

以文本分析呈現臺灣海外史料政治思想輪廓

劉吉軒*、柯雲娥**、張惠真***、譚修雯****、

黃瑞期*****、甯格致*****

摘要

政治制度的發展決定國家利益與人民福祉，臺灣從專制到民主的演變過程中，海外臺灣人民在僑居地資訊充分與言論自由的條件下，曾經對臺灣政治制度的改革扮演著重要的角色，包括思考、辯論、發聲、啟迪，這一段歷史的縮影記錄於當時海外臺灣人民所發行的政論性刊物。然而這些刊物大部分為非正式出版，發行量少、資料散佚、國內罕見，隨著時空環境的轉變，已經面臨逐漸流失的嚴重危機。因此，這批刊物的徵集、整理、數位化、全文化，不僅幫助保存這一段歷史紀錄，做為國內外學界進行臺灣歷史、政治、社會等面向的重要研究資產，也提供了數位人文研究的珍貴素材。本研究以左派刊物為例嘗試探討史料分析的二個面向的議題：第一個議題為人工關鍵詞與全文自動斷詞篩選詞語的比較，通常人工關鍵詞以單篇內容為範圍、呈現人事時地物等面向的核心資訊，而全文自動斷詞篩選詞語則以文字符號組成單元的計算方式，以詞語出現次數與使用重複性的比較為基礎，挑選出代表性詞語。本研究探討二者間的差異及其意涵。第二個議題為社會網路模型的應用，本研究以社會網路模型建構史料中人物、事件、理念之互動情形，透過社會網路的視覺化工具，觀察個別人物、事件、理念的角色及彼此的關聯，提供一種新型態的解讀與分析結果。整體而言，本研究透過數位人文的研究方法，呈現出海外臺灣人民部分的政治思想輪廓，並希望能促進後續更全面、更深入的研究成果。

關鍵字：政治思想、海外史料、數位人文、詞語分析、社會網路分析

* 國立政治大學資訊科學系教授兼圖書館館長，E-mail: liuj@nccu.edu.tw。

** 國立政治大學圖書館編審。

*** 國立政治大學圖書館組員。

**** 國立政治大學圖書館組長。

***** 國立政治大學圖書館專案助理。

***** 國立政治大學資訊科學研究所博士生。

Text Analysis on Overseas Taiwanese Journals for Political Thought Profiling

Jyi-shane Liu^{*}、Yun-er Ke^{**}、Hui-chen Chang^{***}、
Hsiu-wen Tan^{****}、Ruei-chi Huang^{*****}、Ke-chih Ning^{*****}

Abstract

A nation's political system has a profound influence on its growth and its people's welfare. Under the favorable conditions of abundant information and free speech in foreign residency, many overseas Taiwanese intellectuals are enthusiastic in deliberating, debating, advocating, and inspiring political thought, thus, playing an important role in the process of Taiwan's political transformation from anarchy to democracy. A part of this history was captured and reflected in political journals published overseas. These journals were mostly informally published with limited circulation, scattered and lost, and rarely seen in Taiwan. There is a serious fear of forever losing these journals. Therefore, the work of collecting, organizing, and digitizing these journals will help preserve important historical records and provide valuable materials for academic studies of Taiwan's history, as well as its political and social evolution.

The current research attempts to investigate two approaches to text analysis using left-wing journals as an example. The first approach compares human annotated keywords with computer-extracted keywords. Human annotated keywords usually represent core ideas of an article interpreted by human readers, whereas computer extracted keywords are generated by statistical selection of frequently used words. Differences between the two sets of keywords and their implications are analyzed and discussed. The second approach involves the use of social network modeling, and visualization in representing and observing entity relations in text. Overall, as an empirical study in the paradigm of digital humanity, the current research helps illustrate

* Professor, Department of Computer Science, and Director, University Libraries, National Chengchi University. E-mail: liujs@nccu.edu.tw

**Executive Officer, Digital Preservation Section, National Chengchi University Libraries.

***Officer, Digital Preservation Section, National Chengchi University Libraries

****Section Chief, Digital Preservation Section, National Chengchi University Libraries

*****Project Assistant, Digital Preservation Section, National Chengchi University Libraries

*****Ph.D. student, Department of Computer Science, National Chengchi University.

the profile of political thought expressed by overseas Taiwanese intellectuals and hopes to facilitate further investigation for in-depth results.

Keywords: political thought, overseas journal, digital humanity, term analysis, social network analysis

2011

DADHIC