

行政院國家科學委員會 數位典藏國家型科技計畫
九十五年度成果報告

新詩改罷自長吟-唐詩吟唱虛擬實境教學環境建置計畫

計畫類別： 個別型計畫 整合型計畫 國家型科技計畫
計畫編號：NSC 95-2422-H-155-001
執行期間： 95 年 3 月 1 日至 96 年 2 月 28 日

計畫主持人： 羅鳳珠
共同主持人： 張智星、許介彥
專任助理：吳侑倣、黃詩媛
計畫參與人員：李元萍、劉怡姍、詹錦麗、陳培芸、蔡宜蓉
兼任助理（研究生）：陳憶玟、許肇凌、李念容、鄒銘軒
特別協助：清蔚科技公司、優特科技股份有限公司

執行單位：元智大學中國語文學系

中 華 民 國 96 年 4 月 2 日

目錄

中文摘要	I
英文摘要	I
前言	1
研究方法	1
(一) 在音訊處理方面	1
(二) 詩意象資料庫	5
(三) 建立唐詩詞彙切分及語意標註	8
進行步驟及執行成果	20
教學觀摩會	20
網站呈現	20
網站基本架構	23
網站架構圖	25
對於學術研究、國家發展及其他應用方面預期之貢獻	26
成果發表	26
結果與討論	26
參考文獻	26
研究成果附件：〈植基於語意學及使用者認知行為的資訊檢索系統設計：以全唐詩網站為例〉	

摘要

本計畫主要的目的是引用資訊科技，在學術網路上建立唐詩吟唱的多媒體數位學習環境，並與唐詩結合，讓最古雅的中國詩樂與最新的資訊科技結合，以推廣並傳承日益衰微的唐詩吟唱古曲與吟唱方法，免使成為絕響。

音樂與文學是人類各種藝術中最具有抒情性，最易於感人的藝術形式。詩與樂本來就是同源的藝術，都是人心感於外物的情意表現。張騫通西域，引進很多西方樂器，為中國詩樂結合藝術邁入新的階段，唐詩與音樂的結合，也因為新樂器的加入而有新的面貌。

以吟調吟誦唐詩，可以欣賞唐詩聲情之美，還可以協助背誦，提高學習的興趣，本計畫在九十四年度「忍把浮名換了淺斟低唱：宋詞古唱虛擬實境教學環境建置計畫」的基礎，以及「卡拉迷：淺斟低唱」的技術與模式，結合「全唐詩」網站、「依韻入詩」網站，建置「新詩改罷自長吟：唐詩吟唱虛擬實境教學環境建置計畫」，提供唐詩及唐詩吟唱教學使用。

關鍵字：唐詩、唐詩吟唱、音訊處理、網路教學、數位典藏

英文摘要

The goal of this project is to combine the latest technologies from digital signal processing, music information retrieval, multimedia rendering, and speech recognition to construct an integrated environment for teaching/learning the chanting of Tang poetry over the Internet. The constructed system can effectively preserve the music information (ancient music scores) of Tang poetry. More importantly, it can promote the chanting of Tang poetry for the young generation.

Music and literature are two of the most heart-touching parts of all kinds of art formats. The chanting of Tang poetry is the combination of these two art formats, which represent an important milestone of traditional Chinese art for many generations.

The use of multimedia tools for the chanting of Tang poetry can not only lead the young generation to appreciate the beauty of Chinese literatures, but also help them to memorize the Tang poetry with great interests. With the foundation of previous related projects, we are aiming at constructing an integrated environment for preserving the music scores of Tang poetry, and for learning/teaching the chanting of Tang poetry. This is an important task toward preserving the traditional Chinese art of music and literature, and presenting it to the young generation in a totally new manner to let them appreciate the beauty of Tang poetry.

Keywords: Tang Poetry、Chanting of Tang Poetry、Signal Processing、Web-based Learning、Digital Archives

前言

音樂與文學是人類各種藝術中最具有抒情性，最易於感人的藝術形式。詩與樂都是人心感於外物的情意表現。詩與樂本來就是同源的藝術，從《詩經》開始，中國詩歌就建立了與音樂結合的傳統。文學與音樂的三次結合，分別是：

- (一) 先秦以前之《詩經》、《楚辭》：《詩經》的風、雅、頌以音樂分類，《詩經》三百篇，孔子皆弦歌之；屈原創作合樂的〈九歌〉，另一篇〈九章〉也以樂為名，這時期以雅樂為主。
- (二) 漢魏時期的樂府詩：無論是巷末歌謠或吳聲西曲，有很多樂府詩以聲辭相雜的形式傳世，這時期以清商樂為主。
- (三) 唐宋詩詞：以燕樂為主的唐宋詩詞，與音樂的結合就更為密切，更為完美。

詩詞韻文與語體文不同的地方是詩詞有平仄押韻等格律的限制，唐詩分為近體詩與古體詩，近體詩有嚴格的押韻及平仄限制，聲調分為平上去入四聲，每一種聲調的長短抑揚頓挫都不同，交織出詩的聲情之美，唐代之後，各地方也發展出與當地民歌結合，而可以吟誦唐詩的曲調。古體詩的樂府，是當時的琴歌，也都有古調留下。

本計畫主要的目的是引用資訊科技，在學術網路上建立唐詩吟唱的數位學習環境，讓最古雅的中國詩樂與最新的資訊科技結合，以推廣並傳承唐詩及唐詩吟唱的教學。

研究方法

唐詩都是可以吟唱的，唐代樂府詩或保有早期琴歌的曲調，或者從南北朝吳歌西曲等民歌一脈相傳，近體詩除了在後世與地方民歌小調結合之外，可以使用文人讀書調吟唱。

本計畫在保存古譜文獻之外，以輔助唐詩教學以及推廣唐詩吟唱為主要目的，引用資訊科技之最新軟體與音訊技術結合，將唐詩古譜之旋律重現，除使自學、教學兩便之外，與全唐詩資料庫 (<http://cls.hs.yzu.edu.tw/QTS/HOME.HTM>) 連結，讓使用者可以自選吟唱之詩作，不必受限於只能吟唱某些詩作，使教學與自學的自主性更高。再由唐詩之研究與教學者撰寫示範曲的詩作賞析，幫助使用者從體會詩意與詩境之中提升其音樂性的表達能力，可使聲情的詮釋更好。

(一) 在音訊處理方面，主要的研究方法有下列四項：

1. 唐詩吟唱與樂譜的數位典藏

將現存各種唐詩吟唱之樂譜數位化，並抽取必要的資訊，以便進行數位典藏的檢索，所抽取之資訊包含文字及注音（以便進行語音檢索）以及主旋律樂譜資訊（以便進行旋律檢索）。在歌曲的數位檔案方面，我們使用一般通用的標準 KAR 檔案為主，此類檔案可以看成是 MIDI 檔案的延伸，不但包含了樂譜資訊，也包含了歌詞資訊以及對應的時間資訊，所以非常適合唐詩與樂譜的共同保存。此外，此種檔案格式可以通用於數種常用的音樂播放軟體，並可同步呈現歌詞隨旋律的前進而變色，適合唐詩吟唱的學習。

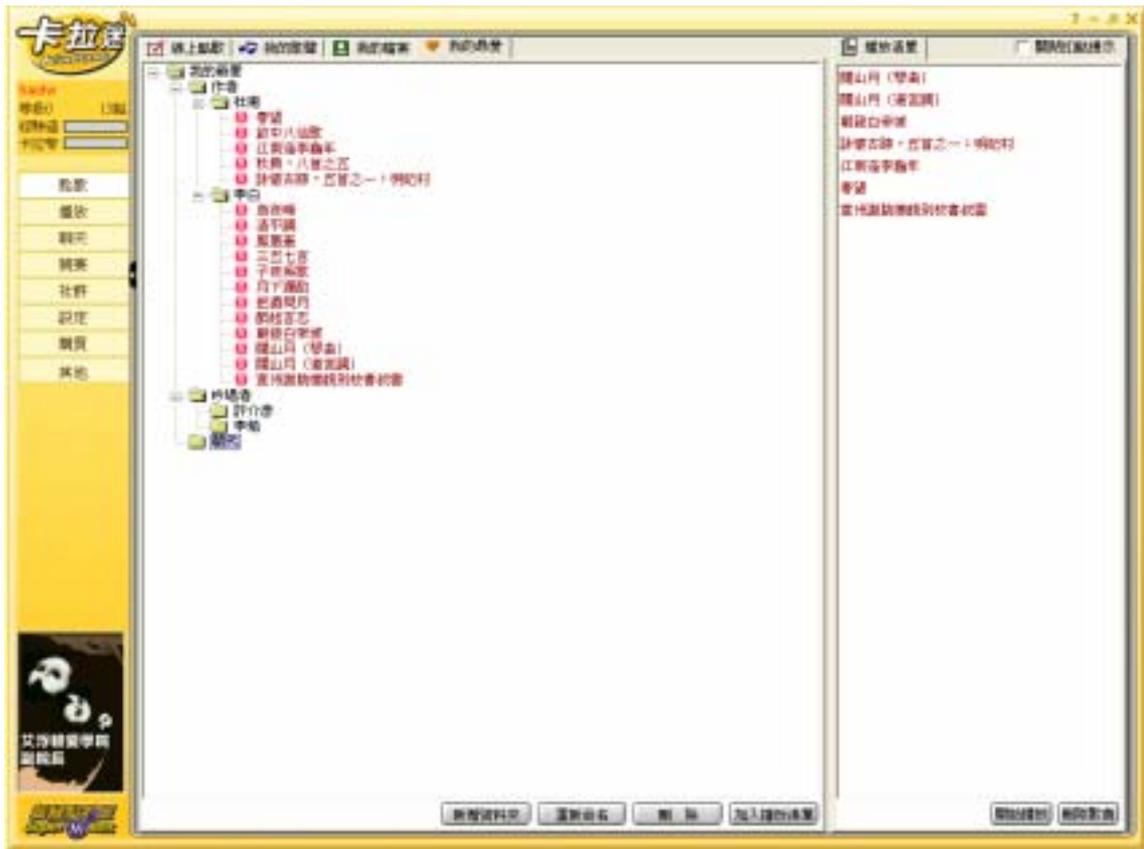
在本計畫中，我們使用與清蔚科技公司合作的「忍把浮名換了淺斟低唱：宋詞古唱虛擬實境教學環境建置計畫」，及清蔚科技授權使用的卡拉 OK 軟體「卡拉迷」，製作唐詩吟唱卡拉迷，直接播放 KAR 檔案。「新詩改罷自長吟卡拉迷」功能畫面如下：



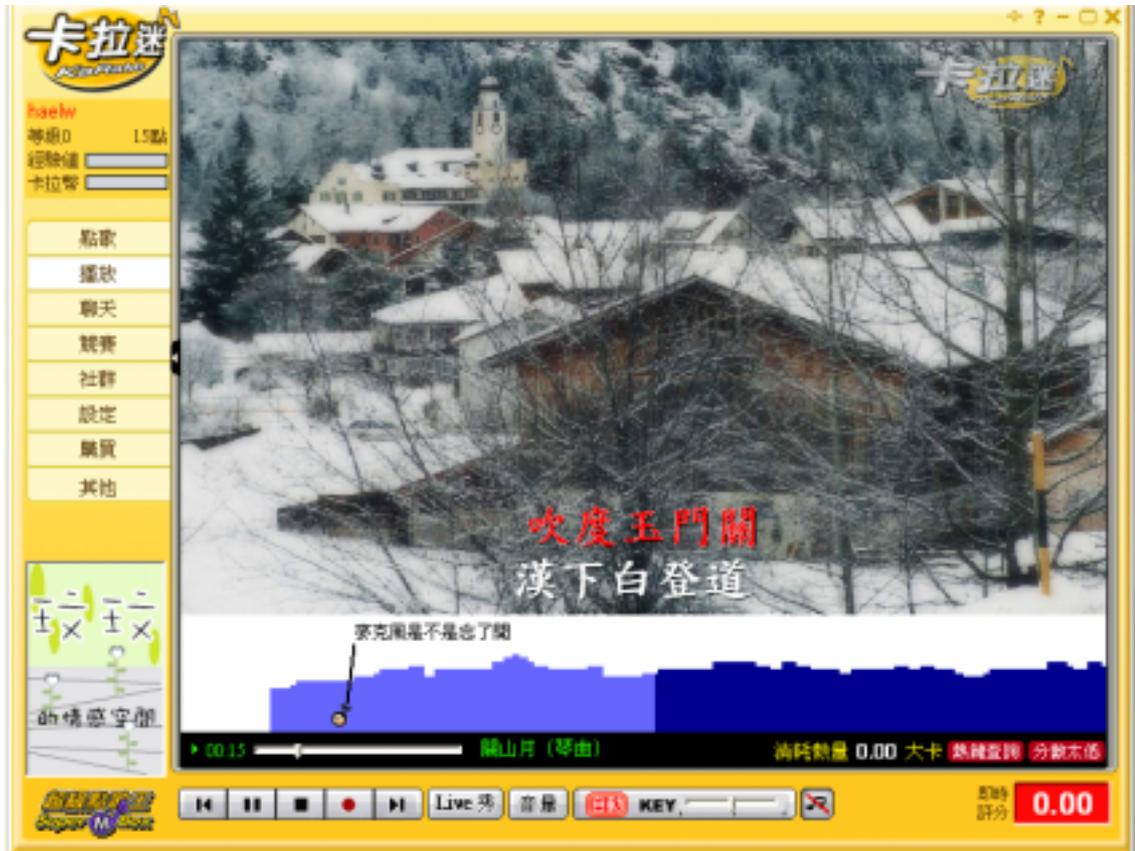
圖一：新詩改罷自長吟卡拉迷畫面



圖二：文字、語音、旋律檢索功能



圖三：出處、朝代、吟唱者索引功能



圖四：伴唱資料播放畫面

吟唱等聲音媒體以上述方式建立素材資料庫典藏，作為教學教材，仍須有適當的教學平台輔助。以年輕學子所熟悉的 Karaoke（卡拉 OK）模式建置，以求貼近使用者的使用習慣，吸引使用者。網站的功能包括下列各項，本計畫所製作的唐詩吟唱也具備相同的功能：

- ✓ 文字檢索：輸入文字，以檢索需要的資料。
- ✓ 語音檢索：對於使用鍵盤輸入不便的使用者，提供語音檢索功能，使用者只要以麥克風輸入語音，即可辨識、檢索。
- ✓ 旋律檢索：使用者只要以麥克風輸入四至八秒的旋律，系統即可比對資料庫的音檔，依符合度高低順序列出檢索的結果。
- ✓ 多種音檔資料：提供真人吟唱、一百二十八種不同樂器的電腦合成音樂之 MIDI 等。
- ✓ 卡拉 OK 伴唱功能：同步顯示歌詞，還能夠根據音高曲線進行即時顯示與評分、根據語音辨識來進行咬字評分，同時也可以進行錄音、廣播、調整音高、音量、節奏、錄音、選擇伴唱音樂、設定及更換背景圖檔、評分、顯示音準、將錄音資料存檔、上傳到伺服器、郵件寄送等功能，可以提高使用者在進行唐詩吟唱的趣味性。

此外，在 KAR 檔案的製作方面，清蔚科技公司提供相關的軟體，以便本計畫人員進行音樂之編輯，並加入歌詞（即唐詩）與變色時間資訊。

2. 唐詩與樂譜的檢索

除了提供一般的文字檢索外，我們也發展語音與旋律哼唱檢索系統。在語音檢索方面，使用者可以直接說出詩題或詩文中任何字句，此系統即可根據語音辨識找出最相近的一首詩。這方面我們所採取的技術是與語者無關（Speaker-independent）的大詞彙中文語音辨識引擎，所有聲學模型（Acoustic Models）的訓練和應用都是根據 HMM（Hidden Markov Models），訓練語料則是來自於長庚大學資訊系呂仁園老師所收集的 300 人語料。

在旋律哼唱檢索方面，使用者只要對麥克風唱一句，此系統即可根據音高的變化曲線，自動從資料庫中找出最相近的樂曲。這方面我們使用的音高追蹤方法是來自相關函數（Auto-correlation Functions），而比對方法可以是 DTW（Dynamic Time Warping）或是 HMM。為了增進比對的效率，我們採用兩階段式的方法，換句話說，先用線性伸縮（Linear Scaling）的方法來比對，剔除 90% 的不可能歌曲，然後再用比較複雜的方法來比對剩下的 10% 歌曲，此種做法可以大幅降低比對時間，辨識率並不會下降太多。

為考慮計算量的分散，此檢索系統必須採取 Client-Server 架構，換句話說，當使用者下載此軟體後，所有的比對計算都在 Client 端完成，軟體只負責將比對結果送回伺服器，再抓回必要的資訊，呈現於 Client 端。若比對資料有所增刪，軟體也必須定期從伺服器自動更新，以便維持資料的同步性。若情況需要，我們也可以將此軟體寫成 ActiveX，直接內嵌於網頁，提供友善的使用介面與使用流程。

3. 唐詩吟唱與樂譜的古音重現

如何以傳統樂器的音色來播放唐詩吟唱的 MIDI 檔案，就要靠軟體音源了。比較軟體音源和硬體音源，前者的優點是擴展性強，音色也很逼真，但是缺點是需要佔用 CPU

資源，所以不能用在速度較慢的電腦。而硬體音源的優點是不占用 CPU 資源，音色儲存於音效卡內部的晶片，但缺點是不能更新，價格較昂貴。若要開發完整的國樂軟體音源，必須進行傳統音樂的音色收集與處理，非常耗費人力物力，但這是一項長期的工作，值得長期投入，但是本計畫在前次之宋詞吟唱計畫所發展的胡琴音效果有限，因此在唐詩吟唱計畫便沒有繼續處理。

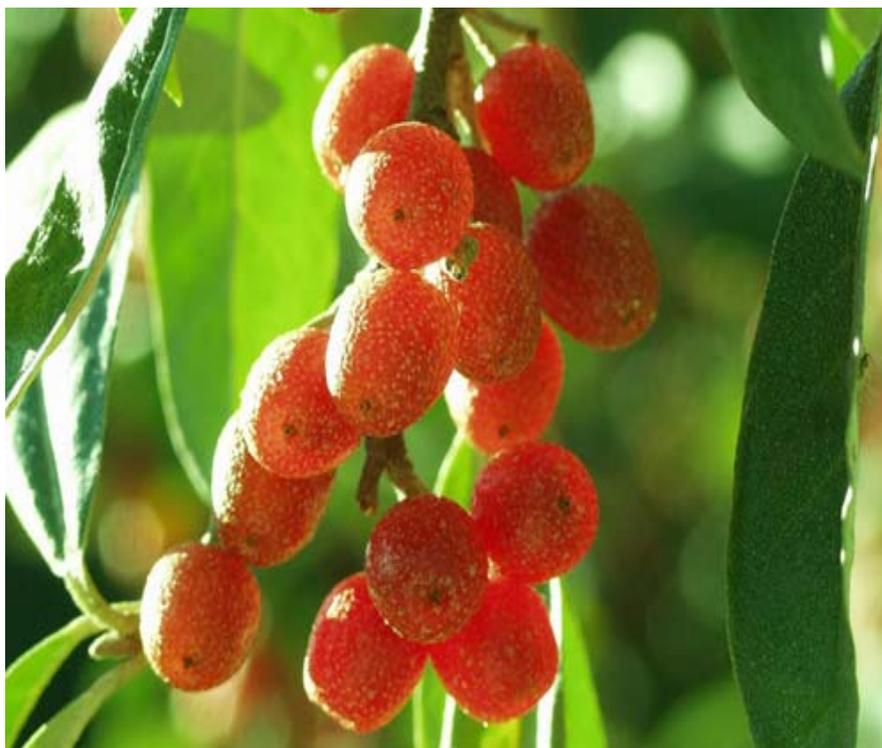
(二) 詩意象資料庫：

文學作品裡以植物與景觀作為物象的比例非常高，植物在文學裡的意義與在生物學裡的意義不一樣，描述的方式也不一樣。作為文學物象與意象使用的植物，其描述方式需從滿足文學應用的角度思考，描述欄位如下表：

圖檔 類別 及 ID	照片	ID	每一張照片或畫作可能提供給多處文本使用，為每一張圖檔建立一個專屬的 ID，以方便電腦程式識別。		
	畫作	ID			
圖檔 描述	作者	記錄作者姓名，以明其著作權。			
	翻拍者	記錄翻拍者姓名，以明其著作權。			
	產出地點	如為植物及景觀照片，尤須記錄地點。			
	產出時間	記錄創作及翻拍時間			
	作品主題	描述最大範圍的主題資料			
人	姓名	畫作裡有一些是真實人物畫像，如蘇軾的自畫像、梁楷的〈太白行吟圖〉。			
	性別	標示圖像中人物性別			
	人數	有些畫作描繪當時的文人生活，如〈西園雅集〉，畫裡是蘇軾、王詵等人。			
植物	學名（可以與植物學領域所建置的植物資料庫連結，以閱讀該植物的詳細資料）	通俗名（文學作品有時使用通俗名）	花 葉 枝 葉 全株	比 喻 象 徵 意 象	文學上描寫植物的比例非常高，經過長期的積累醞釀，不同的植物，在文學作品裡有不同的意義，文人筆下，或以整株植物，或以局部部位描寫，各有不同含義，有些甚至於已經成為固定共認的意象，例如陶潛寫「採菊東籬下」之後，菊花與東籬已經成為公認的隱居意象。
動物	學名（可以與動物學領域所	通俗名（文	禽 鳥	比 喻	與植物相同的情況，文學作品裡描寫動物的比例非常高，不同文體，不同文人，不同風格的作品，所側重的動物也有所不同，全唐詩的字數約為全宋詞

	建置的動物資料庫連結，以閱讀該動物的詳細資料)	學作品有時使用通俗名)	獸 蟲 魚	象 徵 意 象	的二倍，以字頻統計，以文體區分，全唐詩「馬」的使用頻率是4,850次，「鶯」1,034次、「燕」1,556次，全宋詞「馬」1,360次、「鶯」1,236次、「燕」1,952次。很多動物在文人筆下都有固定的意象。
景觀	地景	山、水、原、隰、關、津			文人在不同季節、不同天候、不同時間，面對不同的景觀，所引發的情感也有所不同，在傷春悲秋的文學傳統裡，春天傷失偶，秋天悲失路，各有寄託。
	地貌	樹、草、花			
	季節	春、夏、秋、冬			
	天候	陰、晴、雨、雪、霧			
	時間	清晨、白天、黃昏、晚上			
相關作品	知識性	作者	題目	作品	
		客觀知識性的描述與呈現（連結到該作品的資料庫）			
	情境性	作者	題目	作品	
		主觀情意性的引導與感發（連結到該作品的資料庫）			

圖檔資料經過描述，使用者如需尋找適當的圖檔作為輔助教材，即可依照描述欄位所列出的概念搜尋。屬於知識性的資料，是客觀的，比較單純，容易找到對應資料；以王維〈九月九日憶山東兄弟〉：「獨在異鄉為異客，每逢佳節倍思親。遙知兄弟登高處，遍插茱萸少一人。」為例，可以從「植物類」尋找「茱萸」的圖檔，如下：



