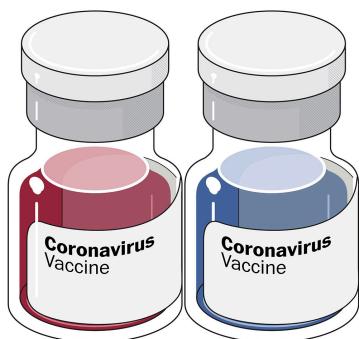


從PTT看疫情爆發前後大眾對疫苗的看法

第七組 Sasuke的水獺

林子茹、郭容羽、許祐璿、黃建智

2021.06.24



目錄

1. 研究動機
2. 研究方法與流程
 - a. 資料取得
 - b. 詞頻
 - c. 情緒分析
3. 研究結果
4. 討論及限制
5. 附錄
 - a. 分工表

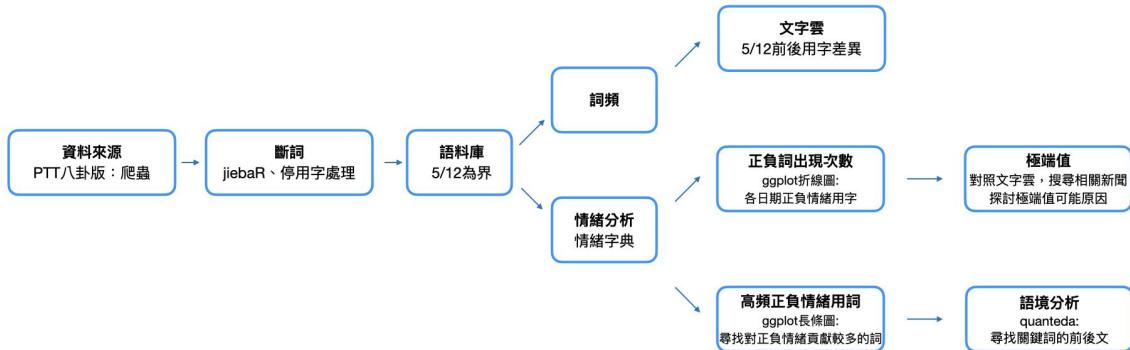
研究動機

Covid 19 從 2019年底開始肆虐全球，至今全球疫情仍處於緊繃狀態。台灣也在 2021 年五月首度進入全國三級警戒，社交距離、W F H、收看每日防疫記者會也成為台灣人的日常。防疫政策、醫療資源及疫苗施打都成為民眾的談論的焦點。其中疫苗與大眾的生活最息息相關，因無論確診與否，大部分的人最終都會接種，我們也因此將其作為我們的主題。

本組利用在課堂中學習到的文字分析及數據視覺化工具，將從ptt上獲取的語料加以整理、分析。探討疫情爆發前後，大眾對於疫苗討論度、討論面向及情緒的差異與改變。藉此，我們希望能夠透過這個分析來反映一班大眾對於疫苗的看法及態度。

研究方法與流程

研究流程表



a. 資料取得

以網頁爬蟲的方式爬取PTT八卦版上，標題含有「疫苗」二字，5/1~5/31之間的所有文章。

b. 詞頻

以jiebaR套件斷詞，並經過停用詞及自定義字典的處理，避免疫情相關重要詞彙被斷詞。接著經過dplyr套件的處理，將處理過後的文本整理成以日期及文字為單位的詞頻表，再以此詞頻表為根據繪製文字雲。

c. 情緒分析

以教授給我們的情緒字典作為基礎，將其與字頻表inner_join後計算每日的正負面用詞出現次數作為當日發文情緒的依據。使用ggplot繪製折線圖後，對於情緒特別的激昂的日子再進一步使用文字雲分析，找出當天的關鍵字以推論當天情緒用字較頻繁的可能原因。

