




道教開光儀式疏文之文本探勘與 數位人文探索：以府城延陵道壇為例

林敬智* 

【摘要】

本文利用文本探勘與 GIS 地理資訊技術，對臺南府城延陵道壇吳政憲道長的神像開光儀式疏文進行數位人文探索。文中將以 Python 擷取開光疏文中所蘊含之神明、宮廟、信徒資訊，接著以 GIS 時空分析與社會網絡分析對於延陵道壇儀式服務進行初步探索，以視覺化方式探究開光儀式的地理空間分佈；亦展示開光地點與神像間的網絡關係，藉此觀察開光神明與宮廟間之關聯性。本文旨在展示如何利用數位人文技術與工具協助人文研究，善用資訊技術可快速而大量處理具有模式化的文獻（如道士疏文），擷取其中關鍵性資料內容，再利用 GIS 與網絡分析探索開光疏文之空間意涵與網絡關係，發揮數位人文擅長之「鳥瞰」（distant reading），提出新問題意識，以作為下一步「近讀」（close reading）之基礎。

* 國立政治大學宗教研究所助理教授

ORCID 0000-0002-6791-6561

E-mail: cclin52@gmail.com

作者感謝詹進發教授、詹琦研究助理在 Python 技術上之指導與協助。

本研究受科技部專題計畫「全球脈絡下的臺灣宗教：國族認同、宗教流動、文化跨界比較研究」子計畫「臺灣財神信仰的在地化與全球化：以北港武德宮的分靈與臉書網絡為例」，計畫編號：MOST 106-2420-H-004 -005 -MY2，在研究計畫期間進行田野調查，獲得本研究所分析之資料。筆者亦參與教育部深耕計畫國立政治大學「華人文化主體性研究中心」，獲得中心補助研究助理費用整理田野調查所獲得之資料，進行數位人文分析。

關鍵詞

數位人文 文本探勘 資訊視覺化 GIS 地理資訊系統 社會網絡分析

壹、緒論

臺南府城延陵道壇吳政憲道長，法名羅錠，傳承家族深厚的道脈傳統，又能利用現代科技的社群媒體（如臉書、Line 群組）與各界互動，素以神像開光儀式聞名於世。自 2007 年於開基玉皇宮閱籙登梯奏職正式陞職為高功道長以來（洪瑩發、林長正，2013），歷經十多年的儀式服務已將其知名度與活動範圍遍及全臺各地，甚至亦遠赴海外為信眾與宮廟舉行齋醮與神像開光儀式，2017 年更舉辦開光神尊大會師慶祝登梯奏職十週年。其所開光之神尊數目與分佈，近年逐漸加速擴大與擴散。作者獲得吳政憲道長之信任，取得甲午年（2014 年）至戊戌年（2018 年）之開光疏文，利用數位人文之資料擷取技術與工具，對於開光疏文中之神明、信徒與宮廟之間的資訊進行數位人文探索，包括 GIS 地理空間分析與網絡分析，應用數位人文擅長處理數量龐大且快速的資訊技術（王汎森，2014），將具有特定模式化的道教開光疏文，擷取出相關資訊，進而利用 GIS 相關平臺工具與技術，將開光信徒之地理資訊於數位地圖上標示定位，同時所擷取出之神明、信徒與宮廟等資訊，建置成表格後，亦能進行網絡分析，探索開光地點與神像之間的網絡關係，藉此觀察開光神明與宮廟之間的關聯性。

道士開光疏文具有高度模式化的傾向，正好適合以資訊技術擷取其中之關鍵內容，包括開光神像、分靈祖廟或子廟、信徒等方面的資訊，藉由數位人文資料擷取技術，可以將相關資訊提取出來製成表格，進一步藉由 GIS 相關技術與工具，轉化成為地圖視覺化呈現開光之地理空間分佈，同時神像與宮廟之間的網絡關係亦能夠利用社會網絡分析之相關技術與平臺工具，予以視覺化呈現。既有利用數位人文技術分析道教之研究，多以道教經典文本作為文本分析或文本探勘之對象（Stanley-Baker, 2019; Stanley-Baker & Chong, 2019; Stanley-Baker & Ho, 2015; Sundararajan, Ting, Hsieh, & Kim, 2020; Yang, Jin, Wang, & Wang, 2015），並未處理過具有模式化的儀式文本，而外文的文本探勘技術則多半不適用於中文格式內容。由

於開光疏文有基本之套路，道士撰寫時填入幾項基本的欄位，相對於一般文獻更為單純，可以利用 Python 程式語言來擷取其中的內容，尚不需要應用到「同位詞夾子」萃取演算法（謝育平，2011）、「半自動詞彙擷取」（杜協昌，2016；彭維謙、程卉、陳詩沛，2018），或對龐大文獻進行之資料擷取技術（張素玢、李鈺淳，2014）。本文聚焦於如何利用數位人文技術與工具來進行宗教文本之文本探勘，以及從擷取之地理資訊進一步繪製 GIS 地圖，展示空間視覺化，同時也將擷取之神像與宮廟間的網絡關係藉由社會網絡分析技術進行初步的探索。至於更進一步的宗教詮釋，則須另文處理，不在本文討論範圍之中。

貳、研究程序

一、研究程序說明

本文將以五年上千筆之開光疏文作為主要文本探勘之對象，利用 Python 程式語言擷取疏文中之地理資訊（地址）、開光神明、神明祖廟、新設宮廟或神壇、信徒姓名、開光日期等。其次以內政部國家地理資訊圖資雲服務平臺（簡稱 TGOS）中之「全國門牌地址定位服務」進行「批次門牌地址比對服務」，一次查詢大筆地址之經緯度，最後利用幾種不同之 GIS 平臺繪製開光儀式之空間分佈地圖，以及將祖廟與神像之網絡關係進行初步之探索。



圖 1 研究程序

在府城延陵道壇吳政憲道長的開光疏文中，大致可以歸納為下圖之模式：

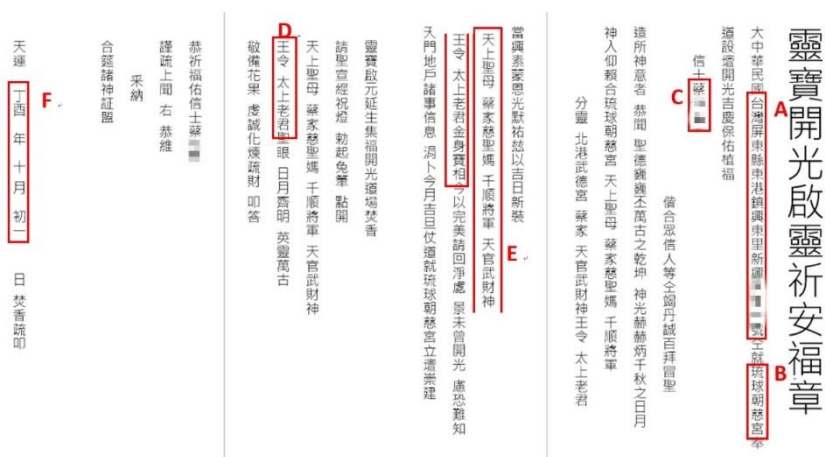


圖 2 延陵道壇吳政憲道長開光疏文之 Word 文件範例

從疏文中可以分別提取出：A 信徒地址；B 分靈祖廟；C 開光信徒姓名；D 開光之神明；E 新宮壇與神明尊稱；F 開光儀式之年月日日期等六種資訊項目。

本研究所蒐集之道士開光疏文的資料共有五年近 1,300 篇，若要透過人工逐一判讀後從中取出需要的文字建立資料庫，勢必須要花費相當多的人力與時間，且人工判讀輸入資料難以避免因為操作者疲憊造成的人為錯誤。若藉由人工處理須以 Word 檢視內容後擷取所需的文字，將其複製到以 Microsoft Excel 應用程式 (Excel) 建立的資料表中，由於操作步驟繁瑣，需耗費許多時間，而且重複多次後極易產生錯誤。因此本研究利用文本探勘之技術，以 Python 語言撰寫程式從開光疏文檔案中自動擷取資料，再進行人工資料校對，能快速處理大量之文獻，示範利用資訊技術輔助人文研究，大幅節省研究時間成本。

