Ryerson University

TED ROGERS SCHOOL OF MANAGEMENT

SocialMediaLab.ca
"Making sense of a networked world"

Session 6:

1) Data Collection using CrowdTangle

(Facebook & Instagram URL-Search Data)

2) Two-mode Semantic Networks

3) Twitter Bot Detection with Botometer API

July 22, 2021

SocialMediaLab.ca

Ryerson University
TED ROGERS SCHOOL OF MANAGEMENT
SocialMediaLab.ca
"Making sense of a networked world"

The Social Media Lab is a multidisciplinary research laboratory at Ryerson University in Toronto, Canada

The lab studies how social media is changing the way people communicate, share information and form communities online, and how these changes impact society.

Videos and slides from past sessions are available online at: communalytic.com

communalytic

FAQ TUTORIALS PUBLICATIONS CSS BOOTCAMP

APR 07 2021

Social Media Lab's Computational Social Science (CSS) Bootcamp – Summer 2021

Off

By COMMUNALYTIC

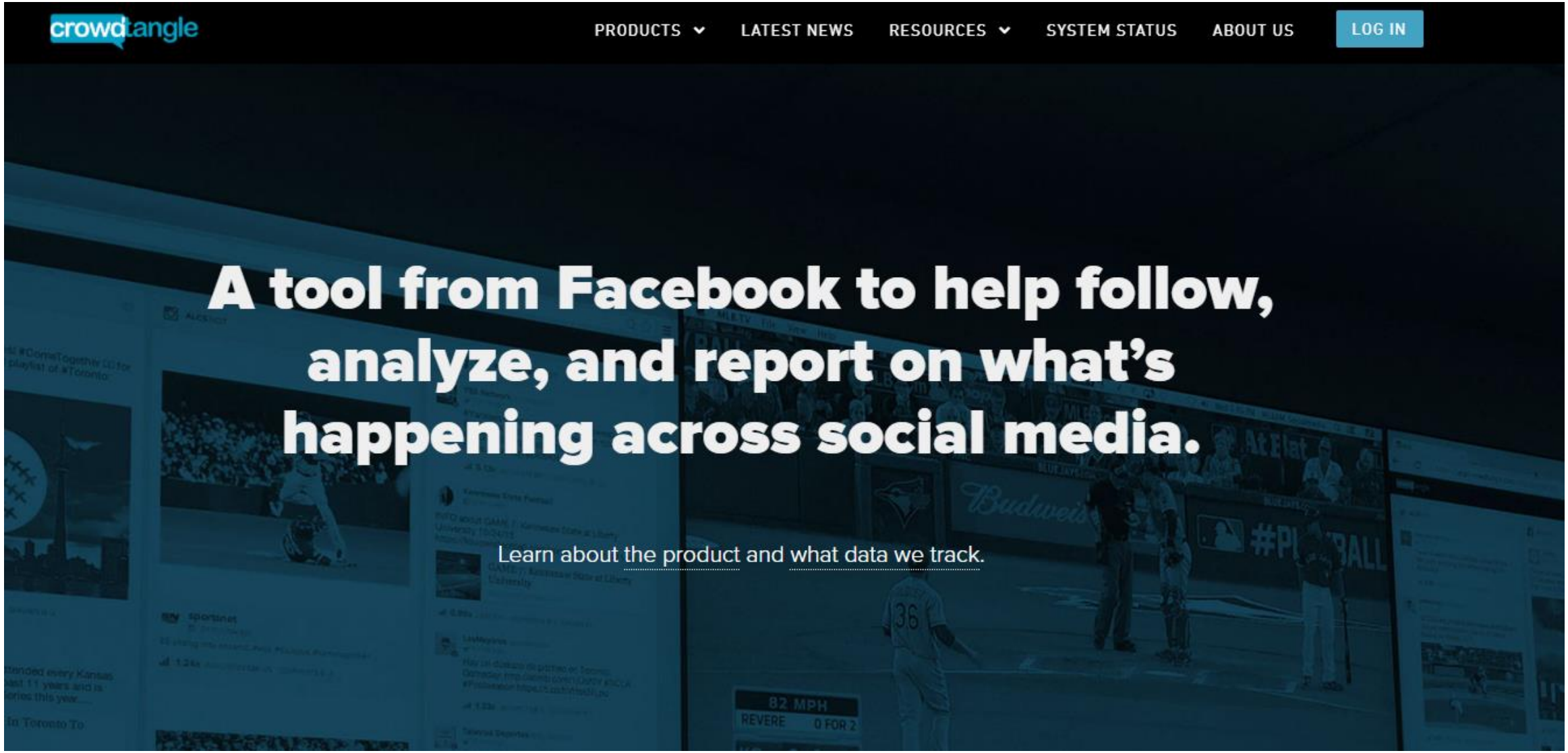
What: CSS Bootcamp on Examining Online Discourse & Networks with Communalytic

When: 2nd and 4th Thursday of the month at 10 am -11:30 am (ET) between May and July, 2021.

Where: Zoom (see details below)
Free Registration via Zoom

Outline

1. Data Collection using CrowdTangle (Facebook & Instagram URL-Search Data)
2. Two-mode Semantic Networks
3. Twitter Bot Detection with Botometer API *(Tiago Ribeiro)*



https://www.crowdtangle.com/

What Facebook/Instagram data does CrowdTangle cover?

- **Facebook: 7M+ Facebook pages, groups, and verified profiles.** This includes all public Facebook pages with more than 25K Page Likes or Followers (automated via API), all public Facebook groups with 95k+ members, all US-based public groups with 2K+ members, and all verified profiles.
- **Instagram: 2M+ public Instagram accounts.** This includes all public Instagram accounts with more than 50K followers, as well as all verified accounts.

https://help.crowdtangle.com/en/articles/1140930-what-data-is-crowdtangle-tracking

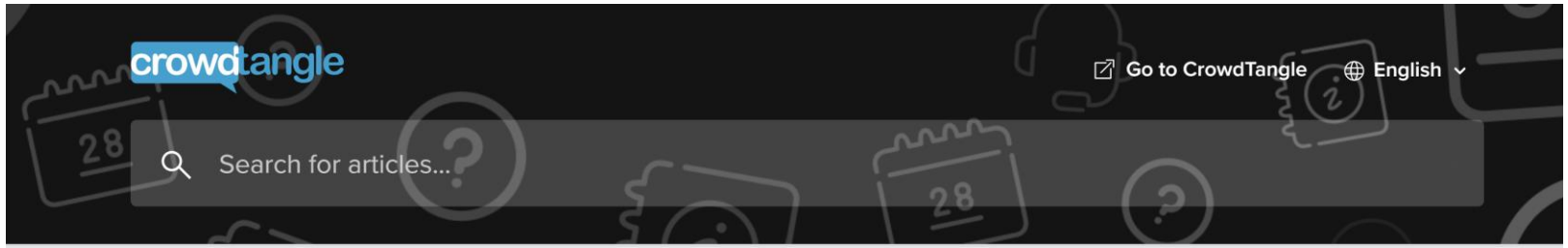
https://www.facebook.com/business/learn/lessons/verify-facebook-instagram-account

What public Facebook Pages does CrowdTangle cover?

The table below summarizes the percentage of **Facebook Pages active in the last 28 days** that CrowdTangle tracks, broken down by number of Page Likes or Followers **(last updated 6/08/2021)**. It does not include places (like this example), which do not contain any posts. The table will be updated every 30 days.

Number of Page Likes/Followers	Percentage of Pages CrowdTangle Tracks
>50,000	99.83%
>30,000	86.21%
>10,000	48.18%
>7,500	40.35%
>5,500	33.24%
>2,500	19.40%
>500	7.18%
ALL	1.72%

https://help.crowdtangle.com/en/articles/1140930-what-data-is-crowdtangle-tracking



All Collections > Partner Hubs > Academics and Researchers > CrowdTangle for Academics and Researchers

CrowdTangle for Academics and Researchers

Learn more and apply to join our program

Written by Naomi Shiffman
Updated over a week ago

CrowdTangle started a pilot program in 2019 to partner with researchers and academics to help them study the spread of public content on Facebook and

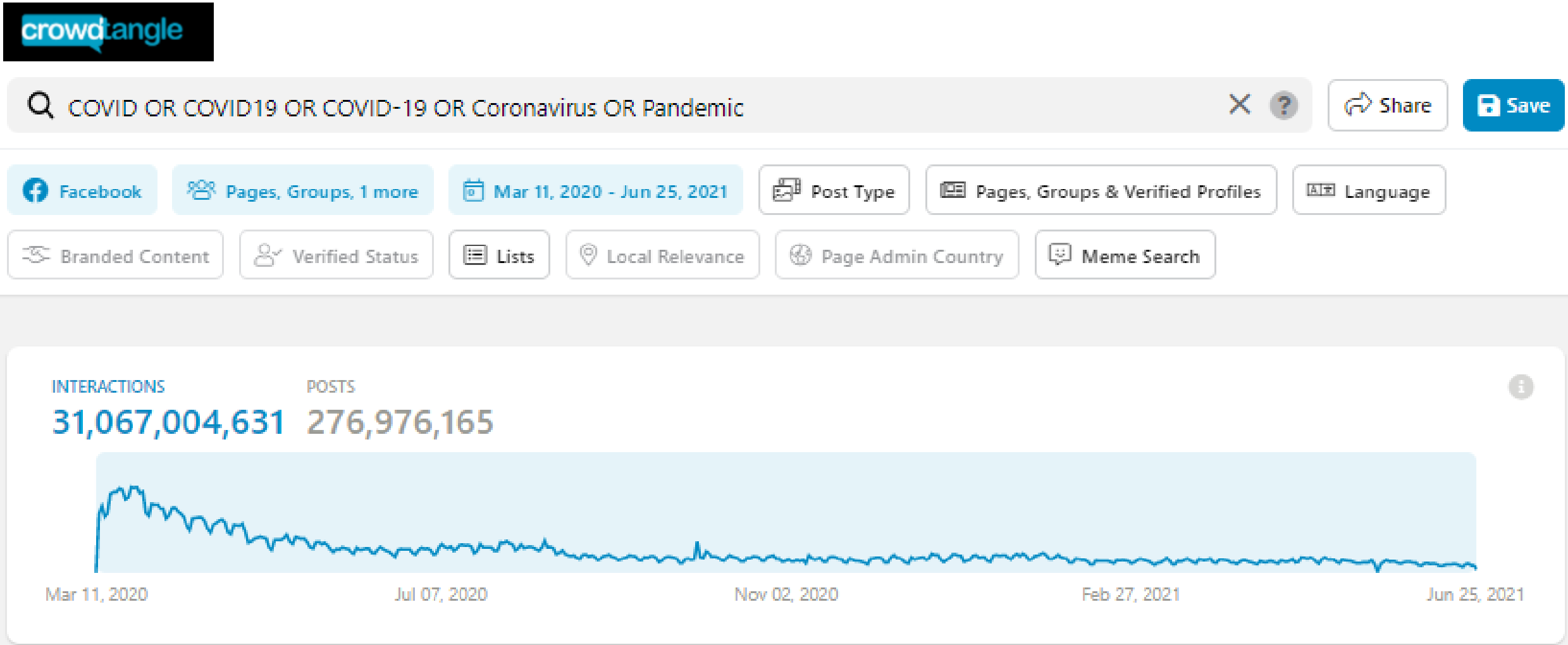
Apply to Access CrowdTangle

Apply to be part of CrowdTangle's Academics & Researchers program and get access to the CrowdTangle interface and API, as well as training and resources. Before applying, learn more about how Academics & Researchers use CrowdTangle here. We're currently prioritizing **university researchers (faculty, PhD students, post-doctoral research fellows)** focused on:

- Misinformation
- Elections
- COVID-19
- Racial justice
- Well-being

Apply Here: https://help.crowdtangle.com/en/articles/4302208-crowdtangle-for-academics-and-researchers

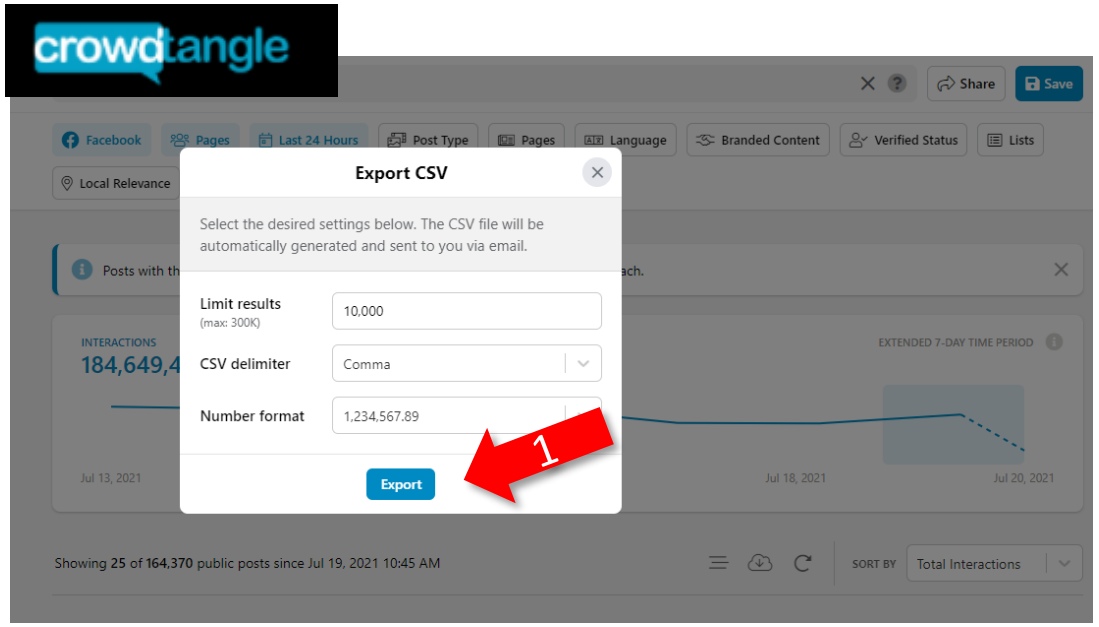
CrowdTange Search Example: COVID-19 mentions on Facebook (across public pages, groups and accounts)