屬於情境性的圖檔，容許主觀的判斷與選擇，比較複雜，需要以概念搜尋，以劉禹錫〈竹枝詞〉：「春江一曲柳千條，二十年前舊板橋。曾與美人橋上別，恨無消息到今朝。」為例，若要尋找符合〈竹枝詞〉情境的圖檔，可以從「景觀」類搜尋。〈竹枝詞〉寫劉禹錫於春天獨自來到二十年前與美人分別的板橋，看到景物依舊，而人事已非，美人迄今都無消息。從經過描述的圖檔，逐項循「地景」選擇「水」（春江）、「橋」（板橋）、「地貌」選擇「樹」（春江一曲柳千條）、「季節」選擇「春天」（春江），即可將圖檔縮小在符合條件的範圍裡，迅速找到適當的圖檔。換言之，每一張圖檔，適合詮釋那些文學作品，在設計時也保留了由使用者自由搭配的空間。以劉禹錫〈竹枝詞〉為例，建立的資料描述如下表：

圖檔類別及ID	照片	ID		
	圖檔描述	作者	王履冰	
		翻拍者	原件	
		產出地點	紐約州阿帕拉契山	
		產出時間	2003年初春	
		作品主題	春愁	
景觀	地景	山、水、原、隰、關、津	板橋、水	
	地貌	樹、草、花	樹	
	季節	春、夏、秋、冬	春	
	天候	陰、晴、雨、雪、霧、雲	晴	
	時間			

相關作品	情境性	作者	題目	作品
		劉禹錫	竹枝詞	春江一曲柳千條，二十年前舊板橋。曾與美人橋上別，恨無消息到今朝。
		白居易	板橋路	梁苑城西二十里，一渠春水柳千條。若為此路今重過，十五年前舊板橋。曾共玉顏橋上別，恨無消息到今朝。
		馮延巳	鵲踏枝	誰道閑情拋棄久，每到春來，惆悵還依舊。日日花前常病酒，不辭鏡裡朱顏瘦。○河畔青蕪堤上柳，為問新愁，何事年年有。獨立小橋風滿袖，平林新月人歸後。
		詩經·小雅	采薇	昔我往矣，楊柳依依。今我來思，雨雪靡靡。行道遲遲，載渴載飢。我心傷悲，莫知我哀。
		韓翃	章台柳	章台柳、章台柳，昔日青青今在否？縱使長條似舊垂，也應攀折他人手。
		柳姬	章台柳	楊柳枝、芳菲節，苦恨年年贈離別，一夜隨風忽報秋，縱使君來豈堪折。
		庾信	枯樹賦	昔年種柳，依依漢南。今看搖落，悽愴江潭。樹猶如此，人何以堪。

(三) 建立唐詩詞彙切分及語意標註：(詞群內容含有唐宋詩詞等使用的詞彙，以求建立詞彙切分及語意標註的資料庫，延伸至其他方面的研究。)

大類	小類	詞群 (部分內容)
人名	本名	牛希濟 (唐)、王充 (漢)、王弘 (晉)、王安石 (宋)、王昌齡 (唐)、任昉 (南朝梁)、列禦寇 (戰國)、匡衡 (漢)、張騫 (漢)、曹丕 (三國魏) ……
人名	別名字號	包公 {包拯 (宋)}、長卿 {司馬相如 (漢)}、史遷 {司馬遷 (漢)}、太沖 {左思 (晉)}、白香山 {白居易 (唐)}、白樂天 {白居易 (唐)}、香山 {白居易 (唐)} ……
人名	他稱	曹氏 {曹植 (三國魏)}、曹奸 {曹操 (三國魏)}、曹瞞 {曹操 (三國魏)}、莊子 {莊周 (戰國) 【莊子】}、莊生 {莊周 (戰國) 【莊子】} ……
人名	法號	目蓮 {目蓮 (佛經人物)}、生公 {竺道生 (晉)}、契虛 {契虛 (唐)}、永禪師 {智永禪師 (唐)}、圓澤 {圓澤 (唐)}、佛印 {釋了元 (宋)}、參寥 {釋道潛 (宋)}、釋參寥 {釋道潛 (宋)}
人名	合稱	二喬 {大喬 (三國)、小喬 (三國)}、尹邢 {尹夫人 (漢)、邢夫人 (漢)}、元白 {元稹 (唐)、白居易 (唐)}、白

		元 { 元稹 (唐)、白居易 (唐) }、孔孟 { 孔丘 (春秋) ……
人名	外國人名	久保天隨 (日本)、山口淑子 (日本) 【李香蘭】、川上金十郎 (日本)、米拉 (法國)、艾森豪 (美國)、伯利 (法國)、狄渥 (匈牙利) ……
人名	待考	子貞、子高、子章、子璋、尹公、元王醴、孔禰、王昌、石舟、朱蹶、何將軍、何孺人、余蘇、吾宗英、吳王、宋邑侯、宋真、李友諒 ……
行政體系	部門名稱	三臺、公車、仙曹、吏部、尚書省、軍部、秦廷、御史臺、荷廳、朝廷、楚庭、虞庭、賊庭、綠莎廳、樞密院、翰林學士院、禮部
朝代君王	朝代民族名號	三晉 (春秋)、三國、大宛、女真、五胡、五獠、元、元清、六朝、巴子國、犬戎、冉駝、北狄、北朝、北燕、仡佬、匈奴、回紇、夷 ……
朝代君王	君王帝號	三皇、五帝、太宗 (待考)、文公 (待考)、文宗 (待考)、世祖 (待考)、宋太祖、宋世宗、宋神宗、李後主、周文王、周平王、周成王 ……
朝代君王	君王年號	大順、大曆、元和、元武、元封、元祐、元嘉、元豐、天寶、太元、文景、正始、民國、永和、永昌、永明、永曆、永興、光緒、同光、同治 ……
朝代君王	君王陵墓	孝陵
神鬼仙人	神鬼專稱	大梵天王、女媧、山神、山鬼、山魃、天人、天女、天公、天王、天仙、天后、天吳、天君、天帝、天皇、天神、太乙、月老、木魅、水神 ……
神鬼仙人	神鬼泛稱	天、木客、仙、玉女、先人、先世、先君、列真、羽人、羽士、羽化、羽客、佛、佛仙、妖、妖孽、妖魔、妖魃、夜叉、怪、祖、神、神仙 ……
神鬼仙人	菩薩尊號	三祖、文殊、佛陀、伽藍、金剛、活佛、修羅、菩薩、羨門、達摩、摩訶迦葉、彌勒、應真、釋迦牟尼、觀世音
身體器官	五官顏面	口、口 (嘴)、目、目 (眼)、目 (眶)、目 (眸)、目 (眇)、目 (睥)、目 (睨)、目 (睛)、目 (瞳)、耳、眉、眉 (睫)、唇、鼻、頸、頸 (項) ……
身體器官	鬚鬢髮膚	毛、毛 (髻)、膚、膚 (皮)、膚 (肌)、髮、髮 (絲)、髮 (髻)、髮 (鬢)、髮 (鬢)、鬚、鬚 (髯)、鬚 (髭)、鬚 (鬍)、鬚 (鬚)、鬚
身體器官	身體四肢	手、手 (拳)、手 (掌)、肘、身、身 (軀)、身 (體)、乳、拇、肢、肱、股、肩、指、指甲、背、胛、脅、胸、胸 (腔)、胸 (臆) ……
身體器官	五臟器官	丹田、心、肉、舌、血、肝、肺、胃、胛、骨、骨 (骨骼)、骨 (髀)、骨 (髓)、喉、腑、腎、脾、腸、腸 (盲腸)、腦、膜、齒、齒 (牙)、齒 (齦)、骸、魑、膽、髓、臟

身體器官	分泌物	汗、沫、津、淚、涕、液、淚
身體器官	合稱	口心、口目、口耳、口舌、口腹、口鼻、五官、六根、孔竅、心目、心血、心肝、心肺、心胸、心骨、心眼、心脾、心腹、心膽、手眼、牙鬚……
地名	行政地名	丁字門、七崁頭、七鯤身、九江、九芎林、九連城、二林、八尺門、八里、八罩、十三灣、卜其吉、三台、三角莊、三原、三蜀、下灰窯……
地名	山嶺地名	七星山、九日山、九華山、九疑山、九龍山、八卦山、十八尖山、三垂岡、三游洞、三貂嶺、丈人峰、上黨山、千佛山、千佛嶺、大小山……
地名	河湖地名	七洲、九曲塘、九曲溪、九龍淵、九鯉湖、二層溪、八卦潭、力力溪、十八重溪、下雙溪、大甲溪、大西洋、大夏河、大湖、大漲海、大龍湫……
地名	島嶼地名	四角嶼、布嶼、目嶼、西吉嶼、西嶼、花嶼、虎嶼、南嶼、紅毛嶼、桃花島、浪嶼、荔島、浯島、球嶼、荷苞嶼、陰陽嶼、湄島……
地名	關津地名	丁卯橋、七盤關、二重橋、八通關、三林港、三港、久里濱、大東溝、大磴門、丹鳳門、五里橋、午門、午橋、天門、天津橋、天關、太平橋……
地名	外國地名	巴黎（法國）、日光（日本）、爪哇島（印尼）、平壤（韓國）、伊豆山（日本）、江戶（日本）、西貢（越南）、呂宋島（菲律賓）、京都（日本）、咖哩（呂宋群島）、咖蠟吧（印尼）……
地名	地名代稱	小蓬萊、小瀛洲、天柱（不周山）、上都（北京）、五城（北京）、六街（北京）、天都（北京）、京城（北京）、京師（北京）……
地名	地名合稱	九河、三川、三江、三洲、五嶺（大庾嶺、始安嶺、臨賀嶺、桂陽嶺、揭陽嶺）、嶺嶠（大庾嶺、始安嶺、臨賀嶺、桂陽嶺、揭陽嶺）、六郡、平津（北京、天津）、京綏（北京、綏遠）、伊洛（伊水、洛水）、伊雒（伊水、洛水）……
國名	國名	中國、日內瓦、日本、北韓、伊朗、印度、波蘭、法國、西班牙、英國、阿拉伯、俄國、南韓、美國、德國、埃及、泰國、真臘、荷蘭、朝鮮、越南、義大利、韓國、羅馬、蘇丹
地理	區域範圍	九方、九夷、九州、九垓、九原、九邊、八州、八表、八荒、八埏、八極、三吳、三秦、三湘、大陸、中土、中州、中西、中東、中原……
地理	邦國都城	州、州縣、庄、邑、邦、邦縣、里、里州、里社、京、府城、社稷、南柯國、城、城鄉、省、郡、國、國土、域、都、堡、鄉、鄉里、鄉國、鄉縣、畿、縣、鎮、疆
地理	郊原村野	市、市井、市朝、田、田園、地、址、村、村市、村社、村

		莊、村郭、甸、町、坪、社、垌、厰、垓、畎畝、畝、郊、郊郭……
地理	街道巷弄	阡、阡陌、巷、陌、逕、途、街、街巷、街、路、道、道路、塍、塍、隧、蹊、衢
地理	特殊場域	牛欄、外家、市舶、市廛、伏櫪、杏壇、災區、阜、京師、庖、家鄉、酒家、酒肆、酒樓、盜藪、菜市、詞林、鄉關、賊窟
自然景觀	山峰崖嶺	山、丘、岬、坏、岑、阪、坡、岡、岫、岳、阜、峙、坨、垓、峭、峨、峰、崖、崙、崗、陵、嶺、崱、塹、崑、嶂、墩、嶠、壁、嶺、嶽、麓、巔、巒、巖、巘
自然景觀	水澤湖泊	川、水、圳、江、池、汐、汛、汎、河、沼、波、沫、泛、泊、泖、派、泉、洋、洪、流、洄、洄、浪、浦、海、海嶠、涓、涯、淵……
自然景觀	沙洲島嶼	汀、冰山、沚、岸、坻、洲、原隰、島嶼、皋、渚、磯、礁、灘
自然景觀	洞穴坑谷	穴、坑、坎、谷、坳、洞、峽、硿、塢、窟、塹、窪、壑、谿、竇
自然景觀	林原莽漠	山林、戈壁、平楚、石林、林、草原、草莽、莽、森、漠、瑤林、叢
自然景觀	山川泛稱	山川、山水、山河、山海、江山、河嶽、溪山
自然景觀	關津渡口	江口、津、埠、梁、埭、堤、堰、港、渡、棧、隈、溪口、閘、隘、橋、關、壩
建築物	建築專名	九成宮、八詠樓、三仁祠、三槐堂、上林苑、上清宮、乞巧樓、夕陽樓、大阪城（日本）、大明樓、女懷清臺、子陵臺、子雲亭……
建築物	宮室屋廬	王庭、北闕、宇、宅、朱邸、行窩、坊、廷、居、店、府、房、舍、邸、困、室、屋、倉、厝、家、宮、庫、旅舍、軒、堂、宿、第……
建築物	亭臺樓閣	亭、亭臺、榭、臺、臺閣、閣、樓、閣、閨、闈、闕、譙
建築物	園林院落	公園、園、苑、庭、庭院、院、園
建築物	寺廟道觀	寺、佛門、佛堂、佛殿、佛閣、法門、剎、宮、祠、祖堂、神社、庵、梵宇、梵宮、道場、僧舍、僧庵、精舍、廟、壇、禪室、禪宮
建築物	碑塔陵墓	孔林、宅兆、京觀、孤拔墓、長陵、安陵、陽陵、茂陵、平陵、青塚、香冢、陵、塚、塔、塋、碑、碣、墳墓、壙、壟
建築物	官署建築	中軍府、公府、公堂、太倉、太學、考場、行臺、局、制府、官府、官舍、官倉、官第、官署、官衙、官閣、帑庫、府庫、明堂……
建築物	建築部件	井、戶、瓦、瓦溝、杆、宗、板、門、門窗、門榜、門磚、

		門牆、垣、柱、架、柵、砌、門、扃、宮門、梯、廊、扉、棟、窗、階……
建築物	軍事攻防	牙門、阱、砲臺、堠、營、壘、甯
閼苑仙境	閼苑	天府、天庭、天曹、天闕、月府、月宮、月殿、水晶宮、仙府、仙宮、仙臺、玉宇、玉京、玉樓、兜率天宮、清虛府、紫府、……
閼苑仙境	仙境	三山、大羅天、丹丘、化城、天池、天堂、方丈、方壺、仙山、仙海、仙鄉、仙源、北溟、玄圃、玉霄、西天、西方、佛界、佛國……
幽冥之境	幽冥	下泉、地下、地獄、阿鼻、幽冥、幽都、泉下、泉路、泉壤、重泉、冥府、冥曹、冥間、鬼府、鬼門關、鬼國、鬼域、鬼境、黃泉
天候氣象	風霜雪露	冰、風、雪、飈、霜、颶、颶、霰、露、颯
天候氣象	雲霧煙霞	氛、氣、嵐、雯、雲、煙、煙雲、氤、霞、霧、霾、靄
天候氣象	雷電電霓	光、虹、虹電、風雷、雷、電、雹、蛻、輝、霆、霓、霓虹
天候氣象	冷寒	冰、冷、冽、凍、涼、寒
天候氣象	熱旱	早、炎、暑、暖、溫、熱、燠
天候氣象	晴天	日麗、朗、氣爽、晴、暘、霽
天候氣象	陰雨	雨、淋、陰、澍、霖、霖、霏
天候氣象	合稱	天氣、冰炭、冷暖、冷熱、炎涼、雨雪、雨晴、雨露、雨霽、風月、風雨、陰晴、雪月、寒暖、寒溫、煙雨、霜雪
天文	日	日、日車、日烏、火傘、火輪、白虹、光、光輪、旭、扶桑、赤丸、金烏、昉、紅輪、飛輪、陽、暉、照、曉色、羲車、燭龍、曜、織烏、曦
天文	月	太陰、月、月華、玉兔、玉鉤、玉盤、玉輪、玉蟾、生魄、冰壺、冰輪、孤輪、明蟾、金波、金輪、秋蟾、桂宮、桂窟、桂魄、桂蟾、海蟾……
天文	星	七曜、三星、三臺、丈人、大星、女宿、天狼、天樞、太白、熒惑、鎮星、太微、太歲、少微、文曲、文昌、斗、斗杓、斗柄、斗魁……
天文	天空	九重、天、天空、天象、月地、玄穹、宇、宇宙、昊蒼、穹、青冥、沕寥、凌霄、虛、蒼昊、蒼垠、蒼冥、霄漢
天文	銀河	天河、天漢、河漢、星河、星漢、星橋、雲漢、漢津、遠漢、銀河、銀漢、層漢、鵲橋
天文	合稱	天文、日月、日星
自然資源	玉石	介圭、玉、圭、圭璋、玖、金玉、玦、珀、珂、琉璃、珠、珠（夜明珠）、珠（珍珠）、珠（驪珠）、珠玉、珠璣、珪、珩、琅、晶、琛……
自然資源	火	火、炎、炙、炬、焚、焰、燐、燒、燃、燄、燧、燹

自然資源	金屬	汞、泥金、金、金粉、金銀、金鋼、金寶、金鐵、青銅、鉛、鉅、箔、銀、銅、銅鐵、鍍、鋼、錫、鐵
自然資源	礦物	土、丹砂、水晶、石、石（三生石）、石（大理石）、石（文石）、石（鍾乳石）、灰、灰塵、沙、岩、泥、泥沙、玻璃、砂、埃、硫磺、硃、硝、雄黃、塵、磧、磷、礫、鑽
自然資源	燃料	瓦斯、油、炭、柴、煤、樵、薪
自然資源	合稱	土木、木石、金石、鐵石
飲食	飲品	玉液、玉漿、甘露、乳、茶、茶（雀舌）、茶（碧螺春）、茶（蒙頂）、茶（龍團）、茗、酒、酒（竹葉青）、酒（柏酒）、酒（屠蘇）……
飲食	食品	丸、土豆、丹、五辛盤、五侯鯖、月餅、火棗、牛奶、半菽、年糕、束脩、角黍、豆花、豆腐、金膏、侯鯖、食、宵夜、草具、脯……
飲食	食材	米、肉、豆麥、豆麻、肥甘、粃糠、蛋、豚羔、野味、野菜、魚子、棗脯、粟麥、粟粃、腴、菽粟、菜、菜根、棗、葷、稬種、蔬、豬肉、豬肝……
飲食	零食	瓜子、石蜜、豆餅、豆糖、豆蓴、紅龜粿、胡餅、核餅、茶果、酒果、鹿糕、湯餅、飴、餅、鳳眼糕、糕、餈
飲食	調味品	油、脂、麻油、蜜、醋、糖、醬、醢、齏鹽、齏、齏鹽、鹽、鹽梅
飲食	菸草	巴菰、水煙、洋煙、菸、煙、鴉片
飲食	宴席	文宴、吟宴、郇廚、宴、席、奠觴、壽宴、歌筵、筵、餞、饗
飲食	合稱	米魚、米糖、米鹽、果蔬、油鹽、柴米、海產、茶酒、茶湯、酒肉、酒肴、酒脯、魚米、魚稻、魚鹽、飲食、梁肉、蔬果、蝦筍、錢米、薪米、雞黍、麵米
四時節氣	節氣	立春、雨水、驚蟄、春分、清明、穀雨、立夏、小滿、芒種、夏至、小暑、大暑、立秋、處暑、白露、秋分、寒露、霜降、立冬、小雪、大雪、冬至、小寒、大寒
四時節氣	節慶	七夕、上巳、上元、中元、中秋、元宵、令節、年節、孟蘭、社日、花朝、重陽、修禊、除夕、寒食、新年、聖誕、詩人節、端午、臘八
四時節氣	四季	冬、冬春、四季、四時、春、春夏、秋、秋冬、秋春、夏、夏冬、夏秋
時間	明確時間	乙卯、乙未、丁子、丁巳、丁丑、丁卯、丁未、丁酉、七月、七更、九月、二月、二更、八月、八時、十一月、十二月、十二月二十四日……
時間	範圍時間（今昔）	今、今昔、兒時、彼時、於後、昔、近來、前期、後時、昨今、邇來

時間	範圍時間 (日)	一日、七日、九日、二十日、二日、八日、十二日、十日、三日、千日、五日、六日、日、半日、生日、伏日、吉日、早日、百日、何日……
時間	範圍時間 (月)	一月、二月、月、月底、月初、月餘、半月、佇月、何季、來月、初旬
時間	範圍時間 (世代)	一世、一代、八代、十世、三世、三代、三朝、千代、千古、千秋、五代、今世、六代、太初、廿紀、世、世紀、代、半世、四代、末葉……
時間	範圍時間 (生)	一生、三生、半生、前生
時間	範圍時間 (年)	一年、七十年、七年、九年、二十年、二年、八十年、八年、十年、三十年、三千年、三年、三百年、千年、五年、五百年、今年……
時間	範圍時間 (晨)	日間、平旦、平明、旦、早、昕、侵晨、凌晨、晝、晨、曉、曙
時間	範圍時間 (夜)	一夜、二夜、十夜、三夜、夕、子夜、中宵、五夜、今夜、今宵、午夜、日旰、日暝、日暮、日薄、半夜、未央、向晚、夙、此夜……
時間	範圍時間 (時刻)	一霎、一瞬、分、分陰、月日、光陰、即刻、良久、辰、其後、刻、初更、初時、近時、剎那、秒、時、晌、破曉、陰、晷、跳丸、韶光、轉眼、轉瞬
時間	範圍時間 (其他)	干支、午、天干、日期、伏臘、合朔、地支、早晚、旬、明朝、春秋、期、朝、殘臘、週
時間	年齡	半百、古稀、年少、年幼、年光、年老、年命、年庚、年長、年紀、年時、年高、年華、年歲、成丁、成年、艾、行年、庚、冠、歲
文藝	書籍	千佛經、小學、山海經、五經、五緯、六經、六藝、六韜、六籍、公羊傳、太玄經、文選、文韜、方書、水經注、水滸傳、史書……
文藝	文具	丹黃、丹鉛、文具、印、書版、書箱、烏絲、紙、紙筆、紙墨、硯、筆、筆架、筆硯、筒、牋、管、箋、墨、墨水
文藝	書篇名	七發、九章、九歌、九辨、九疇、九辯、八哀、八陣圖、十香詞、三頌、三墳、大東、大風歌、大雅、子虛賦、小弁、小戎、小星……
文藝	文體	七律、七絕、八股、小令、元曲、六義、天章、文、文史、文字、文契、文宣、文風、文書、文案、文義、文學、文藝、比興……
文藝	引文	韋曲
技藝	畫	九老圖、十眉圖、丹青、太極圖、水墨、百壽圖、宋搨、卦圖、河圖、版圖、流民圖、粉本、幅、畫、畫品、畫骨、畫

