第四次讀書會

議題分析

指導老師：黃三益 教授

|  |
| --- |

第11組

N104220004 莊佳宜

N104220006 吳孟恬

N104220012 寧家榮

N104220013 倪家瑋

N104220014 蘇千華

N104220017 胡冠甄

N104220025 姜翰良

N104220026 陳嬿雅

目錄

[一、分析議題簡介 P.01](#_heading=h.2s8eyo1)~P.01

[二、工作流程設計](#_heading=h.17dp8vu) P.02~P.03

[三、](#_heading=h.3rdcrjn)LDA主題模型 P.04~P.09

四[、](#_heading=h.3rdcrjn)GuidedLDA主題模型 P.10~P.13

五[、](#_heading=h.26in1rg)字詞網路圖 [P.13~P.](#_heading=h.26in1rg)14

六[、](#_heading=h.lnxbz9)[結論](#_heading=h.1ksv4uv)P.15~P.16

一、分析議題簡介

議題發想：

畢業季是一個重要的人生里程碑，畢業季代表了學生們從一個階段進入另一個階段的轉變，對未來的期望和計劃等，讓人思考未來。

畢業季也代表著學生可能會對於未來的不確定性和未知的挑戰感到焦慮，並尋求建議和支持。他們可能會請教關於就業、進修、職業發展等方面的問題，或者分享他們的計劃和目標。

透過分析他們在網路上對於職業發展和就業問題的討論，我們可以瞭解到他們對於就業市場的觀察和理解、對於未來的期望和擔憂、對於職業發展的規劃和策略等。

瞭解他們在網路上討論的話題和內容，能夠幫助社會和媒體行業更好地了解畢業生群體的需求和興趣，從而提供更適切和具有吸引力的內容和服務。

利用工作流程系統PTT爬蟲找出今年五月一日起至五月二十日，近三週的留言進行下列分析與確認，具體操作如下：

分析期間：2023.5.1-2023.05.20

資料來源：PTT

使用工具：工作流程系統

資料筆數：1024

二、工作流程設計

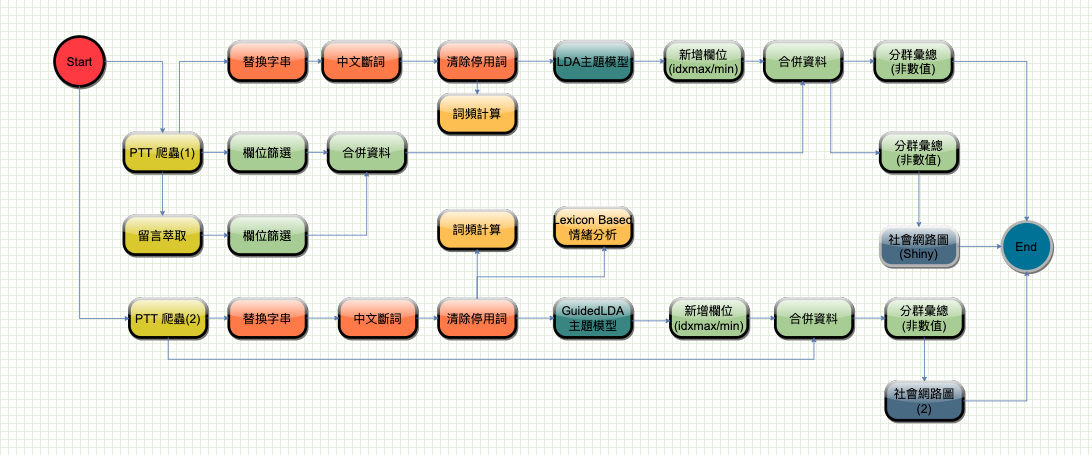
工作流程平台建立的工作流程為「第四次讀書會」



設計執行的步驟依序為：

1. 第一類讀入「PPT各版」資料（包含金融業、研究所、工作、台灣大學、公職、職場、高中、軟體工作、實習教師、教師、科技工作等），指定「畢業」關鍵字與設定排除字。利用LDA主題模型，進行分析。
2. 第二類不分版，設定「就業、職業發展、求職、實習、友情、回憶、社團、留學、交換、論文、畢業典禮、校友、未來目標、畢業典禮、謝師宴、教育資源」。利用GuidedLDA主題模型設定種子字，進行分析。
3. 使用替換字功能去除(\\n)、去除特殊符號字元、取代全形的分號、逗號、句號、問號以及驚嘆號、去除空白。
4. 針對不具分析效益的詞彙，使用清除停用詞設定停用的詞彙。
5. 利用LDA及GuidedLDA主題模型，進行分析。
6. 產出社會網路圖。
7. 用「情緒分析」產生視覺化圖形。

