

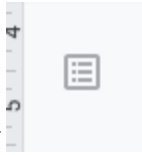
踏出同溫層，多立場資訊平台

目錄：

- 一、開發流程
- 二、系統開發細節
- 三、發想
- 四、會議記錄

黑客松短講簡報

開啟本文件大綱的方法：(方便你到該區域，討論閱讀)



左側有 的符號，按下去就會打開大綱，

會議：日常討論telegram(邀請常常參與任務的夥伴) + 線上開會<https://meet.jit.si/> + UI 呈現
figma + 程式 GitHub

加Line: eliw1006

或Telegram @eli_h1010 or https://t.me/eli_h1010 (eli會提醒你相關事宜)

每週三

晚 21:30-22:00。當天21:28在此放上會議連結

<https://meet.jit.si/fdsjlkfdklkafjdsjfkdsakfjksd>

共筆連結：(資料會在本文件更新，偶爾在slack做更新)

1.Slack: g0v-tw.slack.com #neutral-info

<https://app.slack.com/client/T02G2SXKM/CUVD7R29W>

沒有slack帳號的，要先註冊與登入(網頁版要填入g0v-tw)

一、開發流程

1. 說明：

- a. 確定需求, 目標客群
- b. 後端功能發想
 - i. 各管道資訊擷取
 1. [臉書](#)
 - ii. 辨識中立程度或是立場(語意分析?)
 - iii. 聲量數據分析
- c. 前端功能發想
 - i. 資訊瀏覽
- d. 第一版:資訊頁面呈現, 資訊分組, 聲量
- e. 第二版:立場互動, 增加資訊來源
- f. 第三版:資訊立場自動分析

2. 夥伴自介表：

夥伴暱稱	可能參與方向	本坑的負責方向	對本專案的想像
eli		ui/ux、行銷推廣	1.快速瀏覽同議題不同立場的講法 2.找出偏頗資訊的散佈路徑
邦宇	資料工程師	爬蟲、後端	
kiko555	後端(ELK、爬蟲)、前端(Ionic)	爬蟲、後端(資料Metadata)	
ABow	前端(HTML / CSS / RWD / Bootstrap / JavaScript新手)、平面設計(也許可以支援 UI)	前端優先, 設計方面可嘗試支援	
Stimim	演算法、後端 > 前端		

Wei	前端 > 後端	以前端為主, 後端嘗試支援	
Afra	...不會什麼技術, 但可以參加討論發想, 貢獻一點點個人意見		
ellie		協助測試、推廣	
Jason	使用Python做機器學習影像辨識	圖片/影片 實例分割, OCR	
Simon	立場 thoughts.blueplanet.com.tw	從講清楚立場的 spec 開始	
YenChen		API	
Steven	C++/Python相關	爬蟲	在日常常用的媒體上, 能夠看看不同立場的報導及聲音
PeterWolf	NLP 技術議題	若可以脫離 CKIP 的話, 也許可以提供額外的 NLU 技術與應用方向	
Xwei	後端(爬蟲, Data pipeline)、立場		
夥伴暱稱	可能參與方向	本坑的負責方向	對本專案的想像
1			
2			
3			
4			
5			
6			
7			
8			
9			

10			
11			
12			

二、系統開發細節與進度

1. 資訊擷取來源與負責夥伴：

類別	名稱	方法	負責夥伴
聊天	Line wechat messenger		Z.J
社群	臉書粉絲頁	已完成 相關臉書爬蟲(未維護) https://github.com/padraig7a/fbcrawl 『聲量看政治』x『CIH看粉絲名單』 https://www.facebook.com/cih.tw/posts/4180422561983110	kiko
社群	臉書社團	初步完成, 調整中 相關臉書社團爬蟲 https://freelancerlife.info/zh/blog/python%E7%B6%B2%E8%B7%AF%E7%88%AC%E8%9F%B2%E6%87%89%E7%94%A8-facebook%E7%A4%BE%E5%9C%98%E6%88%90%E5%93%A1%E5%8F%83%E8%88%87%E5%BA%A6%E5%88%86%E6%9E%90/ 臉書社團, 2個小時更新一次, 因為我們不是要即時, 而是要提供資訊。	shuo-ren

		看可以留言可以抓到什麼程度 可以抓貼文的圖片，貼文的情緒，留言的id和文字(留言的圖片目前不抓)。 有用ocr	
論壇	PTT	已完成	邦宇
論壇	Dcard	已完成	邦宇
部落格	Pixnet	API Doc 讀完	邦宇
部落格	blogger		
社群	Instagram	能抓取公開帳號的所有資料 已將抓取的資料存成 json 檔 已完成	Wei
影音	youtube	聲音轉逐字稿	
api	0archive	了解 0archive 裡有什麼可以用	stimim
新聞	自由 政治新聞 東森 政治新聞 中時 政治新聞 臉書社團-1.zip	已完成 以news api 來做為新聞來源 先進行 中文斷詞	kiko
搜尋	google		

2. 類似資訊判別：

- a. 文字語意分析：判斷文章的正面反面。情緒種類。親中親民等特定標籤。找出畫分族群的模式。
- b. 圖形辨識：(可先不要開發)
 - i. tag圖片中的特徵：他是屬於哪一類的，那就有可能做歸類，認為是類似的資訊在流
 - ii. (json)分割圖形：ps加圖刪圖組合過，把這個repository 的圖形分割 run起來 <https://github.com/dbolya/yolact>
 - iii. 圖形相似度辨識：如果僅是簡單修改
 - iv. 分辨圖中的中文文字內容

類別	方法	進度	負責人
----	----	----	-----

圖形辨識	yolact實例分割	修改了一下code resize 圖片防止OOM、判斷圖片檔案正確格式與處理gif(目前對gif處理是將第一幀轉存為jpg做辨識), 人物、動物、交通工具等常見物件較佳, 需進一步處理卡通圖等虛擬圖片。bug待修	Jason
相似度辨識	cnn特徵抽取比較、hash、sift分析	幾種hash感覺不太可行;cnn 特徵抽取嘗試了vgg16、resnet50, 效果還不理想。預計將判斷不佳的部份用sift 分析再做一次比較結果。	Jason
文字辨識	tesseract	待提升精確度。	Jason

3. 資料儲存共通Metadata定義:

文章:

- a. 文章 ID
- b. 文章標題
- c. 作者 ID
- d. 板名
- e. 文章內容
- f. 發文時間
- g. 發文 IP
- h. 情緒
 - i. 推(臉書是否要有差別)
 - ii. 噓

推文:

- i. 推文內容
- j. 推文時間
- k. 推文者 ID
- l. 推文IP
- m. 情緒:
 - i. 推(臉書是否要有差別)
 - ii. 噓

4. 資訊立場分析

- A.預設立場:對發佈管道, 發布者所發佈的歷史文章, 先做人工立場判斷
- B.立場互動:使用者更改該資訊的立場
- C.自動分析:分析標題與文本前50字

三、發想

第一版連結：

<http://neutral-info.herokuapp.com/>

專案簡介：

「一個能簡便篩選資訊立場屬性的平台」

- 1.快速瀏覽同議題不同立場的資訊
- 2.找出偏頗資訊的散佈路徑
- 3.借平台的資訊，供作異溫層溝通的工具之一

專案架構：

第一層：資訊擷取。利用爬蟲程式，抓取不同論壇粉絲頁社團line等管道的資訊

第二層：賦予立場屬性：另用立場互動等方式，給予資訊諸如民進黨、資方、勞方等立場屬性

第三層：平台運用。藉由以上兩層處理後的資料，不同使用者可有各種使用方式，像是與「真的假的」假資訊line機器人合作，對其擁有的資訊給予「立場屬性」；給媒體網站上的文章，類似facebook like的運作機制般，給予文章「立場屬性」的互動功能外掛。

