

香港城市大學
中文、翻譯及語言學系

資訊科技在人文領域的應用與發展

元智大學・中國語文學系

羅鳳珠

2009年3月25日

大綱

- 一. 前言
- 二. 回顧：資訊科技在人文領域的發展
- 三. 如何讓電腦更接近人腦
- 四. 語意標記及語意概念分類類別
- 五. 語意標記及語意概念分類的應用
 1. 提升全文檢索質量：以唐宋詞網站為例
 2. 應用於文學情感計算的研究
 3. 應用於文學地理資訊的研究
- 六. 應用於文學計量研究：唐宋詞領字研究
- 七. 未來的發展：建立語言知識庫

前 言

電腦：具備了比人腦更強的「廣讀群書，博聞強記、搜尋檢索、分析統計」能力。

但是，電腦只能分辨字形的形符，無法解讀字詞含義的意符。不能理解語義，不具有思考能力和人的認知、感發、思維能力，不具備人的知識體系，所以無法「旁徵博引」，無法「觸類旁通」。

回顧：資訊科技在人文領域的發展

- 1987年，應《國文天地》編輯的邀請，以《探一探文史資料自動化的路》為題，訪問資訊科學領域張仲陶教授，文史領域周何教授、毛漢光教授、王邦雄教授、王熙元教授等五位學者，從資訊科學及經、史、子、集不同領域的角度，提出他們對文史古籍數位化的看法，並從中探索發展的方向。

回顧：資訊科技在人文領域的發展

一、資訊科學領域

- 張仲陶教授：「不要問電腦能做什麼？而是問你要電腦做什麼？」的看法，主張由文史界負責提出需要，由資訊界負責滿足需要：「電腦做得比人好的，交給電腦做，人去作電腦還做不到的事。」

二、經學領域

- 周何教授：「電腦是很呆板的東西，但怎樣使他具有高層次的功能，幫助人腦體會，這是我所期望的。」（經學）

回顧：資訊科技在人文領域的發展

三、史學領域

- 毛漢光教授：「就個人經驗言，在文史自動化的過程中，成敗的關鍵在文史界，不在電腦界。依我的經驗，個人想到的，只要分析出來，電腦都可以做到。所以文史自動化成敗的關鍵在文史界，不在電腦界。」（毛漢光教授參與中央研究院第一期史籍自動化計畫，負責《食貨志》資料分析、規劃、系統分析，比其他幾位文史領域的教授，多了實際參與文史資料數位化的實務經驗。）

四、哲學領域

- 王邦雄教授：「文史自動化不能失去人的主導地位。電腦畢竟不是人，無法做創發性的工作。」

五、文學領域

- 王熙元教授：「研究工作最重要的是資料的運用，將研究資料經過分析、整理、歸納，分門別類建立資料庫，才能符合文史研究的需要。」

回顧：資訊科技在人文領域的發展

臺灣中央研究院於1984年7月開始的「史籍自動化計畫」，除了參與該計畫的張仲陶、毛漢光二位教授之外，其餘三位教授未曾使用過數位資料庫，在當時也不會用電腦。

彙整五位學者所提出的觀點，有三項共同處：

- 一. 電腦不能取代人腦；
- 二. 引用電腦節省處理資料的時間，人可以做更多思考性、創發性的工作；
- 三. 古籍數位化成敗的關鍵在人文學界，不在電腦學界，需要由人文學界提出需要，電腦學界滿足需要，二者的通力合作才是古籍數位化成功的關鍵。

而今時隔23年，張仲陶、周何、王熙元三位教授均已作古多年，他們所提出來的觀點，仍然是人文學界與資訊學界努力的方向。

回顧：資訊科技在人文領域的發展

- 1996年在元智大學講授詩詞課程，每每苦於詩詞習作批改的負擔，突然體會到張仲陶教授所說「電腦做得比人好的，交給電腦做，人去作電腦還做不到的事。」
- 今人作詩填詞最大的困難是無法熟記詩詞格律譜以及平仄聲調、韻目韻字，而這些都是固定的，正是電腦可以做得比人好的部分，於是申請經費，開發詩詞格律自動檢測索引教學系統。

采桑子—平韻格 一格一

+-+ | -- | , + | -- ° (韻) + | -- ° (韻) + | --+ | - °
(韻) & +-+ | -- | , + | -- ° (韻) + | -- ° (韻) + | --+ |
- ° (韻)

本篇韻目: 第十二部 輸入文字: 收 查詢韻目

自訂詞題:

最難消受簷前雨，朝也難收。暮也難休。雁喚蛩鳴添客愁。&當年一別音書絕，人也難留。夢也難求。獨擁孤衾憶舊遊。

輸入標點: ° , ` & 檢查格律 作品存檔

最難消受簷前雨，朝也難收。暮也難收。雁喚蛩鳴添客愁。&當年一別音書絕，人也難留。夢也難求。獨擁孤衾憶舊遊。

錯誤1(遊): 此字必須是平聲
錯誤2(遊): 此字必須押韻(第12部)

填詞輔助工具

- 查詢牌詞韻
- 查詢唐宋詞
- 查自建詞庫
- 查訪客作品
- 發表作品
- 使用者註冊
- 給我們意見

回顧：資訊科技在人文領域的發展

- 1997年8月赴北京大學訪問，得有機會持此系統向袁行霈教授請教，袁教授問我，**這個系統能將用錯的字改成對的，能不能將用得不好的字改得更好？**
- 誠如袁教授所說的，這個系統能將用錯的字改成對的字（完全符合平仄押韻），但是不但不能改成語義情境更好的字詞，甚至於在作自動檢測時，對於一首平仄格律完全正確，但是文句語義全然不通的作品束手無策，**關鍵在於電腦不具有理解語義的能力，更不具有思考、感發的能力，甚至於連面對多音字的平仄判斷能力都沒有，而人腦具備這些能力。**

如何讓電腦更接近人腦

- 如果讓電腦具備理解語義的能力，就必須讓電腦的運算更接近人腦的思維過程，需要建立符合人類知識體系的語義概念資料庫，使電腦初步具備人的知識概念體系（Ontology）。
- 1945年麻省理工學院副校長Vannevar Bush構思著建立一個儲存及檢索的機器，能儲備所有文獻，檢索時能符合人類的思維邏輯，建立知識連結的軌跡以進行聯想索引，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。（As We May Think）
- Bush嘗試要解決的便是以電腦幫助及代替人腦做某些事。

如何讓電腦更接近人腦

- 1956年電腦與人腦比較的議題開始受到關注，美籍匈牙利科學家John von Neumann應耶魯大學邀請演講，演講的主題是電腦與人腦，並與合作夥伴開始研究建構一個與活腦極為相似的人造機器的可能性，但是由於對「人腦的語言」（language of the human brain）還是很不清楚，以致于並未成功。

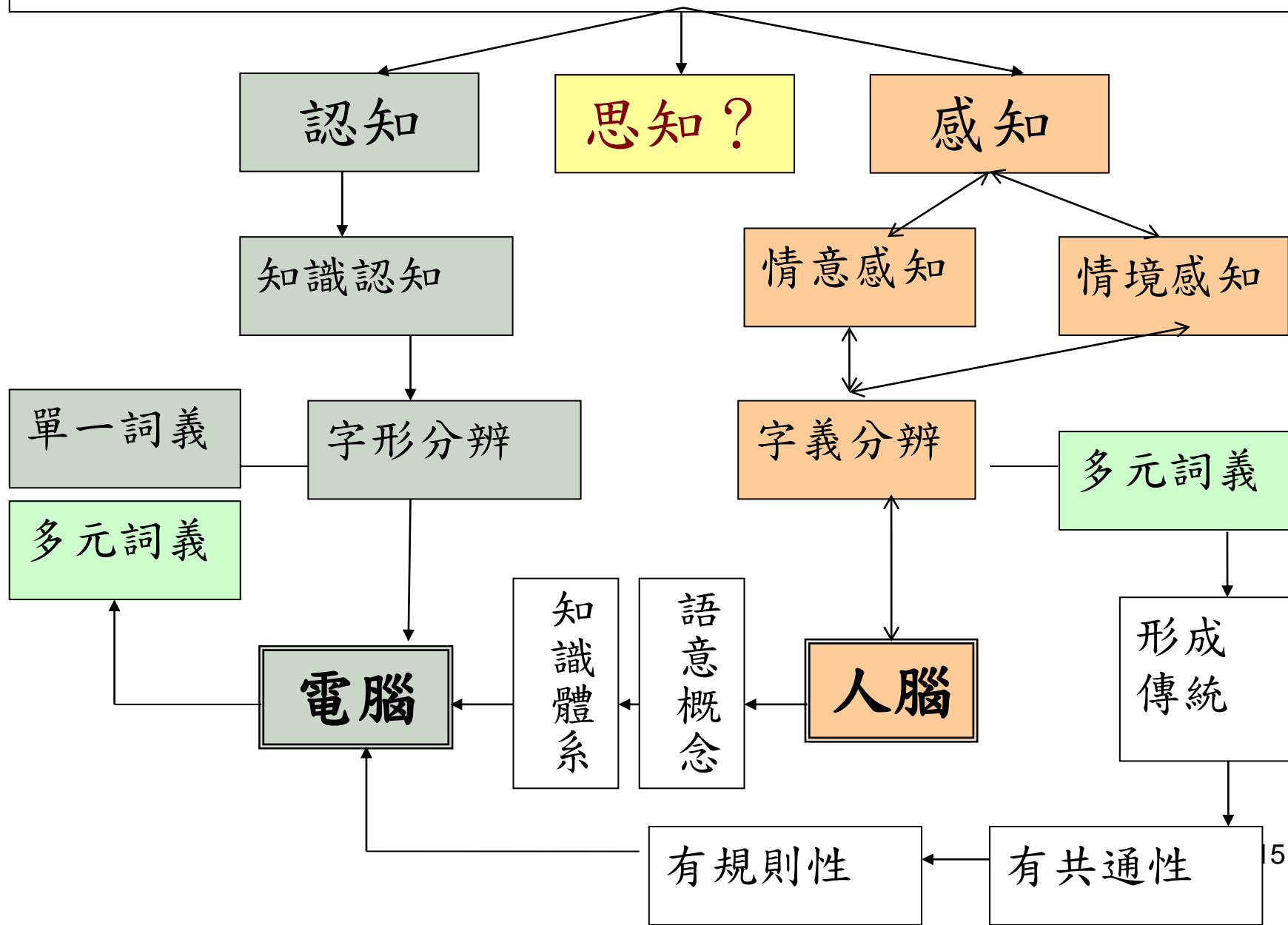
如何讓電腦更接近人腦

- 1970年代美國哲學家弗德（J. Fodor）提出了「思想語言的假設」（Hypothesis of Language of Thought），引起廣泛的討論。
- 翻譯《電腦與人腦》一書的譯者蔡耀明在譯者序文《電腦科學的瞻前顧後》指出：「人腦（心智）的語言到底是如何，仍是未定之數。在方興未艾的認知科學，以結合哲學、語言學、認知心理學、腦神經科學、人工智慧等跨學科共同研究的努力下，相信總有一天會解開人類心智的奧秘，屆時要實現以人類心智為模型的電腦，就不再是難事了。」（引自約翰·馮諾曼（John von Neumann）著，蔡耀明譯，《電腦與人腦》，臺灣商務印書館，2000年初版，〈電腦科學的瞻前顧後〉（譯者序），第7頁。）

