

國立臺灣大學電機資訊學院資訊工程研究所

碩士論文

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

使用者取向之歷史地理資訊系統—古契書與統計資料呈現

User-Oriented HGIS on Old Land Deeds and Statistical Data



Chung-Hsiang Ou

指導教授：項潔 教授

Advisor: Jieh Hsiang, Professor

中華民國 100 年 6 月

June, 2011

誌謝

歲月的巨輪太快又太深，兩年碩士生活一轉眼就過去了，才剛剛熟悉這個環境，即將就要劃下句點。兩年前決定轉換環境選擇到台大就讀，並毫不猶豫選擇了項潔老師的實驗室，非常的幸運！感謝項老師的指導，老師有著清楚的邏輯與對研究的熱情，在這橫跨歷史與資訊的學門中帶領我如何釐清問題與思考問題，迅速的看出問題癥結所在，兩年下來獲益良多。

本論文的完成首先感謝數位典藏中心的蔡炯民學長，炯民學長花了很多時間陪我討論論文的走向、整理資料與修改論文，非常感謝學長的協助；杜協昌博士的見解非常獨到，常常可以看出我論文上的盲點，聽杜博士一席話勝讀十年書呀！；感謝豐恩，可以馬上看出我的論點哪裡有缺失，耐心的陪我討論並修正；謝謝稷安，由使用者的角度給了我很多修改系統的意見，在論文徬徨找方向時，陪我討論了很多可行方案，翻了許多的書；謝謝農堯在 GIS 技術上提供了我好多解決辦法，節省了很多自己摸索的時間；謝謝詩沛學姐與浩洋學長，時常關心我的論文狀況，適時補充論文可以加強的地方；謝謝 cobain 學長，幫忙整理好 Turing 的事務，讓它順利移到資工機房代管；宋浩有不錯的 Computer Science (CS) 底子，可以跟我討論許多技術上的問題，特別是在網頁技術上；感謝最常陪我在實驗室「認真」做研究小黑，常常提供我不錯的意見也幫了我很多生活上的忙，並且讓實驗室無聊的夜晚非常精彩；感謝線線家族的柏淳與士剛帶來許多歡樂？也都是健身房的好夥伴；感謝恐龍在一起修課的時候，幫了我許多忙，將 Project 順利完成；感謝坐在我後面的光哲，常會跟我分享新奇事；感謝鈺淳在我煩躁的時候還會跟我喇賽；感謝嘉文幫忙我接手管理了很多實驗室的機器；感謝數典中心的逸群、弼達、屹靈的協助；感謝 Lab303 的大家，有了你們讓我碩士兩年十分難忘！

最後，謹將本論文獻給我最愛的家人，感謝爸媽在背後無條件的支持我攻讀碩士班！感謝所有我愛的人以及愛我的人給我的支持與鼓勵，將此榮耀與喜悅分享給大家！

歐仲翔 民國一〇〇年八月

中文摘要

歷史資料在經過數位化的整理後，可以結合歷史地圖，並運用地理資訊系統，以空間的方式呈現出人文、社會、經濟等的史料分佈，提供歷史學家一個觀察的環境，更快掌握到歷史發展的脈絡，為人文研究帶來很大的便利。

然而，地理資訊系統雖是一項便利的研究工具，卻有一定程度的技術門檻，往往令歷史研究者望之怯步，使得人文研究和地理資訊系統之間，無法發揮相輔相成的效果。因此，如何能打造一個更親和、更直覺的地理資訊系統，令使用者能因應研究需要，自由地上載資料、進行觀察，便成為本研究所希望達成的目標。

本研究以西元 1904 的年臺灣堡圖為基礎，將地圖中大量的行政區域資訊進行向量化處理。透過 Web GIS 的方式，結合臺灣堡圖與日治時期的史料，提供輔助臺灣歷史研究的工具。本研究提出兩個 Web GIS 的設計，一為建置以地理為基礎的古契書資料庫，內容為 1898 年臺灣總督府展開地及整理時所抄錄保存的民間契約文書，其收錄時期約和臺灣堡圖建立時間相當，因此這些抄錄文書所使用的地名與臺灣堡圖中的地名可作對應，利用資訊技術擷取出其中的地理資訊。本論文由 15899 件總督府抄錄契約文書當中找出 12502 件的地理位置，以 Web GIS 工具呈現於堡圖的行政區域上，可以與當代通用的地圖(Google Map)套疊，讓研究者能以地方名稱或地理範圍進行資料搜尋與觀察。

另一設計則是以臺灣堡圖行政區域圖作為繪製主題圖的底圖，使用者可將自己的研究成果，整理成 Excel 格式的表格資料後上傳，透過我們所設計的工具，依臺灣堡圖的不同行政階層，如州廳、堡里、街庄等，繪製出可以呈現各區域的數量分佈等主題圖。以往的 Web GIS 工具，由於歷史地圖資料量龐大，在使用者端往往需要大量的時間來運算與呈現，更無法輕易地進行不同尺度轉換，此工具利用 HTML5 撰寫程式，使得地圖顯示與縮放更為效率化，節省大量的時間。

希望能以相對簡便的方式，讓更多歷史研究者能夠充分地運用地理資訊系統的強大功能，找出單純文本不容易看出的脈絡，開啟歷史研究一個新的面向。

關鍵字：HGIS、GIS、臺灣堡圖、古契書、日治時期、主題圖

ABSTRACT

Through integrating historical maps and historical data, the Geographic Information System (GIS) can help historian observe and track historical phenomena that are not easily found from historical data alone. However, for scholars without much knowledge in geography, mastering GIS can be a daunting task. There is, therefore, a barrier between humanities research and GIS. Hence, our target is how to build a user-friendly and intuitive GIS tool for scholars to use.

In this thesis, we present two WebGIS systems. The first one built on a database of land deeds. Land deeds are important primary documents that are used by historians in the study of land and social development Qing era Taiwan. We start by identifying the exact latitude and longitude of the land deeds, upon which a retrieval system is built. A user can find land deeds either through typing in words or drawing an area. The deeds retrieved are then presented on the map. Several layers of maps, of different nature and time periods, are also incorporated. The second system presented uses an administrative area map to provide a convenient tool for researchers to generate theme maps. By uploading statistical data in Excel format, the system will automatically generate a map with the data in different regions presented in different color according to density. The system is a very simple visual aid for the user to observe geographical data.

Keywords: HGIS, GIS, The Taiwan Baotu (臺灣堡圖), Land Deeds, Japanese-ruled Period, Statistical Theme Map

CONTENTS

口試委員會審定書	#
誌謝	i
中文摘要	ii
ABSTRACT	iii
CONTENTS	iv
LIST OF FIGURES	vii
LIST OF TABLES	ix
Chapter 1 緒論	1
1.1 研究背景	1
1.2 研究動機	1
1.3 研究回顧	3
1.3.1 GBHGIS	4
1.3.2 The Harvard World Map Project	4
1.3.3 Historical GIS for Germany	5
1.3.4 臺灣歷史文化地圖	5
1.4 論文架構	5
Chapter 2 使用之史料介紹	7
2.1 臺灣總督府抄錄契約文書	7
2.2 臺灣堡圖	9
2.3 日治統計資料庫	11
Chapter 3 現有之地理資訊系統技術回顧	12
3.1 GIS 簡介	12
3.2 GIS 資料格式與標準	12
3.2.1 Shapefile	12
3.2.2 KML	12
3.2.3 GeoJSON	12
3.3 常用 GIS 軟體與工具	13
3.3.1 個人電腦 GIS 軟體	13
3.3.2 伺服器 GIS 軟體	13

3.3.3	網頁地圖服務.....	13
Chapter 4	總督府抄錄契約文書 GIS 工具.....	15
4.1	系統概述.....	15
4.1.1	系統架構.....	16
4.2	對應契書地號流程.....	17
4.3	系統功能.....	21
4.3.1	圈選檢索.....	22
4.3.2	條件過濾與時間軸.....	23
4.3.3	地圖圖層.....	24
4.3.4	檢索呈現.....	25
4.4	操作實例與討論.....	27
4.4.1	金廣福墾隘.....	27
Chapter 5	歷史行政區主題圖繪製工具.....	31
5.1	系統概述.....	31
5.1.1	行政單元.....	32
5.1.2	系統架構.....	33
5.2	系統功能.....	34
5.2.1	Excel 上傳.....	34
5.2.2	功能選單.....	35
5.2.3	主題圖圖例設計.....	35
5.3	圖資整理.....	36
5.3.1	圖資整理流程.....	36
5.3.2	Douglas-Peucker 演算法.....	36
5.4	操作實例與討論.....	39
5.4.1	工具操作流程.....	39
5.4.2	日治統計資料呈現：犯罪統計.....	39
5.4.3	日治統計資料呈現：現住人結婚離婚及配偶.....	42
Chapter 6	結論與未來工作.....	45
6.1	結論.....	45

6.2 未來工作	46
6.2.1 匯入方式	47
REFERENCE	49
附錄一	51



LIST OF FIGURES

Fig. 2-1	總督府抄錄契約文書年代分佈圖	7
Fig. 2-2	契書地號位於右上角紅色字處	8
Fig. 2-3	臺灣堡圖—臺北城	10
Fig. 2-4	臺灣堡圖—新竹街	10
Fig. 2-5	日治統計資料庫	11
Fig. 4-1	歷史行政區主題圖繪製工具系統架構圖	16
Fig. 4-2	歷史行政區主題圖繪製工具資料流程圖	16
Fig. 4-3	比對地號過濾流程	18
Fig. 4-4	以 ArcGIS 呈現之靜態契書分佈圖	21
Fig. 4-5	判斷契書分佈點是否在多邊形內部	23
Fig. 4-6	圈選出新竹頭前溪流流域觀察契書分佈	23
Fig. 4-7	(a)原本契書分佈 (b)調整縮放比例後，合併契書分佈點	25
Fig. 4-8	契書分佈圖，數字為該地點契書的數量	26
Fig. 4-9	每個點均可開啟觀察該點年代分佈	26
Fig. 4-10	群集合併契書分佈	27
Fig. 4-11	1835 年至 1836 年，金廣福開墾初期	29
Fig. 4-12	1835-1838 年，開始拓展至峨眉鄉	29
Fig. 4-13	1835 年至 1850 年	30
Fig. 4-14	1835 年至 1895 年，可看出金廣福開發範圍	30
Fig. 5-1	歷史行政區主題圖繪製工具系統架構圖	33
Fig. 5-2	歷史行政區主題圖繪製工具資料流程圖	33
Fig. 5-3	Excel 表格格式—單層屬性	34
Fig. 5-4	Excel 表格格式—雙層屬性	35
Fig. 5-5	圖資整理流程	36
Fig. 5-6	使用 Douglas-Peucker 演算法	37
Fig. 5-7	規律的省略座標點，失真嚴重	37
Fig. 5-8	Douglas-Peucker 演算法	38

Fig. 5-9	日治統計資料庫—明治 38 年犯罪統計類之廳別檢舉件數 p2-p3.....	40
Fig. 5-10	將表格轉換成 Microsoft Excel 檔	40
Fig. 5-11	匯入系統中選擇欄位畫圖	41
Fig. 5-12	主題圖呈現 - 賭博案件數分佈(廳界).....	41
Fig. 5-13	主題圖下載 - 竊盜案件數分佈(廳界).....	42
Fig. 5-14	左圖為明治 33 年、右圖為明治 34 年五區之年末配偶數	44
Fig. 5-15	左圖為明治 35 年、右圖為明治 36 年廳界之年末配偶數	44



LIST OF TABLES

Table 2-1	臺灣總督府檔案抄錄契約文書具體來源	8
Table 4-1	臺灣總督府抄錄契約文書類別分佈數量	24
Table 5-1	臺灣堡圖圖資，幾何資訊	31
Table 5-2	臺灣行政區階層沿革	32
Table 5-3	行政區域對應(北區).....	43



Chapter 1 緒論

1.1 研究背景

近年來，數位人文已成為蓄勢待發的重要場域，無論保存或研發，都已經有了長足的進展，許多歷史典籍、檔案也都完成全文數位化的工作，可以直接透過資訊科技挖掘出以往逐字閱讀所不容易觀察到的現象。除了文字的部分外，越來越多的歷史地圖被數位化成電子資料，除了單純掃描地圖影像以外，很多計畫花了大量的人力，將歷史地圖上的圖例、河川、行政區等地形地物之資訊用點、線與面標示出來成為向量化(vector)的數位地圖。向量圖具有精準的點線面資料，可對於每筆幾何地理資料增加描述的欄位；並可疊合多種圖層，例如疊合交通網與行政區的圖層製作出台灣主要道路圖；彈性的編輯地圖所需要的部份等，並可以匯入 GIS 軟體中做更多的計算與資料處理……等等。凡此，皆提昇了歷史地圖的可用性。

目前個人電腦的計算能力已足夠處理複雜的向量圖形資料，地理資訊系統(Geographic Information System, GIS)因而日益普及，相關應用大量增加。在國土利用、人口普查、交通運輸管理、傳染病分佈、自然資源分佈、社會學研究等領域，都可以發現 GIS 的應用。利用電腦快速計算的能力，可以在短時間內計算大量資料。透過地圖視覺化的觀察，並運用一些地理分析方法，可以發現表格式資料所無法看到的地理脈絡。在研究中用地圖的方式呈現，能夠更直覺、容易的傳達繪圖者想表達的概念與想法。同時，隨著網頁技術的演進以及瀏覽器的進步，其中網頁版的 GIS (WebGIS)也越來越蓬勃發展，從 Google Map 開始、Microsoft Bing Map、Yahoo! Map，甚至有組織開發了開放式原始碼且開放圖資的 OpenStreetMap、OpenLayers 等。除了給使用者多元化的選擇外，系統開發者也可依照不同的需求選擇合適的平台開發系統。

1.2 研究動機

中央研究院歷史語言所范毅軍研究員於〈試論地理資訊系統在歷史研究上的應用〉一文中提到[7]：

歷史上的戰爭、移民、區域發展、人口分佈、人物的地理分佈、疾病流傳、政區沿革等這些司空見慣的議題，在在具有強烈的空間屬性，

其連帶引起錯綜複雜關係，往往具有豐富的歷史意義，一般研究者多著重文字敘述，而疏於根據地圖做具體的空間分析，此事實上就等於平白忽略了一大部分史實，或者是錯失了瞭解或分析史實的一個面向。地理資訊系統作為一個應用性的工具，正可以對這方面的缺失有所補正。

所謂「史地不分家」，在探討歷史的過程中，總是離不開「人事時地物」等要素，而人、事、時與物都必須以某塊土地作為發生的舞台，幾乎所有的人類活動都與地域有著密切的關係。因此，如何透過 GIS 技術同時結合各種人文因素，有效的結合時間與空間，藉由視覺化的方式，讓歷史研究者可以輕易經由地理的角度，觀察歷史事件的起承轉合，從中觀察出複雜的歷史因果關係，並方便說明與講解史實，變成為學界和開發者所努力的目標。

是以，這十年來歷史研究也逐漸開始運用 GIS，稱為 HGIS (Historical GIS)。HGIS 是個高度跨學門的領域，同時橫跨了地理資訊系統與歷史學，看似相輔相成的兩個學門，卻有著極為不同的思維：GIS 是個講求精準的科學，運用許多科學化的測量與計算，分析地理現象與環境的互動；歷史則存在許多模糊性，尤其是傳統的史料記錄，諸如方志等等，相對位置往往比絕對位置更為重要。因此，要利用一個需要高度精準的工具來呈現模糊的歷史物件，似乎有些隔閡存在。

除了學科思維方式的不同外，地理資訊系統和歷史研究之間還有一些更具體、更實際的阻礙存在其間。荷蘭 Radboud University 的歷史學家 Onno Boonstra 在 International Journal of Humanities and Arts Computing 期刊 Barriers Between Historical GIS And Historical Scholarship 文章[9]中提到，為什麼歷史學家不使用 HGIS：

荷蘭 Radboud University 的歷史學家 Onno Boonstra 在 International Journal of Humanities and Arts Computing 期刊[9]中提到，為什麼歷史學家不使用 HGIS：

- GIS 是個相對新的研究工具，研究者需要多花時間學習適應，有一定的技術門檻在，且設計者多為非歷史研究相關人員，無法迫切的掌握到研究者所需要的是甚麼。
- 歷史學著重在文字的思考，圖表式的表達所帶來的重要性似乎不是那麼被看重，圖表單純只是視覺化文字的內容，HGIS 呈現是屬於附加的價值，歷史學還是以文字闡述為主。