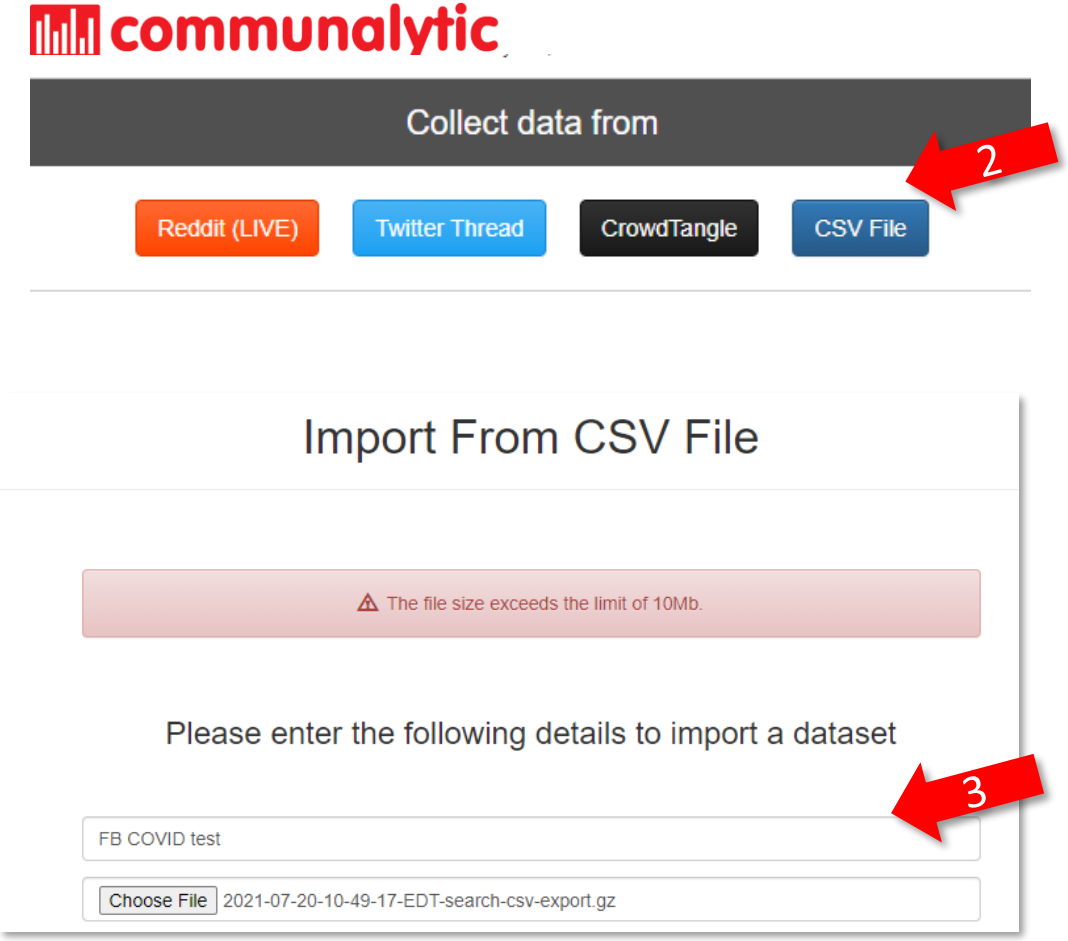


Source: [Crowdtangle]

Two options to import data from CrowdTangle into Communalytic

Option 1: Export -> Import as a CSV file

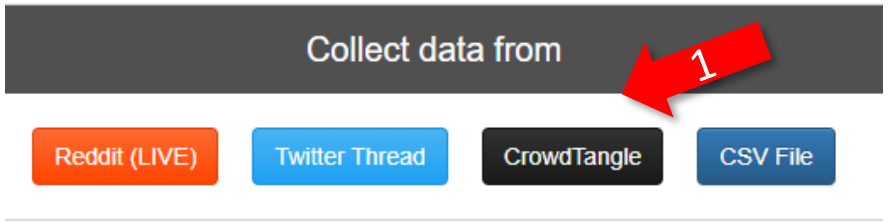




Two options to import data from CrowdTangle into Communalytic

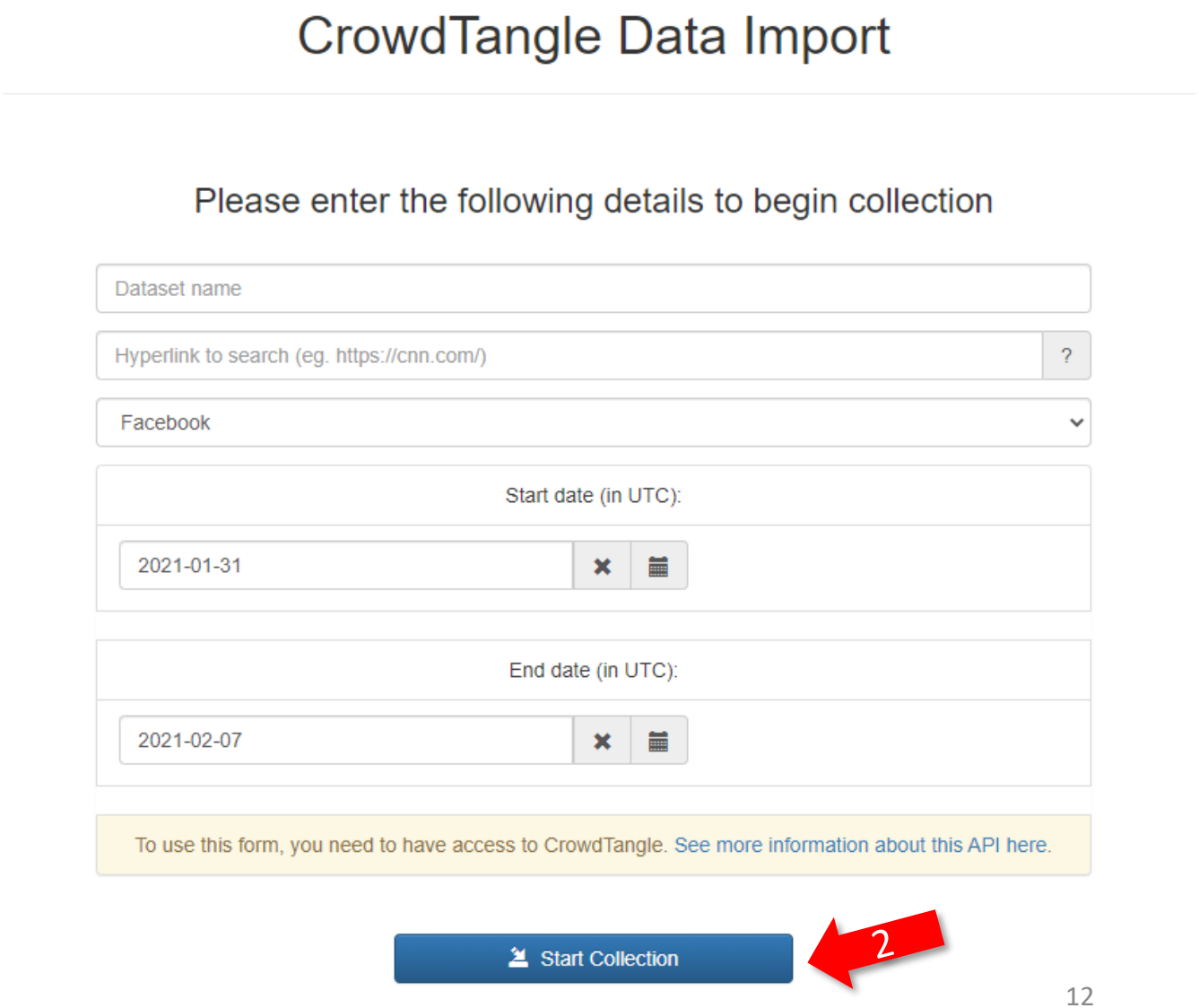
Option 2: Direct Import via CrowdTangle's URL-search API

communalytic



Via CrowdTangle, Communalytic can collect and analyze public Facebook and Instagram posts that **shared the same URL.**

Example: A URL to a New York Times story or the URL to any domain name



Sample URL-based Search with CrowdTangle

A news story about a possible side-effect of the Pfizer-BioNTech COVID-19 vaccine

COVID-19: Bell's palsy in four Pfizer volunteers not due to vaccine, says US FDA

There is no evidence supporting a causal relationship between the Bell's palsy cases and the injection of the vaccine

As the United Kingdom started inoculating people with the coronavirus vaccine developed by Pfizer-BioNTech, four volunteers who were administered the vaccine in the trial stage developed Bell's palsy, according to US Food and Drug Administration (FDA) regulators. Bell's palsy is a form of temporary facial paralysis.

According to healthfeedback.org, "four cases in the vaccine group do not represent a frequency above that expected in the general population."

https://zeenews.india.com/world/covid-19-four-pfizer-vaccine-volunteers-develop-bells-palsy-read-details-here-2329924.html

Facebook posts sharing and commenting on this news story

	A	B	C	D	E	F	G	H
1	postid	date	updated	message	platform	postUrl	postScore	author
2	12801832	########	########	And the FDA says	Facebook	https://wv	-3.375	Dr. Green
3	8115262|7	########	########	◆ Após terem sidc	Facebook	https://wv	-0.75	Portugues
4	12034948	########	########	Quatro voluntários	Facebook	https://wv	-40	Fechado c
5	11286838	########	########	# **"Ma Jimpurtha:	Facebook	https://wv	-26	Coronavir
6	12916638	########	########	https://zeenews.inc	Facebook	https://wv	-16	Titletown
7	12036379	########	########	Still want it?	Facebook	https://wv	-12	Pure Hem
8	10837622	########	########	COVID-19: Quatro v	Facebook	https://wv	-2.6	Passageirc
9	9843570|8	########	########	Quatro voluntários	Facebook	https://wv	-40	2 milhões
10	12893895	########	########	Quatro voluntários	Facebook	https://wv	-2.5	TAUBATE

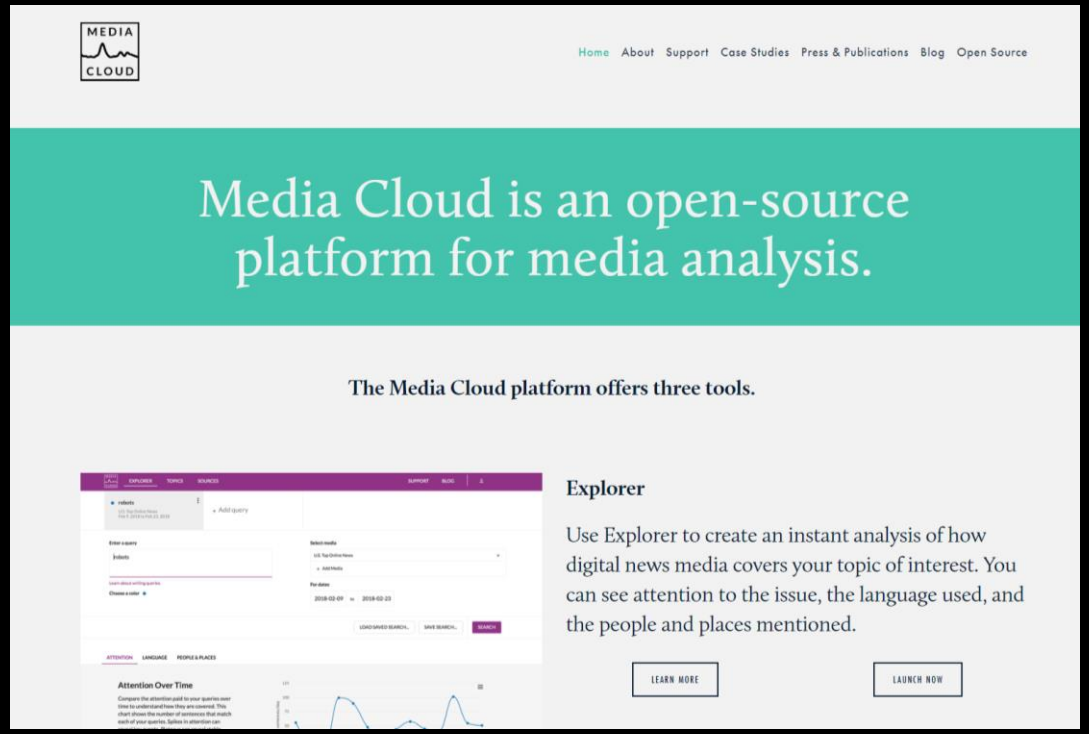
	I	J	K	L	M	N	O
	handle	accountId	accountPlatformId	subscriberCoun	verified	link	accountTyp
	DrGreenM	12801832	5.42303E+14	3839	FALSE	https://zeenews.india.	
	es no reinc	8115262	1.43779E+15	17460	FALSE	https://www.facebook	
	om Jesus e	12034948	2.92951E+15	3890	FALSE	https://www.facebook	
	s News Ma	11286838	2.63281E+15	943	FALSE	https://zeenews.india.	
	Cheesehea	12916638	1.09483E+14	1083	FALSE	https://www.facebook	
	PureHemp	12036379	2.05578E+14	145	FALSE	https://zeenews.india.	
	s do Trans	10837622	4.56768E+14	7221	FALSE	https://zeenews.india.	
	de mulhere	9843570	2.9583E+14	5669	FALSE	https://www.facebook	
	ALERTA	12893895	3.31977E+15	1746	FALSE	https://www.facebook	

Research Projects using CrowdTangle's URL-based Search

Sample Research Topics

- Identify signs of possible coordination among seemingly disparate accounts on platforms like Facebook and Instagram
- Identify who is sharing URLs to low credibility websites, known to spread misinformation on different topics

Ex. Database of Media Coverage



https://mediacloud.org/

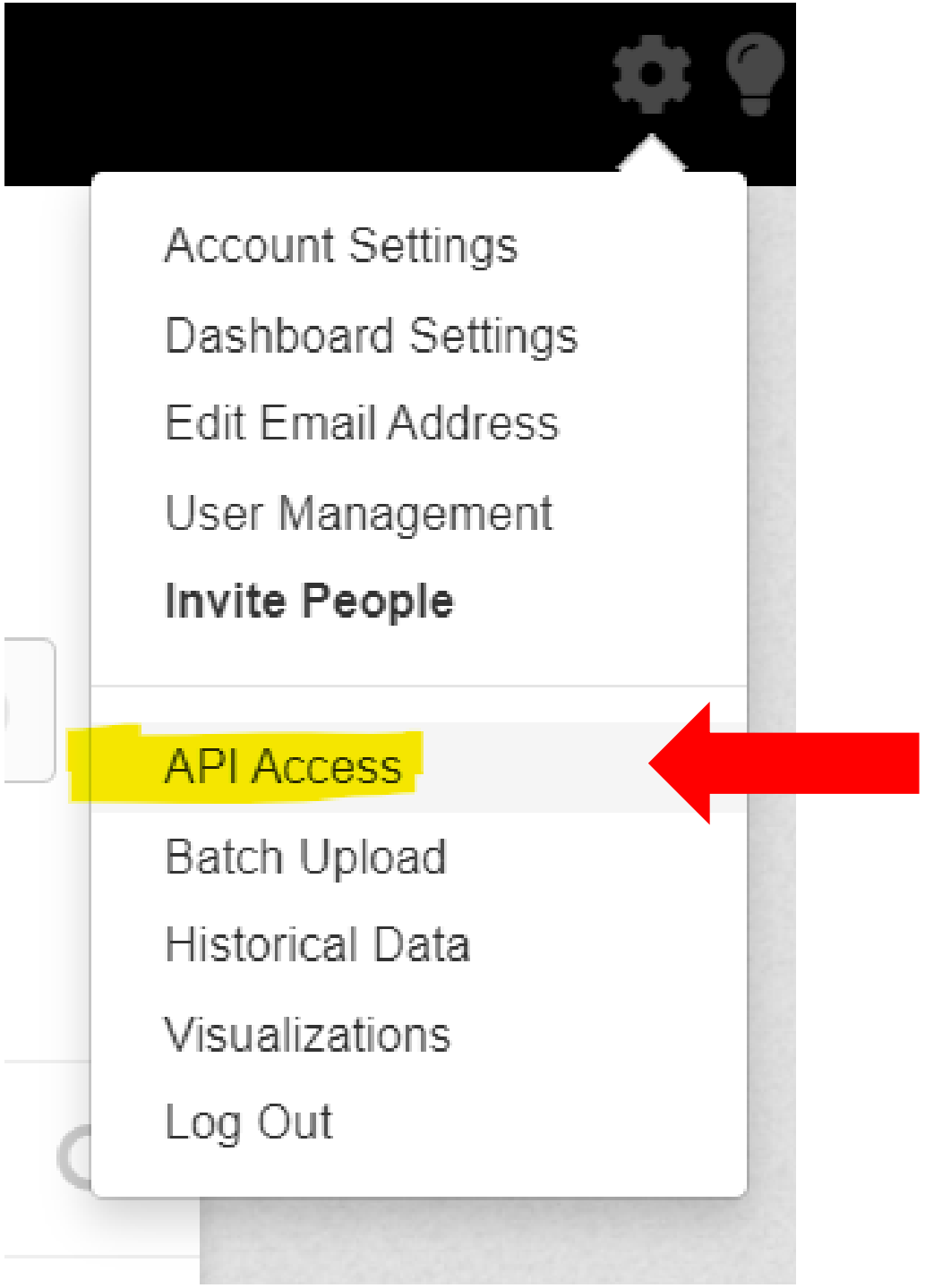
Research Projects using CrowdTangle's URL-based Search