任何想法：

1. 所有文章或影像等等的內容，都在標題上附註各種標籤
 - a. 來源：資訊來源，如自由時報、中國時報、社群文章
 - b. 立場傾向：提醒讀者這個內容具有什麼樣的傾向，
 - i. 政治傾向：如偏紅、偏綠、偏藍
 - ii. 意識形態傾向：如偏左、偏右；自由、專制
 - iii.
 - c. 發文者身份別差異：
 - i. 年齡

- ii. 國家:如IP來源地
 - iii. 產業:如新聞業、自媒體、個人
 - iv. 學經歷:如學歷、總經理
 - v. 宗教
- d. 網路聲量統計:標明議題的熱度(是否需要與時間做綁定)
- e. 其它輔助標籤:供使用者篩選
- ~~2. 踏出同溫層:同議題的資訊,如何依年齡、國家、產業、職業地位、政治傾向、宗教等角度去劃分。國家、政治傾向比較好做~~
- 3. 現在的社會很容易會出現資訊爆炸的狀況,這個平台如何解決資訊爆炸的問題?
- 4. 真的有「中立」的資訊嗎?
 - a. 資訊應該被選擇揭露到怎樣的地步,才算是「中立」呢?
 - b. 不同的「觀點」或「視角」又該怎麼選擇呢?
- 5. 務實上,踏出同溫層有個難點,就算我們認為我們的資訊相對正確,像「美玉姨」這樣的事實查核機器人被踢出長輩群組也是時有所聞,所以如何透過「良好」的溝通讓異溫層的民眾或是長輩可以了解,是我認為目前遭遇到的最大難處。
 - a. 「美玉姨」的表達的確較強烈,有可能造成情緒的問題讓群內不同立場的人群反彈。
 - b. 在這樣的概念上,提供立場較往中間靠攏的資訊源,例:某新聞靠右,就提供靠中的給他參考

專案目標與使用情境:

1. 提供好用的工具辨別「立場」與「中立」資訊
 2. 提供好用的工具讓使用者能透過「良好」的溝通,讓異溫層的民眾或是長輩可了解不同立場的資訊:自動把文章修改成該同溫層的用語;給予立場光譜類似的資訊,而不是給對立的;透過line bot 幫使用者轉傳相關關鍵字的三篇文章給目標人
 3. 幫使用者想到好用的功能和完整的使用情境
 4. 分析異常散佈資訊:ex社群過熱度、
 5. 真的假的或各媒體平台,資訊的「立場」欄位
 6. 專訪或研究所使用的數據
 7. 直接iframe呈現在某些平台。例如「萊豬」相似資訊的列表,可觀察各篇立場或聲量等數據
-
1. 提供資訊列表,有各種filter(聲量 爆發力 管道 發布者 立場...)
 2. 提供類似資訊列表,運用關鍵字或來源網址等方式把相似資訊分群
 3. 分析各管道與發布者的立場,還有類似立場資訊的散佈途徑
-
1. 花力氣搜尋非同溫層的文章不太符合人的惰性,希望可以讓使用者以最日常或最小改變,就能看到推薦的非同溫層文章。例如moptt能夠提供一個非同溫層文章的選項, moptt的使用者應該會比較願意瀏覽

目標客群：

1. 想改變同溫層的
2. 想看到非同溫層資訊的

UI：

1. prototype：
<https://www.figma.com/proto/ZxHOSfVJY4bkqUDzwG7IBU/neutral-info?node-id=151%3A4&scaling=scale-down>
2. Ui 展開連結：
<https://www.figma.com/file/ZxHOSfVJY4bkqUDzwG7IBU/neutral-info?node-id=0%3A1>
3. 草稿：
<https://www.figma.com/file/ZxHOSfVJY4bkqUDzwG7IBU/neutral-info?node-id=0%3A1>
4. 示意圖

◀ 口罩地圖懶人包 ▶

口罩地圖 🔍 📏 🔥 ⚙️

標籤	立場	聲量	事
口罩地圖 台中 <> 🔄 民進黨:70% 國民黨:65% 時力:45% 無立場:44% 武漢肺炎》即時口罩地圖台中市有升級版- 「台中即時口罩地圖」優化30秒更新一次- 台中即時口罩地圖再升級可連結導航察詢表單-自由時報 台中市研考會主委吳皇昇表示，口罩實名制實施後，市府已於2月7日介接健保署「健保特約機構口罩剩餘數量明細清單」資料，建立即時口罩地圖... ▼ 展開更多相關資訊	民進 🇦🇵 國民 🇦🇵 時力 🇦🇵	130 102 90 72	📱 📱 📱

更多相關...

各國數據 中國 防疫措施 WHO

5.

聲量運算邏輯：

- 就目前已擷取的數據，設定權重
- 擷取到的數據 *n *t
- 一天以前擷取的 讚10 $10 \times 10 \times 0.5$
- 一天內擷取的 讚20 ->10各 $10 \times 10 \times 1$

n	閱覽	讚 推 愛心	中立	留言	分享	回文	收藏	
臉書		10		50	100			分享目前未抓
ptt		200	50			200		回文目前無法分辨
新聞網	1	10		50	100			
dcard		20		100		200	50	
Pixnet	1			50				
IG	1	10		50	100		70	

t權重	-24hr	1-7d	7-40d		
	1	0.5	0.1		

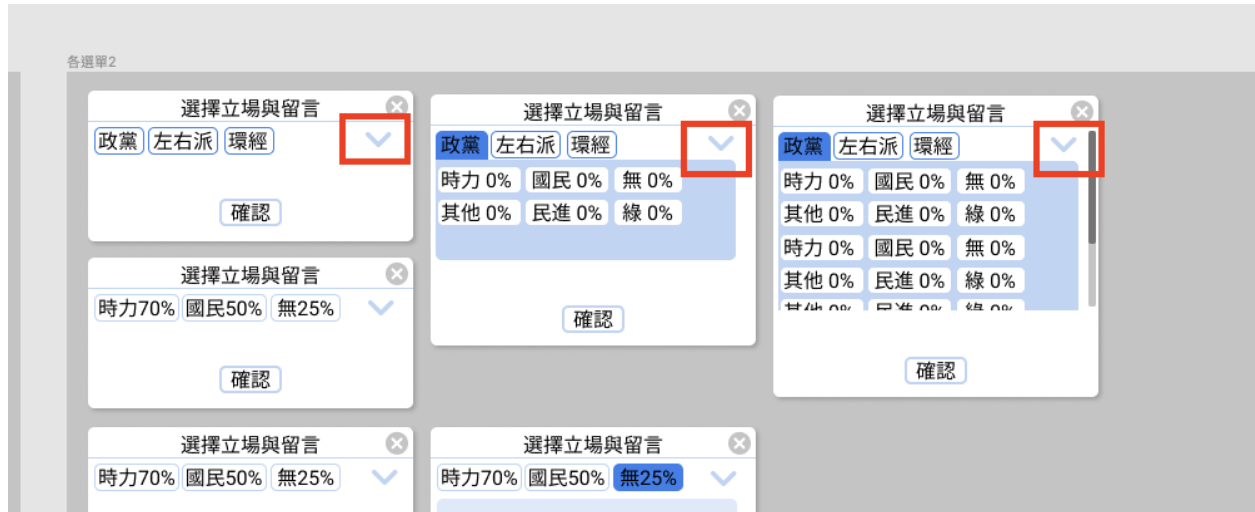
爆發力運算邏輯：

1. 聲量對時間的微分

4/14

<https://www.figma.com/file/ZxHOSfVJY4bkqUDzwG7IBU/neutral-info?node-id=0%3A13/31>

只有立場類別的情境下，不需要有開合的按鈕



<https://www.figma.com/file/ZxHOSfVJY4bkqUDzwG7IBU/neutral-info?node-id=0%3A13/31>