整理相關留言分析流程圖如下：



自定義停止詞：

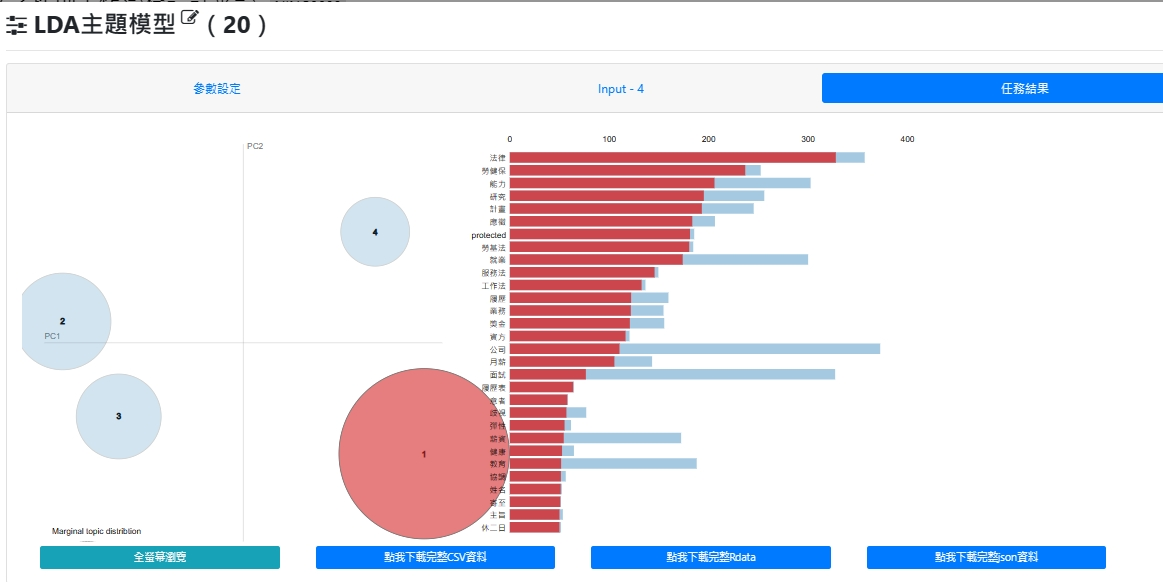
正面、有錢、成功、冷靜、娛樂、簡單、謝謝、信心、幫助、付出、優勢、容易、喜歡、興趣、願意、朋友、OK、安心、哈哈、典範、值錢、幫忙、不錯、笑死、感謝、認可、興旺、重要、可以、分享、自由、降低、垃圾、不要、問題、隨便、不爽、擔心、自我、沒用、白癡、無聊、抱歉、後悔、衝突、痛苦、不足、操心、羞辱、廢物、可怕、智障、可悲、自責、不好、缺點、處罰、八卦、愚蠢、降低、相關、時間、學生、通知、文章、資料、刪除、老師、經驗、學校、性別、email、內容、協助、助理、管理、配合、目前、確實、發文、人員、行政、平等、job、同意、現行、違反、碩士、課程、禁止、自負、文責、確知、張貼、科系、台灣、學習、題目、處理、方式、表示、員工、需要、福利、兩行、小時、主管、服務、障礙、遵守、身心、語言、今年、台大、年終、資訊、執行、知道、準備、系統、國際、現在、最後、學歷、未來、證明、台北市、主要、溝通、電話、發展、活動、上課、事項、機會、學士、作業、交辦、希望、軟體、一下、報名、網路、教授、電腦、之後、尤佳、教育部、證書、負責、專案、結構、小姐、訓練、教師、部分、具備、週休、文件、上班、團隊、需求、提供、自傳、選擇、以往、國立、註明、專任、身分、新聞、辦理、合作、專業、中心、二日、文書、分析、歡迎、教學、規定、建議、資格、刪文、條件、社會、科技、開發、薪水、工時、年齡、產業、規劃、設計、畢業生、技術、加班、為由、會員、予以、婚姻、調整、一年、職缺、mail、人才、進行、介紹、工會、補助、五官、容貌、性傾向、出生地、籍貫、黨派、宗教、狀況、操作、地點、同學、統計、考古題、認為、材料、台北、下午、思想、環境、高中、使用、熟悉、標題

三、LDA主題模型

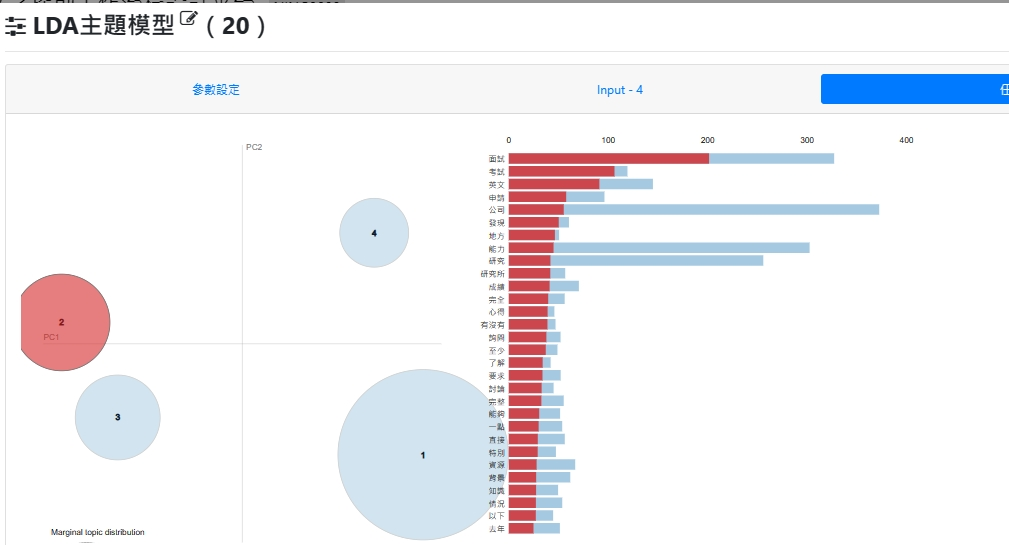
LDA主題模型可以協助分析畢業相關版中的文章、評論等文字資料，從中挖掘出潛在的主題，以進一步了解社群中的討論熱點與用戶需求。詞頻頻率上限調整為0.49。



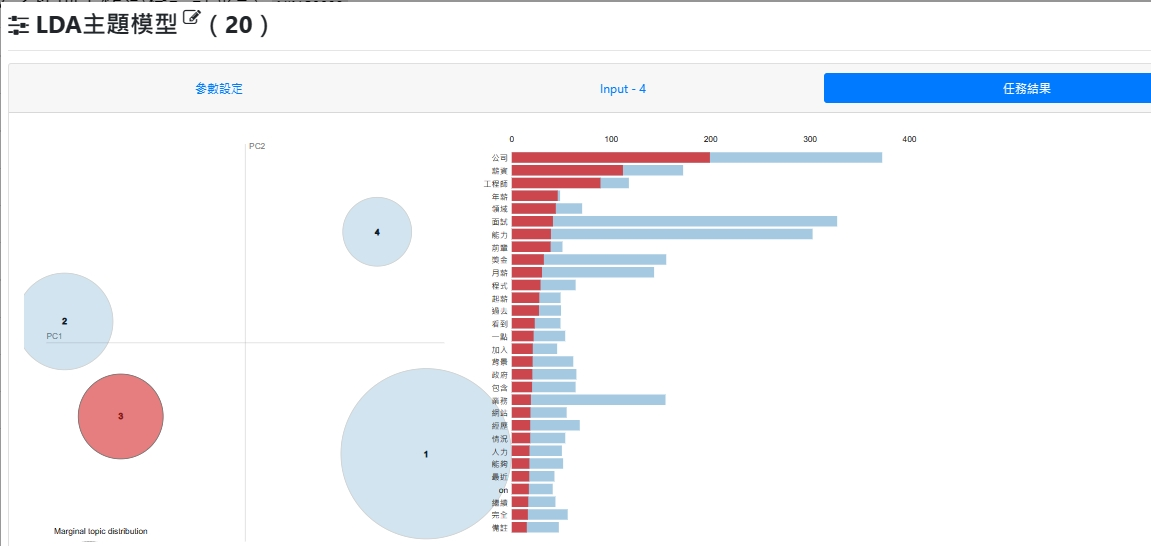
主題模型1.根據關鍵詞，可以將討論主題分為三類「就業法律和保障」（勞基法、勞健保、就業服務法、彈性、獎金、資方、公司、月薪等詞彙可能與就業條件、薪資結構有關）、「應徵與面試過程」（應徵、面試、履歷、履歷表、意者等詞彙指向求職過程，主旨、寄至等詞彙可能與求職信或面試通知等相關）、「教育與能力提升」（教育、研究、計畫等詞彙指向學習和專業能力提升有關，健康、協調、休二日等詞彙可能與工作與生活平衡、工作壓力管理等相關）。現在的求職者對勞工的權益和就業保障的關注度提高。