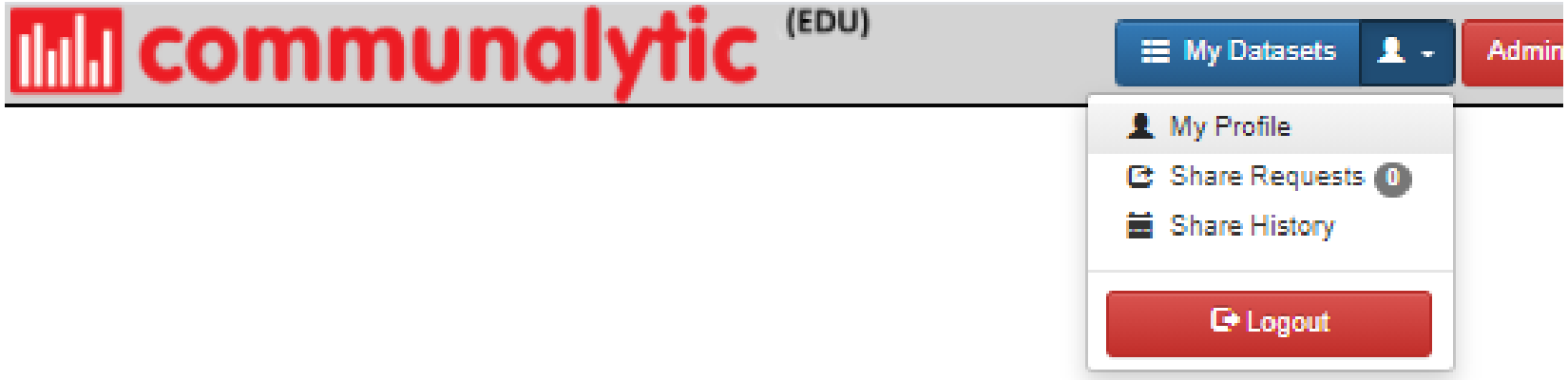
Sample Research Topics

- Identify signs of possible coordination among seemingly disparate accounts on platforms like Facebook and Instagram
- Identify who is sharing URLs to low credibility websites, known to spread misinformation on different topics

Ex. Database of Low Credibility Websites

https://www.newsguardtech.com/newsguard-for-researchers/


Account Settings
Dashboard Settings
Edit Email Address
User Management
Invite People
API Access
Batch Upload
Historical Data
Visualizations
Log Out



Personal Details

Email:	email@address.com
Datasets:	20/30

API Keys

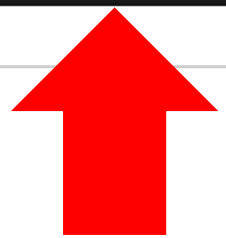
CrowdTangle API:	[key]	Remove Key
Twitter Bearer Token:	[key]	Remove Key
Perspective API:	[key]	Remove Key

Perspective is a machine learning API by Jigsaw and Google designed to conduct a 'toxicity' analysis of online comments. To use this API within Communalytic, please follow these steps to generate an API key. Once generated, enter your API key in the text field above.

communalytic

Collect data from

Reddit (LIVE) | Twitter Thread | CrowdTangle | CSV File



Search dataset list...

communalytic

CrowdTangle Data Import

A news story about a possible side-effect of the Pfizer-BioNTech COVID-19 vaccine

COVID-19: Bell's palsy in four Pfizer volunteers not due to vaccine, says US FDA

There is no evidence supporting a causal relationship between the Bell's palsy cases and the injection of the vaccine

As the United Kingdom started inoculating people with the coronavirus vaccine developed by Pfizer-BioNTech, four volunteers who were administered the vaccine in the trial stage developed Bell's palsy, according to US Food and Drug Administration (FDA) regulators. Bell's palsy is a form of temporary facial paralysis.

According to healthfeedback.org, "four cases in the vaccine group do not represent a frequency above that expected in the general population."

Facebook posts sharing and commenting on this news story

Please enter the following details to begin collection

Dataset name

https://zeenews.india.com/world/covid-19-four-pfizer-vaccine-volunteers-develop-bells-palsy-read-details-here-2329924.html ?

Facebook

Start date (in UTC):

2021-01-31

End date (in UTC):

2021-02-07

To use this form, you need to have access to CrowdTangle. See more information about this API here.

Start Collection

communalytic

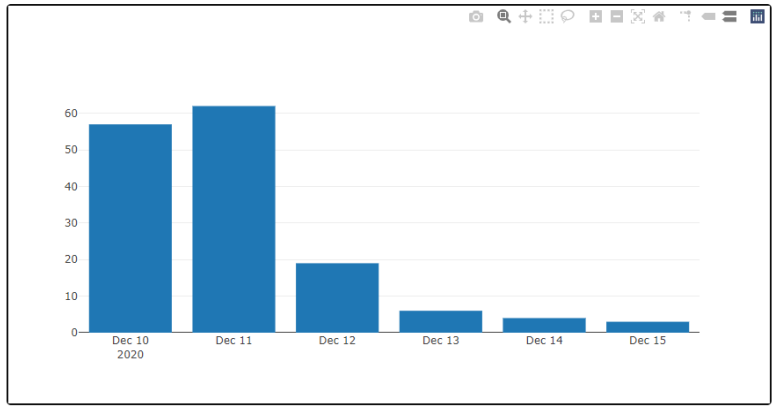
My Datasets | Shared With Me

Total records in the account: 15.1 million / 10M

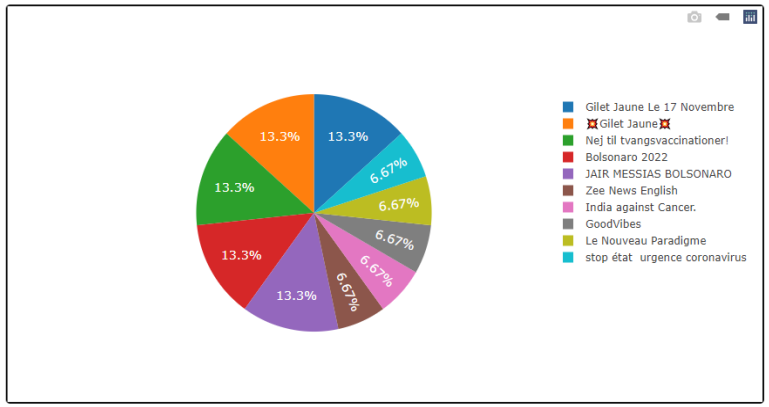
100%

Dataset Name	Number of Records	Toxicity Analysis	Network Analysis	Download Dataset	Collaborators	Delete
COVID-19: Bell's palsy in four Pfizer volunteers not due to vaccine, says US FDA facebook : https://zeenews.india.com/world/covid-19-four-pfizer-vaccine-volunteers-develop-bells-palsy-read-details-here-2329924.html	151					

Posts Per Day

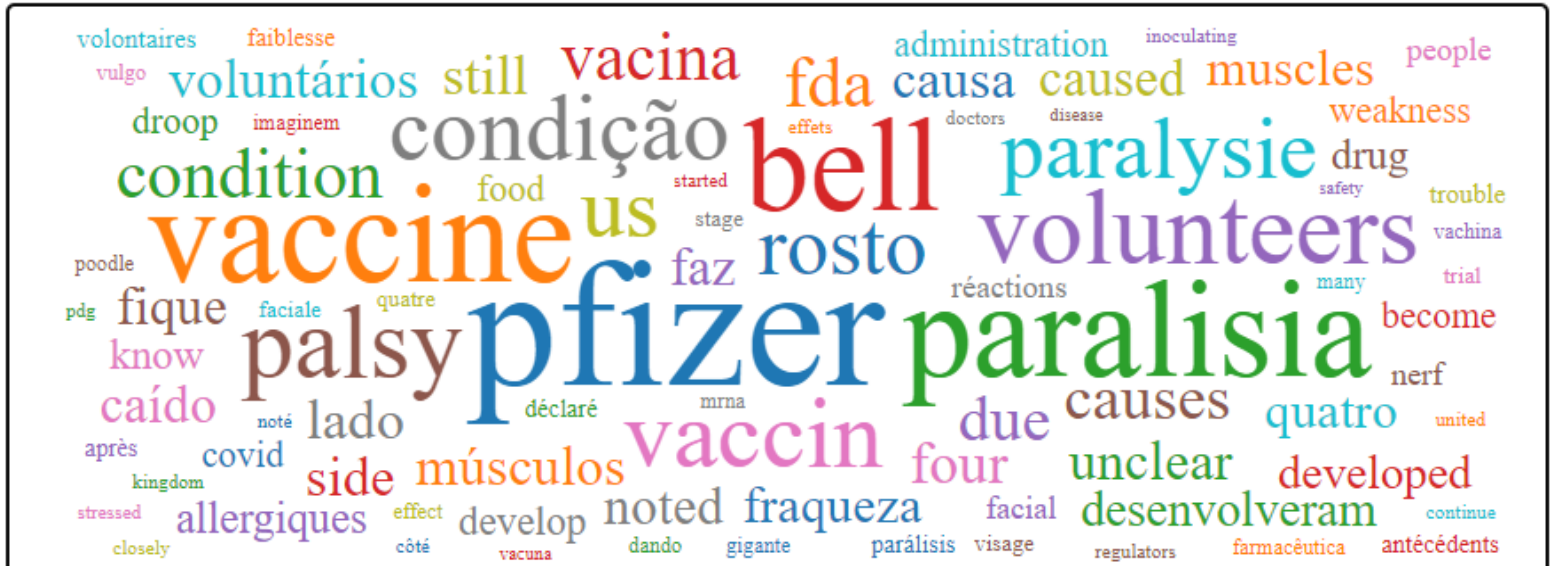


Top Ten Posters



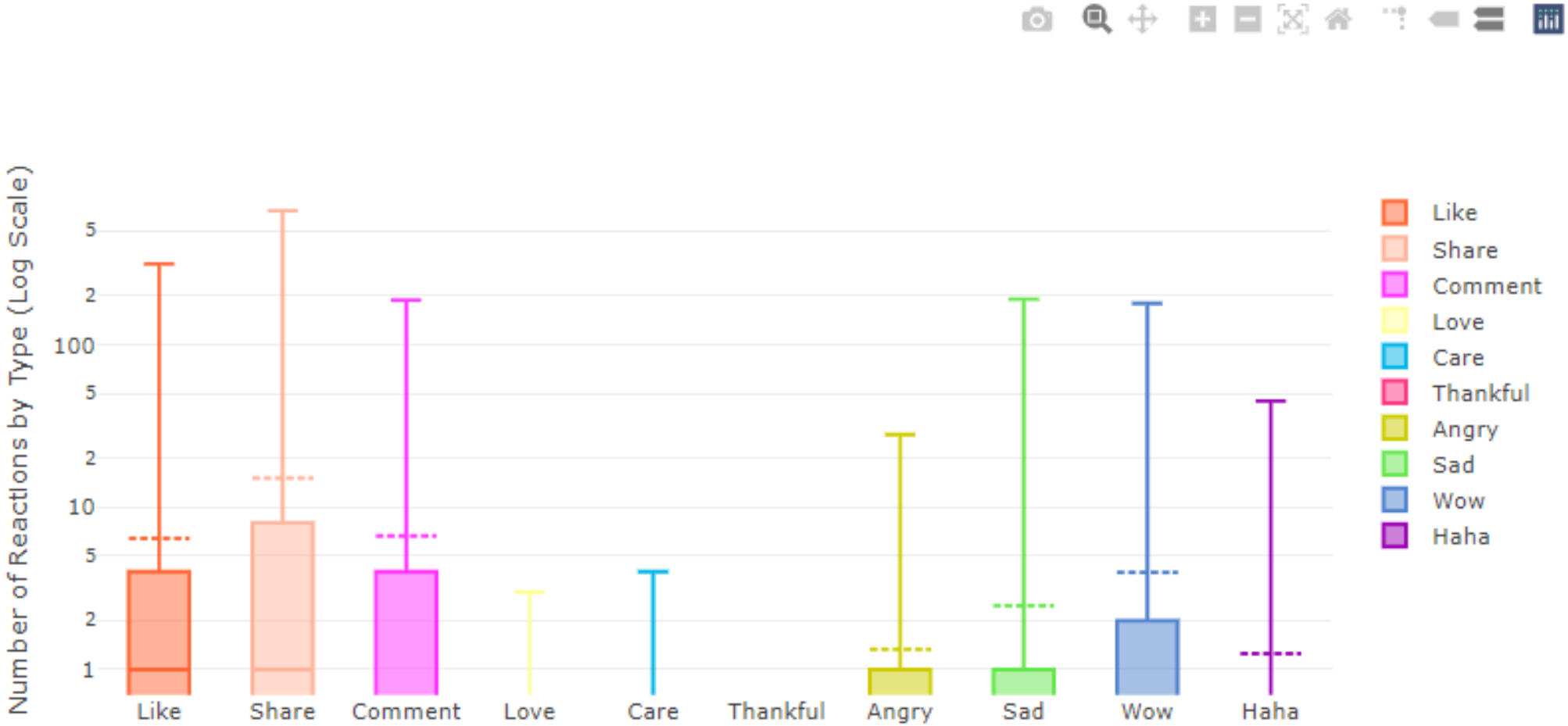
Word Cloud

CSV EXCEL PNG



This chart shows up to 100 most frequently used words in your dataset , excluding numbers, URLs, and stop words in 15 different languages.

Reaction Count Boxplots



Toxicity Analysis

There were **151** comments analyzed in **English**

Print | Reset Analysis

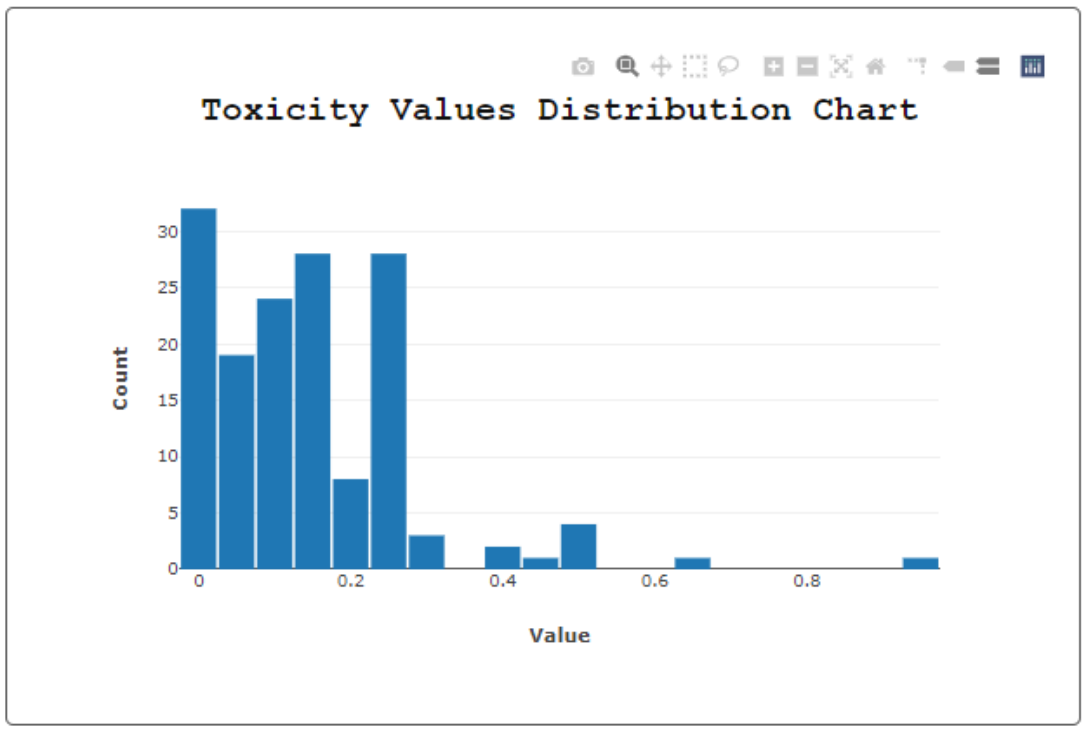
⚠ The Toxicity analysis can only run with one language mode at a time. To change the primary language selection for this dataset, click on the Reset Analysis button and rerun the analysis as needed.