另外，我們也情緒用字詞頻表找出前30的情緒用字，並篩選一些詞彙，使用quanteda套件建立的corpus功能來找尋關鍵詞的前後語境。

討論與結果

1. 疫情爆發前與爆發後討論熱度落差:
 - a.爆發前(5/1-5/12) 405篇相關文章
 - b.爆發後(5/13-5/31) 7170篇相關文章
2. 詞頻分析(文字雲分析)

爆發前：重點聚焦「副作用」



(圖一)

爆發後：重點聚焦「打不打的到疫苗」

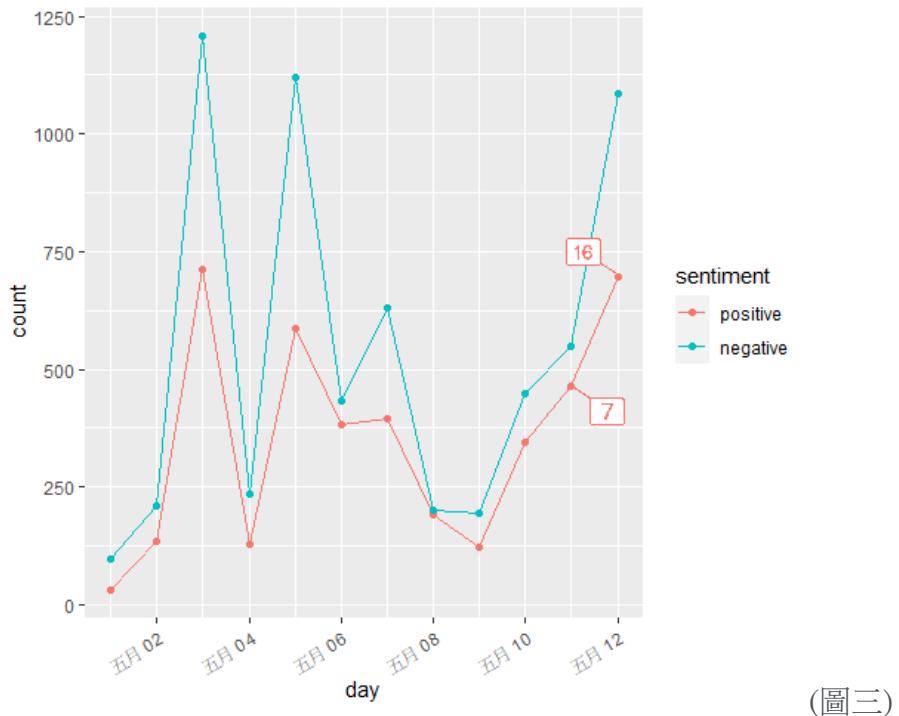


(圖二)

觀察文字雲可看到，疫情爆發前，疫苗的「副作用」詞頻比例高，顯示疫情大規模擴散前，副作用為民眾討論核心之一；反觀疫情爆發後，民眾討論核心轉為關注如「國產疫苗」、「BNT疫苗」、「AZ疫苗」各式疫苗，且「採購」及「買」的疫苗相關取得動詞也成為討論度高的詞彙。「施打」及「接種」亦為大眾所關注的焦點，反倒副作用相關討論減少，可從以上觀察大致推測出，民眾在疫情爆發後，更在乎的是「打不打得到疫苗」。再加上時事新聞也常在報導台灣疫苗數量不足的消息，因此，可發現疫情爆發後，民眾開始討論起國產疫苗，且討論頻率高，與疫情爆發前呈現強烈對比。