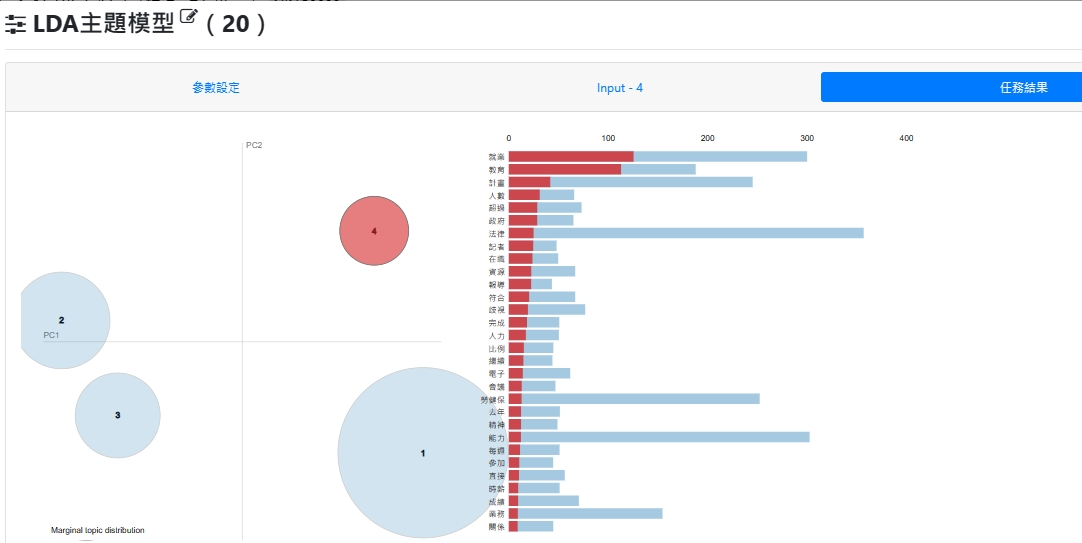
主題模型2.與升學或就業有關，可分為「考試和申請」（可能討論與如何提升能力、優化背景和知識、有效應對考試和申請等議題有關）、「面試和公司」（可能討論與面試技巧、求職歧視、企業文化等相關議題有關）、「研究和知識提升」（可能討論與國內外研究所升學等應考相關知識有關議題）。



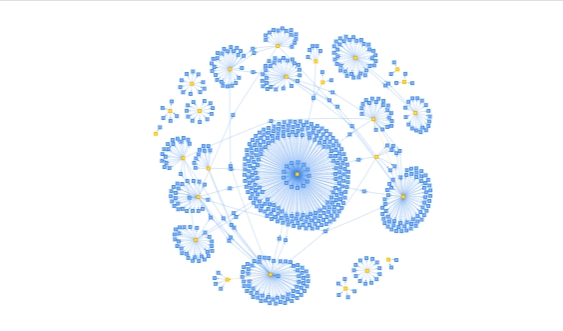
主題模型3.相關詞彙有「公司、薪資、工程師、能力、獎金、月薪、程式、起薪、背景、業務、網站、經歷」應該與討論工作職位、薪資待遇、技能要求、職涯發展機會、工作經驗等相關話題有關。一般對工程師的職業的起薪與薪資待遇感到好奇。

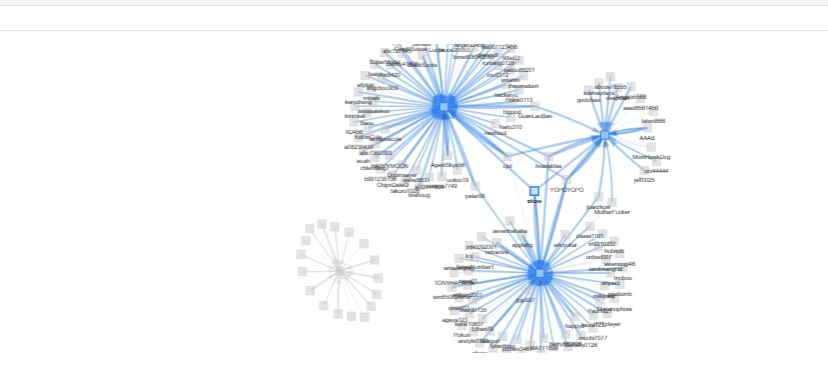


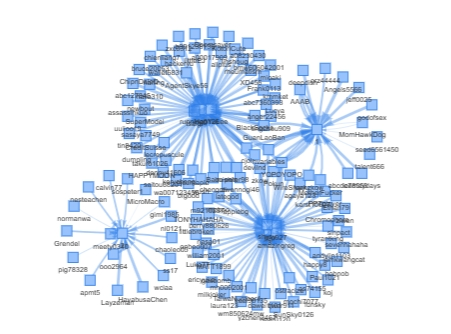
主題模型4.相關詞彙可以分成「就業與勞動力市場」（包含就業、人數、在職、人力、資源、符合、完成、比例、去年、能力、時薪、業務等詞彙）討論就業情況、勞動力市場的需求與供應、勞動法規與保障、薪資待遇、就業歧視等相關話題，「政府政策與公共服務」（主題涵蓋了超過、政府、法律、記者、報導、關係等詞彙）可能討論政府政策、法律法規、政府與媒體的關係、公共服務等相關話題。