		像、罨畫、像、圖、圖畫、輞川圖、翰墨、繪、鐵畫
技藝	書法	草隸、隸、小楷、小篆、正書、甲骨文、字帖、行書、宋楷、佉盧、飛白書、書法、草書、鳥篆、楷書、篆、蝌蚪文、遺墨、龍蛇、籀文、鶴書
技藝	碁	方罫、奕、枰、棋、棋盤、碁
技藝	舞蹈	干羽、胡旋舞、柘枝、魚龍舞、楚舞、歌舞、舞、醉舞、驚鴻舞
技藝	戲劇	戲
生物	神物	天狗、天雞、巴蛇、比翼鳥、水怪、仙花、仙桂、仙桃、仙草、仙葩、仙樹、玄武、白澤、交梨、朱雀、犴狴、青鳥、青龍、虯、虯螭、虯蟠、神羊、神物、神狐、神馬、神魚、神獅……
生物	動物專名（走獸）	牛、羊、兔、兕、狗、狐、虎、狢、獬、狼、狸、班馬、豺、豹、馬、鹿、犀、猴、猩、象、貂、獾、獅、猿、貉、鼠、獐、熊、獾、狸、豬……
生物	動物專名（飛禽）	丹頂鶴、孔雀、布穀鳥、白頭翁、伯勞鳥、杜鵑、信天翁、帝雉、烏鴉、班鳩、秦吉了、隼、商羊、梟、雀、斑鳩、雁、黃鶯、雉、鳩……
生物	動物專名（昆蟲爬蟲）	介蟲、尺蠖、甲蟲、虱、青蚨、虺、蚊、蚓、蚤、蚨、蛇、蛆、蚱蜢、蚯蚓、絡繹、蛛、蜈蚣、蛾、蜂、蜉蝣、蜻、蜥蜴、蝴蝶、蝸……
生物	動物專名（水生兩棲）	丁香魚、比目魚、水母、江瑤、牡蠣、貝、泥鰱、河豚、虎鯊、珊瑚、玳瑁、射工、海馬、烏魚、烏賊、珠母、蚌、蛙、蛤、蜃、蜊……
生物	動物專名（合稱）	牛羊、牛狗、牛馬、牛驢、羊豬、兕虎、狗羊、狗兔、狗馬、狗豬、狐兔、狐鼠、狐豬、虎狼、虎豺、虎豹、虎魚、虎鼠、虎熊、虎貔、虎鵠……
生物	動物通稱	毛、毛角、爪、牝、牝牡、羽、羽毛、羽翮、羽翼、羽鱗、卵、尾、牡、角、隹、牲、牯、胚胎、胎、畜、翅、茸、巢、翎、魚、鳥、鳥獸……
生物	植物專名（木本）	丁香、七里香、人參、八角、山桃花、山茶、山椒、山楂、月季、月桂、木瓜、木芙蓉、木棉、木槿、木蘭、木槠茶花、水楊、冬青……
生物	植物專名（草本）	一葉蘭、三友花、女蘿、山芋、山藥、天葩、月桃、水仙、水蘇、牛溲、王蘭、朮、玉蘭、甘薯、白菖、白蘘、石竹、石花、石蓮、向日葵……
生物	植物專名（禾本）	甘蔗、禾、竹、佛眼竹、君子竹、孟筍、茅、桂竹、秧、芻、芻豆、秫、麥、筍、粟、菅、菰菜、黍、稷、稗、稌、荊竹、椰竹……

生物	植物專名（苔菌）	地衣、苔、苔蘚、海苔、紫芝、菌、蘚
生物	植物專名（瓜果）	瓜、瓜果、羊棗、西瓜、佛手柑、亞波羅、波羅蜜、花瓜、南瓜、香蕉、桑椹、匏、瓠、葫蘆、葡萄、鳳梨、橡栗、橡實、龍眼、釋迦梨、櫻桃、瓢、蘿蔔
生物	植物專名（合稱）	瓜李、禾黍、竹柏、李梅、杏桃、杏梅、杏橘、杞梓、杞棘、杞菊、杉松……
生物	植物通稱	木、卉、皮、朵、秀、果、枝、枝幹、枝葉、枝樹、芳、芽、花、花木、……
稱謂	自稱謙稱	小人、己、予、奴、末學、老夫、臣妾、臣僕、余、吾、我、芻蕘、草茅、鯁生
稱謂	尊稱美稱	二老、人瑞、士紳、大俠、女士、女英、巾幗、干城、才、才人、才士、才子、才鬼、才媛、仁人、仁傑、公、公子、夫人……
稱謂	貶稱惡稱	人妖、土番、小人、干謁、斗筲、冬烘、奴顏、生番、奸、奸宄、曲士、老儻、佞、迂才、迂儒、卑下、卑官、昏王、狎客、鬼……
稱謂	親人眷屬	女、子、子女、子孫、公婆、夫、夫妻、父、父子、父兄、父母、父祖、兄、兄弟、兄嫂、外孫、母、次男、考妣、伴、伯、伯兄……
稱謂	指代稱謂	土人、山人、五妃、巴人、文人、日人、氏、他、功狗、古人、布衣、白衣、伊、同胞、同袍、同輩、同學、汝、行人、西人、別人……
稱謂	一般稱謂	人民、丈夫、士、女、子、夫、父老、生、老人、老大、老子、伯、吾輩、君、男、男女、赤子、兒、姑、孤兒、朋友、者、青年、娃……
稱謂	職業身份	乞丐、尸祝、工、方士、木匠、水手、牛郎、主席、代書、占人、囚、囚（虜）、奴、奴（婢）、奴（僮）、奴（僕）、奴（僕婢）、……
稱謂	職官爵位	三公、士大夫、大夫、中丞、中郎、中書、中尉、中臺、尹、元帥、公侯、公卿、太史、太守、太保、太師、太尉、太常、太傅、太僕……
稱謂	皇室稱謂	元首、天子、太子、太后、王、王子、王公、王姬、王孫、王爺、世子、外戚、甲族、后、后王、妃、君、君王、貝子、侯王、帝……
稱謂	宗教稱謂	上人、比丘、出家、尼、全真、佛子、佛家、佛徒、沙門、沙陀、和尚、居士、施主、偃子、苾芻、真人、衲、梵門、喇嘛、道士、僧、緇流、禪、禪師、羅漢
器物	生活用品（飲食器具）	匕箸、卮、叵羅、玉斗、皿、匱、兕觥、杯、杯（盃）、杯（盅）、杯（栴）、杯（琖）、杯（盞）、杯（觴）、杯勺、杯筴、杯觥、杯盤、杯籌、枚、盃、匙……

器物	生活用品（烹飪器具）	勺、杓、灶、砧、釜、鬲、鼎、鼎彝、廚具、鍋、甑、爐、甌、爨
器物	生活用品（裝置容器）	行李、匣、盆、缸、葫蘆、瓮、栲栳、瓶、瓷、盒、鉢、袋、壺、筐、筐筥、筒、甌、奩、槽、箱、箱（笈）、箱（笥）、箱（篋）……
器物	生活用品（起居用品）	几、几席、几案、几榻、床、床帷、床第、枕、枕席、枕衾、枕簟、枕囊、炕、屏風、席、席（虎皮席）、席（簟）、案、茵、被、被（衾）、被（衾裯）……
器物	生活用品（其他用品）	牙籤、杖、杖（筇）、杖履、拂塵、長命縷、門聯、青奴、春聯、架、面具、香、炷、扇、梳、梳（櫛）、脂粉、眼鏡、傘、晷、粧、鉛黛、漏、線、燈籠、襪、鞢、蠟
器物	交通工具（水路）	水犀、甲板、帆、舟、舟（木蘭舟）、舟（維舟）、舟（獨木舟）、舟（龍舟）、舟船、舟楫、舟 _山 召、舠、柁、柁、舳、舳、舳、航……
器物	交通工具（陸路）	六龍、車（七香車）、車（牛車）、車（羊車）、車（鹿車）、車（雷車）、車（蒲車）、車（摩托車）、車（籌車）、車（薄笨車）、車（鹽車）……
器物	交通工具（空路）	木鳶、火箭、直升機、飛機
器物	交通工具（合稱）	舟車、車船
器物	工具用品（兵器）	刀、刀（刃）、刀匕、刀戟、刀鈇、刀劍、刀鋸、刀錐、刀鎗、刀鐮、匕、干、干戈、干戟、干櫓、弓、弓（弩）、弓刀、弓矢、弓馬……
器物	工具用品（戰爭物品）	刁斗、孔明燈、木牛、火箭、火藥、甲、甲戈、甲冑、戍鼓、角、金柝、金革、金勒、冑、盾、盾戈、狼烽、砲、砲（佛郎機）、砲（炮）……
器物	工具用品（農牧漁獵工）	水車、印床、耒、斧、耒、秧馬、耕、置、畚、犁、耜、釣竿、釣絲、魚叉、魚鉤、耒、筌、筌蹄、農具、農器、耨、罨、鋤、漚……
器物	工具用品（生活工具）	刀尺、刀剪、刀筆、尺、支架、日晷、木鐸、白、把、帚、拍、杵、杵臼、杼、杼軸、俎、指南車、柄、柙、柙、竿、枹……
器物	工具用品（刑具）	炮烙、枷
器物	玉帛服飾（衣冠腰帶）	巾、巾（角巾）、巾（幅巾）、巾（葛巾）、巾（綸巾）、巾（儒巾）、巾（羅巾）、布、布（荊布）、衣、衣（毛衣）、衣（牛衣）……
器物	玉帛服飾（鞋襪配飾）	佩、明珠、花黃、金步搖、金飾、流蘇、韋弦、屐、屐（屨）、屐（蠟屐）、珮、紗翎、珥、笄、釵、釵（玉釵）、釵（寶釵）、釧、黃綿、……
器物	玉帛服飾（合稱）	玉帛、衣巾、衣冠、官服（袍笏）、冠衿、冠帶、冠裳、冠

		履、衫履、笠屐、笠簷、釵裙、簑笠、襟袖、襪履
器物	宗教用品	九轉丹、丹符、丹鼎、丹爐、木魚、水燈、仙藥、玉塵、佛冠、佛座、佛象、佛燈、佛髻、佛龕、佛鉢、念珠、法鼓、社鼓、舍利、金丹、盃琖、香、香案、神水、神杖、神斧、神座、神像……
器物	喪葬祭祀	匕鬯、土牛、少牢、包茅、玉魚、石獸、吉金、柩、香煙、香燭、苦塊、素車、紙錢、芻、祭、奠、棺、瑚璉、鼎鐘、槨、簠、禮器、簠簋、總帳、櫬、靈車
器物	娛樂器具	木馬、木偶、火樹、牙牌、古董、奴戲、打馬、竹馬、陀羅、風箏、鬥草、毬、球、煙火、骰、鞠、繡毬、繡球、鞦韆、爆竹
器物	金帛錢幣	元、文、白鏹、帑、金、金帛、財、貨布、貨泉、鈔、貫、酬、幣、餉、錢、鏹
器物	公堂器物	九錫、仗馬、仙仗、玉檢、白筆、兵符、官符、虎符、金牛、金符、金魚、金牌、帥符、軍符、宮花、笏、符、符節、節、節旄、節鉞、榜、蕊榜、鐘鼎
器物	旗幟帳幕	大纛、帘、幟、幟、幟、旆、旆、旆、旆、旆、帳、帳（穹廬）、帳（芙蓉帳）、帳（絳紗帳）、帳帷、帷、帷幄、旌、幟、幟帳、幄、旒、旒、幟……
器物	神器	雲車
樂部	樂器	缶、角、角（畫角）、弦、弦（三弦）、板、板（牙板）、板（紅牙板）、板（檀板）、律呂、竽、笛、笛（羌笛）、笙、絃、笳、笳（胡笳）……
樂部	樂器合稱	缶瑟、角笳、弦管、笙竽、笙匏、笙笳、笙瑟、笙磬、笙簫、笙簧、笙鏞、笳鼓、琴棋、琴瑟、琴鼓、琴箏、琴劍、琴樽、琴簫、絲竹、絲管……
樂部	樂曲	九韶、八音、大呂、小海唱、小調、山歌、五音、巴歛、水仙操、北管、古調、平調、正音、民謠、白紵歌、皮黃、伊涼、曲、行歌……
空間	方向	八方、中南、北、四方、西、西北、西南、東、東北、東西、東南、南、南北、南東
空間	位置	下、上、上下、中、內、內外、右、外、左、左右、尾、角、底、前、前後、後、背、旁、畔、側、頂、傍、間、隅、裡、鄰、頭、邊
顏色	原本顏色	白、白（皎）、白（皓）、白（蒼）、白（皚）、灰、金、青、青（碧）、青（綠）、青（翠）、青（蒼）、紅、紅（丹）、紅（朱）……
顏色	借代顏色	火色、玉色、白（玉）、白（雪）、白（霜）、黃（麴塵）、暮色、鷺色

		憤愁
感官詞	嗅覺	芳、芬、芬芳、郁、香、臭、腥、聞、薰、馥、馨
感官詞	聽覺	丁冬、叫、吠、音、啁啾、聆、喧、啼、喚、啾、寒砧、聒、蛩、聞、語、鳴、嘯、鬧、聲、譁、囂、響、籟、聾、聽、鈎鞞
感官詞	味覺	甘、味、苦、苦甘、苦辣、甜、甜酸、辣、酸、鹹、鹹苦、鹹酸
感官詞	視覺	見、看、盼、眈、眄、盼、晤、望、眺、眺望、覓、視、睇、覘、睹、睨、瞋、賞、閱、瞥、窺、瞰、矚、覷、瞻、覷、覽、顧、觀

進行步驟及執行成果

- (一) 蒐集整理紙本文獻之音譜，建立數位檔，就現有文獻所存詩的吟唱音譜予以整理，並建立數位檔。
- (二) 將音譜轉成 MIDI 檔，建立哼唱比對之旋律檢索功能。
- (三) 建立唐詩讀音資料庫，提供語音檢索功能。
- (四) 建立唐詩語意標記資料庫，提供語意概念檢索功能。
- (五) 連結全唐詩資料庫，以提供跨資料庫之檢索功能。
- (六) 錄製人聲吟唱以作為教學示範。
- (七) 整合上列資料，建構唐詩吟唱之網路教學環境。
- (八) 進行教學測試與實驗。
- (九) 舉辦教學觀摩會。

教學觀摩會：水調聲長醉裡聽-唐宋詩詞多媒體教學工作坊

(http://cls.hs.yzu.edu.tw/imt/introduction_all.htm)

本計畫執行期間，於 96 年 1 月 27、28 日兩天在元智大學舉辦「水調聲長醉裡聽-唐宋詩詞多媒體教學工作坊」。此教學觀摩會開放給全國高中以上、桃園縣各級學校之華語文專任教師、華語文研究所以上研究生參加，課程內容包含多媒體在文學上的設計與應用、唐宋詞吟唱教學與示範、卡拉迷軟體介紹、詩詞資料庫融入教學、資訊融入教學等實作應用。

網站呈現

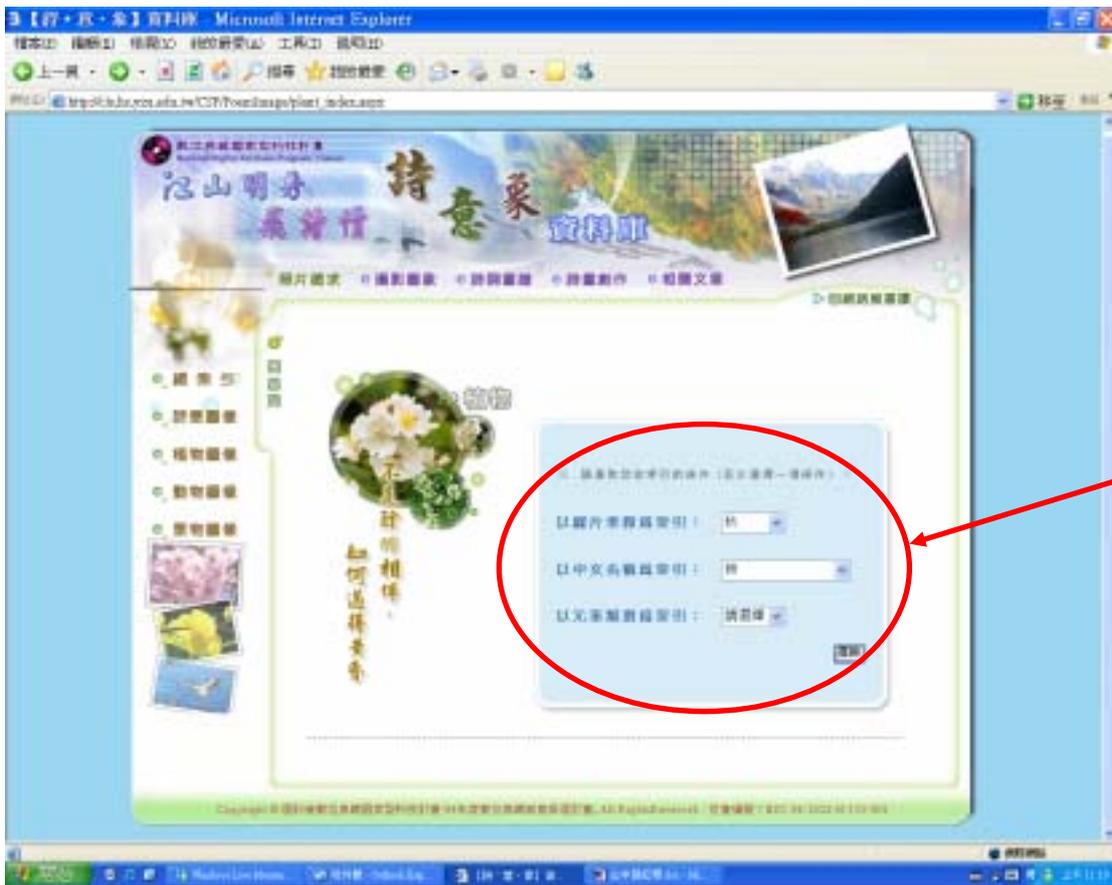
本網站主要分為五大區塊：「聲律對韻」、「唐詩吟唱」、「詩·意·象」、「全唐詩」、「教學區」，方便使用者依照個人需要進入網站。

網站內容包括「全唐詩全文檢索系統」(文字)、「詩意象圖庫」(圖檔)、「唐詩吟唱影音檔」(影音)，使用者可利用索引或者檢索之方式，搜尋所需的資料。網站提供數百首影音檔案，欣賞唐詩同時，更有符合詩文意境的圖檔作配合。預設圖片是由網站製作者主觀意識認定，使用者可依個人意識選擇更為適合之圖檔作搭配。

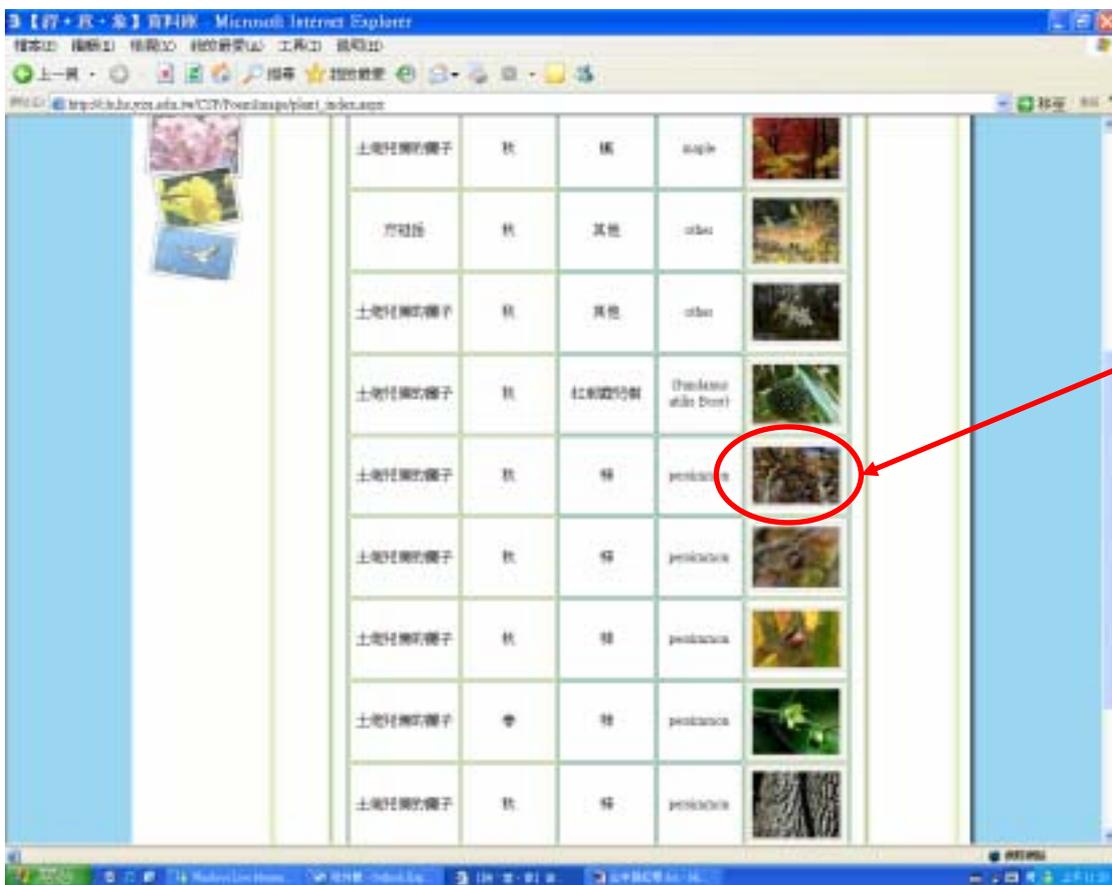
詩意象圖庫提供攝影圖像有 1,774 張，詩詞畫譜有 837 張，圖檔資料經過描述，使用者如需尋找

適當的圖檔作為輔助教材，即可依照描述欄位所列出的概念搜尋。屬於知識性的資料，是客觀的，比較單純，容易找到對應資料；以張籍〈岳州晚景〉：「晚景寒鴉集，秋聲旅雁歸。水光浮日去，霞彩映江飛。洲白蘆花吐，園红柿葉稀。長沙卑濕地，九月未成衣。」為例，可以從「植物類」尋找「柿」圖檔（步驟如下圖所示）；屬於情境性的圖檔，容許主觀的判斷與選擇，比較複雜，需要以概念搜尋，以朱熹〈觀書有感〉：「半畝方塘一鑑開，山光雲影共徘徊。問渠那得清如許，為有源頭活水來。」為例，若要尋找符合〈觀書有感〉情境的圖檔，可以從「景觀」類搜尋。從經過描述的圖檔，逐項循「地景」選擇「山嶺、溪流」（山、水）、「地貌」選擇「倒影」、「天候」選擇「雲」（雲影）、「時間」選擇「白天」（光），即可將圖檔縮小在符合條件的範圍裡，迅速找到適當的圖檔。換言之，每一張圖檔，適合詮釋那些文學作品，在設計時也保留了由使用者自由搭配的空間。





選擇季節(秋)、植物名稱(柿)、元素(果)等。可擇一輸入。



出現的圖庫清單中，選擇最符合的圖片。



左側為檢索出的「園红柿葉稀」。

全唐詩全文資料庫分為「文字檢索」與「語意檢索」兩類，文字檢索採用布林檢索方式，可應用於更強大的資料搜尋功能上，利用「&」=AND、「/」=OR、「—」=NOT 的區分，得到更完整的檢索結果。不僅僅是純文字的檢索，結果頁面更提供「詞彙、詞句的標記」以及「自動詞彙切分與語意標記」等功能，提供了多元化的檢索系統。

