

國立中興大學圖書資訊學研究所

碩士學位論文

應用關注點與自動擷取摘要視覺化使用者意見

研究—以中興大學圖書館為例

The Visualization Research on User Opinion by

Applying Concerns and Automatic Text

Summarization — A Case Study of the National

Chung Hsing University Library User Feedback

指導教授：郭俊桔 June-Jei Kuo

研究生：林美華 Mei-Hua Lin

中華民國一一一年九月

## 摘要

圖書館評鑑係評估圖書館各項業務服務品質是否有達到使用者預期的滿意度，其衡量服務品質主要由圖書館服務滿意度問卷調查做為評鑑結果的指標之一。針對使用者意見方面，使用者可透過圖書館提供的不同管道來表達意見，且實務上圖書館會針對個別留言做立即性處理後即結案。惟滿意度問卷調查僅為涵蓋某一調查期間及受訪樣本之意見表達，故此方式具有較長之時間落後性及樣本受限性的缺點。但為了快速改善缺失、滿足顧客，即時的資料蒐集、輿情分析，甚至資料圖形化呈現，皆有助於企業隨時掌握問題核心，顯見「使用者意見」經由意見累積匯集並加以處理的重要性。然而，「使用者意見」中，存在意見重複和與意見無關的訊息，如何去除將是研究的重要課題。

本研究旨在探討利用匯集後之中興大學圖書館使用者意見，首先將意見內容以人工標記方式決定關注點，並以 SnowNLP 自動情緒分類正評及負評的使用者意見，形成篩選條件。其後利用本研究之五大詞語判斷規則（詢問詞、轉折詞、否定詞、企圖詞、主題分類相關詞）自動化摘要有用的句子，並運用 Python 相關套件產出文字雲及統計圖。最後再採用同儕盲測評分比較本研究方法與 TextRank 演算法、MMR 演算法之優劣。

研究結果顯示本研究方法之自動擷取摘要結果成效最佳，且藉由文字雲亦驗證本研究方法有效刪除沒有用的句子後，能達到大幅度減少無用詞頻對文字雲產出結果所造成的干擾。統計圖表將關注點量化各項主題分類及 TOP 關鍵字詞頻分析，並將處理結果回傳到 Excel 檔，方便讓資訊使用者得依實際業務需求篩選出欲檢討的類別，搭配可彈性選取時間軸（年或季）的方式，有效地呈現各年度圖書館服務滿意度情況與趨勢分析，進而從中探析潛藏的問題點，並有效幫助圖書館主動提前發現問題以做出相關因應與改變，進而達到提升圖書館服務滿意度之目的。

**關鍵字：**意見句子識別、自動擷取摘要、關注點、資訊視覺化

## Abstract

Library evaluation is to evaluate the Service Quality to see if the service of the library meets the expectation of its readers' satisfaction. The assessment of the Service Quality is mainly based on Service Satisfaction Questionnaire Survey as the criteria of the evaluation result. Users can express their opinions through different pathways offered by libraries, and practically libraries respond to individual expressions to deal with the cases immediately. However, Service Satisfaction Questionnaire Survey only covers expressions from a certain period of time and limit of samples, so the survey has the weaknesses of time lags and sample limit. But in order to improve the shortages quickly, satisfy readers, collect real-time information, analyze public opinion, even to visualize data, the method can help to focus on the core of the problems. It clarifies the importance of opinion collection and processing gathered from user opinions. Yet, user opinions contains opinion repeats and irrelevance, and how to eliminate both is essential to do.

The main purpose of the study is to use and collect data from NCHU Library Reader Opinion. First of all, the opinions will be marked manually to decide the Concerns, and select the positive and negative reader opinions with SnowNLP auto sentiment analysis to generate filter points. Second, using the five word-judgement rules (inquiry words, transition words, negative words, intention words, subject classification relevance words) to automatize summaries which are usable sentences and using Python related packages generate word cloud and statistical graph. Last, compare the study method with the algorithm of TextRank and MMR with the peer blind-test score.

The study shows that Automatic Text Summarization does the best in the study and word cloud can testify the method of the study to eliminate useless sentences, even to greater extent reduce useless word frequency. Statistical graphs will quantize Subject Classification and TOP word frequency analysis. The process results will go back to Excel to make it accessible for users to filter by timeline and it will effectively show Trend analysis on reader service satisfaction and remark the potential problems.

**Keywords:** Opinion Sentence Identification; Automatic Text Summarization; Concerns; Information Visualizat