

From CBDB to TBDB

Using the *New Edition of Changhua Gazetteer* as a Starting Point

SU-BING CHANG

NATIONAL TAIWAN NORMAL UNIVERSITY

Outline

Introduction

Taiwan Biographical Database (TBDB)

Textual research

System Architecture of TBDB

Preliminary Results

Automatic Extraction of Poetry Societies

Conclusions

Introduction

Taiwan Biographical Database (TBDB)

A database for Taiwanese historical figures. Based on the concepts and structure of China Biographical Database.

It sets up a database and a set of analytical tools according to the characteristics of modern Taiwanese history and historical figures.

One of the major purposes of developing the TBDB is to model the historical experience of Taiwanese history, either through prosopography, SNA, or GIS analyses.

Since the core participants of the project are also the team members responsible for the biographies of the *New Edition of Zhanghua Gazetteer* and have deep understanding on the categories and attributes of the historical figures recorded in the gazetteer, we decided to begin the TBDB project with the historical figures of Zhanghua.

Inspired by CBDB



Peter Bol

Robert M. Hartwell
(1932-1996)

Michael Fuller



China Biographical Database (CBDB)

*Harvard University, Academia Sinica and Peking University

- * Collect biographical data on figures in Chinese history
 - About 422,600 Chinese figures (2018/09)
 - 7-19 centuries
- * Freely available
 - downloadable, online manual inputting system, online query system
- * Allow users to use big data to explore and to do the research

from CBDB



to TBDB



台灣歷史人物傳記資料庫

Taiwan Biographical Database (TBDB)

Difference Between CBDB and TBDB

Time

- CBDB: before 20th century
- TBDB: from 16th century (late 19th-20th)

Personage category

- CBDB: government officials
- TBDB: various professions

Social status

- CBDB: Government officials with different categories and levels
- Diverse social statuses such as legislators, doctors, and capitalist

Difference Between CBDB and TBDB (Cont.)

Ethnical group

- CBDB: primarily Han people
- TBDB: Han, aborigine and foreigners

Societies: ex. poetry societies and unions

- Imply some kinds of social associations, and exerts their influence in fields such as politics, economy, society, and culture

Overseas experiences

- TBDB stores individuals' overseas experiences, such as schooling, travelling, migration, business activity and work

如何起步？

修訂增補CBDB

建置TBDB

CBDB

TBDB

時間

8 - 15世紀為主
宋代資料尤多

16世紀以降
重心在20世紀前後

奠定社會地位
之途徑

單一

多元

社會流動

固著化

活潑化

社會階層

官宦士人

士農工商、各行各業

資料來源

傳記資料索引、正史
列傳、墓誌銘、文學
作品、書信、官方文
書、其他資料庫

史料、檔案、志書、
報紙、日記、書信、
照片、私文書、族譜、
證書、戶籍資料、人
物鑑.....

Start from *New Edition of Changhua Gazetteer*

Form the project team in 2016/08

- 7 PIs: 5 history, 2 computer science and LIS

Why started from *Changhua Gazetteer*?

- Changhua County has obtained its important position since Qing Dynasty
 - First large irrigation system in Taiwan
 - Most important agriculture area (Rice Growing)
 - First county with population more than one million in 17th century

Changhua County



Start from *New Edition of Changhua Gazetteer*

Data Source: *Changhua Gazetteer*

Four authors are the team member

- The project can use the full texts and photos of biographies without copyright infringement
- helpful in judging the reliability and validity of software services
- The same writing pattern eases the complexity of computerized processing

Pattern of the text

Year of birth and dead

張清華（1884—1948）

First paragraph
Birth place/style name/Parents

員林人，居於今員林市民生里。父張紹乾（另有傳）為員林地區地主，母陳陽。

自幼受到良好的漢學教育，受業於庠生江捷元。明治 32 年（1899）入讀員林公學校，畢業後經營秋田製材商會。

明治 38 年（1905），彰化縣官設鐵道完成，二水站、彰化站、花壇站，員林站、社頭站、田中站通車，捐店面數間供火車站廣場使用；同年 6 月 5 日株式會社彰化銀行創立，為成立委員之一。明治 40 年擔任員林街第一保保正，兼任員林保甲聯合會會長，在地方法上聲望卓著。大正 6 年（1917）11 月，獲臺灣總督府頒授紳章。大正 8 年任大坡心庄長，一年後獲命擔任員林街長，時為員林唯一臺灣人街長，至昭和 10 年（1935）11 月退職為止，共任職 4 屆街長，有「長壽街長」之稱。¹昭和 9 年，員林站站前擬建廣場，便將自家在站前的數間店面拆除，無償捐出土地供廣場之用；翌年獲「從七位」位階榮銜，受



張清華受勳紀念
（張哲豪提供）

他在 16 年的街長任期內，相當注重興辦學校以修興賢書院，大正 12 年協助員林公學校崇頤井分教員林公學校高等科。除公辦教育事業之外，也在自宅張姓子弟，人稱街長宅第漢學堂。

任內也完成多項建設，如大正 13 年（1924）配合工廠，由員林街役場負責管理，使當地衛生條件大幅提升。公園發揮了很好的救災效用，遂極力爭取長期租借興池儲水以備救災之用，即為目前市公所前的員林公園綠地。而在大正 14 年 10 月完工的自來水道建設延伸至員林公園，設立管理站及游泳池，今日仍可見當年的水管遺跡。

政治生涯之外，他也在產業界大展長才，除了在水裡坑開設秋田製材商會，大正 2 年（1913）員林信用購買利用組合設立，出任組合長，並創立員林魚市場外魚類賣買商合資組織。²隔年，擔任羊豬仲買組合長、果物容器販賣信用組合長。大正 10 年，出任員林街農業組合長、員林信用購買組合長。大正 15 年成立員林街玉壺春製冰公司，並出任玉壺春及斗六兩林地區青果業和農產品加工業的競爭力。昭和 12 年合作社組合員、臺中輕鐵會社監察役（監事），以及

此外，建成兩棟西洋樓，前棟落成於明治 43 年（1910），後棟落成於昭和 5 年（1930），乃自日本聘請建築師，以西方文藝復興樣式打造，被稱為「員林街長宅第」，是當時員林所僅見。兩棟樓的建造時間相差 20 年，在造型和風格上各異其趣，建築材料亦不相同，前低（一樓）後高（二樓）。現今張家仍有後代居住於後棟。

妻黃惜，育有 3 子。長男春鐘，日本早稻田大學政治經濟系畢業；次男春申，臺中一中畢業；三子春茂，日本大學法律系畢業；⁴長女春子，彰化高女畢業後，隨長兄留學日本。

Second paragraph
Education

Third and four paragraph
Career/achievement

last paragraph
families

¹ 趙水溝，《員林郡大觀》（臺北：臺灣新民報社，1936），頁 245-246。

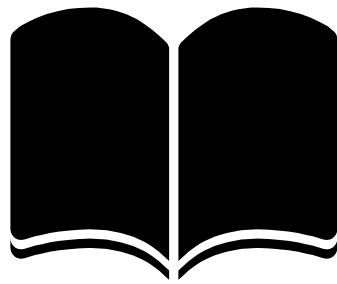
² 內藤素生，《南國之人士》（臺北：臺灣人物社，1922），頁 222。

³ 林連發，《臺灣官紳年鑑》（臺北：民眾公論社，1933），頁 208。

⁴ 臺灣新民報社調查部編，《臺灣人士鑑（日刊一週年版）》（臺北：臺灣新民報社，1934），頁 113。

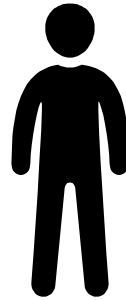
Starting Point

New Edition of Zhanghua Gazetteer



4

volumes



887



103.4

Millions words

1912

1684

1895

1945

2017

Special Figures in *Changhua Gazetteer*

- Some examples

Most influential figure

辜顯榮(1866-1937)



- The first Taiwanese representative in Upper House of Japan
- Exclusive right of Salt, Camphor Opium, Tabaco
- 1915 Business consultor of 段祺瑞
- 1933 Help 蔣介石 to suppress the rebellion of 閩 (閩變)

Highest official (post war)

謝東閔(1908-2001)



- Graduated from Guangdon Zhongshan Univ.
(廣東中山大學)
- 9th Chairman of Taiwan Provincial Government
- 6th Vice president of Republic of China

The first Olympic athlete in Taiwan History

林月雲 (1915-1992)

graduated from Zhanghua girl high School

- 1932 : broken the jump records of Japan in the Olympic athletes election
- 1934 : the champion of triple jump competition in Japanese Empire
- 1935 : the Japanese trainer of 1936 Olympic Game in Berlin
- post war: teaching the “Track and Field” course



我的半生記



The first female Communist in Taiwan 謝雪紅 (1901-1970)

- 1928 co-established "Taiwan Communist Party" in Shanghai French Concession

1947 fled to Xiamen and never came back to Taiwan

- 1949 attended 「第一屆中國人民政治協商會議」 and the member of the presidium





Commandment by Mao



Textual research

台灣總督府職員錄系統 (1896-1944)

who.ith.sinica.edu.tw/mpView.action

您尚未安裝Flash Plugin, 部份頁面將無法順利操作!

回首頁 | 臺史所首頁 | 臺史所檔案館
您的IP是220.129.154.130

字體大小: 小 標準 大 最大

瀏覽服務 進階查詢 專題名錄 電子書櫃 系統公告 常見問題 分享連結

快速查詢 Quick Search

查詢內容:
檢索欄位: 不限欄位
範圍: ☒ 名錄

進階查詢 查詢說明

熱門查詢詞 Popular search terms
書記 臺北州 法院 沖繩
總督官房 安平 文書課 總務長官 民政局

職官錄電子書 Official record books

展示全部51冊

職員錄 明治29年
職員錄 明治30年
職員錄 明治31年
職員錄 明治32年
職員錄 明治33年

組織瀏覽 Organization View

顯示全部組織圖

Get ADOBE FLASH PLAYER

臺灣總督府 展開次層 瀏覽名錄
府內 所屬 地方 展開次層 展開次層 展開次層

上午 09:45 2018/10/1

現在位置 [首頁](#) > 查詢結果

查詢結果

您查詢的條件為：

查詢類別為 不限欄位=辜顯榮

縮小範圍查詢

重新查詢

檢視模式：☒ 條列式 ☐ 目錄式

已選0筆 (上限300筆)：☐ 匯出 ☐ 列印 ☐ 轉寄 ☐ 比較 ☒ 清空

共 0 筆，每頁筆數 ，目前在第 頁

<input type="checkbox"/>	序號	姓名	日本紀年	西元紀年	單位名稱	官職名	頁碼	調閱影像	勘誤	延伸查詢
<input type="checkbox"/>	0	辜顯榮	明治三十一年	1898	臺中縣鹿港辨務署	參事	56			
<input type="checkbox"/>	1	辜顯榮	明治四十三年	1910	臺中廳庶務課	參事	263			
<input type="checkbox"/>	2	辜顯榮	明治四十四年	1911	臺中廳庶務課	參事	272			
<input type="checkbox"/>	3	辜顯榮	明治四十五年	1912	臺中廳庶務課	參事	335			
<input type="checkbox"/>	4	辜顯榮	大正二年	1913	臺中廳庶務課	參事	209			
<input type="checkbox"/>	5	辜顯榮	大正三年	1914	臺中廳庶務課	參事	218			
<input type="checkbox"/>	6	辜顯榮	大正四年	1915	臺中廳庶務課	參事	220			
<input type="checkbox"/>	7	辜顯榮	大正五年	1916	臺中廳庶務課	參事	229			
<input type="checkbox"/>	8	辜顯榮	大正六年	1917	臺中廳庶務課	參事	238			
<input type="checkbox"/>	9	辜顯榮	大正七年	1918	臺中廳庶務課	參事	248			
<input type="checkbox"/>	10	辜顯榮	大正八年	1919	臺中廳庶務課	參事	275			
<input type="checkbox"/>	11	辜顯榮	大正九年八月	192008	臺中廳庶務課	參事	281			
<input type="checkbox"/>	12	辜顯榮	大正九年十二月	192012	臺中州州協議會員	州協議會員	186			
<input type="checkbox"/>	13	辜顯榮	大正十年	1921	各種委員會 臺灣總督府評議會	評議會員	73			
<input type="checkbox"/>	14	辜顯榮	大正十一年	1922	各種委員會 臺灣總督府評議會	評議會員	69			