網站基本架構

首頁以李白的〈關山月〉為構思背景，遠山、蒼雲、瑩月，呈現大唐詩風壯闊的氣象，首頁詩文引自清·趙翼的〈論詩絕句〉：「李杜詩篇萬古傳，至今已覺不新鮮。江山代有才人出，各領風騷數百年。」點出首頁圖中兩位引領唐詩的一代才人-李白、杜甫，作為唐詩內容的表達意象。網站首頁如下。

新詩改羅自長吟 - 唐詩吟唱虛擬環境教學環境 - Microsoft Internet Explorer

檔案(F) 編輯(E) 檢視(V) 我的最愛(S) 工具(T) 說明(H)

地址(A) <http://nl.ku.yu.edu.tw/anglato.html> 轉主 網站

教巨典國學堂科技外意

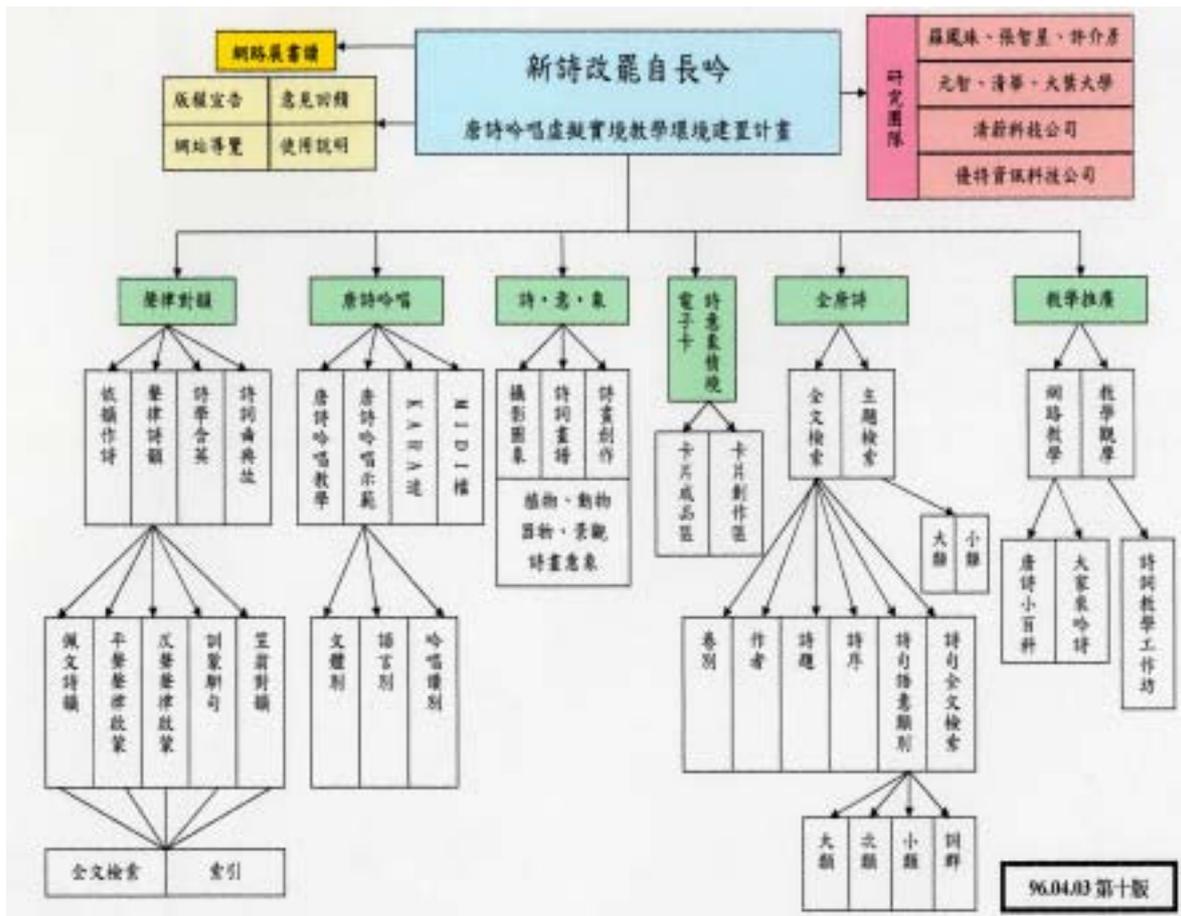
新詩改羅 自長吟

唐詩吟唱虛擬環境教學環境

雲律對韻 唐詩吟唱 詩·意·象 全唐詩 教學區

李杜詩篇萬古傳 至今已覺不新鮮
江山代有才人出 各領風騷數百年

網站架構圖



各主要頁面功能

- 一、聲律對韻：分為依韻作詩、聲律詩韻、詩學含英、詩詞曲典故等部分。依韻作詩收錄五、七言絕句詩譜及律詩詩譜。選擇詩體類別後，依該類的相關韻格點選進入，即可開始詩韻創作，作品完成後，系統將會自動檢查格韻。聲律詩韻收集了《聲韻啟蒙》、《訓蒙駢句》、《笠翁對韻》三本訓練對偶技巧、聲韻格律的工書。按韻分部，使用者依據檢選進入，從中可得語音、修辭的訓練。詩學含英系統根據《增廣詩韻全璧》一書所附〈詩學含英〉建立，旨在提供初習作詩者參考。本系統提供使用者依類別檢索，以做為作詩填詞時應用典故之參考。詩詞曲典故系統，使用者可選擇以索引或檢索方式查詢典故。
- 二、唐詩吟唱：以洪澤南老師的吟讀教學及傳統八音再現兩部分，提供使用者線上學習。
- 三、詩·意·象圖庫：本系統中包含詩意象圖庫，及新開發的詩意象情境卡。
- 四、詩·意·象情境卡：情境卡的設計主要以提供使用者將內心之情與外物之景，配合詩文，傳達言有盡而意無窮的情感，分為「卡片成品區」與「卡片創作區」，前者提供搭配好詩詞圖像的卡片，後者提供使用者自行搭配詩詞圖像以致作卡片的功能與平台。
- 五、全唐詩：即「全唐詩全文資料庫」，提供全唐詩全文檢索系統。
- 六、教學區：內容含有閱讀唐詩的入門小百科，以及線上教唱部分，提供使用者快速進入唐詩世界的環境。
- 七、網站導覽：呈現網站分層架構。
- 八、研究團隊：提供計畫所有參與人員的資訊，包括計畫主持人、共同主持人及研究助理、兼任助理等的基本介紹、個人網頁。
- 九、使用說明：依網站各區域以圖檔方式詳細說明網站的使用方法。

十、意見回饋：使用者可提出任何建議，以供研究小組改進之參考。

十一、版權宣告：說明合理使用的範圍、引用註記規則及版權來源。

十二、各站連結：首頁左上方之「數位典藏國家型科技計畫」及下方「網路展書讀」皆可連結至計畫主持人歷年執行的網站，與數位典藏總網站。

對於學術研究、國家發展及其他應用方面預期之貢獻

國內各級學校的任課老師，都能了解詩的吟唱教學之重要，但苦於缺乏適當的教材可用，能於課堂親自吟唱教學的老師並不多。因此，結合多媒體及最新的軟體與音訊處理技術，於網路上建置唐詩吟唱教學環境，對唐詩的吟唱教學將有很大的幫助。

成果發表

✚ 羅鳳珠、張智星、許介彥（大葉大學）〈植基於語意學及使用者認知行為的資訊檢索系統設計：以全唐詩網站為例〉，第三屆文學與信息科技國際研討會，2007年3月19-20日，日本學藝大學主辦，會議地點：東京。

結果與討論

本計畫主要的目的是引用資訊科技，在學術網路上建立唐詩吟唱的數位學習環境，讓最古雅的中國詩樂與最新的資訊科技結合，以推廣並傳承唐詩及唐詩吟唱的教學，並建立唐詩詞彙切分及語意標註，建立資訊系統在文學領域的自動學習功能，引用資訊學門所發展的專家系統，提高文學引用資訊科技的使用效益，嘗試為文學的研究與資訊科技結合的過程中拓展出新的研究方向，發掘新的研究議題，並為難以單憑人力完成的文學研究議題，從資訊工具中尋找解決的方法。

國內各級學校的任課老師，都能了解詩的吟唱教學之重要，但苦於缺乏適當的教材可用，能於課堂親自吟唱教學的老師並不多。本計畫執行期間，舉辦「水調聲長醉裡聽-唐宋詩詞多媒體教學工作坊」。此教學觀摩會課程內容包含多媒體在文學上的設計與應用、唐宋詞吟唱教學與示範、卡拉迷軟體介紹、詩詞資料庫融入教學、資訊融入教學等實作應用。於網路上建置唐詩吟唱教學環境加上成果推廣活動，相信對唐詩的吟唱教學將有很大的幫助。

參考文獻

專書：

1. 李勉著，《詞曲概論及精選評註》（附唐詩唱譜），台中人文出版社，1982年。
2. 李勉著，《詩詩作法唱法通考》，台中人文出版社，1982年。
3. 李勉著，《唐詩吟唱考定》，金石有聲出版公司，1988年出版。
4. 邱燮友：《品詩吟詩》，台北，東大出版社，1989年。
5. 高嘉穗：《台灣傳統吟詩音樂研究》，台灣師範大學音樂研究所碩士論文，1996年。
6. 傅雪漪著，《中國古典詩曲譜選釋》，北京中國戲劇出版社，1996年出版。
7. 黃冠人：《名詩吟唱》，台北，樟樹出版社，1997年。
8. 傅雪漪著，《傅雪漪古典詩詞配樂吟唱曲》，臺北市臺灣書店，1998年出版。
9. 冉休丹編寫，《唐詩吟誦》，北京中華書局出版，1998年出版。
10. 洪澤南：《大家來吟詩》，台北，萬卷樓圖書公司，1999年。
11. 劉崇德、孫光鈞譯譜，《碎金詞譜今譯》，中國大陸河北大學出版社。2000年1月出版。

12. 劉崇德譯譜，《唐唐詩古樂譜百首》，中國大陸河北大學出版社，2001年5月出版。
13. 潘麗珠：《古韻新聲——潘麗珠吟誦教學》，台北，幼獅出版社，2003年。

論文：

1. 葉嘉瑩：〈談古典詩歌中興發感動之特質與吟誦之傳統〉，《中外文學》21卷11期，1993年4月。
2. 潘麗珠：〈詩歌「吟、唱、誦、讀」的觀念及要領〉，《國文天地》107期，1994年4月。
3. 潘麗珠：〈古典詩歌聲情藝術及其美學義涵〉，國立台灣師範大學國文學報30期，2001年6月。
4. 簡明勇：〈詩歌吟唱教學〉，《國文天地》193期，2001年6月。
5. 潘麗珠：〈古典詩歌聲情的建構教學〉，人文及社會學科教學通訊76期，2002年12月。
6. 莫月娥：〈古典詩吟唱經驗談〉，收錄於《大雅天籟——莫月娥古典詩吟唱專輯》，台北，萬卷樓圖書公司，2003年1月。
7. 陳燕秋著，《資訊科技融入中學國文教師教學現況之研究》，國立交通大學傳播所碩士論文，2003年。
8. 王更生，〈什麼是古典詩詞吟唱藝術〉，2004年9月，《國文天地》
9. 王更生，〈古典詩詞的「吟」和「唱」〉，2004年11月，《國文天地》
10. 王更生，〈古典詩詞吟唱的幾個觀念〉，2004年12月，《國文天地》
11. 王更生，〈古典詩詞吟唱的價值〉，2005年1月，《國文天地》
12. 王更生，〈古典詩詞吟唱前的準備工作〉，2005年2月，《國文天地》
13. 王更生，〈古典詩詞吟唱的三點要素〉，2005年3月，《國文天地》
14. 王更生，〈古典詩詞吟唱的技法 -1〉，2005年4月，《國文天地》
15. 王更生，〈古典詩詞吟唱的技法 -2〉，2005年5月，《國文天地》
16. 王更生，〈古典詩詞吟唱在歷代學者心目中的地位〉，2005年6月，《國文天地》

九十五年度研究成果資料表

日期 96 年 4 月 2 日

國科會補助 計畫	計畫名稱：新詩改罷自長吟-唐詩吟唱虛擬實境教學環境建置計畫 計畫主持人：羅鳳珠 共同主持人：張智星、許介彥 特別協助：優特資訊科技有限公司、清蔚科技公司 計畫編號：NSC 95-2422-H-155-001
研究摘要	<p>本計畫主要的目的是引用資訊科技，在學術網路上建立唐詩吟唱的多媒體數位學習環境，並與唐詩結合，讓最古雅的中國詩樂與最新的資訊科技結合，以推廣並傳承日益衰微的唐詩吟唱古曲與吟唱方法，免使成為絕響。</p> <p>音樂與文學是人類各種藝術中最具有抒情性，最易於感人的藝術形式。詩與樂本來就是同源的藝術，都是人心感於外物的情意表現。張騫通西域，引進很多西方樂器，為中國詩樂結合藝術邁入新的階段，唐詩與音樂的結合，也因為新樂器的加入而有新的面貌。</p> <p>以吟調吟誦唐詩，可以欣賞唐詩聲情之美，還可以協助背誦，提高學習的興趣，本計畫擬在九十四年度「忍把浮名換了淺斟低唱：宋詞古唱虛擬實境教學環境建置計畫」的基礎，以及「卡拉迷：淺斟低唱」的技術與模式，結合「全唐詩」網站、「依韻入詩」網站，建置「新詩改罷自長吟：唐詩吟唱虛擬實境教學環境建置計畫」，提供唐詩教學使用。</p>

	<p>The goal of this project is to combine the latest technologies from digital signal processing, music information retrieval, multimedia rendering, and speech recognition to construct an integrated environment for teaching/learning the chanting of Tang poetry over the Internet. The constructed system can effectively preserve the music information (ancient music scores) of Tang poetry. More importantly, it can promote the chanting of Tang poetry for the young generation.</p> <p>Music and literature are two of the most heart-touching parts of all kinds of art formats. The chanting of Tang poetry is the combination of these two art formats, which represent an important milestone of traditional Chinese art for many generations.</p> <p>The use of multimedia tools for the chanting of Tang poetry can not only lead the young generation to appreciate the beauty of Chinese literatures, but also help them to memorize the Tang poetry with great interests. With the foundation of previous related projects, we are aiming at constructing an integrated environment for preserving the music scores of Tang poetry, and for learning/teaching the chanting of Tang poetry. This is an important task toward preserving the traditional Chinese art of music and literature, and presenting it to the young generation in a totally new manner to let them appreciate the beauty of Tang poetry.</p>
<p>研究特色</p>	<p>本計畫在保存古譜文獻之外，以輔助唐詩教學以及推廣唐詩吟唱為主要目的，引用資訊科技之最新軟體與音訊技術（卡拉迷）結合，將唐詩古譜之旋律重現，除使自學、教學兩便之外，與全唐詩資料庫連結，讓使用者可以自選吟唱之詩作，不必受限於只能吟唱某些詩作，將使教學與自學的自主性更高。</p> <p>國內大學中文系詩選課程與中學國文課程之任課老師，都能了解詩的吟唱教學之重要，但苦於缺乏適當的教材可用，能於課堂親自吟唱教學的老師並不多，能一邊操琴一邊教唱的老師更少，中國大陸之大學與中學也面臨同樣的困窘。因此，結合多媒體及最新的軟體與音訊處理技術，利用時下流行之 KTV 點歌方式，建置唐詩的吟唱教學環境，對唐詩的吟唱教學將有很大的幫助。</p>

<p>研究成果可 應用範圍</p>	<p>可提供各級國文科教師於課堂上利用電腦教唱唐詩，學生也可逕自上網 download 卡拉迷軟體在家中免費歡唱唐詩。詞彙切分以及語意概念更能協助教師、研究者，甚至一般使用者針對唐詩詞彙使用情形，做更進一步的分析、研究。</p>
<p>研究成果預期 效益／商機</p>	<p>於網路上建置具有樂音伴奏及真人吟唱示範之唐詩吟唱多媒體數位教學環境，將現存音譜數位化，建立唐詩音譜資料庫。並建立唐詩讀音音標資料庫，於文字檢索功能之外，提供以語音檢索之功能，使用者輸入任何字詞之文字或讀音即可檢索。整合語音辨識與旋律辨識，以開發唐詩古唱的「智慧型多模式檢索系統」。</p>

附件

第三屆文學與資訊科技國際研討會

日本學藝大學主辦

2007年3月19-20日

植基於語意學及使用者認知觀點的詩詞資訊檢索系統設計：

以全唐詩網站為例

台灣元智大學中國語文學系·羅鳳珠、清華大學·張智星、大葉大學·許介彥

電子郵件：gefjulo@saturn.yzu.edu.tw

摘要

電腦處理有規則及合邏輯的資料，其效率高於人腦何止千萬倍，但是電腦只能分辨字形的符碼，無法解讀字詞含義的意碼，所以引用電腦做為偏重主觀抽象感知的詩詞文學研究輔助工具，比相對客觀的自然科學領域困難得多。

電腦縱然有強大的記憶與搜尋能力，因為只能分辨字形，無法理解語意內涵，不具有思考能力，不具有人的認知、感知能力，不具備人的知識體系，以之建構詩詞網站，做為詩詞研究或教學的輔助工具，無法滿足文學領域使用者的需求。

本文從詩詞語言的特性分析詩詞文體的資訊檢索需求，以全唐詩網站系統的設計為例，探討如何設計一套植基於語意學及符合使用者認知觀點及資訊行為的全文檢索系統，透過語意標記，使電腦從只能分辨字形提升等能理解詞形詞義，建構語意概念分類體系，使電腦從索引式的搜尋提升到更符合人類大腦以思維聯繫尋找資料的模式，使之更接近人的知識體系，滿足不同使用者的需求。

關鍵詞：全唐詩網站、語意標記、語意概念分類、認知觀點、資訊檢索

全文

一、 前言：

詩詞的本質與電腦存在著基本上的差異。詩詞感性主觀，講究「只可意會而不可言傳」(劉大魁〈論文偶記〉)、「不著一字而盡得風流」(司空圖《詩品》)、「言有盡而意無窮」(嚴羽《滄浪詩話·詩辨》)，容許反理而合情；電腦理性客觀，處理有規則而合邏輯的資料。

詩詞所使用的語言與其他文體不同，北京大學中文系袁行霈教授在《中國詩歌藝術研究·原序》論及：「如果從語言學的角度給詩歌下一個定義，不妨說詩歌是語言的變形，他離開了口語和一般的文書語言，成為一種特異的語言形式。」¹。袁行霈教授的論述是從語意、語用及語境的角度提出詩歌語言與其他文體語言的不同，包含詩歌語言的多義性、意象與意境的創造等等。²葉聖陶從詩歌語言形式的角度，認為詩歌的語言純出於人工。³由此可知，無論從語意或語言形式的角度，詩歌非但與一般口語、文書語言不同，與其他文體也存在著差異，再加上詩歌的形式受到詩歌格律之平仄聲調及字數、句法上的限制，語言更精鍊，語意更豐富，修辭更講究，語境更複雜，需要更多聯想，解讀更困難。

電腦處理有規則及合邏輯的資料，其效率高於人腦何止千萬倍，但是電腦只能分辨字形的符碼，無法解讀字詞含義的意碼，所以引用電腦做為偏重主觀抽象感知的詩詞文學研究輔助工具，比客觀知識的自然科學困難得多。

文學的研究，涉及龐大的資料，詩詞亦然。在浩瀚的詩詞文獻中要具備旁徵博引，互相引證發明的能力，往往需要經年累月，皓首窮經才有可能。電腦強大的記憶能力及快速的搜尋檢索、分析統計能力，可以彌補人腦記憶與搜尋能力的不足。但是，電腦只能分辨字形符，無法解讀字詞語意的意符，旁徵博引及引證發明涉及知識的內涵，而不是文字的字形，電腦縱然有強大的記憶與搜尋能力，因為只能以字形分辨，無法辨別語意內涵，不具有思考能力，不具有人的認知、感知能力，不具備人的知識體系，無法滿足使用者的需求。

本研究將以全唐詩網站為例⁴，探討如何設計一套植基於語意學及符合使用者認知觀點及資訊行為的全文檢索系統，以提升檢索效能，使之更接近人的知識體系，滿足不同使用者的需求。

¹ . 袁行霈，《中國詩歌藝術研究》，五南圖書出版公司印行，1989年5月台灣初版，見〈原序〉第2頁。

² . 參見袁行霈，《中國詩歌藝術研究》之〈中國古典詩歌的多義性〉，第3-24頁、〈中國古典詩歌的意境〉第25-56頁、〈中國古典詩歌的意象〉第57-72頁。

³ . 參見王力，《王力詩論》之〈略論語言形式之美·詩的語言〉，引葉聖陶給王力的書信：「詩之句型，大別為二。一為平常的句型，與散文及口頭語言大致不異。一為特殊句型，散文絕不能如是寫，口頭亦絕無此說法，可謂純出於人工。」

⁴ . 參見羅鳳珠、張智星、許介彥，【新詩改罷自長吟—唐詩吟唱虛擬實境教學環境建置計畫】，國科會數位典藏國家型科技計畫：95年度數位典藏創意學習計畫，2006年3月1日至2007年2月28日。計畫編號：NSC95-2422-H-155-001，網址：<http://cls.hs.yzu.edu.tw/tang/index.html>。

二、 資訊檢索 (Information Retrieval) 的發展歷程

很多學者認為：Vannevar Bush (麻省理工學院副校長) 於 1945 年所發表的 “As We May Think” 一文所提出的 Memex 構想，可視為資訊科學發展的起源。⁵ Vannevar Bush 在文中提出：科學使得人際間的交流非常的快速，科學也提供了思想的記錄，使人類能夠運用這些思想，知識就不再限於個人，而是屬於人類的所有生命。人類的經驗正以驚人的速度擴展，這些記錄必須被採用，如何去使用這些快速成長的知識，是 Vannevar Bush 文中所關心的問題。記錄思想的文獻儲存在圖書館，圖書館的資料按照字母或數字排列的索引，只能一層一層往下尋找，不符合人類大腦以思維聯繫尋找資料的方式，因此，Vannevar Bush 便構思著建立一個以聯繫 (association) 代替索引 (indexing) 以選擇資料的機械化個人文件夾和圖書館 (mechanized private file and library)，Vannevar Bush 將之命名為 Memex。Vannevar Bush 期待在這個機器裡可以儲存所有的文獻資料，檢索時能如同人類的思維邏輯，建立知識連結的軌跡以進行聯繫索引，需要引用時，所有的資料都可以隨意檢索出來，並且可以快速的找到另一個相關的資料，也可以二個條件同時並行檢索，這就是 Memex 的基本概念。⁶ Vannevar Bush 的這篇文章，被認為是第一篇論及超文件基本觀念的文獻，啟發資訊科學家對資訊檢索技術之研究。

資訊檢索的產生，導源於對資訊的需求。資訊無論以任何物質／非物質載體、任何媒體形式呈現，從資訊產生者到資訊接收者之間，都會有認知上的落差，提取方法的不同，而產生給予、接受與取得的障礙。人類為了滿足資訊的需求 (need)，所引發的資訊行為包含瀏覽 (Browsing)、搜尋 (Seeking)、檢索 (Retrieval)，為了提升資訊檢索快速、正確、完整的效能，及符合資訊行為的模式，而發展出不同的資訊檢索系統。

瀏覽行為被應用在資訊檢索系統設計的考量因素之中，因而發展出行為取向的檢索系統設計模式。J. Krikelas 在 1983 年所發表的 “Information-seeking behavior: Patterns and concepts”⁷ 文中將瀏覽行為分為「開始」(Starting)、「連結」(Chaining)、「瀏覽」(Browsing)、「辨識」(Differentiating)、「監看」(Monitoring)、「提取」(Extracting) 等六部分。行為人從他人處、文獻評述及被評述的文章、圖書館目錄、摘要和索引等處作為瀏覽的起點，以反向連結 (Backward chaining)、向前連結 (Forward chaining) 的方式進行資訊的連結，並在適當的時候收網 (Closure)，在半指導 (Semi-directed) 或半結構 (semi-structured) 的資料呈現型態中瀏覽檢索，從「獨立的標題」、「整體洞察的方向」、「品質、層級和處理的模式」等三個角度對資料加以辨別，從「非正式的溝通或其他交流管道」、「期刊」、「出版品」監看是否有其他資料，最後從期刊、公共目錄、書目、索引、提要擷取所需要的資料，這個過程構成瀏覽行為的模式。⁸ 爾後，David Ellis 於 1989 年所發表的 “A Behavioral Approach to Information Retrieval System Design”⁹ 論文中，將這六種瀏覽模式發展為資訊檢索系統設計的六個處理步驟。

⁵ . V. Bush, “As we may think” The Atlantic Monthly, Vol.176, No.1 pp.101~108, 1945

⁶ . 從 “As we may think” 原文摘錄重點。

⁷ . J. Krikelas, “Information-seeking behavior: Patterns and concepts”, Drexel Library Quarterly, 19(2), p. 7-17, 1983.

