2024 總統大選網路聲量觀察

#### 背景/動機

2024將迎來第七次總統大選，目前包括新聞、網路媒體已經開始預熱，皆有出現相關的新聞報導與討論，希望能透過資料的討論度了解民眾對於這次大選議題的關切度，以及關心的主題與方向。

#### 資料來源

資料來源：蘋果日報-政治版

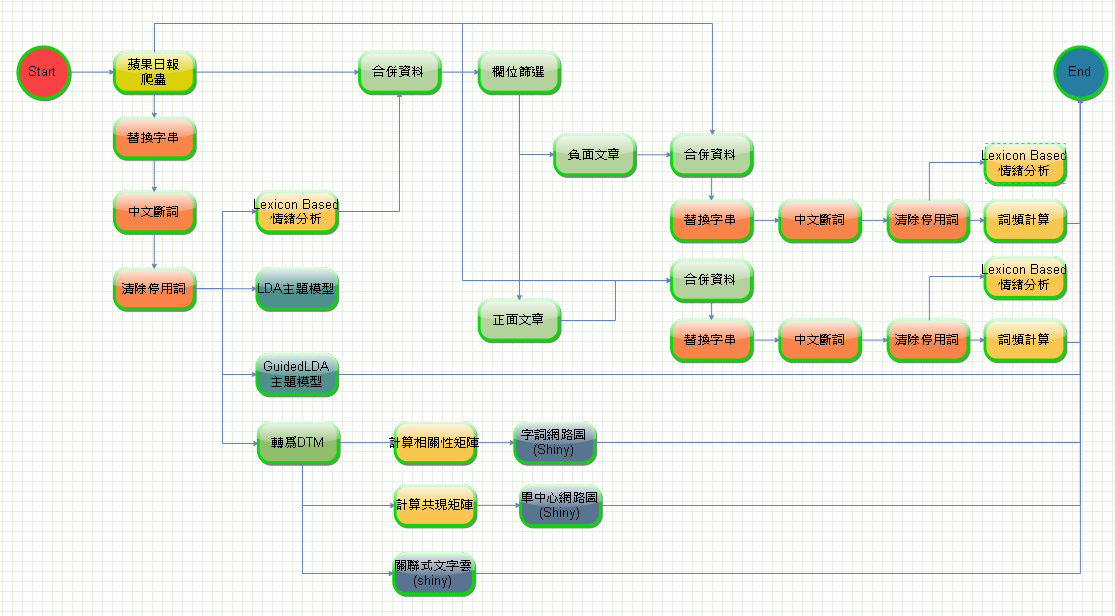
資料範圍：2023/03/01~2023/05/01

搜尋關鍵字：總統大選、總統參選人

資料筆數：205筆

#### 系統流程

流程名稱：Test03



#### 分析過程

* 資料清理

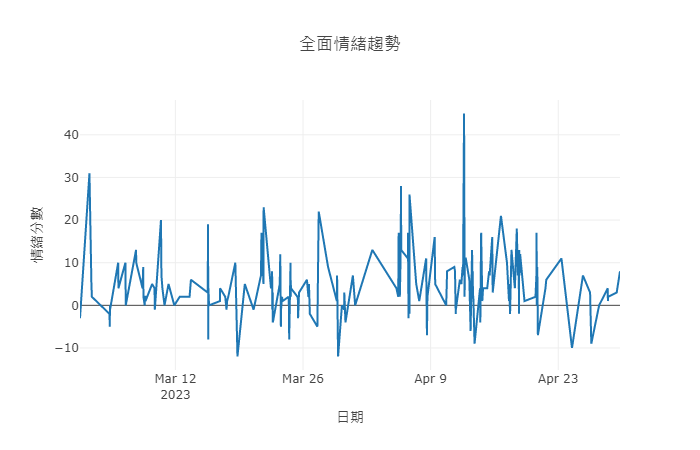
前處理的部分一樣是透過替換字串->中文斷詞->清除停用字的方式，對蘋果新聞爬蟲的資料結果進行初步資料清理。