- HGIS 缺少了足夠的方法論，當然也會缺少訓練，在培養歷史研究者的過程中，使用地理資訊的觀念並不是必備。

綜合上述，本研究即是利用 1904 年臺灣堡圖的向量化地圖為基礎，嘗試將此圖資與歷史資料做整合，運用於 WebGIS 上。並試圖以使用者的角度出發，可以讓研究者主導使用 GIS，而不單純只有被動的接受。希望能夠降低 GIS 使用技術門檻，不需要專業人士的協助，讓研究者能夠有動力去嘗試這新的技術，發現結合歷史地圖與歷史資料，整合時間與空間，可以為歷史研究帶來不一樣的面向。

研究目標為拉近歷史研究者與 GIS 之間的落差，提供歷史研究者一個容易入門的 GIS 工具、一個可以直接與使用者互動的工具，而不僅僅在於瀏覽已整理好的歷史地理資料。從操作資料的過程中，進而觀察出更多歷史現象，能夠在低技術門檻的情況下得到地理資訊系統的協助。專業的 GIS 軟體提供了許多地理分析繪圖的工具，根據帕雷托法則（Pareto 法則），又稱為 80/20 法則，可以發現對於 80% 的歷史研究者來說，大多只使用其中 20% 的功能來製作地圖就已足夠，本研究希望以這種簡潔而有效率的想法出發，提供一些輔助研究的工具，發揮 GIS 系統的潛力，建置一個地理視覺化的觀察並從中誘發歷史思考。

空間中的物件在 GIS 系統中可以以點、線與面的方式做呈現，不同的呈現方式所代表的地理意義也相當不同。點資料通常用於表示發生在某一現象或事件的地點，或是在小比例尺的地圖上無法以線與面來呈現的狀況所使用，例如世界地圖的首都、山峰位置、重要地標等。線資料用於表示該物件在地圖尚可視為無面積存在，常用於線路管道、交通路線等。面資料是由線資料所圍成的一塊區域，在此區域內可視為一樣的性質，例如行政區域、湖泊、土地利用區塊等。本研究將利用 1904 年台灣堡圖的向量式地圖建置兩種 GIS 工具，一為總督府抄錄契約文書 GIS 工具，主要處理的地理物件為點資料，也就是古契約文書的地理位置。另一為歷史行政區主題圖繪製工具，主要處理的地理物件為面資料，以臺灣堡圖行政區域為主，包含廳與堡里。政府施政資料皆以行政區域作為統計的基礎，掌握行政區域可以繪製許多主題地圖。

1.3 研究回顧

隨著 HGIS 的發展，世界各地都有一些相關的計畫值得學習與參考，下面就分別介紹幾個大型的 HGIS 研究：

1.3.1 GBHGIS

網址：<http://www.visionofbritain.org.uk>

GBHGIS (The Great Britain Historical GIS) 計畫於 1994 年由瑪麗王后學院 (Queen Mary, University of London) 開始，2000 年後由英國普茲茅斯大學 (University of Portsmouth) 所接手運作，計畫有設立一網站稱為 A Vision of Britain through Time。所收錄的資料地點包含整個不列顛群島，時間包括 19 到 20 世紀，約兩百年的時間，資料內容有人口普查資料、歷史地圖集和遊記等等。在這計畫中，英國不同時期的行政區邊界都轉成向量的格式，可放到 GIS 上看出不同時間點的行政區域的改變。計畫也收集了大量的統計資料，建成資料庫，所有的統計資料都可以對應到各個行政區域，繪製出各種不同的主題地圖。配合完整的行政區域，讓統計資料精準的反應該時間上的現象。並從史料中整理出不列顛群島上各地點的詳細資料，建成地名辭典 (gazetteer)，可以查詢某個地點在歷史上行政區和地名的變遷、相關歷史與統計資料等等。該計畫更進一步運用其所收藏的資料，編寫一些數位學習的教材，例如農業變遷與土地利用、歷史人口普查概述等。綜合上述，GBHGIS 收集的資料相當豐富，是一個非常完整的英國歷史典藏資料庫，對於研究英國的歷史相當有幫助。

1.3.2 The Harvard World Map Project

網址：<http://isites.harvard.edu/icb/icb.do?keyword=k28501>

World Map 是一基於開放式原始碼軟體所開發的網頁地理資訊系統，主要目的適用於輔助學術研究以及教學，透過 GIS 結合時空把資料整合在一起觀察與分析。系統目前建置了六種地區的地圖資料庫，非洲地圖 (Africa Map)、波士頓研究地圖 (Boston Research Map)、哈佛森林資料圖 (Harvard Forest Data Map)、中國地圖 (China Map)、巴黎地圖 (Paris Map) 和佛蒙特州地圖 (Vermont Map)。利用 GIS 圖層的概念提供數層不同的主題可以疊加上去，其中基本的圖層採用 Google map 的世界地形圖，可於其上再疊上如人口普查、歷史地圖、行政區域等圖層，做進一步的比較與觀察。

建置的方式採用維基百科 (Wikipedia) 的概念，資料由大家貢獻，收集一段時間後更新系統，新增圖層供大家使用。希望讓空間資料與技術能夠由全世界來分享，採用 OGC 所定義的 Web Map Service (WMS) 格式作為標準，如此一來即可讓所製

作出來的圖層讓其他的系統透過網頁 URL (Uniform Resource Locator) 來取得與呈現。此設計方法保持的增加資料的彈性，若依照 WMS 的標準建置的地圖資料，可以自由的放入系統中呈現，但對於歷史研究者來說還是無法即時處理研究需求的資料。

1.3.3 Historical GIS for Germany

網址：<http://www.hgis-germany.de/>

德國的歷史地理資訊系統，所收錄資料的時間介於 1820 至 1914 年，以 ArcGIS 為系統核心建立的空間資料庫，包含人口普查、經濟數據、鐵路車站、道路、河流等地理資訊。有提供使用者介面來操作系統，可疊加圖層、瀏覽地圖、時間軸等分析功能。美中不足的地方是操作介面不甚流暢，可能會對使用者造成困擾。

1.3.4 臺灣歷史文化地圖

網址：<http://thcts.ascc.net/>

臺灣歷史文化地圖計畫由中央研究院執行，核心系統以 ESRI 公司所開發的 ArcIMS4.0 所建置，可提供使用者較完整的 GIS 軟體所包含功能。所收錄的資料相當豐富，行政區包含清領時代縣里界到日治時期廳堡庄界，交通部運輸研究所的現代縣市界、鐵路網、河川、道路，日治時期地方機關、交通設施、土地利用、寺院詞廟等。也有些收錄關於原住民番社、遷徙路線、部落位置。地圖影像包含臺灣航照影像、日治時期所繪製的臺灣堡圖、臺灣地形圖、臺灣番地地形圖等。系統收錄了臺灣數百年來重要的歷史文獻地圖，所有的圖層都可開啟套疊在臺灣地圖上，可同時套疊數種圖層做交叉比對與觀察。使用者介面也提供了一些 GIS 常用的功能，方便研究者觀察。由於採用 ArcIMS 建置，在瀏覽時需要安裝額外的外掛程式才能夠使用系統，對一些研究者來說可能會造成阻礙。

1.4 論文架構

本論文第二章「使用之史料介紹」會說明論文中會用到之史料，以及處理方式。第三章「現有之 GIS 技術回顧」，會討論現有的 GIS 技術，其使用的情形，以及本論文會應用到的相關技術介紹。第四章「總督府抄錄契約文書 GIS 工具」與第五章「歷史行政區主題圖繪製工具」為本論文設計想法之實做，以使用者角度出發建置兩種工具。介紹如何整合古契約文書此歷史資料與臺灣堡圖、系統架構、設計方式和使用實例。以及提出一個通用性的 GIS 工具設計，讓歷史研究者能夠

彈性的運用。第六章為「結論與未來工作」總結論文完成的工具，瞭解其貢獻所在，並思考未來延伸的方向。



Chapter 2 使用之史料介紹

2.1 臺灣總督府抄錄契約文書

台灣歷史數位圖書館(Taiwan History Digital Library, THDL) [19]為集合台灣史一手史料之資料庫，提供全文檢索、詮釋資料檢索、後分類等功能，其中一史料集合為臺灣古契約文書，資料來源包括國立台中圖書館、台大圖書館等，數量約有三萬五千件的全文資料。臺灣在 20 世紀之前，土地在開墾、轉賣等權利移轉過程中都需要訂定契約來證明，而留下這些數量龐大的古文書，是研究清代到日治臺灣土地開發與經濟社會發展珍貴的第一手史料。

本論文所使用的古契書為 THDL 中之「臺灣總督府檔案抄存契約文書」，數量共 15899 件，佔 THDL 所收錄的契書一半左右。這批契書均包含完整的詮釋資料。時間分佈從西元 1701 年到 1945 年，橫跨兩百多年，但大多集中在日本統治時期，約西元 1900 年附近。臺灣總督府於明治三十一年(1898)起，設立「臨時臺灣土地調查局」，展開地籍整理事業，過程中透過抄錄的方式，保存許多民間契約文書。[5]從 Fig. 2-1 中可以大致看出這群契約文書在時間上分佈的狀況。它們大致因為三種原因，而被抄錄：

- (一) 產權糾紛
- (二) 行政需要
- (三) 地籍整理

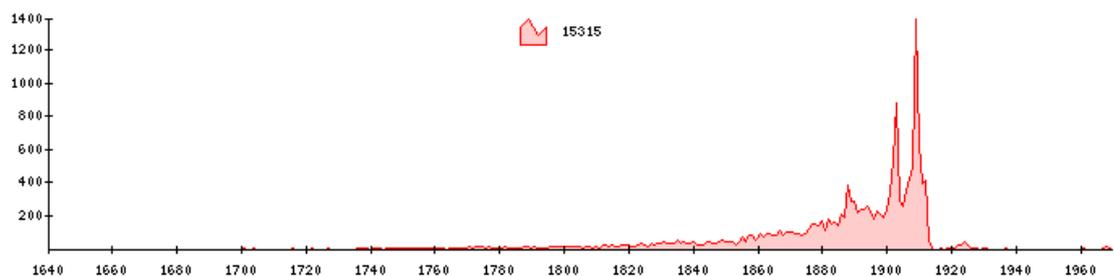


Fig. 2-1 總督府抄錄契約文書年代分佈圖

這批契約雖然統稱為「臺灣總督府檔案抄錄契約文書」，但具體來源則可再細分為四個部份：

Table 2-1 臺灣總督府檔案抄錄契約文書具體來源

來源名稱	數量
臺灣總督府檔案抄錄契約文書・15年保存公文類纂	7297
臺灣總督府檔案抄錄契約文書・永久保存公文類纂	5675
臺灣總督府檔案抄錄契約文書・土地調查公文類纂	2397
臺灣總督府檔案抄錄契約文書・高等林野公文類纂	468

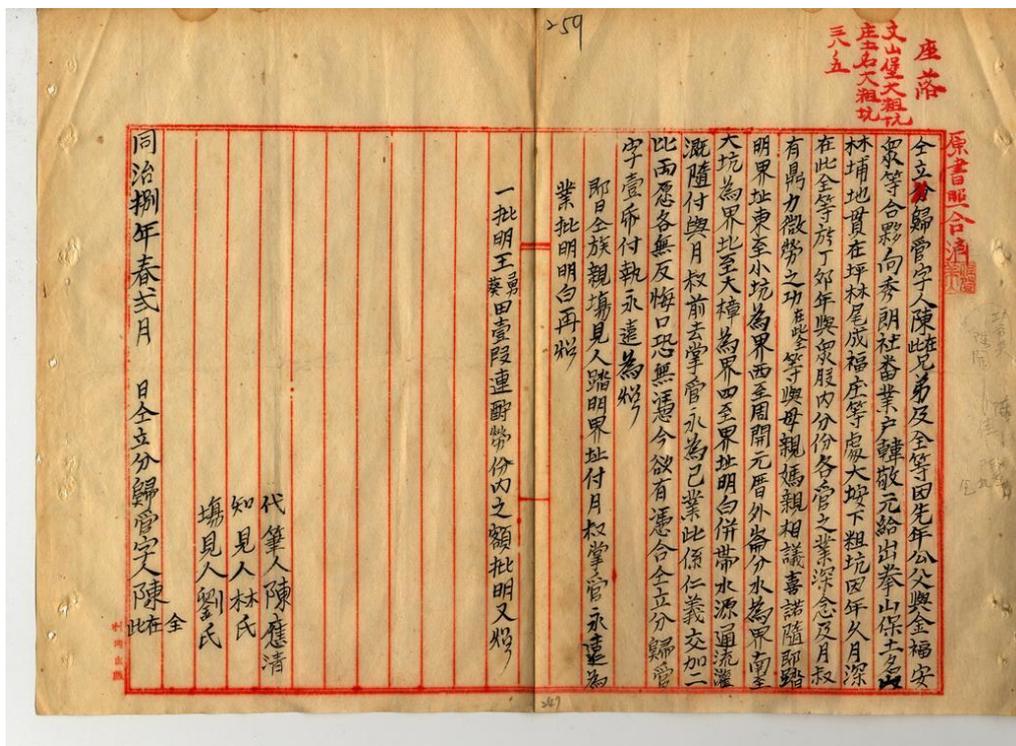


Fig. 2-2 契書地號位於右上角紅色字處

由於這批契約經過總督府抄錄，在現今可見的檔案中，這些契書保留了比較清楚的地理空間資訊，即「地號」。此外，這些地理資訊具有一定的書寫格式，相對容易利用程式批次處理。其他來源的契書地理資訊相當模糊，會有清朝之前的古地名，則需要更多史料輔助比對，才能瞭解其地理位置。這是本研究選擇這一批契約文書作為實驗對象的原因。Fig. 2-2 為契書原件上地號的書寫方式，一些地號例子如下所示：

竹北一堡富興庄土名水流東二一六 | 七二二三 | 一二一六 | 三
 文山堡芎林庄土名鶯仔瀨一七〇之一、一七〇之二

在 15899 件古契書中，有 2931 件契書完全沒有地號資訊，在本研究中先排除這兩千多件的契書，以剩下 12968 件包含地號資訊的契書來標記地理資訊。

2.2 臺灣堡圖

日治時期，臺灣總督府為了有效管理與經營，展開許多調查計畫，地圖的測繪也是其中之一，總共繪製了六種臺灣的全島地圖，分別是（1）1895-1896 年陸地測量部臨時測圖部的「臺灣地形圖」（2）1898-1904 年臨時臺灣土地調查局的「臺灣堡圖」（3）1905 年臨時臺灣土地調查局的「臺灣地圖」（4）1907-1916 年臺灣總督府警察本署的「臺灣蕃地地形圖」（5）1921-1928 年陸地測量部「臺灣地形圖」（6）1924-1938 年「陸地測量部臺灣地形圖」。[2]

本研究所使用的，是 1904 年的臺灣堡圖，20 廳時期的行政區域圖，行政區使用時間為西元 1901 年到 1909 年。此圖比例尺為二萬分一，是上述六種地圖中，比例尺較大的一種，地圖上記載相當多資訊，包括廳堡里街庄行政區界、政府單位、寺廟教堂、河流、山峰等高線、土地利用等等。[1]此外，臺灣堡圖繪製時期與臨時土地調查在時間上相去不遠，比較容易對應到臺灣總督府抄錄契約上出現的行政區地名；另一方面，它也是目前所能使用的數位化圖資中，與清代統治時期最接近的資料。換言之，由於地貌在這段時期的變動，相對於現在，幅度不大，再加上地名的沿革大多有跡可循，臺灣堡圖對於研究日治時代甚至往前至清代，都極具有參考價值。

本研究從中央研究院申請到臺灣堡圖的 Shapfile 中整理出行政區地名清單，但其 Shapefile 屬性資料表為 Big5 編碼，會有一些地名產生缺漏字的情形，需轉換成 Unicode 編碼來儲存並補上缺字(見附錄一)。用 ArcGIS 軟體中空間合併的方式得到行政區域的層級，包括廳、堡里、街庄與土名，其中土名為街庄下層的小區域地名。由於每個行政區塊皆有座標點，因此可以作為對應契書位置的參考。

