

透過 vibe coding 創造自己的文本分析工具

111042019 楊以恩

簡介

本作業紀錄開發一個文本分析網頁的過程，結合 AI 技術與視覺化介面，協助讀者更有效率地理解文本內容。作為外文系學生，在文學分析課程中，花費大量時間在基礎的文本整理工作上，例如人物關係梳理、詞頻統計、情感脈絡分析等，這個工具並非要取代傳統的精讀方法，而是希望透過自動化工具節省整理時間，讓讀者能夠將更多精力投注在文本背後意義的探討上。藉由 Claude Code vibe coding，我初步打造了這個仍在持續測試與迭代的文本分析平台。

一、研究動機與背景

1.1 從文學閱讀到技術實踐

作為外文系的學生，我在課程中經常需要進行深度的文本分析。無論是英美文學、比較文學，還是語言學相關的課程，文本分析都是不可或缺的核心能力，然而，在實際的學習過程中，我發現大量的時間被消耗在基礎的整理工作上——人物、標記情節轉折點、整理意象系統或翻譯等。這些工作雖然重要，但其機械性和重複性往往讓人疲憊，反而壓縮了真正深入思考和討論文本意義的時間。同時，我在課餘時間接觸了 AI，特別是自然語言處理相關的應用，這些技術完全可以協助處理那些繁瑣的基礎工作，更重要的是，透過視覺化的方式呈現分析結果，可以讓文本的結構和模式一目了然，為深度閱讀提供更好的起點。

1.2 技術賦能文學研究的可能性

這個專案的核心理念是「技術服務於人文」。我並不認為 AI 工具會取代理人類對文學的理解和詮釋，相反地，它應該是一個輔助工具，幫助我們更快速地掌握文本的基本資訊，從而有更多時間進行批判性思考和創造性解讀。舉例來說，當我們閱讀一部長篇小說時，人物關係的梳理往往需要反覆翻閱和筆記，但如果工具能自動識別人物、統計其出現頻率、分析其相互關係，並以網狀關係圖方式呈現，

我們就能快速掌握故事的人物結構，進而將注意力集中在人物的心理動機、象徵意義等更深層的分析上。

1.3 Vibe Coding 的學習與實踐

近期，vibe coding 的概念開始在開發社群中興起。這種「憑感覺編程」的方式，特別適合像我這樣有想法，但技術基礎相對薄弱的人，因此 Claude Code 讓我能夠用自然語言描述需求，快速獲得可執行的程式碼。

在多方嘗試和學習的過程中，我逐漸掌握了與 AI 協作開發的技巧，從最初簡單的 HTML 頁面，到後來加入 JavaScript 互動功能，再到整合各種文本分析庫，每一步都是在 Claude Code 的協助下完成的，這種開發方式不僅降低了技術門檻，更重要的是讓我能夠專注於功能設計和使用者的體驗，而不是被繁瑣的語法細節所困擾。

目前，這個網頁工具仍在測試和迭代階段，但已經能夠處理基本的文本分析需求，包括中文分詞、詞頻統計、情感分析、人物關係提取等功能(也在文中提供範例)。

二、開發工具與技術架構

2.1 Claude Code：從想法到實現

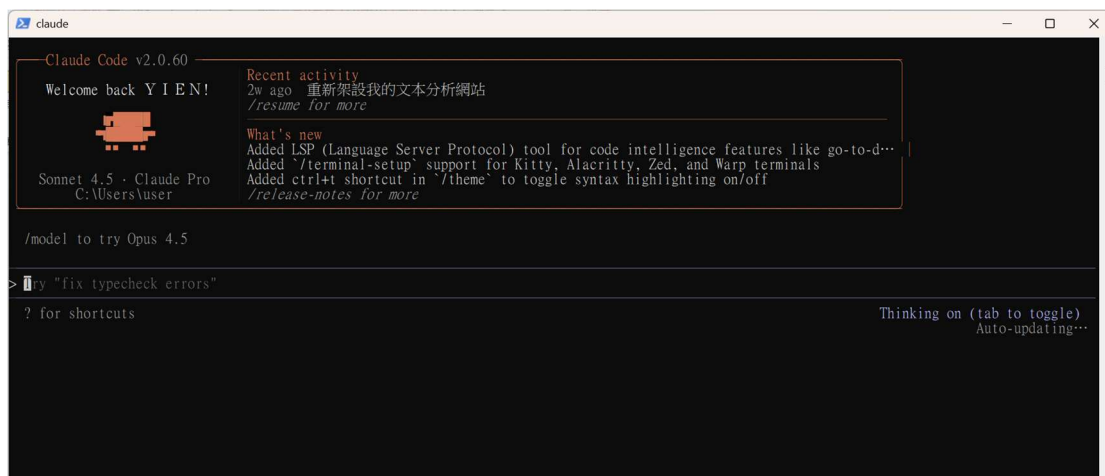
在傳統的程式開發中，從想法到實現往往需要跨越巨大的技術鴻溝。但 Claude Code 改變了這個現狀。作為一個外文系學生，我的程式基礎主要來自於自學，對於複雜的演算法和資料結構並不熟悉，但在 Claude Code 的幫助下讓我能夠：