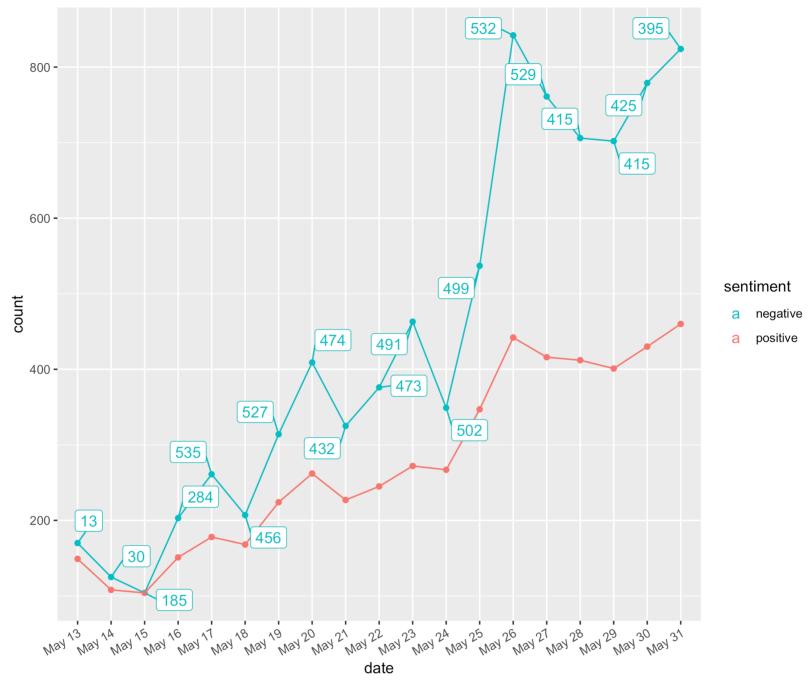
3. 情緒分析

疫情爆發前：負面情緒較高（圖中數字表示當日確診個案）



(圖三)

疫情爆發後：負面情緒仍是高於正面



(圖四)

將情緒分析的結果以折現圖表示，輔以每日確診個案數，讓我們可以對照確診個案的多寡與當日情緒用詞的狀況有無相關性。然而，根據折線圖的走勢，上述兩者無明顯相關性。另外，為了瞭解極端值的可能原因，我們參照當日的文字雲，再進一步搜尋

新聞，探討當時的時空背景是否與極端值生成有關。前期，以5/3為例（圖五），實際探究當天發生的事件後，得知5/3當天政府開始實施1至3類公費對象的「同住者」開放施打疫苗政策，可能是因為關於此事件的討論較多，其中又帶有許多疫情相關詞彙，因而導致負面情緒較高。再者，爆發後以5/26為例（圖六），疫情爆發後的高峰值，文字雲中負面詞彙的比例相對不高，反而出現較多專有名詞，如：民進黨、蔡英文、BNT、AZ、中國等，推估出這些詞彙也許是民眾不滿情緒的起源。



(圖五)



(圖六)