其次從上述擷取獲得之地址資訊，筆者將進行 GIS 定位工作，在地圖上標記開光信徒之地址。由於利用人工一筆一筆在 GIS 地圖上查找其經緯度十分曠日廢時，筆者將利用內政部開發的 TGOS 之「全國門牌地址定位服務」，該平臺不僅可以查找單筆地址之經緯度，亦能在申請核可後，進行「批次門牌地址比對服務」，一次查詢大筆地址之經緯度。



圖 3 TGOS 之全國門牌地址定位服務



圖 4 TGOS 之批次門牌地址比對服務

二、資料擷取與文本探勘

本研究利用文本探勘的技術，以 Python 語言撰寫程式由開光疏文檔案自動擷取資料，再以人工進行資料的檢核校對，能大幅減少所花費的時間成本，並確保資料之品質。資料處理與分析之流程說明如後：

(一) 步驟一：原始資料檔案格式轉換

原始資料是 Word 格式不好處理，因此先透過程式工具將 Word 的“.doc”檔轉成純文字的“.txt”檔，透過以下的 Python 程式可將 Word 檔案的文字內容擷取出來，儲存為純文字的“.txt”檔以便於導入其他程式做後續的處理。

```
#此程式用於列印 Word 檔的內容
import win32com.client # 載入所需程式模組

word = win32com.client.Dispatch("Word.Application")
word.visible = False
in_file = '開光疏文丁酉十月一蔡○○.doc'

doc = word.Documents.Open(in_file) # 開啟 Word 檔案
counter=0
for paragraph in doc.Paragraphs: # 每一個段落逐一處理
    counter += 1
    Text = paragraph.Range.Text # 取得段落的内容
    Print ("%02d: %s"% (counter,Text)) # 列印出文字

doc.Close() # 關閉檔案
```

上述程式執行的結果如圖 5 所示，可供檢視 Word 文件之內容。

```
01:
02: 靈寶開光啟靈祈安福章
03: 大中華民國台灣屏東縣東港鎮興東里新廟 駝全就琉球朝慈宮奉
04: 道設壇開光吉慶保佑植福
05: 信士蔡
06: 偕合眾信人等全竭丹誠百拜冒聖
07: 造所神意者 恭聞 聖德巍巍丕萬古之乾坤 神光赫赫炳千秋之日月
08: 神入仰賴合琉球朝慈宮 天上聖母 蔡家慈聖媽 千順將軍
09: 分靈 北港武德宮 蔡家 天官武財神王令 太上老君
10: 當興素蒙恩光默祐茲以吉日新裝
11: 天上聖母 蔡家慈聖媽 千順將軍 天官武財神
12: 王令 太上老君金身寶相今以完美請回淨處 景未曾開光 慮恐難知
13: 天門地戶諸事信息 涓卜今月吉旦仗道就琉球朝慈宮立壇崇建
14:
15: 靈寶啟元延生集福開光道場焚香
16: 請聖宣經祝燈 勅起免筆 點開
17: 天上聖母 蔡家慈聖媽 千順將軍 天官武財神
18: 王令 太上老君聖眼 日月齊明 英靈萬古
19: 敬備花果 虔誠化煉贖財 叩答
20: 恭祈福佑信士蔡
21: 謹疏上聞 右 恭維
22: 采納
23: 合筵諸神証盟
24:
25:
26: 天運 丁酉 年 十 月 初 一 日 焚香疏叩
```

圖 5 以 Python 程式列印 Word 文件內容的結果

以下的 Python 程式則將 Word 的“.doc”檔案直接轉換為純文字的“.txt”檔案格式。

```
#此程式用於將 Word 檔轉換成純文字的 txt 檔
import win32com.client # 載入所需程式模組

word = win32com.client.Dispatch ("Word.Application")
word.visible = False
in_file = '開光疏文丁酉十月一蔡○○.doc'

doc = word.Documents.Open (in_file)
doc.SaveAs ('開光疏文丁酉十月一蔡○○.txt', FileFormat=4)
doc.Close() # 關閉檔案
```

(二) 步驟二：開光疏文內容分析

如圖 2 所示，本研究欲從數以千計的道士開光疏文檔案自動擷取出六項資料據以建立資料庫。一般在搜尋資料庫的時候會利用關鍵字查詢，例如在查詢圖書館的藏書時，通常會輸入書名、作者姓名或其他可供蒐尋的特定關鍵字詞，便可迅速獲得所需的資料。然而，在開光疏文中每一個資料項目可以有很多種寫法，無法利用關鍵字詞搜尋，須建立一套有效的規則方能擷取出完整而正確的資料。

為了搜尋開光疏文中的資料，本研究使用正規表示式（**regular expression, regex**）建立文本搜索的規則，配合以 Python 撰寫的程式，自動從每一篇開光疏文中擷取出六個資料項目，據以建立資料庫。正規表示式最早於 1950 年代由美國數學家 **Stephen Cole Kleene** 提出，現已廣泛應用於計算機科學。此方法係以文字符號所組成的字串定義搜尋的模式，通常字串搜尋演算法便利用此搜尋模式查找（**find**）或查找並替換（**find and replace**）文本內容，1980 年代以來 **Unix** 作業系統的文本處理工具（**text-processing utilities**）便使用正規表示式，目前此方法已經廣泛應用在很多工具中，例如：搜尋引擎（**search engine**）、文字處理器（**word processor**）、文本編輯器（**text editor**）、詞法分析（**lexical analysis**）（註 1）（註 2）。

以日期為例，傳統是以天干和地支的組合表示年號，十天干為：甲、乙、丙、丁、戊、己、庚、辛、壬、癸，十二地支為：子、丑、寅、卯、辰、巳、午、未、申、酉、戌、亥，構成以「甲子」為首的六十干支循環。年號再加上月、日，以及輸入日期時的空白符號，完整的日期就有無數種表示法，惟日期的表示法仍有規則可循，即以「干支年號」做為開始，以「日」做為結束。欲從文本中找出日期資料，利用正規表示式定義的搜尋模式為：

- “\s〔甲乙丙丁戊己庚辛壬癸〕〔子丑寅卯辰巳午未申酉戌亥〕.+日”。
- “\s”：代表空白字元，包括：空格、跳行、換頁等符號。
- “〔甲乙丙丁戊己庚辛壬癸〕”：代表符合天干的任何一個字元。
- “〔子丑寅卯辰巳午未申酉戌亥〕”：代表符合地支的任何一個字元。
- “.”：代表除跳行符號（“\r”，“\n”）以外的任一個字元。

- “+”：代表符合前面的表示式一次或多次。
- “日”：代表“日”這個字元。

再以搜尋地址這個項目為例，“臺灣屏東縣東港鎮興東里新興○街○號”可以看出是由「臺」與「號」包在一起的，因此匹配條件就設為“臺.+號”，地址為由“臺”開始到“號”結束的字串，但是只靠單一的搜尋模式無法找出全部的地址資料，因為有的地址形式是“臺南市三元慈聖宮全就北極殿奉”或“…淨”，衍生出了更多的地址匹配條件，必須增加“.+淨”和“.+全”表示式。因此，要定義完整的正規表示式必須研究所欲擷取之資料項目的規律性，經過多次測試與修正才能完善搜尋的資料。