選擇立場與留言

政黨 左右派 環經

確認

選擇立場與留言

時力70% 國民50% 無25%

確認

選擇立場與留言

時力70% 國民50% 無25%

確認

userid
我覺得比較偏向國民黨

userid
文章有提到這三個觀點，所以應該比較偏向民進黨

userid
國民黨

userid
偏時代力量

userid
都不像

選擇立場與留言

政黨 左右派 環經

時力0% 國民0% 無0%
其他0% 民進0% 綠0%

確認

選擇立場與留言

時力70% 國民50% 無25%

確認

userid
我覺得比較偏向國民黨

userid
文章有提到這三個觀點，所以應該比較偏向民進黨

userid
國民黨

userid
偏時代力量

userid
都不像

選擇立場與留言

政黨 左右派 環經

時力0% 國民0% 無0%
其他0% 民進0% 綠0%
時力0% 國民0% 無0%
其他0% 民進0% 綠0%

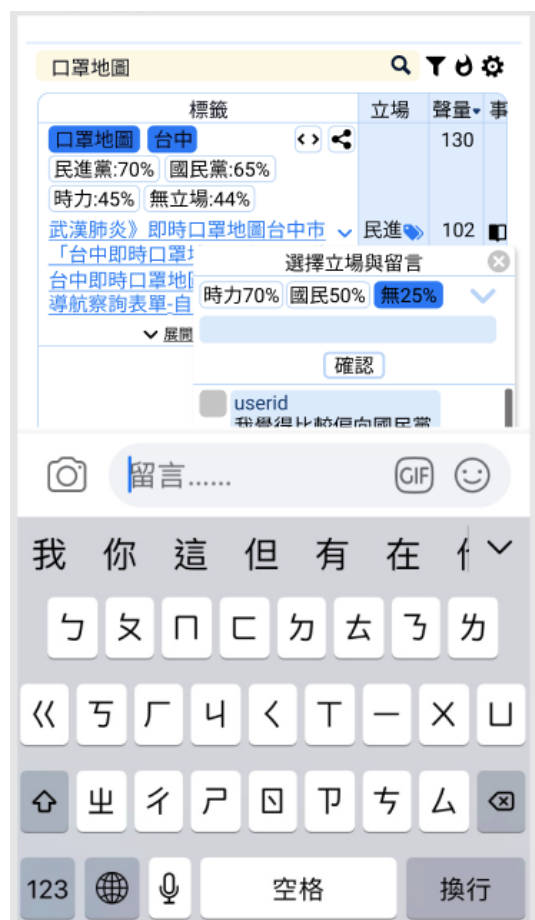
確認

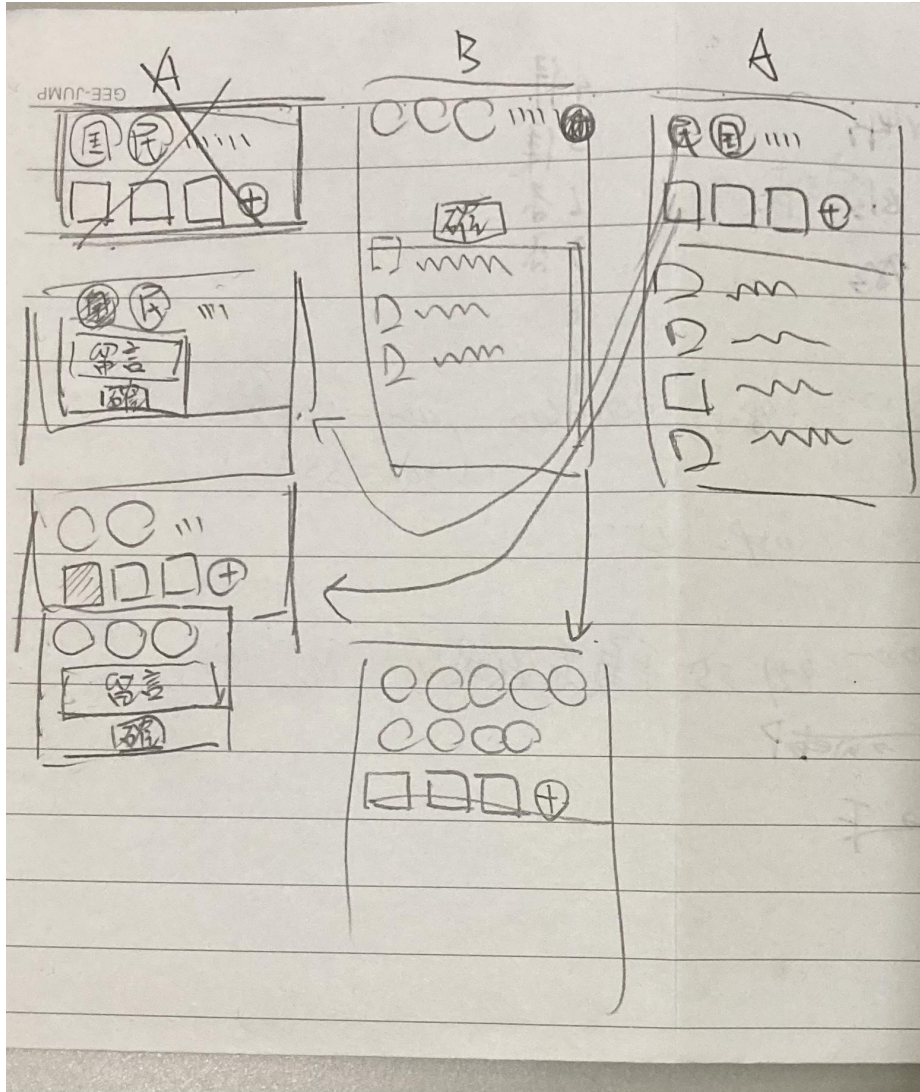
選擇立場與留言

立場
時力70% 國民50% 無25%
其他23% 民進15% 綠13%
...

立場類別
政黨 左右派 環經 ...

確認





第二層改成展開

3/10會議記錄

搜尋萊豬, 100篇

顯示5篇(熱門度排序, 且彼此最高立場不同), 其他要展開

對某一標籤可以贊成或反對 +1 -1

-100就會隱藏, 用民眾刪去立場標籤

立場規則

1. 先選立場類別(你覺得還有什麼立場類別, 像是許願池的概念, 收集到一定數量, 就開放)

,

- 2.再選擇這個類別的立場標籤(是否開放添加自訂標籤;留言;「其他」的標籤, 可以同時留言)
- 3.是否留言

立場類別:(用一個主題議題去區分)

1. **政黨**:藍綠 / 民主黨 共和黨
2. 左派右派:
3. 勞資:勞權 / 資方-慣老闆
4. 環境經濟:環保
5. 開發食安:
6. 統獨:
7. 教育:
8. 能源環境:
9. 選舉:在野 / 執政
10. 藻礁:自然 / 經濟 / 能源
11. 核能:擁核 / 反核
12. 萊豬:食安
13. 路權:四輪 / 二輪 / 行人 / 人本
14. 監控:反監控 / 不在乎
15. 交通:公路派 / 鐵路派
16. 性平:父權 / 女權
17. 同婚:挺同 / 反同
18. 對政府:保皇派-保守派-建制派 / 改革派-民主派
19. 貿易:全球化 / 在地化
20. 語言:便於溝通 / 多樣化

3/3會議記錄

第二版:

1. 熱門頁面的api
2. 立場互動功能(短時間不能操作:1.沒關掉頁面就不能修改)
立場標籤:第二版先做政黨
立場類別:政黨、意識形態、尋找立場有哪些類別。(反對核電 反對核四 反對某種核電廠)
是兩層還是一層, 寫死的還是可以擴充, 還是給活躍使用者選擇然後寫死
3. 會員系統(針對互動。未來個人設定。個人lv增加散佈力)

<http://neutral-info.herokuapp.com/>

2/17會議記錄

1. 是否用google自定義搜尋

<https://developers.google.com/custom-search/docs/element>

<https://search.gamer.com.tw/?q=%E7%A5%9E%E9%AD%94%E4%B9%8B%E5%A1%94#gsc.tab=0&gsc.q=%E7%A5%9E%E9%AD%94%E4%B9%8B%E5%A1%94&gsc.page=1>

2. 做原文連結

版本方向

1. 各資訊來源的聲量
2. 各資訊來源的範圍與抓取頻率
3. 增加資訊來源爬蟲(臉書社團)
4. 立場互動功能
5. 資訊分組

等到立場互動功能完成, 對外找小眾開始測試使用情境

7/29會議記錄

確認第一版api格式

7/8會議記錄

[第40次黑客松](#) 7/25台北

[雙年會](#) 7/15報名截止 12/3-6

6/24會議記錄

國外新聞立場分析的資源 <https://mediabiasfactcheck.com/>
怎麼做的?