#### 分析角度

* 文章主題分析-情緒分析

整體情緒折線圖

從圖中的情緒分數來看，推論目前總統大選的相關報導大多屬於正面的情緒，負面情緒的報導佔少數。

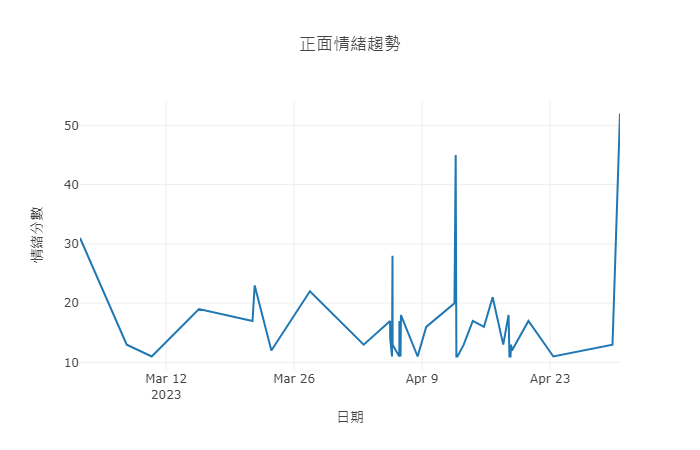


整體詞頻計算

主要的兩大政黨，及討論度高的預期參選人，在文字雲都相當明顯。

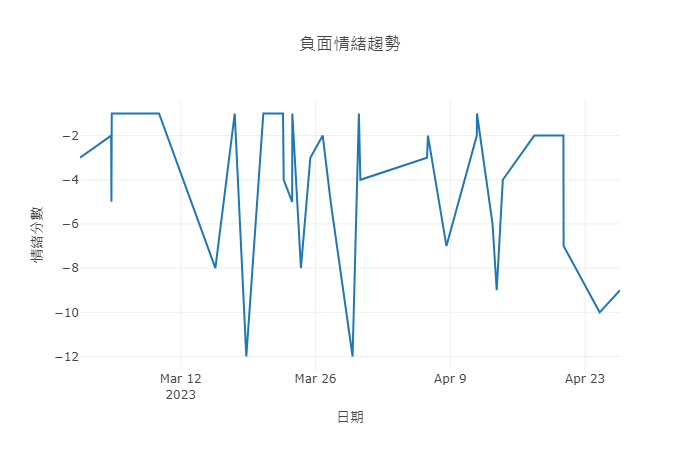


正面情緒折線圖



正面詞頻計算

負面情緒折線圖



負面詞頻計算

不論是在正面或是負面的詞頻上，國民黨、台灣、總統、侯友宜這些字詞都很常被提及。可以了解大眾對於國民黨抱有比較負面的印象。