如何讓電腦更接近人腦

- 以語意標記使電腦從分辨字形的能力提升到分辨字義的能力。
- 以語意概念分類幫助電腦理解人的知識體系，具備認知的能力。
- 以情感計算（Affective Computing）及情境感知（Context Awareness）提升電腦情感識別與情感表達的能力，使電腦具備感知的能力。

如何讓電腦更接近人腦：以詩詞賞析為例

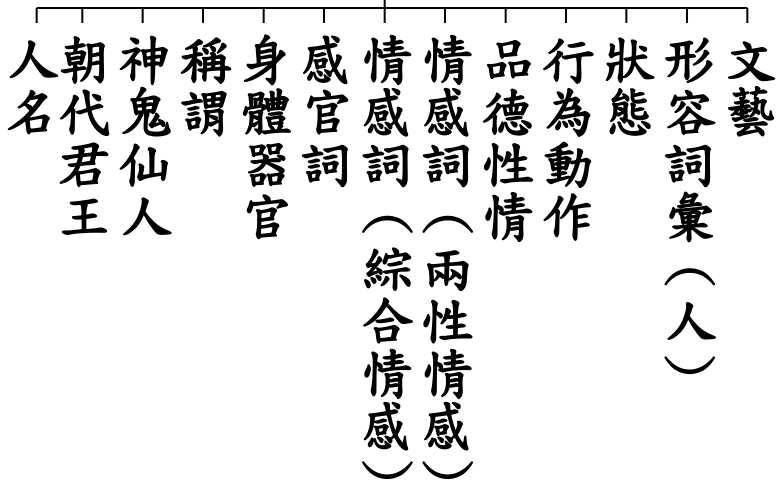


語意標記及語意概念分類

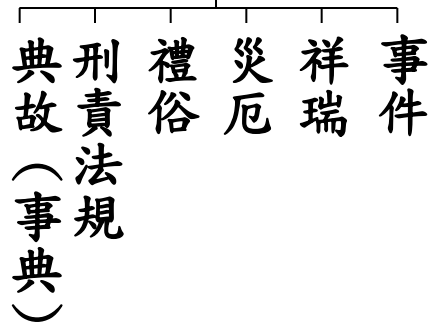
- 在現代漢語方面，中文詞語的語義分類研究成果，紙本文獻以梅家駒主編，上海辭書出版社出版的《**同義詞詞林**》最具代表性。
- 電子本文獻方面有董振東先生于1988年創建的**知網**（HowNet Knowledge database），2003年臺灣中央研究院資訊所詞庫小組陳克健教授與董振東先生展開建構繁體字知網的合作計畫，建立「**概念網**」以及北京大學**語言知識庫**。
- 筆者于2002年起，所建立的**詩詞語意標記與語意概念分類知識庫**。

語意概念分類總表

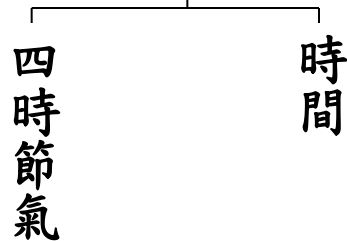
人



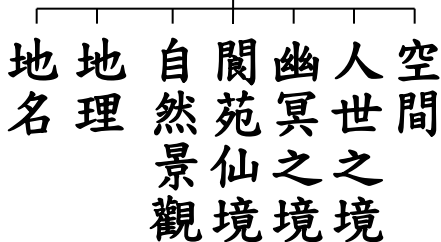
事



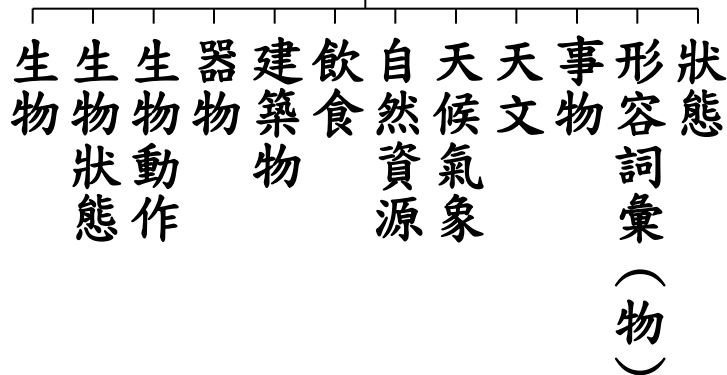
時



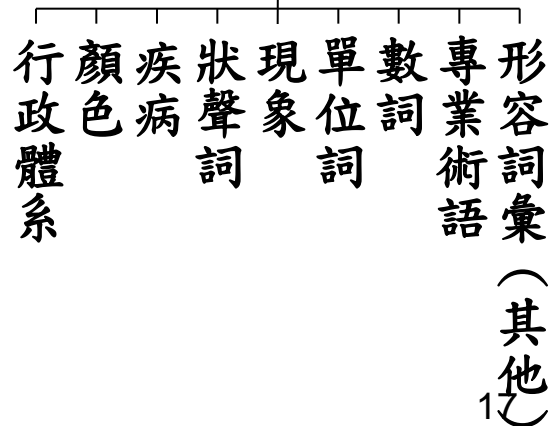
地



物



其他



物

狀態 → 對比狀態

形容詞彙 (物)

↓ 景物形態、植物形態、動物形態、
物品形態、環境形態

事物 → 事物概稱

天文 → 日、月、星、天空、銀河、合稱 (同類、異類)

天候氣象

↓ 風霜雪露、雲霧煙霞、雷電電霓、冷寒、熱
早、晴天、陰雨、合稱 (相反詞、相近詞、
相異詞)、概稱

自然資源

↓ 玉石、火、金屬、礦物、燃料、合稱 (同類、
異類)

飲食

↓ 飲品 (茶、酒、其他)、食品、食材、糕點、調
味品、菸草、宴席、合稱

建築物

↓ 建築專名 (宮室屋廬、亭臺樓閣、園林院落、寺廟道
觀、碑塔陵墓)、宮室屋廬、亭臺樓閣、園林院落、
寺廟道觀、碑塔陵墓、公署建築、建築部件、軍事攻
防

器物

↓ 生活用品 (飲食器具、烹飪器具、裝置容器、起居用
品、其他用品、梳妝用品)、交通工具 (水路、陸路、
空路、合稱)、工具用品 (兵器單稱、兵器合稱、戰
爭物品、農牧漁獵工、生活工具單稱、生活工具合稱、
刑具)、絲帛服飾 (衣冠腰帶、布料、鞋襪、配飾、
合稱)、宗教用品、喪葬祭祀、娛樂器具、金帛錢幣、
公堂器物、旗幟、帳幕、神器、樂器 (單稱、合稱、
泛稱)、文具 (單稱、合稱、泛稱)、器物合稱、器
物泛稱

生物動作

↓ 動物動作、植物動作

生物狀態

↓ 動物生命狀態、植物生命狀態

生物

↓ 神物 (動物、植物、合稱、泛稱)、動物專名 (走獸、
飛禽、昆蟲爬蟲、水生兩棲、合稱)、動物指稱 (皮
毛、泛稱、其他、合稱)、植物專名 (木本、草本、
禾本、苔菌、瓜果、合稱)、植物泛稱 (花、木、葉、
枝、果、草、莖、根、苗、合稱)、異類合稱