6/17會議記錄

1. 臨時檔案局資料研究:
 - a. 狀況
 - i. 資料集並沒有包含留言及情緒資料
 - ii. 目前公開資料集不含臉書
 - iii. 資料集更新不夠即時, 這樣無法達到分析熱門新聞的功能
 - b. 可能的應用
 - i. 資料集可供立場訓練
 - ii. 提供我們回溯補充新聞等資料, 降低重新抓取成本

- c. 新聞補充方法
 - i. 能拿到爬蟲程式原始碼, 所以應可自行抓取, 可以控制資料更新的頻率
 - ii. 或是可以取得檔案局還未匯整成單月前的資料庫權限? (檔案局的設計是先有爬蟲抓取進資料庫, 後再另有一程式產生匯整後的統一格式)
- 2.

6/10會議記錄

6/3會議記錄

各程式邏輯

1.資料對接

各來源資訊擷取

聲量等欄位運算

資訊相似度

視覺ui

前端

資料庫

資訊分組-圖形QAA

資料分組-文字

資訊立場語意分析

2.資訊分組-圖形

a.把圖片中的文字擷取出來。目的是產生文字, 交給「資訊分組-文字」語意分析去確認關鍵字與權重、「資訊立場分析」立場分析等

b.把圖片分割或是擷取特徵。目的是為了比較不同圖片彼此的相似度, 輸出給「資料庫」建立各圖形相似度?

3.資訊分組-文字

A.拆分標題與內文的關鍵字, 依照出現的次數, 直接設權重

b.拆分標題與內文的關鍵字, 語意分析重要的關鍵字後, 再結合是標題、出現頻率較多、較前面的、「」中的、不同語言等判斷, 設權重

1. Input: 標題 & 內文

2. 中文斷詞、先按詞頻排序
3. Output: 前三筆關鍵詞 & 詞頻

4. 資訊立場分析

- A. 從資訊的發布管道預設立場權重
- B. 網友編輯權重
- c. 從標題與內文語意分析

資料分析邏輯：

一、輸入

爬蟲擷取的資訊

- a. 從標題與內文，拆解關鍵字(可以從標題和內文前30個字就好。但大家可以再討論)
- B. 分析出比較重要的關鍵字，並給予該關鍵字在該資料的重要性權重

二、輸出

- a. 網站前端傳來使用者打的關鍵字
- b. 從資料庫比對關鍵字，挑出有關關鍵字且重要性權重比較高，時間彼此比較近的資料
- c. 輸出給前端，這幾個資訊

舉例：以下文件有兩個資訊

<https://docs.google.com/document/d/11GjKf3tyszeGNPm3khWxNrFkIWgDWGlSXjwSoJGB4IQ/edit?usp=sharing>

⇒ 從標題與內文拆解關鍵字後，推出比較重要的
(可能是標題、出現頻率較多、較前面的、「」中的、人名，不同語言的權重比較重)

⇒ 得到各資料重要的關鍵字組

資料1

美國 暴動 庫克 備忘錄 蘋果 黑人 佛洛伊德 George Floyd Tim Cook

資料2

美國 川普 暴動 佛洛伊德 George Floyd 商業 領袖

⇒ 並設定重要性權重，相關性

資料1

美國100 暴動50 庫克50 備忘錄20 蘋果30 黑人30 佛洛伊德10 George10 Floyd10 Tim20
Cook20

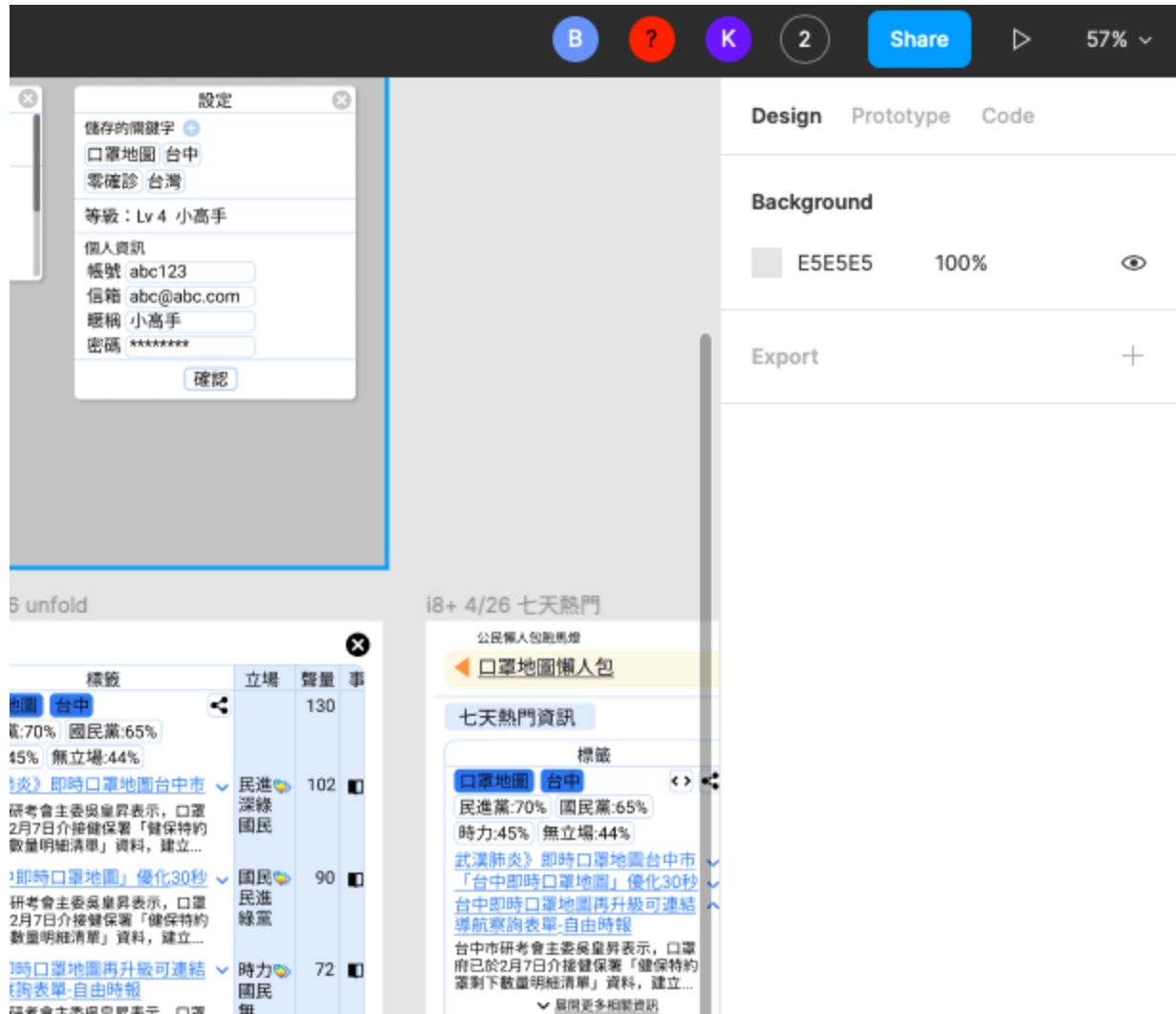
資料2

美國100 川普80 暴動30 佛洛伊德10 George10 Floyd10 商業20 領袖20

⇒ 使用者在網站搜尋「美國 暴動」，從資料庫找有這些關鍵字的資料
假設有很多資料都有「美國」「暴動」。但所有資料中這兩個關鍵字重要性權重最高，且有很多其他相同的關鍵字的資料是「資料1」與「資料2」