⁸ . 從 J. Krikelas, “Information-seeking behavior: Patterns and concepts” 文中摘錄重點。

⁹ . David Ellis, “A Behavioral Approach to Information Retrieval System Design”, The Journal of

三、 詩詞文體語言的特性所衍生的資訊檢索需求

就如詩評家及學者們所提出的觀點，詩詞是特殊的語言形式，以不同的修辭技巧達到精緻凝練、語意無窮，情境高遠、言意與形神兼具的境界。透過修辭的技巧，將作者主觀內化的情思意念，投注在客觀外顯的事物上，作為作者與讀者之間引發共鳴的溝通媒介。詩詞所使用的詞彙，有一些是客觀知識的描述闡釋，有一些是主觀情意的引導感發。以楊花為例，蘇軾〈少年遊〉：「去年相送，餘杭門外，飛雪似楊花。今年春盡，楊花似雪，猶不見還家。」描寫冬季飛雪與春末飛花的景象，將「楊花」與「飛雪」對比呈現，讓只見過其中一種物象的讀者，多出一項理解詞義的管道，這裡的楊花是純粹客觀的景物，不含作者的情意。同是寫楊花，蘇軾寫楊花的名作〈水龍吟〉：「似花還似非花，也無人惜從教墜。拋家傍路，思量卻是，無情有思。縈損柔腸，困酣嬌眼，欲開還閉。夢隨風萬里，尋郎去處，又還被鶯呼起。不恨此花飛盡，恨西園、落紅難綴。曉來雨過，遺蹤何在？一池萍碎。春色三分，二分塵土，一分流水。細看來、不是楊花，點點是離人淚。」多重意象交互呈現，幽居空閨的女子，自傷處境如同隨風飄盪的楊花，拋家傍路，無人憐惜，不知情歸何人，不知郎行何處，多少幽微的心事無處訴說，隨風飄落的片片楊花，看在傷心人眼中，盡是點點離人淚，這是心境、情境、意境。¹⁰

主觀情意的感發，往往因為讀者的學養、經驗、年齡、心理氣質、思想境界等各種背景差異而使感受各有不同，因此相同的詞彙，相同的作品，感受上存在者個別的差異，即便是同一個人，經過不同時空，重讀相同的作品，感受也會不同，這是客觀認知與主觀感知的問題。

資訊的瀏覽、搜尋、檢索等行為，牽涉到行為人的認知，Belkin, N. J. 在1990年發表的“The cognitive viewpoint in information science”¹¹提出以認識論為基礎的認知觀點；Ingwersen, P. 在1992年出版的“Information Retrieval Interaction”¹²書中的“The Cognitive View and Information”、“IR Interaction?the Cognitive Turn”二個章節論及認知觀點；在1996年發表的“Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory”¹³以及2001年發表的“Cognitive information retrieval.”¹⁴進一步將認知觀點與資訊檢索結合，提出資訊檢索的交互作用以認知理論的元素為基礎的觀點。Ingwersen, P. 以“General model of cognitive information seeking and retrieval”架構圖反映出對資訊搜尋和檢索（IS&R）的資訊搜尋之行為者或在搜尋情境裡的群組的認知過程，牽涉到認知和感知的表述，如人類認知、反應、感情、意見的表述，這些都是資訊搜尋和檢索構成的一部份，也是資訊交互影響情境的一種。¹⁵架構圖如下：

Documentation, 45(3), pp. 171-212, 1989

¹⁰ . 參見拙作〈虛擬與實境的交錯：多媒體在文學教學的設計與應用〉，文學數位製作與教學研討會，成功大學主辦，2005年6月12日，收入《文學數位製作與教學》，五南圖書公司，P107~P130，2007年1月出版。

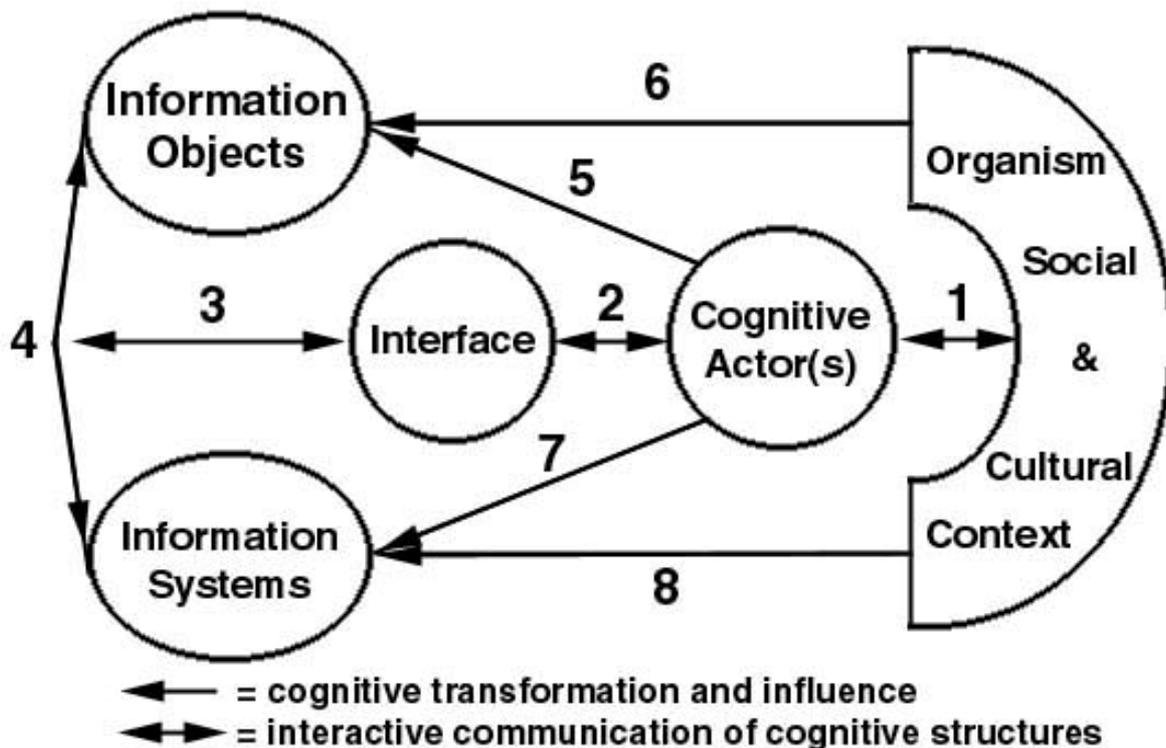
¹¹ . Belkin, N. J., “The cognitive viewpoint in information science”, *Journal of Information Science: Principles and Practice*, 16(1), 11-15. 1990

¹² . Ingwersen, P. “Information Retrieval Interaction”, London: Taylor Graham, 1992. X, 246 p.

¹³ . Ingwersen, P. “Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory”, *Journal of Documentation*, 52(1), 3-50. 1996.

¹⁴ . Ingwersen, P. “Cognitive information retrieval”, In: M. Williams(Ed.), *Annual Review of Information Science and Technology*, 34, 3-51. 2001.

¹⁵ . 引自 Ingwersen, P. “Integrative Framework for Information Seeking and Interactive Information Retrieval”：「The framework reflects the understanding that IS&R is a process of cognition for the



從這個架構圖可知，組織、社會和文化背景，會影響到資訊搜尋及檢索行為人的認知，行為人的認知又會影響到所有的資訊行為過程。余光中的〈隔水書〉詩，正可以用來說明形符與意符的差別，以及對行為人（閱聽者）所產生的認知落差。說明如下：

余光中的〈隔水書〉詩：「長風遠雲，七百里一望是水程，隔水宛轉望一個仙人。島是蓬萊，鷗非青鳥，空郵箋傳幾日的苦眺。委委曲曲一句話，紙摺幾摺就話摺幾摺就心折幾折，要拆要輕輕的拆。那仙人歎口氣，也罷。你心焦水西湄，我心焦東浦，念來逼人似烽火，隔水傳郵隔一宿，看你煉晚霞對我燒日出。」這一首表達愛慕之情的抒情詩，詩中描寫有一個人寫一封信向心中的仙人宛轉表達愛慕之意，將信託給海鷗傳送，幾日的苦眺，盼回的卻是一封空郵箋。寫信人或因含蓄，或因唯恐唐突佳人，不敢將情思直言明說。當對方回以空郵箋，寫信人擔心對方無法理解自己信中曲折委婉之言，於是再寫一封信告訴對方「紙摺幾摺就話摺幾摺就心折幾折，要拆要輕輕的拆。」仙人以空郵箋回覆，原是讓寫信人知道，你我之間原可一切盡在不言中，不必明說，既然你不放心，仙人不說則已，要說則索性清清楚楚的寫著「也罷。你心焦水西湄，我心焦東浦。」用重複的詞彙「心焦」來說明自己的心境與對方完全一樣。

寫信人用「紙」、「話」、「心」三種由具體到抽象的詞彙排比呈現，「紙」是具體客觀的物，認知上不會有落差，相當於形符；「心」是主觀抽象的感受，認知上會有落差，相當於意符；「話」介於「紙」與「心」之間。一段話或一段文字，說者與聽者之間，雖有語音或文字作為具體客觀的表述媒介，但是說者的原意與聽者的感受難免有客觀上的認知與主觀上感知的落差。寫信的人

information-seeking actor(s) or team in context. Algorithmic and IIR, as well as information seeking (IS), involve cognitive and emotional representations are seen as manifestations of human cognition, reflection, emotion, or ideas forming part of IS&R components and kinds of interaction in context', In K. E. Fisher, S. Erdelez & L. Mckechnie (Eds.), *Theories of information behavior* (The Theories 35, pp215-220). Medford, N. J.: ASIST

以這種方式排比呈現，是要告訴仙人，我的「心」折幾折，雖然你看不到，你只要聽（看）「話」摺幾摺，如果還是不夠明確，再看「紙」摺幾摺即可逐一回溯我的「話」摺幾摺，我的「心」折幾折。「紙」是具體客觀的，你看我看他看都不會有誤差，寫信的人以「紙摺幾摺就話摺幾摺就心折幾折」引導收信的仙人，循此路線追溯寫信人千迴百轉的曲折心事。仙人再次回信時，要表達自己「心焦」的心情與對方一樣，但是「心焦」是抽象的感知，每個人對「心焦」的感受不同。詩詞原避免使用重複的詞彙，作者使用完全一樣的詞彙，仙人要表達的是「無論你有多心焦，我就有多心焦」，無論你對「心焦」的感受是甚麼，總之君心似我心，我心如君心。仙人回信使用同樣的詞彙，為的就是要避免彼此認知與感知上的落差。

文學表達的情感通常是主觀而抽象的，即便是藉助客觀的物象作為作者與讀者之間溝通的媒介，仍然存在著感知層次的不同，因此，資訊檢索的技術，應用在文學上，其困難度比自然科學領域高得多，應用在特殊語言形式的詩詞，困難度又高於小說、戲劇、散文等文體。

語意的標記，可以解決電腦只能分辨字形，無法理解字詞義的問題，但仍然與人的判斷解讀存在著相當大的差距。將文本經過詞彙切分及語意標記，使電腦能理解詞彙結構及語義，建立語意概念的分類及上下位詞層級的結構，使之更接近人的知識體系，有助於提升資訊檢索的品質。

四、傳統文學網站資訊檢索功能及缺點

傳統的文學網站檢索都只能比對字形，不能分辨詞形、詞義。文學以文字構成詞組，以詞組構成篇章，中國的文字多數是多義字，構成詞組之後的變化更多，文學講究修辭變化，講究文采意境，一詞多義，一義多詞的現象比比皆是。只能比對字形的檢索，檢索結果的完整性及正確都不足，不能滿足一般的使用需求，對於文學領域以研究及教學為目的的使用者而言，尤其不足。以「月」為例，「月」常用的基本含義有四種，其一是月亮，其二是曆法單位，其三是季節代稱，其四是圓形如月亮的物品，這是一詞多義。「月」若作「月亮」解時，同義詞有「月亮＝玉兔＝銀蟾＝玉杯＝玉鉤＝玉盤＝玉輪＝玉蟾＝生魄＝冰輪＝兔降＝孤光＝明霽＝金波＝金輪＝桂華＝桂魄＝素娥＝掩夜＝清光＝寒暉＝嬋娟＝瓊瑤＝蟾影……」，這是一義多詞。只能比對詞形的全文檢索，既無法分辨一詞多義的關係，也無法分辨一義多詞的關係。使用者檢索資料，以語意為取向，不是以字形為標的，一詞多義會使檢索時出現字形、詞形相符但語意不符的資料，影響正確性；一義多詞會遺漏很多語意相同但字形、詞形不同的資料，影響完整性。

影響「比對字形檢索」品質的另一個因素是電腦對於詞組判斷的正確性不足。在未經過詞彙切分的句子裡，電腦無法分辨詞組的構詞單位。某一個字與前一個字的組合，或與後一個字的組合，會成為不同的詞組，含有不同的詞義，人在閱讀的過程可以依據自己的知識分辨，但是電腦若不具有人的知識體系便無法判斷，所以檢索時會產生因為無法正確判斷詞組而造成檢索結果錯誤的情形。以「春風」為例從全唐詩網站檢索¹⁶，共得 1,182 筆資料：劉長卿〈使回赴蘇州道中作〉詩：「春風何事遠相催，路盡天涯始卻回。」王維〈不遇詠〉詩：「且此登山復臨水，莫問春風動楊柳。」指的都是春風。若是未經過斷詞，則薛稷〈奉和聖製春日幸望春宮應制〉詩：「九春／風景足林泉，四面雲霞敞御筵。」白居易〈喜夢得自馮翊歸洛兼呈令公〉：「已將四海聲名去，

¹⁶ . 同註 4。

又占三春／風景來。」白居易〈三月三日〉：「暮春／風景初三日，流世光陰半百年。」曹唐〈小游仙詩，九十八首之八十〉：「玉洞長春／風景鮮，丈人私宴就芝田。」等 4 筆與春風無關的詩句，都會被檢索出來。

由以上分析可知，電腦除了無法分辨字詞義，也無法判斷詞組的構詞單位，因此，要提升資訊檢索的正確性與完整性，正確的詞彙切分及語意標記是最重要的關鍵。

在進行語意標記之前必須先做詞彙切分，詞彙切分也是非常重要的步驟。詩詞以二字詞居多，五言詩的句法多數是上二下三，下三部分作上二下一，或作上一下二，而成為二二一或二一二的句法；七言詩則多數是二二一二或二二二一句法。五言詩的前二個字及七言詩的前四個字，除非是專有名詞或典故，鮮少使用單字或三字詞¹⁷，因此依據詩的句法做適切的斷詞，對於「正確的檢索出所需要的資料」是非常重要的步驟。

五、全唐詩的分類：以詩的主題分類

在還沒有資訊檢索之前，書籍作者或編纂者因應不同的使用需求，將書籍的內容做不同的分類，來幫助使用者快速方便查檢所需要的資料。孔子刪詩書，訂禮樂，將《詩經》按照風、雅、頌分類，可說是文本篇什分類的起源，《佩文韻府》以韻分類，《唐詩別裁》、《清詩別裁》以詩體分類，分類的用意不外乎為了方便查檢瀏覽。

全唐詩網站的文本內容包括清代彭定求奉敕編纂的《全唐詩》加上陳尚君輯校的《全唐詩補編》，總計收有 2,821 位詩人，51,170 首，4,767,979 字的作品，可謂卷帙繁複。面對這麼繁複的唐詩，使用者要查檢其中的詩作，非常的困難。宋代趙孟堅曾為唐詩做過分類，可惜「佚闕不完，世無刊本」¹⁸，明代張之象「因復有此作（《唐詩類苑二百卷》），凡分三十六部，以類隸詩。」¹⁹1992 年四川辭書出版社以明萬曆刻本的《唐詩類苑二百卷》為底本，「編纂時……以原書為基礎，充分考慮原來的版本、體例；詩文內容有疑誤處參校《全唐詩》等書和作家專集；部類上進行必要的分合、增減、調整，部類名稱既尊重傳統，又兼顧現代，重新編制索引。……得四十部，一千零八十三類」²⁰，名之為《唐詩分類大辭典》，收錄唐詩 10,197 首。分類以題材為標準²¹，所分類目如下：

	大類	中類
1	天文部	日、月、星、河、風、雲、雷、雨、雪、陰、霽、虹、霧、露、霜、冰、火、煙
2	四時節	春、夏、秋、冬、曉、夜、寒、熱、曆、十二月、元正、立春、人日、

¹⁷ . 參見拙作，〈詩詞語言詞彙切分與語意分類標記之系統設計與應用〉，第四屆數位典藏技術研討會，中央研究院主辦，2005 年 9 月 1-2 日。（收於網路展書讀網站：<http://cls.hs.yzu.edu.tw>）

¹⁸ . 參見《武英殿本四庫全書總目提要》，集部總集類存目二，192 卷，第 5 冊，第 156 頁上，《唐詩類苑二百卷》條。

¹⁹ . 同註 18。

²⁰ . 參見《唐詩分類大辭典》，馬東田主編，四川辭書出版社，1992 年成都出版社出版，第 6 頁，前言。

²¹ . 參見《唐詩分類大辭典》，馬東田主編，四川辭書出版社，1992 年成都出版社出版，第 9 頁，凡例五「分類以題材為標準，不按詩的體裁排列。」

	氣部	正月十五、正月晦日、中和節、耗磨日、寒食、清明、三月三日、春社、三月晦日、四月八日、五月五日、伏日、立秋、七夕、中元、中秋、秋社、九月九、小雪、冬至、歲除、臘
3	地理部	郊、原、村、野、林、園、苑、關、峽、石、塵、道路、陌、徑、岸、田、陵墓
4	山部	總山、泰山、衡山、華山、嵩山、王屋山、巫山、終南山、廬山、天台山、龍門山、峴山、荊山、鹿門山、武夷山、鳴皋山、九華山、天目山、青城山、峨眉山、虎丘山、秦望山、交崖山、魚亭山、石門山、天門山、灞陵山、望楚山、皖公山、太行山、茅山、蒙山、五松山、東山、西山、亂石山、石窟山、翅頭山、鹿頭山、松寥山、白鹽山、泥功山、丈人山、中山、三卿山、泉山、桐柏山、峽口山、溫門山、火山、金山、君山、岐山、錫山、天瀟山、地脈山、沃洲山、壺公山、琅琊山、太白山、胡山、八公山、爛柯山、西塞山、麻姑山、望夫山、甌山、月岩山、少華山、四明山、覆釜山、條山、峰、巖、嶺、洞、谷
5	河川部	總水、海水、江水、河水、淮水、濟水、涇水、洛水、渭水、漢水、湘水、汾水、滄水、龍泓水、伊水、隴頭水、湖水、川、浦、溪、澗、石淙、池、津、潭、湍、灘、瀨、洲、渚、源、灣、溝、渡、沙、堤、潮、濤、陂、坡、磯、塘、堆、瀑布、泉、溫泉、井、橋
6	京都部	京都
7	州郡部	州郡
8	邊塞部	總邊、薊門、安西、臨洮、輪臺、隴西、遼、磧、燕支、長城、朔方、河湟、涼州、伊州、陸州、西州、豫章、雲中、并州、梁州、廣州、平城、塞垣、塞上、塞下、塞外、出塞、入塞、巡邊、護邊
9	帝王部	帝王、中宮、太子、諸王
10	皇室部	公主、駙馬
11	職官部	相、少保、僕射、司空、司徒、司直、左丞、右丞、尚書、侍郎、郎中、員外、吏部、南部、黃門、給事、常侍、諫議、補闕、拾遺、舍人、中書、起居、翰林、學士、詹事、庶子、洗馬、秘書、校書、正字、著作、博士、贊善、長史、賓客、中丞、侍御、卿、侍中、評事、金吾將軍、羽林將軍、將軍、都尉、郎將、別將、輕車、執戟、揮使、大總管、節度使、觀察使、朝集使、兵馬使、副使、司馬、判官、書記、掌記、管記、都督、都護、都使、開府、參軍、京兆、少尹、功曹、法曹、兵曹、騎曹、太守、使君、別駕、司錄、司戶、司倉、司農、司法、令、丞、明府、贊府、少府、主簿、尉、押衙、錄事、從事、掾
12	治政部	寵錫、貢獻、除授、薦舉、遷擢、奉使、和蕃、蠻夷、胡兒、招諭、歸降、即敘、風化、掌選、監選、放牒、選補、赴選、求職、入京、入朝、赴闕、入計、計會、赴任、括圖書、巡行、巡農、鎮守、轉運、知貢舉、推事、聽訟、鞠獄、問囚、置倉、覆糧、賑濟、宣慰、採訪、安輯、安擾、市馬、休假、秩滿、囚獄、赦宥、按覆、按獄、刑罰
13	禮部	郊丘、社稷、宗廟、祭祀、釋奠、封禪、明堂、籍田、巡幸、扈從、朝