Click on the highest/lowest values to see the top posts for each category

Download the Toxicity analysis results as a CSV file.

	Average for dataset	Highest value	Lowest value
Toxicity	0.15	0.96	0.01
Severe toxicity	0.10	0.89	0.00
Identity attack	0.13	0.80	0.00
Insult	0.14	0.79	0.01
Profanity	0.12	0.97	0.00
Threat	0.15	0.49	0.01
Sexually Explicit	0.10	0.85	0.01
Flirtation	0.36	0.67	0.18

TOXICITY VALUES DISTRIBUTION CHART



Outline

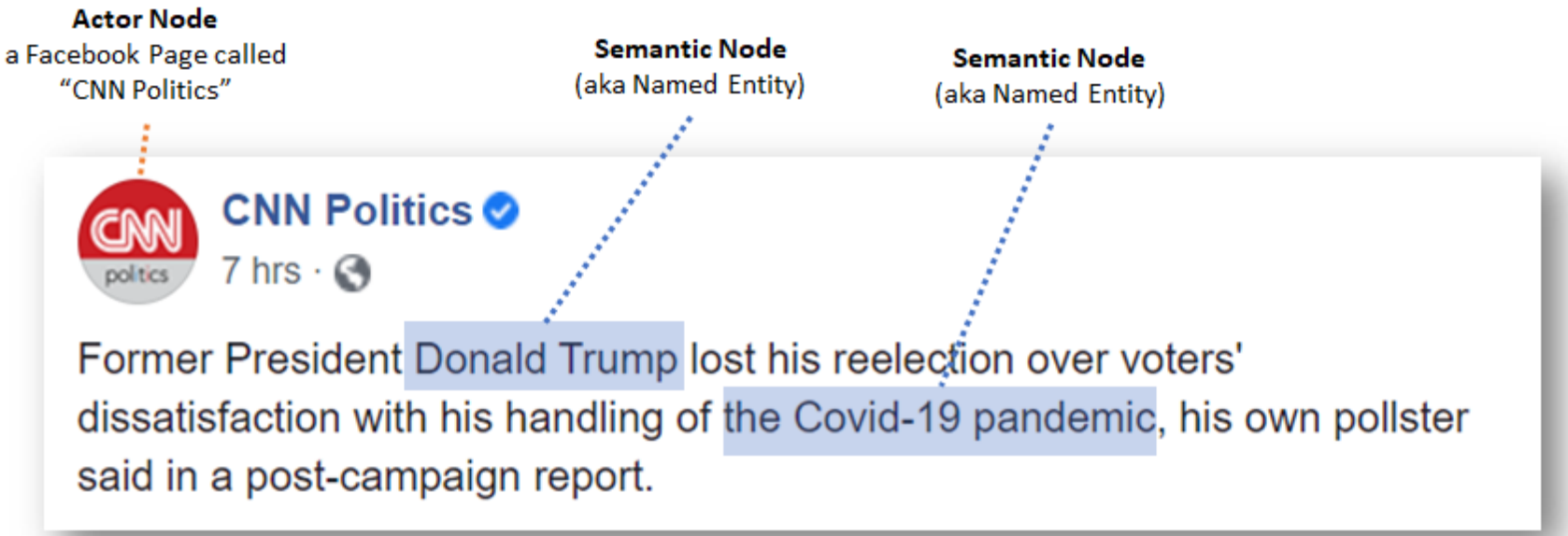
1. Data Collection using CrowdTangle (Facebook & Instagram)
2. **Two-mode Semantic Networks**
3. Twitter Bot Detection

Example of a Two-mode Semantic Network based on a single Facebook post

Step 1: Node Discovery

From the metadata collected, Communalytic will recognize "CNN Politics"- a FB page - as an **Actor Node** and will designated it as such in the network.

Communlaytic then uses natural language processing (NLP) to examine the content of the post and locate any **Semantic Nodes** (Named Entities). In this example, it found two **Semantic Nodes**: "Donald Trump" and "Covid-19 pandemic".

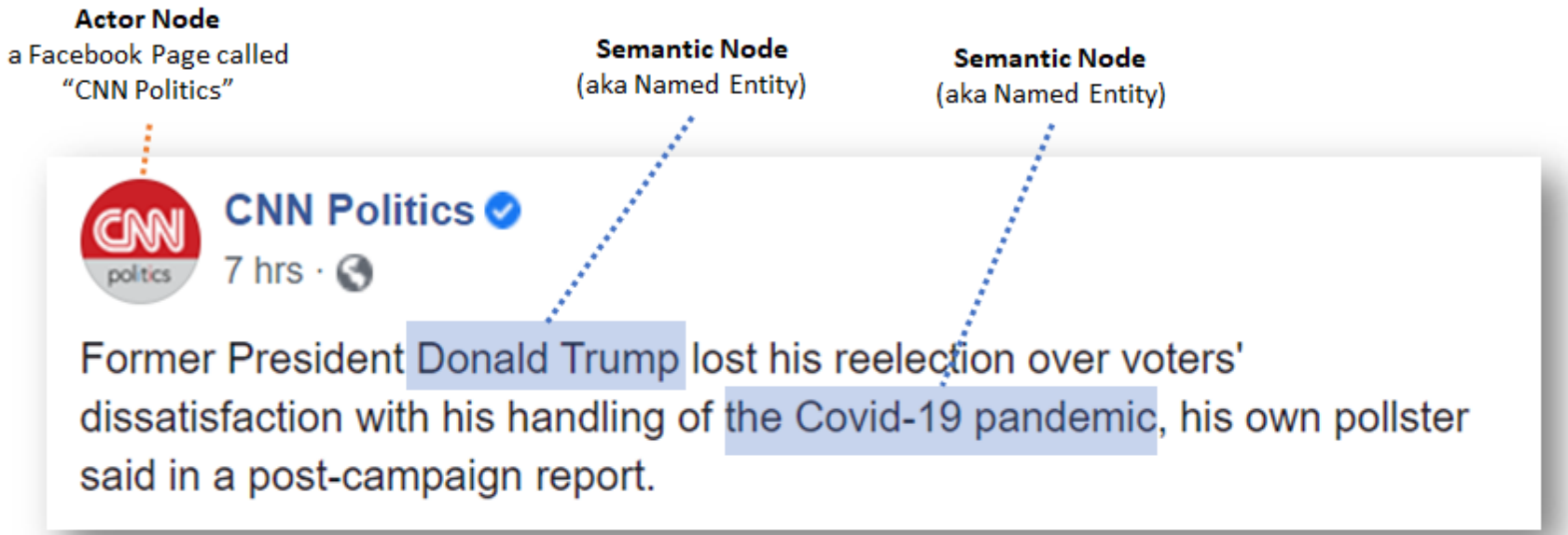


Example of a Two-mode Semantic Network based on a single Facebook post

Step 1: Node Discovery

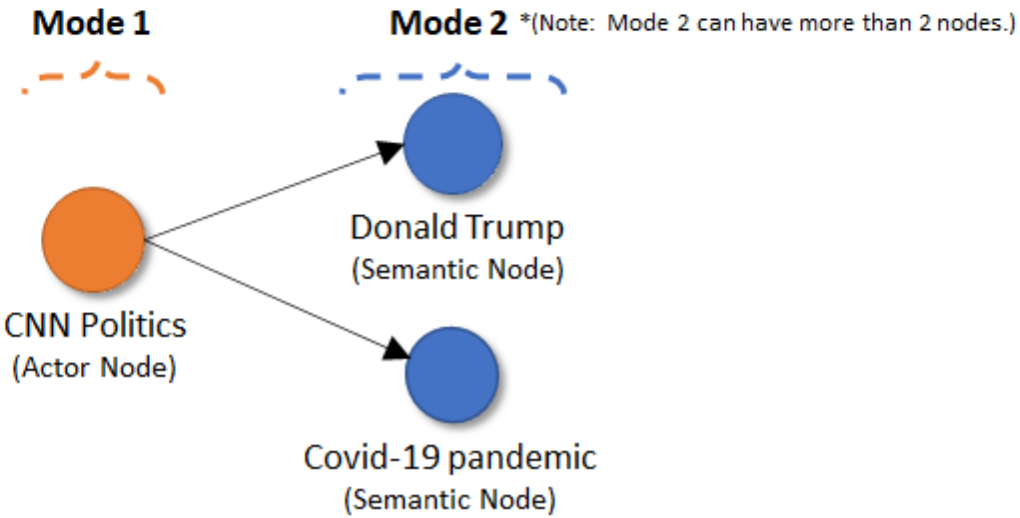
From the metadata collected, Communalytic will recognize "CNN Politics"- a FB page - as an **Actor Node** and will designated it as such in the network.

Communlaytic then uses natural language processing (NLP) to examine the content of the post and locate any **Semantic Nodes** (Named Entities). In this example, it found two **Semantic Nodes**: "Donald Trump" and "Covid-19 pandemic".



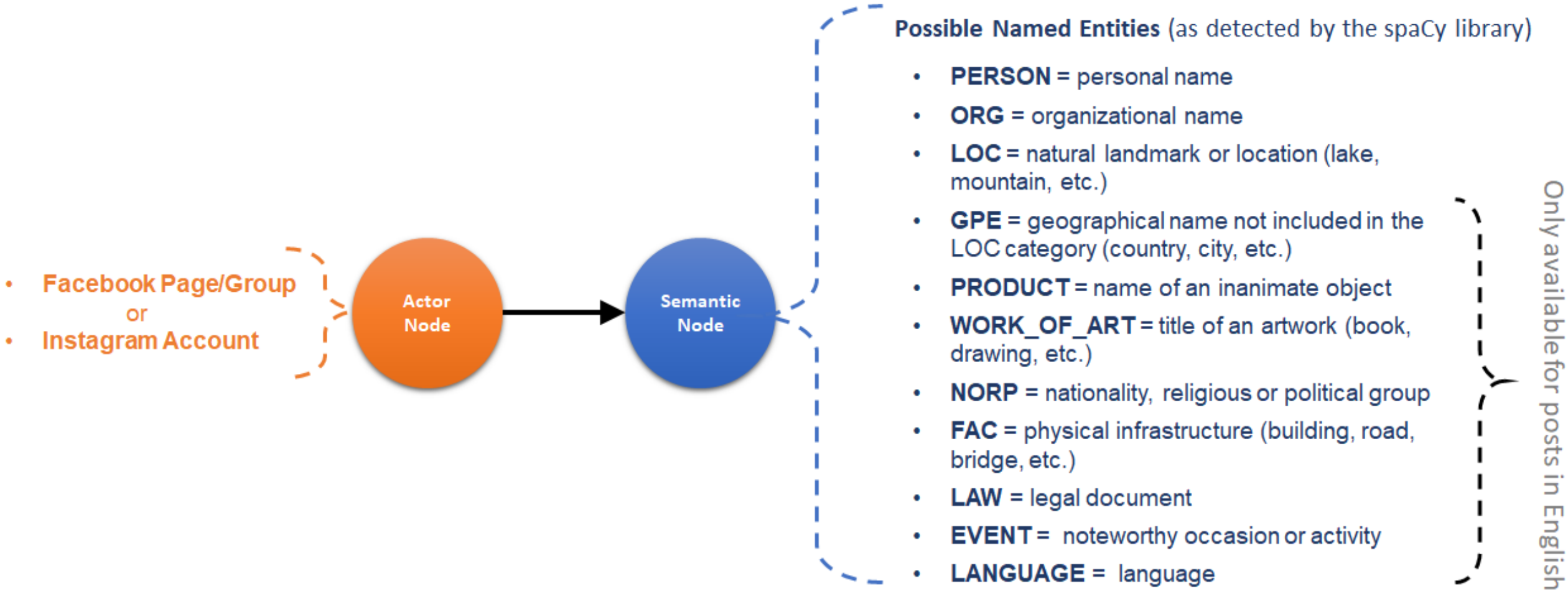
Step 2: Edge Discovery

Next, to create a **two-mode network**, Communalytic automatically connects the **Actor Node** ("CNN Politics") to the two **Semantic Nodes (**"Donald Trump" and "Covid-19 pandemic").



How are ‘Named Entities’ identified?

Communalytic relies on **spaCy,** a natural language processing library, to detect different types of named entities.



Accuracy Measures for Named Entities Detection using spaCy:

Language models	Precision [0.00-1.00]	Recall [0.00-1.00]
English model	0.86	0.85
Japanese model	0.74	0.70
Chinese model	0.74	0.69
Multi-language model	0.84	0.83

https://spacy.io/usage/linguistic-features#named-entities

A news story about a possible side-effect of the Pfizer-BioNTech COVID-19 vaccine

CORONAVIRUS

COVID-19: Bell's palsy in four Pfizer volunteers not due to vaccine, says US FDA

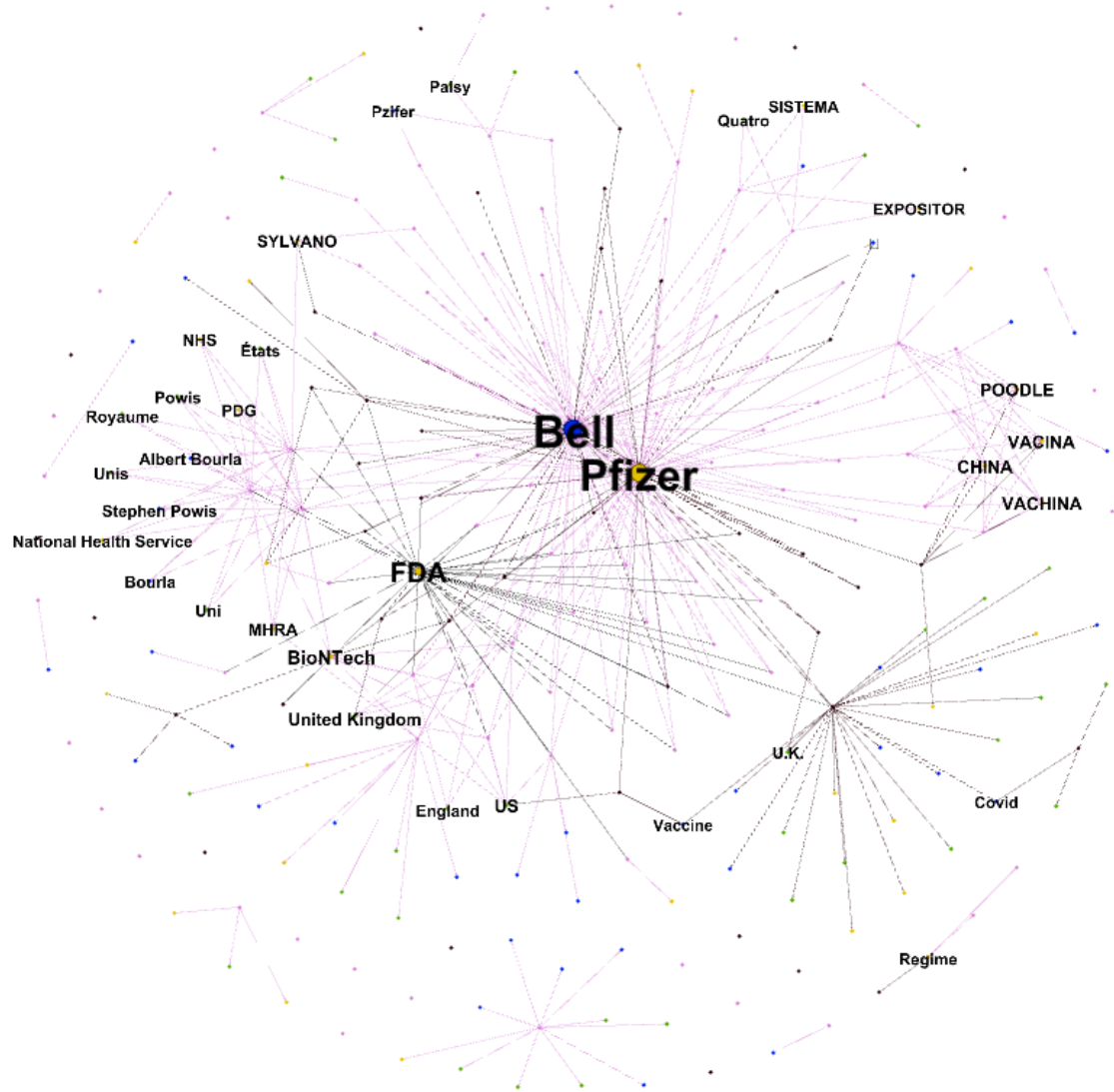
There is no evidence supporting a causal relationship between the Bell's palsy cases and the injection of the vaccine

As the United Kingdom started inoculating people with the coronavirus vaccine developed by Pfizer-BioNTech, four volunteers who were administered the vaccine in the trial stage developed Bell's palsy, according to US Food and Drug Administration (FDA) regulators. Bell's palsy is a form of temporary facial paralysis.