⇒給前端「資料1」與「資料2」的資料庫編號。前端顯示這兩個資料要列在前端的內容

5/27會議記錄



篩選

聲量 -
爆發力 -
立場
時力力量 國民黨 無立場
管道
新聞 社群論壇 聊天 部落格
內容農場 影音

確認

熱門

7天熱門資訊
24hr熱門資訊
立場
民進黨
國民黨
時代力量
綠黨

設定

儲存的關鍵字
口罩地圖 台中
零確診 台灣
等級：Lv 4 小高手
個人資訊
帳號 abc123
信箱 abc@abc.com
暱稱 小高手
密碼 *****

確認

i8+ 4/26 七天熱門

公民懶人包跑馬燈

◀ 口罩地圖懶人包 ▶

七天熱門資訊

標籤

口罩地圖

台中

民進黨:70% 國民黨:65%
時力:45% 無立場:44%

<>

立場	聲量	事
	130	

5/20會議記錄

@ 各資訊管道 [立場分析試算表](#)

1.進度分享

2.5/23六 線上松 09:00-12:30。我們三分鐘短講的時間 **10:08**

報名連結

<https://g0v.hackmd.io/@iothon/g0v-hackath39n/https%3A%2F%2Fg0v.hackmd.io%2F%40iothon%2Fg0v-hackath39n?fbclid=IwAR0M4vKPOrPPqgzz4isX-7uaQK0PbaH3O2MgqQ8ewmBLgr4hUX6D7MQQPKo>

線上松簡報連結

https://docs.google.com/presentation/d/1q_n59mBl0Mp-MWBe4SErQjuqCxGCOiBSBUm6AOXKBRw/edit?usp=sharing

資訊列表欄位：

第一版：

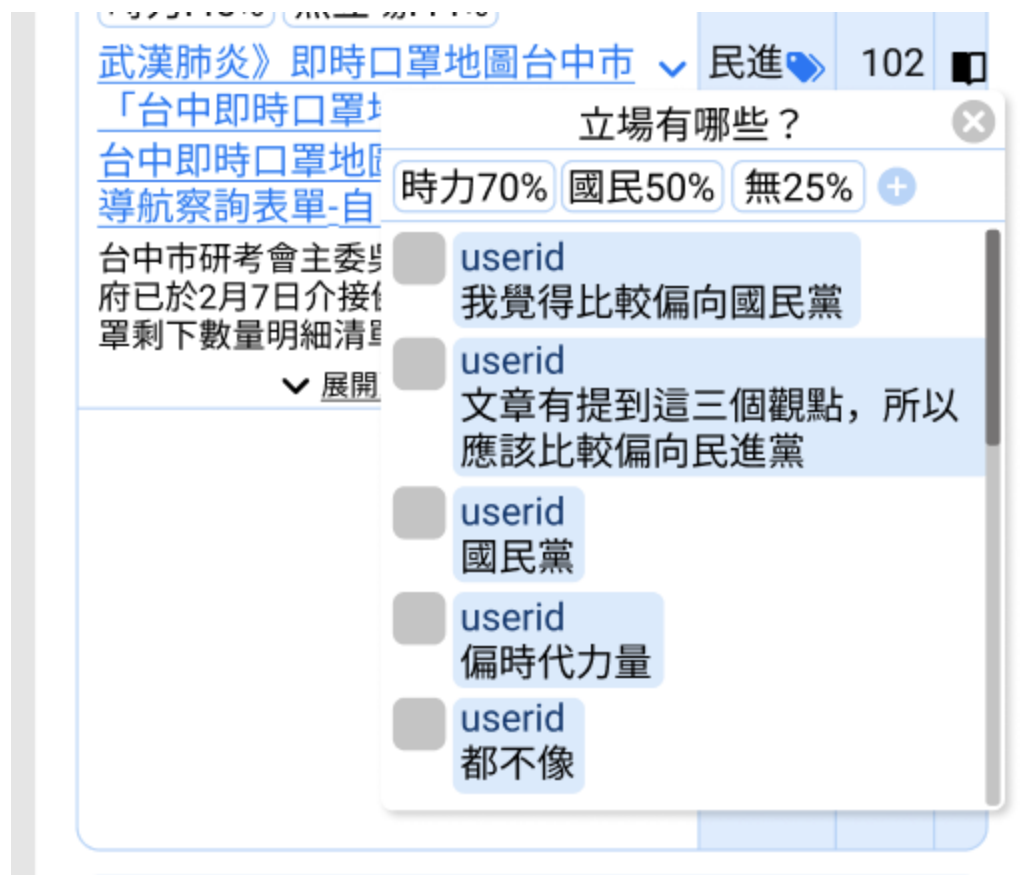
關鍵字標籤。文章標題。發文時間。**發布管道**。發布者。聲量。爆發力。立場標籤(親民進黨。親國民黨...等)。30字文章大綱。

第二版：

真實性(真的假的)。管道類別:新聞、社群論壇 | 、聊天、部落格、內容農場、影音。

第三版：

標題修改記錄。**資訊列表標籤**(社群管道、高機率會熱門、常發文發布者...等)。**過熱度**(不太合理的熱度或聲量爬升)。**特色標籤**(旅遊、政治...等)



臉書社團-1.zip: 公民割草行動, TAG:韓黑; 韓國瑜鋼鐵韓家軍, TAG:國民黨、韓粉、英黑;
韓家軍-永遠忠誠韓國瑜, TAG:國民黨、韓粉、英黑; ●【爆料公社】●, TAG:中國。

5/6會議記錄

下次大松 5/23

| 注意:這裡有公開紀錄, 請留意個資安全 | 自己邀請自己 <https://join.g0v.tw/> | 新加入看教學 <https://g0v.hackmd.io/@jothon/g0v-cowork-guideline/> | 與 [freenode#g0v.tw](#) 同步中

一、議程:

- 各欄位運算邏輯:
 - 聲量:臉書和 ptt 計算聲量 excel 表
 - 爆發:
 - 立場:選擇幾個粉絲頁
 - 可信度改成事實查核數據數 ([台灣事實查核中心](#)、真的假的)
- 給其他平台使用的api

18+ 4/26 index

公民懶人包跑馬燈

◀ 口罩地圖懶人包 ▶

口罩地圖 你覺得本文章的立場有哪些?

