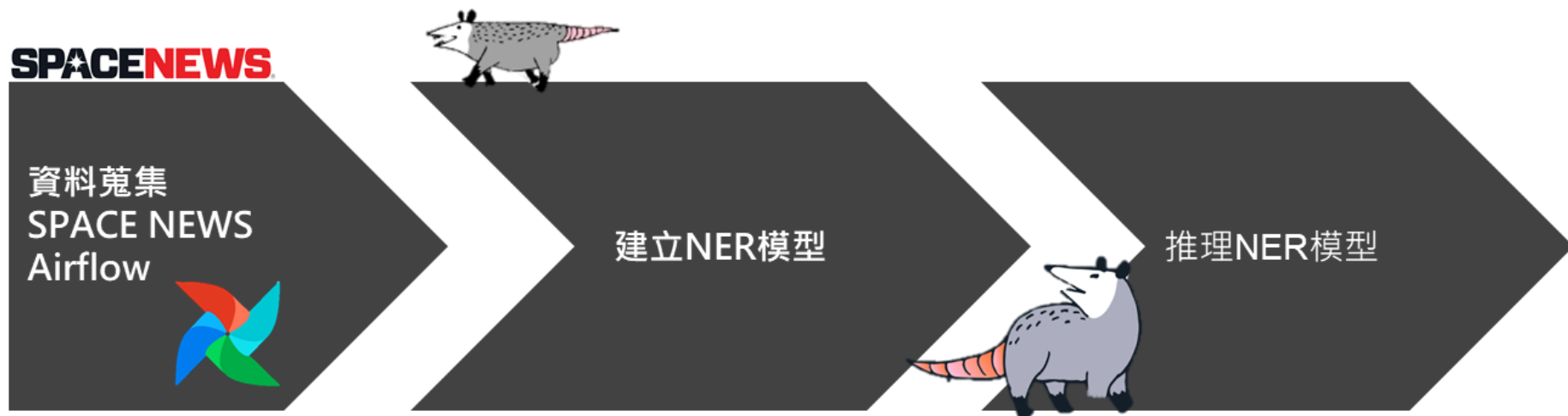


# 專案目標

- 自動化國防新聞監控平台，使得分析工具能保持在最新的趨勢
- 透過命名實體辨識（NER）改善編撰字典花費過多人力時間的困擾



# 自動化分析流程

- 導入工作流程管理系統（**Airflow**）：將過往抓取文集至產生分析資料的流程自動化。
- 使用文字探勘技術幫助建立領域字典：透過命名實體辨識 **Named Entity Recognition (NER)** 任務自動抓取重要的技術、公司、組織、衛星等。