用自然語言表達需求：我可以直接說「幫我做一個可以分析小說人物關係的功能，要能識別人名、統計出現次數，並找出經常一起出現的人物組合」，雖然仍需要建構前後端的架構，但不需要先一直思考要如何打 code。

即時獲得可執行程式碼：Claude Code 不只是給出程式碼片段，而是提供完整的、可以直接運行的解決方案，包括錯誤處理、邊界條件等細節。

學習最佳實踐：透過閱讀 Claude Code 生成的程式碼，我逐漸理解了程式設計的模式和慣例，這種「做中學」的方式比單純閱讀教科書更有效。

快速迭代優化：當功能需要調整時，我只需要描述新的需求，Claude Code 就能幫我修改程式碼，這種快速迭代的能力對於探索性的專案特別重要。



圖一 Claude Code 開發的介面

2.2 開發流程的演進

專案的開發並非一蹴而就，而是經歷了多個階段的演進：

第一階段：原型驗證（第 1-2 週）

- 建立基本的 HTML 介面
- 實現簡單的字數統計功能
- 驗證前端文本處理的可行性

第二階段：核心功能開發（第 3-4 週）

- 整合 Jieba.js 實現中文分詞
- 開發詞頻統計和排序功能
- 加入基本的視覺化圖表

第三階段：功能擴展（第 5-6 週）

- 加入情感分析模組
- 實現人物關係提取
- 支援多種文本類型（現代文、古文、英文）

第四階段：優化與測試（第 7 週至今）

- 效能優化，處理長文本
 - 介面美化，提升使用體驗
 - 實際案例測試，收集回饋
-

三、系統功能設計與實作

3.1 以使用者需求為導向的功能規劃

3.1.1 智慧人物分析系統

這是整個專案最核心的功能之一，在文學作品分析中，人物關係的梳理往往是最基礎但也最耗時的工作。

人名識別：結合規則和統計方法，識別文本中的人物名稱。對於中文，系統在理解文本後，會識別斷句與架構，甚至將代名詞類的詞語，也能識別為相同人物。

頻率統計：統計每個人物在文本中的出現次數，能快速幫助讀者瞭解主要人物和次要人物的分布。

共現分析：分析哪些人物經常在相近的段落中出現，推測其關係的緊密程度，系統會計算共現矩陣，並以網狀關係圖的方式視覺化呈現。

情感關聯：分析人物出現時的語境情感，判斷人物與正面或負面情緒的關聯程度。

3.2 技術實現的關鍵點

作為一個外文系學生，從零開始開發文本分析系統幾乎是不可能的任務。幸運的是 GitHub 上有豐富的開源資源可以參考學習，我的開發過程，很大程度上是一個「讀碼、理解、改寫、整合」的過程。

CharaterGo (<https://github.com/zfsang-zz/CharacterGo>)

- 這是我人物抓取功能的基礎
- 理解其人物關係的原理
- 學習了如何載入自定義詞典，這對處理文學作品來說特別重要

Text Analyzer (<https://github.com/ramgpal/TextAnalyzer>)

- 提供了基本的文本統計框架
- 我仿效其模組化設計，將不同功能拆分為獨立元件

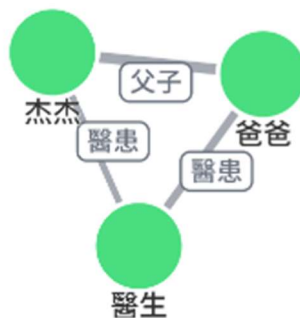
sentiment-chinese (<https://github.com/shibing624/pysenti>)