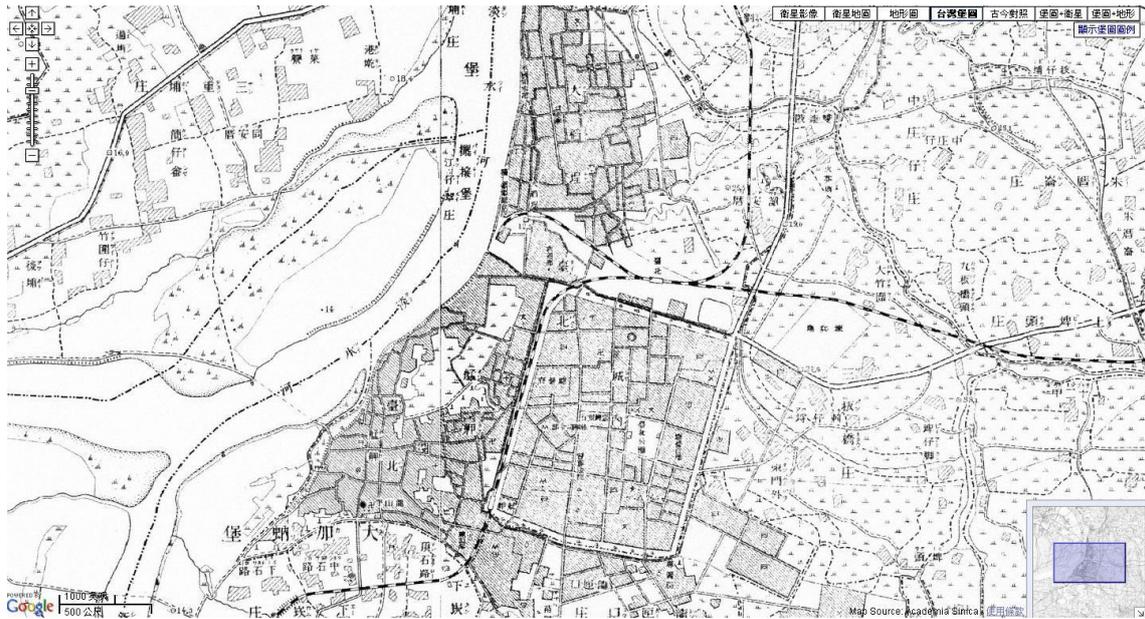


Fig. 2-3 臺灣堡圖—臺北城

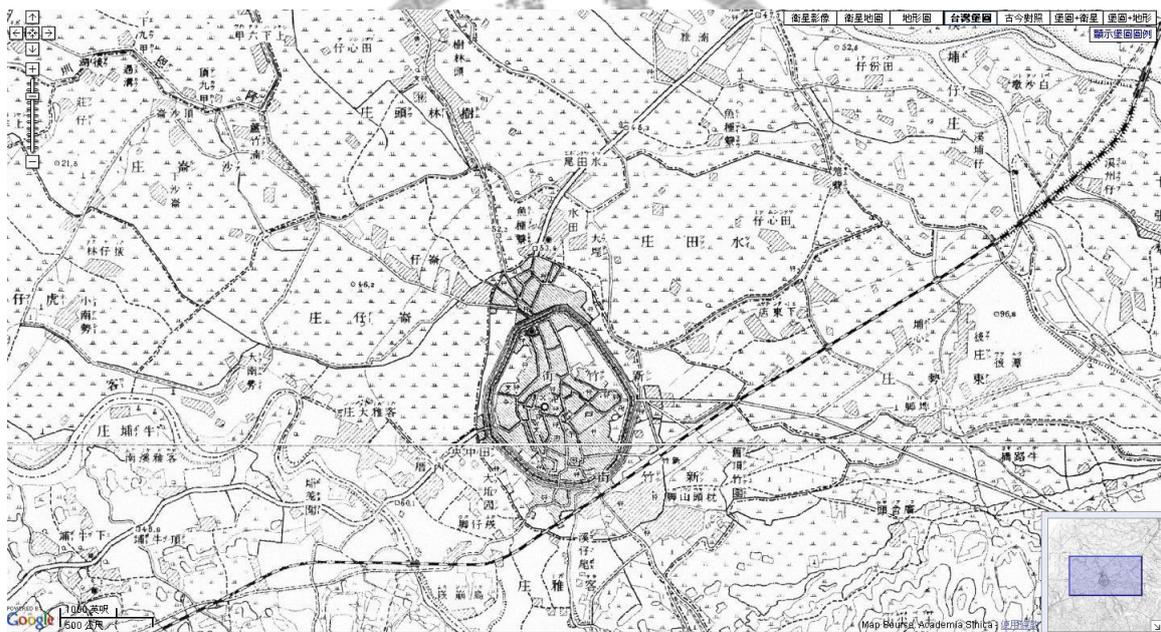


Fig. 2-4 臺灣堡圖—新竹街

2.3 日治統計資料庫

網址：http://tcsd.lib.ntu.edu.tw/main_browse.php



The screenshot shows the website interface for the 'Taiwan Database for Empirical Legal Studies'. At the top, there is a navigation bar with links for 'Search', 'Browse', 'About', 'Help', and 'Home'. Below this, the 'Browse' section is active, displaying a list of categories on the left and a main content area on the right. The categories include: 土地, 土地, 工業, 戶口, 水產業, 司法, 犯罪統計, 生產總額, 交通, 兵事, 官公吏及恩賞、文書, and 林業. The main content area contains a section titled '瀏覽功能說明' (Browse Function Description) and another titled '資料庫主題分類說明' (Database Subject Classification Description). The '瀏覽功能說明' section explains that the database provides '依類別瀏覽' (Browse by Category) and '依書名瀏覽' (Browse by Title) functions. The '資料庫主題分類說明' section explains that the database has 680 subject categories, which are derived from the 'Table Subject' of the 200,000 tables in the database. It also mentions that the database includes 681 volumes of Japanese colonial statistical publications, 194,075 pages of digitized image archives, and 103,732 pages of explanatory materials.

Fig. 2-5 日治統計資料庫

日本統治時期西元 1895 年到 1945 年，為了行政管理需要進行許多大規模的統計調查，例如戶口調查、犯罪、農業生產、財政等等，所調查的範圍涵蓋相當廣泛，是了解日治時期台灣法律、政治、社會、經濟、文化、教育等不可或缺的史料，調查過程累積了大量的統計資料。臺大法律學院與臺大圖書館合作將此珍貴的統計資料統一收錄整理，建置以日本官方統計資料為主的「日治統計資料庫」，共收錄 681 冊日治時期統計出版品，建置 194,075 頁數位化影像檔案，103,732 筆詮釋資料，其中包括書名、表格名稱、統計年份、出版單位等重要資訊，並製作成可供查詢及瀏覽之資料庫。

Chapter 3 現有之地理資訊系統技術回顧

3.1 GIS 簡介

GIS 包含兩種資料型態，向量式地圖(vector)與網格式地圖(raster)，本研究使用以向量式地圖為主。在向量式地圖中，對於地理物件有兩項重要的資訊，一為空間資料，另一個為屬性資料。空間資料是用點、線和面在座標系統上定位，表示其地理位置。屬性資料為描述某空間資料上的特性，例如一個面的屬性為某行政區，人口密度、雨量分佈等各為多少。

儲存資料的格式以表格為主，目前很多資料庫系統已支援空間資料格式，例如 Oracle、PostSQL、MSSQL 等，將空間資料結合屬性資料並透過一些使用者介面或其他方式輸出給使用者觀察。

3.2 GIS 資料格式與標準

OGC (Open Geospatial Consortium)是一個國際性的非營利組織，制定許多關於 GIS 開放的標準，例如如何表達點、線、面，以便更複雜的 GIS 系統開發。

3.2.1 Shapefile

由美國 ESRI(Economic and Social Research Institut)公司所制定來交換地理資訊格式的標準，記載點、線、面空間資料以及屬性資料庫，幾乎可被所有 GIS 軟體所接受。

3.2.2 KML

KML(Keyhole Markup Language)是基於 XML(eXtensible Markup Language)語法標準來交換地理資訊。由 Google 旗下 Keyhole 公司所發展，應用於 Google map, Google Earth 等相關軟體中，也越來越多 GIS 軟體支援此類的格式標準。

3.2.3 GeoJSON

GeoJSON 於 2008 年制定了第一版，尚未正式收錄於 OGC，它是基於 JSON(Javascript Object Notation)所制定交換地理資訊格式的標準，Javascript 乃是開發網頁的基礎程式語言，透過 GeoJSON 的格式，可以方便的在網頁上使用地理資訊，促進 WebGIS 的開發。

3.3 常用 GIS 軟體與工具

3.3.1 個人電腦 GIS 軟體

- ArcGIS -- ESRI 公司 GIS 軟體相關產品，以 ArcGIS 總稱，是企業級 GIS 應用的平台。依不同的應用有分成許多版本：桌面版本 (ArcReader, ArcView, ArcEditor 和 ArcInfo)、伺服器版本 (ArcIMS, ArcGIS Server 和 ArcGIS Image Server)、行動裝置版本 (ArcGIS Mobile 與 ArcPad)，目前最新版本為 ArcGIS 10.0。缺點是針對 Unicode 編碼支援不佳，處理中文的屬性資料庫會有缺漏字的情形。
- Quantum GIS -- 自由軟體的桌面 GIS 軟體，又稱為 QGIS，可運行於多種作業系統。也可以被開發者修改以執行其他 GIS 任務，另外有很多外掛套件可以用來擴展功能。

3.3.2 伺服器 GIS 軟體

伺服器端 GIS 軟體優點在於可以節省用戶端的計算負擔，大部分的 GIS 圖形計算皆於伺服器端完成，透過 HTTP 通訊協定，將計算好的地圖呈現於用戶端。用戶端只要負責處理使用者的輸入以及伺服器端的地圖結果顯示。例如 ArcIMS, 和 ArcGIS Server。開放式原始碼的軟體有 MapServer 和 GeoServer 等。很多網頁地圖應用採用此種方式開發，但當同時多人使用超過伺服器的負擔時，會有嚴重的延遲現象，造成操作不流暢的問題。

3.3.3 網頁地圖服務

基於網頁所開發的 GIS 系統，可以透過瀏覽器使用，目前比較大的入口網站都有提供地圖服務，包括了 Yahoo! Map、Bing Map 與 Google Map 等，這些大型的網頁地圖服務均有提供 API，本研究使用 Google Map 當做工具平台，使用 API 的版本為 version 2。Google Map 除了提供一般使用者瀏覽地圖與街景外，其地圖圖資有三種基本圖：傳統地圖、衛星地圖和地形圖。

除了大型入口網站有 GIS 系統以外，也有其他使用不同技術的地圖服務。豐富網際網路應用 (Rich Internet Application, RIA) 此概念由 Macromedia 公司於 2002 年所提出，希望能夠藉由新的技術，解決網頁瀏覽器所受到的限制，例如處理網路多媒體效果、與使用者互動介面、瀏覽器無法有效利用硬體運算效能等等。

透過此技術的推動，讓使用者可以得到更豐富的上網體驗，絕佳的互動與視覺化呈現。

目前常見的 RIA 技術分成兩種類型，一種為基於現有瀏覽器架構，創造互動性，例如 HTML5、Flash、AJAX 等，另一種為需要安裝額外的軟體，介於單機軟體與網頁內容間，例如 Microsoft SilverLight、JavaFX 等。

本系統採用 HTML5 作為行政區域界與主題圖繪圖的工具，有幾項優點非常適合用於此工具實做：

- 各大瀏覽器皆支援，Chorme、Firefox、IE9、Opera 等等。
- HTML5 不需安裝額外的外掛程式，研究者只要準備瀏覽器即可使用。
- 比起額外安裝的外掛程式，HTML5 所消耗的系統資源相對低，可以得到更佳的效能
- 有足夠的繪圖與多媒體功能



Chapter 4 總督府抄錄契約文書 GIS 工具

4.1 系統概述

此工具為結合古契約文書與臺灣堡圖，面臨的第一個問題是如何從古契約文書中找出其地理資訊並將其對應到地圖上。接著，選定建置系統所使用的工具進行開發，提供使用者介面操作。除了技術工具的選擇外，更重要的是需要思考幾個問題，古契書資料的特性為何，使用者要如何使用這批史料？要選擇怎樣的呈現？如何讓研究者瞭解數位化工具的使用，會不會造成使用上的負擔？

地理資訊系統組成包含地圖圖資與附有屬性的空間要素，一般的網頁地圖服務常使用街道圖來當作空間要素的參考，在圖上可查詢商家資訊、行政機關、交通位置……等等，這些空間座標點均有屬性欄位描述其特徵，例如商家名稱、聯絡電話等。如此的網頁地圖服務即可滿足大部分的使用者需求。但將古契書以此方式標注於地圖上給使用者似乎沒有考量到資料的特性，古契約文書需要考量到不同年代契書會有不同的分佈狀況；契書包含了數種類別，像是杜賣契、開墾契、闢分契等；在同一座標點在同一年代會有多件契約的狀況出現，在觀察或檢索上需要對不同的特性做客製化的功能設計。

此批檔案所收集的古契約文書中以土地買賣契約為大宗，瞭解其土地地理資訊會非常有參考價值，所以在呈現上忠實的反映出其精確的地理位置相對重要，需要盡可能的從古契約文書中擷取出最精準的位置。除了地點以外，古契約文書中還存在有其他不同的資訊，例如立契時間、契約類型等，都需要提供給使用者做不同的運用。同時在呈現上也要考慮到提供全文的方式，資訊技術擷取的東西無法到 100% 的正確率，需要透過人工校正，使用時必須能夠隨時的切換到全文瀏覽，以避免使用者觀察到錯誤的資訊。

4.1.1 系統架構

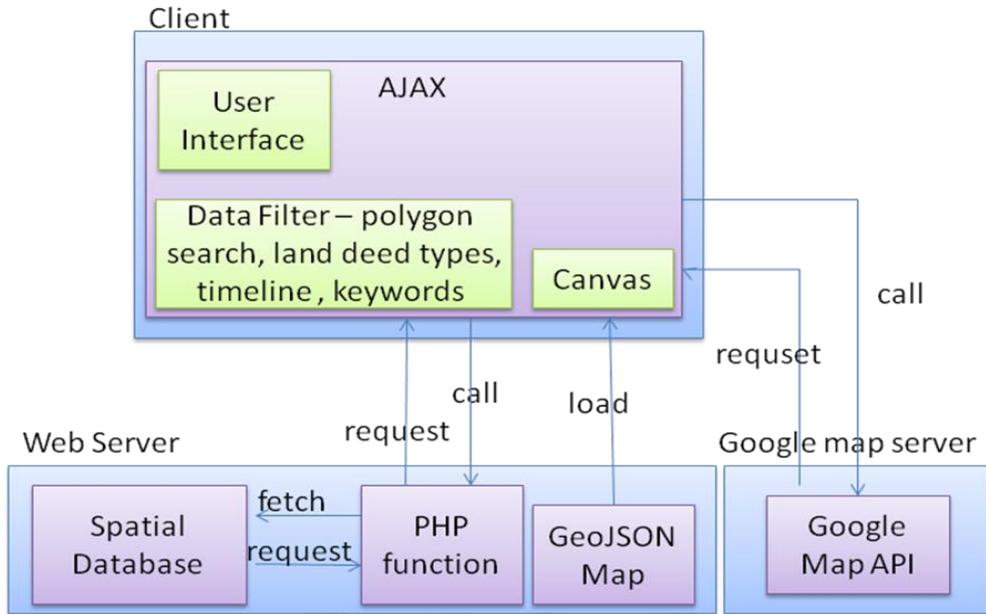


Fig. 4-1 歷史行政區主題圖繪製工具系統架構圖

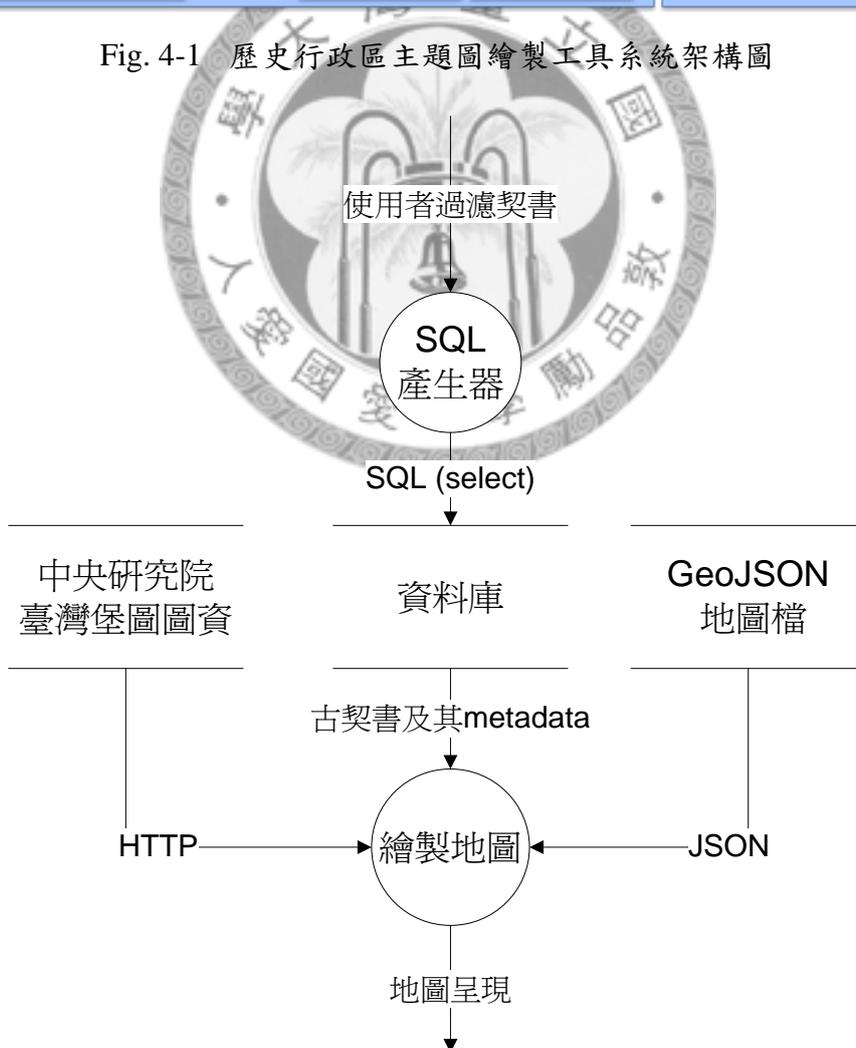


Fig. 4-2 歷史行政區主題圖繪製工具資料流程圖

用戶端以 AJAX(Asynchronous Javascript and XML)為核心，處理資料的讀取，圖層、地圖呈現以及地圖操作均以 Google Map API 來控制。當使用者透過使用者介面(User Interface)做了動作，由 AJAX 向網頁伺服器中的 PHP 函式要求資料，PHP 會依照參數於資料庫中選出符合的契書，並傳回去到用戶端呈現。Canvas 則是負責處理行政區域圖層，此為靜態的圖層。用戶端網頁開啟後直接將 GeoJSON 的行政區域圖資讀取到用戶端，以備所需時開啟。