語意標記及語意概念分類的應用

一、對文學研究與教學的影響

1. 拓展文學研究的議題
2. 以科學數據作為文學研究的立論基礎
3. 提供詩詞語言演變研究
4. 提供詩詞韻文教學及習作使用

二、在文學研究的應用

1. 提升全文檢索質量：以唐宋詞網站為例
2. 應用於文學情感計算研究
3. 應用於文學地理資訊研究
4. 其他文學計量研究：以唐宋詞單字領字研究為例₁₉

語意標記及語意概念分類 對資訊檢索質量與使用需求的影響

一、對資訊檢索質量的影響

1. 提高全文檢索的**完整性**
2. 提高全文檢索的**正確性**
3. 使資料的搜尋**從索引式變為關聯式**

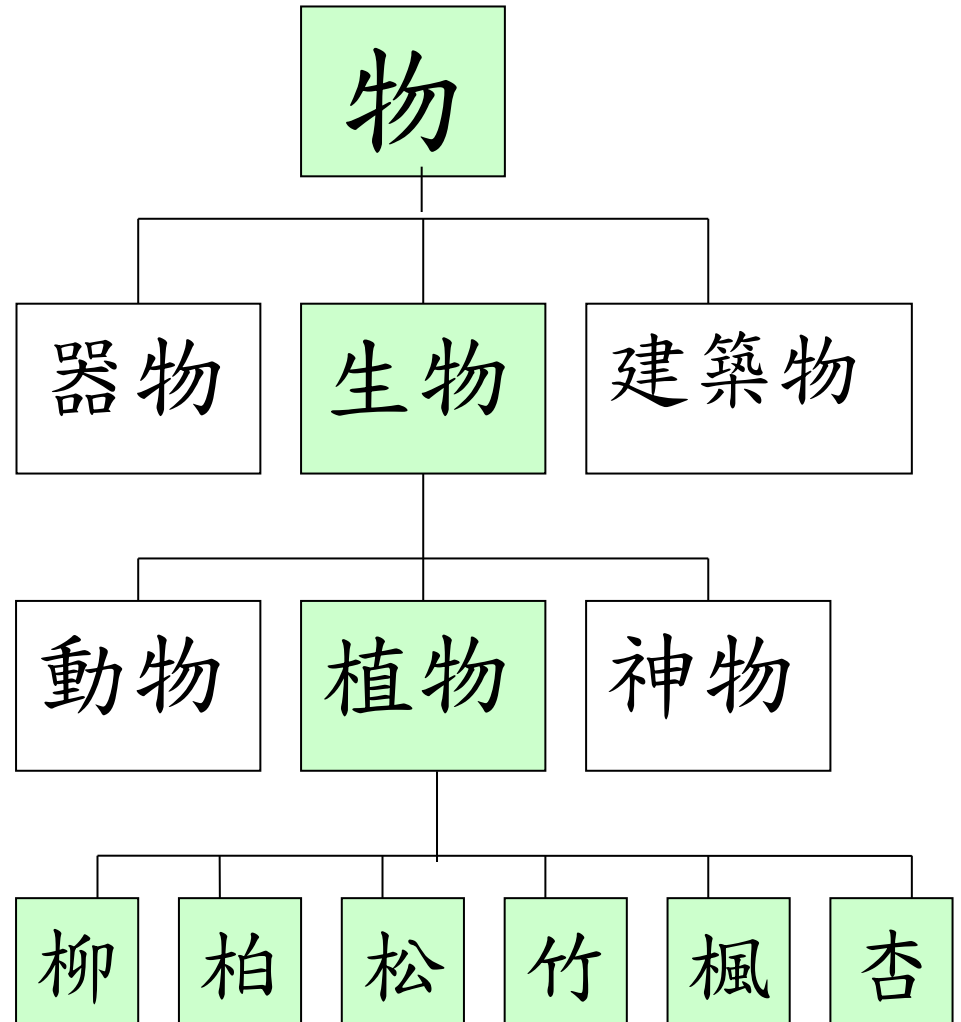
語意概念標記應用於全文檢索之驗證： 以唐宋詞網站為例

語意概念分為總類、大類、次類、元素四層。

往上層擴大搜尋範圍。

往下層縮小搜尋範圍。

同一層為同一知識領域範圍。



語意概念檢索

有布林功能的欄位可使用「&」=AND、「/」=OR、「-」=NOT

作者：	<input type="text"/>	(含別名字號/布林)
朝代：	請選擇-	
詞牌：	<input type="text"/>	
詞題、序：	<input type="text"/>	(含詞序)
詞文：	<input type="text"/>	(字形/布林檢索)
詞句語意 類別：	總類：請選擇	(測試中)
	大類：請選擇	
	次類：人事	
	元素：時地物其他	
宮調：	請選擇-	
目前只有蘇軾(1036~1101)詞含有作品時空標記資料(1060~1101)		
註解：	<input type="text"/>	
寫作時間：	<input type="text"/>	(紀元年ex.宋仁宗嘉祐五年)
寫作時間：	<input type="text"/>	(西元年ex.1060)
寫作年齡：	<input type="text"/>	
寫及地點：	<input type="text"/>	
寫作地點：	<input type="text"/>	
<input checked="" type="radio"/> 字形檢索		
<input type="button" value="開始檢索"/> <input type="button" value="取消"/>		

語意概念標記應用於全文檢索之驗證：

以唐宋詞網站為例

語意概念標記應用於全文檢索之驗證： 以唐宋詞網站為例（物→天文→星）

以語意概念檢索「星」結果：2,076筆

共計2076筆

... 41 42 43 44 45 46 47 48 49 50 ...

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文	詞中辭彙	辭彙總類	辭彙大類	辭彙次類	辭彙元素
401	阮郎歸	王之道	宋代	二首之一：和廣濟王宰		玉繩低轉斗闌干。	玉繩	物	天文	星	玉繩
402	醉蓬萊	王之道	宋代	代人上高御帶，時在太學		長庚誕貴。	長庚	物	天文	星	長庚
403	好事近	王之道	宋代	何希淵生日		曾見少微初降，	少微	物	天文	星	少微
404	好事近	王之道	宋代	董令升生日		歸/去騎箕尾。	箕尾	物	天文	星	尾星、箕星
405	臨江仙	王之道	宋代	和劉南伯		疏星明耿耿，	疏星	物	天文	星	星

以字形比對檢索「星」：1,427筆

http://cls.hs.yzu.edu.tw/CSP/W_DB/index.htm

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H) 連結 >>

【淺顯低唱】若有知音見採—唐宋詞全文資料庫

共計1427筆

... 31 32 33 34 35 36 37 38 39 40 ...

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文
351	浣溪沙	曾覿	宋代	櫻桃		滿林璀璨綴繁星。
352	好事近	王之望	宋代	和榮太監		佳句麗星日。
353	減字木蘭花	王之望	宋代	恭人生日		與爾尊前作壽星。
354	醉花陰	王之望	宋代	生日		休畫老人星，
355	風流子	王之望	宋代	范覺民生日		中有麗天星斗，
356	錦堂春	葛立方	宋代	正旦作		星翁說道宜官。
357	滿庭芳	葛立方	宋代	五姪將赴當塗，自金壇來別		別來幾換星霜。
358	水調歌頭	王識	宋代	觀星		初出極星未遠，
359	水調歌頭	毛并	宋代	上元郡集		十里寒星相照，
360	漁家傲引	洪适	宋代	伏以黃童白叟，皆是煙波之釣徒；青笠綠蓑，不識衣冠之盛事。長浮家而醉月，更輟棹以吟風。樂哉生涯，翻在樂府。相煩女伴，漁父分行。		碧落雲高星爛爛。

... 31 32 33 34 35 36 37 38 39 40 ...

語意概念標記應用於全文檢索之驗證： 以唐宋詞網站為例（物→天文→月）

語意概念檢索

有布林功能的欄位可使用「&」=AND、「|」=OR、「-」=NOT

作者：	<input type="text"/>	(含別名字號/布林)
朝代：	請選擇-	
詞牌：	<input type="text"/>	
詞題、序：	<input type="text"/>	(含詞序)
詞文：	<input type="text"/>	(字形/布林檢索)
詞句語意 類別：	總類： 物	(測試中)
	大類： 天文	
	次類： 請選擇	
	元素： 請選擇	
宮調：	請選擇-	
目前只有蘇軾		空標記資料 (1060~1101)
註解：	<input type="text"/>	
寫作時間：	<input type="text"/>	(紀元年ex:宋仁宗嘉祐五年)
寫作時間：	<input type="text"/>	(西元年ex:1060)
寫作年齡：	<input type="text"/>	
寫及地點：	<input type="text"/>	
寫作地點：	<input type="text"/>	

Dropdown menu options for 詞句語意類別:

- 合稱 (同類)
- 合稱 (異類)
- 星
- 銀河

以語意概念檢索「月」結果：6,734筆

http://cls.hs.yzu.edu.tw/CSP/W_DB/index.htm

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H) 連結 >>

【淺顯低唱】若有知音見採—唐宋詞全文資料庫

回首頁 文字檢索 網路展書

回檢索畫面

共計6734筆

... 11 12 13 14 15 16 17 18 19 20 ...

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文	詞中辭彙	辭彙總類	辭彙大類	辭彙次類	辭彙元素
121	定西番	張先	宋代			小逐畫船風月，	月	物	天文	月	月
122	定西番	張先	宋代	執胡琴者九人		競／嬋娟。	嬋娟	物	天文	月	嬋娟
123	少年遊慢	張先	宋代			梯漢同登蟾窟。	蟾窟	物	天文	月	蟾窟
124	翦牡丹	張先	宋代	舟中間雙琵琶		玉盤大小亂珠迸。	玉盤	物	天文	月	玉盤

以字形比對檢索「月」檢索：7,705筆

回檢索畫面

共計7705筆

... [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) ...

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文
161	謁金門	晏殊	宋代			風月宛然無異。
162	破陣子	晏殊	宋代			人間好月長圓。
163	浣溪沙	晏殊	宋代			月好謾成孤枕夢，
164	浣溪沙	晏殊	宋代			三月和風滿上林。
165	浣溪沙	晏殊	宋代			鬢髮欲迎眉際月，
166	鵲踏枝	晏殊	宋代			明月不諳離恨苦。
167	紅窗聽	晏殊	宋代			風前月底，
168	採桑子	晏殊	宋代			陽和二月芳菲遍，
169	採桑子	晏殊	宋代			淡月朧明。
170	喜遷鶯	晏殊	宋代			堯冀隨月欲團圓。

... [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) ...

語意概念標記應用於全文檢索之驗證： 以唐宋詞網站為例

- 一、以「星」、「月」為例
- 二、比較以「字形比對」及「語意概念」檢索結果

項目	語意概念檢索	字形比對檢索
星	2,076筆	1,427筆
月	6,734筆	7,705筆

詩詞文學的情感分類及情感計算

情感詞彙分類與計量：情感詞彙

大類	中類	小類之分項	數量
綜合情感	正面情感	憐惜愛寵、景仰羨慕、喜悅欣樂、悠然自適、自我肯定	5
	負面情感	擔驚受怕、悲哀傷痛、憂愁掛慮、惱怒怨妒、厭煩憎惡、慚愧自責、辛酸艱苦、感懷惆悵、思鄉念舊、驕矜自恃、懊惱憾惜、自我否定	12
	其他	情感（概稱）、綜合情感（相反合稱）、綜合情感（相近合稱）	3
兩性情感	傾心愛戀		1
	相思怨懟		1