透過一系列的正規表示式，可以過濾文本中的內容，擷取出所需的資料項目，以〈開光疏文丁酉十月一蔡○○〉這個檔案做示例。讀取這個檔案之後，①搜尋到「新裝」擷取下一行的字串：“天上聖母 蔡家慈聖媽 千順將軍 天宮武財神 王令 太上老君金身寶相今以”，去掉「今以」、跳行符號、空格，可得到**(E)新宮廟**。②找到[甲乙丙丁戊己庚辛壬癸][子丑寅卯辰巳午未申酉戌亥]的組合並擷取到「日」結束：“丁酉 年 十 月 初 一 日”，去掉空格，找出以年、月、日分隔的資料，便可獲得**(F)年月日**。③擷取「臺」到「號」的字串：“臺灣屏東縣東港鎮興東里新興○街○號”，可得到**(A)地址**。④擷取「全」到「奉」的字串：“全就琉球朝慈宮奉”，去掉「全」、「就」、「奉」、空格，可得到**(B)祖廟**。⑤擷取「信士」開始的字串：“信士蔡○○”，去掉「信士」、空格，可得到**(C)請神人**。⑥擷取「日月齊明」以前的字串：“王令 太上老君聖眼 日月齊明”，去掉「日月齊明」、「聖眼」、空格，可得到**(D)神明**。

表 1

擷取項目輸入與輸出結果

項目	匹配條件	輸入資料	處理	輸出結果
A 地址	1.擷取「臺」到「號」 2.擷取到「淨」 3.擷取到「全」	臺灣屏東縣東港鎮興東里新興○街○號	1.無進一步處理 2.去掉「淨」、空格 3.去掉「全」、空格	臺灣屏東縣東港鎮興東里新興○街○號

(續下表)

(接上表)

項目	匹配條件	輸入資料	處理	輸出結果
B 祖廟	1.擷取「仝」到「奉」 2.擷取到「奉」	仝就琉球朝慈宮 奉	1. 去掉「仝」、「就」、「奉」、空格 2.無進一步處理	琉球朝慈宮
C 請神人	1.擷取「信士」開始的字串 2.擷取到「奉祀人」開始的字串	信士蔡○○	1.去掉「信士」、空格 2.去掉「奉祀人」、空格	蔡○○
D 神明	擷取「日月齊明」以前的字串	王令 太上老君 聖眼 日月齊明	去掉「日月齊明」、「聖眼」、空格	王令 太上老君
E 新宮廟/壇與神明尊稱	出現「新裝」擷取下一行的字串直到「今以」結束	天上聖母 蔡家 慈聖媽 千順將軍 天宮武財神 王令 太上老君 金身寶相今以	去掉「今以」、跳行符號、空格	天上聖母 蔡家 慈聖媽 千順將軍 天宮武財神 王令 太上老君 金身寶相
F 年月日	找到〔甲乙丙丁戊己庚辛壬癸〕子丑寅卯辰巳午未申酉戌亥〕的組合並擷取到「日」結束	丁酉 年 十 初 一 日	去掉空格,分配到年月日	丁酉年十月初一日

(三) 步驟三：建立資料庫

原始道士開光疏文資料分不同年度存放在以干支年號命名的資料夾中，本研究開發之 Python 程式會自動搜尋資料夾中的每一個檔案，將前述六項資料的擷取結果以 CSV (Comma-separated Values) 格式寫入以年號為名的 CSV 檔案中，例如：「丁酉.csv」。CSV 的檔案格式為將欄位值以逗號分隔的純文字檔，由於沒有特殊的控制符號，可以利用一般的文字編輯器進行檢視和編輯（例如：Microsoft Windows 內建的記事本程式），適合用於不同系統之間的資料交換。CSV 檔案亦可利用 Excel 進行檢視與編輯(圖 6)，並可供匯入資料庫管理系統中，對於資料之人工檢核校對與建置資料庫均甚為方便。

行號	A 地址	B 祖廟	C 請神人
1			
5	台灣台中市大甲區興安	下林玉聖宮	大甲聖澤堂
6	台灣彰化縣田尾鄉正義村中正路一段	松柏嶺受天宮	盧 陳 盧 劉
7	台灣嘉義市西區下埤里下埤	松柏嶺受天宮	盧 盧
8	台灣屏東縣東港鎮興東里新	琉球朝慈宮	蔡
	台灣台南市新市區民族路	開基玉皇宮	張 張 穆 吳
9			
10	台灣台北市士林區延平北路	雲林縣元長合和村三房路48號	社 府 王
11	台南市台南市新市區	頂太子沙淘宮	林
12	台灣高雄市茄定區金	下茄定金鑾宮	董 董 董
13	台灣高雄化龍宮	高雄化龍宮	顏
14	台灣台中市豐原區田	南廠保安宮	陳

圖 6 以 Excel 程式檢視道士開光疏文資料分析的結果

整體資料處理分析流程如圖 7 所示。

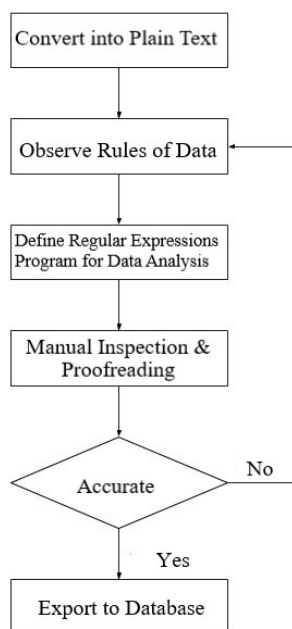


圖 7 道士開光疏文資料處理分析流程圖

三、GIS 地理空間分析與網絡分析

經過上述的文本探勘擷取出具有地址資訊的資料庫，接下來便是利用這些地址資訊查找開光信徒所在之經緯度，以便在地圖上標記其地理位置。

(一) 步驟一：門牌地址批次比對

由於資料庫數目上千筆以上，若以人工一筆一筆查詢，不僅曠日廢時，也可能因人為因素而產生誤差值。所幸內政部的 TGOS 提供官方權威之門牌定位服務，經過申請審查通過，得以進行批次門牌地址比對服務，一般用戶每日比對查詢服務的上限為一萬筆，而政府機關、法人機構與學術單位則有每月 30 萬筆的查詢上限。申請網址為：https://www.tgos.tw/TGOS/Web/Service/Apply/TGOS_Apply_Service.aspx?keyword=批次地址比對服務。經過內政部資訊中心審核通過，可以獲得一組 API 位置、ID 及 APIKey，如下圖所示：



圖 8 取得 API 資訊

使用者必須將 APIKey 複製貼上在「基本參數設定」欄位，並選取擬將地址轉換為哪個坐標系統，選擇包括 EPSG:4326 (WGS84) 國際上通用之系統，包括 Google Maps，以及臺灣常用之澎湖及金馬適用 EPSG:3825 (TWD97TM119)、臺灣地區適用 EPSG:3826 (TWD97TM121)、澎湖及金馬適用 EPSG:3827 (TWD67TM119)、臺灣地區適用 EPSG:3828 (TWD67TM121)等坐標系統。考量之後擬在國際通用地圖上繪圖與標記，因此選用 EPSG:4326 (WGS84) 國際上通用之系統。



圖 9 輸入 APIKey，選擇坐標系統

所需轉換為經緯度之地址，必須依系統所規範之表格格式與欄位設定，系統方能讀取所需要之資訊，包括 id、Address、Response_Address、Response_X、Response_Y (見圖 10)，使用者僅需填寫前兩欄位，將擬查詢經緯度之地址自前述擷取完成之資料庫中複製貼上於此表格，並儲存為 CSV 格式。然後在圖 11 所顯示之網頁上傳新建之地址表格檔案。

