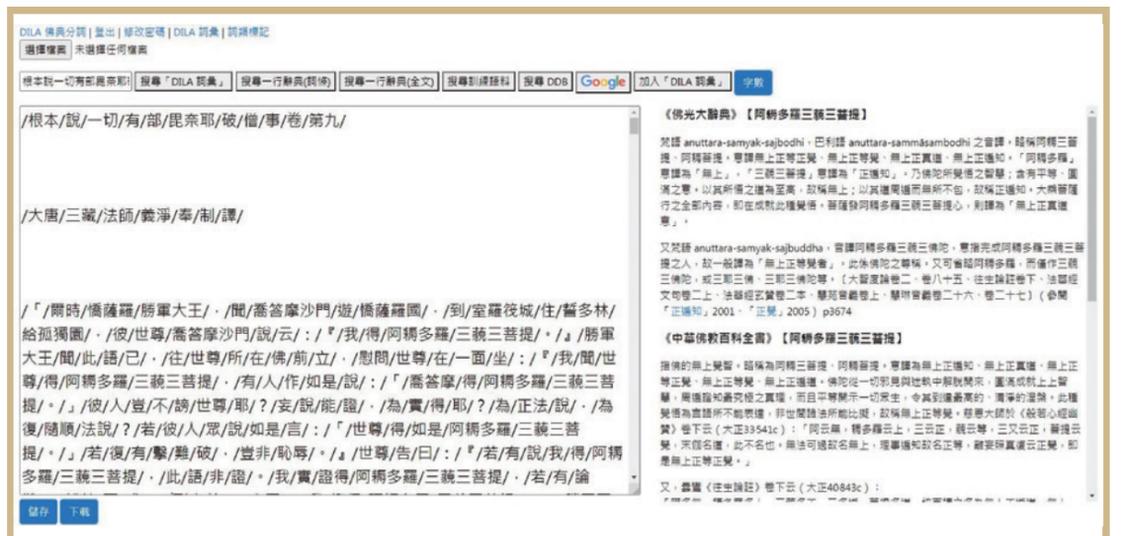
Background

Chinese word segmentation is an important task for Chinese language processing. In the past decade or so, many research teams have developed various Chinese word segmentation methods and systems for modern Chinese, many of which are available for free, such as the Jieba (結巴) system developed in China and the CkipTagger of the Academia Sinica. However, in the processing of traditional Chinese, an effective and accurate word segmentation tool is still lacking. In terms of Chinese Buddhist scriptures, the vocabulary is quite different from general Chinese documents. The lack of appropriate word segmentation tools makes processing and researching Chinese Buddhist scriptures rather limited. In the past, CBETA have invested considerable efforts to produce a complete and high-quality electronic Buddhist scripture integration. In order to allow CBETA data to be used more widely, and to further provide word-segmented texts, this project will begin to plan the word segmentation work of the Buddhist scripture texts in 2018.

In order to construct an automatic word segmentation model using machine learning methods, we need a sufficient amount of manually labeled segmentation data. However, before manual labelling, we must first formulate the criteria for word segmentation to determine what a word is. In terms of modern Chinese word segmentation standards, the Academia Sinica Chinese Knowledge and Information Processing group (CKIP Sinica) established the "Chinese Word Segmentation Processing Principles" (CNS14366) national standard in 1999, which is currently the most widely adopted standard in automatic word segmentation systems in Taiwan. However, the foregoing word segmentation labels are all designed for modern Chinese, and modern Chinese is different to ancient grammar, and is even more different to Buddhist grammar. For example, in terms of vocabulary, two-syllable words are commonly used in modern times, while monosyllable words are mainly used in ancient times; ancient nouns are often used as verbs, and the lexical ambiguity is much higher than that of modern Chinese. In terms of syntax, there are also many differences in word order, and the Buddhist scriptures contain many compound words that are rare in ancient Chinese. The transliterated words and technical terms of the Buddhist scriptures are also very different from traditional terms. These differences in vocabulary and syntax make it impossible for us to directly adopt the word segmentation standards and tools of modern Chinese to carry out the word segmentation of the Chinese Buddhist scriptures. Therefore, in order to apply CBETA data, we must redefine the principles of word segmentation in Buddhist scriptures, and build a sufficient training corpus by manual labelling, in order to gradually complete our goal of constructing an automatic word segmentation system suitable for Chinese Buddhist scriptures.