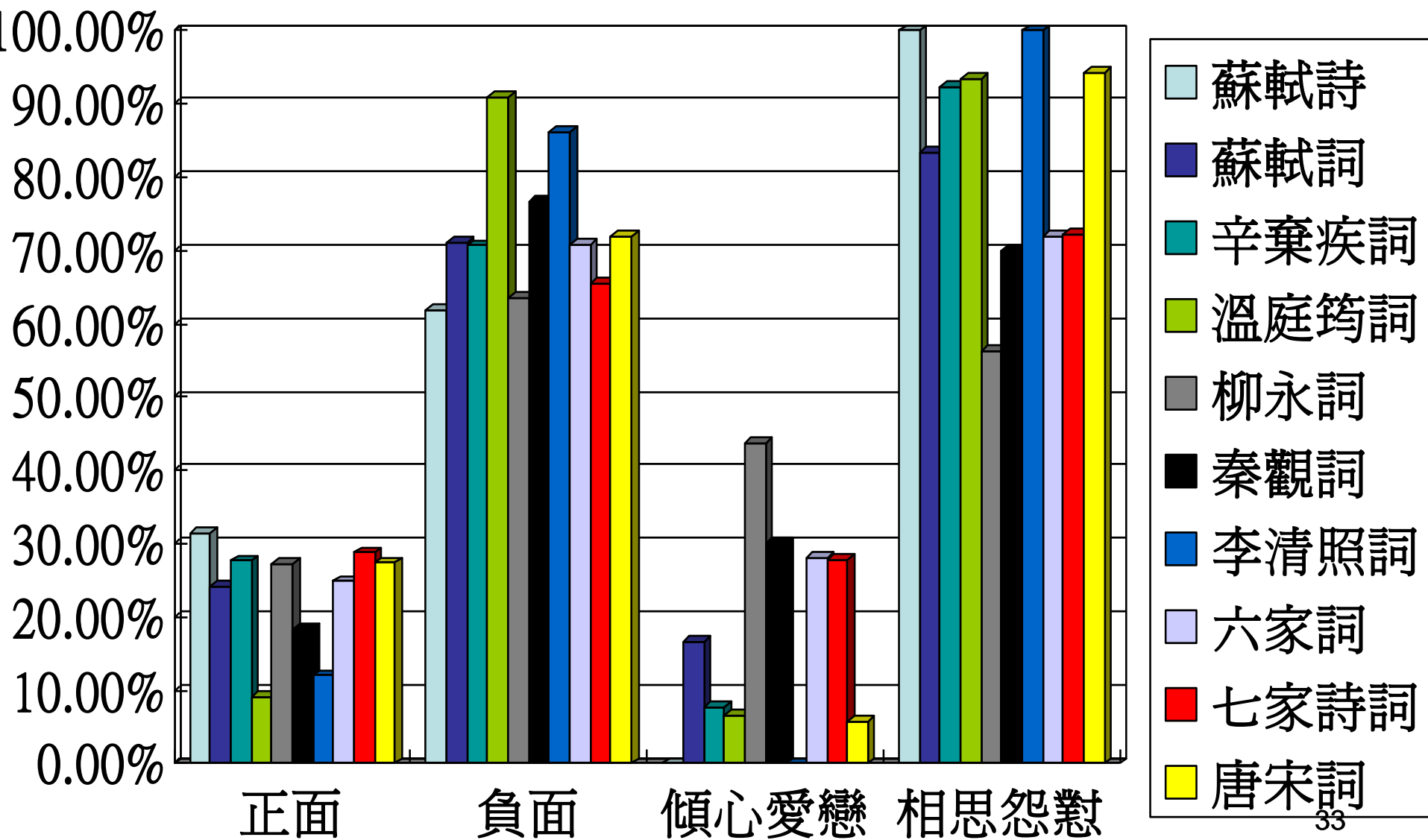
文學裡的情感表達

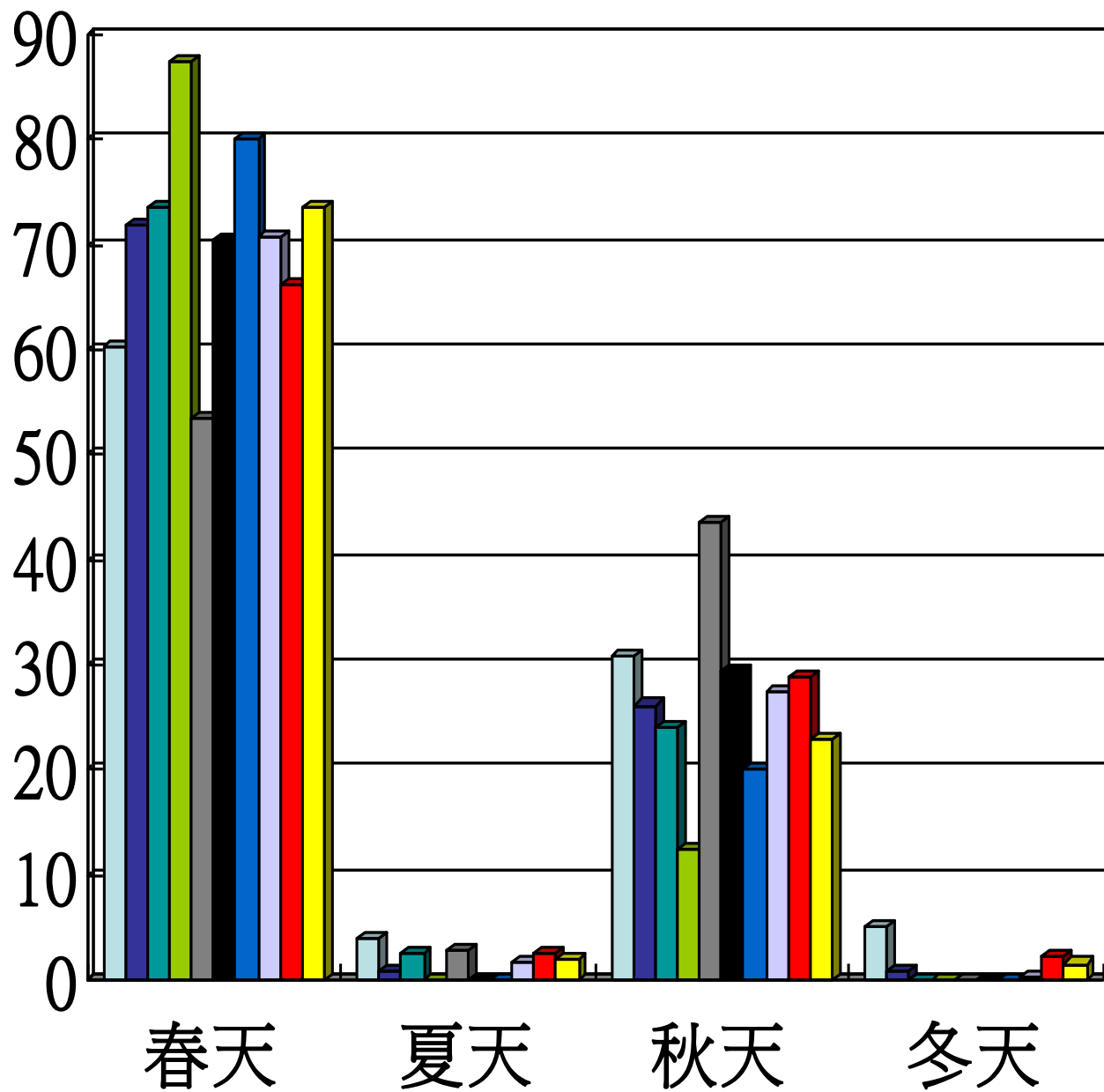
- 一. 文學裡的情感表達方式：以情襯景、以景襯情、情景交融
- 二. 影響情感的外在因素：季節（傷春悲秋）、天候（晴雨冷暖）、時間（晝夜晨昏）
- 三. 用以寄託情感的外在事物：天文、動物、植物、器物、景觀

詩詞文學的情感分類及情感計算 研究範圍

作者/項目	文體別	詞作風格	作品數量	詞彙數量
蘇軾	詩		2,855	95,297
蘇軾	詞	豪放	379	12,418
辛棄疾	詞	豪放	629	26,608
溫庭筠	詞	婉約	74	1,685
柳永	詞	婉約	218	10,670
秦觀	詞	婉約	148	5,815
李清照	詞	婉約	53	1,783
總計			4,356	154,276 ₃₂