查詢後分類

最多僅提供 50 筆

◎ 日本紀年

- ◎ 昭和五年 (3筆)
- ◎ 大正十一年 (2筆)
- ◎ 大正十二年 (2筆)

OPEN/CLOSE

◎ 官職名

- ◎ 參事 (12筆)
- ◎ 評議會員 (12筆)
- ◎ 會員 (10筆)

OPEN/CLOSE

◎ 單位名稱

- ◎ 各種委員會 臺灣總督府評議會 (12筆)
- ◎ 臺中廳庶務課 (11筆)
- ◎ 各種委員會 建功神社祭祀者資格審查會

中華民國政府官職資料庫

1911~ now



15,708 digital image items

85 million records

中華民國政府官職資料庫

任意詞

考試院秘書 經濟部 新聞局

姓名

馬英九 唐飛 蕭萬長 雷震

命令日期 年 月 日 到 年 月 日

出刊日期 年 月 日 到 年 月 日

公報期數 至

公報來源 說明

送出查詢

版權聲明：本資料庫之結構化欄位資料，為國立政治大學所擁有，未經授權，不得擅自複製使用

00066880



name:蔣中正 檢索 簡易查詢 進階查詢 查詢歷史



name:蔣中正的查詢結果

顯示1~20筆 / 共有61筆

姓名	命令日期	任/免	原因	機關 官/職等 職務	原稿
蔣中正	012-03-17	任	特任	大本營參謀長	海陸軍大元帥大本營No.3
蔣中正	012-04-10	免	呈請辭職	大本營參謀長	海陸軍大元帥大本營No.7
蔣中正	012-06-17	任	特任	大元帥行營參謀長	海陸軍大元帥大本營No.16
蔣中正	013-03-20	任	派	陸軍軍官學校入學試驗委員長	海陸軍大元帥大本營No.9
蔣中正	013-05-03	任	特任	陸軍軍官學校校長	海陸軍大元帥大本營No.13

Household Register under Japanese ruled (1906-1945)

Marriage

The date of birth

address

parents

The order in family

第008/册 第00108頁
本戶計 7 頁

第 0232 頁
浮籤最多 / 張

嘉義廳打猫南堡打猫街番地不詳
台烟户

嘉義廳他里霧堡蘇園庄

大正四年四月十七日陽曆

福 福

天 二 天 二

母伯 太 主 戶

天傷四 拾貳 拾貳

林 氏

太伯呂瑞喜

高 氏

水生

明治拾貳年七月七日

呂 氏

水生

明治廿年七月一日

母生 母死 付產 相續

現住所

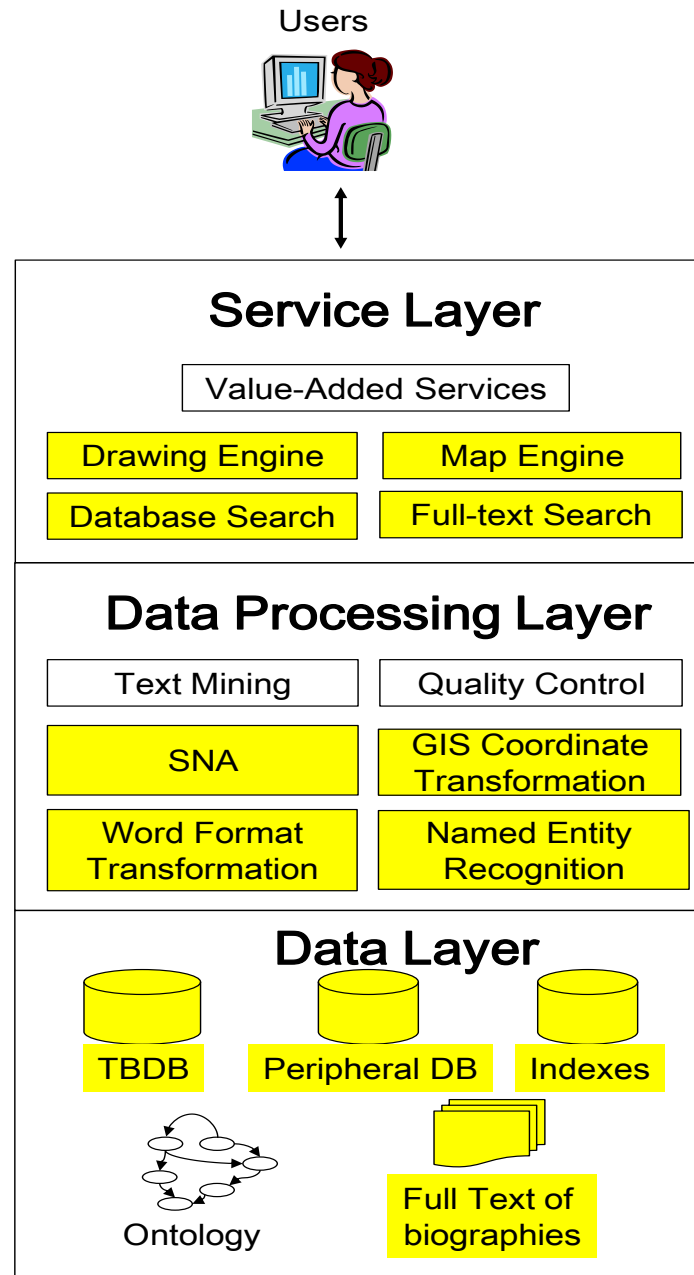
甲 號

Other resources

- * 《府報》 《臺灣總督府公文類纂》
- * 《臺灣總督府專賣局檔案》
- * 《臺灣總督府職員錄》
- * 《臺灣日日新報》 《臺灣士商名鑑》
- * 《臺灣商工人名錄》
- 《臺灣實業名鑑》 《臺灣名人傳》
- 《臺灣風雲人物》 《臺灣時人誌》
- 《臺灣醫師名鑑》

System Architecture of TBDB

System Architecture of TBDB



System Layers

Data Layer

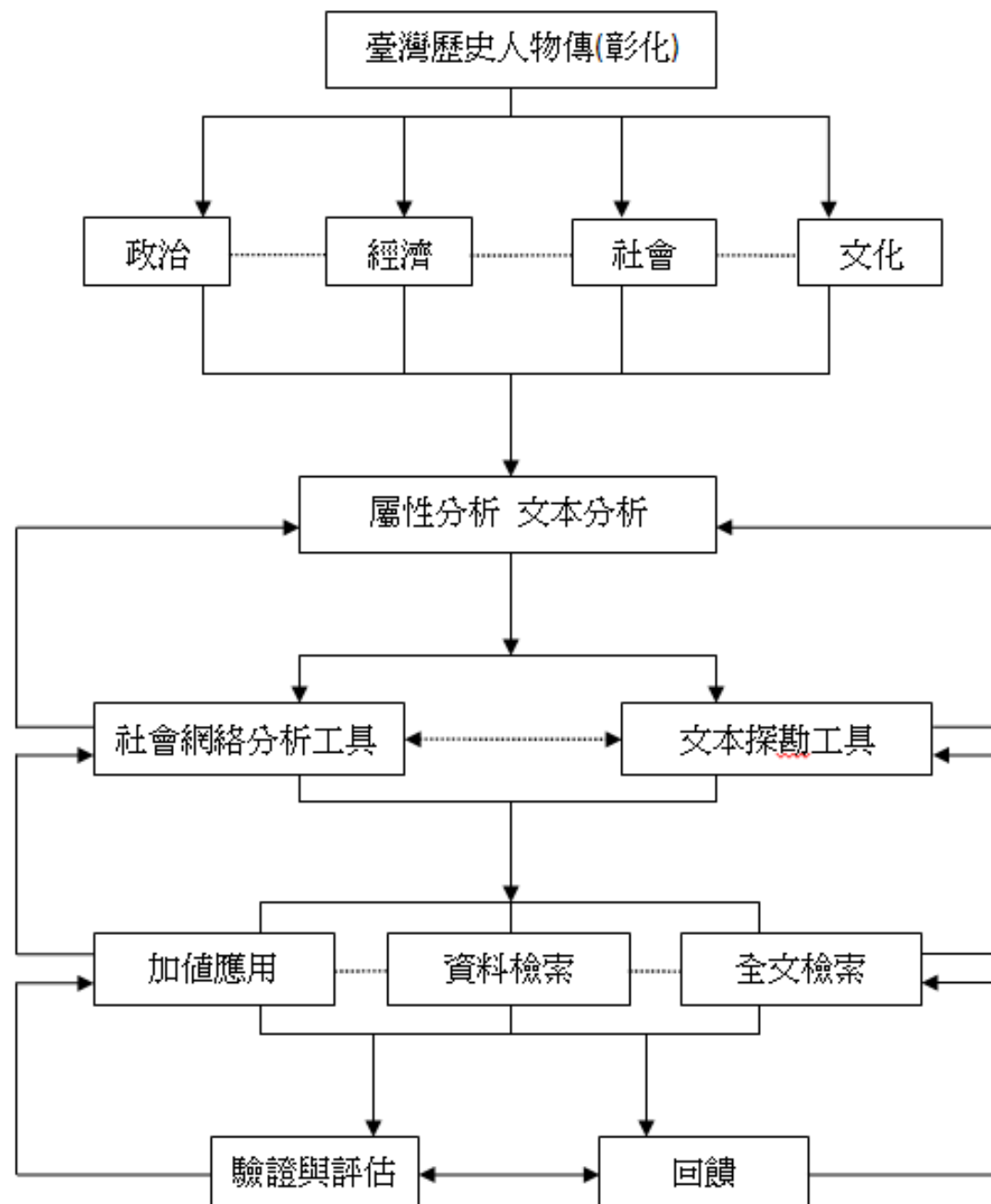
- MySQL to store information on figures
 - Lucene for text processing and full-text search
 - Peripheral DB: tables for authority control
-

Data Processing Layer

- Word format transformation, named entity recognition, social network analysis (SNA), GIS coordinate transformation, text mining, quality control

Service Layer

- database search, full-text search, drawing engine, map engine, and value-added services



TBDB系統建置流程圖

Preliminary results

<http://tbdb.ntnu.edu.tw/search.jsp>

Home Page and Search Results



類別

經濟人物 (19)
文化人物 (18)
政治人物 (18)
社會人物 (13)