	A	B	C	D	E
1	id	Address	Response_Address	Response_X	Response_Y
2	1	新北市中和區平河里1段44巷1號1樓			
3	2	新北市三峽區龍埔里5段1巷1號			
4	3	新北市新莊區中信里13段1巷1號			
5					
6					
7					
8					
9					
10					

圖 10 上傳門牌檔案 CSV 格式之範本

圖 11 比對參數設定

上傳門牌地址檔案後，便須設定比對參數設定。可以選擇「進行完全比對」，比對結果必須與輸入的地址相同，由於此一設定所產生出來的結果往往過於嚴格，地址的書寫方式相當多元，容易有許多失敗的結果，因此建議使用「允許模糊比對」，在找不到輸入的地址而完全比對失敗時，會進行模糊比對，以供使用者參考選擇。

模糊比對規則也有幾項設定，包括是否區分單雙號；誤差範圍在幾號以內；比對是否鎖定於特定縣市、鄉鎮市區、村里、路段或巷弄；當比對不到時是否忽略村里與鄰再進行搜尋；是否限制當比對結果為多筆時，回傳所有的筆數或特定的數量；以及是否一併查詢舊門牌資料，這對於查詢具有歷史之地址有一定的幫助。進行門牌地址比對服務之前，宜對所有門牌地址先行檢視校對一番，減少錯別字、缺漏字、重複字等，有時從 Python 自動擷取之「地址」欄位，實際上僅呈現宮廟、道壇或社團名稱，必須儘可能先在網上查找相關地址，以便比對經緯度的成功率得以提升。經過測試，數字寫成國字數字或全形數字，並不會影響比對的結果。

經過上傳系統進行批次比對，會接到系統回覆之郵件說明已收到申請服務之請求，在幾小時內便能獲得比對結果的通知郵件（見圖 12）。依地址的繁簡程度及所在的區域不同，鄉村地區的地址較常出現找不到指定門牌地址的情況。儘管仍有部分比例的比對失敗結果，但已經為研究者減省許多時間，接下來便是研究者必須自行利用其他 GIS 地圖平臺進行門牌地址的比對工作。



序號	門牌地址	經緯度	比對結果
1	臺南市安平區	22.991111, 120.155556	成功
2	臺南市安平區	22.991111, 120.155556	成功
3	臺南市安平區	22.991111, 120.155556	成功
4	臺南市安平區	22.991111, 120.155556	成功
5	臺南市安平區	22.991111, 120.155556	成功
6	臺南市安平區	22.991111, 120.155556	成功
7	臺南市安平區	22.991111, 120.155556	成功
8	臺南市安平區	22.991111, 120.155556	成功
9	臺南市安平區	22.991111, 120.155556	成功
10	臺南市安平區	22.991111, 120.155556	成功
11	臺南市安平區	22.991111, 120.155556	成功
12	臺南市安平區	22.991111, 120.155556	成功
13	臺南市安平區	22.991111, 120.155556	成功
14	臺南市安平區	22.991111, 120.155556	成功
15	臺南市安平區	22.991111, 120.155556	成功
16	臺南市安平區	22.991111, 120.155556	成功
17	臺南市安平區	22.991111, 120.155556	成功
18	臺南市安平區	22.991111, 120.155556	成功
19	臺南市安平區	22.991111, 120.155556	成功
20	臺南市安平區	22.991111, 120.155556	成功
21	臺南市安平區	22.991111, 120.155556	成功
22	臺南市安平區	22.991111, 120.155556	成功
23	臺南市安平區	22.991111, 120.155556	成功
24	臺南市安平區	22.991111, 120.155556	成功
25	臺南市安平區	22.991111, 120.155556	成功
26	臺南市安平區	22.991111, 120.155556	成功
27	臺南市安平區	22.991111, 120.155556	成功
28	臺南市安平區	22.991111, 120.155556	成功
29	臺南市安平區	22.991111, 120.155556	成功
30	臺南市安平區	22.991111, 120.155556	成功

圖 12 門牌地址比對結果

筆者較常使用的是 Google Maps，除了有時可以直接查詢到 TGOS 全國門牌定位服務平臺所查不到的地址外，還能夠結合街景實境的影像輔助確認地址，宛如親臨實境的實地考察。當然若能夠親自確認，會更為準確，但所花費之時間與經濟成本過高。利用 Google Maps 查找到地址後，在該位置上點選滑鼠右鍵，會出現下拉式選單，點選「這是哪裡？」（圖 13），便會出現該地點之經緯度資訊，若再點選其經緯度之連結，會進一步顯示該地點相關資訊與照片（圖 14），包括景觀實境照片，可供比對確認。此

一階段之校正工作其實是全部流程中最花費時間之瑣碎工作，但在 TGOS 全國門牌定位服務平臺已經提供許多經緯度資訊，為研究者節省了大量的寶貴時間，不必一筆筆查詢，加上現今有 Google Maps 平臺的輔助，使用者可以利用虛擬實境的 360 度 VR 照片，也為查找地址提供了許多的便利性。



圖 13 點擊滑鼠右鍵獲取該點經緯度



圖 14 該經緯度之相關資訊與照片

(二) 步驟二：繪製GIS地圖

經過上述對於道士開光儀式疏文之文本探勘，將所擷取之地址轉化為經緯度，得以在 GIS 地圖上為每次開光的地點定位，繪製出甲午年至戊戌年 (2014-2018 年) 歷年開光地圖。筆者分別利用 QGIS、DocuGIS、TGOS 之地圖協作平臺等 GIS 相關軟體與 WebGIS 平臺進行開光地圖之繪製，由於 GIS 相關之論文與書籍已汗牛充棟 (註 3)，不在本文詳細之範圍，僅藉由地圖視覺化來呈現歷年開光儀式之整體趨勢和區域特徵。

限於篇幅，以甲午年至丙申年的變化趨勢來觀察 (圖 15)，吳道長從南部為主的開光儀式服務，逐漸向中北部及東部開拓。而丙申年開始，更向海外擴展 (圖 16)，包括日本、南京、新加坡，與吳政憲道長擅用網絡與信徒溝通的做法有著密切的關聯，海外信徒透過網路平臺而知道其開光儀式服務，進而邀請吳道長專程前往，信徒還需支付其住宿與交通費用，成本並不低廉，可以見得吳道長之開光儀式口碑已經建立。再觀察丁酉年與戊戌年的變化 (圖 17)，可以發現數量上持續增加外，版圖也不斷擴張，包括新竹、北海岸、屏東皆有新增地點。