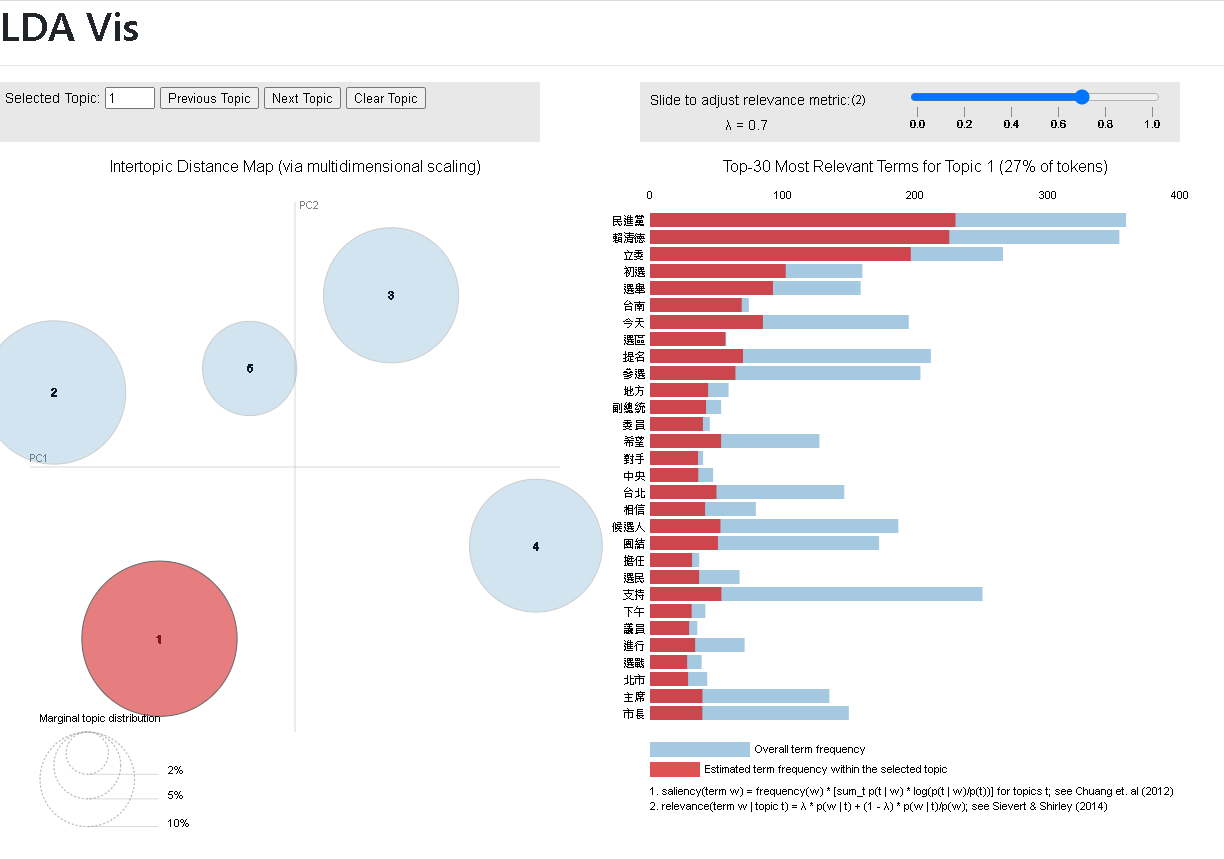


* 文章主題分析-LDA主題模型

LDA

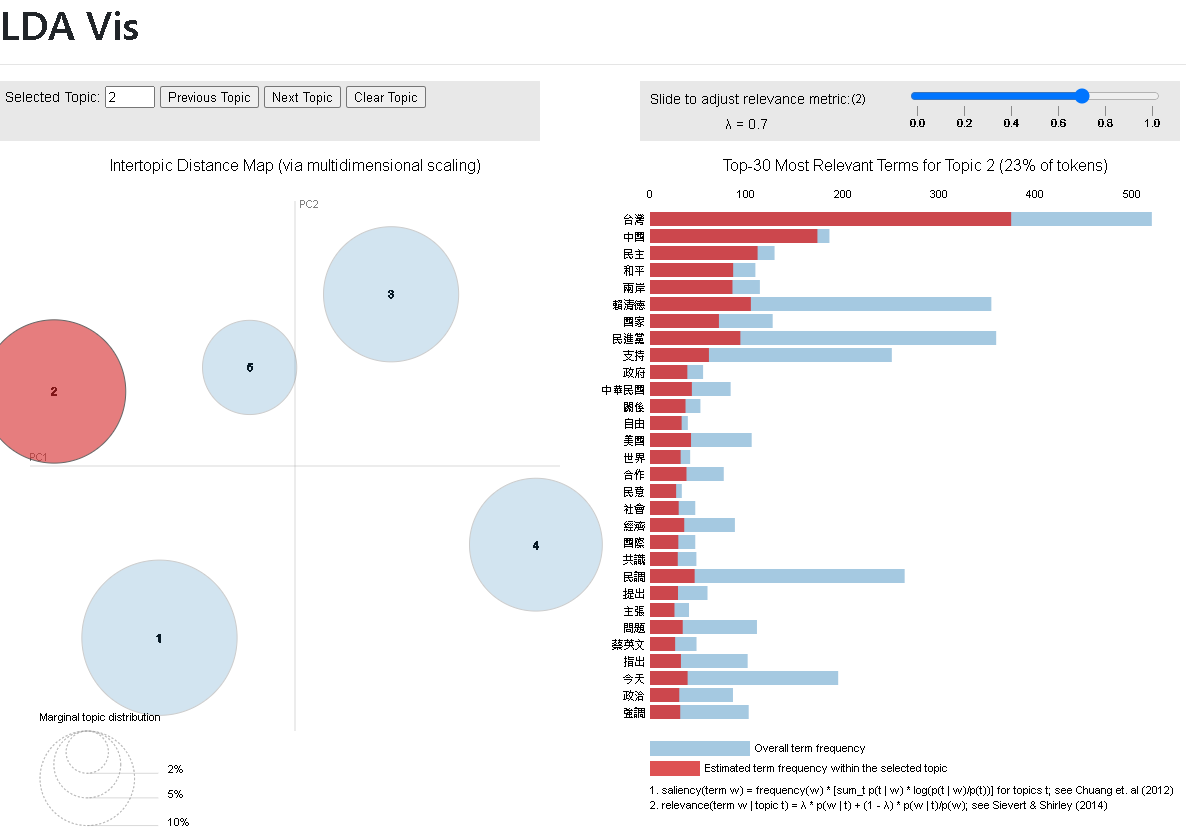
1.民進黨參選狀況

在第一個主題中，可以看到前三名字詞為民進黨、賴清德、立委，因此初步認為是與民進黨相關的選舉主題，而後又見初選一詞，故認為此主題是與民進黨的黨內參選裝況。



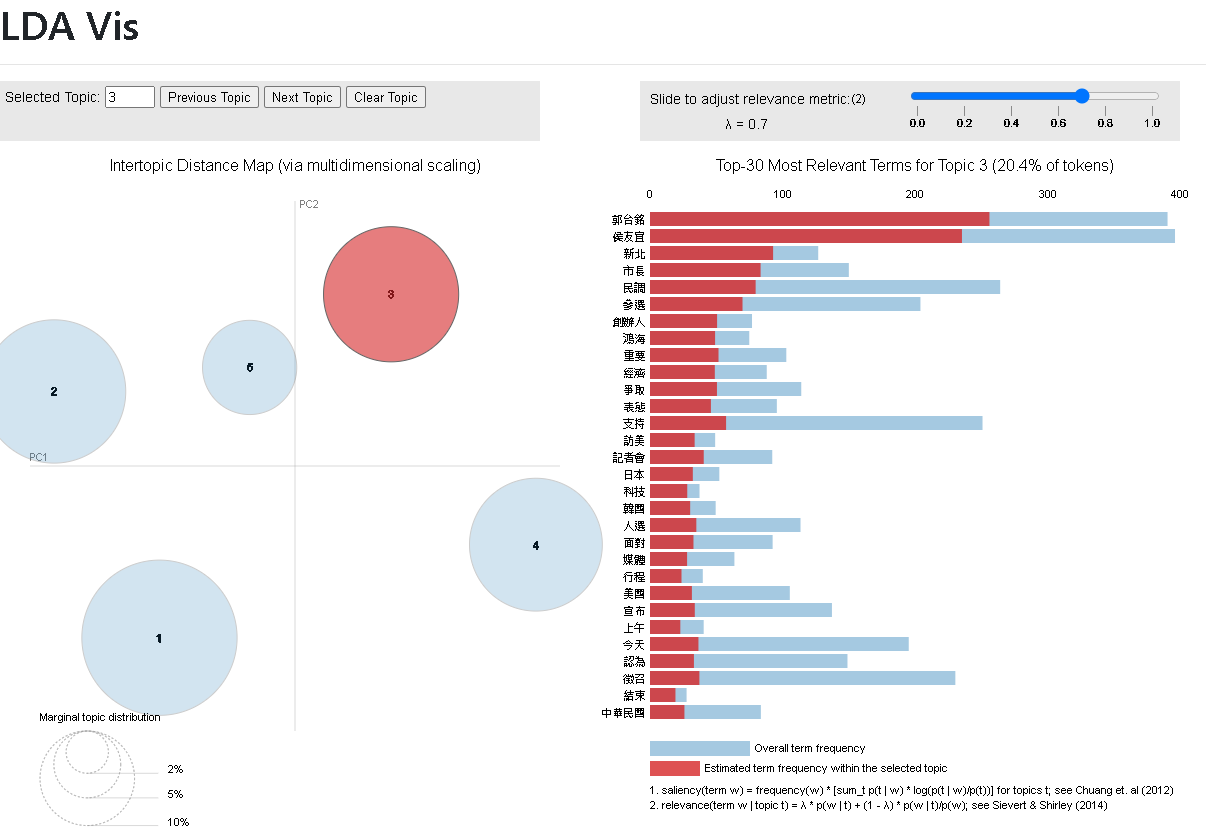
2.兩岸關係

第二個主題當中，台灣、中國、民主這些字詞在為前三名，後續又見到和平、兩岸的字詞，因此推估此主題應是聚焦在兩岸關係的討論上。