According to healthfeedback.org, "four cases in the vaccine group do not represent a frequency above that expected in the general population."

Facebook posts sharing and commenting on this news story

Two-mode Semantic Network connecting Actor Nodes to Semantic Nodes



communalytic

My Datasets | Shared With Me

Total records in the account: 15.1 million / 10M

100%

Dataset Name	Number of Records	Toxicity Analysis	Network Analysis	Download Dataset	Collaborators	Delete
COVID-19: Bell's palsy in four Pfizer volunteers not due to vaccine, says US FDA facebook : https://zeenews.india.com/world/covid-19-four-pfizer-vaccine-volunteers-develop-bells-palsy-read-details-here-2329924.html	151					

Network Analysis

This module automatically creates a two-mode semantic network from your CrowdTangle (FB or IG) dataset. If you are new to Social network Analysis (SNA), check out **this tutorial** to learn more about how two-mode semantic networks are built and how to analyze them.

Start by selecting the desired network type, then click on the "Generate Network":

- ○ **Reply-To Network (not available for this dataset type)**
- ◉ **Two-Mode Semantic Network (learn more)**

Generate Network

communalytic

Network Analysis

This step automatically creates a 'who replies to whom' communication network from your Twitter dataset. If you are new to Social Network Analysis (SNA), check out **this SNA primer** to learn the fundamentals of Network Analysis.

28%

Check progress in 29 sec

Network generation in progress! The completion time depends on the network size and may take up to few hours for large networks. Feel free to close the window and come back later.

communalytic

Network Analysis

This module automatically creates a two-mode semantic network from your CrowdTangle (FB or IG) dataset. If you are new to Social network Analysis (SNA), check out **this tutorial** to learn more about how two-mode semantic networks are built and how to analyze them.

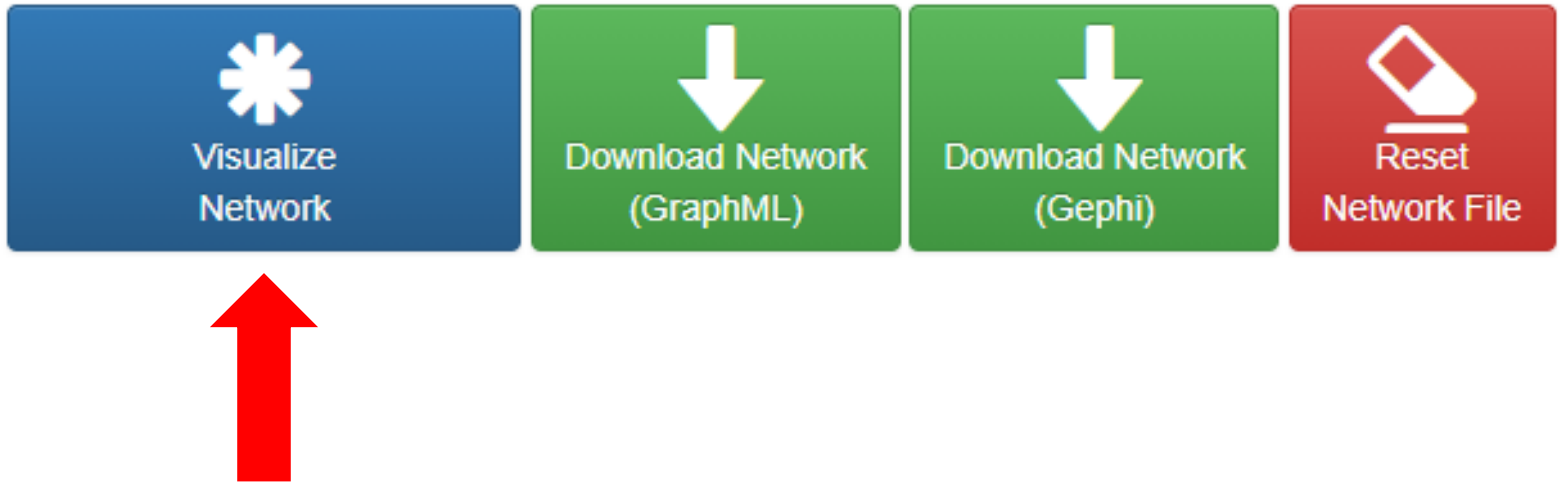
communalytic

Two-Mode Semantic Network Visualizer in Communalytic

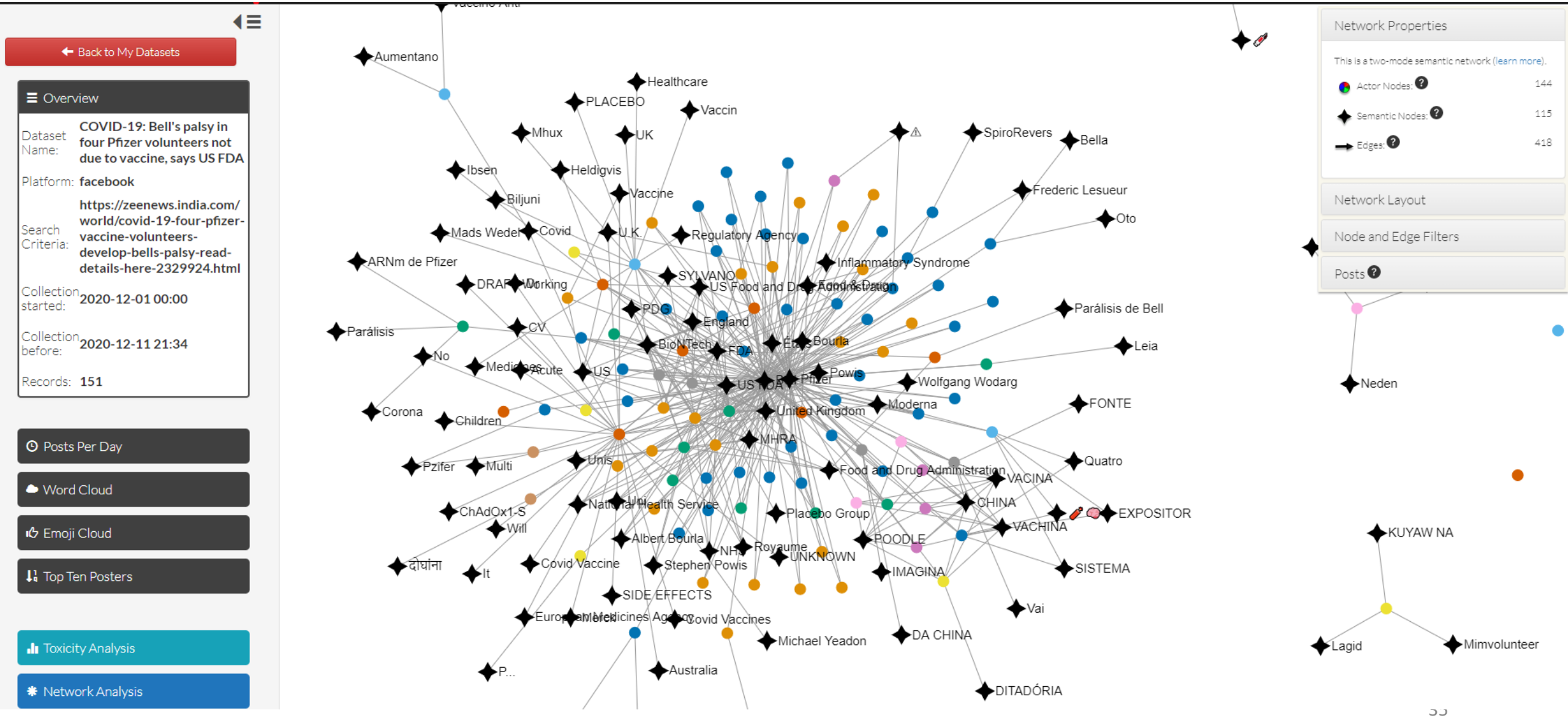
communalytic

Network Analysis

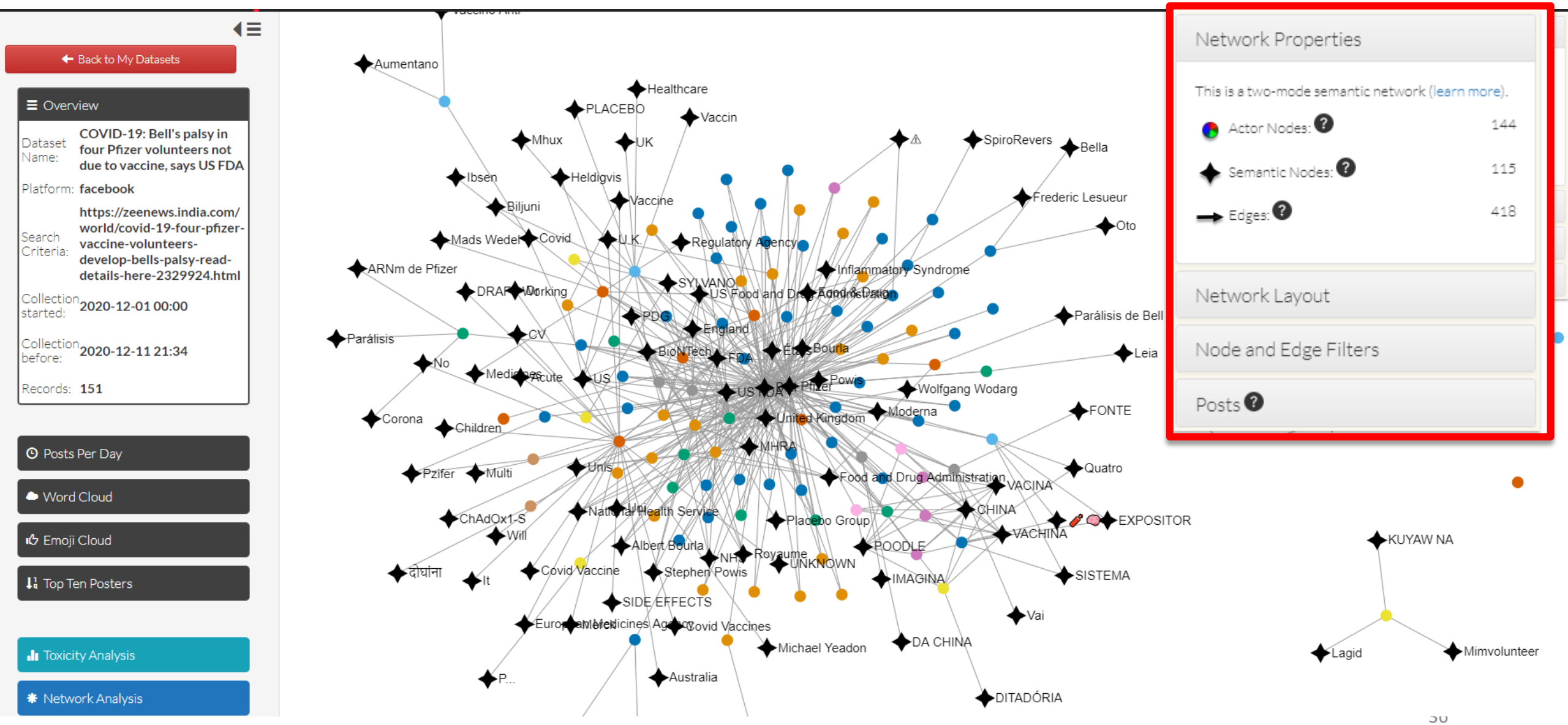
This module automatically creates a two-mode semantic network from your CrowdTangle (FB or IG) dataset. If you are new to Social network Analysis (SNA), check out **this tutorial** to learn more about how two-mode semantic networks are built and how to analyze them.