		會、禮見、召對、省直、誕辰、拜表、生日、滿月、晬日、宴饗、酺宴、婚姻、拜掃、送葬、輓歌
14	樂部	總樂、樂府、歌、舞、琴、胡琴、箏、瑟、阮咸、琵琶、箜篌、鐘、鼓、笙、簫、笛、胡笳、箜篌、管、磬、越器、雲和、角、拍板
15	文藝部	經、書、子、史、詩、賦、字、文、碑、內狀、檄、朱版、紙、箋、墨、筆、硯、書筒、詩筒、書船、文集櫃
16	軍事部	戰伐、幕府、行營、監軍、送兵、乞師、教戰、觀兵、赴軍、從軍、入軍、獻兵書、軍中、行軍、告捷、凱旋、回軍、八陣圖、屯田、戰場、烽、旌、旗、戈、劍、刀、吳鉤、弓、箭、箭鏃、弩、彈、馬鞭
17	人事部	覲省、扶侍、迎親、歸養、兄弟、子嗣、女、童子、神童、宗族、鄰里、姻戚、交游、僚友、耆老、懷思、尋訪、參謁、招致、會遇、話舊、遇宿、待後、問候、宴集、行樂、俠少、豪華、醉、夢、沐浴、應辟、還山、山中、隱逸、閒適、寄、贈、答、贈答、代書、問訊、得書、獻詩、示諭、規諷、言語、詠笑、嘲戲、送、別、送別、留別、贈別、話別、酬別、寄別、宴別、游覽、眺望、臨泛、懷古、羈旅、移家、避地、歸國、還鄉、感遇、古風、雜興、述懷、喜幸、感慨、時事、宦情、遭讒、貶謫、流徙、棄官、罷官、疾病、艱危、貧窮、感舊、懷賢、哀傷、淚、哭、賢媛、美人、寵幸、閨情、幽怨
18	儒家部	秀才、文學、文士、宏詞、有道、孝廉、明經、貢士、赴舉、御試、放榜、落第、擢第、及第、座主、先輩、同年
19	釋佛部	僧、衲子、尼師、尼子、小尼、出家、經典、禮拜、行香、施香、受戒、齋戒、出齋、談玄、問道、飯僧、道場、錫杖、佛鉢、幡、念珠、銅碗、佛、還俗
20	道家道教部	仙、道士、女道士、入道、齋戒、誦經、受籙、步虛、投龍、焚修、奠簡、修養、學道、訪道、味道、問易、種藥、求藥、謝惠藥、買藥、採藥、洗藥、燒金、藥轉、大還、得道、避穀、丹灶、道場、道書
21	居處部	宮、殿、闕、樓、臺、榭、閣、亭、院、官舍、第、宅、堂、齋、軒、屋、房、室、廊、廬、窩、居、舊居、新居、卜居、移居、別墅、精舍、書院、題壁、題扉、題門、題居、山居、郊居、林居、村居、溪居、隱居、莊、山莊、山墅、茅茨、山家、田家、酒家、城、郭、市、街、門、柱、館、客舍、郊舍、店、驛
22	寺觀部	寺、浮圖、樓、閣、院、禪居、禪房、佛龕、精廬、精舍、禪堂、法堂、經堂、講堂、影堂、玉像堂、禪室、石室、浴室、僧舍、齋、軒、窗、林泉、亭、池、蘭若、臺、題壁、觀、壇、院、堂、房、道居、道室、柱
23	祠廟部	祠、廟、神
24	產業部	農、賈、織、漁、釣、樵、牧、獵
25	器用部	銅儀、漏、舟、帆、車、鏡、帘、篔、帷、繡障、帳、屏、香、香毬、香印、燈、燈杖、燈缸、燈檠、燈花、燭、燭花、炭、扇、印囊、藥盒、鼎、爐、床、茵、簟、罽、浦團、枕、席、杖、竹夾膝、拂子、樽、卮、

		盞、杯、瓢、酒胡子、榼、匙、碗、瓶、筒、砧、秤、莎柵
26	服飾部	冠、巾、帽、簪、釵、章服、袍、衣、襖、衫、裘、襪、屐、鞋、襦、被
27	飲食部	酒、食、餅、麵、粥、湯、鱠、素饌、蔬、茶、小饅味、藥
28	玉帛部	寶、玉、珠、珊瑚、水精、琉璃、琥珀、金、銀、錢、絲、錦、綺、繡、綾、羅、布、素
29	技藝部	畫、射、棋、彈棋、風箏、打球、鞦韆、竟渡、傀儡、戴竿、拔河、龜背戲、繩伎
30	方術部	醫、卜、陰陽、推命
31	花部	牡丹、芍藥、桂花、蘭花、木蘭花、蕙草、菊花、合歡花、辛夷花、躑躅花、葵花、海紅花、木香花、紫薇花、紅薇花、薔薇花、木槿花、紫陽花、梔子花、海棠花、丁香花、麗春花、玉蕊花、紫荊花、萱花、迎春花、夜合花、金錢花、雞冠花、杜鵑花、玫瑰花、玉簪花、優鉢羅花、芙蓉花、蓮花、藕、菱花、雜花
32	草部	紅豆、萵苣、菜、韭、石芥、筍、瓜、枸杞、恆春草、菖蒲、靈仙毘、蒲、術、白蓂荷、訶梨勒、黃精、地黃、白頭翁、蒼耳草、車前子、紫參、人參、茱萸、菟絲、葛、蘆、葦、棕、茅、蒹葭、荻、苔、萍、蓼花、芭蕉、藤、莎、蓬、蕓、瓦松、垣衣、雛、雜草、雜藥
33	果部	梅、桃、杏、棗、李、郁李、梨、櫻桃、榴、荔枝、橙、橘、枇杷、林檎、胡桃、葡萄、雜果
34	木部	槐、楸、靈壽木、柳、竹、石竹、桐、琪樹、端正樹、松、松英、柏、杉、檜、楠、石楠、檜、榆、桑、情深樹、冬青、桤、木蓮樹、椰、柑樹、櫻樹、槿、椴子樹、雜木、雜樹、葉、雜植
35	鳥部	鳳、鶴、鵠、孔雀、鶯、鸚鵡、伯勞、鳩鵲、百舌、反舌、戴勝、烏、鴉、鵲、鳩、鷓鴣、杜鵑、燕、雉、雀、鷹、隼、鶻、鴛鴦、白鷗、鷗、鳧、雁、鳶、鷺鷥、鸕鶿、鷓鴣、提壺、啄木、浴浪鳥、精衛、鷓鴣、訓狐、雜禽鳥、山雞、雞、鴨、鵝
36	獸部	麟、象、豹、虎、熊、馬、牛、驢、羊、鹿、兔、麕、猿、犬、獺、狐兔、鼠
37	鱗介部	龍、魚、龜、蟹、蛤
38	虫豸部	蟬、蜂、蝶、蜘蛛、蠶、螢、蚊、蚤、蟲、蝦蟆、蛙、蝌蚪、促織、壁魚、蚯蚓、蠟蛾、雜虫豸
39	祥異部	甘露、慶雲、大有年、祥光、洛書、芝草、瑞蓮、鸚鵡貓兒、時樂鳥、白雉、水、旱、火
40	雜部	

除了《唐詩分類大辭典》之外，另有以題畫詩為範圍編纂的《御定歷代題畫詩類一百二十卷》。詩畫藝術相通，是唐宋代的藝文特色，詩人為畫題詩，名為題畫詩，畫家從詩作裡尋找畫作的題材與靈感，名為詩意畫，前者如杜甫的〈戲為雙松圖歌〉，後者如宋代馬遠〈寒江獨釣〉、范寬的〈寒江釣雪〉畫都是取柳宗元〈江雪〉：「千山鳥飛絕，萬徑人蹤滅。孤舟蓑笠翁，獨釣寒江

雪。」詩的意境所作的畫。清代陳邦彥奉旨校刊的《御定歷代題畫詩類一百二十卷》²²，依內容主題將題畫詩分為：天文、地理、山水、名勝、古蹟、故實、閒適、古像、寫真、行旅、羽獵、仕女、仙佛、神鬼、漁樵、耕織、牧養、樹石、竹蘭、花卉、禾麥蔬果、禽、獸、鱗介、花鳥合景、草蟲、宮室、器用、人事、雜題等三十類，也可作為查檢的分類選項。

全唐詩網站在《唐詩分類大辭典》原有的類目分類基礎上，修訂分類類目，將該辭典未收的詩作予以增補，提供使用者依「主題分類」查檢的選項。也將所有詩作之題畫詩，依據《御定歷代題畫詩類》之類目分類，提供查檢。

然而，根據《唐詩分類大辭典》凡例所言「分類以題材為標準」²³，檢視全書，有幾個問題值得商榷，舉其要者說明如下：

（一） 分類與詩作內容不合：

依凡例所言「分類以題材為標準」，檢視全書，絕大多數都以詩題之字面含義分類，不是以詩作內容之主題分類，使得有些分類與詩作內容的主題未必相關，產生分類類別與主題不合的現象，例如將杜牧〈金谷園〉：「繁華事散逐香塵，流水無情草自春。日暮東風怨啼鳥，落花猶似墜樓人。」歸在地理部的「園」類，與詩的主題不合。金谷園是石崇所建，這首詩寫的是綠珠為石崇墜樓而死的殉情故事，不是金谷園的園景，歸類於此，並不適當。又如〈關山月〉是漢樂府〈橫吹曲〉曲調名，多用來寫邊塞士兵因久戍不歸，思念家人，和家人互傷離別之情的作品，《唐詩分類大辭典》所收十一首〈關山月〉，其中儲光義〈關山月〉：「一雁過連營，繁霜覆古城。胡笳在何處，半夜起邊聲。」的「繁霜」，可解為月光或霜雪的霜，如解為月光，這十一首詩都寫到月亮，在戰地的征人雖因見到月亮而想念家鄉，但主要的重點是描寫邊塞征人思念家人之苦，月亮只是因發思鄉情懷的觸媒劑，依該書分類凡例，這些詩應該歸入「人事部」。事實上《全唐詩》共收有二十一首〈關山月〉，只有皎然及霍總二人所寫的〈關山月〉與邊塞征戰無關，皎然的詩寫月亮²⁴，霍總的詩寫男女之情²⁵。

（二） 標舉小類的體例不一

有些大類之下的小類，有普通名詞、專有名詞及其他詞類，含多種詞類的類別，有些只列出普通名詞，有些增列專有名詞，體例不統一，例如「河川部」的詩，有些詩題上只泛寫普通名詞，如劉長卿〈湖上遇鄭田〉、李德裕〈潭上喜見新月〉，有些以河川的專有名詞當作詩題，如宋之問〈洞庭湖〉、杜甫〈登白馬潭〉，「洞庭湖」與「白馬潭」都是專名詞，所以「河川部」的小類，列有專有名詞「淮水、濟水、涇水、洛水……」及普通名詞「湖、池、津、潭、湍、灘……」等，「山部」亦如是。「居處部」的詩，與「山部」、「河川部」一樣，有些詩題直接以專有名

²² . 清代陳邦彥奉旨校刊，《御定歷代題畫詩類一百二十卷》，台北市世界書局根據《景印搗藻堂四庫全書薈要》，456-457，集部，第109-110冊，總集類重印，1988年出版。

²³ . 同註20。

²⁴ . 皎然〈關山月〉：「家家望秋月，不及秋山望。山中萬境長寂寥，夜夜孤明我山上。海人皆言生海東，山人自謂出山中。憂虞歡樂皆占月，月本無心同不同。自從有月山不改，古人望盡今人在。不知萬世今夜時，孤月將誰更待。」

²⁵ . 霍總〈關山月〉：「珠璣翡翠床，白皙侍中郎。五日來花下，雙童問道傍。到門車馬狹，連夜管弦長。每笑東家子，窺他宋玉牆。」

詞為題，如李白〈長信宮〉、崔顥〈黃鶴樓〉等，有些只泛寫亭臺樓閣，如劉長卿〈同諸公登樓〉、邱為〈題農父廬舍〉，「居處部」的小類卻只列出「宮、殿、闕、樓、臺、榭……」等普通名詞，未列出「長信宮」、「黃鶴樓」等專有名詞，「寺觀部」情況也相同，可見體例不統一。

（三） 專名與通稱、俗名混合

除了大類之下的小類列法體例不一，在同一小類中，也有將專有名詞與通俗名稱混合的情況，該書的分類，既不是將所有詩題裡出現的專有名詞一一列舉，又未曾將通俗專名做詞彙控制，如「草部」列有「苔」及「垣衣」、「瓦松」，全唐詩詩題有「苔」字者共十七首，寫「蘚」者二首，「苔、蘚」並列者一首，寫「垣衣」者一首，寫「瓦松」者一首。事實上「垣衣」是苔的一種，《廣群芳譜》之〈苔譜〉：「苔，一名圓蘚，……苔在牆曰垣衣，在屋曰瓦松」。²⁶但小類中列了「苔」、「瓦松」及「垣衣」，未列「蘚」。在「木部」列有「情深樹」，李白有〈題情深樹寄象公〉詩：「腸斷枝上猿，淚添山下樽。白雲見我去，亦為我飛翻。」李白詩題裡的「情深樹」只是李白看到的某一棵樹，李白隨意名之曰「情深樹」，在全唐詩之詩題及詩句裡，也只有李白這首詩這麼寫，當作一個小類，顯然不合適。又如「木部」列有「端正樹」，全唐詩收有溫庭筠〈題端正樹〉、貞元文士〈題端正樹〉二首詩，貞元文士詩之詩序曰：「酉陽雜俎云：長安西端正樹，去馬嵬一舍之程，迺德宗皇帝幸奉天，睹其蔽芾，錫以美名。有文士經過，題詩逆旅，不顯姓名。」可見「端正樹」是唐德宗錫以美名的樹，不是該樹種的學名，與本小類其他以「柏、杉、檜、楠……」等樹種的學名為類別的體例不同。此項分類，另列有「雜樹」，應將「情深樹」、「端正樹」列在「雜樹」。其他如「鳥部」所列「反舌」即「百舌」鳥，未做詞彙控制而將二者並列為小類，徒增繁瑣與混淆。

從以上所敘述的情況看來，《唐詩分類大辭典》所區分的大類或小類，其分類的方法及訂定詞目的方式，與語意的概念及使用者的認知不盡相同，也不符合學術分類的規則與慣例，使用者無論從「客觀知識性取向」或「主觀情意性取向」判斷，《唐詩分類大辭典》的分類，都不盡符合語意概念及使用者認知，再加上分類的結果與凡例所說的「分類以題材為標準」不盡相合，使用者的認知與該書分類有落差，如能以詩作之內容作語意標記與分類，作為判斷詩作主題的依據，其準確度將會比《唐詩分類大辭典》的正確度高。

本研究將透過對全唐詩詩作內容的語意標記及分類，嘗試從不同角度建立唐詩主題分類，以與《唐詩分類大辭典》的分類相比較，並期待在未來的研究裡，能進一步從語意標記及概念分類、詞彙共現三個角度分析、彙整、歸納唐詩主題。

六、 詩詞語言語意標記對全文檢索的重要性

中國的韻文發展到唐代，可說是達到了顛峰，成為中國文學史上燦爛的一顆明珠。唐詩成就攀登高峰，有很多原因，舉其要者言之，一方面是因為文體發展過程的自然演變現象，一方面是唐代科舉考試以詩取士，文人寫詩成為通往利祿之路的終南捷徑，加上在位者的喜好與提倡，在在都促進了唐詩的發達，而其中最重要的因素是唐代以詩取士，使得很多文人因為詩寫得好而能

²⁶ . 參見清汪灝等撰，《佩文齋索引本廣群芳譜》第九冊卉譜：苔譜，新文豐出版社，1980年出版，第5054-5055頁。

透過科舉考試取得官位，發揮了影響力。詩人來自四面八方，來自各個不同的出身背景，使得唐詩的數量繁多，題材豐富多變，呈現萬千氣象。

清代彭定求奉敕編纂的《全唐詩》以明代胡震亨《唐音統籤》及季振宜《唐詩》為底本編纂而成，全書分為九百卷，將帝王后妃的作品列於首，其次是樂章、樂府，接著依作者時代先後編排，附以作者小傳，最後是聯句、逸句……等等。作者之後繫上詩題，使用者只能依照作者或詩題逐一查檢資料，非常不便。《唐詩分類大辭典》雖然存在著上述問題，但是初步達到了提供使用者從內容主題尋找資料的功能，比起只從作者、詩題尋找資料的編輯方式方便許多。但是儘管《唐詩分類大辭典》聲稱是以主題分類，事實上有很多詩作除了文字表面的含義，還含有「雙關義」、「情韻義」、「象徵義」、「深層義」、「言外義」²⁷，不是單一主題，使得主題分類的困難度非常高。

要使全文檢索從傳統只能分辨字形提升到能分辨字詞義，需要建立詞彙語意標記，讓電腦能理解字詞的語意。詞彙是構成文學作品的基本元素，作者以各種不同的詞彙，透過修辭技巧、想像力及藝術創造表達自己的情思意念，讀者對作品解讀及體會的深度，因自身學養、年齡、經驗等等條件不同而有不同。詞彙的語意標記，可以幫助讀者解讀詞彙、體會詩作，與詩人的情思意念共鳴；語意標記資料庫的建立，也可以使電腦從只能分辨字形提升到可以理解字詞義，進而達到提升全文檢索質量的目標。

前已述及，詩是特殊的語言，特別講究聲律及形式之美，受到字數及平仄押韻等格律的限制，詩的用語與語體文不同。因字數的限制，一首詩短則 20 字（五言絕句），多則不過數百字（排律或古體詩），短短篇幅要包含豐富的內涵，修辭的技巧自然比語體文更為講究，因此也存在著更多一詞多義，一義多詞的現象，除了一般用語之外，使用了更多典故等特殊用語，使得詞彙的含義比語體文更豐富，更轉折，甚至於脫離了原有的語意。這種特殊文體的特殊用語，若以傳統比對字形的方式開發檢索系統，檢索的質量勢必無法滿足使用者的需求。以「酒」為例，唐代詩人不乏喜歡飲酒者，唐詩裡寫到酒的作品非常多，如果我們要查詢唐詩裡對於酒的描寫，在傳統的全文檢索模式裡，使用者直接的認知是輸入「酒」檢索，但是寫酒的詩，不一定使用「酒」字，舉凡：瓊漿、玉液、濁醪、花釀、新醅等詞彙指的都是酒，李白「舉杯邀明月，對影成三人」（〈月下獨酌〉詩）、「飛蓬各自遠，且盡手中杯」（〈魯郡東石門送杜二甫〉詩），雖然沒有寫到酒字，杯中盛的都是酒。王翰〈涼州詞，二首之一〉詩：「醉臥沙場君莫笑，古來征戰幾人回。」使之醉臥沙場的是葡萄美酒。白居易〈酬夢得比萱草見贈〉詩：「杜康能散悶，萱草解忘憂。」的「杜康」是酒，白居易〈效陶潛體詩〉詩：「清光入杯杓，白露生衣巾。乃知陰與晴，安可無此君。」的「此君」也是酒。然而宋之問〈綠竹引〉詩：「含情傲睨慰心目，何可一日無此君。」的「此君」是竹子²⁸，李瀚〈蒙求〉詩：「杜康造酒。蒼頡制字。」的「杜康」指的是造酒的杜康這個人，而不是酒。由此可知，若直接輸入「酒」查詢，不但會遺漏了很多同義詞，也會將與

²⁷ . 參見袁行霈著，〈中國古典詩歌的多義性〉，收於《中國詩歌藝術研究》，五南圖書出版公司印行，1989年5月台灣出版，第6-23頁。

²⁸ . 此君在詩中有三個含義，其中一個含義是如字面語意，指的是「這一個人」，一個含義指「酒」，典故出自文中所引之白居易〈效陶潛體詩〉詩：「清光入杯杓，白露生衣巾。乃知陰與晴，安可無此君。」一個含義指「竹子」，典故出自《世說新語》下卷上〈任誕〉～760～「王子猷嘗暫寄人空宅住，便令種竹。或問：『暫住何煩爾？』王嘯詠良久，直指竹曰：『何可一日無此君？』」。

酒無關的「酒虎」（鹽漬海鮮）、「酒泉」（地名）檢索出來。由「酒」的例子可知，詩中寫到酒時，或直接用酒字，或用同義詞，或用典故，或用「杯」，用「醉」表示。唐代詩人杯中所盛的固然以酒居多，但是宋代詩人杯中所盛，酒與茶都有，如陸游〈入梅〉詩：「墨試小螺看斗硯，茶分細乳玩毫杯。」杯中盛的是茶。被杜甫稱之為「斗酒詩百篇」的李白，寫〈和盧侍御通塘曲〉詩：「偶逢佳境心已醉，忽有一鳥從天來。」使他醉的，有時是佳境不是酒。由此可知，對詞彙進行語意標記，使電腦能分辨語意，是提升檢索正確性及完整性的重要關鍵。

但是詩句的語意不一定在單一詞彙顯示，或說從單一詞彙解讀，不見得能夠完整理解整首詩的含義。整首詩的含義，有時需要從共現的詞彙、從作者及寫作的背景綜合判斷，非常複雜，也非常困難。以杜甫〈江南逢李龜年〉：「落花時節又逢君」²⁹為例，「落花時節」字面的含義是指春末百花凋零的季節，以整首詩分析，這裡的「落花時節」，除了字面的含義之外，還包含杜甫與李龜年人生的「落花時節」及唐代國勢的「落花時節」。杜甫與李龜年在長安相識時，二人都是青年時期，也正值盛唐國勢的顛峰，經過安史之亂，國勢已由盛轉衰，重逢於江南，都成為他鄉異客，唐代的國勢、二人的人生，都已如「落花時節」進入中晚年的凋零衰敗。

本研究嘗試把詩的語言，區分為一般用語、特殊用語、情境用語三種。一般用語指的是詞彙的語意與其他文體類似，多數是客觀知識性的描述用語；特殊用語指的是含有典故、象徵、比喻、意象等等修辭技巧含義的詞彙，多數具有主觀情意性的引導與感發功能，詞彙的含義不只是字面的含義；情境用語指的是不同詞彙共現所烘托出來的情境，這些詞彙的本身，也許是一般用語，也許是特殊用語，但是因為共現而產生的情境，往往就是詩的主題，如本文後段所舉的苔蘚與門階共現的例子即是。由此可知，植基於語意學的詞彙語意標記，可以使電腦對文本內容含義的理解與判斷更接近人的認知觀點。如前所述，詩詞語言是特殊的語言形式，比起其他文體，存在著更多一詞多義，一義多詞的現象，加上詩句語意或詩作主題不一定在單一詞彙顯示，而是以共現的詞彙整體烘托出來，增加語意標記的困難，但相對而言，也可以從中歸納、推斷詩的主題。今就詩詞的特殊用語及情境用語分述如下：