## 自動化產業新聞監控更新平台

爬蟲程式抓取

資料清理與自然語言處理

分析資料轉換

視覺化互動工具



## 自動化產業新聞監控更新平台

### 趨勢分析系統

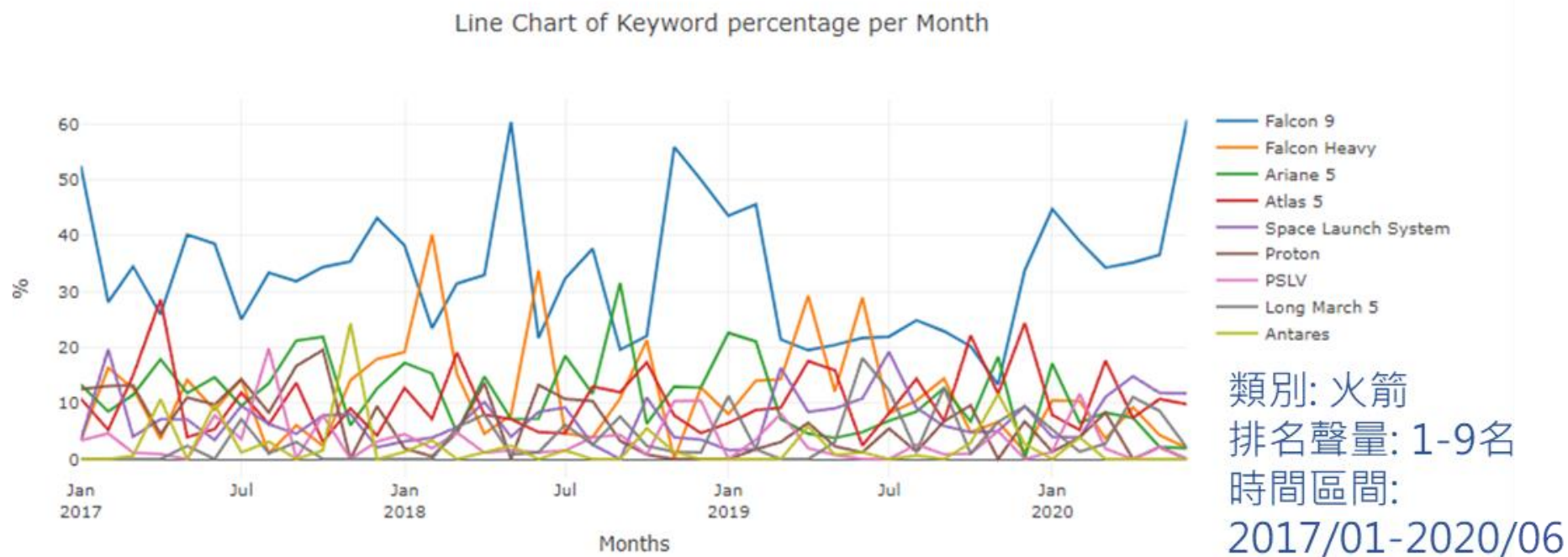
資料來源：SpaceNews

聲量趨勢

關鍵字排名

字詞網路圖

這個介面我們呈現特定類別的關鍵字，在不同時間區間，所出現的次數及佔比。  
另外有一個名次篩選功能，藉由該功能我們可以縮小觀察的範圍



## 國防 SpaceNews 關鍵字詞聲量趨勢查詢

### 關鍵字趨勢

選擇類別

☐ 一般術語 ☐ 公司 ☐ 組織 ☐ 衛星 ☒ 國家地區 ☐ 火箭

選擇關鍵字排名區間



起始日

2017

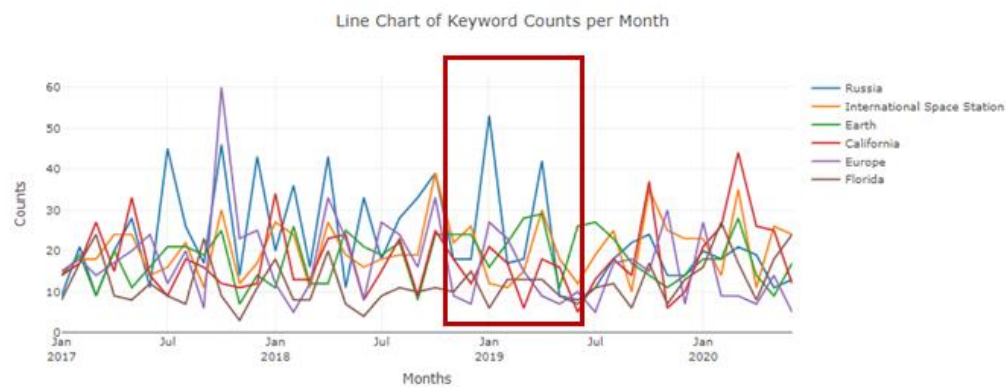
一月

終止日

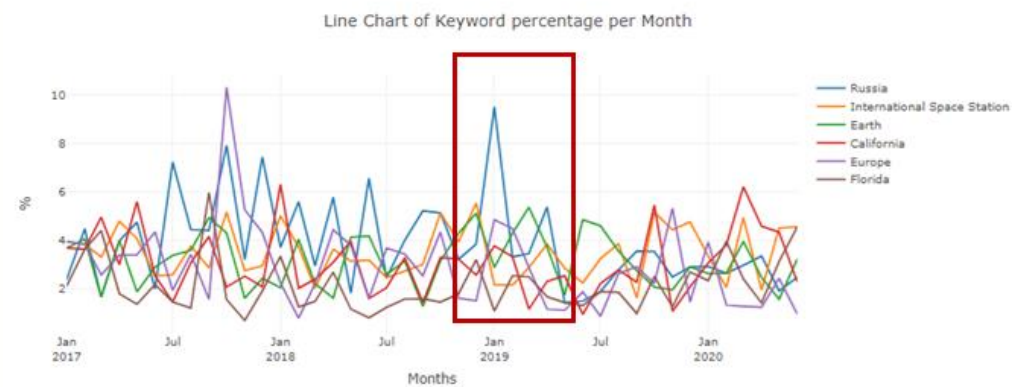
2020

六月

#### 每月文章篇數



#### 每月文章比重



左圖以次數作為基本單位

右圖為所佔之百分比(特定關鍵字的當月出現次數/當月同類關鍵字的總出現次數)