- 學習了基於詞典的情感分析方法
 - 參考其情感詞彙分類方式，建立了自己的詞典
 - 加入了針對古詩詞的情感詞彙
-

四、實作成果與案例分析

4.1 實際應用案例

4.1.1 床邊故事分析：《我不要爸爸》(<https://www.storymami.com/story/i-dont-want-daddy/>)



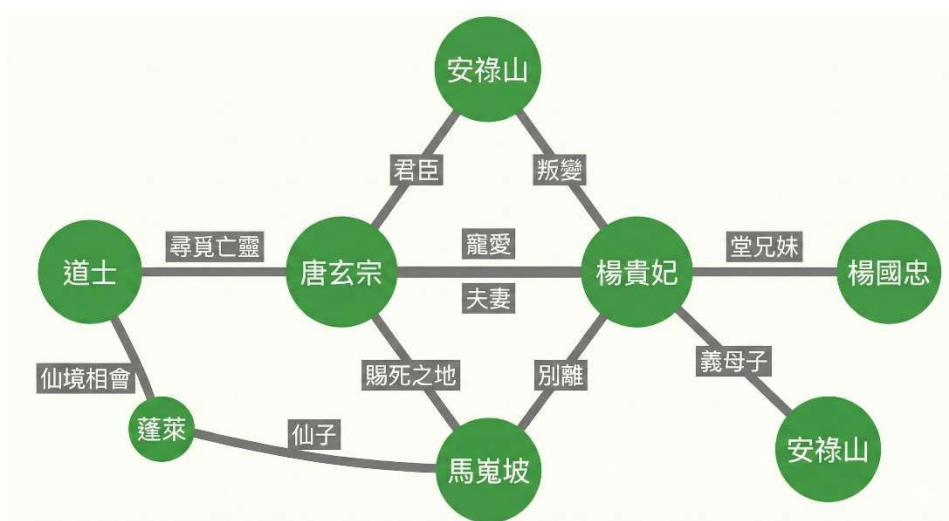
圖二 使用《我不要爸爸》作為範例分析之關係圖

按重要性排序

人物名稱	描述	人物行為
杰杰	五歲的小男孩，有明亮的大眼睛和可愛的小酒窩	經常說「我不要爸爸」來逃避責任 不願意按時睡覺、刷牙、收拾玩具和做功課 爸爸住院後每天去醫院探望 學會自己做功課、刷牙、收拾玩具 主動學習洗碗
爸爸	杰杰的父親，一位溫柔有耐心的父親	經常和杰杰玩拋高高遊戲 溫柔地提醒杰杰該做的事情 因車禍住院成為植物人 在昏迷中能聽到杰杰的聲音 最終康復出院
醫生	照顧爸爸的醫生	告知爸爸可能成為植物人 預測爸爸可能不會醒來 與杰杰和爸爸的關係：醫生與病人及家屬