LDA主題模型社會網路圖

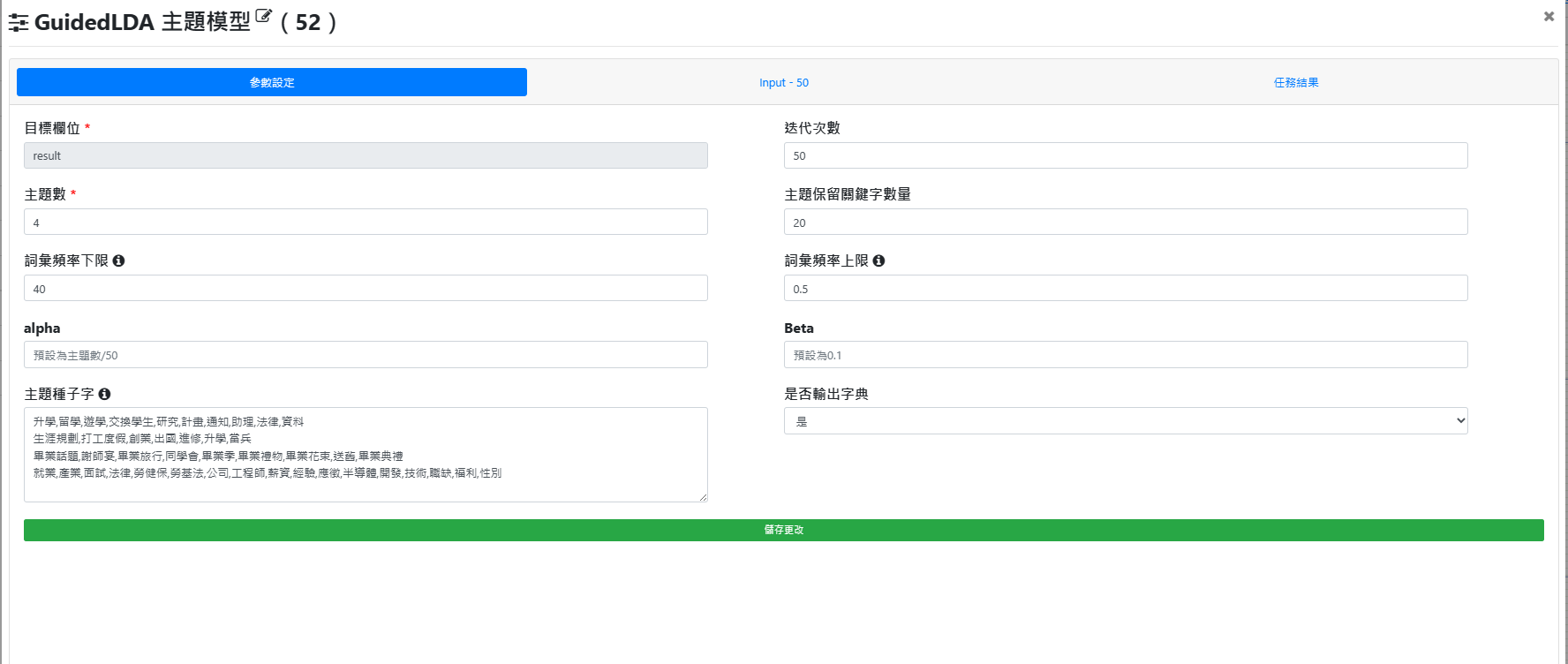




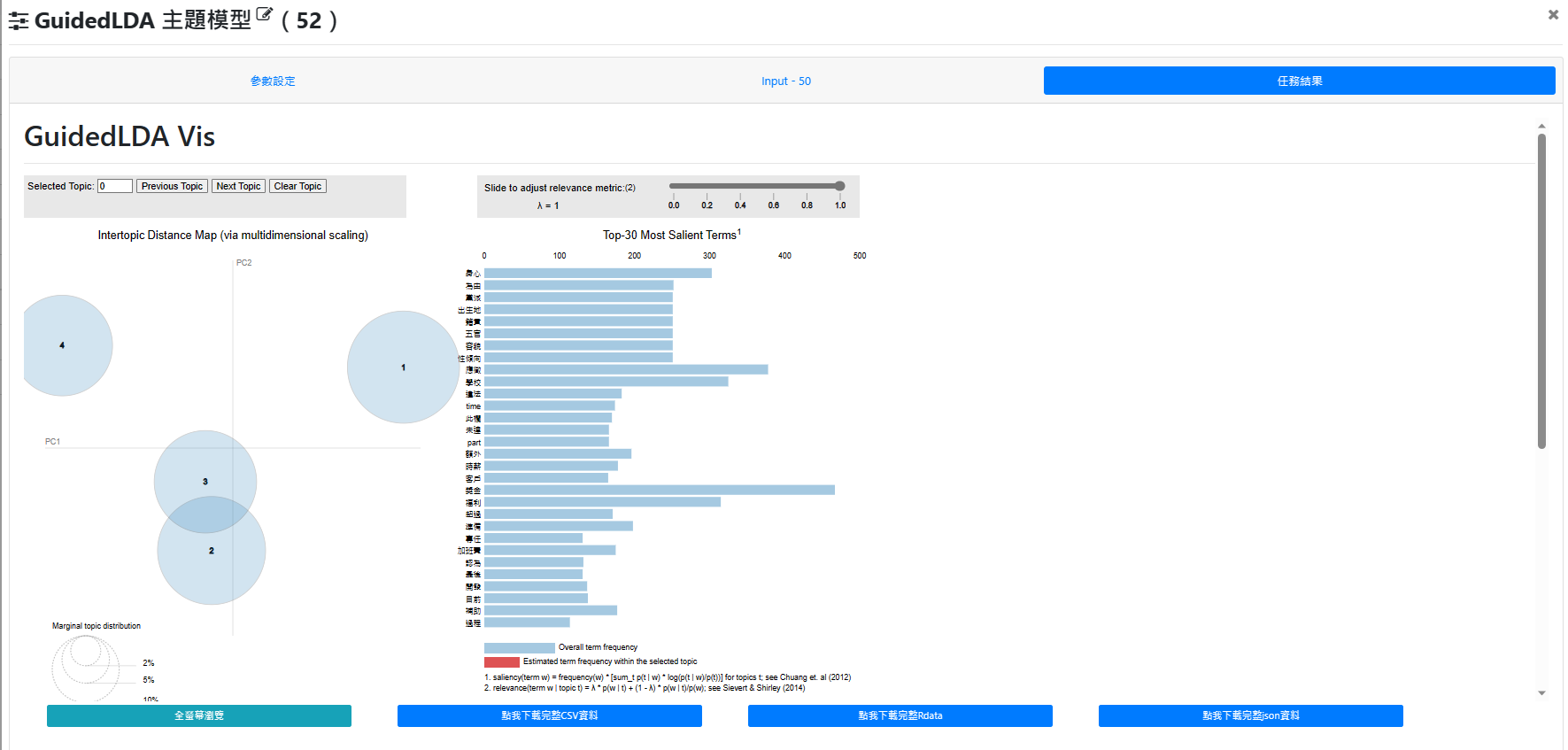


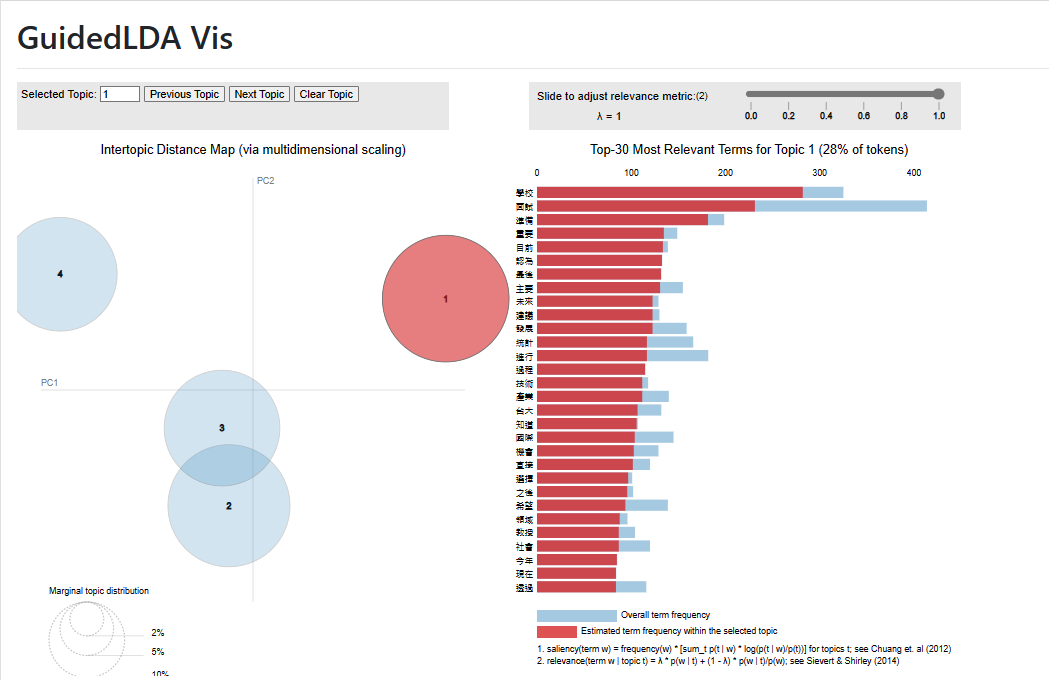
四、主題模型