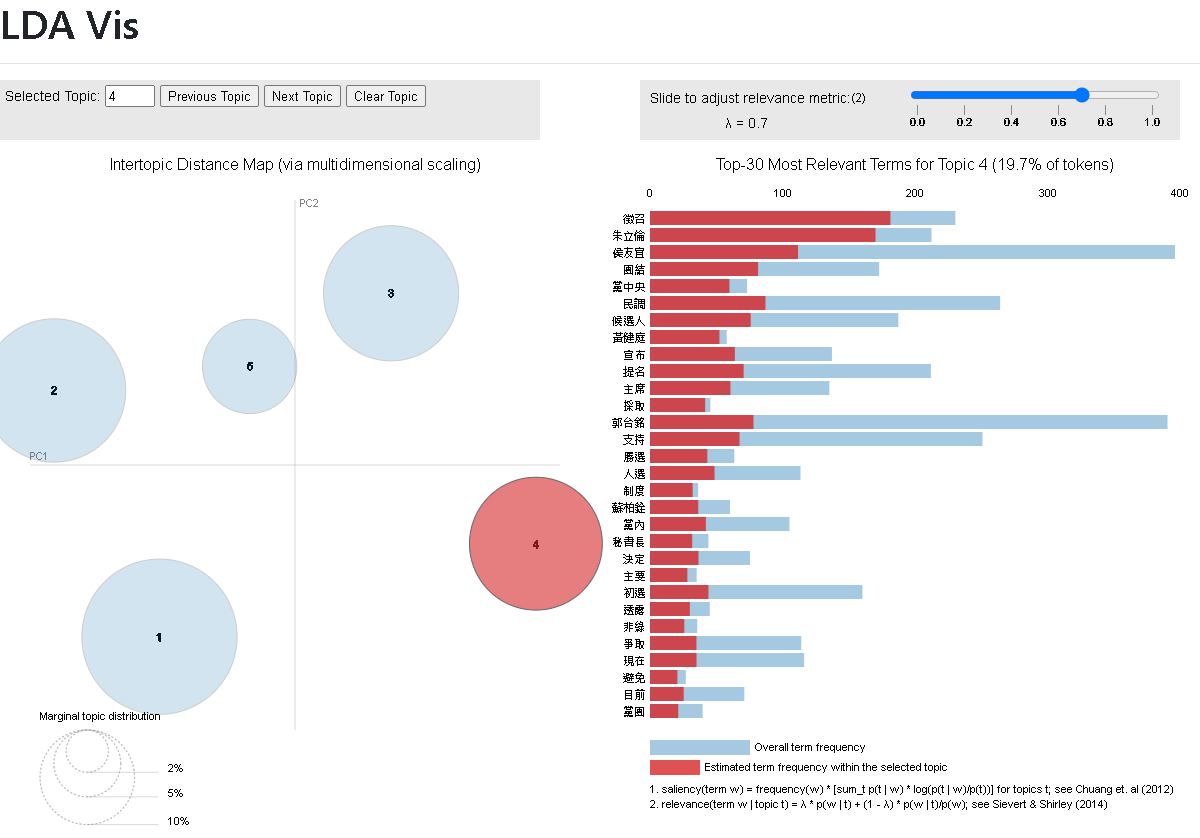
3.國民黨參選狀況

在第三個主題當中，郭台銘、侯友宜這兩個名字出現頻率高於其字詞非常多，並且民調、參選的字詞也出現的頻率排名也較靠前，因此認為此主題主要是與國民黨的參選狀況有關。



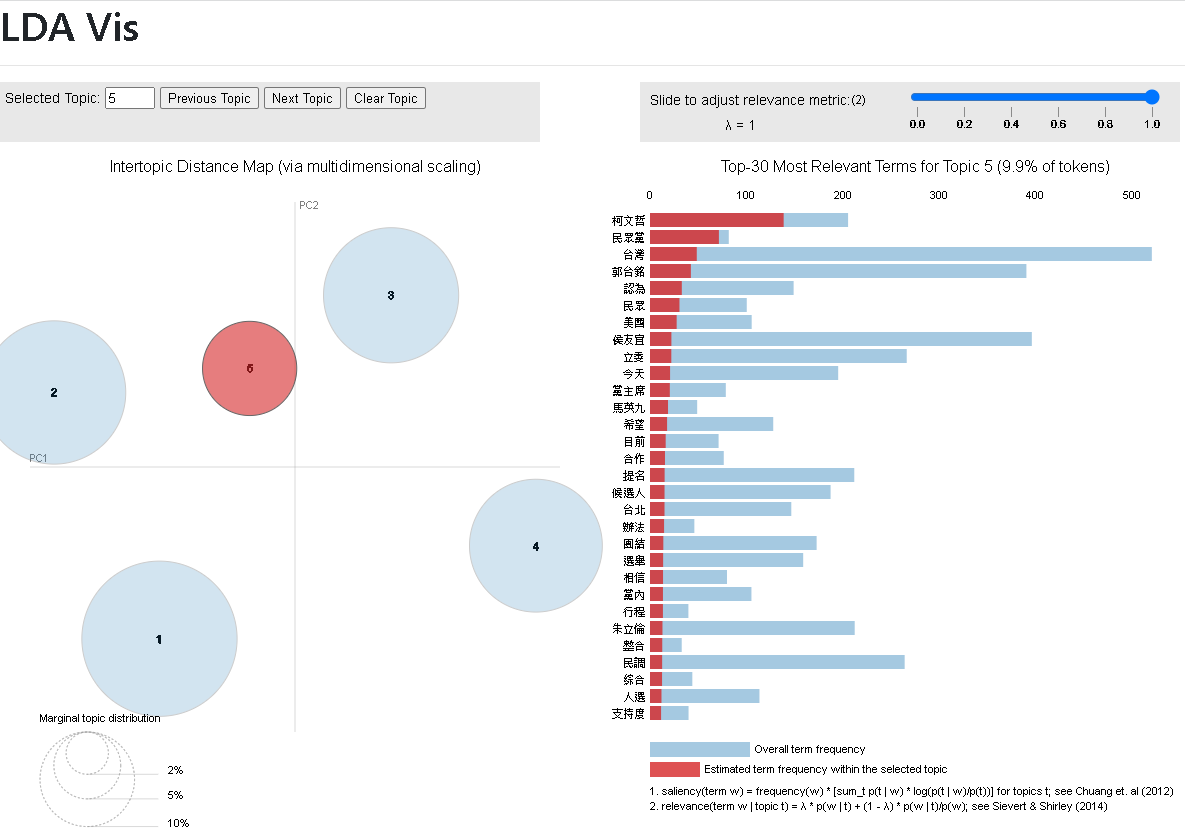
4.國民黨黨內初選民調

於第四個主題中，徵召、朱立倫、侯友宜、民調這些字詞為前幾個出現頻率較高的字詞，朱立倫為現行的國民黨黨主席，再加上徵召一詞，因此認為此主題是在討論國民的黨內初選。



5.民眾黨