圖三 使用《我不要爸爸》作為範例分析之人物分析

4.1.2 古典詩詞分析：白居易《長恨歌》

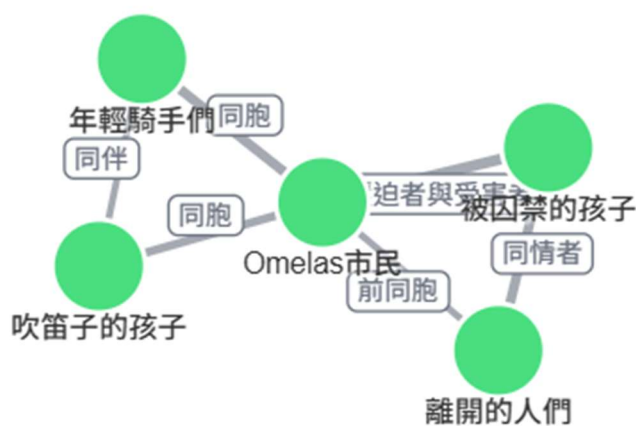


圖四 使用白居易《長恨歌》作為範例分析之關係圖

人物名稱	描述	人物行為
唐玄宗(君王/漢皇)	唐朝皇帝，楊貴妃的愛人	<ul style="list-style-type: none"> 多年尋找傾國美人 因楊貴妃廢朝政，整日觀看歌舞 安史之亂時倉皇西逃 在馬嵬坡被迫同意處死楊貴妃 思念楊貴妃，派方士尋找她的魂魄
楊貴妃(太真)	唐玄宗寵妃，絕世美人	<ul style="list-style-type: none"> 在深閨長大 被選入宮成為貴妃 回眸一笑驚艷眾人 在華清池沐浴 與君王朝夕相處
侍兒	宮中服侍的宮女	<ul style="list-style-type: none"> 扶起沐浴後的楊貴妃 與楊貴妃：主僕關係
道士/方士	奉命尋找楊貴妃魂魄的修道之人	<ul style="list-style-type: none"> 四處尋找楊貴妃的魂魄 上天入地尋找 與唐玄宗：受命尋找楊貴妃魂魄

圖五 使用白居易《長恨歌》作為範例分析之人物關係

4.1.3 英文文本分析：*The Ones Who Walk Away from Omelas*

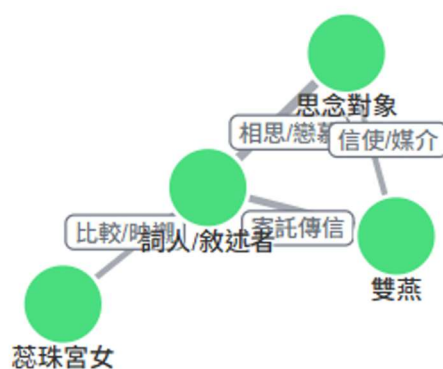


圖六 使用 *The Ones Who Walk Away from Omelas* 作為範例分析之關係圖

人物名稱	描述	人物行為
Omelas市民	Omelas城市的一般居民	參加夏季節慶活動 跳舞、騎馬、遊行 知道地下室孩子的存在 享受繁榮和快樂的生活 向8-12歲的孩子解釋這個‘交易’
被囚禁的孩子	被關在地下室的可憐孩子，約10歲	被鎖在骯髒的地下室中 生活在惡劣的環境中 乞求被放出來 漸漸失去說話能力 生活在極度痛苦中
離開的人們	選擇離開Omelas的人	參觀被囚禁的孩子 選擇離開城市 獨自向山區或北方走去 永不回返 與Omelas：選擇與整個社會決裂
年輕騎手們	參加賽馬的年輕人	準備參加賽馬比賽 安撫他們的馬匹 排列在起跑線上
吹笛子的孩子	9或10歲的獨自演奏者	獨自坐在人群邊緣 專注地吹奏木笛 沉浸在音樂中 與圍觀者：被欣賞但保持距離

圖七 使用 *The Ones Who Walk Away from Omelas* 作為範例分析之人物關係

4.1.4 宋詞分析：范仲淹「御街行」



圖八 使用范仲淹《御街行》作為範例分析之關係圖

人物名稱	描述	人物行為
詞人/敘述者	思念故人的主人公	裁剪冰綃製作衣裳 輕輕疊放布料 淡淡塗抹胭脂 打扮妝容 觀察院落景色
蕊珠宮女	故宮中的宮女，作為美貌的參照	被提及作為美貌的比較對象 與詞人：被用作美貌的對比
雙燕	院中的燕子，象徵傳情使者	在院中飛翔 無法傳達人言 與詞人：被寄予傳達相思之情的希望，但未能實現

圖九 使用范仲淹《御街行》作為範例分析之人物關係

4.2 使用者回饋與改進

在初步測試階段，我邀請了幾位同學試用系統，收集到寶貴的回饋：

正面回饋：

- 「視覺化效果很好，特別是關係圖，一眼就能看出文本重點」
- 「人物關係網絡圖對理解小說很有幫助」
- 「分析速度很快，比自己手動整理節省很多時間」

改進建議：

- 「希望能支援更長的文本，比如整本小說」
 - 「英文的詞形還原還不夠準確」
 - 「未來可以加入寫作輔助工具」
-

五、開發挑戰與解決方案

5.1 技術層面的挑戰

5.1.1 中文的複雜性

中文 NLP 相比英文有其獨特的挑戰，首先是分詞問題，中文沒有空格分隔，詞彙邊界模糊，其次是多義詞問題，同一個詞在不同語境下可能有完全不同的含義。其次是語句理解的功能。