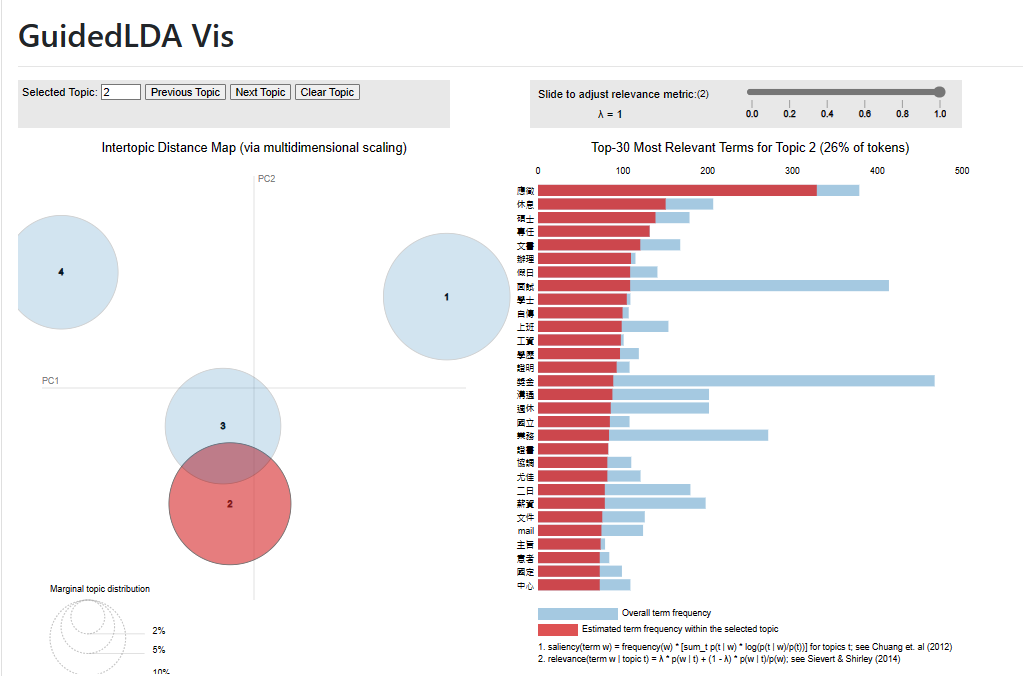
GuidedLDA種子字先從升學、就業、畢業、生涯規劃等四大面向進行設定。



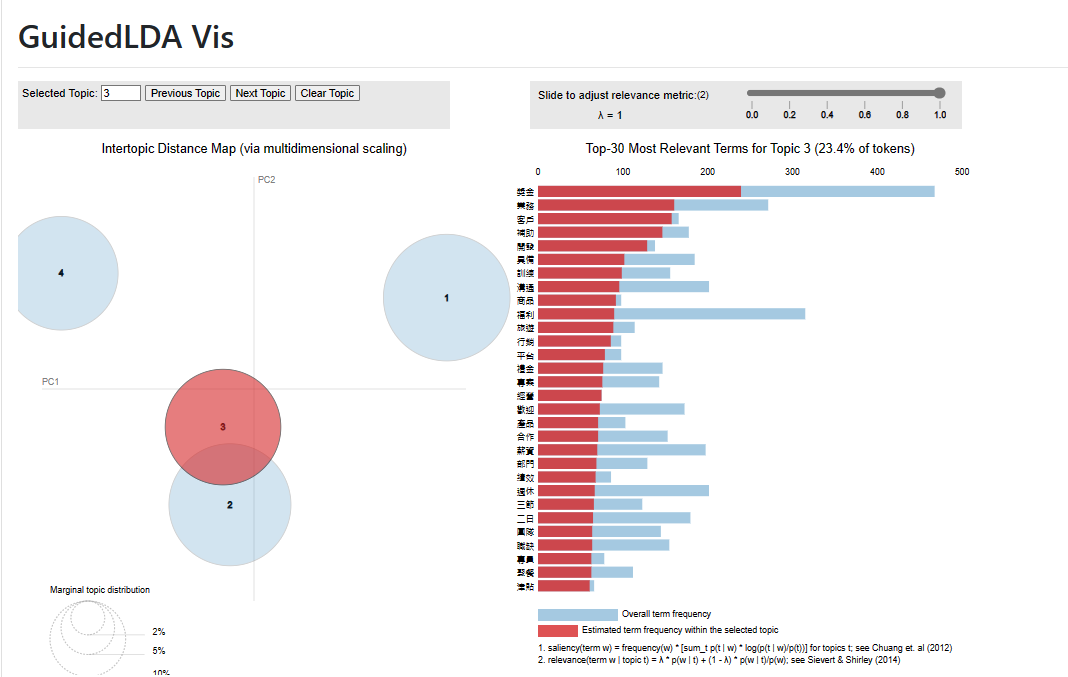
得出四個主題分布如：

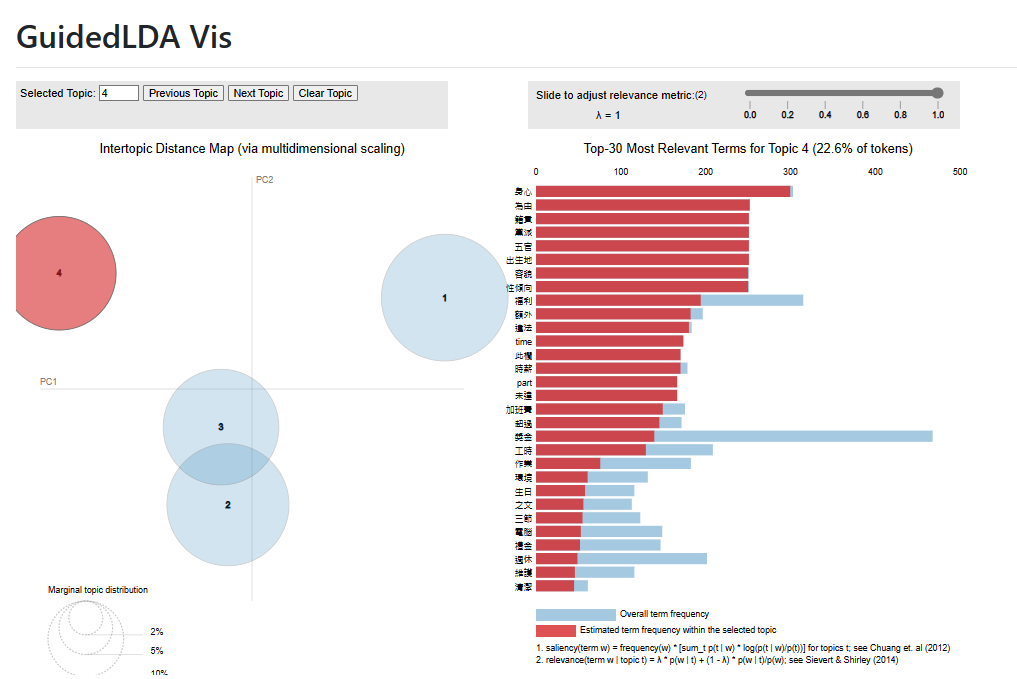