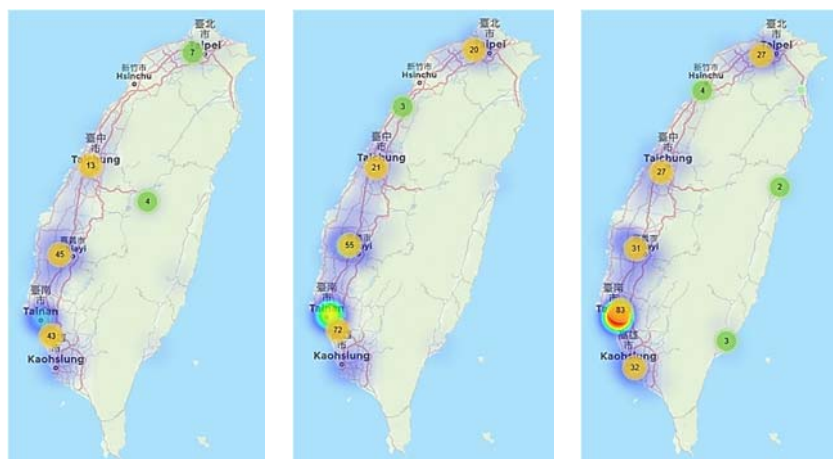


圖 15 甲午、乙未、丙申年開光主家空間分佈

道教開光儀式疏文之文本探勘與數位人文探索：以府城延陵道壇為例

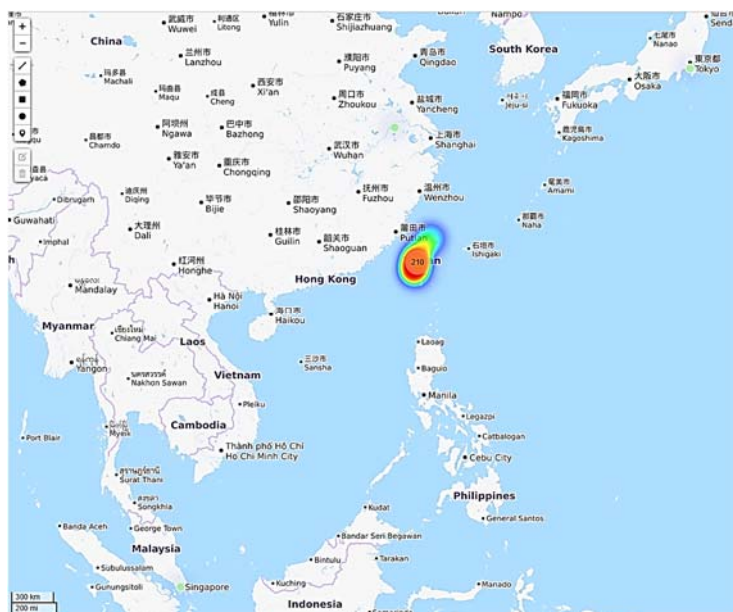


圖 16 丙申年開光主家空間分佈：海外日本東京、南京、新加坡



圖 17 甲午至戊戌年延陵道壇開光儀式空間分佈

(三) 開光疏文中之網絡關係探索

從上述文本探勘所擷取出之資料庫表格中，亦可以進行不同欄位間的網絡關係探索。以甲午年為例，筆者將開光之神明與開光地點進行關係網絡之視覺化呈現，利用 Stanford 大學所開發之 Palladio 平臺，繪製網絡關係圖，由於網絡分析之相關文章與論文亦有不少，本文不再詳述流程，僅舉例說明開光疏文中的資訊經過數位人文之文本探勘，可以展現哪些新的可能性，並指引新的研究議題。

從完整的網絡關係圖(圖 18)來觀察，吳政憲道長於甲午年所開光之神尊，以中壇元帥、天上聖母、廣澤尊王、天官武財神為大宗，中壇元帥以新營太子宮為主要地點，但也包括在臺南的其他太子宮廟及自宅(圖左較大之深色點)。再細看簡化後的關係網絡(圖 19)，天上聖母卻不是以某間宮廟為主，而是分散於不同的天上聖母宮廟，顯示出不一樣的型態。有別於天官武財神集中於北港武德宮的現象，廣澤尊王則分別在臺南兩大廣澤尊王信仰中心：開基永華宮與西羅殿。

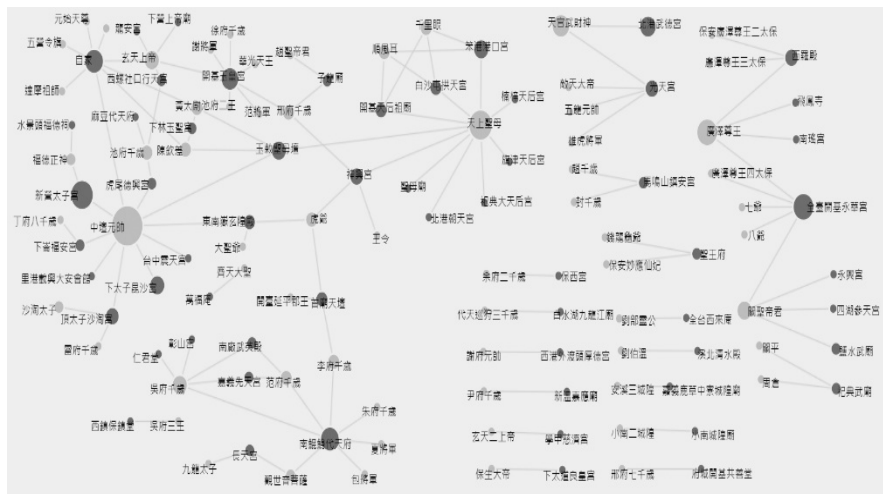


圖 18 甲午年開光地點與神尊之關係網絡
(淺色為開光神明，深色為開光地點)

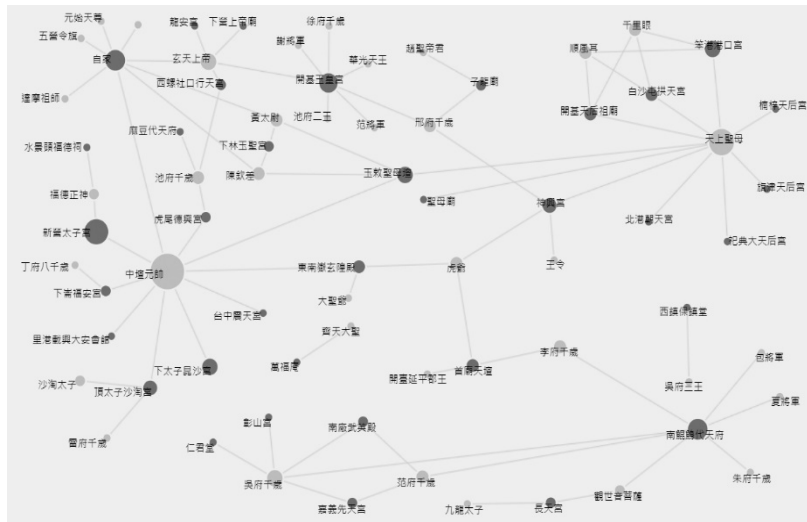


圖 19 甲午年開光地點與神尊之關係網絡：簡化版
(淺色為開光神明，深色為開光地點)

參、結語

藉由資訊技術與工具的協助，將橫跨五年、上千人之開光疏文，進行大規模之整理，並從中擷取出開光神明、祖廟、新宮廟、開光信徒及其地址等資訊，便能進一步開展更多研究的可能性。本文前半段詳細展示研究之流程，後半段則限於篇幅僅簡單舉例說明，可以如何應用 GIS 和社會網絡分析來進一步處理資料，進行數位人文探索。透過 Python 程式語言和 TGOS 的批次門牌地址定位之協助，大量節省整理疏文之時間，並建置成可以進一步分析之資料庫，應可作為資訊技術輔助人文研究之示範。

其他相似的道教科儀疏文，如禮斗、建醮、年例祭典等，亦可以利用本文之相關文本探勘技術來擷取其中的內容資訊，建置成該科儀之信仰資料庫，再利用數位人文方法進行探索。

數位人文的探索往往能夠協助研究者掌握資料整體的趨勢與圖像，避免見樹不見林之弊病。從本次處理的道士開光疏文中，還發現神明開光日期有兩大特色，農曆七月不舉行開光儀式，以被稱為「鬼月」之民俗月份

不進行神明開光儀式，並不難理解，但透過大量文獻也提供了實證的基礎。此外還發現延陵道壇每月農曆十五日亦很少開光，這究竟是普遍現象，抑或是吳政憲道長個人的時間安排，還有待進一步深入訪談，並觀察其他道壇是否也有此現象，來解開這個謎題。