活動地

彰化 (48)
臺中 (31)
日本 (22)
員林 (22)
臺北 (21)
鹿港 (15)
埔心 (15)
北斗 (12)
臺南 (10)
東京 (8)

類別

詩社

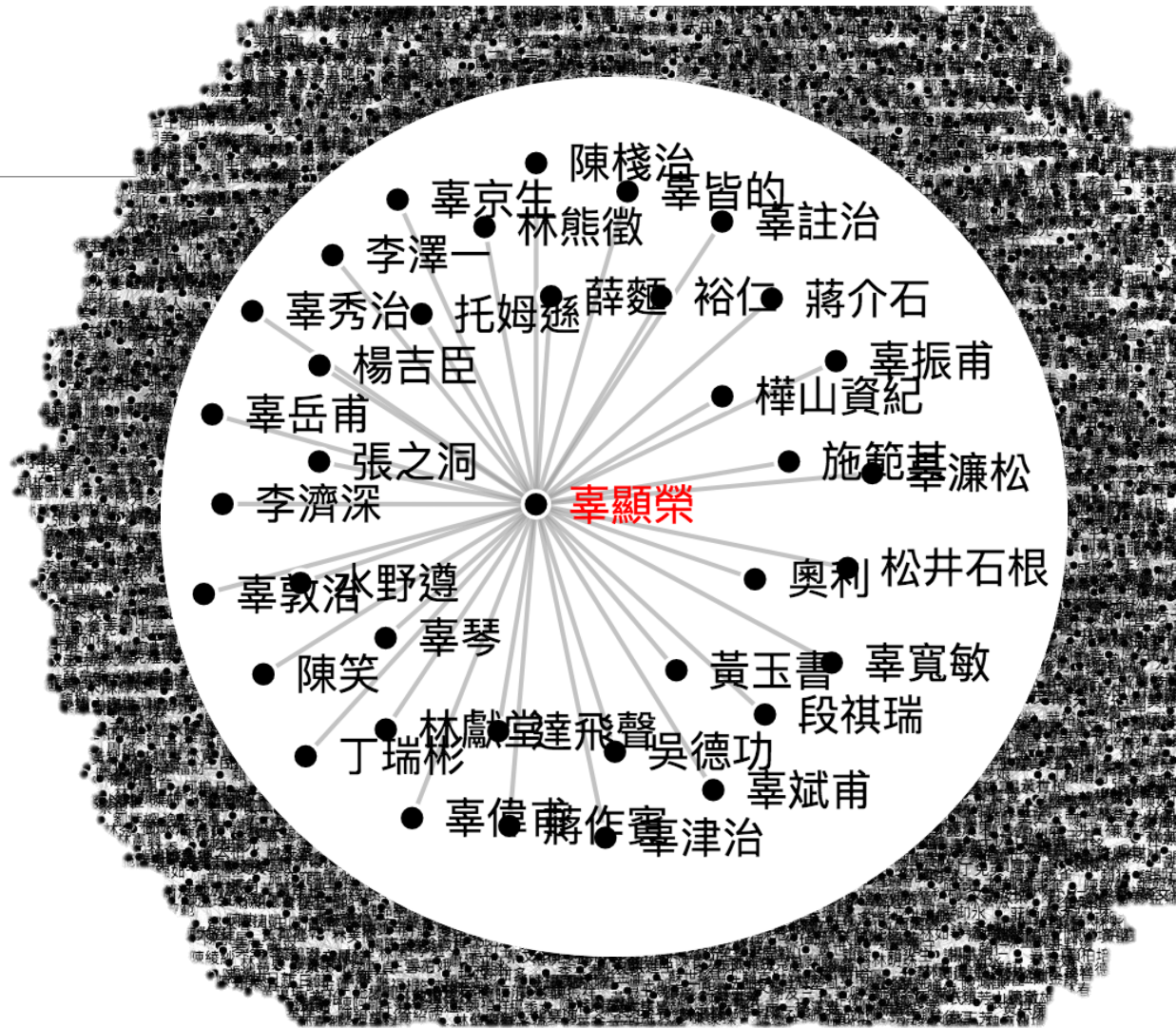
崇文社 (2)
興賢吟社 (2)
櫟社 (1)
蘭社 (1)
詩社

總筆數：68筆

姓名	出生地	類別	生卒年	查看
黃天	溪州	政治人物	1908-1950	檢視傳文
黃旭	花壇	經濟人物	1873-1952	檢視傳文
黃服	線西	經濟人物	1908-1978	檢視傳文
黃秋	鹿港	經濟人物	1887-1959	檢視傳文
黃龍	竹塘	經濟人物	1865-1932	檢視傳文
黃周	和美	文化人物	1899-1957	檢視傳文
黃圖	田尾	文化人物	1902-1976	檢視傳文
黃況	線西	社會人物	1895-1979	檢視傳文
黃石柱	花壇	政治人物	1922-1980	檢視傳文

辜顯榮

• SNA(beta)

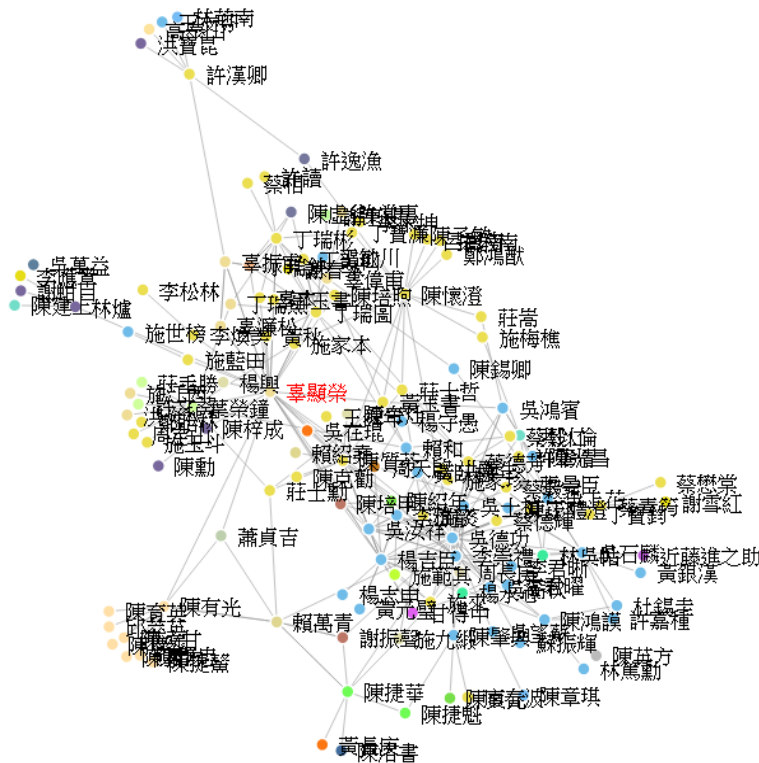


SNA – 另有傳

另有傳

- Someone was mentioned in another biography

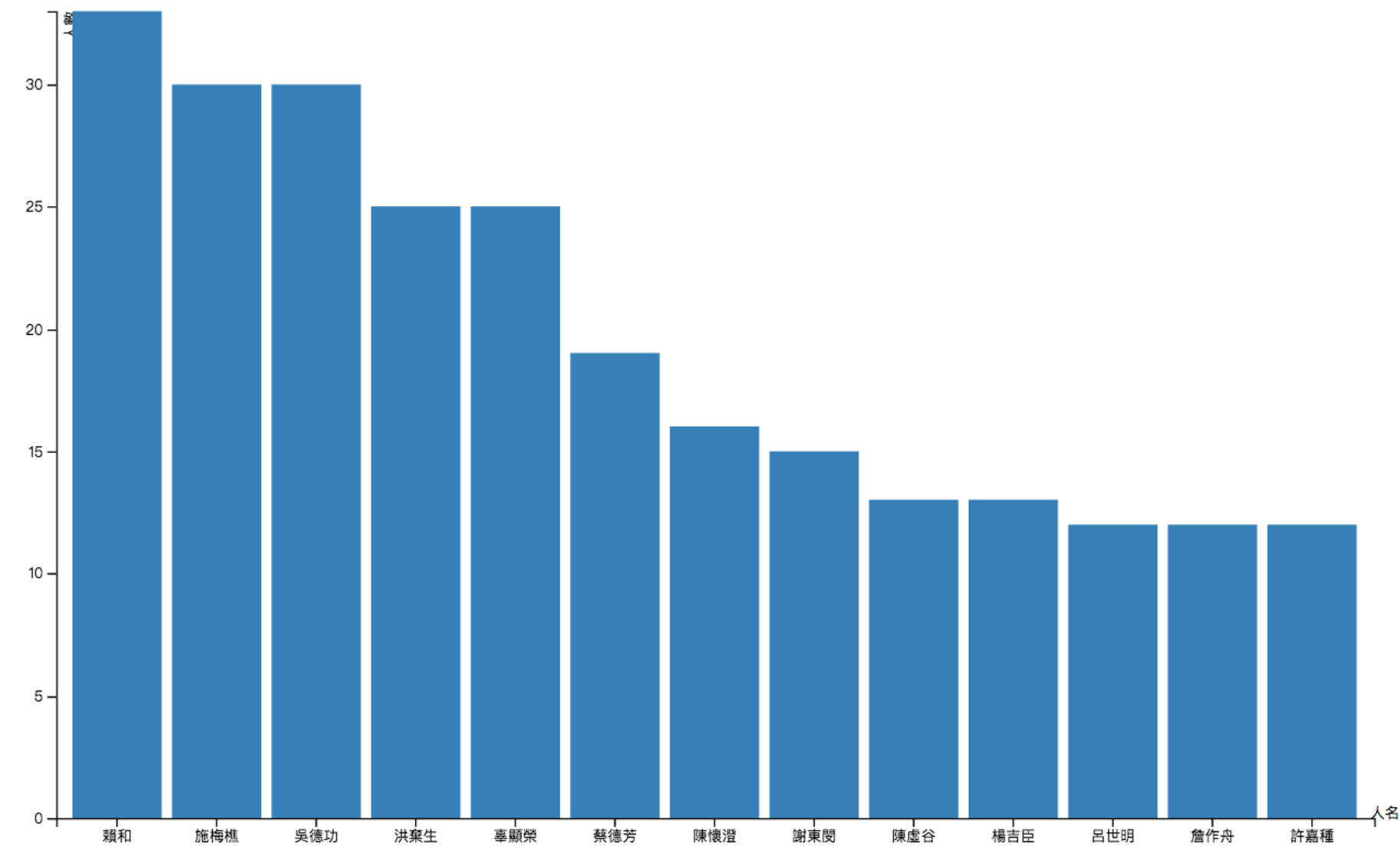
Help identify the relationship between them