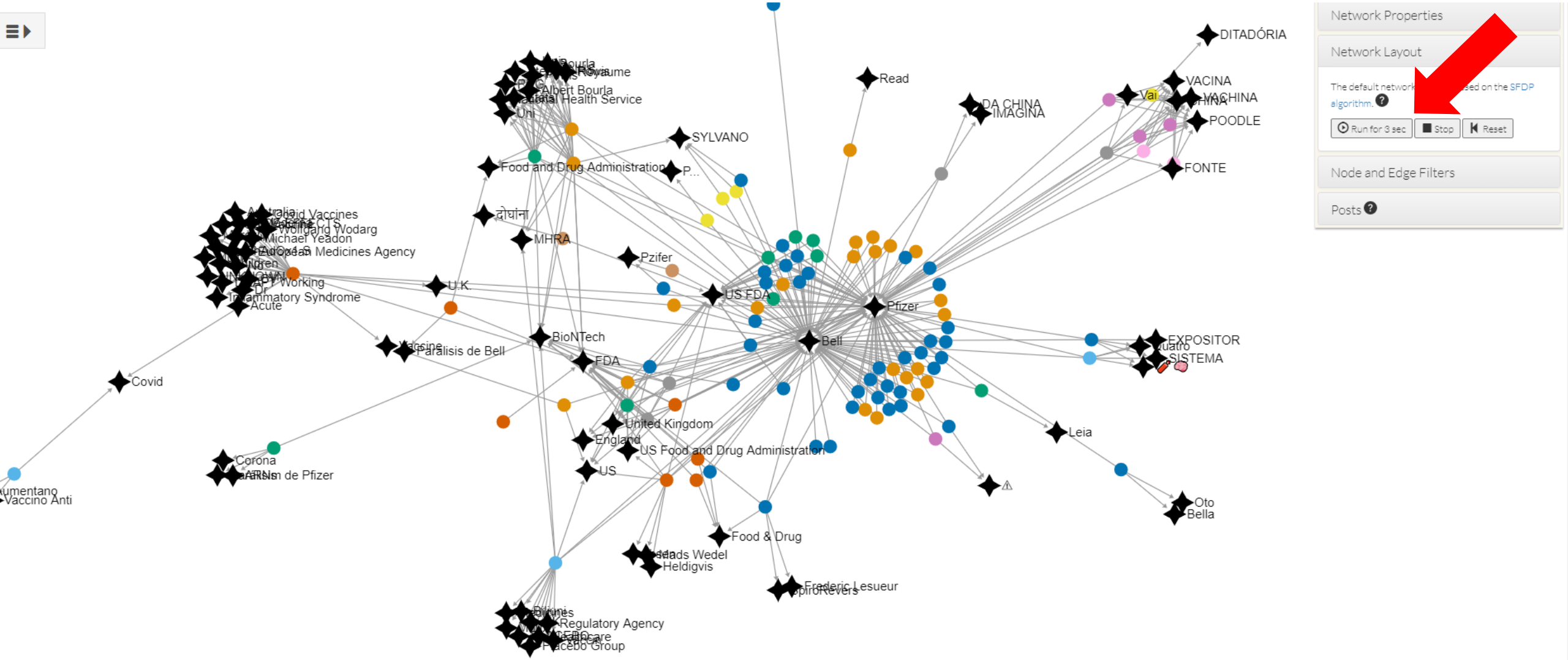
communalytic



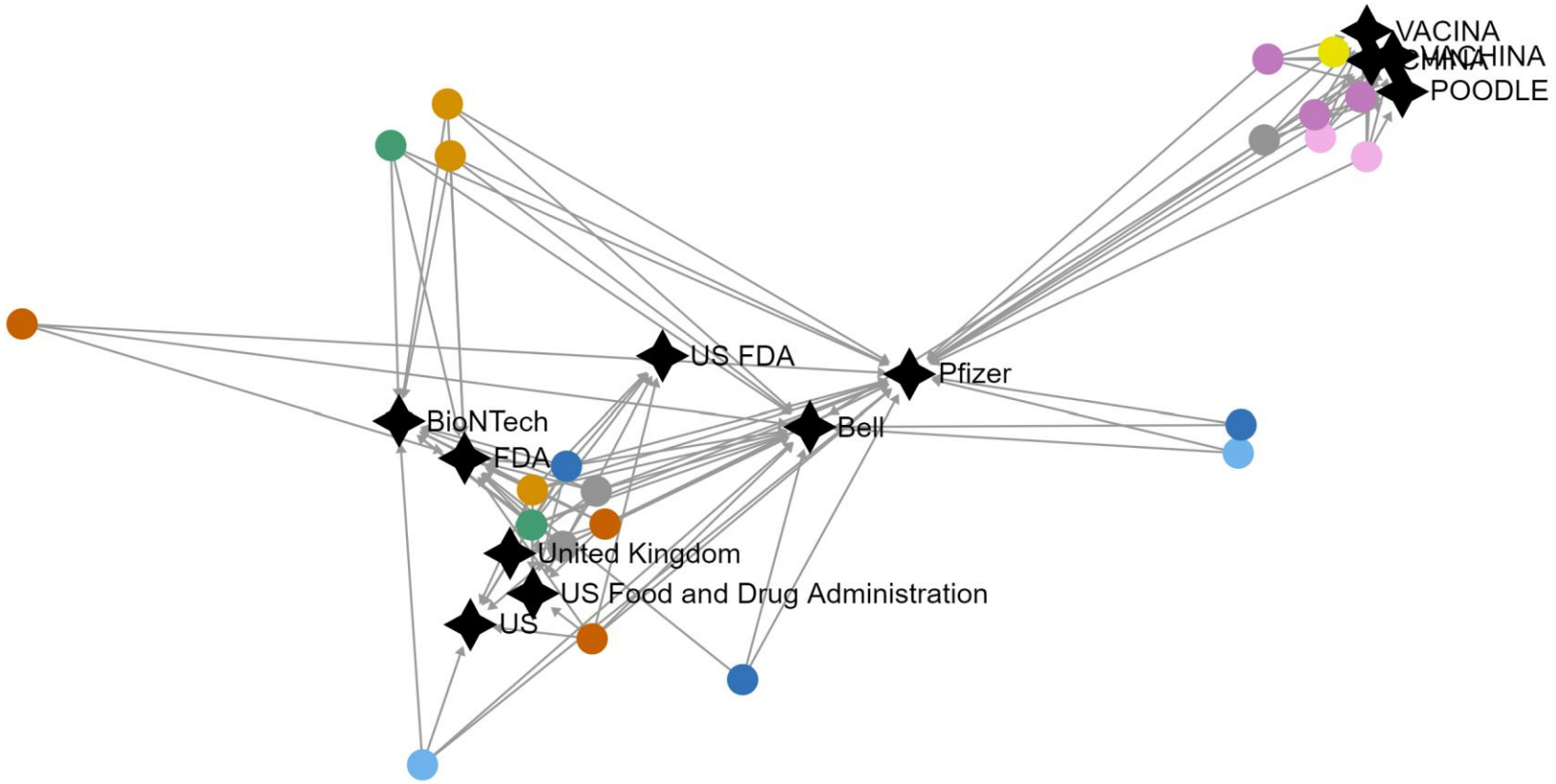
communalytic



communalytic



communalytic



Network Properties

Network Layout

Node and Edge Filters (Applied)

Community Clusters (>=10 nod

All nodes

Min Degree: 5

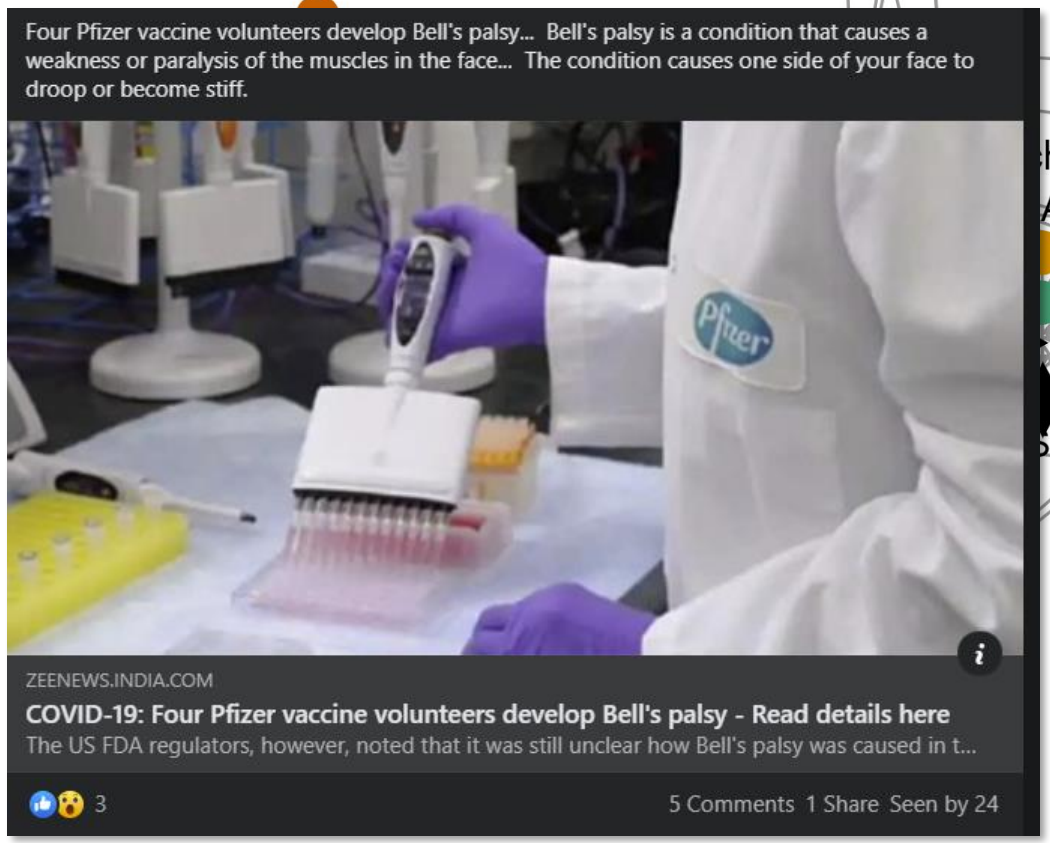
0 86

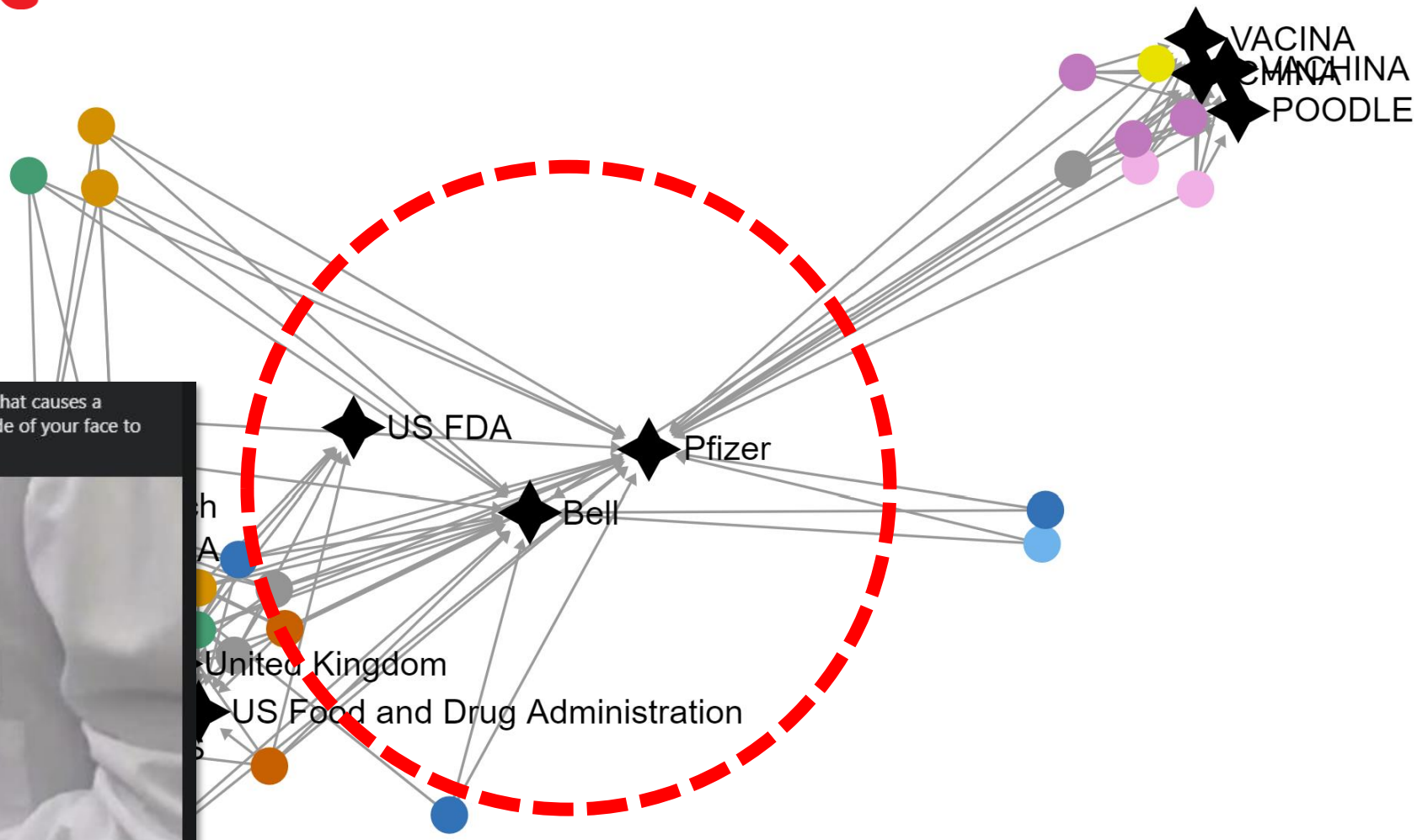
Min Toxicity: 0

0 100

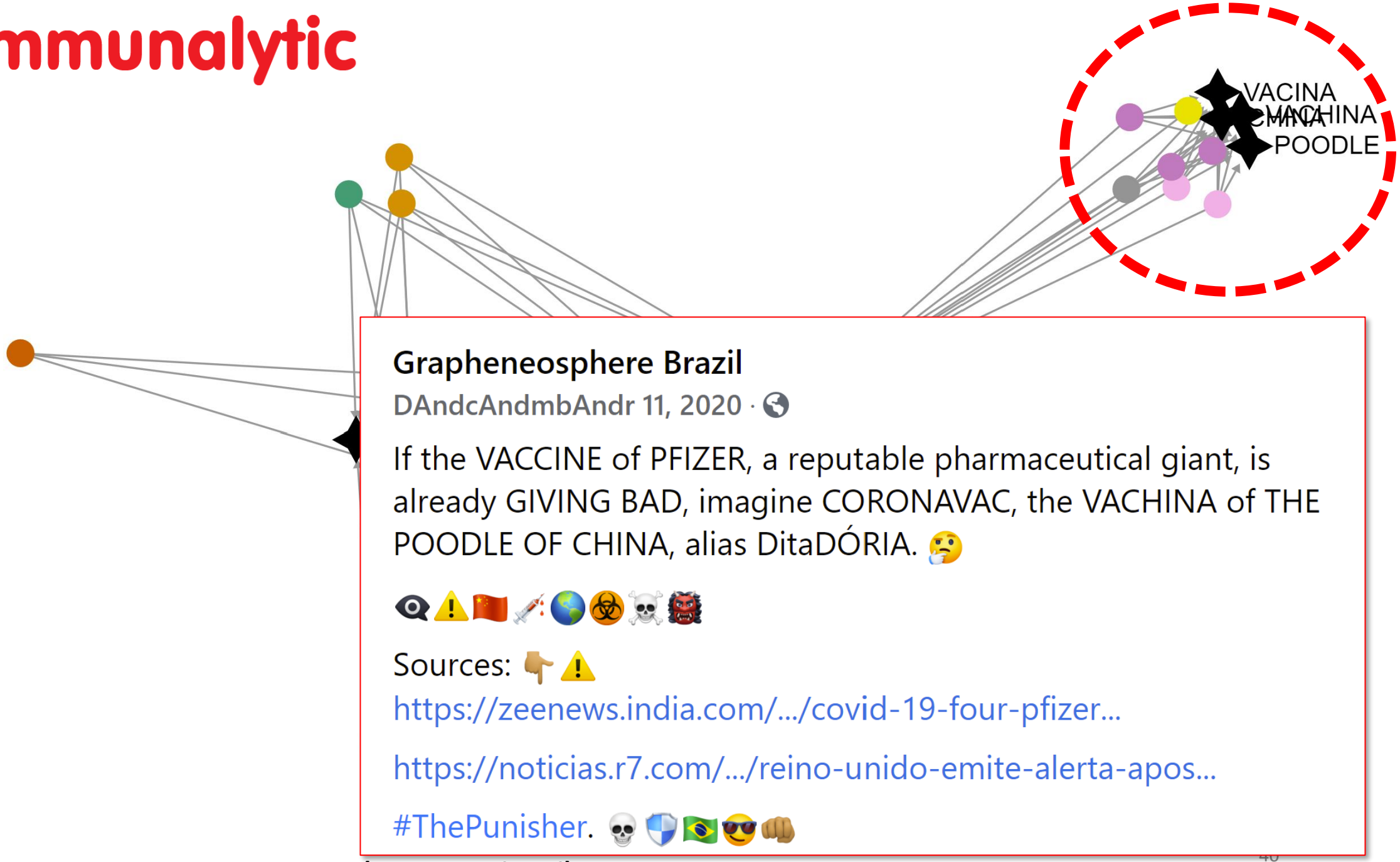
Posts

communalytic





communalytic



Grapheneosphere Brazil

DAndcAndmbAndr 11, 2020 ·

If the VACCINE of PFIZER, a reputable pharmaceutical giant, is already GIVING BAD, imagine CORONAVAC, the VACHINA of THE POODLE OF CHINA, alias DitaDÓRIA. 🤔