Fig. 4-2 資料流程圖(Data Flow Diagram)中，使用者透過使用者介面選擇想觀察的契書，系統會將過濾條件轉成 SQL 語法於古契書資料庫中檢索，選出符合條件的契書，傳送至 Google Map 繪製函式，將契書分佈於地圖上點出來。當使用者想更換底圖時，系統視需要會讀取以 GeoJSON 方式儲存的行政區塊圖，或是使用中研院台灣堡圖圖資，此圖資會經 HTTP 由中研院伺服器取得。

4.2 對應契書地號流程

如前所述，總督府抄錄契約文書中，大多有比較清楚的地理空間資訊，即地號。因此，本研究首先利用自動方式，將這些資訊擷取出來，以便後續處理。地號是臺灣總督府政府在做地籍清查時將每塊土地所賦予的編號。從地號中可以找到這篇契書所在的區域與位置。完整地號如下所示：

竹北一堡富興庄土名水流東二一六 | 七二二三 | 一二一六 | 三

文山堡芎林庄土名鶯仔瀨一七〇之一、一七〇之二

上述地號中包含堡名、庄名、土名以及號碼，其中號碼代表該契書所包含土地的編號，可能不只一塊。這是用來判斷契書地點很重要的資訊。但是在總督府抄錄契約文書中，並非每一筆契書的地號都很完整。有時原始文件上並未記載地號，或地號不夠完整，或是字跡模糊等問題，後兩者在數位化的過程中會出現缺漏字、錯字或一字多型各種不同的狀況。這些問題都會增加自動化判斷契書地點的難度。例如：

南勢打鐵坑尾 (不完整地號)

文山堡□□坑庄土名□□□□ (原始文件字跡模糊無法數位化)

竹北一堡員窠仔庄土名員窠仔 (窠為罕見字，也寫做窠)

將把每一篇契書的地號萃取出來，整理成一份契書與地號的對應清單。在這個步驟完成後，就可以過濾掉不含有地號的契書，留下有機會找到地理資訊的部份。結果如下：臺灣總督府抄錄契約文書共有 15,899 件，其中有 2,931 件沒有書寫地號（其中包括並非契約文書的公文書類）；將這些沒有地號的契書去除後，剩下 12,968 件。其中有 466 件的地號不完整，包含漏字或錯字，因此可以確定契書地點的，共計有 12,502 件。這些數目也就是以下所使用的資料數量。

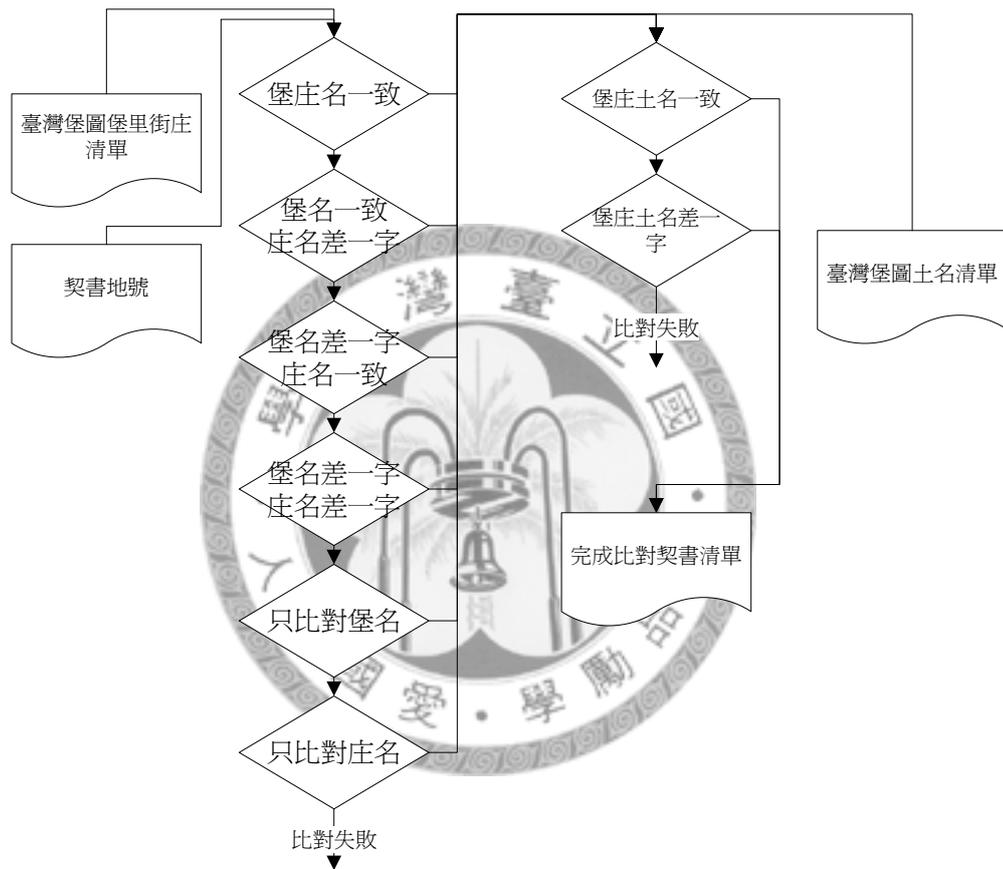


Fig. 4-3 比對地號過濾流程

接下來要將上述一萬多件契約文書的空間資料，與地圖對應。將臺灣堡圖地名清單以及契書中的地號，以程式自動比對。每篇契書會標上一個臺灣堡圖的行政區，同時也標示行政區的不同層級。含有土名的契書可以在地圖上點出比較精細的位置，但只含有庄名的契書就只能以庄為精準度來定位，甚至可能只有堡名的資訊，程式目標為盡可能找出越底層的行政區域。

此外，為了解決錯別字與漏字等地號資訊不全的狀況，在處理的過程中利用字串模糊比對，希望盡可能把每一篇契書的地區找到。首先，利用臺灣堡圖行政

區地名清單中的三種階層，上層為堡，中層為庄，最下層為土名。如 Fig. 4-3 所示，輸入包含兩種資料，一為臺灣堡圖地名清單，此清單由 Shapefile 圖資所匯出，並轉換成為 Unicode 格式；另一為國立台中圖書館 92,93 年所數位化的總督府契書以 XML 格式儲存。程式可以分成兩大階段來處理，第一階段為比對堡里街庄地理位置，實做幾層地名過濾器逐步比對下來，第一步為完全比對，若古契書中的地號可以完全符合堡名、庄名，即可直接定位和取得座標。接著開始模糊比對，比對地號與堡里和街庄名相差一個字的狀況，也就是說可以經由一次字的新增、刪除或取代而相同的地名。最後，只比對庄名或堡名。因為最後一個階段處理的地號幾乎都是很不完全的，經過上面的過濾後直接比對庄名。以下為比對到的例子：

三貂堡雞母嶺庄土名巫里岸一四〇 | 一 (三貂堡, 雞母嶺庄)

竹北一堡竹北一堡寶斗仁庄土名寶斗仁八九番ノ二 (竹北一堡, 寶斗仁庄)

文山堡松柏崎庄土名耳空龜五ノ一 (文山堡, 松柏崎庄)

三貂堡三□港□土地公嶺五二 | 二四六 | 一五六 | 四五六 | □七五 | 四七三 | 三 (三貂堡, 三叉港)

墩仔腳庄六四七-一 (苗栗三堡, 墩仔腳庄)

龜殼庄一八一之一 (苗栗三堡, 龜殼庄)

三貂堡□□庄土名□□四六 | 一 (三貂堡, NULL)

第二階段為試著將古契書上的土名標注於臺灣堡圖上，在第一階段中已把堡庄的位置定下來，在這階段輸入為臺灣堡圖土名清單，會先將該堡庄下的土名列出並一一與地號比對，抓出包含有土名的地號，得到更精準的契書地理位置。

竹北一堡南埔庄土名大份林五〇番ノ四 (竹北一堡, 南埔庄, 大份林)

文山堡員潭仔庄土名小粗坑一五ノ一 (文山堡, 員潭仔坑庄, 小粗坑)

文山堡原德崗坑土名藤寮坑三一番 (文山堡, 厚德崗坑, 藤寮坑)

經過這幾道地名過濾後，可以定位出約九成的地名。有了這些位置後，先利用 GIS 的工具，繪製各種主題的靜態地圖。在此使用 ESRI 公司授權臺灣大學地理資源學系的 ArcGIS 9 來操作。首先，將臺灣堡圖圖資中堡里行政區域的 Shapefile 匯入 ArcGIS 中，新增一地圖圖層。接著整理出比對好的契書與地名對應清單，製作成 csv 逗點分隔檔案。從圖層屬性設定中，選擇 JOIN 的動作，這個動作將會對於 csv 檔中的地名欄位編號與向量圖中相同的欄位編號作對應，如果比對一致的

話，軟體會將此兩筆紀錄合併在一起，也就是說賦予每筆多邊形行政區資料契書數量的屬性。最後，根據數量調整其行政區多邊形顏色，Fig. 4-4 中顏色分隔採用巢狀平均(Nested Mean)處理，並加上圖例製作出一靜態契書分佈地圖。

但操作大型 GIS 軟體對於研究者來說存在有些障礙，大多都需要專業人士的協助，才能產生出所需要的地圖。本研究將古契約文書資料與臺灣堡圖做整合，放在容易存取的 WebGIS 環境上，提供一些簡單操作的查詢功能，建置一研究者觀察契書的平台。

本研究總共將 12,502 件契約文書加以定位。有了這些資料後，即可繪製靜態的地圖，也可以建置動態的觀察系統。關於前者，先將目前所能掌握空間資訊的所有契約文書匯入 ArcGIS 後，會呈現如 Fig. 4-4 所示。本研究的底圖採用臺灣堡圖，以貼近台灣總督府執行土地調查時的行政區域的畫分，來觀察契書的分佈。

從這張圖中，可以簡單觀察出一些現象。首先，契約大多集中在台灣北部包含文山堡、石碇堡、竹北一堡等，相形之下，南部某些區域甚至沒有契約文書，整體而言呈現出北多南少的地理分佈。南部唯有在善化里西堡有大量契書分佈，查看後發現幾乎都是有關地主林人文的相關契書，內容多是因為河流改道，而發生地權歸屬的問題，其中也包含了一些理由書。日本總督府執行土地調查應該是全台灣平均抄錄才是，會呈現出此現象值得研究者進一步討論。

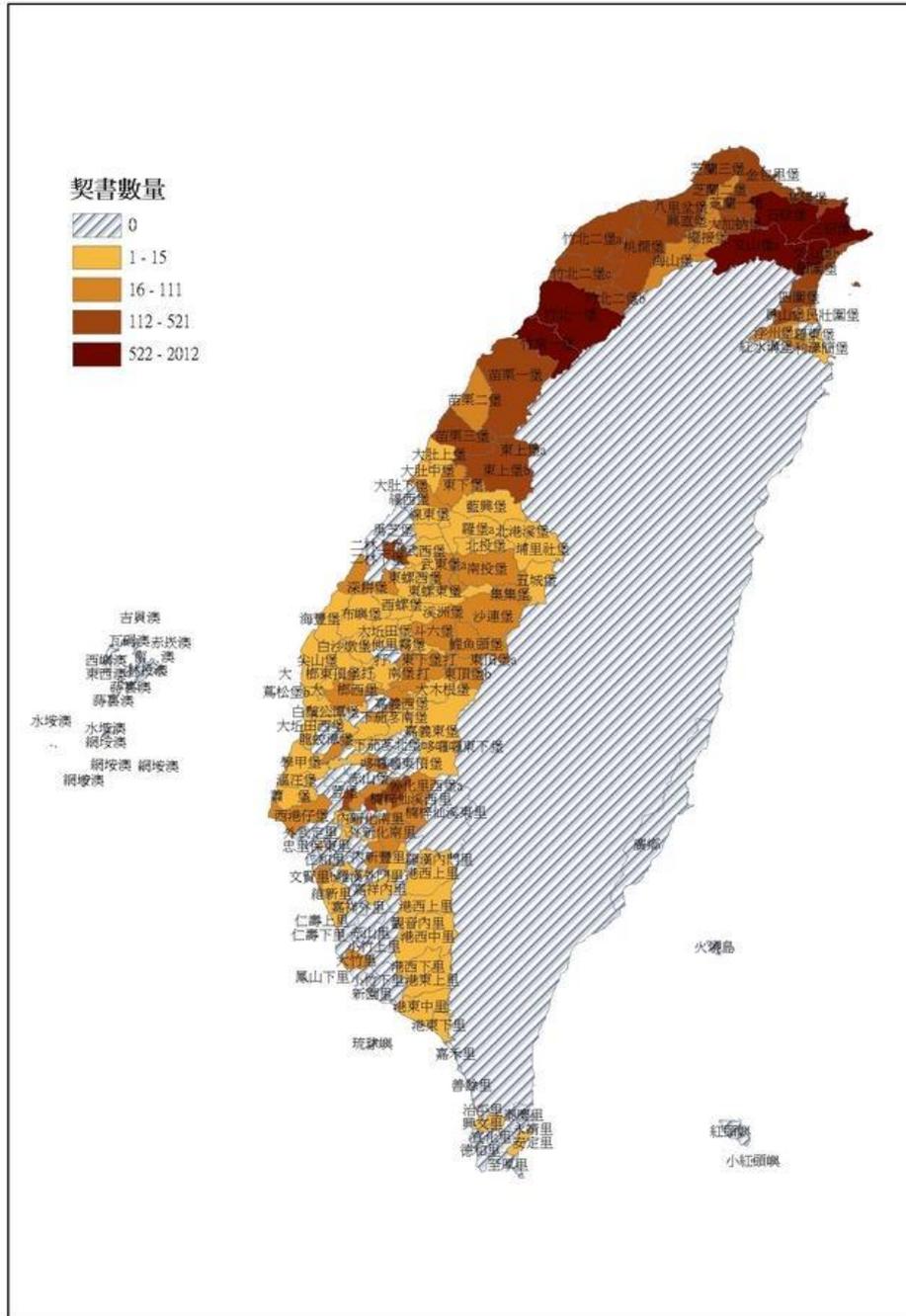


Fig. 4-4 以 ArcGIS 呈現之靜態契書分佈圖

4.3 系統功能

一個完整的 GIS 必需提供一個好的使用者介面，讓使用者一目了然，用最少的滑鼠鍵盤的動作完成所需的任務。總督府抄錄契約文書 GIS 工具也是朝這方向設計，此工具針對的使用者為歷史研究者，而通常研究者對於新技術大多不熟悉，

工具的操作方式要符合傳統的操作習慣並融合進新的 IT (Information Technology) 技術，存在一些困難，以下介紹此工具所提供的功能及其設計理念。

4.3.1 圈選檢索

在使用文字來檢索的資料庫上，欲找出存在於某地區的資料有其困難存在。文件中不一定會存在完整的空間資訊，無法以關鍵字檢索找出關於某地區的資料。透過 GIS 的輔助，從空間的方式來檢索，可以解決此一問題。