Ranking of 「另有傳」

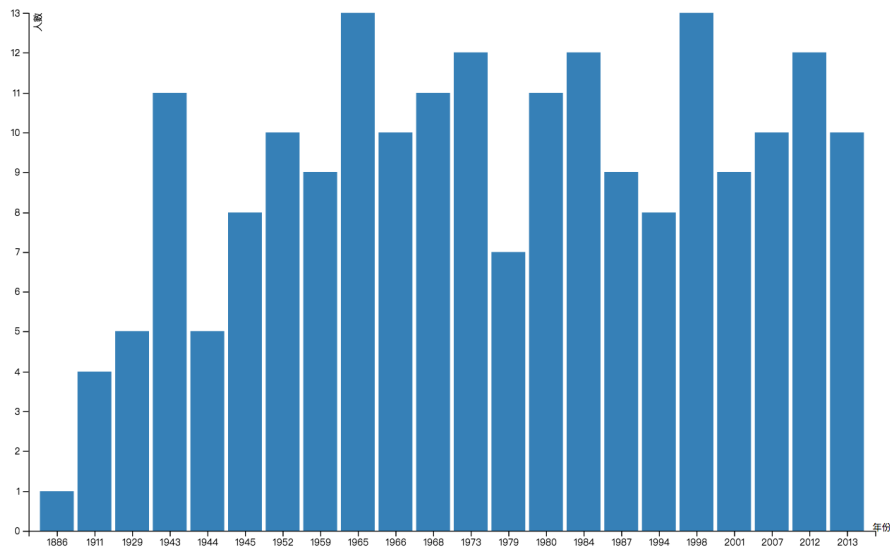
限制數量

查詢

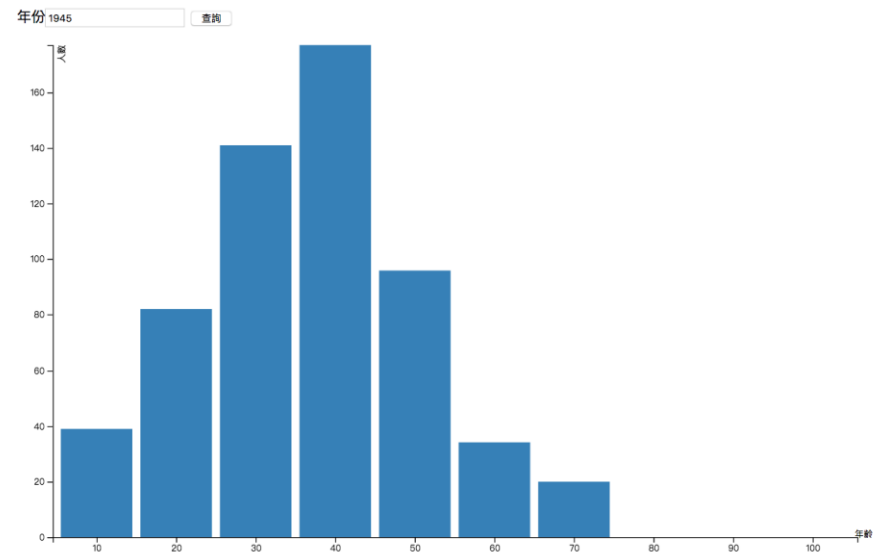


Life span

Distribution of death

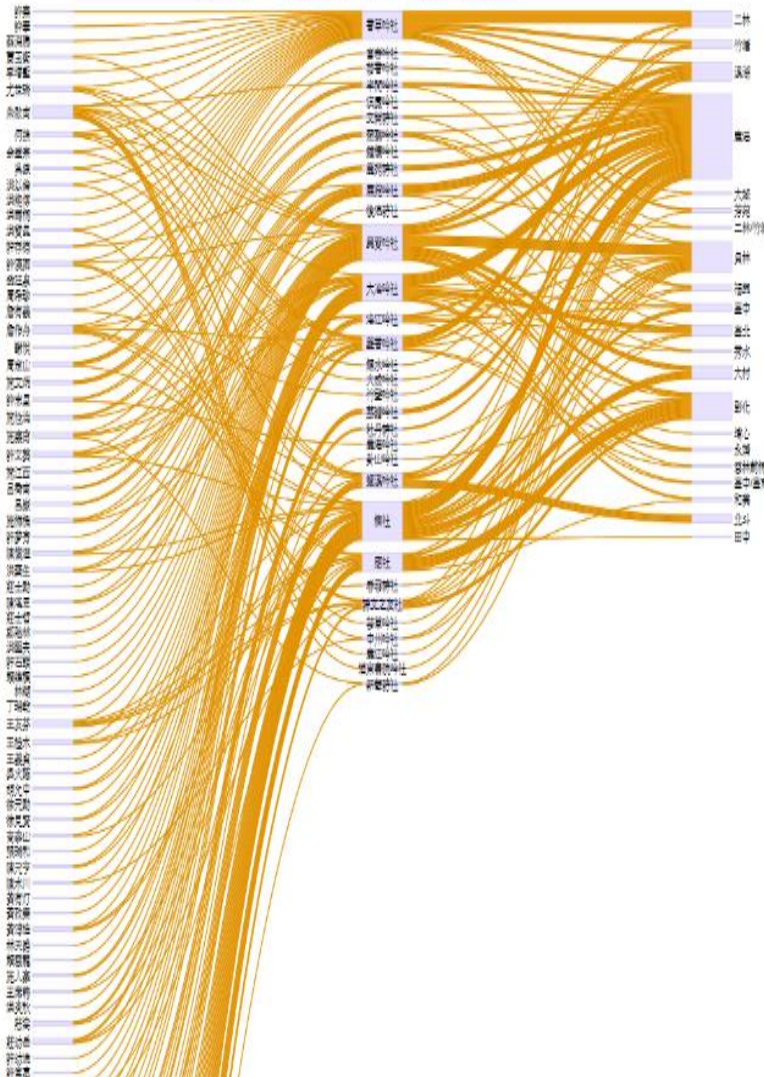


Distribution of birth



SNA – Poetry Societies, Places, and Personages

詩社、地點及參與者關係(III)

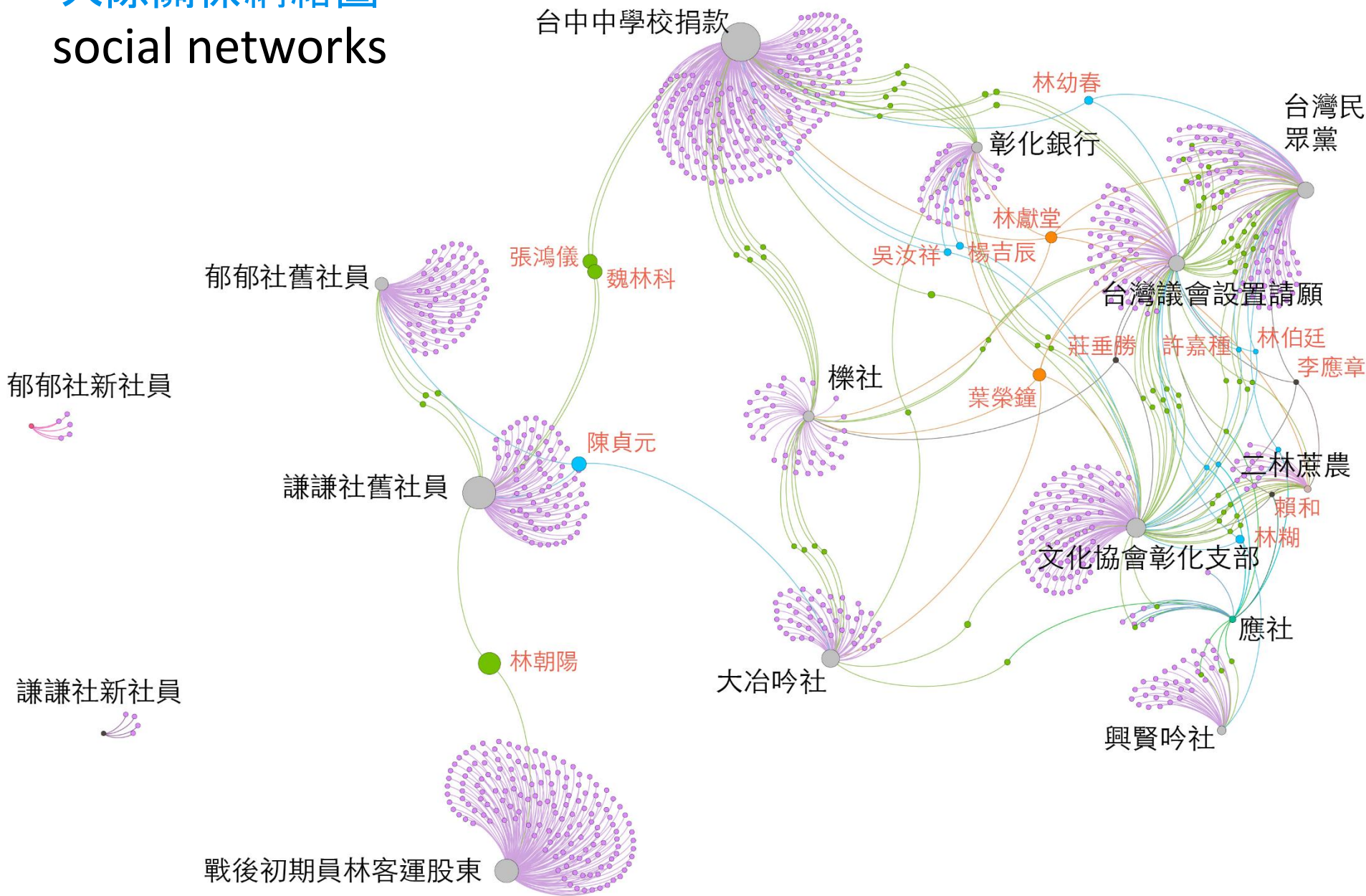


Highlights the relationship between poetry societies, places, and personages participating the poetry societies