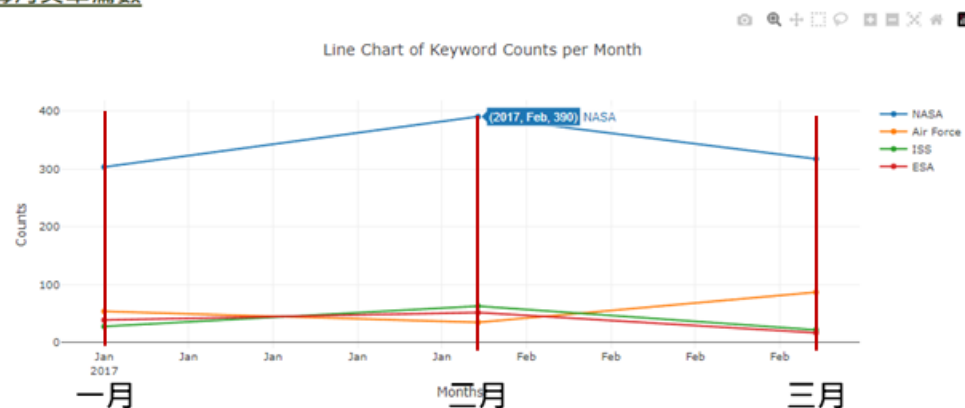
兩途乍看之下相同，實際上則有所差異

選擇類別  
☐ 一般術語 ☐ 公司 ☒ 組織 ☐ 衛星 ☐ 國家地區 ☐ 火箭

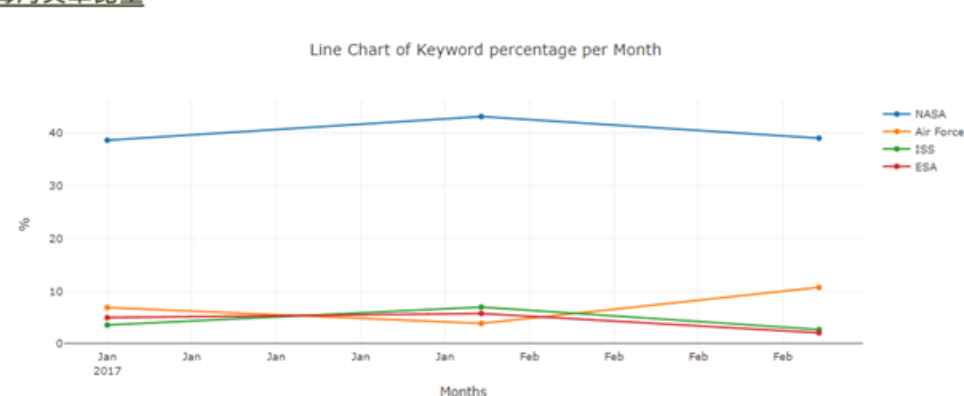
選擇關鍵字排名區間  
☒ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

起始日 2017 一月 終止日 2017 二月

每月文章篇數



每月文章比重

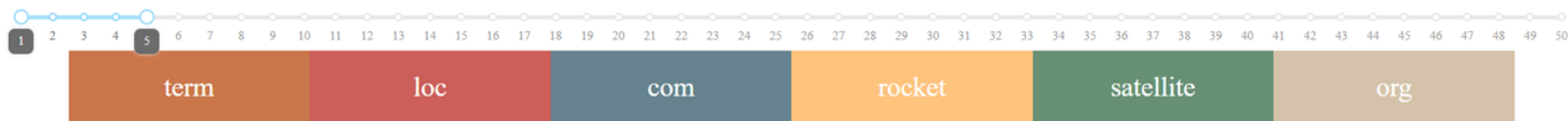


Ex. 我們現在所檢視的是在當期(2017/1-2017/3)所屬組織關鍵字中出現頻率最高的1-4名歸類於組織的關鍵字 在2017年二月區間總共出現 904 次  
 其中出現次數最多的NASA佔390次 故其在2017年二月的組織佔比為 $390/904 = 43.14159\%$   
 同個月組織關鍵字出現次數第二名Air Force 共出現63次 佔 $6.969027\%$