4.高頻正負詞統計

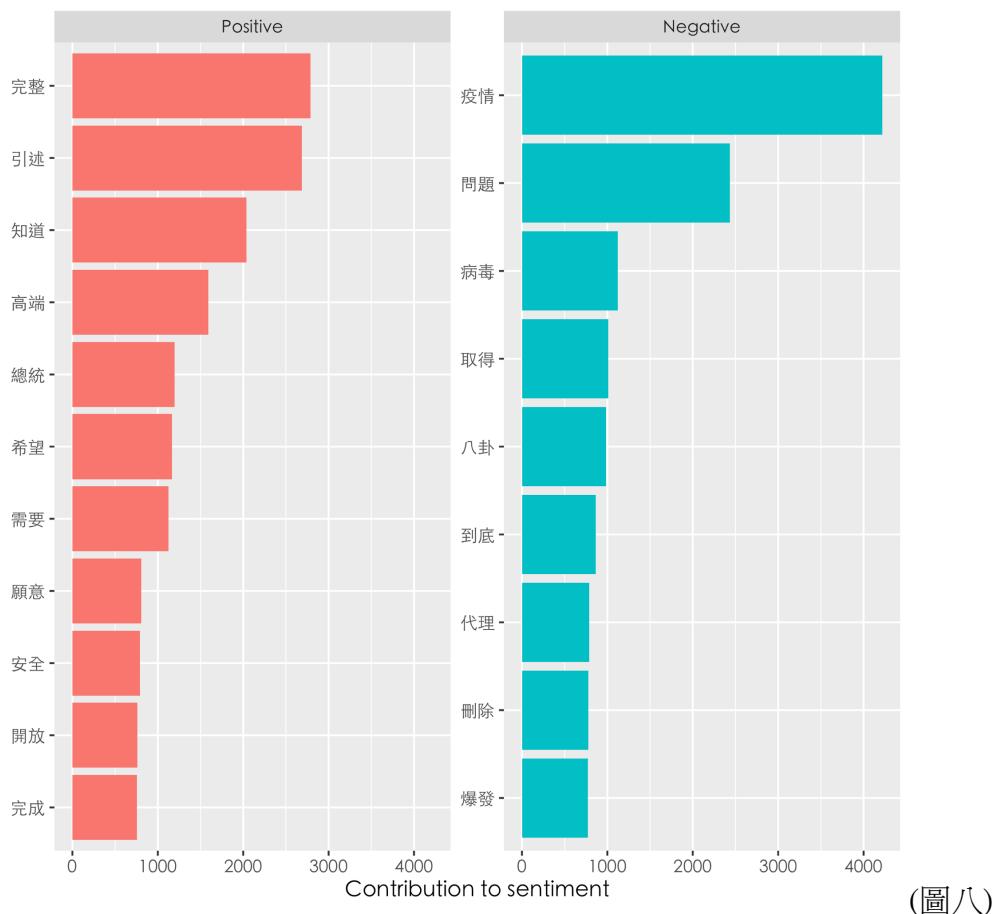
在看完圖表的情緒分析之後，我們進一步統計哪些詞對於情緒貢獻較高，以(圖七)與(圖八)呈現。

疫情爆發前：



(圖七)

疫情爆發後：



(圖八)

透過這兩張圖，我們可以找出貢獻較高的情緒用詞，進而分析我們感到有興趣的高頻詞，分析結果如下：

4 - 1. 疫情爆發前：分別以「副作用」、「感染」、「自由」、「嚴重」為例

副作用：

docname	from	to	pre	keyword	post	pattern
content4	42	42	，受到國際傳出血栓	副作用	而乏人問津，如今國內	副作用
content21	126	126	要打疫苗的當	副作用	大於 covid19 發生率的	副作用
content21	313	313	差嗎確實 AZ 有	副作用	血栓也確實英國澳洲	副作用
content21	432	432	也沒有血栓之類的嚴重	副作用	打這篇只是希望大家	副作用
content32	144	144	了出問題只要有一個人	副作用	強烈一點你他媽的就	副作用
content50	14	14	中心回答各國沒有因為	副作用	而暫停使用 AZ 疫苗	副作用
content69	59	59	美國的疫苗一樣有	副作用	一樣有人會不舒服	副作用
content72	104	104	挨個幾針可能有些	副作用	總比中標傳染給你身邊	副作用
content72	146	146	可能很衰疫苗打下去	副作用	出事會是自己明顯	副作用
content106	64	64	很多但相對的 AZ	副作用	也比較小另外對	副作用
content108	47	47	國外的疫苗更不會有	副作用	而且效果一樣好那麼	副作用
content126	95	95	統計結果是 AZ 嚴重	副作用	比輝瑞跟莫	副作用

常與副作用一起出現的高頻字為「AZ」、「血栓」、「不舒服」，可以看出民眾疫情爆發前，對於副作用的討論度聚焦在AZ上，呼應當時台灣只有AZ疫苗可施打的狀況。另外，也與第一部分文字雲的結果相呼應。