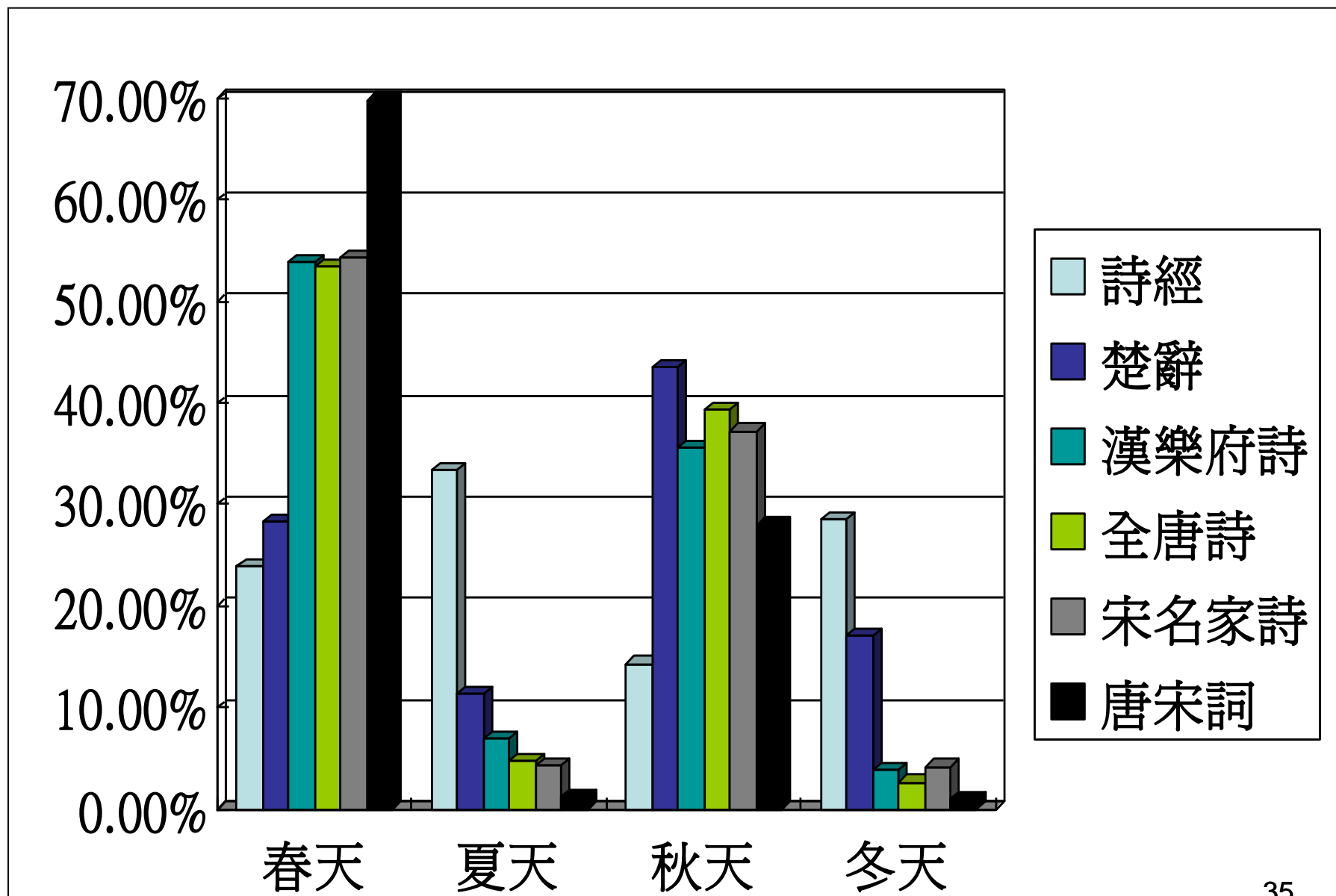
六家七種詩詞的情感詞彙，其情感之不同類別（中類）



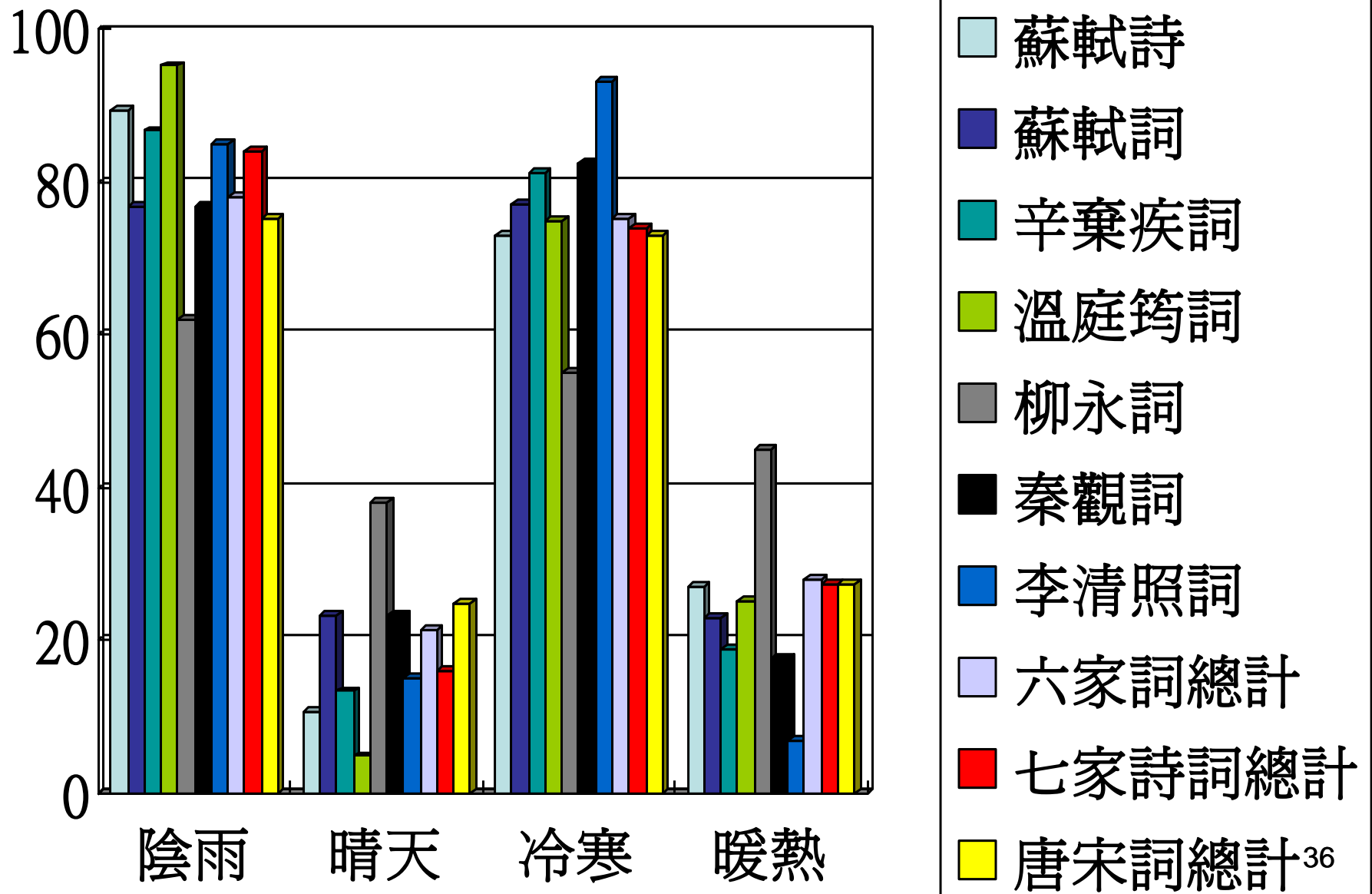


- 蘇軾詩
- 蘇軾詞
- 辛棄疾詞
- 溫庭筠詞
- 柳永詞
- 秦觀詞
- 李清照詞
- 六家詞
- 七家詩詞
- 唐宋詞⁴

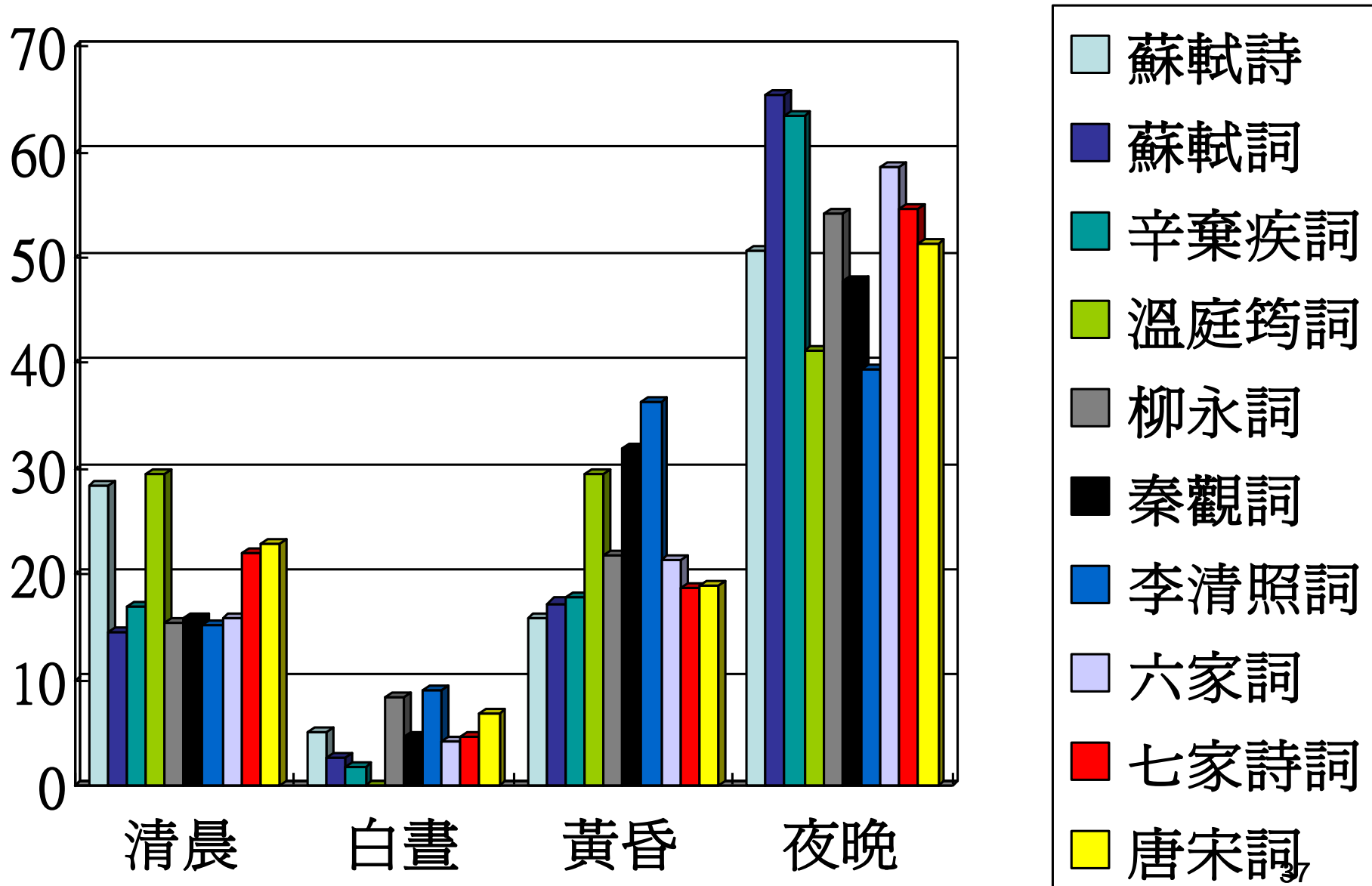
詩經、楚辭、漢樂府詩、全唐詩、宋名家詩、唐宋詞的季節描寫比較



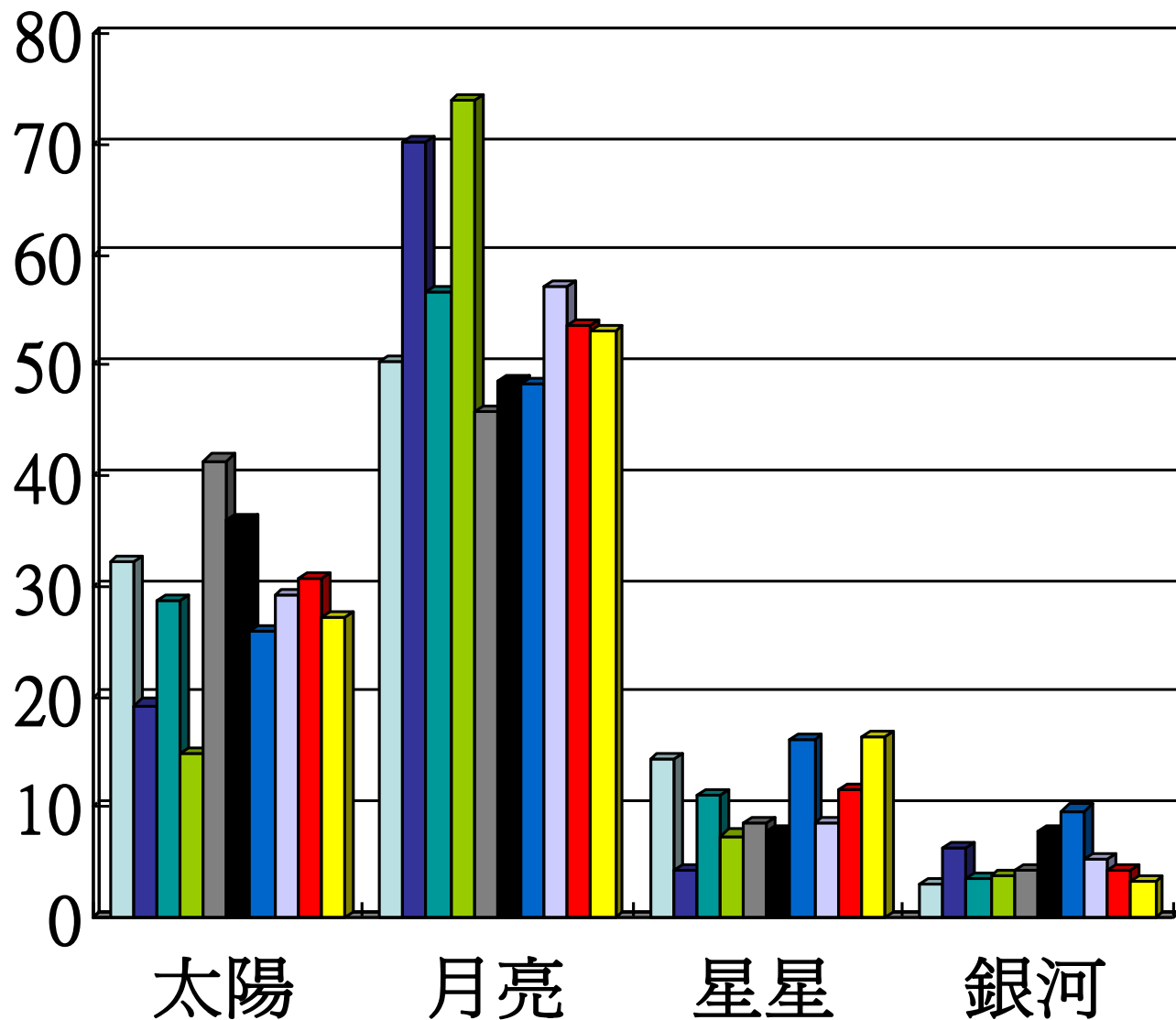
各家詩詞描寫天候氣溫的比例



晝夜時間光線明暗的變化（清晨、白晝、黃昏、夜晚）



日、月、星、銀河等天文景觀比例



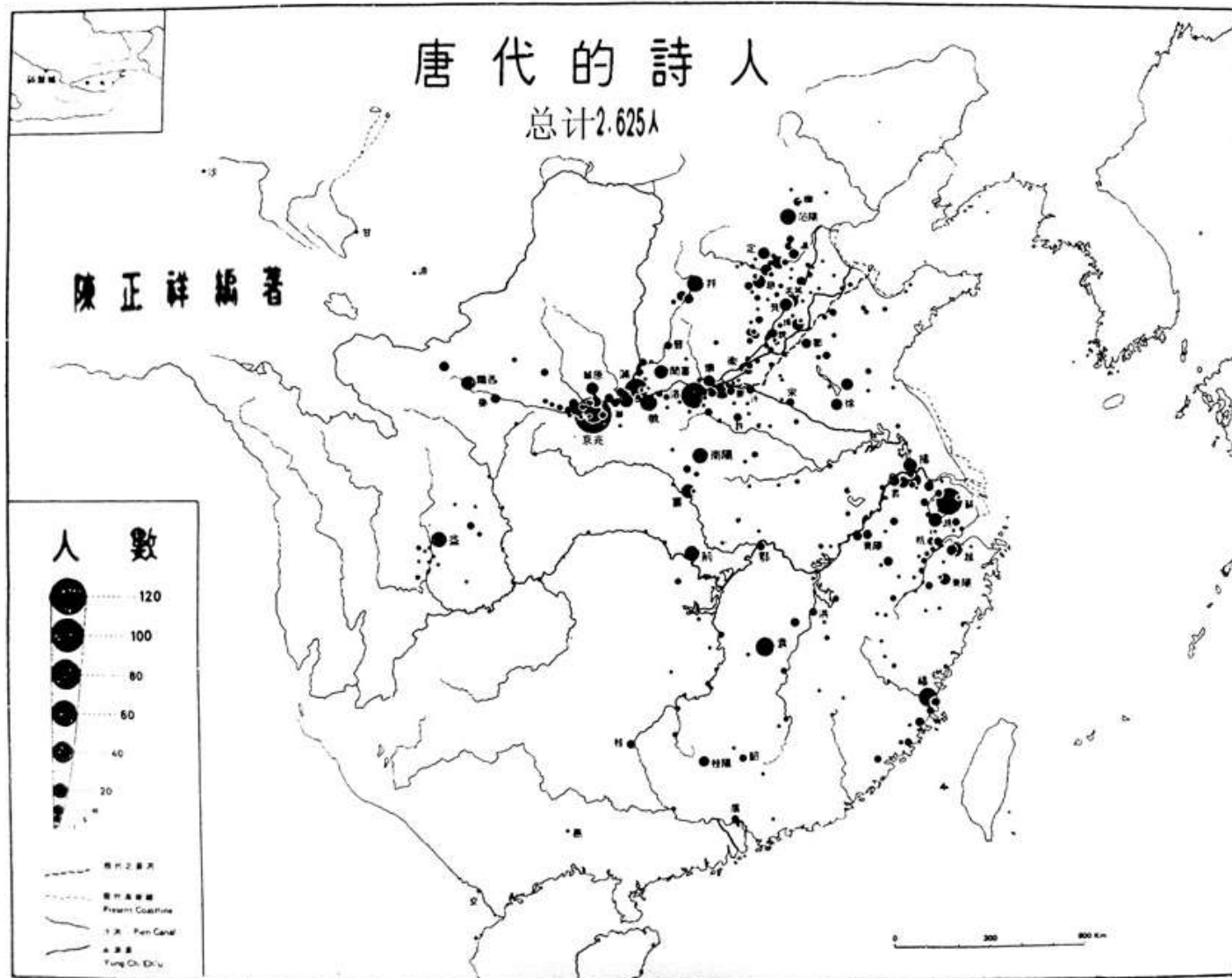
- 蘇軾詩
- 蘇軾詞
- 辛棄疾詞
- 溫庭筠詞
- 柳永詞
- 秦觀詞
- 李清照詞
- 六家詞
- 七家詩詞
- 唐宋詞

應用於文學地理資訊的研究

陳正祥教授的文化地圖：唐代詩人籍貫分佈圖（引自陳正祥《中國文化地理》）

本圖統計唐代詩人：2,625人

唐代詩人共計：2,712人



引自1979年在日本廣島大學及國土地理學院的演講稿
《中國文化中心的遷移》

應用於文學地理資訊的研究

陳正祥教授的文化地圖：北宋詞人籍貫分佈圖（引自陳正祥《中國文化地理》）

本圖統計北宋詞人：347人

唐宋詞人共計：1,440人



應用於文學地理資訊的研究

陳正祥教授的文化地圖：宋代詩人籍貫分佈圖（引自陳正祥《中國文化地理》）

本圖統計宋遼金詩人：？人 宋代詩人共計：9,215人