本文旨在展示利用數位人文技術與工具，如何協助人文與社會研究，善用資訊技術可以快速而大量處理具有模式化的文獻(如道士疏文)，擷取其中關鍵性的資料內容，再利用 GIS 與網絡分析之數位人文技術與工具之新視角，探索開光疏文之空間意涵與網絡關係，發揮數位人文擅長之「鳥瞰」(distant reading)，觀察文獻之整體趨勢，即使無法直接解答研究問題，亦能提出新的研究問題意識，以作為下一步「近讀」(close reading)之基礎，結合歷史研究與人類學田野調查，與數位人文相輔相成，而且誠如祝平次引述林富士在發表《數位人文白皮書》時的說法：「在數位分析結果的基礎上，進一步進行的人文研究。就實際研究進程中，數位的分析和人文學的探問，常常要來回往覆，不止一次。」(祝平次，2019)若以三軍聯合作戰來作比喻，數位人文提供空軍的鳥瞰視野，歷史與文獻研究可以像陸軍作戰必須腳踏實地勤查資料，並廣泛閱讀，而田野調查與深入訪談，有時像海軍潛水一般，更加深入其中，若能結合三者之長處，截長補短，相信研究成果可以更為「見樹又見林」。

(接受日期：2020 年 12 月 21 日)

註釋

註 1：參見國家教育研究院，雙語詞彙、學術名詞暨辭書資訊網，「正規表式；正規表示式；常規表示式」辭條

<http://terms.naer.edu.tw/detail/11989920/?index=1>；

維基百科全書，「正規表示式」辭條

<https://zh.wikipedia.org/wiki/%E6%AD%A3%E5%88%99%E8%A1%A8%E8%BE%BE%E5%BC%8F>

註 2：The GreenEnd Organisation，<http://www.greenend.org.uk/rjk/tech/regex.html>
Wikipedia, Regular expression, https://en.wikipedia.org/wiki/Regular_expression

註 3：利用 GIS 於宗教民俗研究，可參見洪瑩發、范毅軍、張智傑、廖泫銘(2020)，空間視野與地方知識：臺灣宗教與時空分析的回顧與展望。《華人宗教研究》，16，99-136。

參考文獻

- 王汎森(2014)。數位人文學之可能性及限制：一個歷史學者的觀察。在項潔編，*數位人文研究與技藝*(頁 25-35)。臺北市：國立臺灣大學出版中心。
- 杜協昌(2016)。半自動詞彙擷取：簡化的詞夾子方法以及其 JavaScript 元件的開發與應用。在項潔編，*數位人文：在過去、現在和未來之間*(頁 171-206)。臺北市：國立臺灣大學出版中心。
- 祝平次(2019)。數位人文學專刊前言。《清華中文學報》，22，5-8。
- 洪瑩發、林長正(2013)。《臺南傳統道壇研究》。臺南市：臺南市政府文化局。
- 洪瑩發、范毅軍、張智傑、廖泫銘(2020)。空間視野與地方知識：臺灣宗教與時空分析的回顧與展望。《華人宗教研究》，16，99-136。
- 彭維謙、程卉、陳詩沛(2018)。從全文到表格：地方志職官志中職官資料之半自動擷取。《數位典藏與數位人文》，1，79-125。
doi:10.6853/DADH.201804_1.0004
- 張素玢、李鈺淳(2014)。資料擷取與描述比對：清代臺灣方志物產篇分析系統之建置與解釋。在項潔編，*數位人文研究與技藝*(頁 67-89)。臺北市：國立臺灣大學出版中心。
- 謝育平(2011)。同位詞夾子：主題式分類詞庫萃取演算法。在項潔編，*數位人文研究的新視野：基礎與想像*(頁 133-162)。臺北市：國立臺灣大學出版中心。
- Stanley-Baker, M., & Ho, B. H.-L. (2015). Mapping drugs across epistemic and geographic domains: A case study for early medieval China. *6th International Conference of Digital Archives and Digital Humanities*. Symposium conducted at the meeting of Research Center for Digital Humanities. National Taiwan University, Taipei. Retrieved from <https://hdl.handle.net/10356/88481>
- Stanley-Baker, M., & Chong Eng Keat (2019). Materia medica in Chinese religious sources: Towards a critical digital philology for modelling knowledge distribution

- in early Chinese texts. *2019 Pacific Neighborhood Consortium Annual Conference*. Symposium conducted at the meeting of NTU Digital Humanities Research Cluster. Nanyang Technological University, Singapore. Retrieved from https://www.researchgate.net/profile/Michael_Stanley-Baker/publication/338065561_Materia_Medica_in_Chinese_Religious_Sources_Towards_A_Critical_Digital_Philology_for_modelling_Knowledge_Distribution_in_Early_Chinese_texts/links/5dfc50b4299bf10bc368e6fe/Materia-Medica-in-Chinese-Religious-Sources-Towards-A-Critical-Digital-Philology-for-modelling-Knowledge-Distribution-in-Early-Chinese-texts.pdf
- Stanley-Baker, M. (2019). Daoing medicine: Practice theory for considering religion and medicine in early Imperial China. *East Asian Science, Technology, and Medicine (EASTM)*, 50, 21-66.
- Sundararajan, L., Ting, R. S.-K., Hsieh, S.-K., & Kim, S.-H. (2020). Religion, cognition, and emotion: What can automated text analysis tell us about culture? *The Humanistic Psychologist*. Advance online publication. Retrieved from <https://doi.org/10.1037/hum0000201>
- Yang, Y. B., Jin, J., Wang, S., & Wang, L. (2015). Secrets of the secrets: Text mining on Dao Canon. *The Open Cybernetics & Systemics Journal*, 9, 2536-2544. doi:10.2174/1874110X01509012536



Text-Mining and Digital Humanities Analysis of “Eye-Opening” Consecration Ritual of Contemporary Yanling Daoist Altar of Tainan

Ching-Chih Lin*

【Abstract】

This article uses Digital Humanities methods and technologies, Text mining and GIS (geographic information system) to explore the ritual texts of the Deity “Eye-Opening” Consecration Ritual of Daoist Wu Zhengxian at the Yanling Daoist Altar in Tainan. The author uses the Python data extraction technology to extract the information of deities, temples, and believers contained in the consecration ritual texts, and then uses GIS time-space analysis and social network analysis to examine the rituals of the Yanling Daoist altar. The article demonstrates the visualization of the spatial distribution of the consecration rituals of the Yanling Daoist Altar. In addition, it also illustrates the network relationship between the deities and the ritual locations, mostly temples. This article aims to show how the use of digital humanities technologies and tools can assist humanities research. The proper use of information technology can quickly and massively process patterned documents (such as Daoist ritual texts), extract key data content, and then use GIS spatial and

* Assistant Professor, Graduate Institute of Religious Studies, National Chengchi University
ORCID 0000-0002-6791-6561
E-mail: cclin52@gmail.com

network analysis to explore the spatial significance and network relationship behind the ritual texts. The author tries to make use of the “distant reading” of digital humanities and proposes new research questions for future studies combined with the “close reading” of traditional research approaches.

Keywords

Digital Humanities, Text-mining, Information visualization, GIS (Geographic Information System), Social network analysis

【Summary】

This article uses Digital Humanities methods and technologies, Text mining and GIS (geographic information system), to explore the ritual texts of the Deity “Eye-Opening” Consecration Ritual of Daoist Wu Zhengxian at the Yanling Daoist Altar in Tainan. The author will use the Python data extraction technology to extract the information of deities, temples, and believers contained in the consecration ritual texts, and then use GIS time-space analysis and social network analysis to examine the rituals of the Yanling Daoist altar. The article will demonstrate the visualization of the spatial distribution of the consecration rituals of the Yanling Daoist Altar. In addition, it will also illustrate the network relationship between the deities and the ritual locations, mostly temples. This article aims to show how the use of digital humanities technologies and tools can assist humanities research. The proper use of information technology can quickly and massively process patterned documents (such as Daoist ritual texts), extract key data content, and then use GIS spatial and network analysis to explore the spatial significance and network relationship behind the ritual texts. The author tries to make use of the “distant reading” of digital humanities and proposes new research questions for future studies combining with the “close reading” of traditional research approaches.