主題一：升學、出國留學、研究所等主題，「學校」與「面試」都是這些瀏覽者比較關注的，如:如何準備學校面試、未來的發展建議等評論。

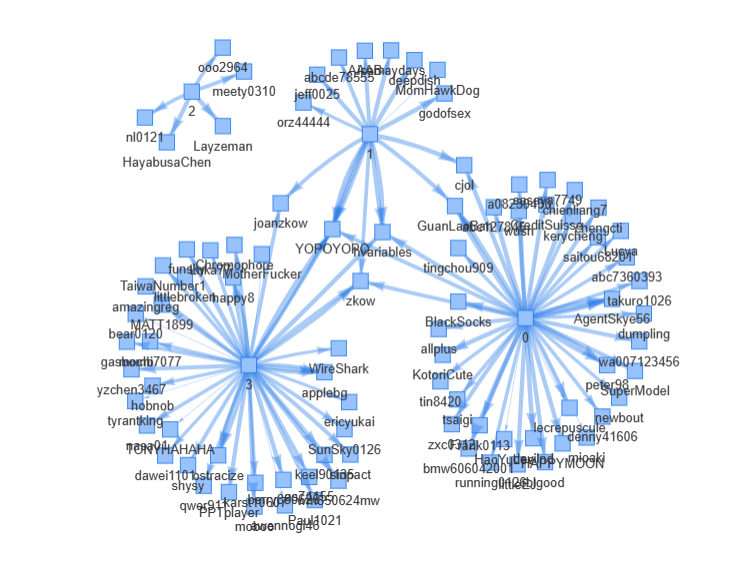
主題二：應屆畢業、就業相關主題，以應屆畢業生的規劃為主，有應徵、休息、碩士、面試，討論面向為畢業生尋找工作時的各種需求條件或福利。