👁️⚠️🇨🇳💉🌎☣️☠️👹

Sources: 👇⚠️

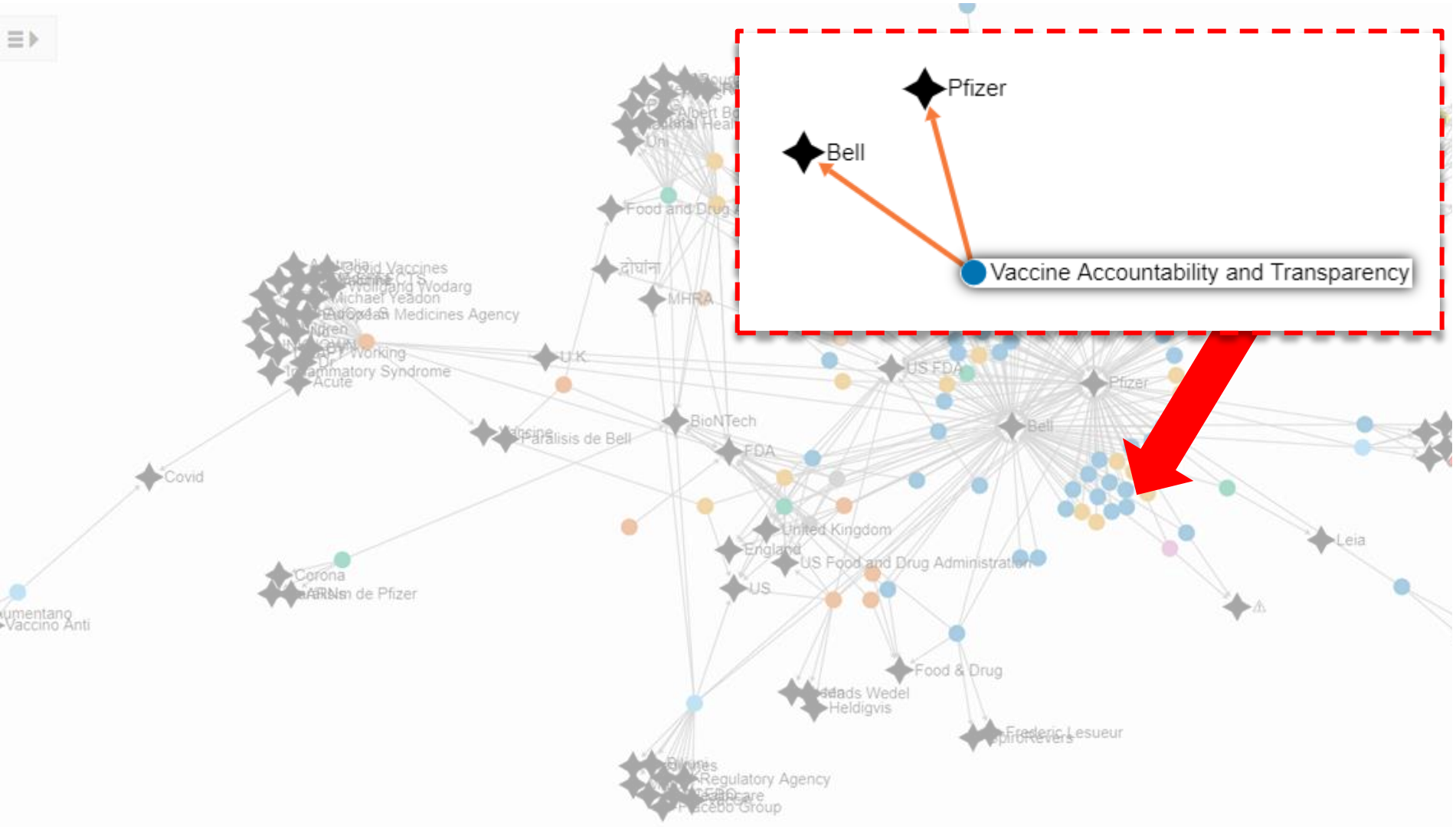
https://zeenews.india.com/.../covid-19-four-pfizer...

https://noticias.r7.com/.../reino-unido-emite-alerta-apos...

#ThePunisher. 💀🛡️🇧🇷😎👊

(auto-translated)

communalytic



Network Properties

Network Layout

Node and Edge Filters

Posts

When you click on a node, only node(s) that are directly connected to the selected node are shown. (To undo the selection, click on an empty space in the network visualization.)

From 2 To 0

The following are posts from **Vaccine Accountability and Transparency:**

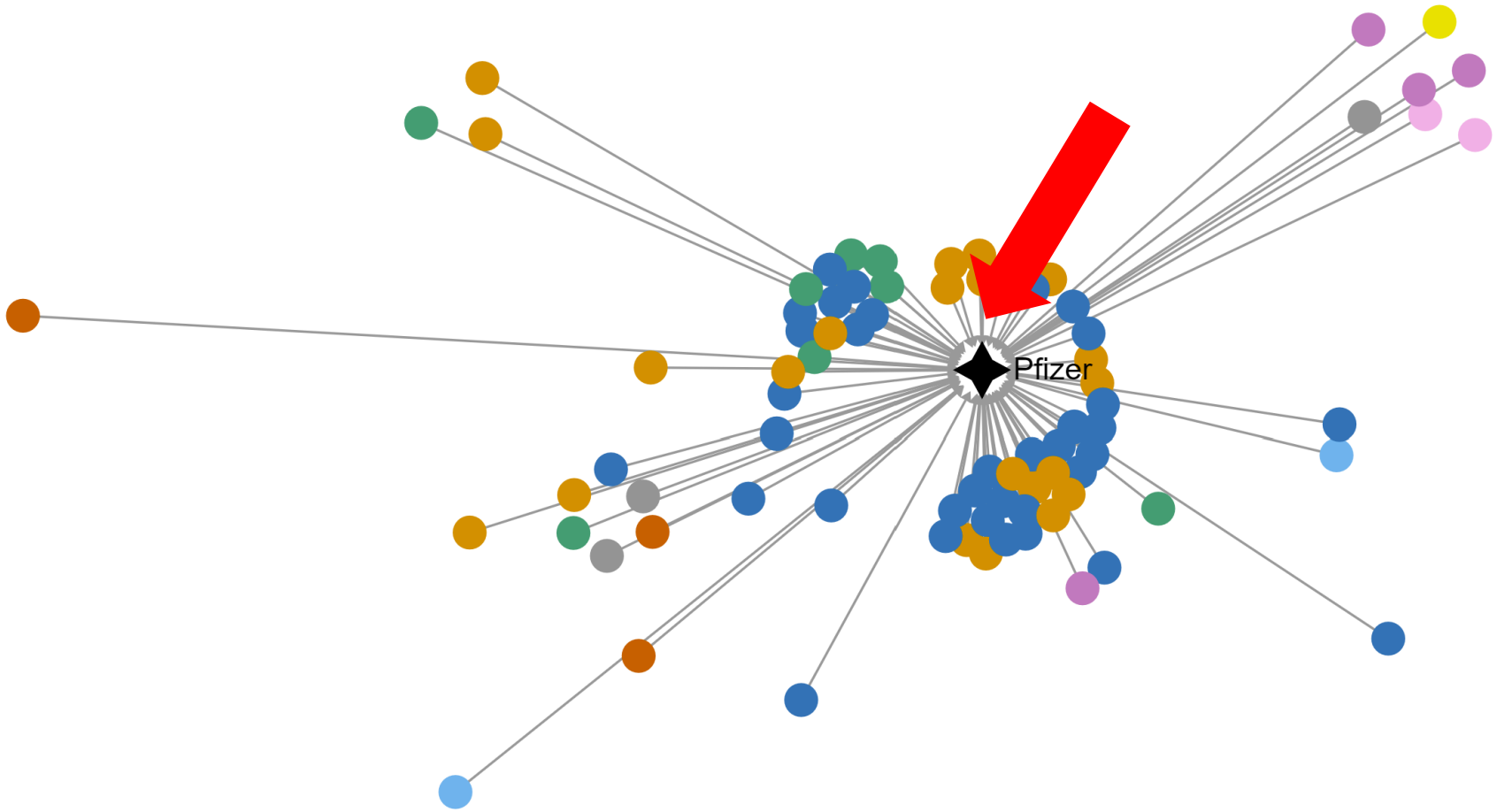
Mentions: Pfizer

Four Pfizer vaccine volunteers develop Bell's palsy... Bell's palsy is a condition that causes a weakness or paralysis of the muscles in the face... The condition causes one side of your face to droop or become stiff. https://zeenews.india.com/world/covid-19-four-pfizer-vaccine-volunteers-develop-bells-palsy-read-details-here-2329924.html

Mentions: Bell

Four Pfizer vaccine volunteers develop Bell's palsy... Bell's palsy is a condition that causes a weakness or paralysis of

communalytic



Network Properties

Network Layout

Node and Edge Filters

Posts ?

When you click on a node, only node(s) that are directly connected to the selected node are shown. (To undo the selection, click on an empty space in the network visualization.)

From 0 To 81

From: Coronavirus News Malta

"Ma Jimpurthax Jekk Xi Numru Ta' Nies Jizviluppaw Mard Serju Bil-Vaccin !!!!!" **Mhux ovvja, jafu jwaqqfu qlieh ta' Biljuni ghax il-Vaccin tikkawza xi mard serju jew mhux f'ammont ta' nies li jkunu haduh. Il-fatt li hadd minn dawk li nghataw il-PLACEBO minflok il-vaccin ma zviluppa Face Paralysis, hija indikazzjoni li l-vaccin aktarx kawzaha

Coronavirus News Malta

Public group · 1.2K members

About | Discussion | Members | Events | Media

[redacted] shared a link.
December 14, 2020 ·

"Ma Jimpurthax Jekk Xi Numru Ta' Nies Jizviluppaw Mard Serju Bil-Vaccin !!!!!"

Mhux ovvja, jafu jwaqqfu qlieh ta' Biljuni ghax il-Vaccin tikkawza xi mard serju jew mhux f'ammont ta' nies li jkunu haduh. Il-fatt li hadd minn dawk li nghataw il-PLACEBO minflok il-vaccin ma zviluppa Face Paralysis, hija indikazzjoni li l-vaccin aktarx kawzaha fuq dawk li hargitilhom.

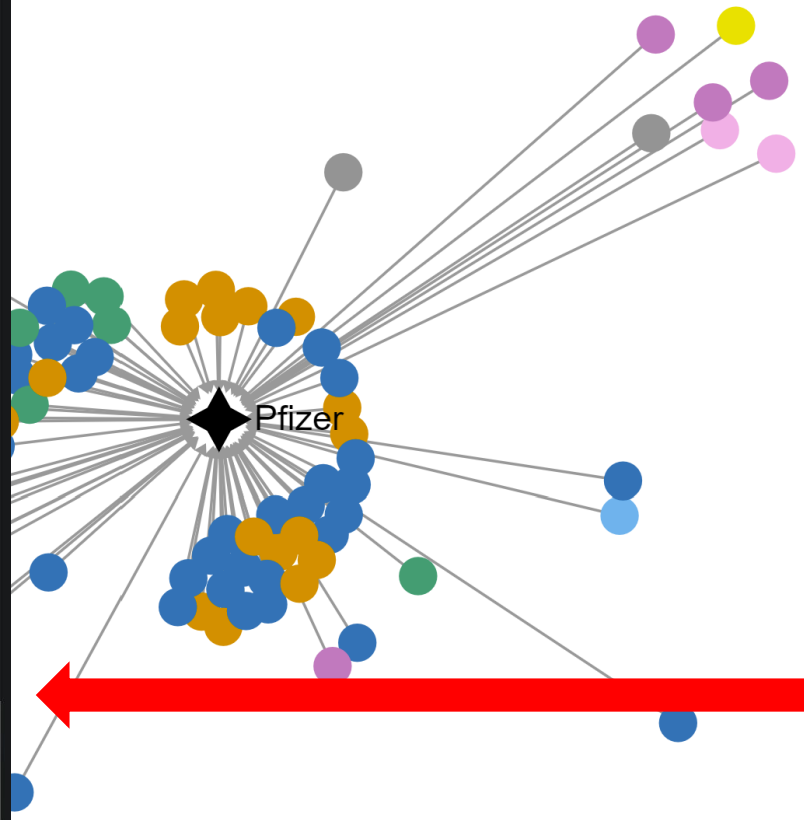
COVID-19: UK regulator makes big claim after 4 Pfizer vaccine volunteers develop Bell's palsy in US

UK drug regulator has allayed fears of Bell's palsy being associated with corona virus vaccine developed by Pfizer.

> UK drug regulator Medicines and Healthcare products Regulatory Agency (MHRA) has allayed fears of Bell's palsy being associated with coronavirus vaccine developed by Pfizer-BioNTech after 4 participants of a US trials suffered from the facial paralysis.

> The vaccine has been developed by US-based pharmaceutical company Pfizer and German biotechnology company BioNTech.

According To Reports, 4 Out Of 21,720 Participants In The US Suffered Bell's Palsy After receiving Pfizer Vaccine But None Out Of The 21,728 Subjects Faced Any Problem In Placebo Group.



Network Properties

Network Layout

Node and Edge Filters

Posts ?

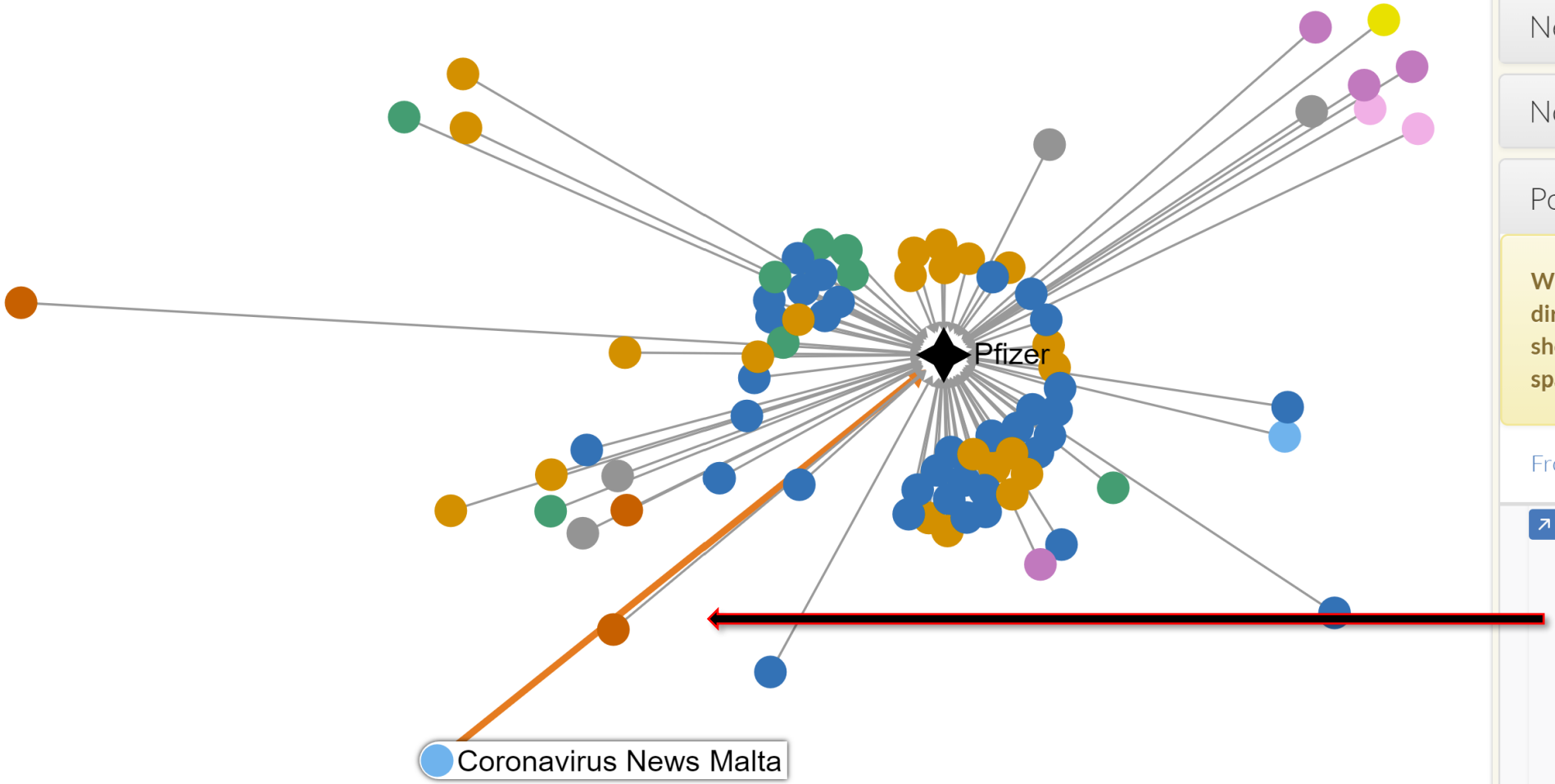
When you click on a node, only node(s) that are directly connected to the selected node are shown. (To undo the selection, click on an empty space in the network visualization.)