本工具可直接在地圖上圈選欲觀察的區域，以多邊形的方式圍起來。由於所呈現的資料類型為古契書，研究者也許會想觀察某區域的契書分佈，例如用多邊形圈選出河流流域可觀察在流域上的契書分佈情形，如 Fig. 4-6 所示，即可檢索出該流域的契書分佈，再搭配上時間的變遷，可以得到一些有趣的觀察。此功能預設是關閉的狀態，需要從左上方功能選單區啟動，即可開始圈選。

系統會紀錄使用者所點下的位置，放置於一陣列之中，使用 Google Map API 之 GPolygon 來繪製出所圈選的多邊形。當送出查詢後，先取出多邊形的四個端點，取得一最大的矩形包圍此多邊形。所有超過此矩形的契書先過濾掉，接著把剩下的契書由其座標點往東做一條無窮盡的直線，檢查此直線與多邊形的交點個數有多少，若有奇數個交點則在多邊形內(點 A)，若有偶數個交點則落在多邊形外(點 B)，如 Fig. 4-5 所示。程式會做此演算法判斷，再把檢索結果呈現於地圖上。

Pesudocode 如下所示：

```
function pointInPolygon(Polygon poly, Point pt)
    count = 0
    取得一直線方程式 L 通過 pt 並與 x 軸平行
    // L: y = pt.y
    for each edge E in poly
        if L 與 E 有交點
            count = count + 1
        end
    end

    if count % 2 = 1
        return true //多邊形內
```

```

else
    return false //多邊形外
end

```

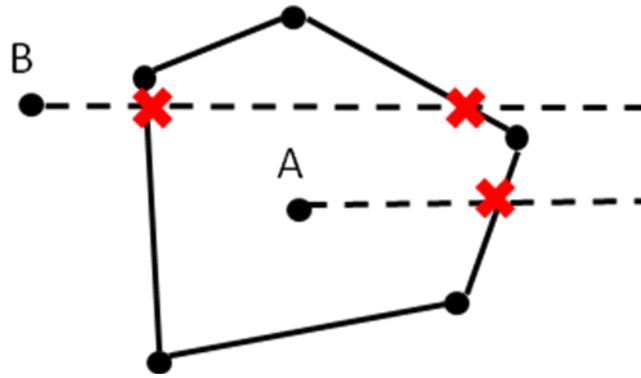


Fig. 4-5 判斷契書分佈點是否在多邊形內部

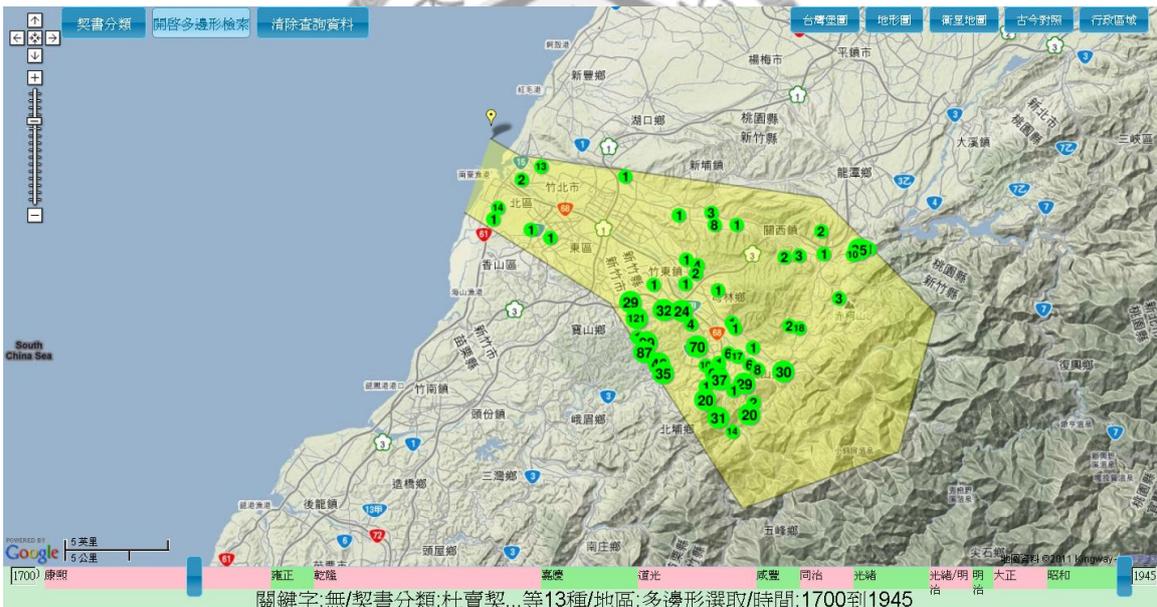


Fig. 4-6 圈選出新竹頭前溪流域觀察契書分佈

4.3.2 條件過濾與時間軸

契書本身具有一些屬性，例如地點、時間、內文關鍵字與契書類類，研究者可能會對某些類型的契書感興趣，這時候其他類型的就成為噪音，為了使研究者能夠專注於欲觀察的對象，本工具提供了一些過濾的功能。

契書類別分成以下這幾種杜賣契、鬮分契、開墾契、贖耕字、胎借字、典契、

添典/找洗、公文書類、丈單/執照、合約字、合股、證明/協議/決議、目錄和其他，此分類方法由盧家慶《台灣古契書自動分類與依分類定義契書角色》所分類[4]。總督府抄錄契約文書中詳細各類別件數如 Table 4-1 所示。

「時間」是 HGIS 中不可或缺的一項要素，所以時間軸的功能特別將它放在最顯眼的位置，讓使用者清楚瞭解到自己所觀察的時間點為何。時間軸的尺度單位為西元年，同時有提供清朝帝號作為參考對應，直到台灣日治時期。時間軸可以設定起訖年代，在軸上拖拉橫桿設定，或是直接輸入年代的數值，做一時段的觀察。

地點提供廳跟堡里為選項來過濾契書。也可以在一般檢索的地方輸入關鍵字，找出全文中包含的詞做過濾，例如可輸入林本源、金廣福、姜秀鑾等關鍵詞，找出某些人或是組織的開發活動。所有的條件均可合併一起檢索，例如同時使用時間、關鍵字與圈選多邊形。

Table 4-1 臺灣總督府抄錄契約文書類別分佈數量

契書類別	件數	契書類別	件數
杜賣契	4765	公文書類	120
鬮分契	3820	丈單/執照	336
開墾契	140	合約字	832
贖耕字	38	合股	7
胎借字	11	證明/協議/ 決議	416
典契	161	目錄	3
添典/找洗	30	其他	53

4.3.3 地圖圖層

底圖提供幾種不同的圖層可抽換觀察，包括網格式臺灣堡圖、臺灣地形圖、

臺灣衛星地圖、堡圖與現代街道圖之古今對照與臺灣堡圖行政區塊。臺灣堡圖圖資與古今對照圖，來自中央研究院地理資訊系統聯合實驗室。可以於系統中配合臺灣堡圖原圖觀察契書分佈，對照當時的地形地物，可以得到更完整的參考資訊。臺灣行政區域對照有廳、堡里與街庄，此圖資以 GeoJSON 的格式處理，使用者需要時由 HTML5 動態繪出，以隨機不同顏色分隔出行政區，滑鼠指標滑過其行政區域會有文字提示，方便使用者瞭解 1904 年臺灣行政區界。如何處理此 GeoJSON 圖資會在下面一章節介紹。

4.3.4 檢索呈現

檢索出來的結果以群集的方式呈現，同一個點可能會出現多筆契書，利用群集的方式並在上面標注數字，可以一眼就瞭解契書分佈的情形與數量，搭配地形圖或是其他的圖層，即可觀察契書分佈現象與其他地理環境的交互關係。契書分佈點會根據地圖的縮放比例做合併的動作，當使用者改變縮放比例時，程式會於地圖上做假想網格，同一格子內的契書分佈點會被集成一群集，更新地圖呈現分佈，如此一來可以避免相鄰的點疊在一起遮住數值。每個群集點均可點擊，會呈現出該地點各年數量分佈圖，並可進一步於台灣歷史數位圖書館(THDL)中檢視全文，如 Fig. 4-7 所示。Google Map API 透過 `fromLatLngToPixel` 函式可將地理座標換成以地圖上像素(pixel)為單位的座標，將地圖上每 100 像素畫成正方形格子，判斷格子內包含幾筆契書，合併加總起來成為一個新的群集點並附上加總的數量，不同的數量層級用不同顏色來區隔，藍色圓圈表示 1 到 10 件，黃色圓圈表示 11 到 100 件，紅色圓圈表示 101 件以上，此方法可以有效的增加查詢大量契書的呈現速度與方便使用者觀察。

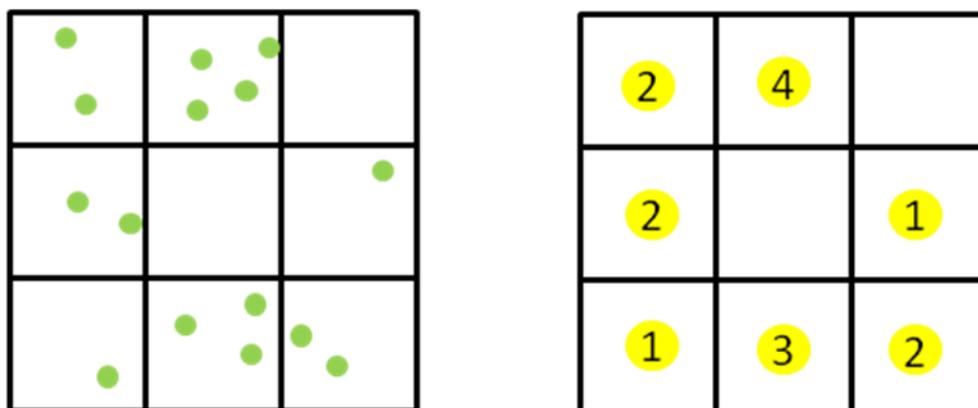


Fig. 4-7 (a)原本契書分佈 (b)調整縮放比例後，合併契書分佈點

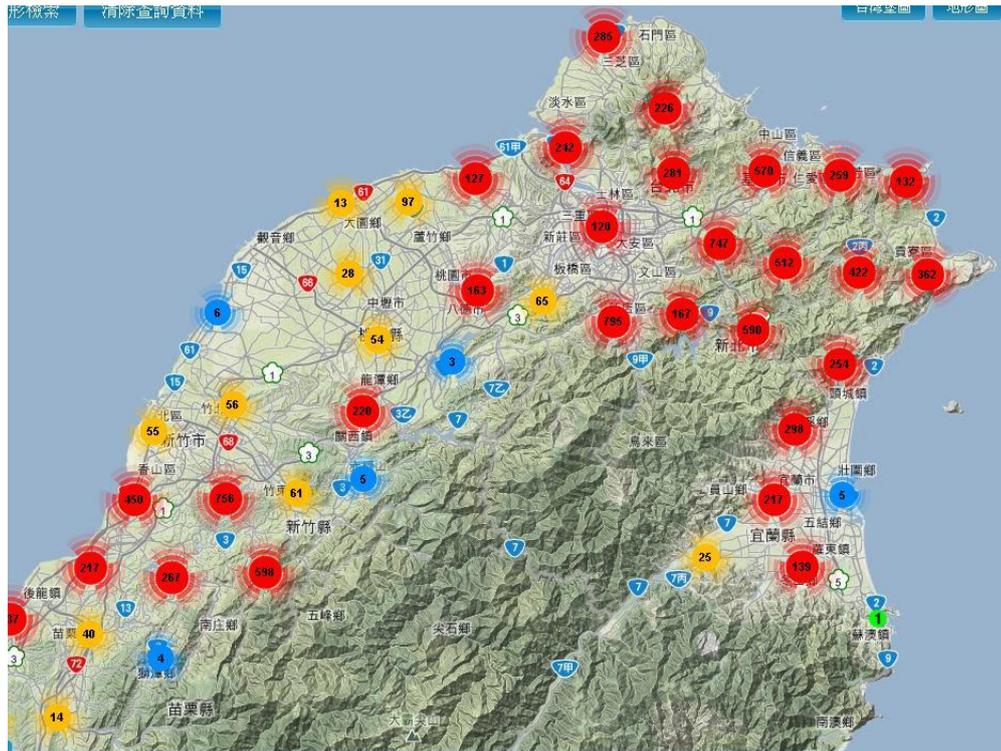


Fig. 4-10 群集合併契書分佈

4.4 操作實例與討論

4.4.1 金廣福墾隘

同治十年(1871年)陳培桂《淡水廳志》記載：

金廣福大隘，民隘。此隘本分設於城東廂之鹽水港、南隘、茄冬湖、石碎崙、雙坑、大崎、金山面，圓山仔、大北埔、小銅鑼圈等十處；其小銅鑼圈即舊之中港尖山隘也。各隘因地日闢，已越舊址；乃裁撤歸併為一，移入內山五指山右角。沿山十餘里均設銃櫃，為各隘之最大者；距城東三十五里小銅鑼圈之北。其石碎崙隘原設隘丁四十名，隘糧由官撥給充公租稅以補不敷。其大北埔中港、尖山二隘亦官奏設，民自給糧。其鹽水港、南隘、茄冬湖、小銅鑼圈四處，原設隘丁各二十名，雙坑原設隘丁十四名，大崎、金山面二處原設隘丁各十八名，圓山仔原設隘丁六名；隘糧均民自給。今統設一百二十名，就地取糧不敷，同知黃開基籌撥充公租穀四百餘石，每年發印串一百張，給令自行催收。

金廣福墾隘為新竹城東南山區的重要開墾組織，拓墾範圍包括今日之北埔鄉、寶山鄉、峨眉鄉等地，至清代後期已擴展至苗栗南庄、三灣鄉一帶。新竹東南山區生番猖獗，需要大量的人力與資金設隘防番，以往小規模的墾號很難進入此地開墾。金廣福墾隘就在這樣的時空背景設立起來，在清朝政府的支持下由粵

籍姜秀鑾及閩籍林德修、周邦正所設立。

吳學明於金廣福墾隘研究[10]中提到，金廣福的拓墾在時間上分成兩期：道光末年(1851年)以前為第一期，北埔、南埔、中興、月眉等庄中樞地帶的開墾設隘，和外圍舊隘線連成一氣，使寶斗仁與北埔間的土著前後兩面受到壓迫，於是全部敗退瓦解至內層，此間的土地悉歸金廣福所佔有。咸豐、同治年間(1851-1875年)為第二期，逐漸逼近內山地區的開墾，金廣福已管理北埔一帶五、六十庄地區。

1933年《北埔開闢史》記載「道光十四年甲午，先由三角城(竹東鎮三重里)及寶斗仁方面著手開墾，繼開大隘北埔一帶」。接著，道光14年(1834年)12月，淡水廳同知李嗣業任命金廣福墾隘的粵籍頭人姜秀鑾防番，所以金廣福入墾北埔很有可能在道光十五年(1835年)。以「姜秀鑾」為關鍵字，分別檢索1835年至1836年、1835-1838年、1835-1850年、與1835-1895年這四個時間區段，觀察與姜秀鑾有關之契約文書在「總督府抄錄契書地理資訊」離型系統上的座落分佈。

- 1835年至1836年(Fig. 4-11)，在地圖上查詢結果只有一篇契書，位於今竹東鎮三重埔(前稱為三角城)，為道光十五年的開墾契，可以發現姜秀鑾由此地附近開始開發。
- 1835年至1838年(Fig. 4-12)，契書分佈往現今峨眉鄉區域發展，且契書類型多為開墾契。
- 1835年至1850年(Fig. 4-13)，道光末年，金廣福開墾第一期約略結束，準備更深入內山開墾。
- 1835年至1895年(Fig. 4-14)，經過劉銘傳裁隘政策後，金廣福也被裁撤，此時的契書鮮少開墾契，多為杜賣契與鬮分契，分佈範圍遍佈今新竹寶山、北埔、峨眉等地。

從圖中可以看出這些契約文書的座落分佈，最早由今日的新竹縣的竹東市開始，逐漸向北埔、寶山、及峨眉三個鄉拓展。除了讓使用者可以依時間序列與地理分佈兩者關係，約略對照出當年金廣福整合新竹東南山區的官民隘，向新竹東南山區推進的拓殖開發之情形，可進一步探索其拓展情形是否受其他地理環境或政經社會條件所影響。

若依文書種類進行檢索後分類，也可以看出不同文書種類在時間序列上所顯示的差異，尤其西元1886年(光緒12年)以後，福建臺灣巡撫劉銘傳推行清賦，

裁廢墾隘，金廣福正式走入歷史，雖然還有與姜秀鑾相關的契約文書出現，但大概為與分割或轉手相關種類，而非首墾契書了。

綜合上述系統的呈現，可知其結果大致和吳學明的觀察大致相類，在考量到吳學明所採用的為姜家鑾家族所收藏的契約文書，不同的史料來源，卻有著類似的發展趨向，正說明了本系統在資料性質的代表性。而直覺式、可自由圈選範圍的觀察平台，更是早期台灣區域史學者所難以想像的優勢。是以，以金廣福墾隘為例，可以證明本系統對於歷史研究有著極大的幫助。