（一）詩詞特殊用語

以詩詞語言而論，典故與意象都是特殊用語，都是意在言外，層層轉折。典故與成語雖然都是引用前人之「人」或「言」或「事」以豐富詞義及情境的修辭技巧，成語的用語經過約定俗成的轉化過程，所使用的詞彙是固定的，但是典故的詞彙容許詩人個別創造和變化，使用的變化可歸納為三種：其一是典故出處含義都相同，而詞彙不同的同義典故；其二是典故出處相同，但側重的含義不同，用詞或相同，或不同；其三是典故出處不同，所取的含義相同，用詞或相同，或不同³⁰，例如，自從陶淵明寫下〈飲酒〉詩：「採菊東籬下，悠然見南山」之後，東籬與菊花都附上一層隱逸的含義，後代詩人引用這個典故時，所使用的詞彙包含「東籬」、「東籬菊」、「見南山」、「採菊」、「看山悠悠」、「結廬人境」、「晉菊」、「人境」、「陶令菊」、「黃花東籬」等等，不勝枚舉。

意象指的是詩人將主觀的情意投注在客觀的物象上，使得物象附著了詩人的情感而成為意象。袁行霈教授在〈中國古典詩歌的意象〉文中說：「古人所謂意象，儘管有種種不同的用法，

²⁹ . 杜甫〈江南逢李龜年〉：「岐王宅裡尋常見，崔九堂前幾度聞。正是江南好風景，落花時節又逢君。」

³⁰ . 參見范之麟、吳庚舜主編，《全唐詩典故辭典》，湖北辭書出版社，1989年出版。

但有一點是共同的，就是必須呈現為象……因此可以說，意象賴以存在的要素是象，是物象。……意象是融入了主觀情意的客觀物象，或是借助客觀物象表現出來的主觀情意。」³¹

從歷代對於「意象」的定義看來，「意象」產生的先決條件是詩人「主觀的情意」，這麼說來，「意象」的個人色彩濃厚，詞彙之語意是否會因意象的使用而氾濫不可收拾，實則不然。

詩人寫詩，會創造新的詞彙，也會在已有的詞彙上賦予新的詞義，二者皆有可能為後代詩人沿用。那些詩人創造那些詞彙或創造那些詞義，這是詞彙語意學的範疇，後者一旦為後人沿用，便成為典故。例如「南浦」，最早出現於《楚辭》卷第二〈九歌·河伯〉：「與女遊兮河之渚，流澌紛兮將來下。子交手兮東行，送美人兮南浦。」其後江淹〈別賦〉：「春草碧色，春水淥波。送君南浦，傷如之何！」這二處的「南浦」都附著了離別感傷色彩。

「浦」的本義是水邊或大河流的小叉口，「南浦」泛指面南的水邊。古人外出，如果有水路，通常走水路，沒有水路才會改走陸路，因此常在水邊送別。水面有東西南北四個方向，自從唐代王維寫下〈齊州送祖二〉：「送君南浦淚如絲，君向東州使我悲。為報故人憔悴盡，如今不似洛陽時。」之後，詩人寫水邊送行的感傷，都使用「南浦」，從此之後，「南浦」已經牢牢的與離別感傷結合在一起。檢視全唐詩，寫到「南浦」的詩句總計101句，這101句作者的時代在王維之前的有16人，其餘都在王維之後，內容幾乎都與離別有關。寫「西浦」的只有杜甫〈梅雨〉：「南京西浦道，四月熟黃梅。」及張籍〈宿江店〉：「野店臨西浦，門前有橘花。」2句；寫「北浦」的只有盧尚書〈哭李遠〉：「纔收北浦一竿釣，未了西齋半局棋。」1句；「東浦」0句，這些句子都與離別無關。唐宋詞寫到「南浦」的詞有166句，「西浦」2句，「北浦」0句，「東浦」0句，而使用「南浦」的詞句，幾乎都與離別有關。由此可見，寫「南浦」送行而感傷的作品雖在屈原及江淹時就已出現，但是王維「送君南浦淚如絲」的詩句感動後人，而使「南浦」附著了「離別感傷」的固定意象，為後世詩人沿用，而其原來的語意已經被隱藏在意象之後，成為次要的意義。也可以說屈原、江淹、王維創造了新的詞義，就如同「東籬」「菊花」在陶淵明之後，在本義之外多了隱居的含義一樣。

凡此，因某位詩人所使用的意象，廣為後世詩人沿用而成為典故的詞彙，都有固定的詞義，在語意標記時，可以標記其典故詞義。

（二）詩詞情境用語

所謂情境用語，指的是詞彙中含有意在言外的含義。詩詞中物象的描述，有些是客觀的陳述，如王維〈積雨輞川莊作〉：「積雨空林煙火遲，蒸藜炊黍餉東菑。漠漠水田飛白鷺，陰陰夏木轉黃鸝。山中習靜觀朝槿，松下清齋折露葵。野老與人爭席罷，海鷗何事更相疑。」詩中「白鷺」、「黃鸝」、「槿」、「葵」只有植物及動物物象本身的含義，語意標記時只需要標註物象本身的含義。至於附著了詩人主觀情意並為後世所沿用而成為典故者，已如前述。

³¹ . 袁行霈，〈中國古典詩歌的意象〉，收於《中國詩歌藝術研究》，五南圖書出版公司印行，1989年5月台灣初版，61頁。

詩詞中有些物象的描述，雖然也只是客觀的呈現，但是因為詩詞裡的修辭技巧，而使得這些客觀的物象，縱然不直接附著詩人主觀的情意，卻間接的引發整首詩的情境，婉轉引出詩人的心境。而使整首詩同樣具有詩人主觀的情意心境在內，只是這種情意的呈現往往需要與其他的詞彙互相襯托才能顯現。

詩詞的描寫，可以大別為三類，其一寫詩人主觀內在的心靈活動，其二寫客觀外在的景物，其三寫人與外在景物的互動，而以第三種所佔的數量最多。描寫人與外物的互動，通常使用以情襯景，如李商隱〈落花〉詩：「高閣客竟去，小園花亂飛。參差連曲陌，迢遞送斜暉。腸斷未忍掃，眼穿仍欲歸。芳心向春盡，所得是沾衣。」或用以景襯情，如杜甫〈登高〉詩：「風急天高猿嘯哀，渚清沙白鳥飛迴。無邊落木蕭蕭下，不盡長江滾滾來。萬里悲秋常作客，百年多病獨登臺。艱難苦恨繁霜鬢，潦倒新停濁酒杯。」或情景交融，如劉禹錫〈楊柳枝〉：「春江一曲柳千條，二十年前舊板橋。曾與美人橋上別，恨無消息到今朝。」三種描寫方法，所要表達的，都是詩人的情感。言情與寫景的交互襯托，需要依賴外物作為作者與讀者之間溝通的媒介。以景襯情之作，詩人寫景物時，並未就任何單獨的詞彙附著上個人主觀情感的色彩，作者與讀者都以視覺、聽覺、嗅覺、味覺、觸覺等感官感受外物。作者以文字引導讀者對生理感官感受的想像，以共現的詞彙營造出某種情境。這些想像所引發的感覺必須是符合多數人的認知基礎與感知經驗，才能引起雙方的共鳴，讀者才能體會作者心中所要營造的情境。例如賀鑄〈青玉案〉：「若問閒愁都幾許。一川煙草，滿城風絮。梅子黃時雨。」用具體的物象，用多數人生活中共同有的經驗「一川煙草，滿城風絮。梅子黃時雨」，做為作者與讀者共鳴的媒介，將作者心中抽象的「閒愁」透過共有的認知媒介，傳遞到讀者心中。又如秦觀〈減字木蘭花〉：「欲見回腸。斷盡金釭小篆香。」也是利用「斷盡金釭小篆香」的具體物象傳遞「回腸」的幽微心事。

詩人們以某些詞彙共現在同一首詩中，以達到傳遞心境，烘托情境的效果，所使用的詞彙，或所選擇共現的詞彙，與所要烘托的情境，必然具有物象本身的特徵及某些人際間共有的認知脈絡可以依循。這個共同的認知脈絡，引發讀者的認知、感知與聯想，成為作者與讀者之間共鳴的媒介，詩人因心境的感知，以聯想創造文學藝術，讀者以聯想、感知體會詩人的創意與心境、情境。以「苔蘚」為例，苔蘚植物是由水生生活趨向陸生生活的植物族群，生長於極端潮濕的向陽地區或是陰暗的環境中，以配合其親水性結構，不然便很快枯死。³²這些地方往往也是人跡罕至之處，這是我們對於「苔蘚」的認知基礎。由於苔蘚性喜陰暗潮濕的特性，潮濕陰暗在感官上給人清冷幽深，人跡罕至的感覺，如李白〈長干行，二首之一〉：「門前遲行跡，一一生綠苔。苔深不能掃，落葉秋風早。」使得詩詞中寫到苔蘚時，往往是「江邊道路多苔蘚，塵土無由得上衣。」³³的失意文人，或是「玉階陰陰苔蘚色，君王履綦難再得。」³⁴的失寵佳人，或是「天花寂寂香深殿，苔蘚蒼蒼闕虛院。」³⁵的塵外僧人，或是「肌膚銷盡雪霜色，羅綺點成苔蘚斑。」³⁶的空閨女

³² . 引自延陵科學綜合室建置的「延陵植物標本閣」網站：<http://www.ngensis.com/flora/fn01.htm>，苔蘚類單元：<http://ngensis.com/flora-b/L&M.htm>

³³ . 姚合，〈送崔約下第歸揚州〉：「滿座詩人吟送酒，離城此會亦應稀。春風下第時稱屈，秋卷呈親自束歸。日晚山花當馬落，天陰水鳥傍船飛。江邊道路多苔蘚，塵土無由得上衣。」

³⁴ . 沈佺期，〈古歌〉：「落葉流風向玉臺，夜寒秋思洞房開。水晶簾外金波下，雲母窗前銀漢回。玉階陰陰苔蘚色，君王履綦難再得。璇閨窈窕秋夜長，繡戶徘徊明月光。燕姬彩帳芙蓉色，秦女金爐蘭麝香。北斗七星橫夜半，清歌一曲斷君腸。」

³⁵ . 劉長卿，〈齊一和尚影堂〉：「一公住世忘世紛，暫來復去誰能分。身寄虛空如過客，心將生滅是浮雲。蕭散浮雲往不還，淒涼遺教歿仍傳。舊地愁看雙樹在，空堂只是一燈懸。一燈長照恆河沙，雙樹猶落諸天花。天花寂寂香深殿，苔蘚蒼蒼闕虛院。昔余精念訪禪扉，常接微言清道機。今來寂寞無所得，唯共門人淚滿衣。」

子，或是「花木閒門苔蘚生，澹川特去得吟情。」³⁷的山林雅士。「苔蘚」若是長於門前或台階，詩人將之與「門」、「階」並列呈現時，更核心的集中在描寫因失寵而幽居空閨女子的淒清冷落。中國詩詞素來即有以男女關係影射君臣關係的傳統，因此也常用來烘托不得君王重用的失意臣子困頓孤絕的心境，所以也常用來描寫失寵的臣子門前冷落車馬稀的荒涼寂靜，如鄭谷〈長門怨，二首之二〉：「流水君恩共不回，杏花爭忍掃成堆。殘春未必多煙雨，淚滴閒階長綠苔。」以不再使用而長滿綠苔的閒階，婉轉呈現幽居於此的空閨女子，因君恩不回，沒有人前來，階閒無用而長滿綠苔的蒼涼景象；或如胡曾〈薄命妾〉：「阿嬌初失漢皇恩，舊賜羅衣亦罷薰。倚枕夜悲金屋雨，卷簾朝泣玉樓雲。宮前葉落鴛鴦瓦，架上塵生翡翠裙。龍騎不巡時漸久，長門空掩綠苔紋。」寫的是因「龍騎不巡時漸久」，而使得「長門空掩綠苔紋」的悲涼景象。

凡是這一類以景襯情的描寫技巧，藉以寄懷的景物並未直接明顯的附著詩人的情感，難以成為公認的意象，也未能成為典故用語，但是因為物象本身的特性，使得不同的詩人無論是從認知上的取向或是從感知上的感受都有相似之處，而使得以景襯情時，所藉以襯情的物象具有認知與感知的相通性，整首詩的主題便藉著物象的現象，意在言外的呈現出來，達到「不著一字而盡得風流」的境界，若拆解成字詞，卻又無一字觸及詩意。這一類的描寫技巧，單獨詞彙的語意標記，難以表達，需要再從共現的詞彙烘托，有待更多的研究來尋找規則。

七、詩詞語言的語意分類對全文檢索的重要性

使用者的認知行為，除了前述所論及對詞彙的知識認知、情感感知之外，使用者還具備由各種單一詞彙的知識所建構的知識體系，這是只能分辨字形的電腦所無法具備的。詞彙經過語意標記，雖然已經能夠使電腦初步具有分辨語意的能力，但也只是零星片段沒有組織的語意，不符合人的知識體系。具有語意的詞彙若只是各自獨立的詞組，全文檢索時，雖然可以檢索出同近義詞，但是仍然無法符合人類以「思維聯繫」搜尋資料的行為模式。詞彙經過語意標記之後，需要建立符合知識結構及人的知識體系的知識架構，才能夠達到 Vannevar Bush 所構思的 Memex 的基本概念：「檢索時能如同人類的思維邏輯，建立知識連結的軌跡以進行聯繫索引，需要引用時，所有的資料都可以隨意檢索出來，並且可以快速的找到另一個相關的資料，也可以二個條件同時並行檢索。」³⁸全文檢索從「索引式」提升為「思維聯繫式」。全唐詩網站的全文檢索，在這些基礎下，除了提供傳統以比對字形為主的全文檢索，並提供詩的主題分類及詩句的語意概念檢索功能。

若要提供「思維聯繫式」的語意概念檢索功能，還需要將單一的語意概念建構為符合本體知識 (ontology) 的語意概念架構。中文詞語的語意分類研究成果，紙本文獻方面以梅家駒主編，上海辭書出版社出版的《同義詞詞林》最具代表性。本書第一版發表於 1983 年，第二版發表於 1996 年，收錄詞彙 54,235 筆，將詞義分為 12 大類，95 中類，1,428 小類，3,925 詞群。每一次類所包含的詞彙以及該詞彙的同近義詞依序列出，成為研究詞彙語意學及語意概念分類的重要工具書。電子本文獻方面有董振東先生於 1988 年創建的知網 (HowNet Knowledge Database)，2003

³⁶ . 嚴粲，〈望夫石〉：「何代提戈去不還，獨留形影白云間。肌膚銷盡雪霜色，羅綺點成苔蘚斑。江燕不能傳遠信，野花空解妒愁顏。近來豈少征人婦，笑採蘼蕪上北山。」

³⁷ . 張喬，〈城東寓居寄知己〉：「花木閒門苔蘚生，澹川特去得吟情。病來久絕洞庭信，年長欲思廬岳耕。落日獨歸林下宿，暮雲多繞水邊行。干時退出長如此，頻愧相憂道姓名。」

³⁸ . 同註 5，從 "As we may think" 原文摘錄。

年中央研究院資訊所詞庫小組與董振東先生展開建構繁體字知網的合作計畫，建立「概念網」，收錄 55,912 筆，分析歧義後包含概念 60,482 筆。³⁹，以上語料都是以現代漢語為範圍。以文學文本建立語意網，目前有中央研究院藉由「數位典藏國家型計畫」透過其中的「語言座標——參考資源建置與服務」⁴⁰與「溫柔在頌——唐詩三百首」網站⁴¹，參考「中央研究院中英雙語知識本體詞網」(Sinica BOW)⁴²和 SUMO (Suggested Upper Merged Ontology, 建議上層共用知識本體)⁴³資料所合作建置的「唐詩三百首知識本體」網站⁴⁴以及「蘇軾詩知識本體」網站⁴⁵。這二個網站僅先將文本中的動物、植物、人造物分類。筆者於 2002 年 7 月所主持的「以 XML(eXtensible Markup Language)可延伸式標注語言建立文章標記(Content Markup)系統研究——以蘇軾詩詞為範圍」計畫網站⁴⁶，將蘇軾詩全面標記與分類；爾後再以蘇軾詩所建置的詞彙語意標記資料庫為基礎，於 2005 年 3 月以所主持的「淺斟低唱——宋詞古唱虛擬實境教學環境建置計畫」⁴⁷選出柳永 (218 闋詞)、蘇軾 (379)、秦觀 (148)、周邦彥 (187)、李清照 (53)、姜夔 (87)、辛棄疾 (627) 等七家詞為範圍做語意標記與分類；2006 年 1 月所主持的「智慧型全台詩知識庫計畫第二階段建置計畫」⁴⁸，以國家台灣文學館委託施懿琳教授編纂的《全台詩》前三年成果為範圍，建置全台詩語意標記與分類；2006 年 3 月繼續擴充，於「新詩改罷自長吟——唐詩吟唱虛擬實境教學環境建置計畫」⁴⁹，以全唐詩為範圍，建置全唐詩詩句語意標記與分類，並增加詩的主題分類。

由前述可知，目前有關於語言之自動化處理技術及語意概念、語意標記等相關研究，現代漢語的研究比較多，語料的來源雖也包含了文學典籍的語料，但是以語體文為主，並未包含韻文。由於韻文的語言表達與構詞方式與語體文不同，目前所發展的語言處理工具及語意標記及分類項目，勉強套用於韻文，除了扞格不入之外，也抹煞了韻文的語言特徵。好的語言處理工具，需要兼顧其使用目的，用之於圖書資訊檢索，用之於語言學研究、用之於文學語言研究，用之於文學

-
- ³⁹ . 參見中央研究院資訊科學研究所陳克主持，「概念網網站」
(<http://ckip.iis.sinica.edu.tw/CKIP/conceptnet.htm>)，以及陳克健、黃淑齡、施悅音、陳怡君，2005，多層次概念定義與複雜關係表達——繁體字知網的新增架構，漢語詞彙語意研究的現狀與發展趨勢國際學術研討會，北京大學。
- ⁴⁰ . 中央研究院語言研究所所執行的「數位典藏國家型科技計畫」語言座標計畫，網址：
<http://linganchor.sinica.edu.tw/>
- ⁴¹ . 元智大學羅鳳珠、胡順萍、歐慶亨主持，【數位博物館計畫——中小學語言文學知識網(搜文解字二)——生活語文與唐宋流行歌——唐宋流行歌——溫柔在頌——唐詩三百首】網站，網址：<http://cls.hs.yzu.edu.tw/300/HOME.HTM>
- ⁴² . 中央研究院語言研究所與遠見科技股份有限公司共同開發，「中央研究院中英雙語知識本體詞網」(The Academia Sinica Bilingual Ontological Wordnet)，網址：<http://bow.sinica.edu.tw/>
- ⁴³ . 參見中央研究院中英雙語知識本體詞網，網址：http://bow.sinica.edu.tw/ont/bow_sumo_cont.html；SUMO 簡介：SUMO (Suggested Upper Merged Ontology, 建議上層共用知識本體)，網址：
<http://www.ontologyportal.org/>，由 IEEE 標準上層知識本體工作小組所建置。這工作小組的目的是發展標準的上層知識本體，這將促進資料互通性、資訊搜尋和檢索、自動推理和自然語言處理。
- ⁴⁴ . 「唐詩三百首知識本體」網站，網址：http://bow.sinica.edu.tw/ont/ts300_ont.html
- ⁴⁵ . 「蘇軾詩知識本體」網站，網址：http://bow.sinica.edu.tw/ont/ShuShi_ont.html
- ⁴⁶ . 羅鳳珠、李錫捷，【以 XML(eXtensible Markup Language)可延伸式標注語言建立文章標記(Content Markup)系統研究——以蘇軾詩詞為範圍】，國科會，2002 年 7 月 1 日至 2003 年 6 月 30 日(計畫編號：91-2422-H-155-3401-)，網址：<http://cls.hs.yzu.edu.tw/CM/login.htm>
- ⁴⁷ . 羅鳳珠、張智星、許介彥，【淺斟低唱——宋詞古唱虛擬實境教學環境建置計畫】，國科會數位典藏國家型科技計畫：94 年度數位典藏創意學習計畫，2005 年 3 月 1 日至 2006 年 2 月 28 日。計畫編號：NSC 94-2422-H-155-001，網址：<http://cls.hs.yzu.edu.tw/CSP/index.html>。
- ⁴⁸ . 羅鳳珠，【智慧型全台詩知識庫計畫第二階段建置計畫】，國家台灣文學館，2006 年 1 月 1 日至 2006 年 12 月 31 日。網址：<http://cls.hs.yzu.edu.tw/TWP/index.asp>
- ⁴⁹ . 羅鳳珠、張智星、許介彥，【新詩改罷自長吟——唐詩吟唱虛擬實境教學環境建置計畫】，國科會數位典藏國家型科技計畫：95 年度數位典藏創意學習計畫，2006 年 3 月 1 日至 2007 年 2 月 28 日。計畫編號：NSC95-2422-H-155-001。網址：<http://cls.hs.yzu.edu.tw/tang/index.html>

研究，都會因為需求不同而有不同的分類方法。今比較《同義詞詞林》與詩詞韻文語意概念分類如下：

《同義詞詞林》語意概念分類表：

序號	類別	總目
一	人	泛稱、男女老少、體態、籍屬、職業、身份、狀況、親人（眷屬）、輩次、關係、品性、才識、信仰、丑類
二	物	統稱、擬狀物、物體的部分、天體、地貌、氣象、自然物、植物、動物、微生物、全身、排泄物（分泌物）、材料、建築物、機具、用品、衣物、食品（藥品、毒品）
三	時間 空間	時間、空間
四	抽象 事物	事情（情況）、事理、外貌、性能、性格（才能）、意識、比喻物、臆想物、社會（政法）、經濟、文教、疾病、機構、數量（單位）
五	特徵	外形、表象、顏色（味道）、性質、德才、境況
六	動作	上肢動作、下肢動作、頭部動作、全身動作
七	心理 活動	心理狀態、心理活動、能願
八	活動	政治活動、軍事活動、行政管理、生產、經濟活動、交通運輸、教衛 科研、文體活動、社交、生活、宗教生活、迷信活動、司法、惡行
九	現象 與狀 態	自然現象、生理現象、表情、物體狀態、事態、境遇、始末、變化
十	關聯	聯繫、異同、配合、存在、影響
十一	助語	疏狀、中介、聯接、輔助、呼嘆、擬聲
十二	敬語	敬語
總計	12 大 類	95 中類