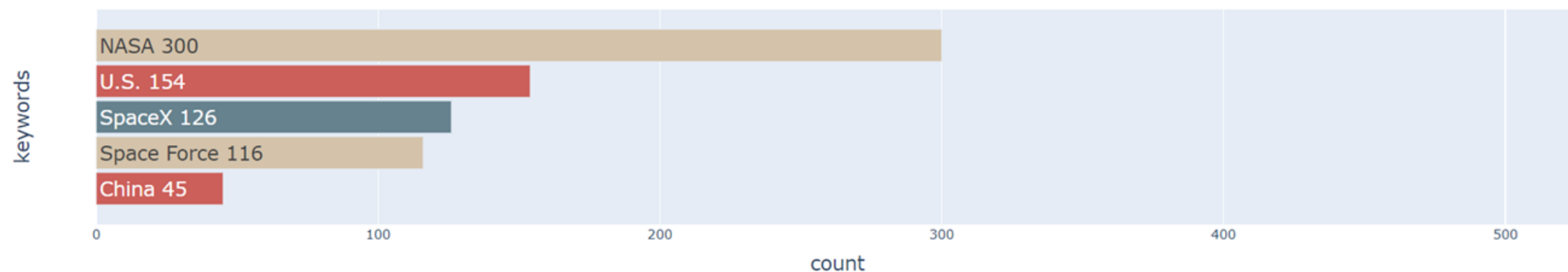
# 關鍵字聲量排名

## 國防SpaceNews關鍵字詞頻率排名

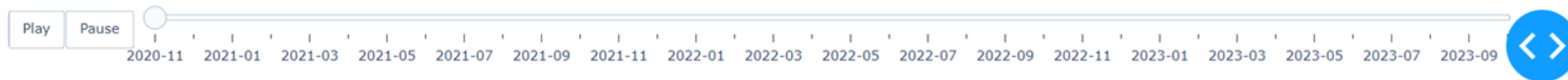
選擇關鍵字排名區間



Amounts of top 1 to 5 keywords appear per month

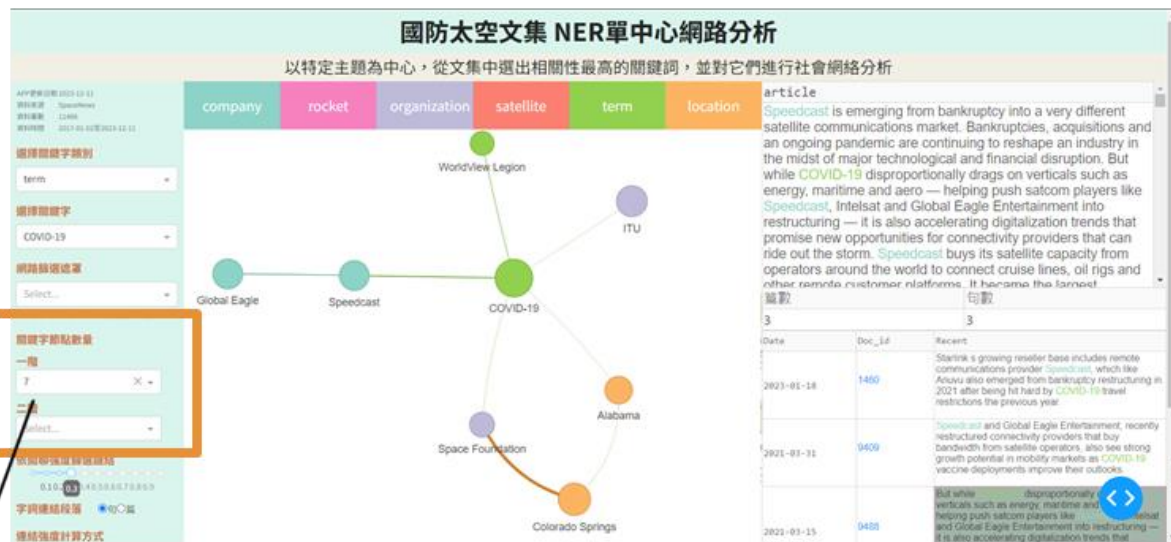


Year-Month: 2020-11





# 單中心網路



點選邊或節點，會出現該詞彙的實際補助案內容

目標：了解關鍵字之間的關聯性，或是選擇特定詞類的關聯。