感染：

docname	from	to	pre	keyword	post	pattern
content2	110	110	病例，其中 6 例	感染	源不明。台大醫院安全	感染
content2	179	179	會遞減。振興醫院	感染	科主任李明政	感染
content3	45	45	本土案例暴增且	感染	源不明而嚴峻，	感染
content11	12	12	打疫苗造成多起社區	感染	源不明現在是誰要	感染
content13	119	119	完整新聞標題：：	感染	危機台灣卻沒疫苗吳	感染
content13	165	165	出現 6 例無法追蹤	感染	源病例，讓台灣	感染
content13	519	519	因為：台灣幾乎沒有	感染	的歷史，台灣人	感染
content13	542	542	就更容易形成：	感染	危機，就是前面做	感染
content14	29	29	.完整新聞標題：	感染	危機台灣卻沒疫苗吳	感染
content14	73	73	出現 6 例無法追蹤	感染	源病例，讓台灣	感染
content14	412	412	，因為台灣幾乎沒有	感染	的歷史，台灣人	感染
content14	434	434	，就更容易形成	感染	危機，就是前面做	感染
content21	39	39	打疫苗風險還大過	感染	：至今還是整個 FB	感染

常與感染一起出現的高頻字為「醫院」、「社區」、「感染源不明」，可以看出隨著5月中旬確診案例數得出現，找出感染源是民眾主要擔心的重點。

自由：

docname	from	to	pre	keyword	post	pattern
content45	3	3	台灣 太	自由	了， 明明 有 疫苗	自由
content144	69	69	人家 咩 每個人 都有 他的	自由	不喜歡 他 講 的 話	自由
content163	121	121	* 例如 蘋果 日報 、	自由	時報 (請 參考 版	自由
content164	91	91	* 例如 蘋果 日報 、	自由	時報 (請 參考 版	自由
content172	177	177	外國 疫苗 。 果然 民主	自由	之 都 ， 真 新聞	自由
content172	293	293	啊， 還是要 學 民主	自由	的 台灣 ， 不喜歡 一樣	自由
content183	35	35	了 好像 年底 就 能	自由	出國 玩 惹 打 不	自由
content205	48	48	國會 山 都能 是 美國	自由	民主 的 展現 了 。	自由
content251	418	418	表示 「這就是 民眾 的	自由	選擇 。」 5 .	自由
content260	15	15	1. 媒體 來源：	自由	時報 2. 記者 署名	自由
content288	196	196	沒有 任何 一天 能夠 完全	自由	行動 (所有 時間 都	自由
content325	8	8	. 媒體 來源 ::	自由	時報 :2. 記者	自由

常與**自由**一起出現的高頻字為「日報」、「媒體」、「時報」，此類報章雜誌的出現，可知道PTT上轉載許多網路新聞。由於「**自由**」一詞為正面用詞，因此成為高頻的情緒用字，也造成情緒分析上的誤判。

嚴重：

docname	from	to	pre	keyword	post	pattern
content3	395	395	企圖， 對 疫情	嚴重	的 友邦 更要 不吝 適時	嚴重
content19	331	331	到底有 多少 我們 不知道 的	嚴重	程度 阿 ? -	嚴重
content21	372	372	想想 有些 疫情 更	嚴重	的 國家 連 疫苗 都沒有	嚴重
content21	431	431	所知 也沒有 血栓 之類 的	嚴重	副作用 打 這篇 只是 希望	嚴重
content22	145	145	想起 去年 美國 疫情	嚴重	時， 口罩 不夠 ，	嚴重
content69	70	70	發燒 頭痛 漚身 無力 還有	嚴重	過敏 的 案例 研究 數據	嚴重
content69	91	91	五萬 分 之一 輝 瑞	嚴重	過敏 的 比例 是 十萬	嚴重
content72	96	96	全部 打完 變成 台灣 最	嚴重	就好 笑了 挨 個 幾針	嚴重
content81	3	3	AZ 的	嚴重	反應 機率 是 百萬 分	嚴重
content81	35	35	有 350 人 會 出現	嚴重	反應 那 不打 的 話	嚴重
content86	13	13	最近 疫情 好像 變	嚴重	有點 擔心 之後 大家 會	嚴重