Help identify the geographical distribution of the members of a poetry society

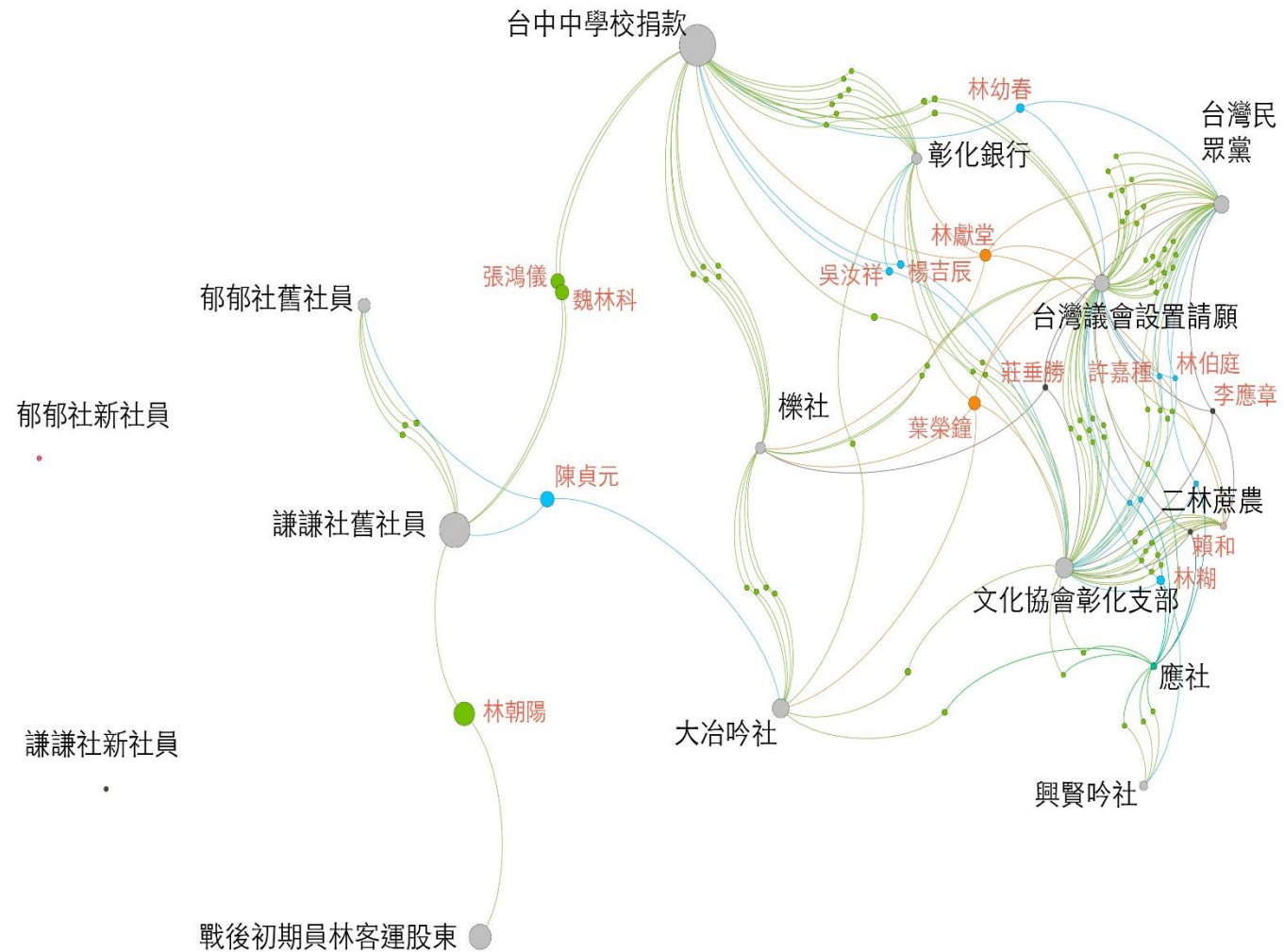
人際關係網絡圖

social networks



social networks

poetry club
economy
politics



Flow of figures

Space Analysis

GIS & Digital tools

- space material in biography
birth place / career's place
- tool : GIS

Biography vs. prosopography



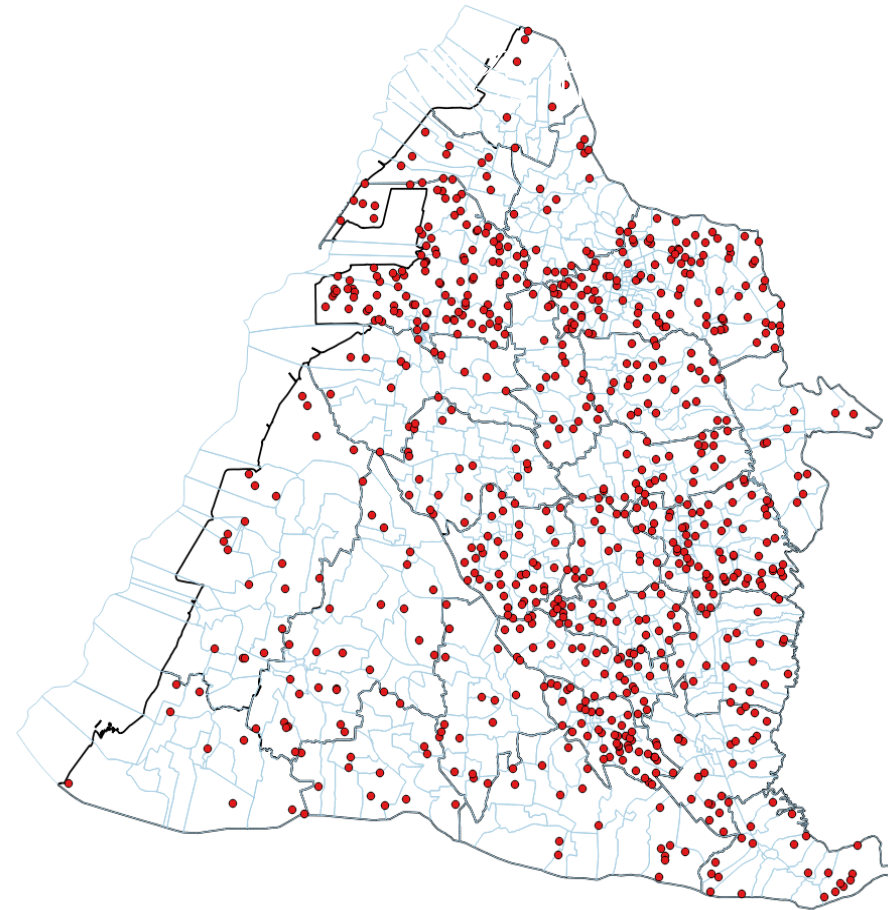
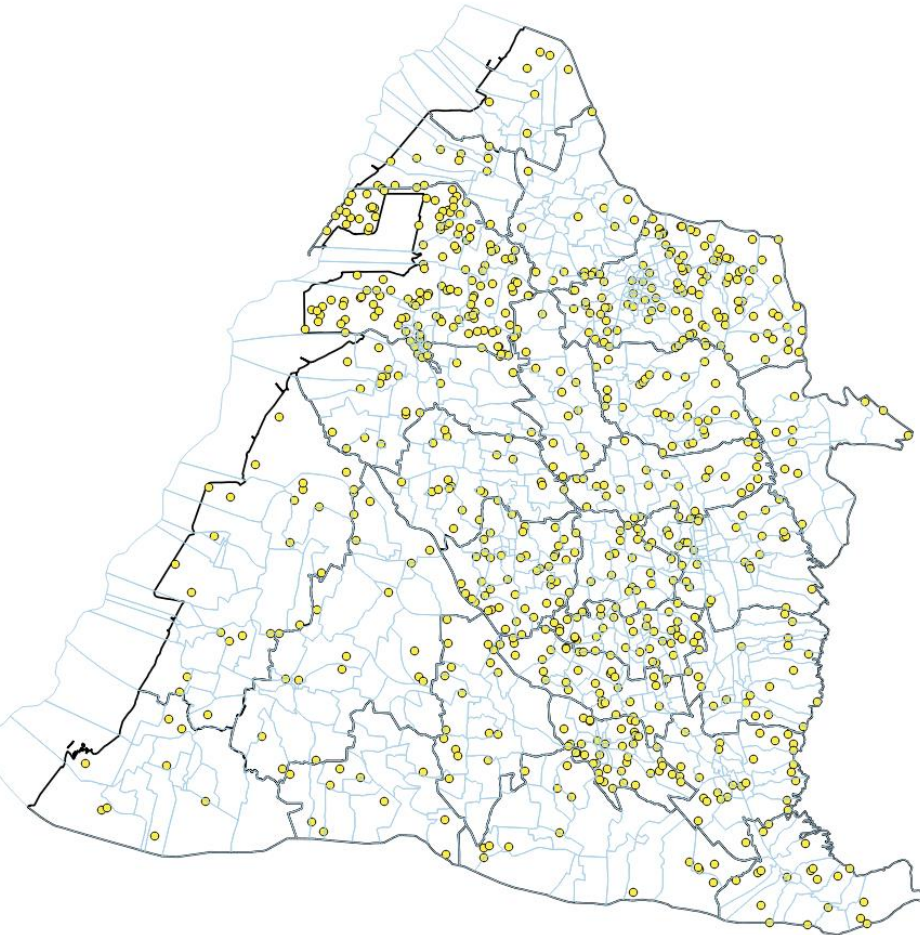
GIS