最後第五個主題當中，很明顯的可以看到柯文哲、民眾黨的是出現頻率高的字詞，因此不難得知此主題是與民眾黨的參選相關。

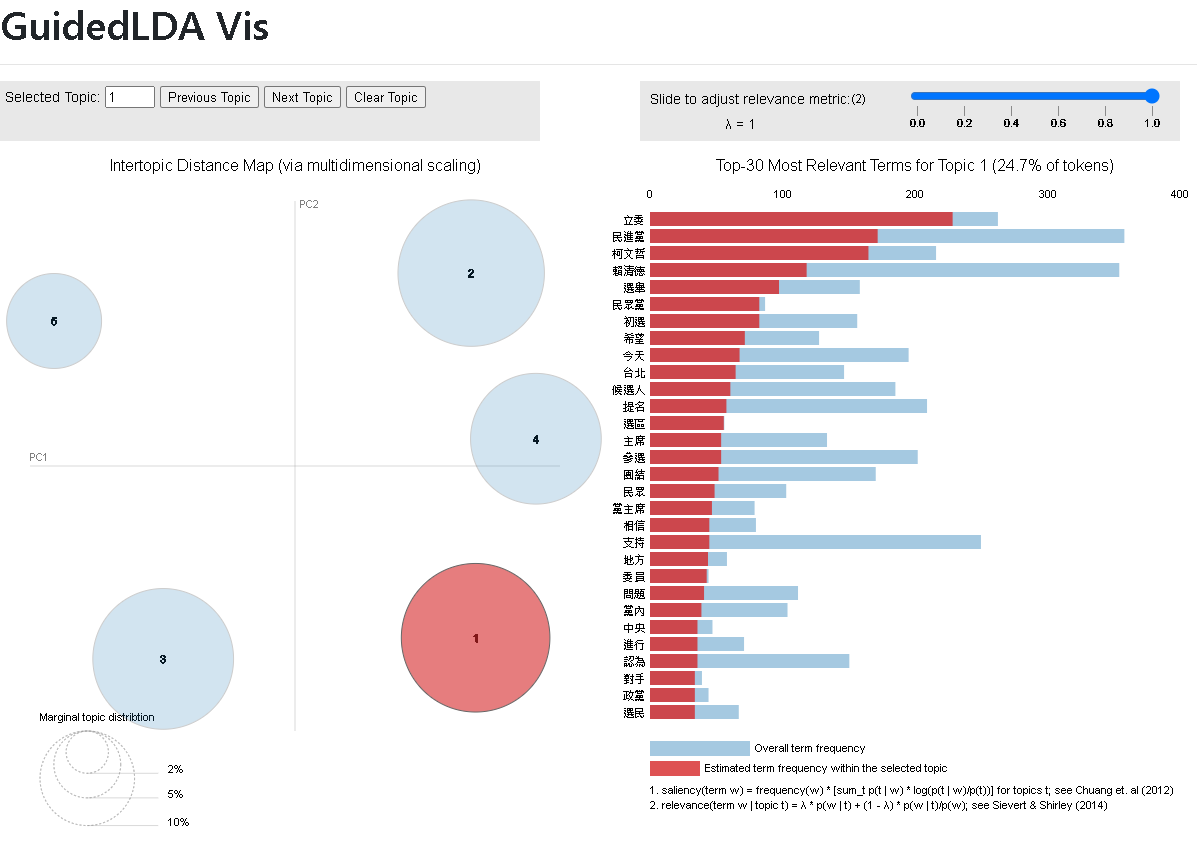


* 文章主題分析-GuidedLDA主題模型

GuidedLDA

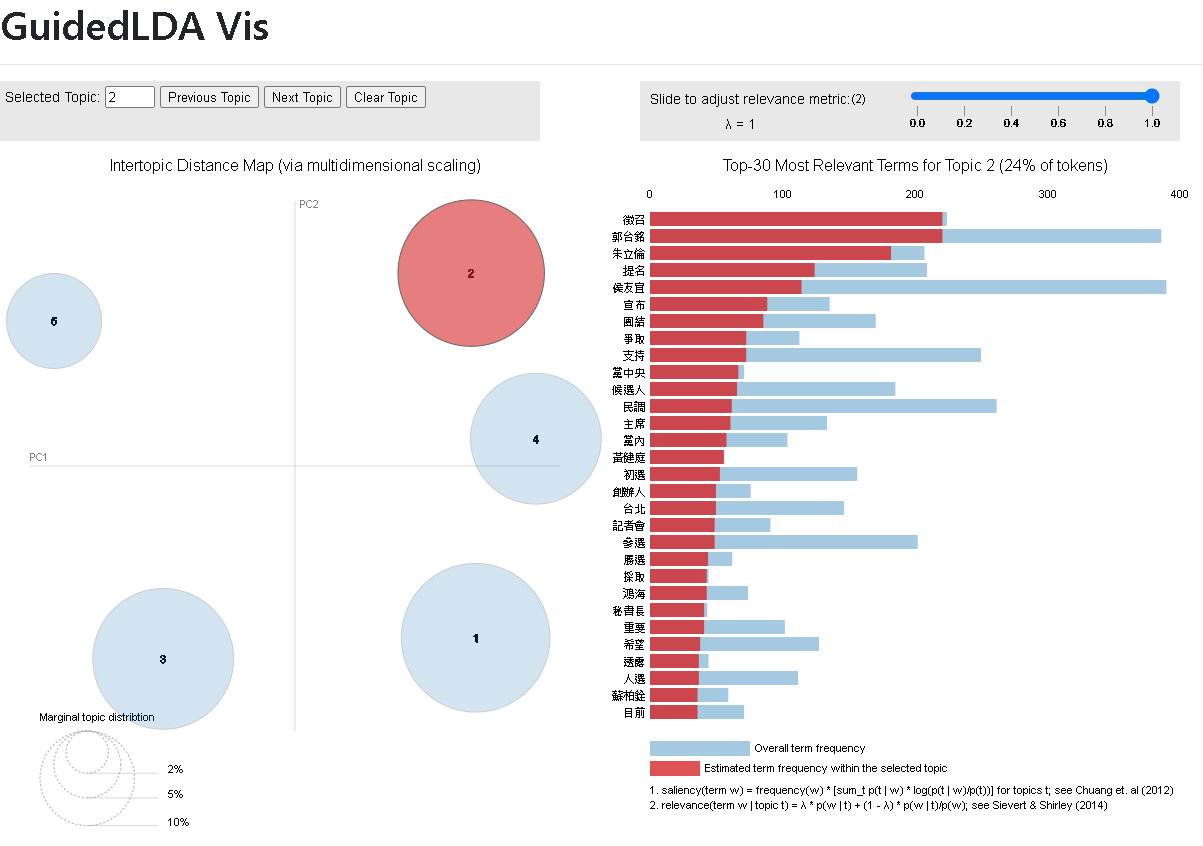
1.民進黨VS民眾黨

在這個主題中，我們可以看到代表的民眾黨的柯文哲以及民進黨與賴清德的字詞頻率較高，而立委這個詞則是排名第一，由於2024 年總統大選與立委選舉是以一起的，因此可知這個主題是與立委選舉相關，並主要聚焦在民進黨與民眾黨上。