Copyright, 1977, by the International House for China Studies

文學地理資訊的研究與應用

陳正祥教授研究不足之處：

- 據以統計的資料不完整。
- 以人工統計分析，難以駕馭大量的資料。

文學地理資訊的研究與應用

- 2001年南京大學胡阿祥教授以魏晉時代文學家和文學作品的多寡為主要指標探討文學發展的地區差異。胡文的統計基礎是作家的籍貫地。
- 胡阿祥教授研究不足之處：只以籍貫統計，客觀性與正確性均不足。
- 參見胡阿祥著，《魏晉本土文學地理研究》，南京大學出版社，2001年5月第一版

文學地理資訊的研究與應用

陳教授以文人分佈建立文化變遷研究，胡教授建立了文史計量研究的方法，其貢獻及前瞻性均不容置疑。

但由於所據以統計的文獻不完整，勢必影響結論的正確性，再者，以唐宋代詩人籍貫分佈作為文化中心分佈變遷的依據，失之粗率，主要的原因有三：

- 一. 詩人的籍貫地未必是出生地；
- 二. 詩人的籍貫或出生地未必是其文學的寫作地，例如辛棄疾與李清照都是山東濟南人，但寫詞時都在南方；
- 三. 每位詩人作品數量差異懸殊，以作者作為統計單位，而未考慮作品數量，其客觀性與正確性均不足。

文學地理資訊的研究與應用

- 以全宋詩為例：全宋詩9, 215位詩人，籍貫可考的詩人5, 953人，作品總數量254, 240首（最少1首～最多9, 271首）。

作品數	作者人數	佔總作者人數百分比
1首	4, 975人	53. 99%
2首	1, 366人	14. 82%
3-20首	2, 083人	22. 60%
21-500首	673人	07. 30%
501首以上	118人	01. 28%