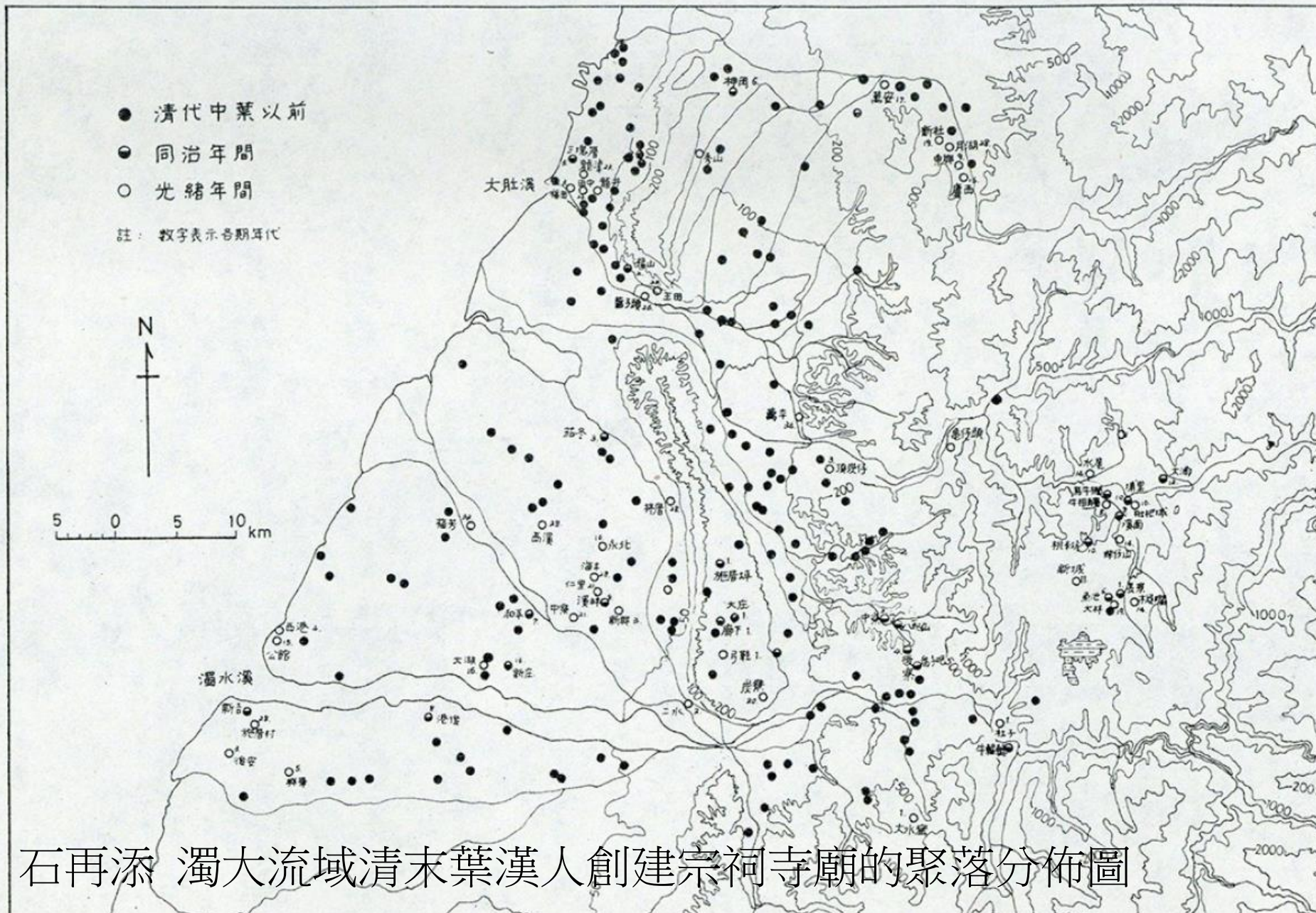
→ Spatial characteristics

Geographical distribution of figures

birth place

career's place





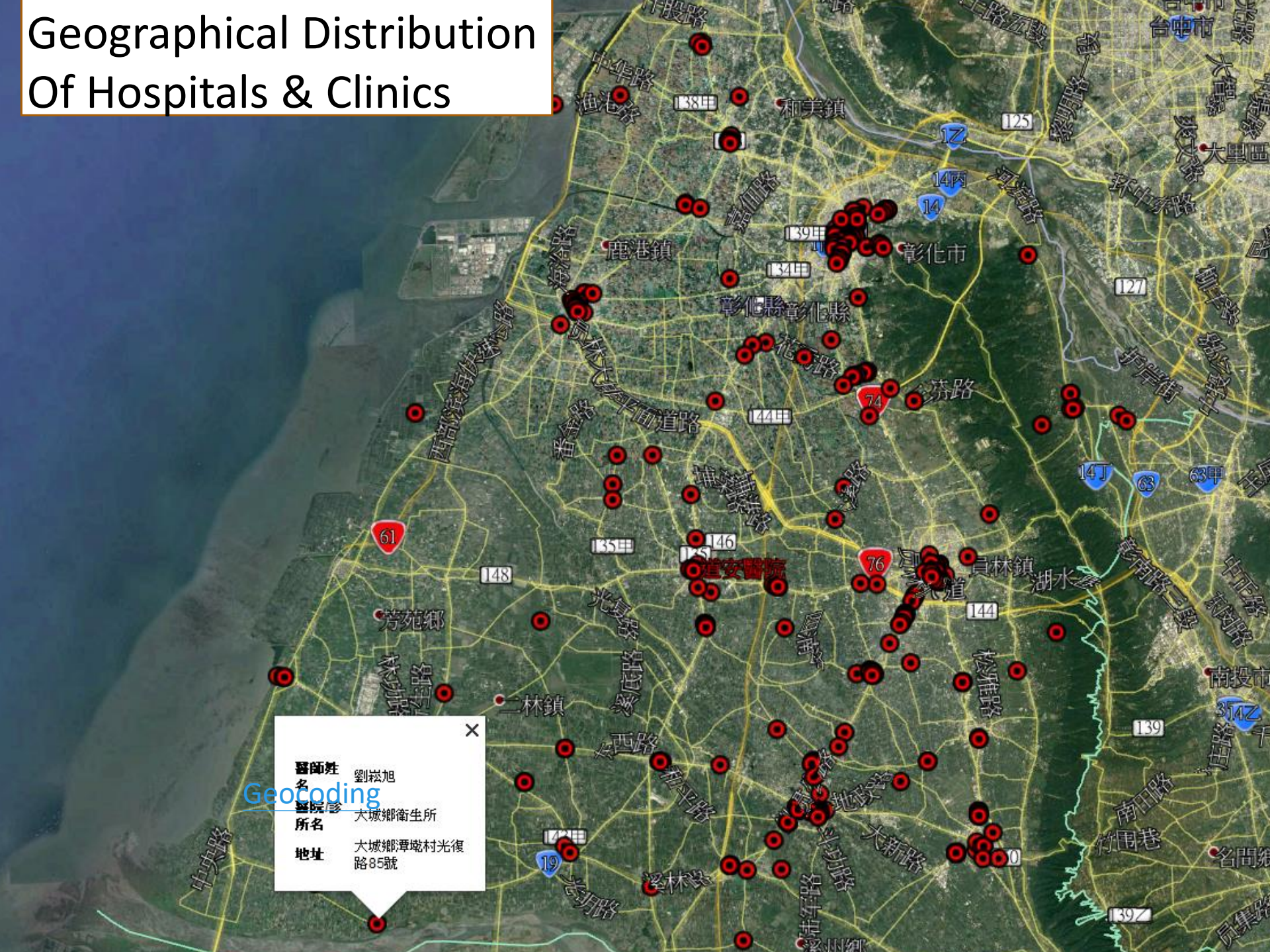
石再添 濁大流域清末葉漢人創建宗祠寺廟的聚落分佈圖

Geographical distribution of temples and villages of late 19th century
(investigated in 1972-1976)