時力70% 國民50% 無25% +

口罩地圖 台中

我覺得比較偏向國民黨

文章有提到這三個觀點, 所以應該比較偏向民進黨

民進黨:70% 國

時力:45% 無立

武漢肺炎》即時口罩地圖台中市 ✓ 民進 102

「台中即時口罩地圖」優化30秒 ✓ 國民 90

台中即時口罩地圖再升級可連結 ^ 時力 72

公民懶人包跑馬燈

◀ 口罩地圖懶人包 ▶

口罩地圖 你覺得本文章的立場有哪些？

時力70% 國民50% 無25% +

留言簡述理由，或直接按確認

確認

我覺得比較偏向國民黨

文章有提到這三個觀點，所以應該比較偏向民進黨

▼

國民 無

台中市研考會主委吳皇昇表示，口罩府已於2月7日介接健保署「健保特約

導航察詢表單-自由時報

台中市研考會主委吳皇昇表示，口罩府已於2月7日介接健保署「健保特約

📷 留言..... GIF 😊

我 你 這 但 有 在 ↑ ↓

ㄅ ㄆ ㄇ ㄋ ㄌ ㄍ ㄆ ㄇ

4/29會議記錄

一、議程：

- 各欄位運算邏輯：聲量。爆發。立場，可信度改成事實查核數據數([台灣事實查核中心](#)、真的假的)
- 資訊的立場前處理：
- 立場互動程式邏輯
- 給其他平台使用的api？

立場與權重是由發佈管道立場分析，
與網友增減編輯而成。

你贊同本文章的立場分析嗎？



時力70%

國民50%

無25%



我覺得比較偏向國民黨



文章有提到這三個觀點，所以應該
比較偏向民進黨



無25%

無25%↑

無28%↓

無23%⊗

無25% ↑ 👍 ↓

無25% ↑ 👍 ↓

◀ 口罩地圖懶人包 ▶

口罩地圖



標籤	立場	聲量	事
口罩地圖 台中 <> 分享 民進黨:70% 國民黨:65% 時力:45% 無立場:44% 武漢肺炎》即時口罩地圖台中市有升級版- 「台中即時口罩地圖」優化30秒更新一次- 台中即時口罩地圖再升級可連結導航察詢表單-自由時報 台中市研考會主委吳皇昇表示，口罩實名制實施後，市府已於2月7日介接健保署「健保特約機構口罩剩下數量明細清單」資料，建立即時口罩地圖... 展開更多相關資訊	民進 國民 時力 無 102 90 72 70% 50% 25%	130 102 90 72	事 事 事 事

更多相關...

各國數據
中國 防疫措施
WHO

4/22會議記錄

一、議程：

7. 各欄位運算邏輯：聲量。爆發。立場，可信度改成事實查核數據數([台灣事實查核中心](#)、真的假的)
8. 資訊的立場前處理：
9. 立場互動程式邏輯
10. 給其他平台使用的api？

立場邏輯：

1. 初始立場數據：發佈粉絲頁與資訊來源，各來源設為100%，依據散佈的管道計算立場數據。
2. 變動立場數據：網友共同編輯，對每個標籤可 +5% (如果覺得該數據是正確的，是否有同意的選項？)。
3. 網友共同編輯：有作加減的，可選擇留言。有留言的操作改為 +8%

立場粉絲頁與資訊：

1. 確定來源與其立場，設為100%
2. 發布該資訊，時間最前面的*100%，其他*90%，最後除以來源個數

欄位運算邏輯：

聲量運算邏輯：

就目前可擷取的數據，設定權重

擷取到的數據 *n *t

一天以前擷取的 讚10 $10 \times 10 \times 0.5$

一天內擷取的 讚20 ->10各 $10 \times 10 \times 1$

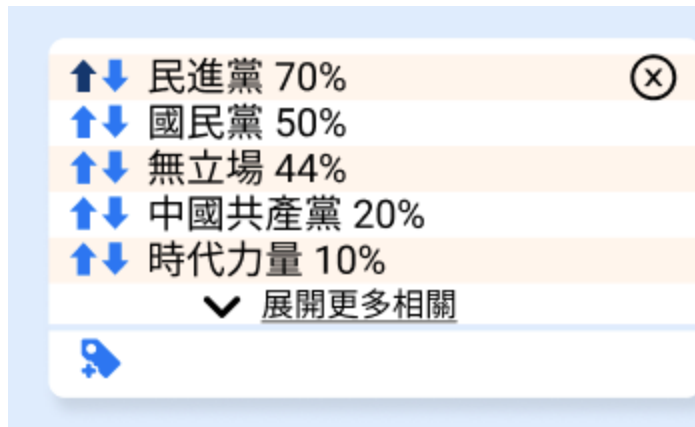
n	閱覽	讚 推 愛心	留言	分享	回文	收藏	
臉書		10	50	100			

ptt		20			200		
新聞網	1	10	50	100			
dcard		20	100		200	50	
Pixnet	1		50				
IG	1	10	50	100		70	

t權重	-24hr	1-7d	7-40d		
	1	0.5	0.1		

爆發力運算邏輯：

聲量對時間的微分



標籤	立場	聲量	爆發	事實	
口罩地圖	台中	130	95		
「台中即時口罩地圖」優化	▼ 民進	102	60		🗨️
即時口罩地圖，台中很多	▼ 國民	90	55		🗨️
即時口罩地圖，台中很多	▼ 時力	72	41		🗨️
▼ 展開更多相關					

4/15會議記錄

一、議程：

1. 對立場的判斷，

- A.對發佈粉絲頁與資訊來源來作標籤:設100%
- B.網友共同編輯 對每個標籤+/-5%
- C.這個平台的互動就是立場選擇或覺得無立場, 反串, 選擇理由

ABow:

- A.5個顏色(藍、綠、橘、紅、白)
- B.認為有特定立場、無特定立場
 - B-1:有特定立場>偏藍、綠、橘、紅、白>原因...
 - B-2:無特定立場>>原因...

C.需要揪出反串嗎？

C-1:設計問題揪反串, 例蘋果日報的老實樹揪出反串韓粉

2. 老實樹揪假

■支持蔡又兼具2條件 韓粉反串疑佔3百分點

由於韓國瑜陣營以所謂的「空城計」，呼籲支持選民接到民調電話時，一律回答「民調唯一支持蔡英文」，為找到有多少韓粉在民調中反串為蔡英文陣營支持者，《蘋果》在民調特別安排2棵「老實樹」問題，第一棵「老實樹」問題是詢問民眾對「韓國瑜要支持者接到民調電話時一律回答『唯一支持蔡英文』」的看法，在提示選項情況下，有42.4%回答是「愚蠢的政治騙術」，20.3%回答是「高明的選舉策略」，但也有37.3%未明確回答。

第二棵「老實樹」問題是詢問民眾「如果沒有發生影響選情的重大事件，情勢也保持穩定沒太大變化，您認為目前媒體公布的民調結果，和實際投票結果會不會有很大差距？」，結果有58.1%認為「民調和選舉結果差距不大／很接近」，有24.8%則認為「民調和選舉結果差距很大」，17.1%未明確回答。

進一步交叉分析，在三組候選人中，「英德配」支持者有57.2%認為韓的反串策略是「愚蠢的政治騙術」，16.5%回答是「高明的選舉策略」。「國政配」支持者中有41.8%認為是「高明的選舉策略」，18.3%認為是「愚蠢的政治騙術」。親民黨「宋余配」支持者中有62%說是「愚蠢的政治騙術」、27.8%認為是「高明的選舉策略」。

因此若將「英德配」支持者中，認為反串是「高明的選舉策略」的16.5%，再加上同時回答認為「民調和選舉結果差距很大」後，即可推論可能是韓粉反串英粉，經電腦程式估算後約佔整體比率的3%，亦即在支持「英德配」的50.8%支持者中，可能有3%是韓粉反串，減掉反串的3個百分點後，「英德配」的支持度變成47.8%（50.8%-3百分點），「國政配」則成為18.2%（15.2%+3百分點），兩組差距從35.6個百分點，縮小為29.6個百分點。而根據中選會資料，去年九合一選舉，有投票權人數約1920萬人，3%約為58萬人，也就有約有58萬的韓國瑜支持度聽從指示，在接到民調電話時回答「唯一支持蔡英文」。

C-2:如需要, 推算出的反串的比例該如何計算？(比如綠反串藍, 所以歸類為立場綠, 或是不計算)

2. 共同編輯規則,

- a.可非登入互動。會員互動。會員等級。可信任會員。

ABow:

前題是釐清可以編輯的資料為何？

A.評價審核:

A-1:可以使用者互評, 如果被許多使用者評為黑名單(如:貼文帶有極為強烈的偏頗立場、轉貼十分明顯的假訊息圖文、或惡意、人身攻擊別人), 可能會被封鎖權限。

A-2:如果資料被多次且惡意修改, 會封鎖資料的編輯功能。

3. 與操作細節

eli:ux操作可以直接選擇「國民黨」直接選擇這個關鍵字組合的國民黨文章

4/8會議記錄

一、議程：

1. 上週任務報告(每人3-5min)
任務進度可以放在「系統開發細節與進度」的各自表格中
2. 使用者情境排序與新增(討論完後, 建立投票) 1 2 3 8
3. 依重要性排序欄位(討論完後, 建立投票)立場 可信度 / 聲量 爆發力 / 瀏覽數 互動
關鍵字文章數聲量
4. 各欄位運算邏輯:聲量。立場, 可信度改成事實查核數據數([台灣事實查核中心](#)、真的假的)
5. 網友共同編輯:評價評論。闢謠(導去真的假的)。錯誤回報。編輯立場與特色標籤(除了選擇與新建標籤外, 寫下理由與參考連結會增加其建議的權重)。未來事件交易所的期貨概念。共同編輯公民懶人包。帳號登入投票
6. 定義何種立場:「吃益生菌可以提升免疫力嗎。元老級抗衰老診所醫師柯威旭指出...」
→ [負面](#) (42%);可從一些藍綠媒體的非中立標籤新聞開始建立。關鍵字擷取:
[TF-IDF](#).