作品數量、詩人籍貫、作品總數比較

作品數量最多前十人	詩人籍貫最多前十處	作品總數最多前十處
越州山陰 (陸游/9271首)	莆田 (119人/6161首/3名)	越州山陰 (9626首/5人)
福建莆田 (劉克莊/4557首)	永嘉 (110人/3175首/11名)	眉州眉山 (6316首/52人)
吉州吉水 (楊萬里/4284首)	鄞縣 (104人/3204首/10名)	莆田 (6161首/121人)
鄭州 (趙蕃/3735首)	錢塘 (81人/4024首/7名)	吉州吉水 (5289首/4人)
安徽宣城 (梅堯臣/2933首)	晉江 (59人/141首/271名)	宣城 (5013首/22人)
安徽歙縣 (方回/2859首)	崇安 (59人/1021首/56名)	開封 (4390首/58人)
四川眉山 (蘇軾/2824首)	建安 (58人/1628首/32名)	錢塘 (4024首/81人)
河南開封 (韓滉/2624首)	開封 (58人/4390首/6名)	鄭州 (3740首/3人)
亳州譙縣 (張耒/2268首)	平陽 (50人/95首/319名)	歙州歙縣 (3480首/32人)
江西分寧 (黃庭堅/2204首)	邵武 (49人/2086首/22名)	鄞縣 (3204首/104人) 46

文學地理資訊的研究與應用

- 一. 詩作最多的前十名只有劉克莊的籍貫在詩人分佈最多的前十處。
- 二. 可見以詩人籍貫分佈作為文學分佈或文化變遷的立論依據，其正確性與客觀性均不足。
- 三. 文獻資料的數位化及資訊科技工具可以幫助學者作更精確的量化統計分析研究。


文學地理資訊的研究與應用

http://cls.hs.yzu.edu.tw/sung/sung/SMP_GIS.html

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

【宋人與宋詩】地理資訊系統

使用說明



圖層

- 宋人傳記之籍貫 (北宋)
- 宋人傳記之籍貫 (南宋)
- 宋人傳記出生地 (北宋)
- 宋人傳記出生地 (南宋)
- 北宋地名(1111年)
- 南宋地名(1208年)
- 北宋州界(1111年)
- 南宋州界(1208年)
- 蘇軾黃州古今行吟圖
- 北宋路界(1111年)
- 南宋路界(1208年)
- 世界地圖

重畫地圖

查詢作者：

查詢籍貫：
-----籍貫-----

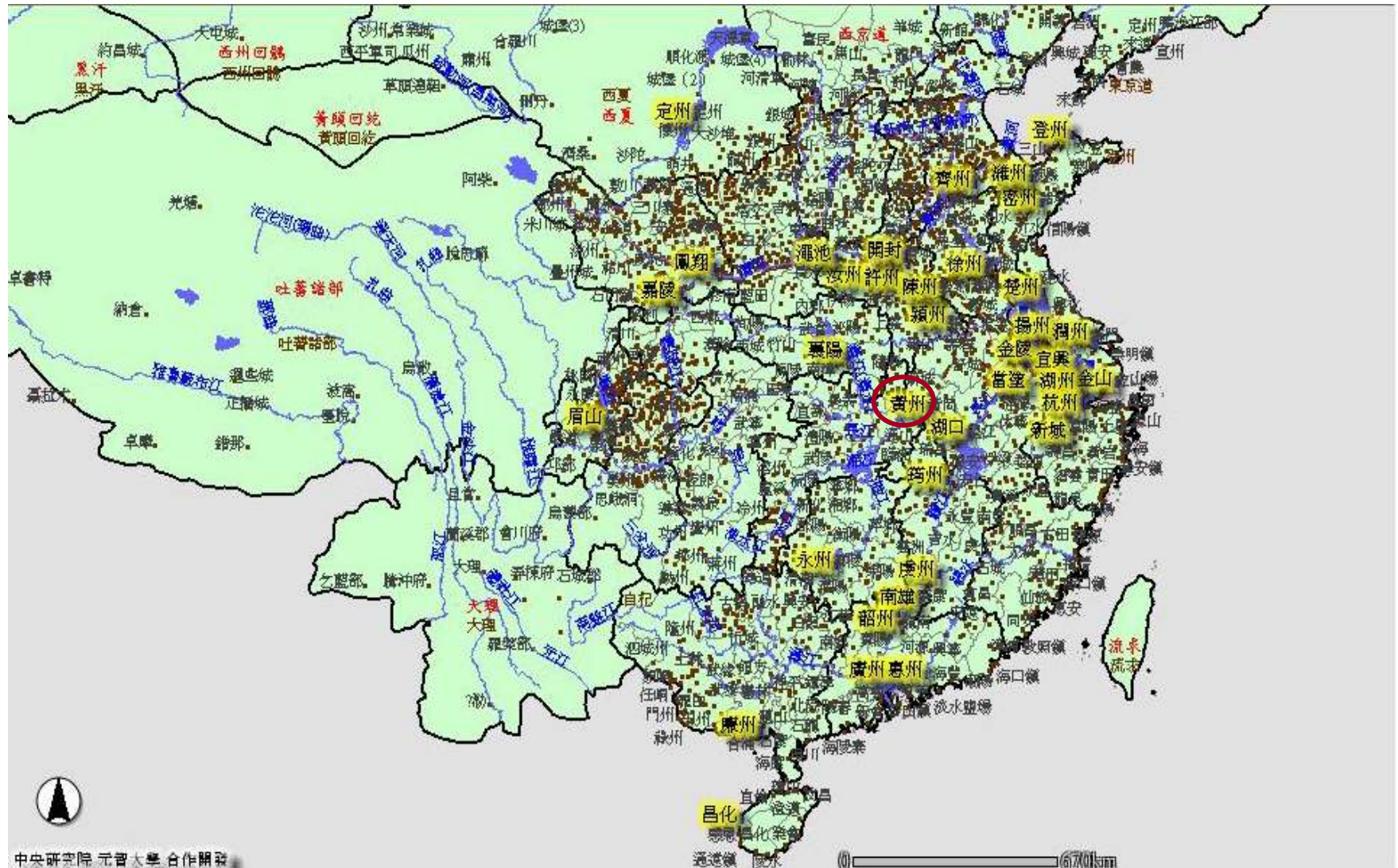
查詢出生地：
-----出生地-----

全文檢索：

完成 網際網路 100%

蘇軾仕宦遊歷寫作地圖

依據蘇軾生平及詩詞文作品繫年繪製蘇軾仕宦遊歷寫作地圖



宋人與宋詩文學地理資訊系統

蘇仕宦遊行跡地圖（從眉山到開封）

蘇軾文學地圖 | 蘇軾文學時空繫年 | 宋代文學家宦遊行跡圖

從眉山到開封

行經：眉山（眉州，成都府路）→嘉州（嘉州，成都府路）→犍為（嘉州，成都府路）→宜賓（戎州，梓州路）→牛口（宜賓至戎州之間）→戎州（梓州路）→江安縣（潼州，梓州路）→合江縣（潼州，梓州路）→巴縣（恭州，夔州路）→涪州（涪州，夔州路）→豐都（忠州，夔州路）→峽江→忠州（忠州，夔州路）→武寧縣（萬州，夔州路）→夔州（夔州，夔州路）→奉節縣白帝城（夔州，夔州路）→巫峽→巫山→巴東縣（歸州，荊湖北路）→興山縣（歸

筆數 | 詩題

1	郭綸
2	初發嘉州
3	犍為王氏書樓
4	過宜賓見夷中亂山
5	夜泊牛口
6	牛口見月
7	戎州
8	舟中聽大人彈琴
9	泊南井口期任遵聖長官，到曉不及見，復來
10	過安樂山，閬山上木葉有文，如道士篆符，二
11	過安樂山，閬山上木葉有文，如道士篆符，二
12	渝州寄王道矩
13	江上看山
14	涪州得山胡次子由龍
15	留題仙都觀
16	仙都山鹿
17	江上值雪，效歐陽體，限不以鹽玉鶴鷺絮絳等字，次子由龍
18	屈原塔
19	望夫臺
20	竹枝歌，並引
21	過木樨觀
22	八陣碓

【蘇軾宦遊行跡_所有時期】為現在啟動的圖層

座標: X:385470.87, Y:3172497.64

繼續深化的工作

宋代黃州地圖（譚其驤《中國歷史地圖集》）

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

上一頁 搜尋 我的最愛

網址(D) <http://gis210.sinica.edu.tw/bsgis/default.htm> 移至 連結

蘇軾文學地圖 蘇軾文學時空繫年 宋人地理資訊系統 操作說明 網路展書讀

放大 縮小 平移 連結詩詞 地名檢索 地圖全覽 縮小

【北宋_州_縣】為現在啟動的圖層

國際網路

繼續深化的工作

〈北宋黃州地圖〉（中研院史語所范毅軍教授研究團隊製圖）
在譚其驤《中國歷史地圖集》嵌入太空航照圖



繼續深化的工作6 / 15

黃州衛星地圖 (中研院史語所范毅軍教授研究團隊製圖，2007年5月)