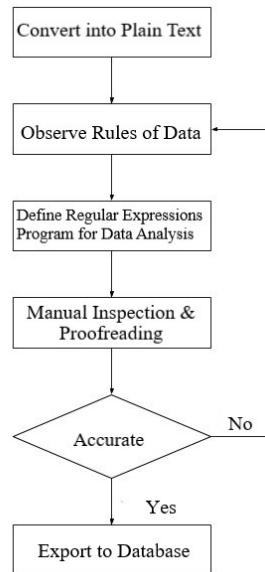
The Daoist Master Wu Zhengxian of Yanling Daoist Altar in Tainan,

inherits the family’s long Daoist tradition, and is good at utilizing modern technology and social media (such as Facebook, Line group) to interact with followers. He is renowned for his Deity “Eye-Opening” Consecration Ritual. He has formally promoted to Grand Daoist Master since the Jiao Ceremony of Kaiji Jade Emperor Temple in 2007. After more than ten years of ceremonial services, his popularity and ritual activities have spread all over Taiwan, and he even went abroad to serve believers and temples for ritual services. In 2017, Daoist Master Wu held a ceremony to reunite the clients of the consecration rituals of the deities celebrate the tenth anniversary of his promotion ceremony. The number and distribution of the deities consecrated by Daoist Master Wu have gradually accelerated its expansion and spread in recent years. The author has obtained the trust of Daoist Master Wu Zhengxian, and acquired the consecration ritual texts from 2014 through 2018. Each year, Daoist Master Wu Zhengxian presides consecration rituals for hundreds times at various locations in Taiwan and is sometimes invited to provide ritual services abroad. The numbers of deity statues he consecrated in the past few years have exceeded one thousand per year.

Existing researches on Daoism using technologies of digital humanities mostly analyze Daoist classics (Stanley-Baker, 2019; Stanley-Baker & Chong, 2019; Stanley-Baker & Ho, 2015; Sundararajan, Ting, Hsieh, & Kim, 2020; Yang, Jin, Wang, & Wang, 2015). Rare research has dealt with modularized ritual texts. Researches using text-mining technologies on non-Chinese religious texts are mostly not applicable to Chinese format. Since there are basic modularity in the consecration ritual texts, Daoist priests fill in several basic fields, which is simpler than ordinary documents. Python programming language can be used to retrieve the content and extract basic data from the consecration ritual texts. Some technologies of digital humanities, such as “term clip” extraction algorithm, “semi-automatic vocabulary extraction”, or data extraction technology for huge documents, are not applicable in this study (Chang & Li, 2014; Du, 2016; Hsieh, 2011; Peng, Cheng, & Chen, 2018).

This article focuses on how to use digital humanities technologies and tools

to conduct text exploration of religious texts, using the Python programming language to extract the content of the consecration ritual texts. Then the author will build a database of consecration rituals, and further draw GIS maps and visualize the spatial distribution of the consecration rituals. Meanwhile, the author will also visualize the network relationship between the consecrated deities and the temples by using the Palladio platform social network analysis technology developed by Stanford University.

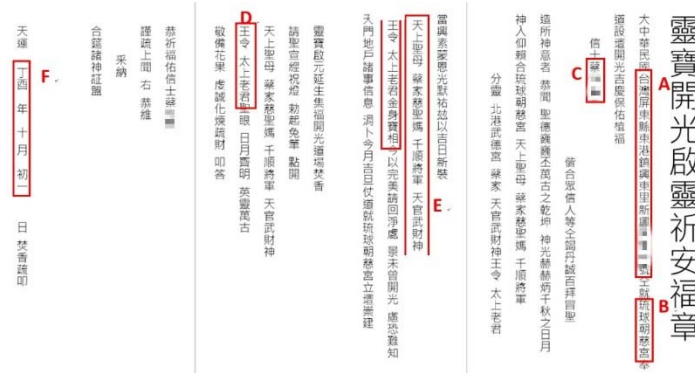


Data Extraction Process

By using the data extraction technologies and tools of digital humanities, the author explores the information of the deities, believers and temples in the consecration ritual texts and utilizes GIS Geospatial analysis and network analysis of digital humanities which is good at processing a large number of information speedily. This essay utilizes text-mining technology to extract the names and addresses of the clients and the titles of the consecrated deities from the highly modularized ritual texts, and then uses GIS-related platform tools and technologies to transform geographical information into digital maps. This

Text-Mining and Digital Humanities Analysis of “Eye-Opening” Consecration Ritual of Contemporary Yanling Daoist Altar of Tainan

essay can also analyze the network relationship between the consecration locations and the deities, so the author can observe the connection between the deities and the temples. First, the author converts all the consecration ritual texts into plain texts, and then observe the modular rules of the data. By defining the regular expressions, the author utilizes Python to write the program for data extraction analysis. After the result comes out, the author needs to sift and proofread errors. All accurate data later will be exported to create a database for further processing.

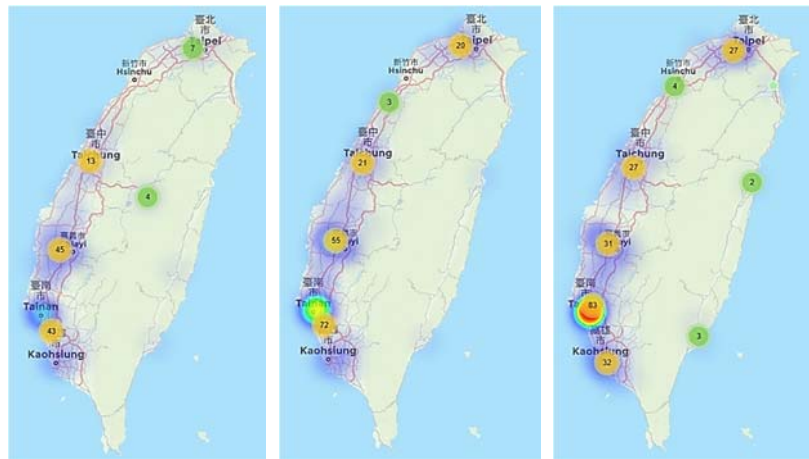


Modularity of the Consecration Ritual Text

地址	祖廟	請神人
台灣台中市大甲區興安	下林玉聖宮	大甲聖澤堂
台灣彰化縣田尾鄉正義村中正路一段	松柏嶺受天宮	盧 陳 盧 劉
台灣嘉義市西區下埤里下埤	松柏嶺受天宮	盧 盧 盧
台灣屏東縣東港鎮興東里新	琉球朝慈宮	張 吳
台灣台南市新市區民族新	開基玉皇宮	張 樞 吳
台灣台北市士林區延平北路	雲林縣元長合和村三房路48號	社 府 王
台南市台南市新市區	頂太子沙淘宮	林
台灣高雄市茄萣區金	下茄萣金鑾宮	董 董 董 董 董
台灣高雄化龍宮	高雄化龍宮	顏
台灣台中市豐原	南廠保安宮	陳