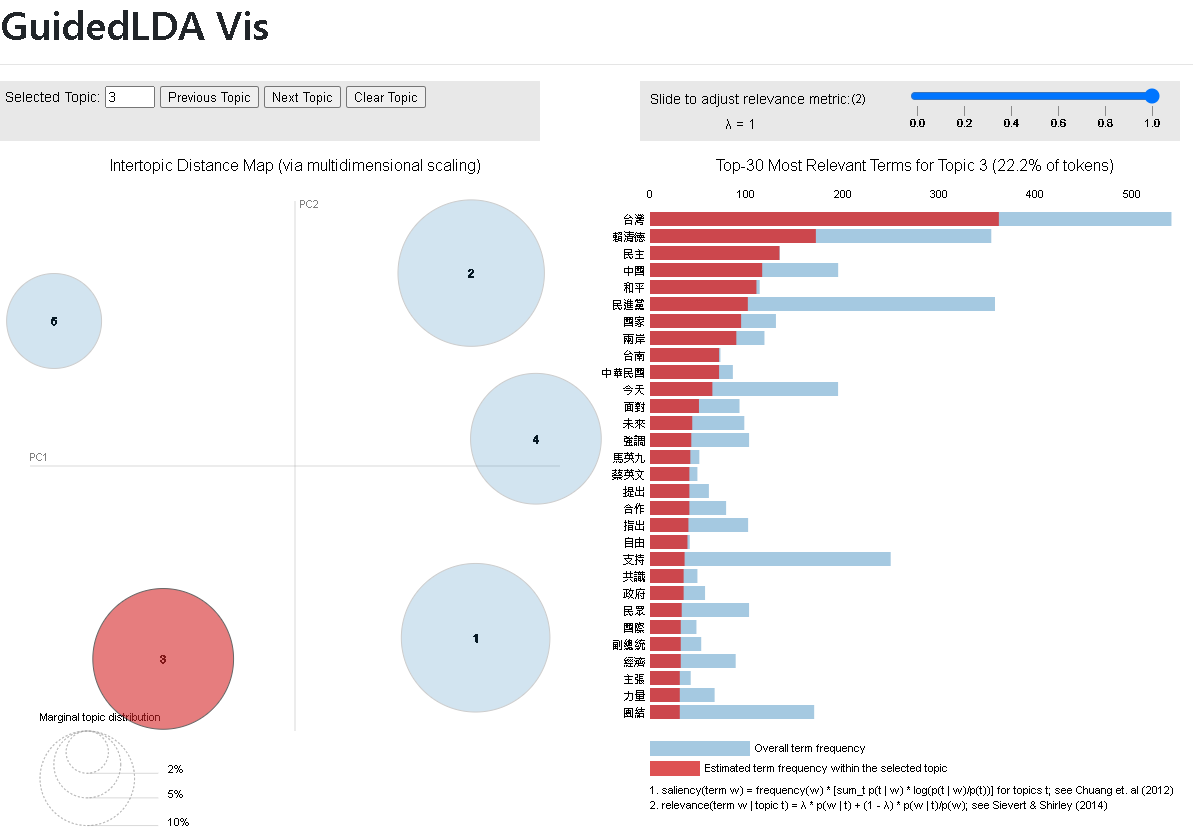
2.國民黨黨內初選提名

在第二個主題當中，可以看到徵召、郭台銘、朱立倫這些討論度高的字詞，因此從與身為黨主席的朱立倫相關且又與徵召相關聯，來判定此主題是在討論國民黨的黨內初選提名。



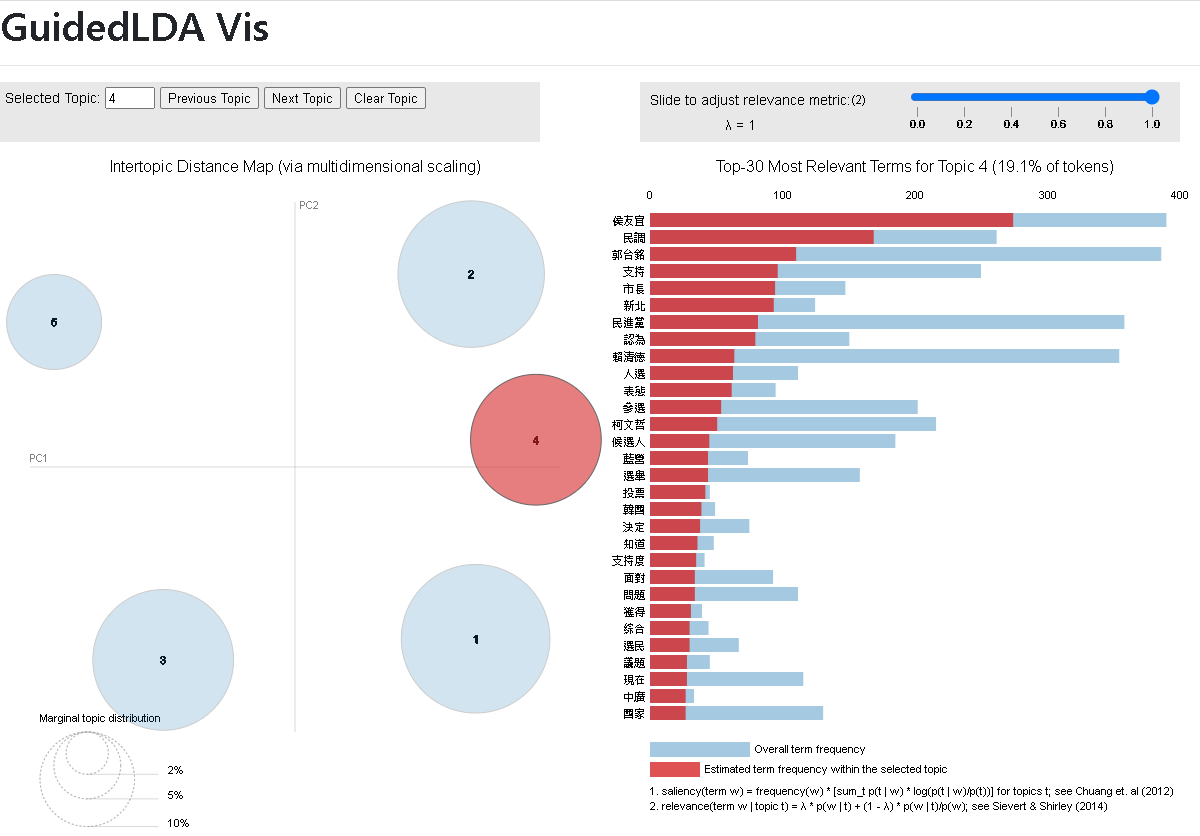
3.兩岸關係

在第三個主題裡面，台灣這字詞出現頻率最高，而中國、兩岸、台商的詞頻也偏高，因此認為此主題應為兩岸關係的討論相關。



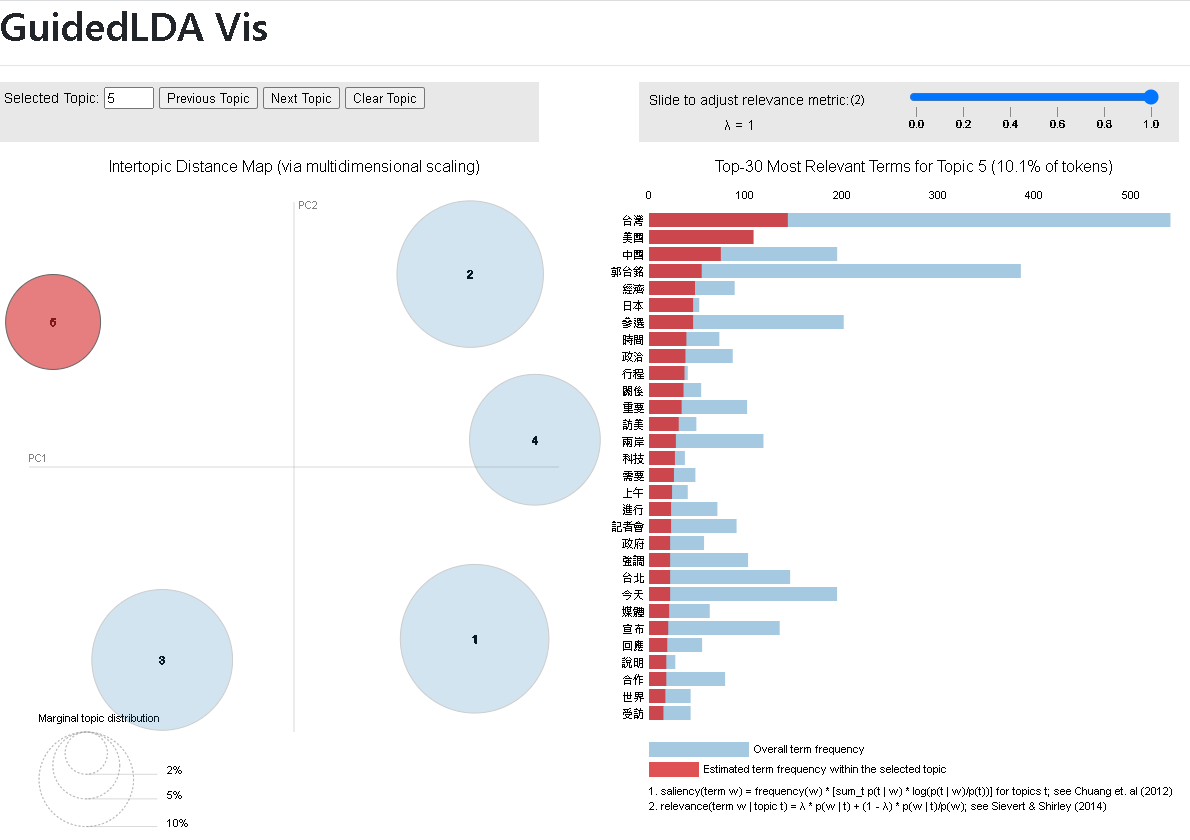
4.國民黨黨內初選