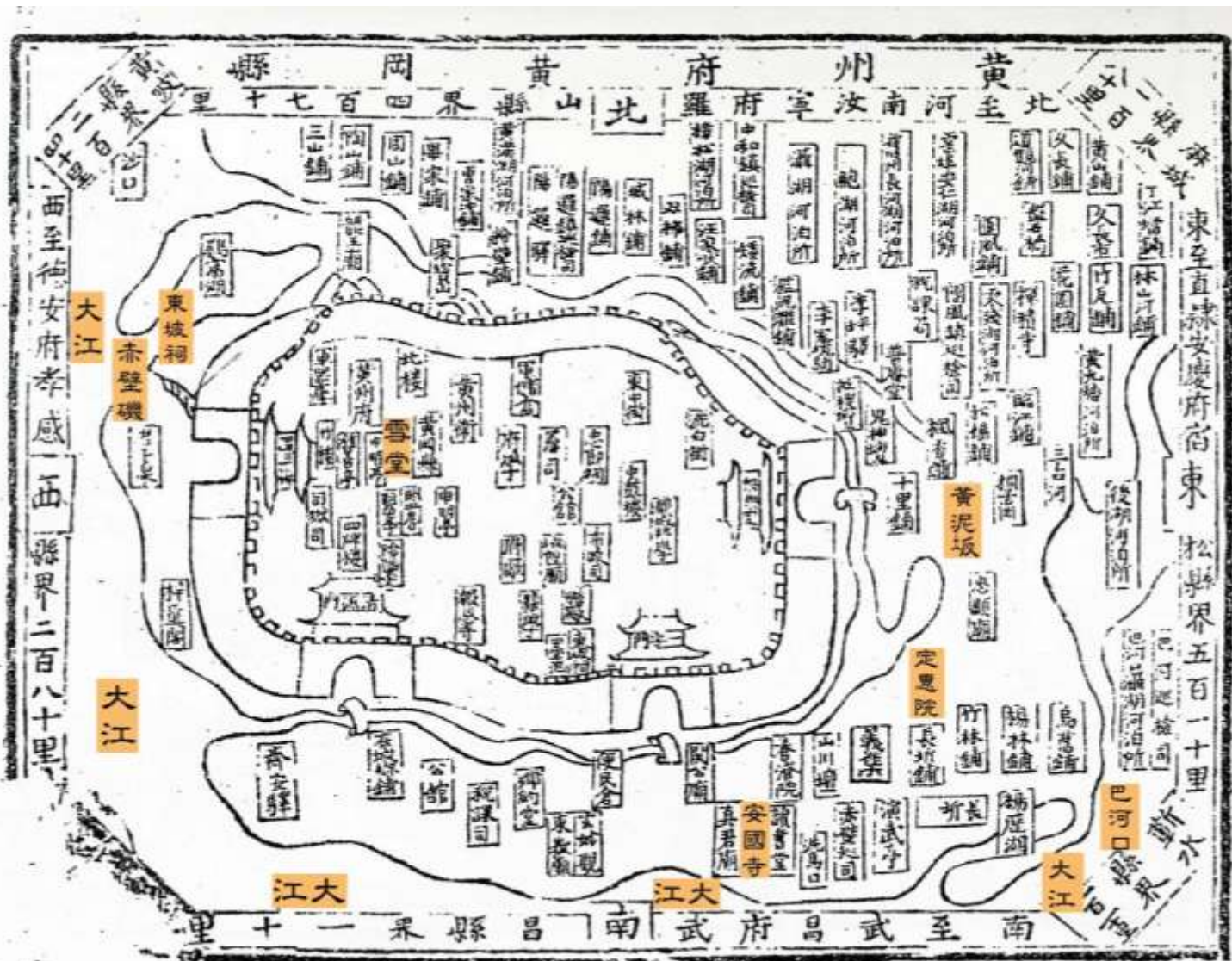
繼續深化的工作

google 地圖，黃州 (2007年10月) 7 / 15



繼續深化的工作

〈弘治黃州府志黃岡縣圖〉（蘇軾行跡） 8 / 15



文學地理資訊的研究與應用

台灣民間文學地理資訊網站 (台中縣石岡鄉歌謠)

The screenshot shows a web browser displaying the GIS of Taiwan Folk Literature website. The URL is <http://cls.hs.yzu.edu.tw/TFL/GIS/index.aspx>. The page title is "台灣民間文學地理資訊系統 GIS of Taiwan Folk Literature". The map shows Shigang Township with various streets and landmarks. A red circle highlights "梅子巷" (Meizi Lane), with a speech bubble saying "我的家在這裡" (My home is here). The right sidebar shows search results for "童謠" (Folk Songs), listing 126 items. The bottom status bar shows "完成" (Done) and "網際網路" (Internet).

台灣民間文學地理資訊系統 GIS of Taiwan Folk Literature

地圖 衛星 混合地圖 地形 日治時期臺灣堡圖 古今對照

地圖功能表

查詢

查詢結果統計

清除查詢結果

童謠: 126篇
全部共126篇

查詢結果

- 童謠
- 石岡鄉-ABCD a1
- 石岡鄉-ha2 kat4 a2 sanq3 a2
- 石岡鄉-khi2 小沙袋遊
- 石岡鄉-一一一 a1
- 石岡鄉-一一一 a2
- 石岡鄉-一一一 a3
- 石岡鄉-一二三 a1
- 石岡鄉-一二三 a2
- 石岡鄉-一群鳥仔飛過?
- 石岡鄉-一陣鳥仔叫 be be1 a1
- 石岡鄉-一隻鳥仔叫咧
- 石岡鄉-三妹姑

石岡鄉 點仔謠

現今行政區：台中縣 石岡鄉 金星村
日據行政區：臺中廳 東上堡b 石岡仔庄
採錄地點：石岡鄉金星村
文體：童謠

全文

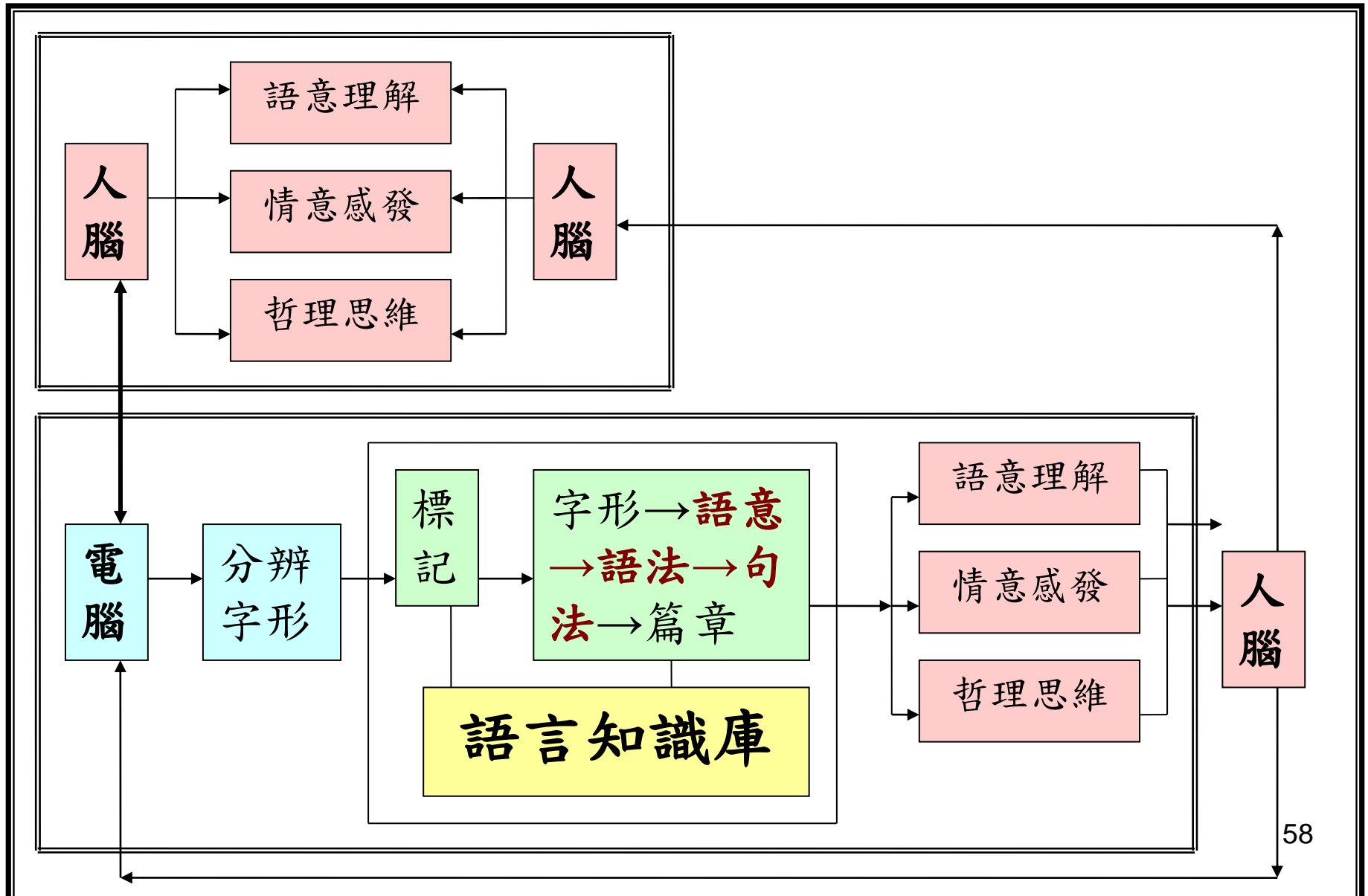
1000 英尺

完成 網際網路 100%

應用於文學計量研究：唐宋詞領字研究

項目	以部分宋詞為範圍	以全部唐宋詞為範圍
領字的起源	柳永	柳永之前已有27位詞人在52種詞牌使用過領字
領字的聲調	去聲、仄聲	<p>平聲490字種3,951字 + 仄聲892字種23,113字（上聲261字+去聲497字+入聲134字）=1382字（平聲／仄聲=35.46%／64.54%）</p> <p>四個聲調使用次數最高的前120字，平聲23字、上聲21字、去聲71字、入聲5字。</p>
領字的詞類	副詞、動詞、連接詞、形容詞、介詞	唐宋詞有以名詞（柳15燕8鶯6馬2雁9酒13）及代名詞（我35你18他2翁6公14君92）當作領字者（唐宋詞使用領字總字數27,144字）
領字所使用的詞體	一般用在長調，小令較少，引、近也少用	<p>長調755（長調533種+中調222種）+小令188種=943種（755/188=4/1）</p> <p>唐宋詞引體36種，含有領字的有27種（75%）；近體20種，含有領字的有14種（70%）；慢體共有86種，含有領字的有80種（93%）。</p>

未來的發展：建立語言知識庫



參考資料

一、參考網站

網路展書讀：<http://cls.hs.yzu.edu.tw/>

二、參考論文

1. 羅鳳珠，〈引信息的「術」入文學的「心」——談情感計算和語義研究在文史領域的應用〉，《文學遺產》（中國社會科學院文學研究所主辦，《文學遺產》編輯部編輯出版），2009年第1期，2009年1月15日，第138-141頁。
2. 羅鳳珠、曹偉政，〈唐宋詞單字領字研究〉，《語言暨語言學》，第九卷第二期，中央研究院語言學研究所，2008年4月出版，第189~220頁。
3. 羅鳳珠，〈植基於中國詩詞語言特性所建構之語意概念分類體系研究〉，《2008年海峽兩岸圖書資訊學學術研討會論文集》，中華圖書資訊學教育學會出版，2008年，第385~407頁。
4. 羅鳳珠、范毅軍、鄭錦全〈宋人與宋詩地理資訊系統之設計與應用〉，第五屆數位地球國際研討會，文化大學主辦，2007年5月15-18日。
5. 羅鳳珠，〈探一探文史資料自動化的路〉，《國文天地》第3卷第3期，1987年8月，頁33-43。

謹以本報告紀念
張仲陶教授及先師周何教授、
王熙元教授

謝謝聆聽，敬請指教