Geographical Distribution Of Hospitals & Clinics

Geocoding

醫師姓名 劉崧旭
醫院診所名 大城鄉衛生所
地址 大城鄉潭墘村光復路85號

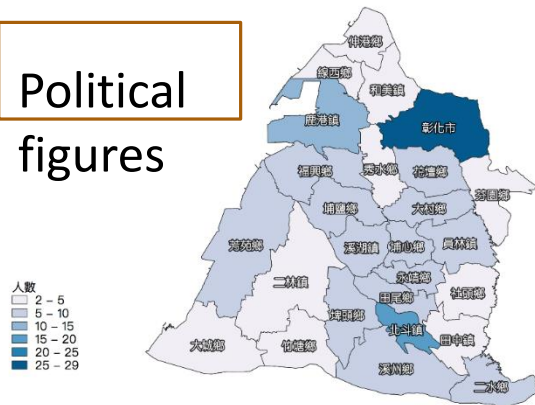


Geographical Distribution

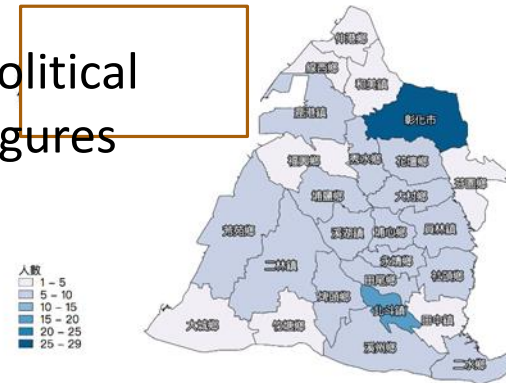
birth place

career's place

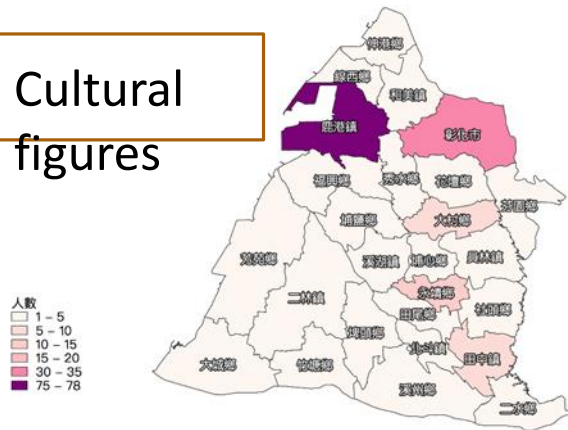
Political
figures



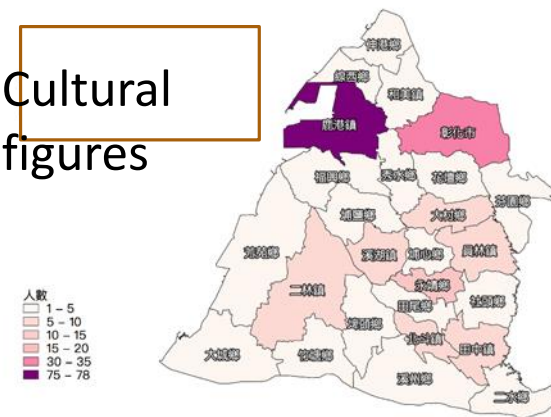
Political
figures



Cultural
figures



Cultural
figures



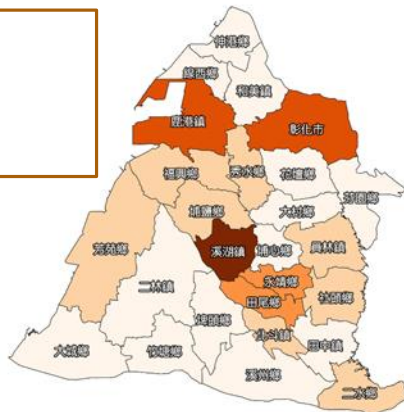
Geographical Distribution

birth place

career's place

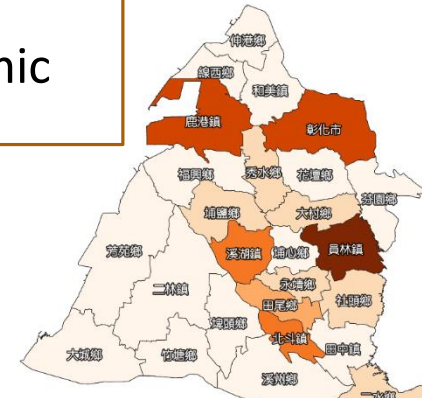
Economic figures

人數
1 - 5
5 - 10
10 - 15
15 - 20
20 - 22



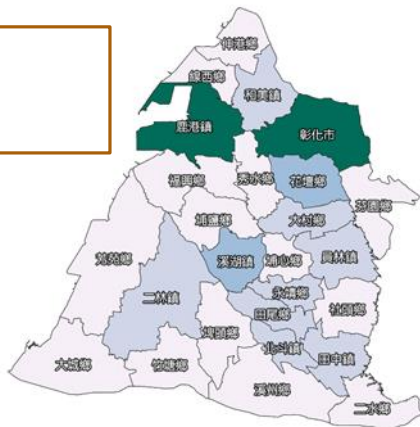
Economic figures

人數
1 - 5
5 - 10
10 - 15
15 - 20
20 - 25
25 - 28



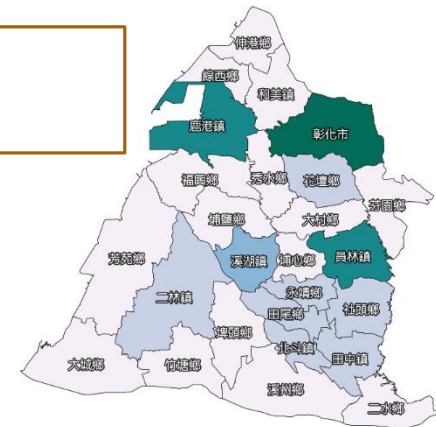
Social figures

人數
0 - 5
5 - 10
10 - 15
15 - 20
20 - 25
25 - 30
30 - 33

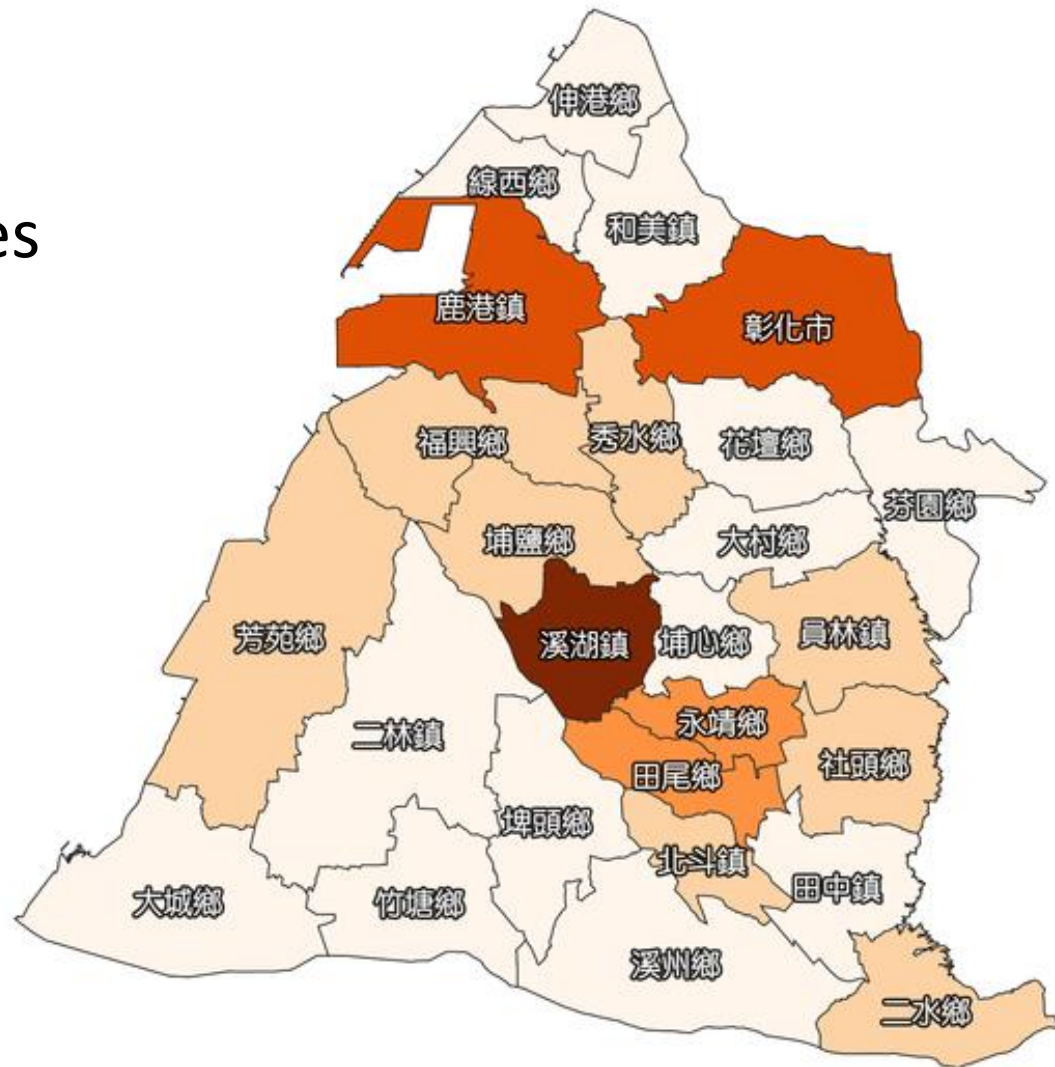


Social figures

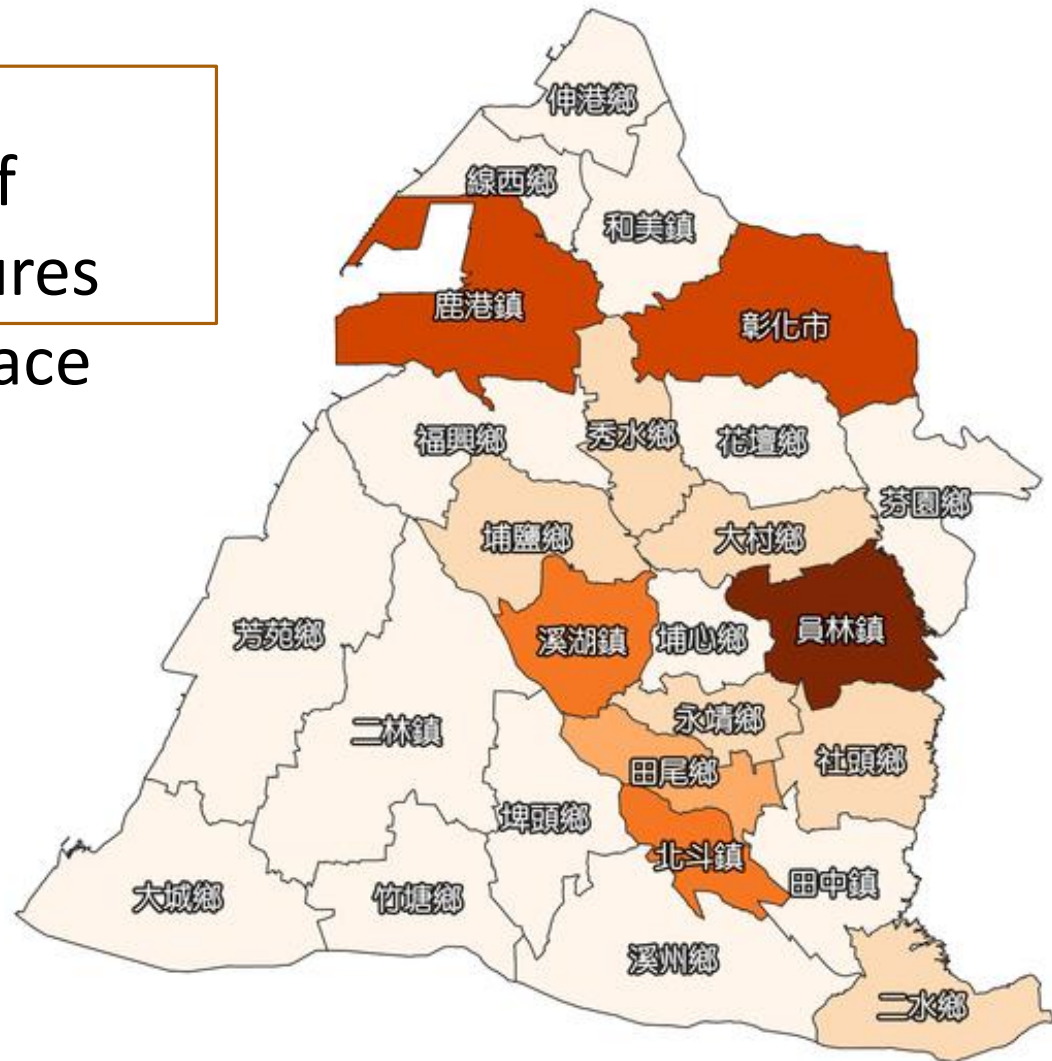
人數
1 - 5
5 - 10
10 - 15
15 - 20
20 - 25
25 - 28



Geographical distribution of Economic figures By birth place



Geographical
distribution of
Economic figures
By career's place



Conclusions

Lesson Learned

1. Discussion with an open mind is necessary for mutual understanding between different disciplines
2. Fanciness is not the most important factor of a good tool, but the capability to support the research of humanists is
3. Only when humanists realized what purpose of the tools , what problems need to be **solved** , databases and digital-humanities tools can exert their largest effect

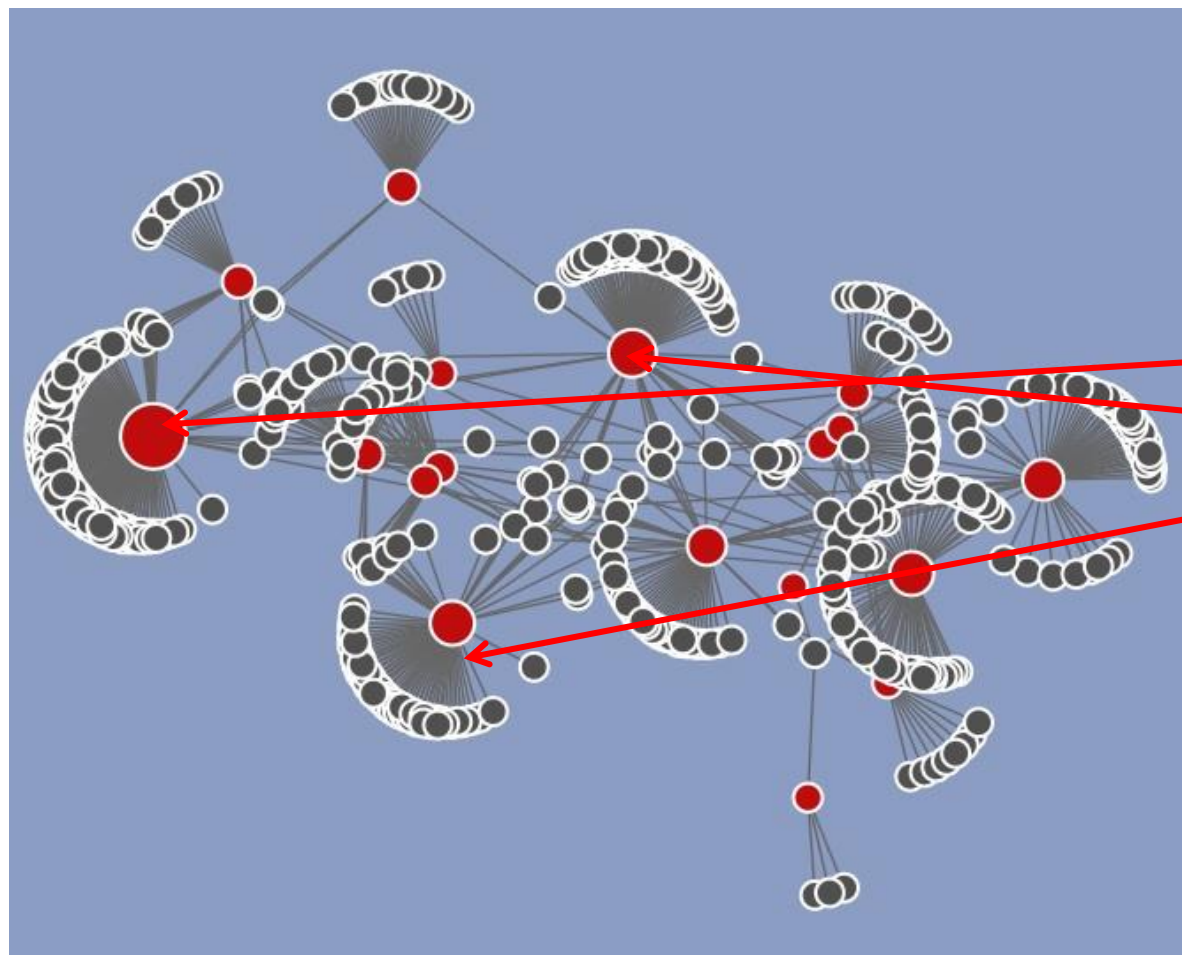
The prospect of TBDB

1. Improve the quality of tools

2. Expand the content of the database
 - biographies of other gazetteers
 - dictionaries of historical figures
 - diaries available
3. Apply face recognition technology to old photographs
4. provide API to connect with other platforms