Database of Extracted Information from the Consecration Ritual Texts

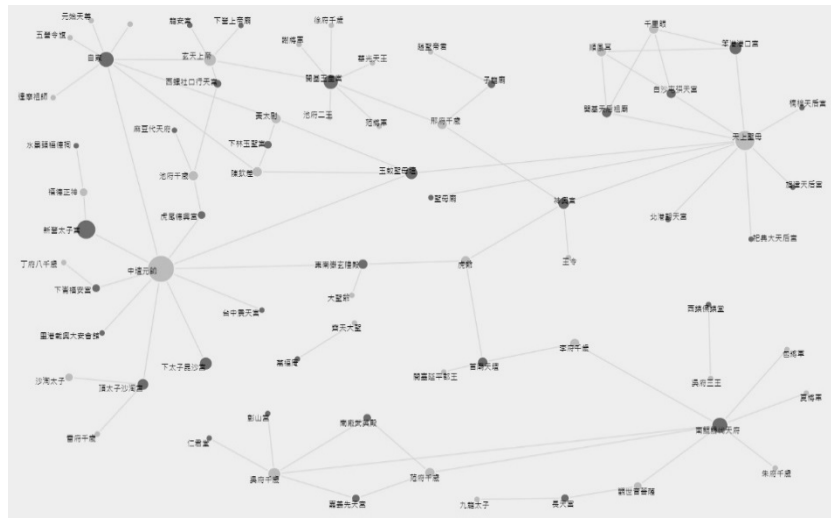
Daoist consecration ritual texts have a highly modular tendency, and it is just suitable to use information extraction technology to capture key contents, including titles of consecrated deities, the temple names of the original or subsidiary deity statues, and the names of clients. The relevant information is extracted and transformed into a table and creates a database. By using the addresses of the clients, we can obtain latitudes and longitudes for coordination. The author utilizes the address-transfer platform of Taiwan Geospatial One Stop (TGOS) of Taiwan's Ministry of Interior. If TGOS platform cannot find the locations, the author will use Google Map and the Street View function to obtain the coordination information. The coordination data can be further visualized into maps to demonstrate the geographical spatial distributions of the consecrated deities by GIS-related technologies and tools. The author will use the extracted data and visualization obtained from this study to compose another essay focusing on the interpretation in the field of religious studies later.



Spatial Distribution of the Consecration Rituals

In the meantime, the network relationship between the deities and the temples can also be visualized by using the technologies, platforms, and tools of social network analysis.

Text-Mining and Digital Humanities Analysis of “Eye-Opening” Consecration Ritual of Contemporary Yanling Daoist Altar of Tainan



Network of Consecrated Deities and Temples

Romanized & Translated Reference for Original Text

王汎森 (2014)。數位人文學之可能性及限制：一個歷史學者的觀察。在項潔編，*數位人文研究與技藝* (頁 25-38)。臺北市：國立臺灣大學出版中心。

【Wang, Fan-Sen (2014). The possibilities and limitations of digital humanities: A historian's perspective. In Hsiang, Jieh (Ed.), *Digital humanities and craft: Technological change* (pp. 25-38). Taipei: National Taiwan University Press. (in Chinese)】

杜協昌 (2016)。半自動詞彙擷取：簡化的詞夾子方法以及其 JavaScript 元件的開發與應用。在項潔編，*數位人文：在過去、現在和未來之間* (頁 171-206)。臺北市：國立臺灣大學出版中心。

【Tu, Hsieh-Chang (2016). Semi-Automatic term extraction with simplified term-clips method: Development and applications of a JavaScript component. In Hsiang, Jieh (Ed.), *Digital humanities: Between past, present, and future* (pp. 171-206). Taipei: National Taiwan University Press. (in Chinese)】

祝平次 (2019)。數位人文學專刊前言。*清華中文學報*，22，5-8。【Chu, Ping-

- Tzu (2019). Shu wei ren wen xue zhuan kan qian yan. *Tsing Hua Journal of Chinese Literature*, 22, 5-8. (in Chinese)】
- 洪瑩發、林長正 (2013)。臺南傳統道壇研究。臺南市：臺南市政府文化局。
【Hung, Ying-Fa, & Lin, Chang-Zheng (2013). *Tainan chuan tong dao tan yan jiu*. Tainan: Cultural Affairs Bureau, Tainan City Government. (in Chinese)】
- 洪瑩發、范毅軍、張智傑、廖泫銘 (2020)。空間視野與地方知識：臺灣宗教與時空分析的回顧與展望。《華人宗教研究》，16，99-136。【Hung, Ying-Fa, Fan, I-Chun, Zhang, Zhi-Jie, & Liao, Hsiung-Ming (2020). Kong jian shi ye yu di fang zhi shi: Taiwan zong jiao yu shi kong fen xi de hui gu yu zhan wang. *Studies in Chinese Religions*, 16, 99-136. (in Chinese)】
- 彭維謙、程卉、陳詩沛 (2018)。從全文到表格：地方志職官志中職官資料之半自動擷取。《數位典藏與數位人文》，1，79-125。【Pang, Wai-Him, Cheng, Hui, Chen, Shih-Pei (2018). From text to data: Extracting posting data from Chinese local gazetteers. *Journal of Digital Archives and Digital Humanities*, 1, 79-125. (in Chinese)】 doi:10.6853/DADH.201804_1.0004
- 張素玢、李鈺淳 (2014)。資料擷取與描述比對：清代臺灣方志物產篇分析系統之建置與解釋。在項潔編，*數位人文研究與技藝* (頁 67-92)。臺北市：國立臺灣大學出版中心。【Chang, Su-Bing, & Li, Yu-Chuen (2014). Data extraction and mapping: The taxonomy of fauna and flora of Taiwan local gazetteers in the Qing dynasty. In Hsiang, Jieh (Ed.), *Digital humanities and craft: Technological change* (pp. 67-92). Taipei: National Taiwan University Press. (in Chinese)】
- 謝育平 (2011)。同位詞夾子：主題式分類詞庫萃取演算法。在項潔編，*數位人文研究的新視野：基礎與想像* (頁 133-162)。臺北市：國立臺灣大學出版中心。【Shieh, Yuh-Pyng (2011). Appositional term clip: A subject-oriented appositional term extraction algorithm. In Hsiang, Jieh (Ed.), *New eyes for discovery: Foundations imaginations of digital humanities* (pp. 133-162). Taipei: National Taiwan University Press. (in Chinese)】
- Stanley-Baker, M., & Ho, B. H.-L. (2015). Mapping drugs across epistemic and geographic domains: A case study for early medieval China. *6th International Conference of Digital Archives and Digital Humanities*. Symposium conducted at the meeting of Research Center for Digital Humanities. National Taiwan

Text-Mining and Digital Humanities Analysis of “Eye-Opening” Consecration
Ritual of Contemporary Yanling Daoist Altar of Tainan

University, Taipei. Retrieved from <https://hdl.handle.net/10356/88481>

- Stanley-Baker, M., & Chong, Eng Keat (2019). *Materia medica in Chinese religious sources: Towards a critical digital philology for modelling knowledge distribution in early Chinese texts*. *2019 Pacific Neighborhood Consortium Annual Conference*. Symposium conducted at the meeting of NTU Digital Humanities Research Cluster. Nanyang Technological University, Singapore. Retrieved from https://www.researchgate.net/profile/Michael_Stanley-Baker/publication/338065561_Materia_Medica_in_Chinese_Religious_Sources_Towards_A_Critical_Digital_Philology_for_modelling_Knowledge_Distribution_in_Early_Chinese_texts/links/5dfc50b4299bf10bc368e6fe/Materia-Medica-in-Chinese-Religious-Sources-Towards-A-Critical-Digital-Philology-for-modelling-Knowledge-Distribution-in-Early-Chinese-texts.pdf
- Stanley-Baker, M. (2019). Daoing medicine: Practice theory for considering religion and medicine in early Imperial China. *East Asian Science, Technology, and Medicine (EASTM)*, 50, 21-66.
- Sundararajan, L., Ting, R. S.-K., Hsieh, S.-K., & Kim, S.-H. (2020). Religion, cognition, and emotion: What can automated text analysis tell us about culture? *The Humanistic Psychologist*. Advance online publication. Retrieved from <https://doi.org/10.1037/hum0000201>
- Yang, Y. B., Jin, J., Wang, S., & Wang, L. (2015). Secrets of the secrets: Text mining on Dao Canon. *The Open Cybernetics & Systemics Journal*, 9, 2536-2544. doi:10.2174/1874110X01509012536