Fig. 4-11 1835 年至 1836 年，金廣福開墾初期



Fig. 4-12 1835-1838 年，開始拓展至峨眉鄉



Fig. 4-13 1835 年至 1850 年



Fig. 4-14 1835 年至 1895 年，可看出金廣福開發範圍

Chapter 5 歷史行政區主題圖繪製工具

5.1 系統概述

有許多的歷史資料均含有地理的特性，可以放入地圖中觀察。經由地理視覺化資料過後，可以呈現出複雜的表格，就可以此為出發點思考歷史的脈絡。歷史行政區主題圖繪製工具，希望以簡單容易的方式讓歷史研究者帶入地理資訊系統做觀察。

此工具所面臨到的問題會在於繪製的方式，以往，大型的 GIS 軟體以單機的方式執行，可以很有效的利用系統資源來作圖與運算，但卻需要具備一定程度的軟體訓練，這往往會令研究者望之怯步。研究的目的是能夠做成更一般性的工具，讓更多的歷史研究者能夠無負擔的使用，並希望能夠透過網頁的方式，完全開放，提供輔助歷史研究的工具。

原本在大型 GIS 軟體上可以輕易的處理的圖資，移到了網路的環境，會遭遇到一些困難。圖資本身檔案龐大，一個行政區的多邊形由數千個座標點所圍成，全台灣的行政區需要紀錄許多座標點與屬性，見 Table 5-1。大量的幾何資料透過瀏覽器呈現對於瀏覽器來說會是一個很大的負擔，需要新的技術協助才有辦法處理。以 Google Map v2 API 中，包含有繪製多邊形的函式，但處理這麼龐大數量的座標，會造成瀏覽器無法回應的狀況。

Table 5-1 臺灣堡圖圖資，幾何資訊

	行政區(含番界)	多邊形數	總座標點數
廳	21	162	69,493
堡里	204	250	137,885
街庄	2138	2749	429,243

如第三章所述，RIA 網頁技術逐漸興起，透過這些技術已有辦法負擔複雜的網頁繪圖任務。在選擇網頁技術時，希望以對使用者最方便的工具來實做，所以先不考慮需要外掛程式的技術，對於不熟悉電腦操作者可能會遇到一些問題。HTML5 不需額外安裝外掛程式，且大多新版瀏覽器對於 HTML5 均有硬體加速的

功能，效能可有顯著的提昇。

本工具以容易的方式提供使用者上傳 Microsoft Excel 試算表檔案，匯入資料庫以後，可以依照表格資料自由繪製出主題圖，方便觀察統計資料的地理分佈情況，提供另一角度的觀察。利用 HTML5 Canvas 作圖可以即時的將繪製好的面量圖直接呈現給使用者觀察，以非常低的技術進入門檻，讓研究者可以方便的使用此工具。

5.1.1 行政單元

如第二章所介紹，本研究所使用的圖資為西元 1904 年的臺灣堡圖，行政區域包含廳、堡里、街庄與土名，日治時期行政區域基本上由街庄組合而成。「只要從《臺灣堡圖》擷取街庄界線，臺灣大部分地區歷年各層級行政區的範圍，幾乎都可以組合出來。」[8]當我們掌握了行政區域的單元後，若遇到不同時期行政區域的統計資料即可利用行政單元組合成符合的區塊，或是將不同時期的統計資料組合而成共同的大區域，例如分為北、中、南、東和澎湖五區。

Table 5-2 臺灣行政區階層沿革

1901 年起	廳	支廳	區	街/庄/鄉/社	
		(堡/里/鄉/澳)			
1902-1904 年起	廳	支廳	區	街/庄/鄉/社	土名
		(堡/里/鄉/澳)			
1920 年起	州/廳	郡/支廳	街/庄/區	大字	小字
		市		町/大字	
1946 年起	縣	區	鎮/鄉	(段)	(小段)
		縣轄市			
	省轄市		區		
1950 年起	縣/管理局		市/鎮/鄉	(段)	(小段)
	省轄市		區		

5.1.2 系統架構

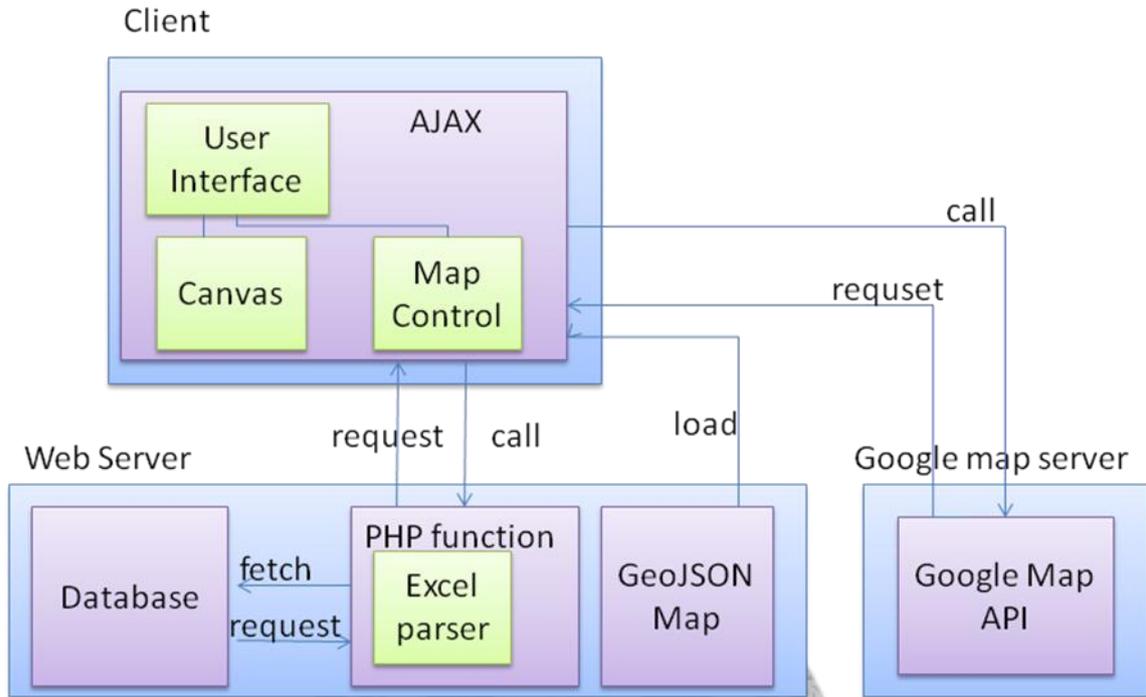


Fig. 5-1 歷史行政區主題圖繪製工具系統架構圖

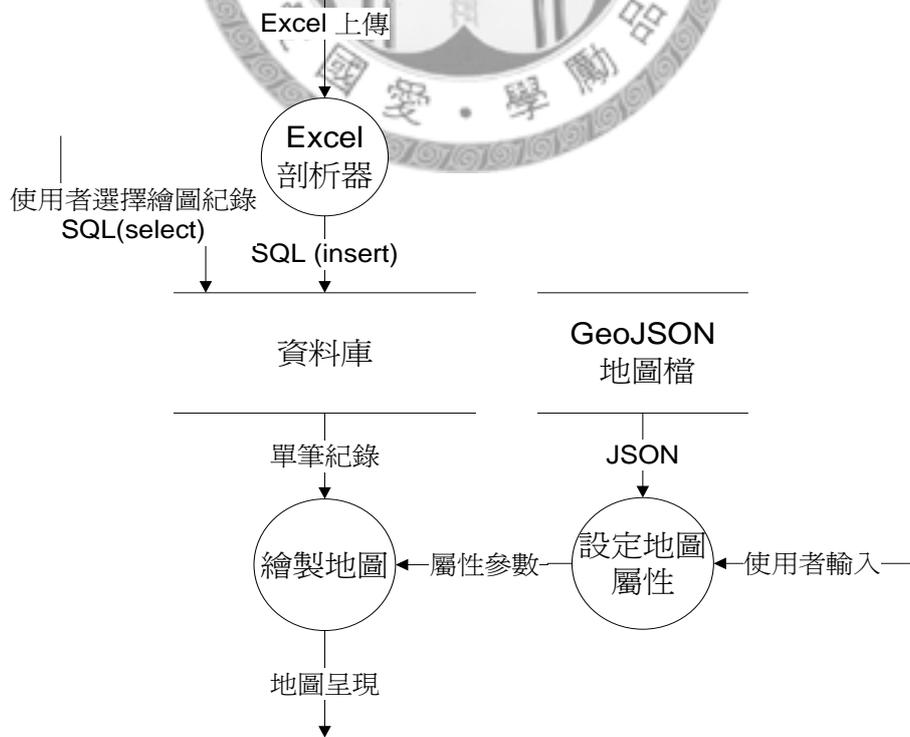


Fig. 5-2 歷史行政區主題圖繪製工具資料流程圖

用戶端仍以 AJAX 為核心，處理地圖的繪製與呼叫 PHP 函式。用戶端會上傳一 Microsoft Excel 檔案副檔名為.xls，由 AJAX 呼叫 Excel Parser 將檔案處理過後匯入資料庫中。當使用者點擊表格繪圖時，地圖用 HTML5 canvas 的方法讀取 GeoJSON 圖資並設定面量圖級距與顏色，然後繪製出來，並利用 Goole Map API 中的 addOverlay 函式貼至地圖上。使用者可自由調整面量圖各顏色的級距與級距個數。

資料流程圖(Data Flow Diagram) 如 Fig. 5-2 所示，使用者透過使用者介面選取本機電腦的 Excel 檔案，透過 AJAX 上傳至伺服器，使用 Execel 剖析器處理成 SQL 指令，將資料匯入資料庫中。當使用者選取 Excel 中某筆紀錄作圖時，會從資料庫中選出該筆紀錄並搭配使用者所選擇的繪圖選項，經由 HTML5 繪圖後輸出到 Google Map 上呈現。

5.2 系統功能

5.2.1 Excel 上傳

本工具可接受 Microsoft Excel 97 – 2003 版本的檔案，副檔名為.xls。內容格式固定，最左欄為地名與編號，程式以編號來辨認地名。上欄為屬性，針對各個地點提供一個屬性的值。程式可以支援一層 Fig. 5-3 或兩層屬性欄位 Fig. 5-4，例如以日治統計資料庫中明治三十八年「廳別檢舉件數表」，即有兩層屬性欄位。罪種別中，妨害風俗為第一層欄位第二層可以再分成賭博、賭場開帳、賭房給與、妨害風俗印刷物及物品販賣和猥褻。本工具會提供 Excel 表格範本，使用者可以直接下載輸入表格上傳製圖。Excel 剖析器(parser)處理後會把每一行剖析成一筆資料庫的紀錄，資料庫的欄位如所設計，一筆紀錄即可畫出一張主題圖。若有兩層的標題，也會做多筆紀錄自動加總的動作，將同樣的 Level2_title 的數值加總，可以直接觀察第二層分類的主題圖。

	A	B	C	D	E	F	G
1	id	loc_name	attribute_1	attribute_2	attribute3	...	attribute_m
2	id_1	name_1	value_11	value_21	value_31	...	value_m1
3	id_2	name_2	value_12	value_22	value_32	...	value_m2
4	id_3	name_3	value_13	value_23	value_33	...	value_m3
5
6	id_n	name_n	value_1n	value_2n	value_3n	...	value_mn

Fig. 5-3 Excel 表格格式—單層屬性

	A	B	C	D	E	F	G
1			title_1		title2	...	
2	id	loc_name	attribute_1	attribute_2	attribute3	...	attribute_m
3	id_1	name_1	value_11	value_21	value_31	...	value_m1
4	id_2	name_2	value_12	value_22	value_32	...	value_m2
5	id_3	name_3	value_13	value_23	value_33	...	value_m3
6
7	id_n	name_n	value_1n	value_2n	value_3n	...	value_mn

Fig. 5-4 Excel 表格格式—雙層屬性

5.2.2 功能選單

繪製主題圖中有提供幾個小功能讓使用者可以方便的使用繪製工具：

- 調整透明度：可設定主題圖的圖層透明度，方便使用者疊合不同背景圖層觀察。
- 下載靜態圖儲存：使用者可將單純的主題圖以 PNG 的格式下載儲存，以便做其他運用。
- 利用面積正規化(normalization)：將每個行政區的屬性資料值除以其面積大小，得到每平方公里的屬性值。這樣的正規化可以避免因為土地大小造成屬性的不均勻狀況。例如：原始為人口統計數字，透過面積正規化後代表的意義即為人口密度。

5.2.3 主題圖圖例設計

圖例是主題圖很重要的部份，如何設計級距及顏色需要謹慎的考慮。同樣的資料不同的級距可以造成不同的視覺效果，處理不妥的話會造成誤解的狀況。通常圖例的級距以不超過十個為準，分出太多級距需要多種顏色會造成觀察者的混亂，一般以 4 到 5 種為最佳，此系統可以讓使用者選擇 4 到 8 種級距數量，保留一些彈性存在。計算出級距也有幾種方式來處理，一般來說觀察資料的分佈情況來做選擇，本工具提供等比序列、等差序列與百分位數。假設資料數列為 a ， L 為數列長度。

等比序列首項為 a_0 ，找出所分佈的資料最小值作為 a_0 ，最大值為 a_n ， d 為公差。則可以得到：

$$a_n = a_0 + (n - 1)d$$

等差序列首項為 a_0 ，找出所分佈的資料最小值作為 a_0 ，最大值為 a_n ， r 為公比。

則可以得到：

$$a_n = a_0 r^{(n-1)}$$

看需要分成幾個級距， $n - 1$ 為級距數。帶入公式即可得到公差 d 或公比 r ，便可依公式推導出各級距的值為何。

百分位數是位於數列中 $p\%$ 位置的值，稱為第 p 百分位數，看所需級距為何先取得 $(100/n)\%$ 位置的值，並依公式類推依序取得各百分位數值為級距。

$$a_{L*(100/n)*i\%}, i = 1 \text{ to } n$$

5.3 圖資整理

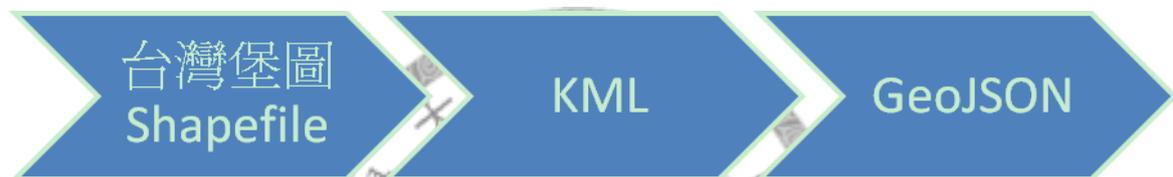


Fig. 5-5 圖資整理流程

5.3.1 圖資整理流程

原始臺灣堡圖行政區圖資為 Shapefile 格式，座標系統為 WGS84，型態為多邊形。第一步先透過 QGIS 的 OGR 檔案轉換工具將 Shapefile 轉換成 KML 檔案格式，方便撰寫程式剖析各個多邊形的地理資訊。第二步需要將 KML 格式轉換成 GeoJSON，HTML5 以 Javascript 為基礎，GeoJSON 的圖資格式可直接讀取繪出多邊形。轉換過程利用 Douglas-Peucker 演算法來壓縮多邊形，減少多邊形端點的數量，使工具初始化讀取圖資時更順暢。

5.3.2 Douglas-Peucker 演算法

有一些方法均可簡化多邊形的端點，例如只留下位於 10 的倍數的點，單純而有規律的省略一些座標點，此方法會造成多邊形的失真 Fig. 5-7。但此為行政區域多邊形，失真太嚴重某些地區會讓使用者無法辨識出行政區 Fig. 5-8，本系統採用 Douglas-Peucker 演算法避免失真。