一樣可以看到侯友宜、郭台銘、民調為前三高的字詞，因此也判定此主題與國民黨的黨內初選相關。

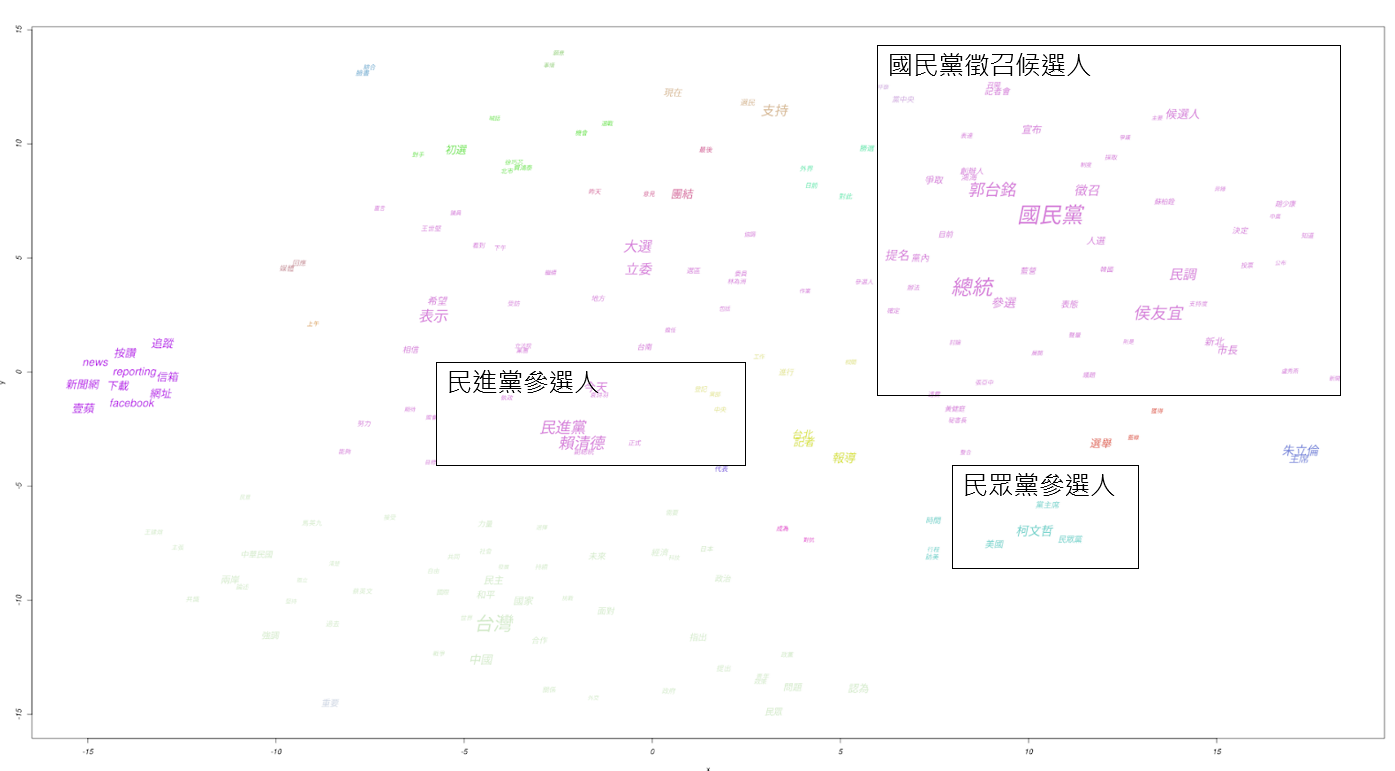


5.美中台角力

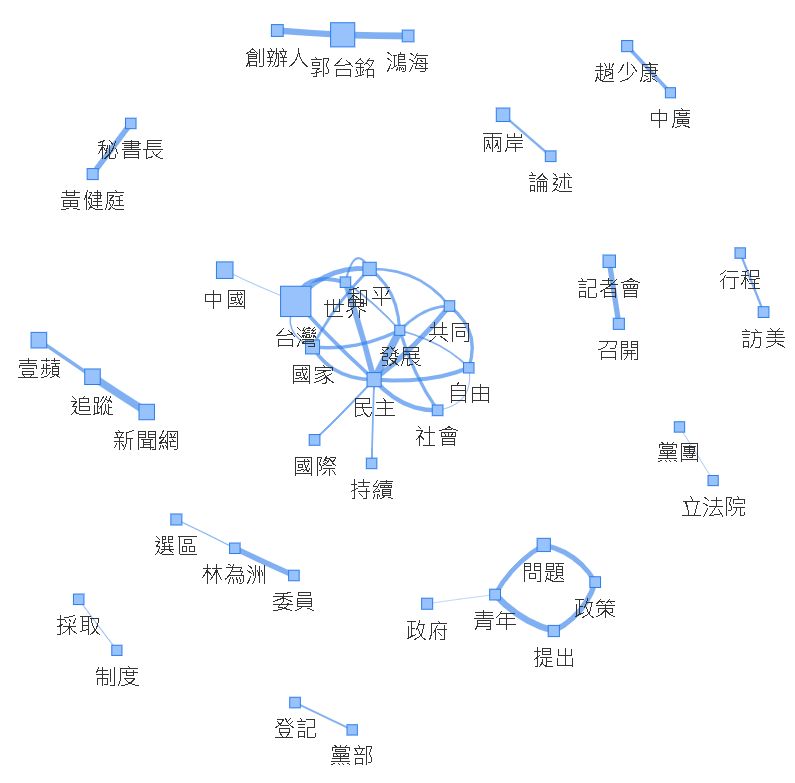
在這個主題當中，可以看台灣、美國、中國、日本這些國家字詞的頻率較高，由於台灣總統大選也會被國際關注，尤其是有利益關係的相關國家，因此這個主題主要是台灣大選與其他國家有關的討論。



* 文章主題分析-關聯式文字雲

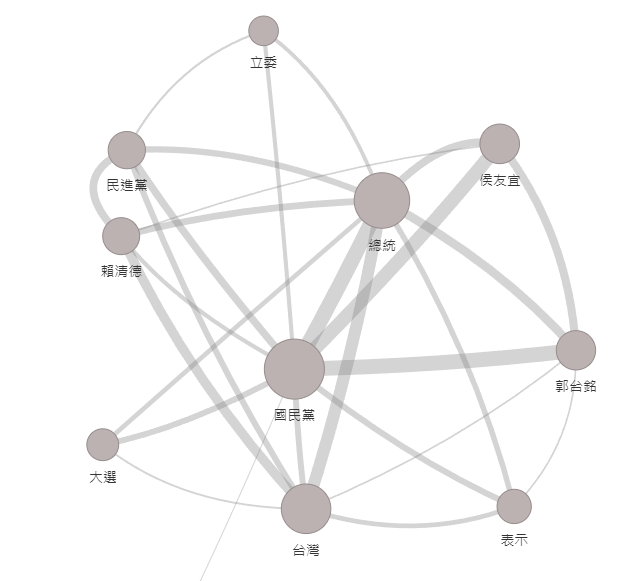


* 文章主題分析-字詞網路圖



由字詞網路圖可以看出討論台灣的字詞關聯性是最多的，其中包含了中國、世界、民主、和平、發展...等，顯示民眾在討論總統大選議題時，還是會認同選舉是民主國家中重要的自由表現，為社會帶來穩定和和平的發展。

* 文章主題分析-單中心網路圖



由單中心網路圖可以看出本次總統大選的議題中，討論關聯程度最高的為國民黨、總統、台灣以及其他參選人等，也符合現在目前的新聞熱度。

第九組成員：

N104020003 王博玄

N104020004 葉子豪

N104020026 林冠宏

N104020028 姜孟嫻

N104020030 林欣民

N104020027 林欣穎

N104020029 陳勇任