主題三：勞工就業市場主題，以具社會經驗勞工為討論對象，出現的詞彙有獎金、福利、業務、周休等，討論大多是公司福利、獎金、輔助開發等，可以說是就業市場主觀的討論。

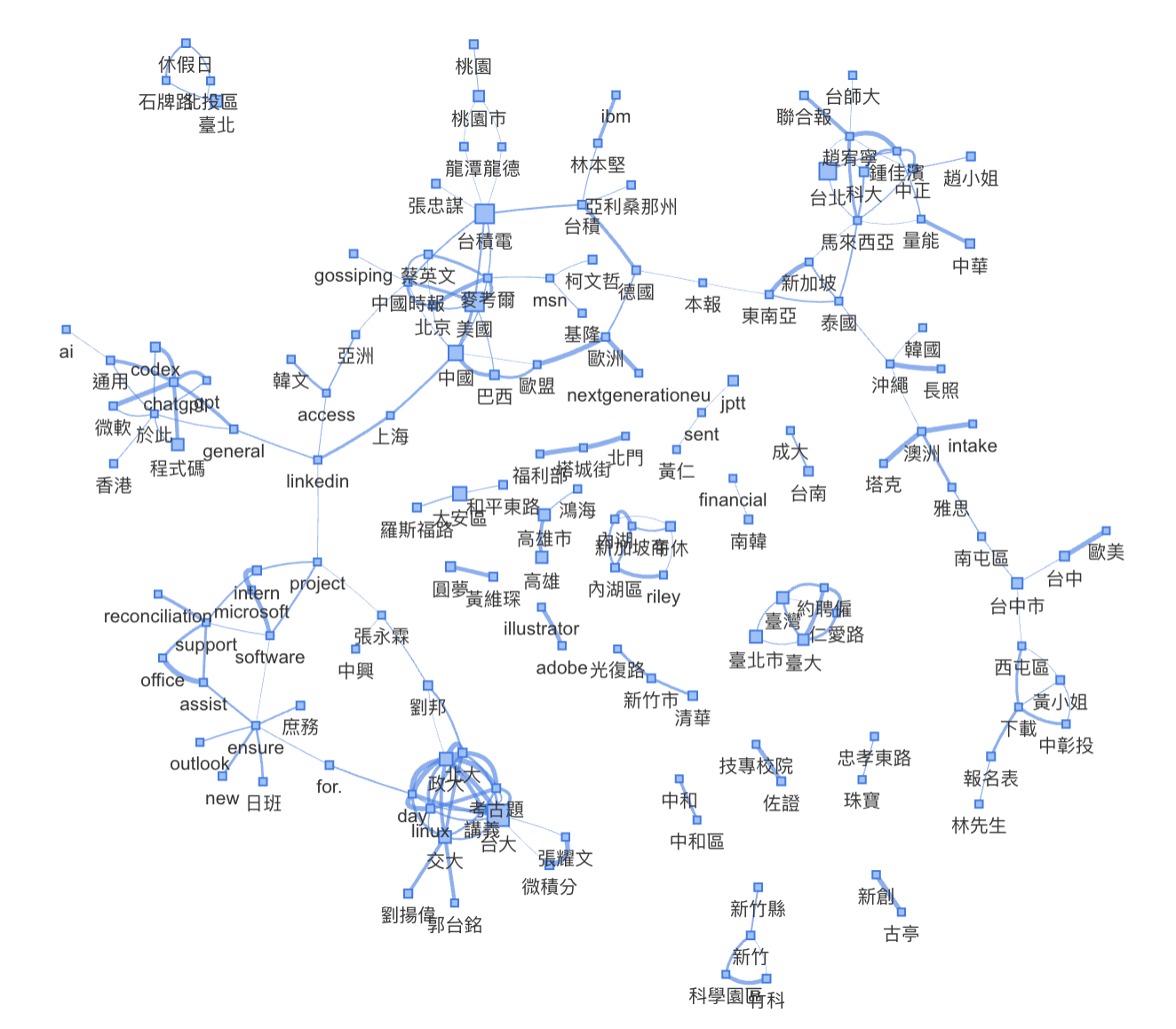


主題四：職缺主題，以就業市場客觀條件為討論，出現詞彙包含獎金、福利、身心等，討論多以就業服務法所規範的細項，顯見是開設職缺為主。

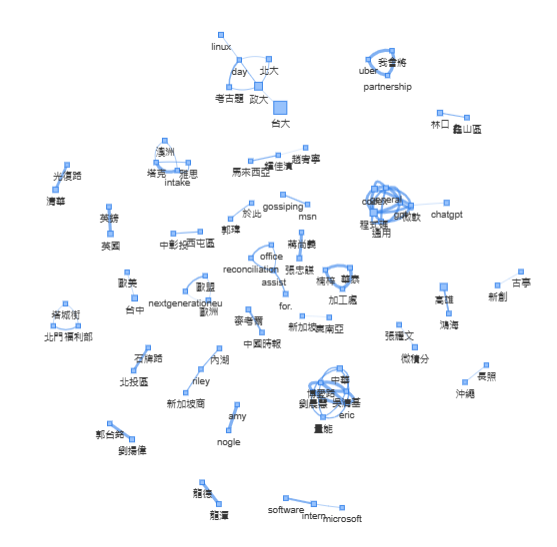
GuidedLDA主題模型社會網路圖



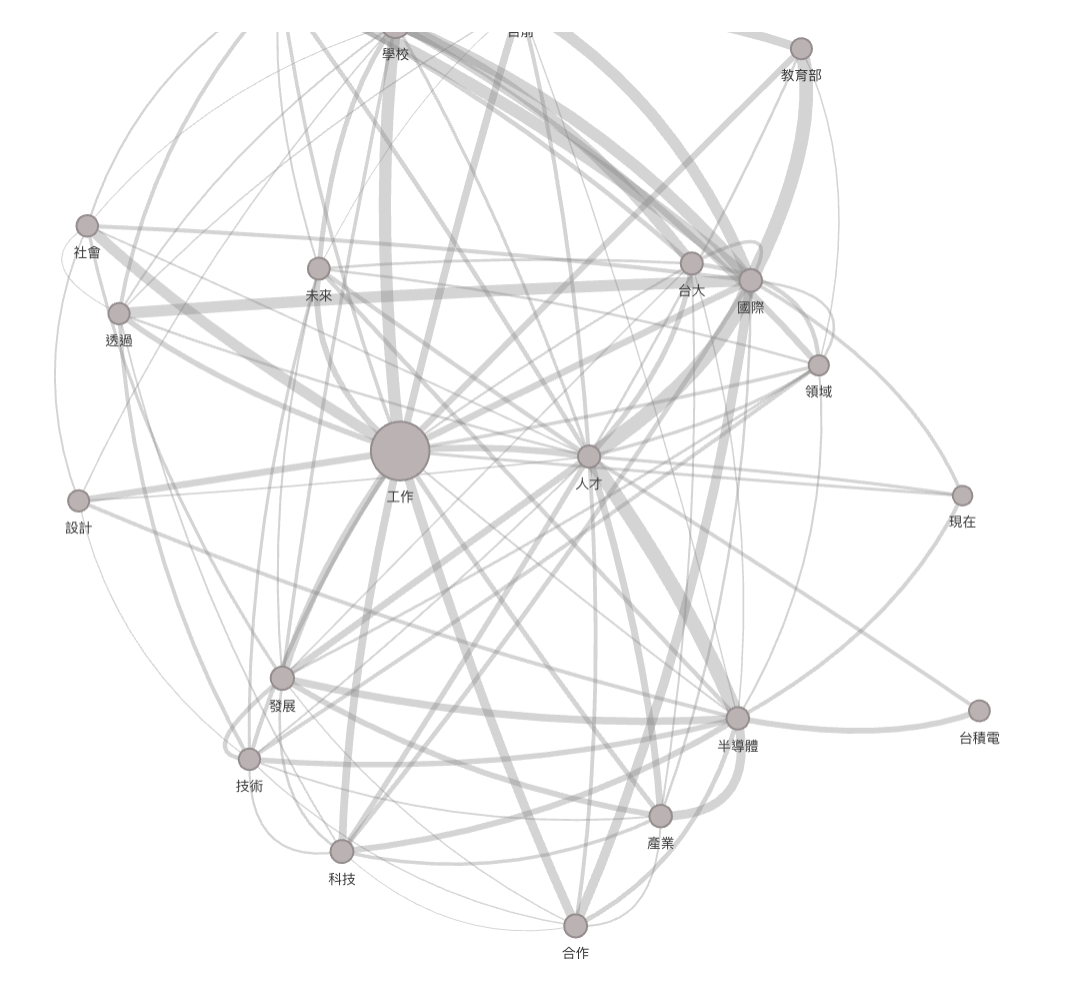
五、字詞網路圖＿篩選名詞



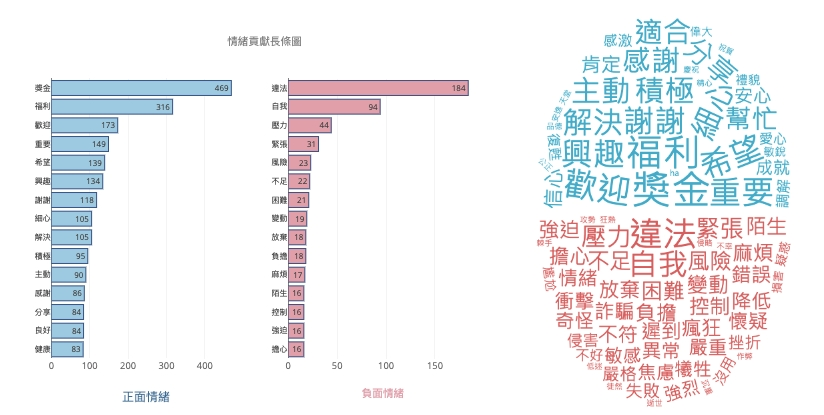
共現字詞網路圖



單中心網路圖-以工作為例



五、Lexicon Based情緒分析



正面情緒顯示：在畢業話題的背景下，成就感和滿足感：喜歡成就、福利重要、希望獎金等詞彙表明對取得成就和獲得獎勵的渴望，以及對自己努力的肯定和滿足感。

積極主動和改善意願：主動積極改善、解決興趣、細心關心等詞彙表明一種積極主動的態度，有意願不斷提升自己、解決問題、關心周圍的事物。

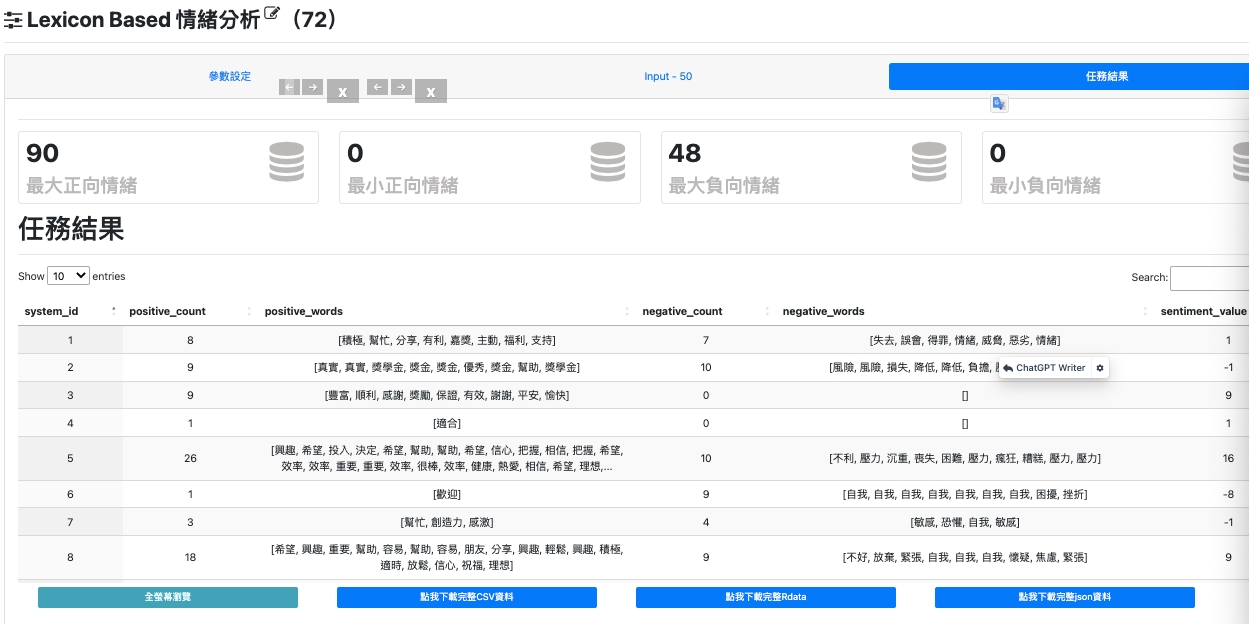
歡迎和慶祝：希望獎金、慶祝等詞彙表明期待獲得獎金，並對達到目標或重要事件的到來感到歡迎和慶祝。

負面情緒

就業壓力：強迫壓力、擔心添足自我風險、麻煩潔等詞彙表明對就業的壓力和擔憂。

職業選擇困難：放棄困難變動錯誤、控制降低、侵害不符、失敗強烈等詞彙暗示對做出正確職業選擇的困難和擔憂。

學業壓力和焦慮：犧牲多、嚴格焦慮、不好敏感異常等詞彙可能表示在畢業階段面臨的學業壓力和焦慮。



正面情緒大於負面情緒