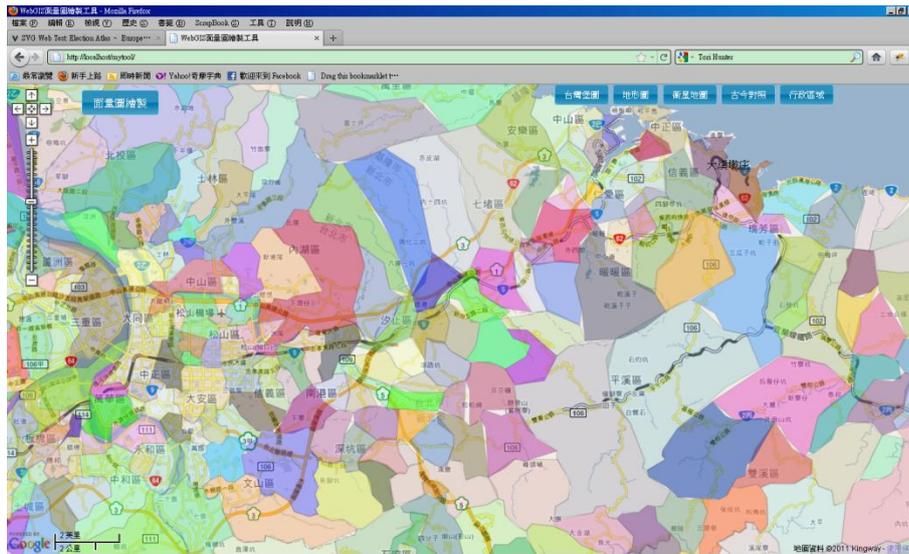


Fig. 5-6 使用 Douglas-Peucker 演算法

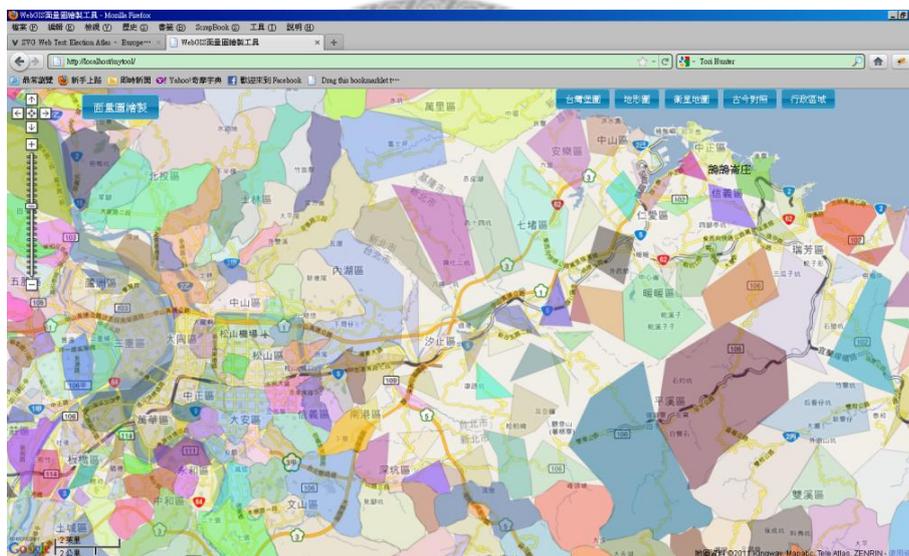


Fig. 5-7 規律的省略座標點，失真嚴重

Douglas-Peucker 演算法步驟 Fig. 5-8：

1. 連結多邊形中最遠兩個端點成一直線，找出其他端點中離此直線最遠的一個端點。
2. 判斷此點與直線的距離是否大於一門檻值 ϵ ，大於則保留此點，小於則捨去。
3. 保留此點，並連出此點與第一個端點的直線，重複第一個步驟。
4. 利用遞迴把所有的保留下的點都遍歷過一次。

最後即可得到刪減過的多邊形，可以盡可能的保持其原本的形狀，比規律的刪除一些端點還不容易失真。以下為 Pseudocode：

```

function DouglasPeucker(PointList[], $\epsilon$ )
//找出最遠的點
distanceMax = 0
index = 0
for i = 2 to (length(PointList) - 1)
  d = OrthogonalDistance(PointList[i], Line(PointList[1], PointList[end]))
  if d > distanceMax
    index = i
    distanceMax = d
  end
end
end
//如果最遠的距離大於門檻值  $\epsilon$ 
if distanceMax >=  $\epsilon$ 
  //Recursive call
  recResults1[] = DouglasPeucker(PointList[1...index],  $\epsilon$ )
  recResults2[] = DouglasPeucker(PointList[index...end],  $\epsilon$ )
  //將剩下的點放在 array 中
  ResultList[] = {recResults1[1...end-1] recResults2[1...end]}
else
  ResultList[] = {PointList[1], PointList[end]}
end
return ResultList[]
end

```

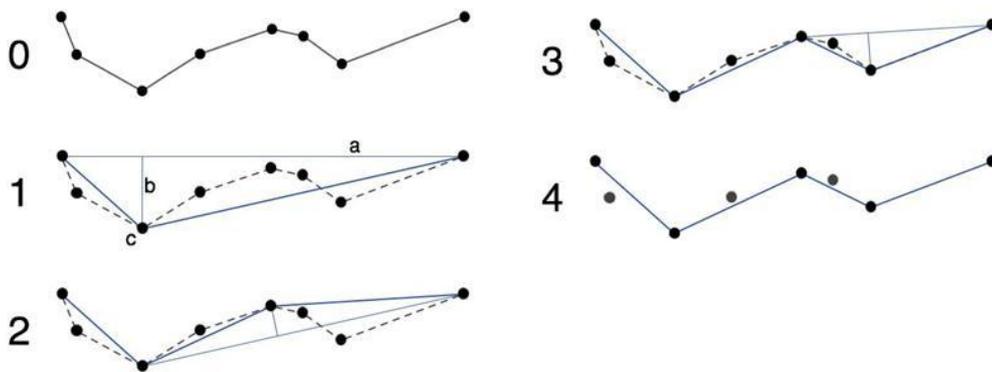


Fig. 5-8 Douglas-Peucker 演算法

5.4 操作實例與討論

5.4.1 工具操作流程

操作流程可大致分為三步驟：

步驟 1: 下載 Excel 樣板，系統會依據此格式來剖析表格。

步驟 2: 填完表格後上傳至工具中。

步驟 3: 可點擊所產生的表格繪製圖表，可調整圖例級距、主題圖透明度、下載主題圖、以面積正規化數量等功能。

5.4.2 日治統計資料呈現：犯罪統計

許多社經統計資料能利用此工具，製作大量的主題地圖，讓成果有更好的呈現。特別是臺大法律學院與圖書館合作之「日治時期統計資料庫」(<http://tcsd.lib.ntu.edu.tw/>)，共收錄 681 冊日治時期官方統計出版品，建置 194,075 頁數位化影像檔案，103,732 筆詮釋資料，可以查詢及瀏覽包括書名、表格名稱、統計年份、出版單位等重要資訊，雖主要目的為提供研究者一個內容具高信度的資料庫，因此所有資料均保留原始素材的面貌，但研究者仍可以自行轉換成可以數字運算的表格(tabular form)。以臺灣總督府出版之 1905 年(明治 38 年)犯罪統計書為例，圖四為依所屬廳別進行不同案件件數統計，若能有一個地理資訊平台系統，可以將這些統計資料轉換成地圖顯示，相信可以更幫助研究者觀察。

使用者可以自行上傳 Excel 檔案，可以依不同層級之行政區界，自行繪製主題圖，使研究成果有更好的呈現。目前僅能提供 1904 年台灣堡圖做為底圖，供使用者自行繪製日治時期台灣州廳與堡里兩個層級的行政區域主題圖，Fig. 5-12 即為依 Fig. 5-9 所繪製之 1905 年臺灣各廳賭博案件統計主題圖，使用者將 Fig. 5-9 轉為 Excel 檔案上傳後(Fig. 5-10)，會出現與 Excel 對應之表格按鈕，圖 Fig. 5-11 所示，即可選取所欲製作之項目(不同犯罪項目)，自行決定資料分級的級距，可以是等比或等差序列，然後選取顏色，即可繪製主題圖，由於圖上由淺到深的色塊即表示數量多寡，易於顯示區域差異，以 Fig. 5-12 為例可以發現賭博案件在各廳的差異，以台北最高，偏遠地區相對較低，可以提供研究者作初步的觀察，並作更深入的研究解析。除此之外使用者可以將其繪製好的主題圖透過工具所設計的功能下載

下來，Fig. 5-13 即為下載下來的地圖呈現，所使用的檔案格式為 PNG，可以提供使用者將此主題圖做其他的運用。

Fig. 5-9 日治統計資料庫—明治 38 年犯罪統計類之廳別檢舉件數 p2-p3

	A	B	妨害安寧				妨害信用								
	id	locname	匪徒刑罰令	官吏侮辱	官吏職務執行妨害	囚徒逃走及罪人藏匿	闖窺違犯	往來通信/家宅侵入	計	貨幣偽造	官x盜用私印	盜用私書偽造	免狀xxx	偽證	證書
3	6	斗六	0	0	1	0	3	1	3	8	0	0	2	0	0
4	3	宜蘭	0	0	0	0	3	0	3	6	0	1	0	0	0
5	18	阿猴	0	0	0	0	3	0	1	4	0	1	1	0	0
6	0	南投	0	0	0	0	1	0	1	2	1	0	1	0	0
7	5	恆春	0	0	0	0	0	0	0	0	0	0	0	0	0
8	16	苗栗	1	0	0	0	10	0	0	10	0	0	6	0	2
9	8	桃園	2	0	2	1	4	0	0	7	0	0	3	0	0
10	2	基隆	1	0	1	4	29	0	0	34	0	2	3	0	0
11	9	深坑	1	1	0	0	3	0	0	4	0	0	0	0	0
12	7	新竹	0	0	1	0	12	0	0	13	2	7	10	1	3
13	1	嘉義	0	0	0	2	28	0	1	31	0	1	2	0	0
14	4	彰化	0	0	1	2	17	0	0	20	1	1	7	0	0
15	12	臺中	1	0	2	0	18	0	0	20	0	0	6	0	0
16	13	臺北	3	0	1	1	89	0	5	96	1	0	5	1	0
17	15	臺東	0	0	0	0	1	0	0	1	0	0	0	0	0
18	14	臺南	0	1	0	0	43	0	4	48	2	2	7	1	0
19	19	鳳山	0	1	0	0	10	0	0	11	0	3	4	0	0
20	10	澎湖	0	0	0	0	3	0	1	4	0	0	0	0	0
21	17	蕃薯寮	1	0	0	0	1	0	0	1	0	0	2	0	0
22	20	鹽水港	0	0	0	0	0	0	1	1	0	0	1	0	0

Fig. 5-10 將表格轉換成 Microsoft Excel 檔



Fig. 5-11 匯入系統中選擇欄位畫圖



Fig. 5-12 主題圖呈現 - 賭博案件數分佈(廳界)

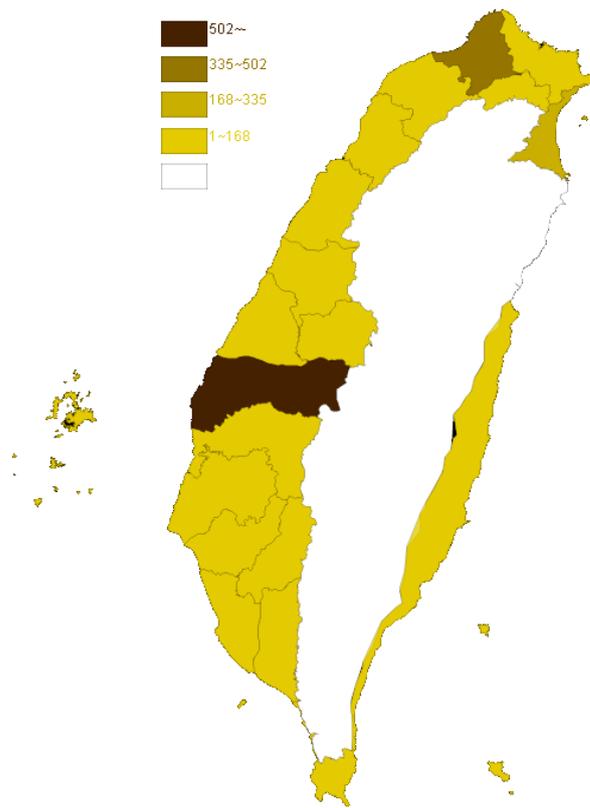


Fig. 5-13 主題圖下載 - 竊盜案件數分佈(廳界)

5.4.3 日治統計資料呈現：現住人結婚離婚及配偶

統計圖表之行政區可能與系統提供的行政區不同，以日治統計資料庫中，現住人結婚離婚及配偶統計表，年代由明治31年(西元1898年)到大正14年(西元1925年)。但這時段行政區有所變革，明治31年時為三縣三廳，明治34年由三縣三廳，廢縣增廳為20廳，明治42年又裁併成為12廳，而所有統計資料均基於行政區做統計，在處理主題圖的過程中瞭解其行政區變革相當重要。

不同行政區單位做比較時，需要做些資料上的調整。明治31年行政區以辦務署當單位，明治34年行政區以廳為當單位，本研究所使用的行政區為明治34年之20廳，需要把辦務署對應至這20廳的區域中。在此例中可將臺灣分為臺北、臺中、臺南、臺東與澎湖，五個大區域，見Table 5-3 即為北區行政區對應，表左邊為所有屬於北區的辦務署，右邊為改廳後的北區，將不同年代的行政區對應於

同樣的基準上即可做主題圖比較。明治 42 年裁廳後做了較大的變革，廳界做了一些重劃[26]：

- 合併台北、基隆二廳及深坑廳之一部分為台北廳
- 合併宜蘭廳及深坑廳之一部分為宜蘭廳
- 桃園廳照舊
- 合併新竹廳及苗栗廳之一部分為新竹廳
- 合併台中、彰化二廳及苗栗廳之一部分為台中廳
- 合併南投廳及斗六廳之一部分為南投廳
- 合併嘉義廳及斗六廳之一部分及鹽水港一部分為嘉義廳
- 合併台南、鳳山二廳及鹽水港一部分為台南廳
- 合併阿猴、蕃薯寮、恆春三廳為阿緞廳
- 將台東廳分為台東廳、花蓮港廳
- 澎湖廳照舊

很多廳部份被裁併至別的廳，可以從文獻中考察再將堡或庄行政區去合併出新的廳界，本研究所使用之圖資可再處理，提供研究者裁廳後的廳界。

Table 5-3 行政區域對應(北區)

明治 31 年行政區(辦務署)		明治 34 年行政區(廳)	
北區	臺北	北區	臺北
北區	基隆	北區	基隆
北區	深坑	北區	宜蘭
北區	滬尾	北區	深坑
北區	大料崁	北區	桃園
北區	桃仔園	北區	新竹
北區	新竹		
宜蘭	宜蘭		
宜蘭	礁溪		
宜蘭	頭圍		
宜蘭	東港		
宜蘭	羅東		
宜蘭	叭哩沙		
宜蘭	利澤簡		
宜蘭	蘇澳		