下週：

預設列表：

1. 公民資訊懶人包
2. 各立場熱門文章(點選火焰icon)
3. 7天熱門關鍵字(點選火焰icon)
4. 24小時熱門關鍵字(點選火焰icon)

漏斗icon:對該關鍵字, 做立場, 管道, 聲量等篩選

齒輪icon: 自己的關鍵字等設定

◀ 口罩地圖懶人包 ▶

口罩地圖



標籤

立場

聲量

爆發

事實



口罩地圖 台中

130 95

[「台中即時口罩地圖」優化](#)



民進

102 60



[即時口罩地圖，台中很多](#)



國民

90 55



[即時口罩地圖，台中很多](#)



時力

72 41



[展開更多相關](#)



更多相關...

各國數據
中國 防疫措施
WHO

4/1會議記錄

一、議程：

1.上週任務報告(每人3-5min)

臉書第一頁情緒可以知道總數 但不能知道留的人的id，留言內容

使用者情境(粗體為比較重要的情境)

1. A看到line群組有偏頗立場文章->轉貼該連結給賴機器人「多立」->「多立」提供該組文章相關的不同立場三篇資訊列表->轉貼到群組

2. A在社群中看到感興趣的資訊->到多立網頁->搜尋相關關鍵字->了解相關資訊->分享到臉書

3. A像讀報一樣, 早上七點先到多立網頁, 閱讀公民懶人包資訊、這七天熱門關鍵字的資訊列表->有時也會閱讀他設定好的關鍵字資訊列表->可以用爆發力做排序

4. A有感於網路資訊太多謾罵->設定資訊情緒為正面作為篩選條件->散佈相關資訊到社群

5. A看到賴群有親藍的資訊->到多立網頁->搜尋相關關鍵字->了解相關資訊->分享立場介於藍綠中間的資訊到群組(也可設定別種立場分類)。試圖以沒那麼對立的資訊, 讓對方願意閱讀

6. A看到line群組疑似假資訊->轉貼該連結給賴機器人「多立」->「多立」提供該組文章相關的三篇資訊, 可供比對判斷, 並有「真的假的」的真實性判斷->A轉貼上述資訊到群組

7. A設定最常逛的ptt dcard 東森新聞為管道來源->每日閱讀該來源資訊列表

8. 看到文章, 來到多力平台, 了解發布者的立場的偏向

7. 使用者情境排序與新增(討論完後, 建立投票)

8. 依重要性排序欄位(討論完後, 建立投票)

文章部分:

a. 需計算的數據: 聲量。爆發力。立場。互動數。主文情緒。主文正反面。留言情緒。留言正反面。關鍵字。關鍵字的文章數與瀏覽數。可信度?。

b. 直接紀錄的數據(時常變化): 瀏覽數。留言數。按讚愛心等互動數。回文數。分享數。紀錄時間。互動者id、個人資訊與內容。

c. 直接紀錄的數據(不變化): 文章標題。文章內容。管道名。版名。發布者。

9. [新文易數](#) 如何做到文章自動標籤關鍵字

10. 聲量運算邏輯

11. 什麼叫做中立或是多立場?

12. 網友共同編輯: 評價評論。闢謠(導去真的假的)。錯誤回報。編輯立場與特色標籤(除了選擇與新建標籤外, 寫下理由與參考連結會增加其建議的權重)。共同編輯公民懶人包。

ABow 發想:

1.多立量表



2.中立值(多立值) :適合分享的建議

3.優質多立媒體精選

4.學經歷:如學歷、總經理 > 階級數值? 階級綜合評分?

5.kol 度:某新聞的 kol 關注度(輔助指標)

6.當日 or 最熱門訊息推播 or 通知

7.今日我最夯 or 今天在大聲什麼啦?(頁面裡可能包含今日最熱門的文章(標題)、最熱門的關鍵字、最多回饋的文章)

8.opview audience 公關危機監測 > 感覺很有趣, 想了解是怎麼運作呢?

9.多立島是個位於世界中心的島, 每天都會因為洋流從世界各處漂來的東西, 就這樣一日一日堆積著, 形成了壯觀如博物館的地方, 吸引許多世界各地的遊客來閱覽...(發想緣由: eli 的多立魚、漂流木、動物森友會)

kiko555:

誰做光譜標定?

自動分類的基準是?

新聞的光譜, 是因為機構本身的過往表現來定義嗎?

或是在社群的引用? 但這樣就變成以社群的過往來定義?

二、本週任務

1. 臉書社團爬蟲

a. 回文的內容是重點, 找幾個蒐集的策略讓大家討論

i. 回文大到某種程度後再進行分析

ii. 或情緒數大到某個程度

3/25會議記錄

一、議程:

1.上週任務報告(每人3-5min)

2.腦力激盪:功能

3.排序功能、資訊來源、欄位重要性

立場分析:

1. 是否被泛藍泛綠的粉絲團, 發表者分享

2. [如何程式與人工](#)最後確認去判別發布者立場(請下次開會前看完)

4.程式開發時程討論

二、本週任務

2. 臉書社團爬蟲

a. 麻煩整理所有資訊欄位

b. 大小規模的社團是否會影響速度

c. Selmium的headless的效能

d. 2小時append一次資料

3/18會議記錄

可能合作方:

1.零時檔案局

<https://g0v.hackmd.io/@chihao/0archive/%2FcdctnMJWQpKWQYhSxB8sCw>

https://airtable.com/shrd0utGHlTWmQsYt/tbl3DrYs5mXgl0EV9/viw2cuXweY8OxNkX6?block_s=hide

2.真的假的

本週任務

1.整理腦力激盪:功能

eli功能:

1. 列表:你搜尋的關鍵字當天文章列表、你搜尋的關鍵字這七天熱門文章列表(資料庫還是會紀錄所有時間的數據, 但不運用到前台)、這七天熱門文章列表、前七天發表的熱門文章列表

2. 列表欄位:文章標題。發文時間。管道。發布者(名稱、id;年齡;國家:如IP來源地;產業:如新聞業、自媒體、個人;學經歷:如學歷、總經理;宗教)。聲量。爆發力(斜率)。過熱度(不

太合理的熱度或聲量爬升)。正反面。情緒。立場標籤(親民進黨。親國民黨...等)。特色標籤(旅遊、政治...等)。資訊列表標籤(社群管道、高機率會熱門、常發文發布者...等)。標題修改記錄。真實性(真的假的)。30字文章大綱。