總目之下再分細目，細目之下條列同義詞（以「心理活動類」為例）：

類 別	總目	細目	同義詞

心理活動	心理狀態：高興	高興	喜歡、快、快樂、愉快、樂、開心、喜、喜悅、歡樂、歡喜、樂意、快活、愉悅、欣然、歡欣、歡愉、欣喜、喜氣洋洋、賞心悅目、興沖沖、怡然、欣、欣欣、歡欣鼓舞、如獲至寶、美滋滋、樂陶陶、歡快、歡然、怡、陶然、美不滋兒、美絲絲、豈樂、陶陶、喜沖沖、喜洋洋、喜滋滋、愉愉、稱快、樂呵呵、樂和、樂悠悠、樂滋滋、樂融融、歡欣雀躍、歡娛、歡虞、喜樂、天倫之樂
		興高采烈	心花怒放、喜出望外、欣喜若狂、樂不可支、狂喜、歡天喜地、大喜過望、銷魂、其樂無窮
		眉飛色舞	喜形於色、滿面春風、春風滿面、面有喜色、開顏
		手舞足蹈	歡騰、鳧趨雀躍
		慶幸	皆大歡喜、慶、幸甚、喜從天降、大快人心、以手加額、幸喜、拍手稱快、欣幸、喜幸、額手稱慶、
		彈冠	彈冠相慶
		又驚又喜	悲喜交集、轉悲為喜、驚喜交集
	心理狀態：悲傷	悲傷	難過、傷心、悲哀、傷感、難受、哀傷、哀愁、憂傷、淒然、哀、傷悲、哀哀、悲愴、惻然、愴然、悲愁、惻惻、不是味兒、不是滋味、如喪考妣、怛怛、哀戚、淒迷、淒惘、淒淒、淒惻、淒惶、淒愴、悲郁、悲戚、惻怛、愀愴、愴惻、愴愴、憂戚、熬心、怛怛
		痛心	悲憤、悲痛、不堪回首、哀痛、痛切、痛定思痛、斷腸、人琴俱亡、五內如焚、五內俱裂、肝腸寸斷、肝腸痛斷、肝腸斷絕、長歌當哭、柔腸百轉、淒戾、創巨痛深、悲切、椎心泣血、腸斷、腸斷魂消、慘切、黯然銷魂……
		痛苦	傷痛、悲苦、黯然神傷、心如刀割、纏綿悱惻、回腸九轉、愧痛、慘苦、慘然
		辛酸	心酸、悲辛、悲酸、酸辛
		淒楚	淒切、慘慘
		兔死狐悲	物傷其類

詩詞韻文語意概念分類表：

一層	二層	三層	四層
人	人名	本名、他稱、別名字號、法號	
	君王	君王年號、君王帝號、君王陵墓	
	神鬼仙人	神鬼泛稱、神鬼專稱、菩薩尊號	
	稱謂	一般稱謂、自稱謙稱、宗教稱謂、指代稱謂、皇室稱謂、尊稱、貶稱、親人眷屬、職位身份	
	身體器官	五官顏面、五臟器官、分泌物、身體四肢、鬚鬢髮膚	
	感官詞	味覺、視覺、嗅覺、聽覺、觸覺	

	情緒詞	怒、怨、恨、苦、惆悵、喜、悲、悲歡、惱、感傷、愛憐、愁、憂、憎	
事	典故	一般典故、佛典	
時	時間	時間單位、明確時間、範圍時間	
地	地名	山嶺地名、外國地名、地名代稱、地名合稱、行政地名、河湖地名、關津地名	
	地理	邦國都城、郊原村野、特殊場域、區域範圍、街道巷弄	
	自然景觀	山川泛稱、山峰崖嶺、水澤湖泊、沙洲島嶼、林原莽漠、洞穴坑谷、專名、關津渡口	
	幽冥之境	幽冥	
	閬苑仙境	閬苑、仙境	
	空間	方向、位置	
物	神物	動物	動物專名（水生兩棲）、動物專名（合稱）、動物專名（走獸）、動物專名（昆蟲爬蟲）、動物專名（飛禽）、動物通稱
		植物	植物專名（木本）、植物專名（瓜類）、植物專名（禾本）、植物專名（合稱）、植物專名（果類）、植物專名（苔蘚）、植物專名（草本）、植物通稱
	生物	動物	動物專名（水生兩棲）、動物專名（合稱）、動物專名（走獸）、動物專名（昆蟲爬蟲）、動物專名（飛禽）、動物通稱
		植物	植物專名（木本）、植物專名（瓜類）、植物專名（禾本）、植物專名（合稱）、植物專名（果類）、植物專名（苔蘚）、植物專名（草本）、植物通稱
	器物	工具用品、公堂器物、水澤湖泊、古董、玉帛服飾、生活用品、交通工具、宗教用品、金帛錢幣、娛樂器具、神器、喪葬祭祀、樂器	
	建築物	寺廟道觀、官署建築、亭臺樓閣、建築專名、建築部件、軍事攻防、宮室屋廬、園林院落、碑塔陵墓	
	食物	合稱、食材、食品、宴席、菸草、飲品、零食、調味品	
	自然資源	火、玉石、合稱、金屬、燃料、礦物、烟火	
	文藝	引文、文具、文體、書篇名、書籍、單位	
	技藝	書法、畫、碁、舞蹈、戲劇、樂曲	
	天文	天空、日、月、合稱、星、銀河	
	天候氣象	合稱、冷寒、熱旱、風霜雪露、陰雨、晴天、雲霧煙霞、雷電雹電	
	四時節氣	四季、節氣、節慶	
	其他	行政體系	國名、部門名稱、朝代民族名號
災厄		火、自然災害、政變、礦物	
疾病		病名癥狀、醫療方法	
顏色		借代顏色、原本顏色	
味道		酸、甜、苦、辣、香、臭	
數量詞		數字	
單位詞		長度、重量、面積、時間、量詞	

專有名詞	田獵農耕、全球地理、年齡、行為活動、宗教、宗教流派、法條、法條法則、風俗禮儀、借代、婚喪喜慶、罪行、德行教條
------	--

《同義詞詞林》從語言學的學理做專家分類，詩詞韻文語意分類，最優先的考慮是文學文本內容的修辭特性與文體語言的特性。文學的內容以人為主軸，人、事、時、地、物五個要素是小說、戲曲內容必備的元素，詩詞因為受到字數的限制，在有限的文字裡時常省略一些用語，所以詩詞內容的文字，雖然不一定直接寫及這五項元素，但整體而言，所描述的內容，仍然包含了這些元素，例如：柳宗元的〈江雪〉：「千山鳥飛絕，萬徑人蹤滅。孤舟蓑笠翁，獨釣寒江雪。」詩中有人（孤舟蓑笠翁）、有事（獨釣）、有時（雪，冬天）、有地（寒江）、有物（孤舟）。李白〈靜夜思〉：「床前明月光，疑是地上霜。舉頭望明月，低頭思故鄉。」沒有直接寫人，但是「舉頭望明月，低頭思故鄉。」的是人。又如杜甫〈八陣圖〉：「功蓋三分國，名成八陣圖。江流石不轉，遺恨失吞吳。」不直寫人事時地，而以典故帶出：人（諸葛亮、劉備、孫權）、事（吳蜀魏三分天下）、時（漢代末年吳蜀魏三分天下時）、地（劉備被陸遜大敗於猇亭，今湖北宜都縣北）、物（八陣圖）。詩詞內容雖然不一定直接寫人，但是仍然以人為主軸，書寫以人為主軸的事、時、地、物，所以將語意類別的最上層，分為人、事、時、地、物，難以歸入這五類者，另歸為其他類。

詩詞文體語言的特性方面，例如文學的風格與內容受到地理環境的影響，因此地理景觀詞彙需要更細緻的分類。又如詩詞裡有傷春悲秋的传统，春天傷失偶，秋天悲失路，季節是文學情感表現非常重要的因素。晨昏晝夜不同的時段，對詩人心境的影響很大，獨守空閨的怨婦，最容易思念丈夫的時間是黃昏及夜晚；浪跡天涯的遊子，最容易思念家鄉的時間也是黃昏及夜晚。抒情婉約的言情詩與豪邁雄壯的言志詩，對時空距離的描寫也有所不同。由此可知，語意的分類，無論是對於解讀文學作品內涵，或是作為文學寫作背景因素、文學風格的研究，都佔有重要的地位。

《同義詞詞林》未將專有名詞定為一類，但是以詩詞韻文文本內容為範圍，以作為「提供文學研究」目的的詩詞語意分類，有必要增加名詞的類別，並且予以區分為「人名」、「地名」、「朝代君王」、「建築專名」、「器物專名」、「書籍篇章專名」……等等，予以分類。以「地名」分類為例，文學作品風格受到地理環境背景的影響，地理環境與地貌景觀受到自然天候影響，北方的地景、地貌與南方不同，形成北方文學（如《詩經》）與南方文學（如《楚辭》）風格不同。地名的類別與標記，即可判斷其地理位置，但是要再進一步區分該地名是作品的「寫作地點」或「寫及地點」⁵⁰。再以「人名」為例，從文學作品所寫及的人名，可以從中分析作者與寫及對象彼此之間的關係、互動與影響，進而作為解讀作品的依據，如前述杜甫〈江南逢李龜年〉詩，需了解杜甫與李龜年的背景與互動關係，才能正確解讀詩意。詩詞韻文語意概念的分類，針對詩詞語言的特性以及詩詞文本內容含義、修辭特色等角度分類，是兼顧語言學及文本內容特性與詩詞韻文修辭技巧、解讀之需的多重考慮所做的分類，除了考慮知識的認知層次，同時考慮情感的感知層次，因此與純粹從語言知識角度所做的分類有所不同。也正因為這些因素，所以本研究的分類未沿用《唐詩分類大辭典》及《同義詞詞林》的分類，而回歸到以文本內容知識內涵，並兼顧文學領域的領域特性分類，以符合文學研究與教學之需。

⁵⁰ . 參見拙作〈文學地理時空資訊系統設計與應用：以蘇軾詩為例〉，第二屆數位地球國際研討會，2004年5月27-28日，中央研究院、國家實驗研究院、中國文化大學主辦。

語意概念分類之前，需先做好語意標記。詩詞以詞語的節奏及平仄押韻表現其音韻聲情之美，以雙字詞為主要的詞組單位，但是有些雙字詞從語言學的角度應該要予以切分，例如：「離恨恰如春草，更行更遠還生」(李煜〈清平樂〉)，從詩詞語言的角度，切分為「離恨／恰如／春草／，更行／更遠／還生」，後句之「更行／更遠／還生」都是「副詞＋動詞」，從語言學的角度，應該予以切分，但是從詩學的角度，便失去了詩詞語言的特性以及詩詞的節奏音步。再從這一句的「春草」來看詩詞語意標記的困難，「春草」顧名思義是指「春天的草」，其語意含有「春」(季節)及「草」(植物)二種概念，其語意的分類應該歸入「春」還是歸入「草」？唐宋词總計 24,226 闕，含有「草」的詞彙有 2,268 筆，含有「春」的詞彙有 10,528 筆。從整闕詞(別來春半。觸目愁腸斷。砌下落梅如雪亂。拂了一身還滿。雁來音信無憑。路遙歸夢難成。離恨恰如春草，更行更遠還生。)來分析，這裡的「春草」側重的詞義是含有離別意象的「草」，而不是「春」，語意的分類應該歸入「草」，而不是「春」，「春草」與「秋草」裡的草，都含有離別的意象，春季與秋季在詩詞裡的含義各有其詞義、意象與意境的作用，不能忽略。但是，若將「春草」切分為二個單字詞，也同樣失去了詩詞語言的特性。詩詞中還有很多「形容詞＋名詞」也有同樣的問題，如「春鴻」、「孤雁」、「秋猿」、「晚蟬」、「翔鸞」、「垂柳」、「幽蘭」、「修竹」……等等，不勝枚舉。這些詩詞用語現象，都使得詩詞語言的詞彙切分與語意概念標記與分類產生很大的困難。

八、植基於語意學及使用者認知觀點的全唐詩全文檢索設計

經過語意標記之後，電腦便初步具備解讀語意的能力，經過語意概念分類，及建立上下層從屬關係的結構之後，語意概念的類別及從屬關係的邏輯已經建立起來，使用者執行全文檢索時，比傳統的全文檢索，多了許多功能。全唐詩全文檢索介面如下：



檢索欄位包含作者、卷別、詩作主題類別、詩題、詩序、詩句語意類別、詩文等七個欄位。作者欄位無論輸入作者本名或別名字號都可以檢索，作者、詩題、詩序、詩文等四個欄位提供布林檢索功能，詩文欄位的布林檢索再區分為「以句為單位」或「以整首詩為單位」；詩作主題類別及詩句語意類別的類別層次以下拉式選單選擇。

以語意概念檢索的結果，與傳統的比對字形的檢索結果有那些差別？以「銀河」為例，使用者想要翻檢全唐詩中有多少詩寫到天上的「銀河」，傳統的全文檢索方法是輸入「銀河」，共搜尋出 100 筆資料，以語意概念檢索，共搜尋出 355 筆資料，與「銀河」同義的「天漢」、「雲漢」、「河漢」、「銀漢」、「天河」、「星橋」、「遠漢」……都可以檢索出來。再以「星」為例，以語意為基礎的全文檢索，除了可以將「天府、天狗、天海、天狼、天淵、天剛、天樞、天廚、太乙、太微、少微、天囷、太白、太歲、斗牛、斗牛箕、斗轉、牛斗、北戶、北斗、北辰、玉繩、列宿、長庚、南箕、昴、參橫、牽牛、眾宿、槍、箕、箕斗、蒼龍、欖、河鼓、紫薇、朱鳥、朱雀、婺女、玉衡……」等等與「星」同義的詞彙檢索出來，還可以先將「福星」、「壽星」等等含有「星」字，但是不是天文概念的「星」的資料剔除。

詩句布林檢索功能，可以做為詞彙共現的研究，提供「以句為單位」或「以整首詩為單位」二種選項，有其必要性，以前述之「苔蘚」或「苔」、「蘚」與「階」、「門」共現的檢索為例，全唐詩寫及「苔蘚」的詩句有 112 筆，寫及「階」的詩句有 960 筆，「苔蘚+階」以句為單位，檢索結果共有 11 筆資料；以整首詩為單位，檢索結果共有 15 筆資料。全唐詩寫及「苔」的詩句有 1,390 筆，以「苔+階」以句為單位，檢索結果共有 44 筆資料；以整首詩為單位，檢索結果共有 94 筆資料。

上下位層的語意概念關聯架構，語意概念的分類對使用者認知觀點的引導，多重欄位選擇搭配的檢索策略設計，提供使用者更多元的檢索方式，完成其檢索任務，滿足研究上的需求。

植基於語意學及使用者認知行為的資訊檢索系統設計，可以降低使用者的知識異常狀態 (Anomalous State of Knowledge)。任何一種檢索系統，永遠都會有新手使用，永遠都會有因為是新手而產生的檢索策略困窘以及對知識內容了解不足的知識異常狀態而降低檢索的質量。檢索策略的困窘，可以透過使用說明及搜尋引擎、功能的設計解決；斷詞、語意標記、語意概念的建立，能夠在檢索時產生「引導」的作用，降低知識異常狀態，彌補對知識內容了解的不足。例如使用者要檢索天文概念的「星」，可能因為知識的不足，無法掌握所有與「星」的概念同義的詞彙，語意的標記與概念的分類，讓使用者可以完整掌握所有的同近義詞詞彙，使檢索的結果更完整。

在資訊檢索的行為過程中，還常出現資訊偶遇 (Information Encountering) 的狀況，既是偶遇，便含有可遇不可求的機率，語意概念的分類以及上下層語意關聯的結構，可以引導使用者往上層語意概念擴大搜尋或檢索的範圍，也可以往下層語意概念縮小搜尋或檢索的範圍，這是化消極的等待資訊偶遇成為積極的創造資訊相遇的搜尋管道。

此外，建立各種不同的索引典，以視覺化的動態方式呈現知識結構，使知識的體系一目了然，提供按圖索驥的功能，既可降低知識落差而產生的知識異常狀態，又可以提高資訊偶遇的機率，甚至於化被動的偶遇而成為主動的知識引導。

九、結論與未來的發展

詩詞語言的特殊性，包含構詞的特殊、詞義的特殊，篇章結構的特殊，這些特殊性增加了詩詞詞彙切分及語意標記、語意分類的困難，也使得詩詞語意概念的分類，除了必須符合知識認知層次的語言知識分類，還需顧及情感認知層次的意象取向分類。詩詞的詞彙以雙字詞居多，但是詩詞句子的字數多數是奇數，所以多數奇數字數的句子之中都會有一個單字詞，單字詞的詞義比雙字詞更豐富多變，這也增加了詞義標記的複雜性。

詩詞裡一般用語及典故用語的標記方式，前已述及，甚至於以某些詞彙共現以烘托某種情境的修辭方式，必然也可以用更多相關的研究，用更多分析、統計、歸納的方法尋找出規則。

詩詞語意標記的困難，還包括詩詞喜用多義字來創造聯想的空間，以引出無窮的趣味及意在言外的情境。袁行霈教授〈中國古典詩歌的多義性〉文中即闡明：「一個多義詞固然有多種意義，歧義是一般情況下使用語言時需要特別避忌的毛病。但是在詩詞裡，恰恰要避免詞義的單一化，總是盡可能地使用詞語帶上多種意義，以造成廣泛的聯想，取得多義的效果。中國古典詩詞的耐人尋味，就在於這種複合的作用。」⁵¹，文中袁教授並將詩詞語言的多義性分為宣示義及啟示義二種，啟示義包含雙關義、情韻義、象徵義、深層義、言外義五種⁵²。宣示義是借助語言明確表達給讀者的語意，如前所述王維〈積雨輞川莊作〉：「漠漠水田飛白鷺，陰陰夏木轉黃鸝。山中習靜觀朝槿，松下清齋折露葵。」只有字面含義。雙關義、情韻義、象徵義能夠從統計學歸納出規則，能夠從單一詞彙標記其多重含義，如前述的「苔、蘚」，但是深層義與言外義的含義隱藏在字詞表面含義之外，以整首詩來烘托，這些詩即是所謂「言有盡而意無窮」、「不著一字而盡得風流」者，讀者要了解這些深層含義，必須要了解作者及寫作對象的背景資料，必須要具備更豐富的學養及詩詞鑑賞能力才能夠體會，以詞彙為單位的語意標記無法標記，想要讓電腦理解，誠非易事，如前述杜甫的〈江南逢李龜年〉詩即是。

宋代以後，受到理學發達及禪宗思想的影響，有很多的哲理詩，如蘇軾〈題沈君琴〉：「若言琴上有琴聲，放在匣中何不鳴？若言聲在指頭上，何不於君指上聽？」楊萬里〈戲筆〉：「野菊荒苔各鑄錢，金黃銅綠兩爭妍。天公支與窮詩客，只買清愁不買田。」這些哲理詩正是劉大魁所說的「只可意會而不可言傳」的作品，需仰賴讀者的哲理思維領悟，當然也不是單獨的詞義標記所能夠清楚描述的，想要借助詞義標記或語意概念分類讓電腦能理解，其難度又比深層義及言外義更高，在分類上，只能以詩的主題作分類。

⁵¹ . 引自袁行霈，〈中國古典詩歌的多義性〉，收於《中國詩歌藝術研究》，五南圖書出版公司印行，1989年5月台灣初版，第6頁。

⁵² . 同註27。

除了構詞與詞義的特殊性之外，詩詞文體的格律限制，句法結構與其他文體不同，詩詞常因受限於平仄格律而使用倒裝句，如元稹〈遣悲懷〉：「謝公最小偏憐女，自嫁黔婁百事乖。」即是「謝公偏憐最小女」的倒裝句，這個句子因為是在同一句倒裝，而且所使用的詞義只有知識認知的層面，不構成語意標記的困難。蘇軾〈張子野年八十五，尚聞買妾，述古今作詩〉：「詩人老去鶯鶯在，公子歸來燕燕忙。」因為對仗的需求，將「鶯鶯燕燕」的詞彙拆解為「鶯鶯」與「燕燕」，而且分置於上下句，便構成詞彙切分及語意標記的困難。又如蘇軾〈和子由柳湖久涸，忽有水，開元寺山茶舊無花，今歲盛開二首：其一〉：「如今勝事無人共，花下壺盧鳥勸提。」這個句子的下句「花下壺盧鳥勸提」，完整的詞句是「花下提壺盧鳥勸人提起葫蘆喝酒」，使用了倒裝、典故、雙關義、縮略語等修辭技巧⁵³，由人解讀尚且困難，更遑論以電腦解讀。

從以上所述可知，可以透過語意標記及語意概念分類，讓電腦從分辨字形提升到理解詞義，使電腦的認知能力更接近人的知識體系，以提升全文檢索的品質。人的知識體系分為「對知識的認知」、「對情意的感知」、「對哲理的辨知」，認知是表述的層次，感知需要聯想輔助，哲理詩的解讀體會，卻不是單純的邏輯思維能解決，也就更難以用語意標記處理。

資訊檢索所涉及的相關議題，包括認知的元素、語意與多重表述的認知觀點，詞語概念及表述的關聯性，資訊行為的研究、使用者的交互作用，檢索的策略、搜尋引擎、檢索方法的應用等等。本研究從語意學及使用者的認知觀點，以全唐詩網站的設計作為實證的研究，初步建立全唐詩語意標記資料庫及全唐詩語意概念分類表，期待後續的研究，能進一步彙整、分析、歸納唐詩語意類別，再與共現詞彙比對，期望能從中建立唐詩主題分類的方法與規則，以彌補《唐詩分類大辭典》的不足。

全唐詩語意標記及分類的架構是以全唐詩的文本內容為範圍，並不包含全唐詩內容之外的其他詞彙或知識，本研究進行的過程是從蘇軾全部詩作為基礎，建立蘇軾詩作的詞彙庫及語意標記及分類資料庫，其後加入了宋代七家詞人⁵⁴、四百萬字的全台詩⁵⁵，以這些文本的詞彙庫及語意標記及分類資料庫為基礎，進行唐詩的詞彙切分及語意標記及分類。整合不同文體（詩、詞）、不同時代（唐、宋、清末至民國）、不同寫作背景（中國大陸及台灣）但是構詞方式相近的韻文文本，所包括的詞彙也因為文體、時代、寫作背景的差異而有些微不同，因此需要再進行人工校對。

全唐詩網站所包含的文本數量龐大，人工校對所需要花費的人力很多，加上虛詞的語意標記及分類的困難度比實詞高，本研究的語意標記及分類只針對表中項目所訂的類別分類，撰寫本文時，還來不及完成人工校對，語意的分類也不完整、不精確，尚待後續的研究修正及補足。

⁵³ . 參見拙作，〈蘇軾詩典故用語研究〉，第五屆漢語詞匯語意學研討會，2004年6月14-17日，新加坡國立大學主辦。

⁵⁴ . 七家詞人包含柳永（218闕詞）、蘇軾（379）、秦觀（148）、周邦彥（187）、李清照（53）、姜夔（87）、辛棄疾（627）等，選擇的標準考量北宋與南宋的時代均衡及婉約派與豪放派的風格均衡。

⁵⁵ . 參見羅鳳珠主持，【智慧型全台詩知識庫計畫第二階段建置計畫】，國家台灣文學館，2006年1月1日至2006年12月31日，網址：<http://cls.hs.yzu.edu.tw/TWP/index.asp>。

致謝：本文由羅鳳珠執筆。感謝國科會贊助本研究計畫：【新詩改罷自長吟—唐詩吟唱虛擬實境教學環境建置計畫】，國科會數位典藏國家型科技計畫：95 年度數位典藏創意學習計畫，2006 年 3 月 1 日至 2007 年 2 月 28 日執行。計畫編號：NSC95-2422-H-155-001。感謝計畫共同主持人張智星、許介彥教授，感謝參與計畫之助理：程式設計陳駿銘、李元萍、曹偉政；資料分析整理：吳侑倣、詹錦麗、黃詩媛、劉怡姍；網頁視覺設計：陳憶玟。