解決方案：透過 Claude Code 的協助與學習，文本分析系統會先進行初步分詞，然後根據詞彙的前後關係進行調整。例如，「研究生命科學」可能被分為「研究生/命科學」或「研究/生命科學」，系統會根據上下文判斷正確的分詞方式，且透過 API 串接現有的成熟語言模型，比如 OpenAI、Google 或台灣本土的 AI 平台提供的 API，既能節省自行訓練模型的時間與成本，也能利用大廠訓練的大量中文語料優勢，快速讓 AI 具備中文斷句與理解能力。

六、學習收穫與反思

透過這個專案，讓我訓練抽象思維的能力，將複雜的文學分析需求，分解為可執行的程式邏輯，這個過程訓練了我的抽象思維能力，同時，問題解決的方法，遇到 bug 時，學會了系統性地排查問題、定位原因、測試解決方案，持續學習的習慣：技術發展迅速，需要不斷學習新的工具和方法，所幸 Claude Code 大大降低了學習門檻，讓我能快速嘗試新技術。

在 AI 逐漸掌管生活各個面向的今日，自己該如何妥善運用也成為思考很重要的一環，這是我將人文素養與技術能力結合的一次嘗試，也證明了文科學生也能成為技術創新的參與者，而不只是使用者，透過 Claude Code 這樣的 AI 工具，所有人都能夠跨越技術門檻，將自己的領域知識轉化為實用的工具，我相信，隨著 AI 技術的發展，會有更多跨領域的創新應用出現，讓各個領域的專業知識都能透過技術的力量，發揮更大的價值。感謝這門課給了我探索的機會，也感謝 Claude Code 讓這一切變成值得嘗試的事。

七、致謝

本專案的完成，離不開師長與同儕的協助與啟發，在此致上誠摯的謝意。

首先，特別感謝 王道維教授 提供本專案的原創發想。在專案執行過程中，教授持續與我進行深度討論，慷慨提供相關資源，並針對系統功能給予具體的改進建議，是引導本專案從構想到落地的關鍵推手。

感謝 羅曄哲同學 在技術實作上的強力支援。感謝你引領我認識 Claude Code 與 GitHub 等開發工具，協助我克服技術門檻；更感謝你在過程中不斷鼓勵我跨領域多方學習，讓我在面對程式挑戰時能保持信心。

此外，也要感謝 人社院接受訪談的同學們。感謝你們願意撥冗參與測試與訪問，讓我能找出文科學生在文本分析上的真實痛點，進而根據你們的回饋不斷改良系統，讓這個工具能真正貼合使用者的需求。

八、參考資料

1. Lane, H., Howard, C., & Hapke, H. (2019). *Natural language processing in action: Understanding, analyzing, and generating text with Python*. Manning Publications. (用於學習 python)
2. Sun, J. (n.d.). *Jieba: Chinese word segmentation module* [Computer software]. GitHub. <https://github.com/fxsjy/jieba> (用於學習中文斷句)
3. zfsang-zz. (n.d.). *CharacterGo* [Computer software]. GitHub. <https://github.com/zfsang-zz/CharacterGo.git> (用於學習編排人物關係)
4. Ramgpal. (n.d.). *Text analyzer: A web-based tool for performing basic text analysis* [Computer software]. GitHub. <https://github.com/ramgpal/TextAnalyzer> (用於學習文字分析)
5. Farmer, J. (n.d.). *Text-analysis-js* [Computer software]. GitHub. <https://github.com/jfarmer/text-analysis-js> (用於學習文字分析)
6. Xu, M. (n.d.). *Sentiment-chinese: Chinese sentiment analysis with dict* [Computer software]. GitHub. <https://github.com/shibing624/pysenti> (用於中文斷句分析)

7. Anthropic. (n.d.). *Claude Code documentation*.
<https://docs.anthropic.com/> (用於 vibe coding)
8. Chart.js Contributors. (n.d.). *Chart.js documentation*.
<https://www.chartjs.org/> (用於編排人物關係)
9. StoryMami. (n.d.). *我不要爸爸*. <https://www.storymami.com/story/i-dont-want-daddy/> (用於文本分析)
10. 白居易 (唐)。長恨歌。 <https://zh.wikipedia.org/zh-tw/%E9%95%B7%E6%81%A8%E6%AD%8C> (用於文本分析)
11. 范仲淹(宋)。御街行。 https://fanti.dugushici.com/ancient_proses/47794
(用於文本分析)
12. Le Guin, U. K. (1973). *The ones who walk away from Omelas*. (用於文本分析)