表明青世代畢業生對於是否能夠順利畢業以及未來就業存在一定的恐慌，但又夾雜著未來可以賺錢，對職業發展有所期許的積極期待。

六、結論

我們此次運用GuideLDA主題模型從升學、就業、畢業、生涯規劃等四大面向進行探討，發現主題一的升學跟其他三個主題雖然有落差，但是主要在於人生未來的方向不同，但是從主題二到四，這三個都會有所關聯，可以從中找到未來面臨畢業的社會新鮮人、找尋新工作的職場老鳥、就讀在職專班想要轉職的職場人對於哪些條件、福利、薪酬的想法。

例如：

主題二的以應屆畢業生的規劃為主，有應徵、休息、碩士、面試，討論面向為畢業生尋找工作時的各種需求條件或獎金、福利等；主題三則是探討勞工就業市場主題，以具社會經驗勞工為討論對象，出現的詞彙有獎金、福利、業務、周休等，討論大多是公司福利、獎金、輔助開發等，可以說是就業市場主觀的討論；而主題四：職缺主題，以就業市場客觀條件為討論，出現詞彙包含獎金、福利、身心等，討論多以就業服務法所規範的細項，顯見是開設職缺為主。

以獎金這個詞彙為例：

獎金一再出現在主題二、主題三和主題四中，那麼這個模型分析的結果可能顯示福利或獎金在就業、畢業和生涯規劃等方面是一個重要的討論要素。這可能意味著獎金在這些主題中扮演著重要角色，無論是作為畢業生尋找工作時的考慮因素，還是公司福利和就業市場條件的一部分。

值得注意的是，獎金的出現頻率可能反映了在這些主題中討論獎金的相對重要性。

因此，根據模型分析，獎金的實際出現頻率比預估的頻率高，這可能意味著在升學、就業、畢業和生涯規劃等方面，這可能顯示人們對於獎金在工作環境中的價值和重要性給予了更多關注。

獎金的高頻出現可能反映了人們對於工作報酬的關注，以及對獎金作為激勵和獎勵的期望。這可能是在討論畢業生找工作時的薪酬條件、勞工在就業市場中薪資和福利的考慮，以及職缺中提供的獎金等方面。