常與**嚴重**一起出現的高頻字為「疫情」、「發燒」、「AZ」，這些相關的討論，除了疫情本身以外，民眾也關心嚴重副作用與過敏反應。

4-2. 疫情爆發後：分別以「問題」、「安全」、「希望」、「取得」為例

問題：

docname	from	to	pre	keyword	post	pattern
4	196	196	民回應政府決策	問題	醫師太遠表示國產疫苗	問題
8	22	22	病症疫苗放到過期假設性	問題	買多會過期來作真實情況訊息	問題
11	25	25	記者會廠商掛保證產品	問題	遮住名字還以法請問	問題
15	83	83	疫苗年前怕國光出	問題	建議國際採購疫苗奇怪	問題
15	92	92	疫苗奇怪年高端聯亞出	問題	在野腦袋正常在野不敢	問題
16	219	219	討厭民進黨股票問題疫苗	問題	民進黨政府做民進黨擋	問題
16	299	299	變BNT中國製疫苗	問題	國外疫苗買不到中央統一	問題
26	366	366	進來疫苗引發不良反應副作用	問題	同樣國家預防接種救濟系統	問題
31	16	16	陳建仁跑出國產疫苗	問題	安慰劑知道場合炒	問題
32	75	75	民眾們不爽關說黃委員	問題	陳市長反應問題醫院端能	問題
32	78	78	黃委員問題陳市長反應	問題	醫院端能關說問題知道	問題

常與**問題**一起出現的高頻字為「廠商」、「高端」、「政府機關相關詞彙」，這邊多半談論國產疫苗的問題，呼應文字雲，可以顯示國人對疫苗的重視提高。同時，由於疫情升溫，對於政府防疫政策的關注度也上升。

安全：

docname	from	to	pre	keyword	post	pattern
4	376	376	疫苗試驗出爐證實有效性	安全性	進行交貨一切順利加速第三期	安全
15	4	4	蔡英文幾天國家	安全	會議召開衛生署長楊志良	安全
15	48	48	藥物增加台灣克流感	安全	存量恐國光疫苗臨床	安全
23	179	179	感不足蔡英文幾天國家	安全	會議召開衛生署長楊志良	安全
23	223	223	藥物以增加台灣克流感	安全	存量恐國光疫苗臨床	安全
24	98	98	一線醫護搶疫苗剝奪	安全	屏障黃昭順接受蘋果	安全
24	279	279	第一線醫護人員搶疫苗剝奪	安全	屏障S566uVe僧多粥少疫苗第一	安全
26	562	562	實驗相關資料證明疫苗	安全性	品質藥效達規定原廠	安全
26	622	622	一紙公文進口確保藥品	安全	有效品質常規審查程序	安全
26	853	853	執行完試驗基本藥物	安全	有效性資料依舊少發生	安全
30	28	28	認可生產證明抽樣檢測	安全	無誤就讓人民能夠施打	安全
41	59	59	台灣他媽的疫苗打啊疫苗	安全	沒事幫忙靠北台灣政府	安全