3. 列表欄位(以2-4個標籤為一組的版本):顯示文章為3則,展開為10則。可選擇以哪3則作為分享或比較用
4. 各欄位可篩選呈現,與設定權重,可分享篩選後的頁面到社群(這邊要加強行動誘因)
5. 評價與評論功能:數據也會被加入聲量中
6. 公民懶人包跑馬燈:
7. 相關資訊
8. 自動擷取文章中100字摘要
9. 自動highlight需要彼此比對的文字片段
10. 關鍵字標籤雲
11. 會員機制:可儲存個人化列表,喜好文章,參與其他開放編輯的功能
12. 可去除類似資訊,類似立場資訊
13. 儲存機制:個人篩選的設定、喜好文章
14. 資訊內容判斷:所有內容文字。聲量。正反面。情緒。自動擷取重要關鍵字(可作立場分析)。
15. 共同編輯的方式與範圍?
16. 讓使用者想用什麼管道,就要用什麼社群帳號登入。這樣才能抓臉書 line和ig的資料?
17. 介面方便比較不同立場或標籤

Eli輿情系統分析

目標	資訊來源	分析	
Google Trend			
keyhole			
engaged		帳號影響力、貼文、競爭者ig、hashtag、限時動態 帳號影響力:貼文數、粉絲數、其他平均粉絲數、增長、時間趨勢圖、性別、年齡 貼文:平均觸及、平均瀏覽、平均互動率、時間段篩選、各貼文發布時間、觸及、瀏覽、愛心數、成效指標、留言、珍藏、每人平均瀏覽次數 競爭者ig:少了性別年齡,有平均愛心,平均留言 限時動態:多了平均看完率、用戶退出率	
q search		監測關鍵字超過聲量提醒:寄到信箱 每日熱門tag 關鍵字的:文章與連結列表。文章數,按讚與心情數,分享數,平均關注度:按讚與心情數/文章數。平均傳播力:分享數/按讚與心情數。各心情時間趨勢圖。前十個有分享的粉絲頁或個人。前十熱門文章列表,前	

		十熱門連結列表。前五熱門文章的前十熱門留言 時間軸看各頻道的頻率 散佈力 貼文內容  與該組關鍵字相關的關鍵字雲 設定的關鍵字所擷取的文章, 可以設定自己新增的標記ex:競品 我家	
qsearch linkbundle 受眾管理 工具		用縮網址, 取得點擊者的數據變成廣告用受眾 對不同縮網址下tag, 之後可分析點擊者的各個標籤 ex購物傾向	
蛛思輿情 分析		選擇關鍵字提供超過某聲量的警報。查出散佈來源與 途徑。文字的情緒分析。競品找文案或是營銷策略(用 樂銷的圖形辨識, 與這些企業合作) 文章數。主文情緒。留言情緒。頻道廣度:新聞頻道 數。社群頻道數。互動數 主文或留言的關鍵字。熱門或竄升 關鍵字網路圖:該關鍵字連到其他相關關鍵字	
透事		文本的自動摘要, 情緒感測, 重要詞彙萃取	
keypo		七天熱門 24小時熱門 分成社群 新聞 討論區 部落格 欄位:三種情緒 回文數 按讚數 分享數 聲量時間趨勢圖 情緒時間趨勢圖 情緒 熱門標籤雲 社群(社群 討論區 部落格)活躍度:與新聞聲量做比 較 競品比較(關鍵字):各管道筆數, 社群活躍度, 網路好 感度, 聲量時間趨勢圖	
fando	臉書	競爭粉絲頁:貼文類型分佈(貼文 圖片 影片 連結 直 播 活動), 粉絲數等時間趨勢圖。粉絲數 回應數 貼文 數 分享數 單一粉絲的相關互動資訊	
opview audience	ptt dcard m01	@opview	

	babyhome 巴哈 新聞 臉書等	報表: 文章列表, 頻道(ex各粉絲頁, 各版)統計, 關鍵領袖, 社群活躍度, 正負面情緒, 聲量趨勢, 資訊來源分佈, @公關危機監測 @opview audience 各標籤族群在各資訊來源的比例 各資訊來源的高聲量文章列表 數據: 聲量, 討論網友數, 關鍵發布者回文數, 擴散頻道數, 好感度, 社群活躍度 如何標記使用者?	

ABow:想像了一下該有什麼功能, 示意圖初稿如下頁:

公民資訊跑馬燈

← 台灣防疫成績全球關注! →

?

⚙

📄 文章標題

🔥 聲量

📍 來源

⚠ 可信度

👁 事實查核

😊 情緒

▶

口罩2.0流程不變'3月25日...

23.5K

衛福部

100%

1則

😊 👍



中央流行疫情指揮中心今(23)日表示,「口罩實名制2.0」將於3月25日起進行第二輪線上預購,歡迎有需求的民眾上網訂購。

指揮中心指出,第二輪預購流程與第一輪相同,期程縮短,使民眾可儘早取得口罩:
3月25日(三)上午8時至3月27日(五)晚上8時,使用健保卡/自然人憑證或健保快易通APP進行認證預購。
成功預購者於3月28日(六)上午8時至3月30日(一)晚上8時內以ATM轉帳、信用卡方式完成付款...

▼

口罩2.0流程不變'3月25日...

23.5K

衛福部

100%

1則

😊 👍

口罩2.0流程不變'3月25日...

23.5K

衛福部

100%

1則

😊 👍

公民資訊跑馬燈

← 台灣防疫成績全球關注! →

文章標題 聲量 來源 可信度 事實查核 情緒

口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
中央流行疫情指揮中心今(23)日表示,「口罩實名制2.0」將於3月25日起進行第二輪線上預購,歡迎有需求的民眾上網訂購。 指揮中心指出,第二輪預購流程與第一輪相同,期程縮短,使民眾可儘早取得口罩: 3月25日(三)上午8時至3月27日(五)晚上8時,使用健保卡/自然人憑證或健保快易通APP進行認證預購。成功預購者於3月28日(六)上午8時至3月30日(一)晚上8時內以ATM轉帳、信用卡方式完成付款...					
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍

公民資訊跑馬燈

← 台灣防疫成績全球關注! →

文章標題 聲量 來源 可信度 事實查核 情緒

口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
中央流行疫情指揮中心今(23)日表示,「口罩實名制2.0」將於3月25日起進行第二輪線上預購,歡迎有需求的民眾上網訂購。 指揮中心指出,第二輪預購流程與第一輪相同,期程縮短,使民眾可儘早取得口罩: 3月25日(三)上午8時至3月27日(五)晚上8時,使用健保卡/自然人憑證或健保快易通APP進行認證預購。成功預購者於3月28日(六)上午8時至3月30日(一)晚上8時內以ATM轉帳、信用卡方式完成付款...					
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍
口罩2.0流程不變*3月25日...	23.5K	衛福部	100%	1則	👍

3/14會議記錄:感謝各位今天新入坑的夥伴

- 1.請各位沒參與到的夥伴,可以到下方填寫自我介紹
- 2.本週任務:
 - a.思考還有哪些需要解決的問題,直接記錄在下方
 - b.搜尋一下可以用到的程式資源
- 3.下週三晚上3/18(三) 21:30-22:00
 - a.參與者可從本文件了解最新更新
 - b.最簡單的聯繫可以
 - >加telegram @eli_h1010 or https://t.me/eli_h1010
 - (直接說「嗨」就可以囉。eli會提醒你參加相關事宜)

-> g0v-tw.slack.com #neutral-info

c.會議軟體Zoom, [官網](#)。每週三21:30會在此更新會議連結

一、共筆連結: (資料會在本文件與slack中做更新)

1.Slack: g0v-tw.slack.com #neutral-info

<https://app.slack.com/client/T02G2SXKM/CUVD7R29W>

沒有slack帳號的, 要先註冊與登入(網頁版要填入g0v-tw)

二、3/14線上松

1.11:00-12:00今天會議連結 <https://intro.g0v.ronny.tw/meet/channel/g0v-hackath38n/14>

2.[簡報連結](#)

3.線上松議程:

a.腦力風暴:像是什麼filter是大家在意的? 聲量。管道。...。功能? 可供借鏡服務?

c.短期任務規劃

d.下次線上開會:3/18(三) 21:30-22:00