圖像意義：

不同顏色顯示不同的關鍵字分類，以節點大小呈現它的出現頻率，而節點與節點之間的鏈結可以表達共同出現次數或是相關性，當數值越大，鏈結就會越厚實。

工具篩選功能：

從關鍵字類別選擇關鍵字、篩選只顯示特定詞類、設定節點數量、依關聯強度篩選鏈結、字詞連結段落、連結強度計算方式、一二階功能整合、字體大小調整

關鍵字節點數量

一階

7

二階

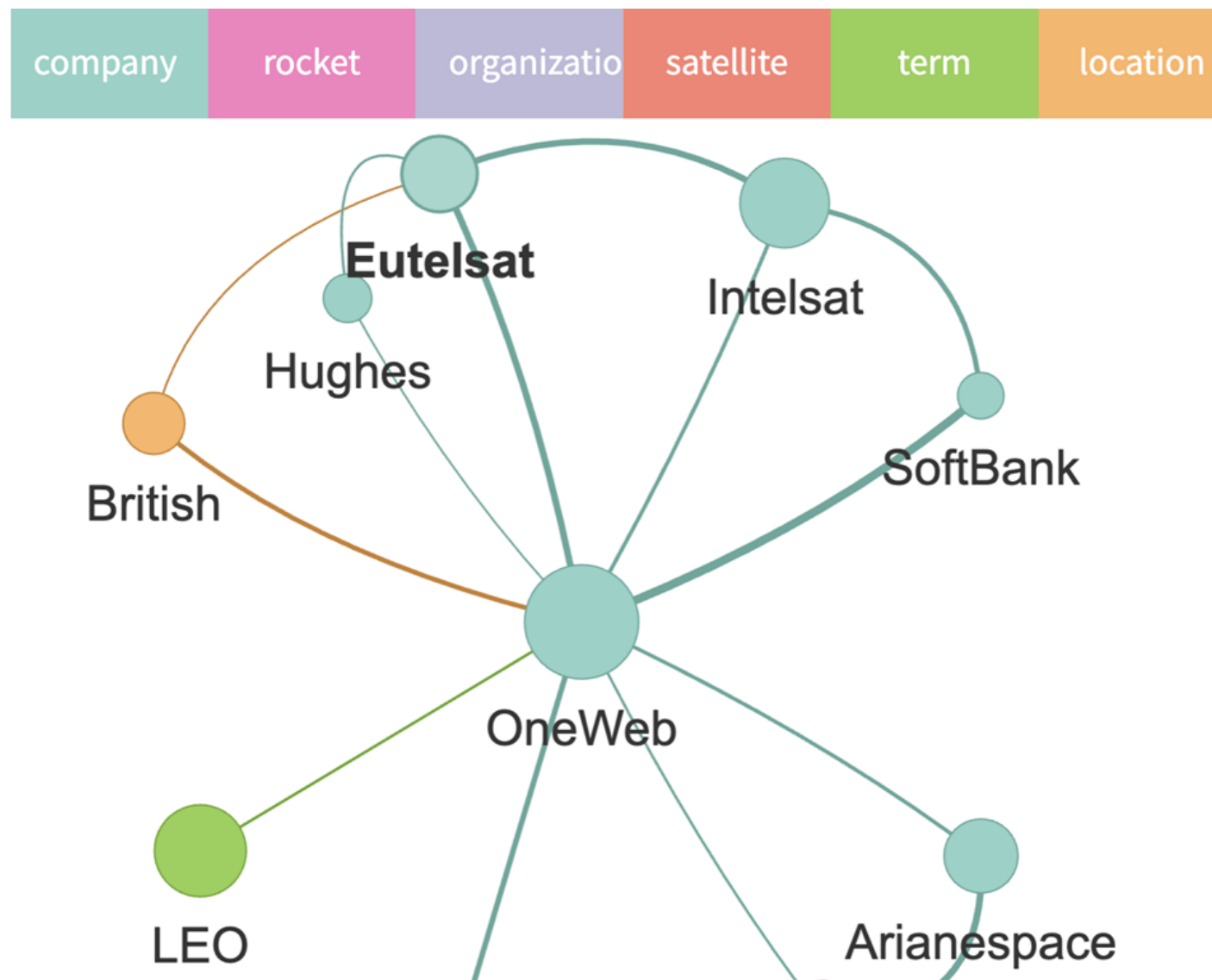
4

左方參數選單，可選擇一二階節點數量

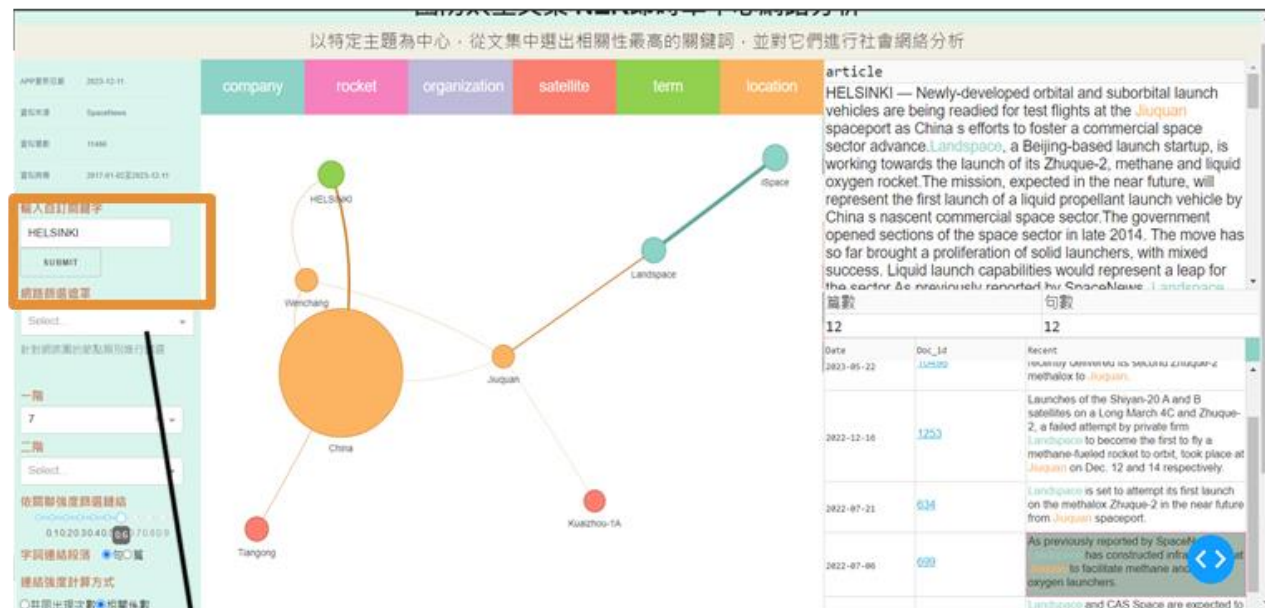


## 案例(OneWeb)

- SoftBank撤資