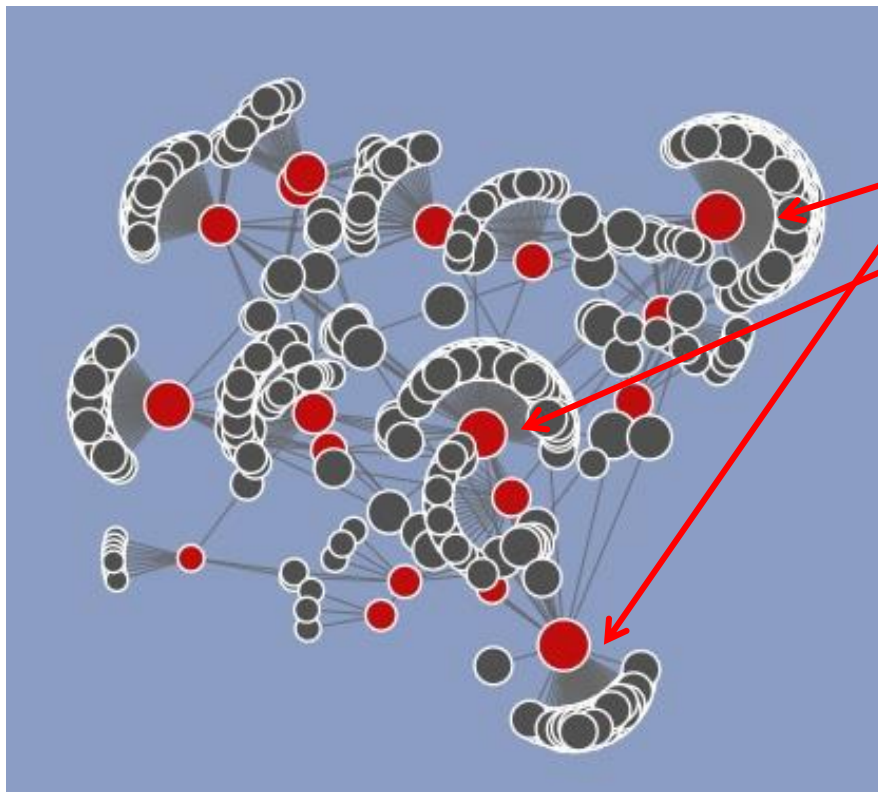
彰化詩社「度中心性」

櫟社、香草吟社、文開詩社



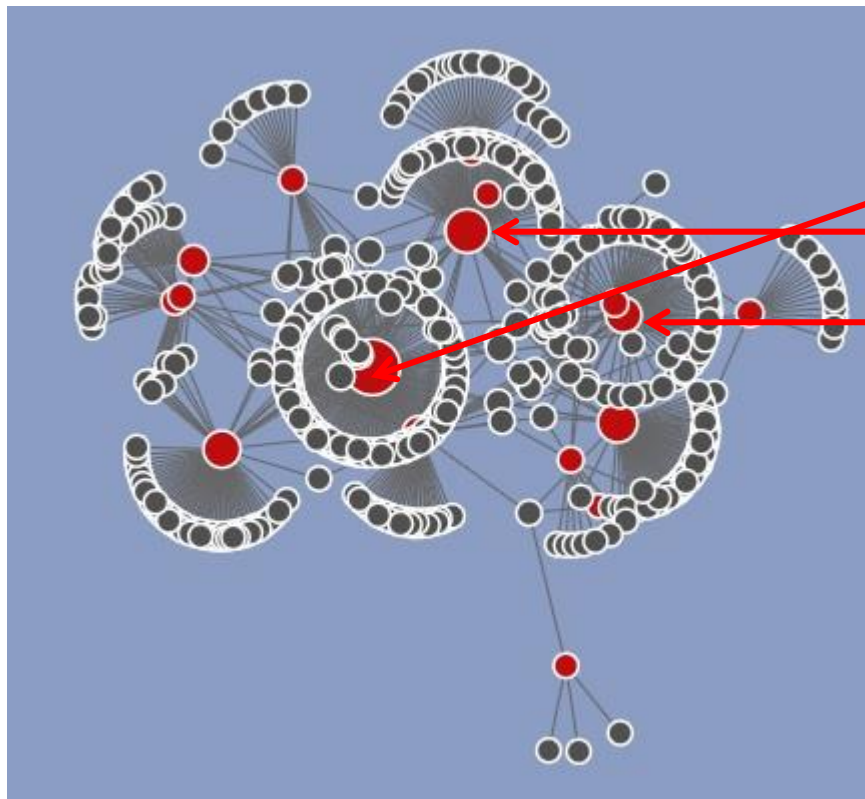
name	degree
櫟社	64.00
香草吟社	17.53
文開詩社	12.12
詩文之友社	11.55
翼賢吟社	7.48
彰化縣詩學研究協會	5.39
洛江吟社	1.87
蘭社	1.25
大冷吟社	1.07
半閒吟社	0.83

彰化詩社的「接近性」



name	closeness
詩文之友社	64.00
櫟社	57.87
香草吟社	53.33
吳燕生	50.92
興賢吟社	47.37
周定山	43.44
朱啟南	40.50
洪以倫	38.18
半閒吟社	36.60
陳逢源	36.53

彰化詩社的「中間性」



name	betweenness
標社	64.00
香草吟社	24.09
詩文之友社	15.27
文開詩社	10.39
興賢吟社	8.68
彰化縣詩學研究協會	8.12
洛江吟社	2.58
周定山	1.83
蘭社	1.56
吳燕生	1.37

Face Recognition Technology



Harvard Library.

**4 Aureomycin Research Group at
Harlem Hospital, New York, NY, 1949.
Louis Tompkins Wright papers, 1879,
1898, 1909-1997 (H MS c56). Center
for the History of Medicine, Francis
A. Countway Library of Medicine,
Harvard Medical School.**



THE ARCHITECTURAL SOCIETY

3 Office of Horace Trumbauer, undated. Julian Abele, Chief Designer, fifth from left, in back row. Image courtesy of the David M. Rubenstein Rare Book & Manuscript Library, Duke University.

4 Architectural Society, **Julian Abele seated in center.** *The Record* (Philadelphia: University of Pennsylvania, 1902), page 112. Image courtesy of the University Archives and Records Center, University of Pennsylvania. | The first African American graduate of the University of Pennsylvania's Department of Architecture, Julian Abele was elected to the prestigious position of President of the Architectural Society in his senior year.

5 Archibald Cary Coolidge to Horace Trumbauer, July 23, 1912. Coolidge

周秀娥(1911-2007)

彰化人

彰化女高

東京女子醫學專門學校

夫周淵昌 眼科醫師

周眼科婦產科診所

馬偕醫院第一任眼科主任

子周神彬為眼科醫師



周秀娥與高慈美等留日學生合影



中研院臺史所檔案館提供



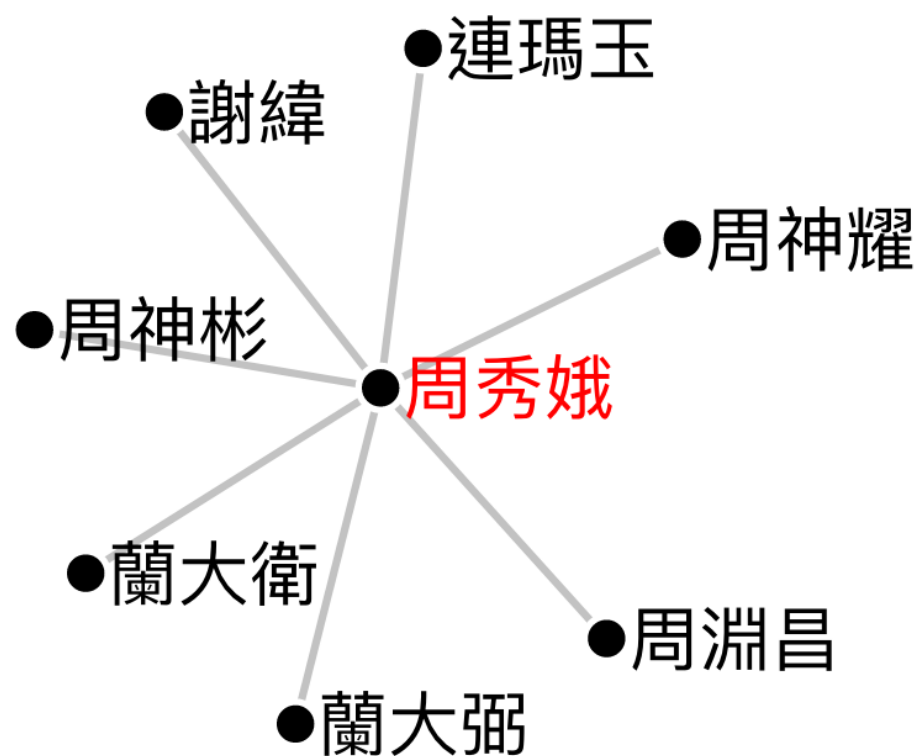
彰化基督教青年會クリスマス祝賀會
1 9 3 6 . 1 . 3

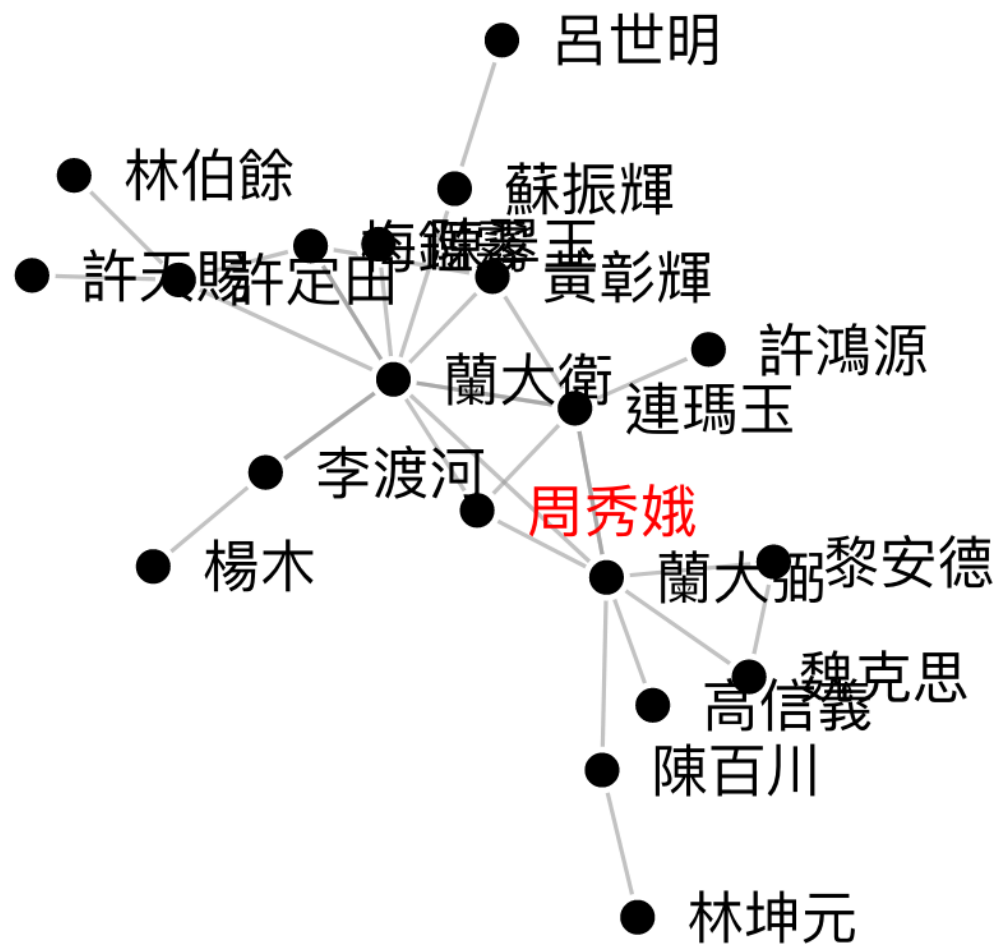
彰化市醫師會創立紀念(1937.8)



賴和紀念館提供

周秀娥傳文提及人物





周秀娥另有傳人物(1層)

Apply face recognition technology to old photographs



1937年，彰化市醫院會館前
連錦玉、李允中、林錦雲、高慈
彰化基督教醫院院史館典藏

Another possibility for
SNA?
社會網絡分析的另一個可能?

Thank you for listening