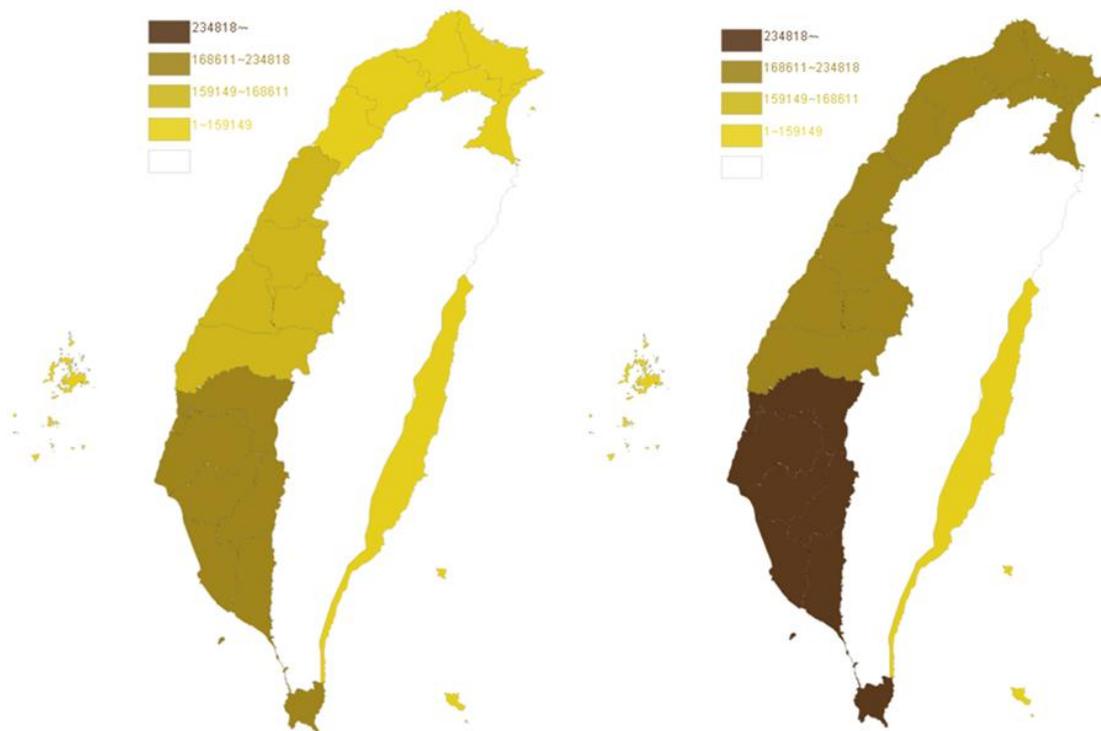


Fig. 5-14 左圖為明治 33 年、右圖為明治 34 年五區之年末配偶數

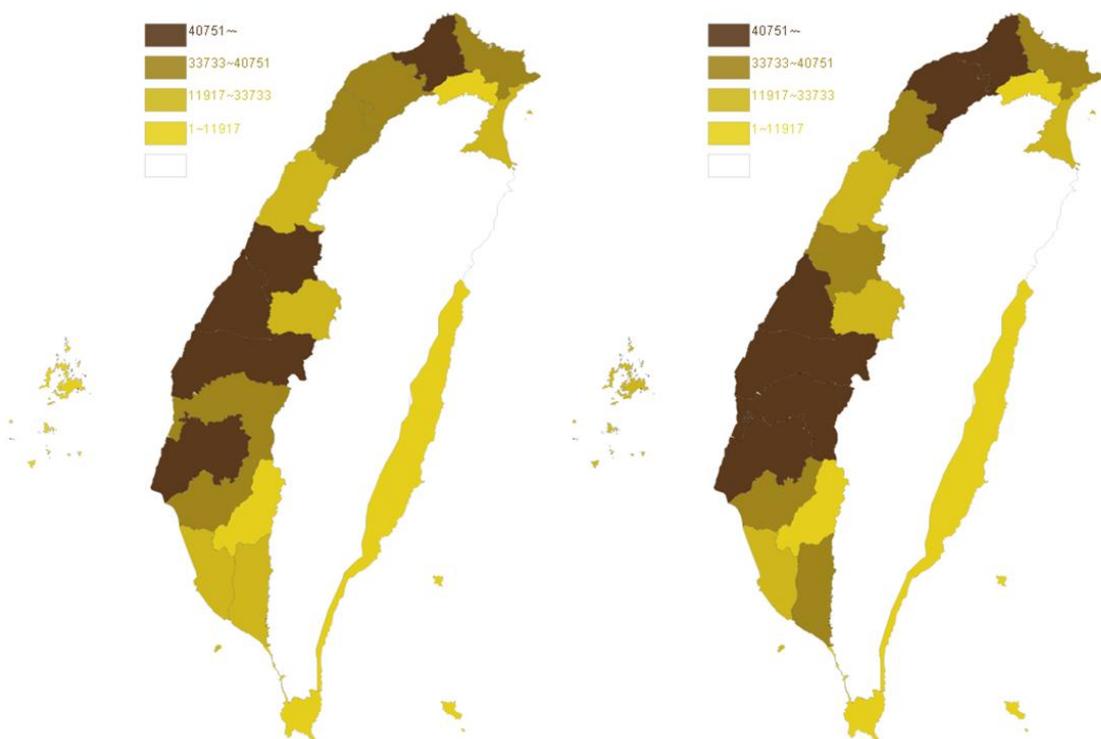


Fig. 5-15 左圖為明治 35 年、右圖為明治 36 年廳界之年末配偶數

Chapter 6 結論與未來工作

6.1 結論

台灣大學歷史系周婉窈教授在近日的一場演講中，強調地理空間對台灣史研究的重要，她從曹永和院士「台灣島史」的概念出發，曹永和指出了台灣作為一個歷史舞台的獨立性，從史前時代起便有各種不同種族、語言、文化的人群在其上活動，凡此都是台灣史的重要構成。以台灣作為空間的疆界，觀察不同時段人民的活動，是未來研究台灣史的重要走向。周教授則更進一步地將台灣的地理條件，歸納成山、海、平原三個意象，她指出：「山、海、平原，是我長期研究臺灣史之後，逐漸浮現出來的意象。我越來越覺得，無論我們研究臺灣史的哪個議題，如果能將這三個意象蓄於心，會有在茫茫大海中航行找到定位的安穩感，也能隨時維持一個鳥瞰的視野。若能這樣，再細瑣的題目，也都能因為連結到整體的大問題而不再顯得細瑣。……山、海、平原，無庸說是臺灣地理的三個重要的構成要素。……我在這裡標舉山、海、平原，主要是著眼於人群的活動，而不是地理環境本身及其對人群的影響。當然地理環境也是我們要注意的，……以人群活動為主的山、海、平原這三個元素，在臺灣歷史的發展過程中，不斷起作用，有時是交錯起作用，有時是分別起作用，或彼消此長，或毫無干係，端看議題而定。……我們就可以看到這種山、海、平原的人群及其文化正在互相滲透的動態過程。歷史引人入勝的地方正是在這裡。」[14]作為一個敏銳的研究者，周婉窈教授可說在不經意之中，為 HGIS 提供了一個來自史學界的最佳詮釋。本研究即企圖由系統建置的角度，為地理資訊研究和人文研究兩者的結合，提出一個更簡易、更直覺的使用方式，讓研究者能自由地悠遊於兩者之間，是史地兩個本該緊密相關學門充分結合。本論文提出 WebGIS 工具設計的想法，希望能夠讓研究者以最簡便的方式操作 GIS 軟體而得到協助，幫助歷史研究更順利的進行。開發過程的所有使用到的軟體工具皆是免費開放使用的，不需要花大筆的金錢購買專業的 GIS 軟體，也能達到所預期的效果。

具有地理資訊的物件除了其地理座標或地理形狀以外，存在有許多屬性。將這些屬性利用不同的方式整理與檢索，合併出某些使用者想關心的屬性，進而觀察出不同屬性間的差異或脈絡。古契約文書是以地契為主，所以契書的地理位置

屬性相當有參考價值。原本尚未做地理資訊標注的時，研究者若要觀察上百上千篇的契書地理分佈時，需要付出相當多的時間與精力，逐篇的將契書整理出其地理位置，並標注於地圖上，做地理資訊的觀察。總督府抄錄契約文書 GIS 觀察工具，收錄了一萬兩千多篇台灣古契約文書，佔了目前所發現的古契約文書將近一半的數量，如此龐大的數量若用人力來處理耗時費力，透過資訊科技的輔助可以節省整理資料的時程，並利用多種過濾的方式讓使用者排除不需要關心的部份，可專注的觀察與自己研究主題相關的契書，進而提昇研究的品質。此工具使用了一相對完整的古契書集合--「臺灣總督府抄錄契約文書」，此一集合分佈的狀況並非平均分配於臺灣各地，而是集中於北部，為何政府抄錄的契書會有不平均調查的現象？此一原因有待歷史學家進一步觀察與研究，本工具可提供最便利的研究協助。

每張行政區圖均有其施政上的意義，政府統計資料也基於行政區來處理。當掌握了行政區域，在配上當時的統計資料，可以繪製出許多方便研究者觀察的主題地圖。歷史行政區主題圖繪製工具可以讓使用者改變一般 GIS 數位典藏計畫的使用方式，由被動的接收資料成為主動的提供資料來使用 GIS。使用者整理好 Excel 表格後可上傳至系統中，馬上可以繪製出對應該表格的主題圖，直接可以由空間的方式觀察歷史統計資料。

除了呈現研究成果外，亦可利用於歷史教學上，不是每所大專院校或高中都有足夠的資源購買 GIS 軟體，透過簡易的上傳 Excel 試算表，降低學生在學習操作軟體的阻礙，達到歷史地理教學的目的，而不會浪費太多時間再於技術的操作。系統在設計上具有彈性擴充的功能，若有其他圖資可以利用，依照上述的方法轉換圖資並匯入，即可做出不同年代或區域的地理分佈面量圖。使用 HTML5 canvas 繪圖可以迅速的將主題圖產生出來，即時回饋給使用者，減少延遲所造成的困擾。

6.2 未來工作

專業 GIS 軟體對於一般沒有相關地理背景的歷史研究者存在一些阻礙，若能仔細思考與討論，將專業 GIS 軟體的部份功能抽取出來，透過簡易的方式設計與包裝，增進其可用性，成為一獨立運作的工具提供給研究者。如此一來可降低使用 GIS 軟體的進入門檻，又能達到 GIS 直接輔助歷史研究的目的。

此外，有許多的歷史資料具備空間屬性，透過地圖觀察這些歷史資料，協助

研究者以空間的方式來視覺化史料甚至觀察出其脈絡。建置此總督府抄錄契約文書 GIS 觀察工具包含三個要素：地理座標及其屬性、合適呈現的地圖與檢索觀察介面。未來若有其他相關的地理特性的資料，並取得符合資料的地圖，可嘗試匯入工具中呈現，利用工具已具備的檢索與瀏覽方式，搭配地圖觀察，更貼近當時的時空背景，對於人文歷史研究相當有助益。

本研究受限於資料的取得，目前只擁有西元 1904 年之臺灣堡圖行政區圖、古契約文書與日治統計資料庫的資料，無法全方位的提供臺灣歷史研究所需之 GIS 協助。但本研究實做之系統均有保留擴充彈性，若有取得其他相關地圖資料。以歷史行政區主題圖繪製工具來說，有兩種維度的擴充方式：

- 將臺灣不同時期的行政區域放入此系統中，可以瞭解到行政區域的變遷。政府施政與統計社經資料均基於行政區域，瞭解其變革對於每時期的社會狀態可以較精準的掌握，於 GIS 呈現上則必須考慮到變革狀況。此外，接收到不同時期的統計資料時，也可以忠實地呈現當時調查的狀況。
- 除了臺灣地區以外也可以加入其他國家的地圖，例如使用中國清代行政區地圖，繪製出中國相關的主題圖，但該國家行政區域變遷越頻繁需要考證的部份越多，需要歷史學家協助解決，以求精準的繪製出主題圖。

6.2.1 匯入方式

總督府抄錄契約文書觀察工具包含兩個部份：空間資料庫與前端操作介面，空間資料庫可利用 MySQL 或 PostgreSQL 等支援空間屬性欄位的資料庫，將具有空間屬性的資料物件建立於資料庫中。以古契書為例，本論文將古契書之詮釋資料、空間座標、地名與契書內文建成資料庫欄位，提供前端操作介面做檢索的動作。第一步先觀察想於地圖上呈現的資料屬性，將這些屬性抽取出來並建立於空間資料庫中。接著可以直接利用前端所提供的關鍵字與區域檢索等功能來查詢資料庫，或是略為增加修改，依照不同資料屬性額外設計新的查詢方式。

歷史行政區主題圖繪製工具只要有新的圖資即可匯入，先將所得到的 Shapefile 依照本文第五章的圖資整理流程轉換成 GeoJSON 地圖檔，放入網頁伺服器中，修改首頁之 Javascript 全域變數 *geo* 動態的指定到此 GeoJSON 地圖檔。完成後，即可在 Google Map 上繪製出行政區域圖。若要使用此行政區圖繪製相關主題圖時，

需再處理以下步驟。由 GeoJSON 檔建立每一塊行政區多邊形(polygon)之編號對應，確認每一行政區域均有獨立不重複的編號，也需處理破碎的行政區，例如澎湖群島於圖資中包含許一百多片多邊形區域，可將這些區域合併起來編號成為一大行政區。接著由圖資中整理 Excel 樣版，樣版包含地名列表與其編號對應，並依照使用者需求客製化 Excel 剖析器(parser)，讓系統可接受此 Excel 樣版。整理好地圖圖資與其編號匯入資料庫，並完成 Excel 剖析器的設計，即可利用新的圖資製作主題圖。



REFERENCE

- [1] 朱瑪瓏，〈用 GIS 重新解讀日治時期兩套中比例尺地圖〉，《地圖》第十三期，p93-103,2004.
- [2] 施添福，〈「臺灣堡圖」日本治臺的基本圖〉，收於《臺灣堡圖集》，臺北：遠流
- [3] 歐仲翔、林農堯、涂豐恩、項潔，〈GIS 與臺灣契約文書的空間分析〉，2010 年數位典藏與數位人文國際研討會論文集，pp.,台北：國立臺灣大學霖澤館，2010 年 11 月 29-30 日
- [4] 盧嘉慶，〈台灣古契書自動分類與依分類定義契書角色〉，臺灣大學資訊工程學研究所碩士論文，2008
- [5] 李文良，〈土地行政與契約文書-台灣總督府檔案抄存契約文書解題〉，《台灣史研究》，11（2）：p223-242, 2004
- [6] 范毅軍、白壁玲、嚴漢偉，〈空間資訊技術應用於漢學研究的價值與作用〉，《漢學研究通訊》，78：75-82, 2001
- [7] 范毅軍，〈試論地理資訊系統在歷史研究上的應用〉，《古今論衡》2：93-96, 1999
- [8] 葉高華、廖法銘，2009，〈歷史人口統計與地圖的關鍵性交會〉，林祥偉編，《台灣人文地理資訊系統的案例與研究》，頁 61-80，花蓮市：國立東華大學鄉土文化系
- [9] Onno Boonstra , Barriers Between Historical GIS And Historical Scholarship ,International Journal of Humanities and Arts Computing, volume 3, no 1-2,2009
- [10] 吳學明，《金廣福墾隘研究(上)》，新竹縣文化局，2000
- [11] Ian Gregory and Paul Ell, Historical GIS, Cambridge University Press, 2007
- [12] Anne Kelly Knowles, Placing history : how maps, spatial data, and GIS are changing historical scholarship, ESRI Press, 2008
- [13] Linda L. Hill, Georeferencing : the geographic associations of information, MIT Press, c2006

- [14] 周婉窈，山、海、平原：臺灣島史的成立與展望，
<http://tw.myblog.yahoo.com/jw!uduCo2SGHRYWlzLEAu0T/article?mid=960&prev=-1&next=956>，檢索日期 2011 年 6 月 30 日
- [15] Jieh Hsiang, Szu-Pei Chen, Hsieh-Chang Tu: On Building a Full-Text Digital Library of Land Deeds of Taiwan. Digital Humanities 2009, University of Maryland, College Park, June 22 – 25, 2009, Conference Abstracts. 85-90
- [16] 臺灣史料集成編輯委員會編，「臺灣總督府檔案抄錄契約文書」，台北市：行政院文化建設委員會、遠流出版公司，民國 95 年

網站

- [17] 日治統計資料庫，< <http://tcsd.lib.ntu.edu.tw/>>，檢索日期 2011 年 6 月 30 日
- [18] 臺灣歷史文化地圖，< <http://thcts.ascc.net/>>，檢索日期 2011 年 6 月 30 日
- [19] Taiwan History Digital Library (THDL) ，< <http://thdl.ntu.edu.tw>>，檢索日期 2011 年 6 月 30 日。
- [20] The Harvard World Map Project ，
< <http://isites.harvard.edu/icb/icb.do?keyword=k28501>>，檢索日期 2011 年 6 月 30 日。
- [21] Historical GIS for Germany ，< <http://www.hgis-germany.de/>>，檢索日期 2011 年 6 月 30 日。
- [22] A Vision of Britain through Time ，< <http://www.visionofbritain.org.uk/>>，檢索日期 2011 年 6 月 30 日。
- [23] Google Map ，< <http://maps.google.com.tw/>>，檢索日期 2011 年 6 月 30 日。
- [24] Google Map API V2 ，
< <http://code.google.com/intl/zh-TW/apis/maps/documentation/Javascript/v2/reference.html>>，檢索日期 2011 年 6 月 30 日。
- [25] MarkerClusterer ，
< <http://gmaps-utility-library.googlecode.com/svn/trunk/markerclusterer/1.0/docs/reference.html>>，檢索日期 2011 年 6 月 30 日。
- [26] Wikipedia ，
<http://zh.wikipedia.org/wiki/臺灣日治時期行政區域>，檢索日期 2011 年 6 月 30 日。

附錄一

缺漏字堡名	Unicode 補齊地名
大 榔西堡	大 糠榔西堡
大 榔東下堡	大 糠榔東下堡
大 榔東頂堡	大 糠榔東頂堡
打 北堡	打猫北堡
打 西堡	打猫西堡
打 東下堡	打猫東下堡
打 東頂堡	打猫東頂堡
打 南堡	打猫南堡
南 澳	南 蔡澳
茅仔 堡	茅仔 蔡堡
蕭 堡	蕭 墘堡
忠里	効忠里
羅堡	猫羅堡
東下堡	棟東下堡
東上堡	棟東上堡
荳堡	蔴荳堡

缺漏字廳名	Unicode 補齊地名
蕃薯 廳	蕃薯 蔡廳