常與**安全**一起出現的高頻字為「藥品」、「疫苗」、「有效」，看出討論趨勢轉為疫苗後，其相關藥物的安全性也成為討論重點之一。

希望：

docname	from	to	pre	keyword	post	pattern
7	130	130	防疫 努力 接觸 多種 管道	希望	協尋 購買 世界衛生組織 認證 疫苗	希望
24	137	137	施打 剩餘 劑量 彙 整後才	希望	剩下 疫苗 浪費 妥善 利用	希望
24	402	402	引發 質疑 低調 道歉 表示	希望	防疫 要減 少 車程 跑來跑去	希望
26	73	73	國外 採購 疫苗 民間 企業	希望	採購 疫苗 全世界 疫情 嚴重	希望
26	1259	1259	疫苗 提升 施打 效率 指揮中心	希望	中央政府 統 採購 分配 此法	希望
35	235	235	發現 家鄉 有難 運送 回來	希望	鴻海 集團 創辦人 郭台銘 拿到	希望
40	68	68	佛光山 所屬 國際 佛光 表示	希望	捐贈 萬至 50 萬劑 疫苗	希望
44	199	199	留得 住嗎 報平安 擔 心	希望	沒事 海邊 有沒有 消波塊 cdn2	希望
44	258	258	若解 盲 疫苗 成功 合約	希望	政治 政治 科學 科學 科學	希望
57	143	143	中央 疫苗 不足 恐慌 急	希望	疫苗 進來 提醒 做 製造	希望
60	148	148	中央 疫苗 不足 恐慌 急	希望	疫苗 進來 提醒 注意 做	希望
79	45	45	綠 蛆 偷偷 打嗎 真	希望	郭 董能 狠狠 綠 蛆	希望

常與**希望**一起出現的高頻字為「**政府機關**」、「**民間企業**」，再次呼應文字雲，民眾把討論焦點放在：希望儘早取得疫苗，而非先前的副作用上。

取得：

docname	from	to	pre	keyword	post	pattern
4	80	80	加速 三期 臨床 試驗 進行	取得	國際 認證 目標 網友 提問	取得
4	309	309	疫苗 做 完 第三期 人體試驗	取得	緊急授權 上市 謝思民	取得
4	320	320	回 應 輝瑞 2020 年	取得	美國 緊急授權 莫德納 2020 年	取得
4	326	326	美國 緊急授權 莫德納 2020 年	取得	美國 緊急 授 權 謝	取得
4	339	339	說明 時間 完成 第三期 試驗	取得	授權 所謂 緊急 授 權	取得
4	413	413	聯繫 NIH 簽署 授權 合約	取得	新冠肺炎 候選 疫苗 相關 生物	取得
7	123	123	慈濟 政策 允許 慈濟 儘速	取得	投入 防疫 努力 接觸 多種	取得
7	147	147	業 流程 多方 努力 中	取得	疫苗 合法 進口 確定 因素	取得
26	227	227	署 申請 疫苗 查驗 登記	取得	藥物 許可證 藥事法	取得
26	269	269	代理商 購買 代理商 透過 付費	取得	代理權 原廠 設 代理商 購買	取得
26	280	280	者則 機會 原廠 洽談 原廠	取得	相關 授權 資料 食藥署 申請	取得
26	356	356	責任 則在 政府 政府 審核	取得	藥證 疫苗 單位 買 進來	取得

常與**取得**一起出現的高頻字為「**緊急授權**」，取得一詞大多提及取得疫苗緊急授權，也揭露出疫情爆發後，台灣疫苗短缺的情況。

結論與展望

1. 限制

情緒分析無法精準探究詞彙在文章中的真實意思，以高頻情緒詞「自由」、「高端」為例。自由：雖然自由被放在正面詞彙，但在這次文章中卻只是一個專有名詞，高端：呼應上述的限制，高端在中國用語為正面的形容詞，在這種一詞多義的情況下，單看圖八的長條圖，我們常常無法去知道這些情緒詞彙在前後文真正的意義。

2. 總結

雖然情緒分析存在一定的限制，但無論疫情爆發前後，民眾的負面情緒皆是負面多過正面，另外也可以看出討論重心的轉移，從疫苗副作用到打不打得到疫苗，相信民眾的擔憂確實存在，期許疫情盡快趨緩。

附錄

1. 分工表：

- 黃建智:爬蟲、書面報告、口頭報告 6
- 林子茹:資料處理、製圖、書面報告 6
- 許祐璿:爬蟲、資料處理、製圖 6
- 郭容羽:靜態網頁製作、文字撰寫 6