Phases and Tools

Generally speaking, word segmentation must meet the following two conditions: (1) the word segment has the correct semantics in the context, and (2) the word segment forms a unit with independent semantics. In order to improve the consistency and efficiency of word segmentation, we must first establish a set of word segmentation principles and practical guidelines, and design



▲ Figure 1: Artificial word segmentation interface

a word segmentation tool that is convenient to query and which can improve consistency.

With respect to the word segmentation principle, we refer to the CNS14366 standard, and gradually established the "DILA Buddhist Scriptures Segmentation Principle" from discussion of its implementation. The content includes: "The semantic meaning cannot be combined by the direct addition of the combined components; the part of speech cannot be combined from the combined components; merging of transliterated words; merging of terms included in dictionaries; segmentation of those words that contain interpolated components; segmentation of complex words with more than three characters, etc." A total of 14 basic and auxiliary principles. All personnel involved in word segmentation must be educated and trained to be familiar with these rules; after actual word segmentation, the results are also reviewed with each other in order to achieve a stable and consistent word segmentation.

For this project, we have developed a simple and easy-to-use web page tagging system (see Figure 1) for efficient manual word segmentation and tagging. In the system, word segmenters can consult various Buddhist dictionaries, such as the *Yi Xing Dictionary of Buddhist Studies*, the *Ministry of Education Mandarin Dictionary*, and our customized *DILA Glossary*. In addition, one can also query the split and combined records of a word from the completed training corpus, which will help to improve the speed and consistency of marking up.

After collecting a sufficient number of artificially labeled and segmented words, we use two machine learning methods of artificial intelligence: Conditional Random Fields (CRF) and Long Short-Term Memory (LSTM)

to train the model of automatic word segmentation. Our word segmentation accuracy is shown in Table 1. The highest accuracy rate has reached over 94%, which shows that our word segmentation system is quite effective. If Jieba and CKIP Tagger, which are designed for modern Chinese, are used for Buddhist scripture segmentation, the accuracy is only 75% and 78% (see Table 1), which is enough to prove the importance of Buddhist scripture segmentation in this project. We have also opened up the automatic word segmentation system for Chinese Buddhist scriptures, and the system for building a web interface can be used by anyone, as shown in Figure 2.

Results and Prospects

At present, we have completed the manual segmentation of Buddhist scriptures of more than 3 million characters, and the accuracy rate of automatic segmentation has reached over 94%. In contrast, the modern Chinese automatic word segmentation system, which has

been developed for many years, generally has an accuracy rate of 96%, which is not far from ours. In the spring of 2021, we have officially opened all CBETA texts after automatic word segmentation. It is hoped that this result can drive more research and text application of Buddhist scriptures. At present, we continue to perform artificial word segmentation for texts of the Chinese writing types with poor word segmentation results, hoping to improve the accuracy rate of word segmentation of classics from different periods (medieval and modern) and different types (sutra, vinaya, shastra, commentary, biography, preface, recorded sayings). In addition, in natural language processing, the next stage of word segmentation is Part-of-Speech tagging. At present, we have started to write the DILA Part-of-Speech Tagging Manual and have initially completed the artificial part-of-speech tagging of 250,000 words. Our hope for the future is to be able to move towards the development of an automatic tagging system for Buddhist scriptures.

Table 1: Project automatized word segmentation system performance

Precision	Recall	F1-score
0.9461	0.9526	0.9493
0.9316	0.9328	0.9322
0.7860	0.7312	0.7576
0.7546	0.7140	0.7338



▲ Figure 2: Buddhist scripture automatized word segmentation system