- British入主OneWeb
- 與Eutelsat合併，整合GEO與LEO的資源
- Hughes提供網關設備



# 即時單中心網路



輸入自訂關鍵字

SUBMIT

輸入有興趣的關鍵字，再按下SUBMIT

## 目標：

輸入有興趣的關鍵字，以此關鍵字向外延伸找到其他與之較常有聯繫的其他字詞網路。在不用擴充關鍵字字典的情況下，一樣可以深入瞭解詞彙和其他關鍵字的關聯。

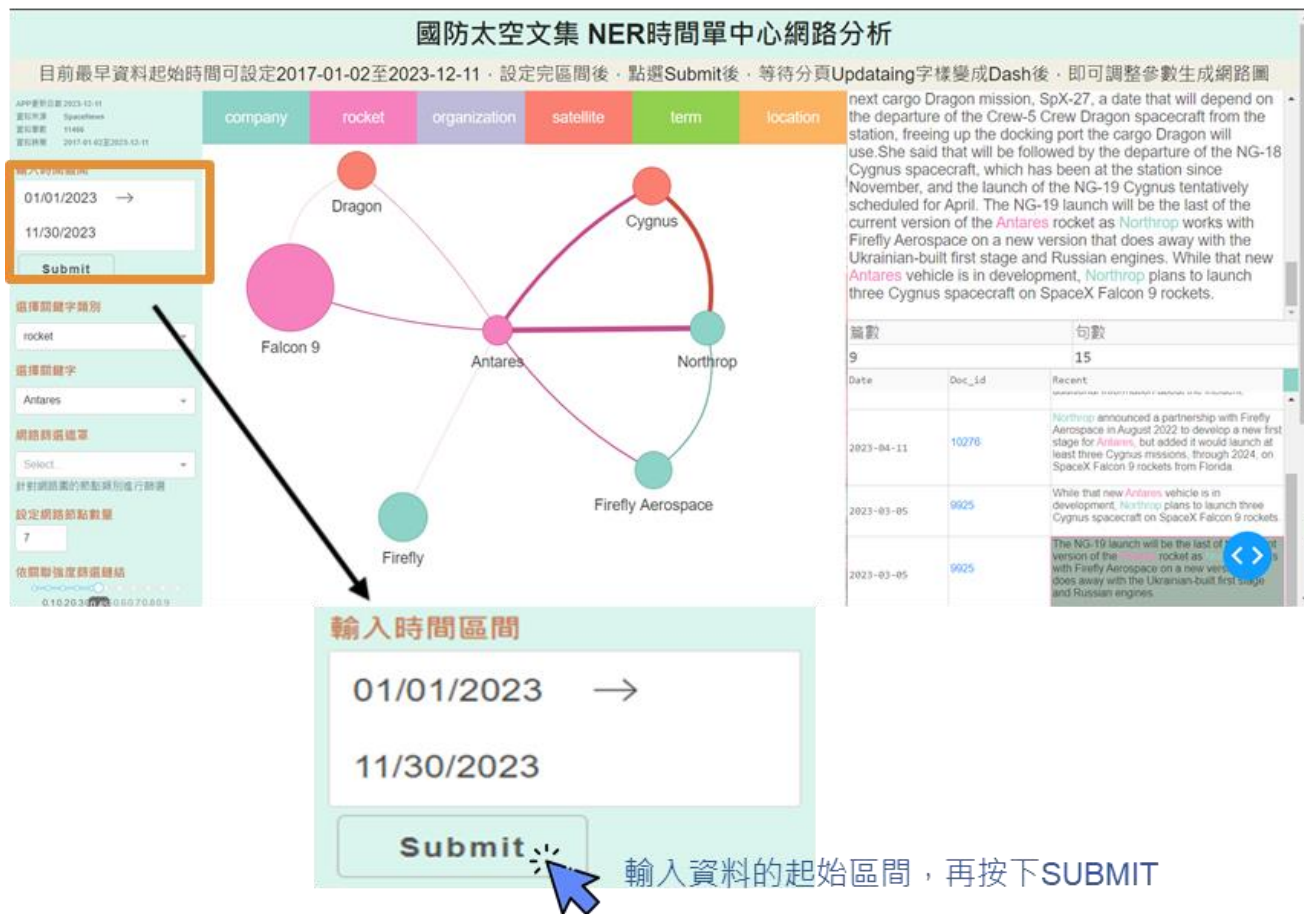
## 圖像意義：

輸入的關鍵字顯示為節點「keyword」，不同顏色顯示不同的字詞分類，以大小呈現它的出現頻率，而節點與節點之間的鏈結可以表達共同出現次數或是相關性，當數值越大，鏈結就會越厚實。

## 工具篩選功能：

篩選詞類、設定節點數量、依關聯強度篩選鏈結、字詞連結段落、連結強度計算方式、一二階功能整合、字體大小調整

# 時間單中心網路



目標：輸入資料的起始區間，了解不同時段的關鍵字之間的關聯性，或是選擇特定詞類的關聯。

## 圖像意義：

特定起始區間之不同顏色顯示不同的關鍵字分類，以節點大小呈現它的出現頻率，而節點與節點之間的鏈結可以表達共同出現次數或是相關性，當數值越大，鏈結就會越厚實。

## 工具篩選功能：

從關鍵字類別選擇關鍵字、篩選只顯示特定詞類、設定節點數量、依關聯強度篩選鏈結、字詞連結段落、連結強度計算方式字體大小調整