From 0 | To 81

From: Coronavirus News Malta

"Ma Jimpurthax Jekk Xi Numru Ta' Nies Jizviluppaw Mard Serju Bil-Vaccin !!!!!" **Mhux ovvja, jafu jwaqqfu qlieh ta' Biljuni ghax il-Vaccin tikkawza xi mard serju jew mhux f'ammont ta' nies li jkunu haduh. Il-fatt li hadd minn dawk li nghataw il-PLACEBO minflok il-vaccin ma zviluppa Face Paralysis, hija indikazzjoni li l-vaccin aktarx kawzaha

communalytic



Network Properties

Network Layout

Node and Edge Filters

Posts ?

When you click on a node, only node(s) that are directly connected to the selected node are shown. (To undo the selection, click on an empty space in the network visualization.)

From 0 To 81

From: Coronavirus News Malta

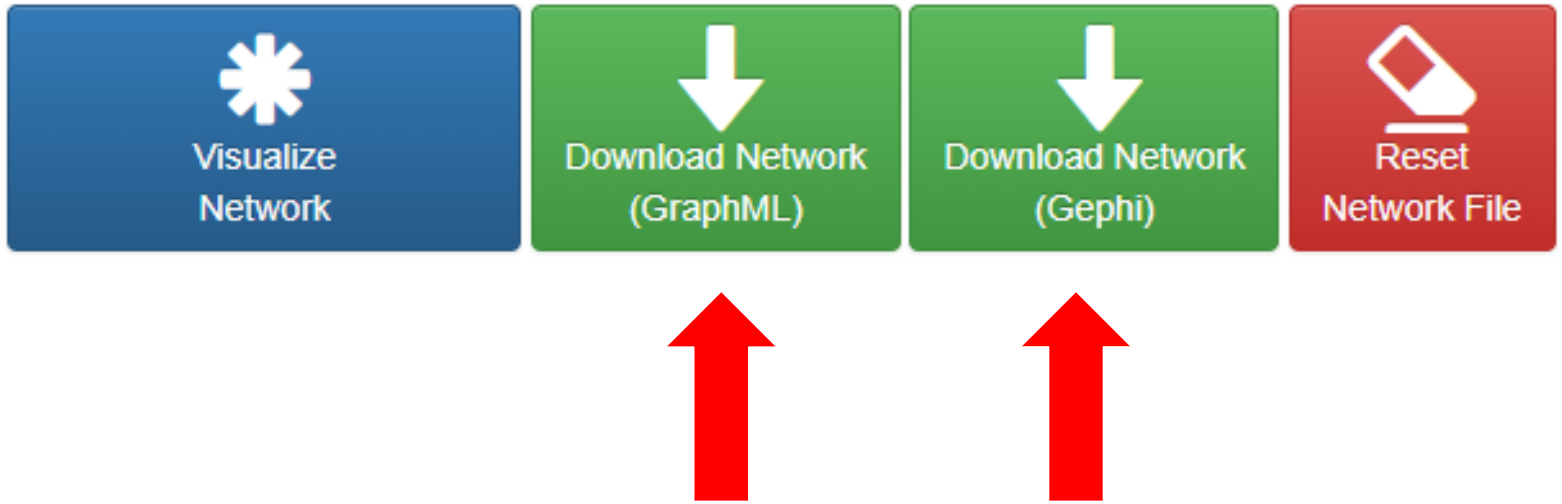
"Ma Jimpurthax Jekk Xi Numru Ta' Nies Jizviluppaw Mard Serju Bil-Vaccin !!!!!" **Mhux ovvja, jafu jwaqqfu qlieh ta' Biljuni ghax il-Vaccin tikkawza xi mard serju jew mhux f'ammont ta' nies li jkunu haduh. Il-fatt li hadd minn dawk li nghataw il-PLACEBO minflok il-vaccin ma zviluppa Face Paralysis, hija indikazzjoni li l-vaccin aktarx kawzaha

Two-Mode Semantic Network Visualizer in Gephi

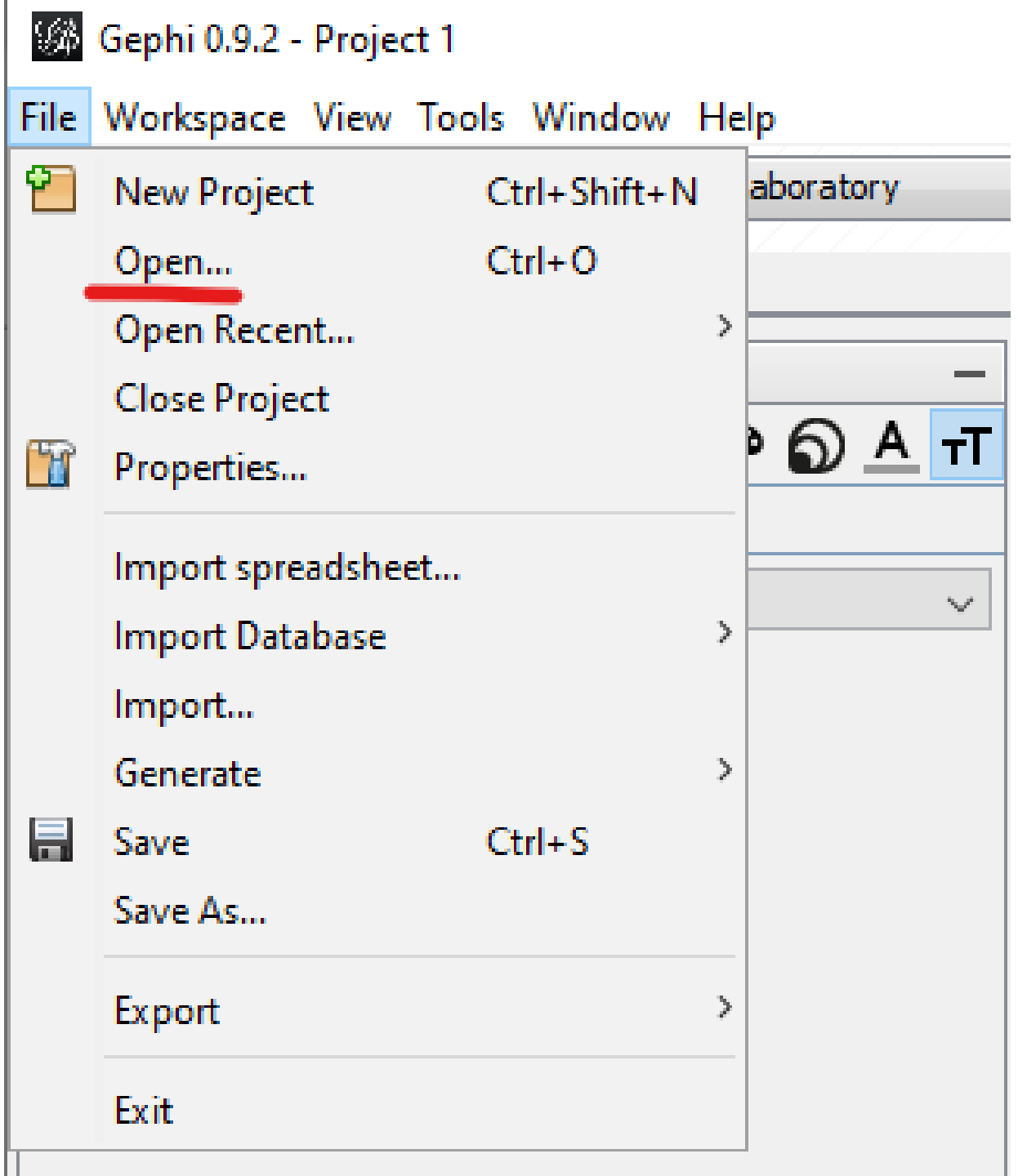
communalytic

Network Analysis

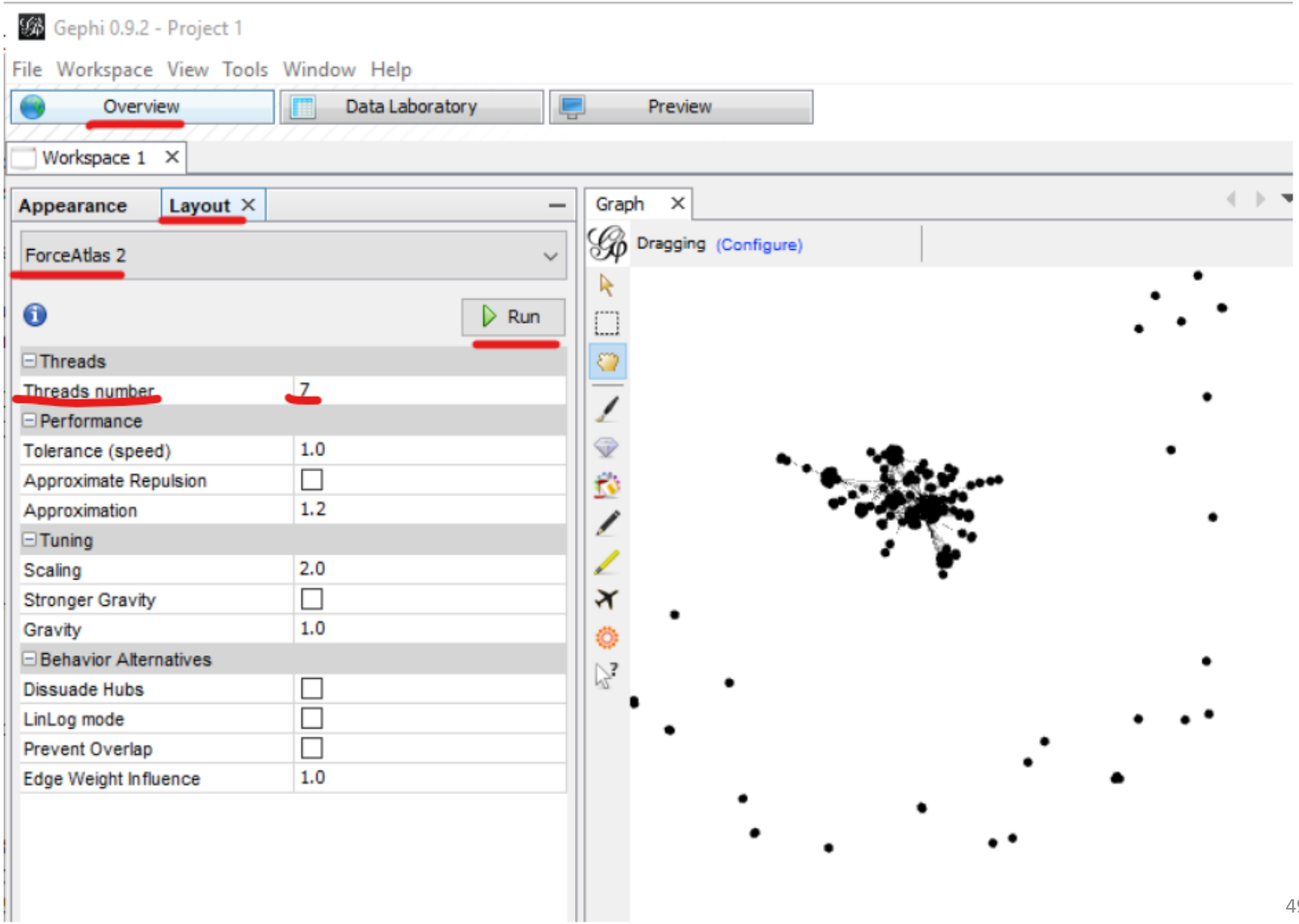
This module automatically creates a two-mode semantic network from your CrowdTangle (FB or IG) dataset. If you are new to Social network Analysis (SNA), check out **this tutorial** to learn more about how two-mode semantic networks are built and how to analyze them.



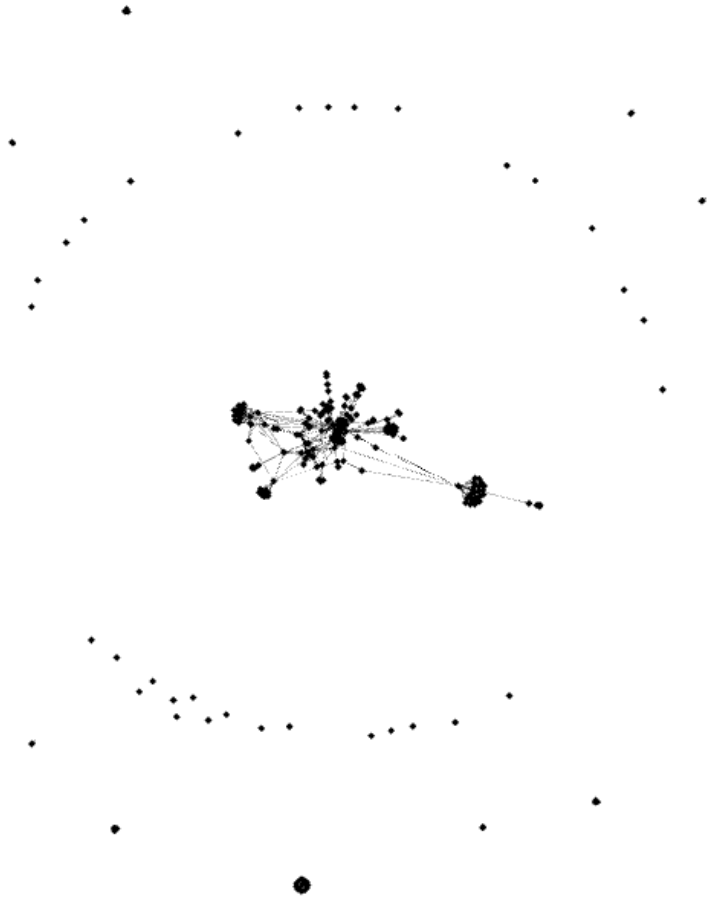
Gephi


Gephi 0.9.2 - Project 1
File
Workspace
View
Tools
Window
Help
New Project Ctrl+Shift+N
Open... Ctrl+O
Open Recent...
Close Project
Properties...
Import spreadsheet...
Import Database
Import...
Generate
Save Ctrl+S
Save As...
Export
Exit
aboratory

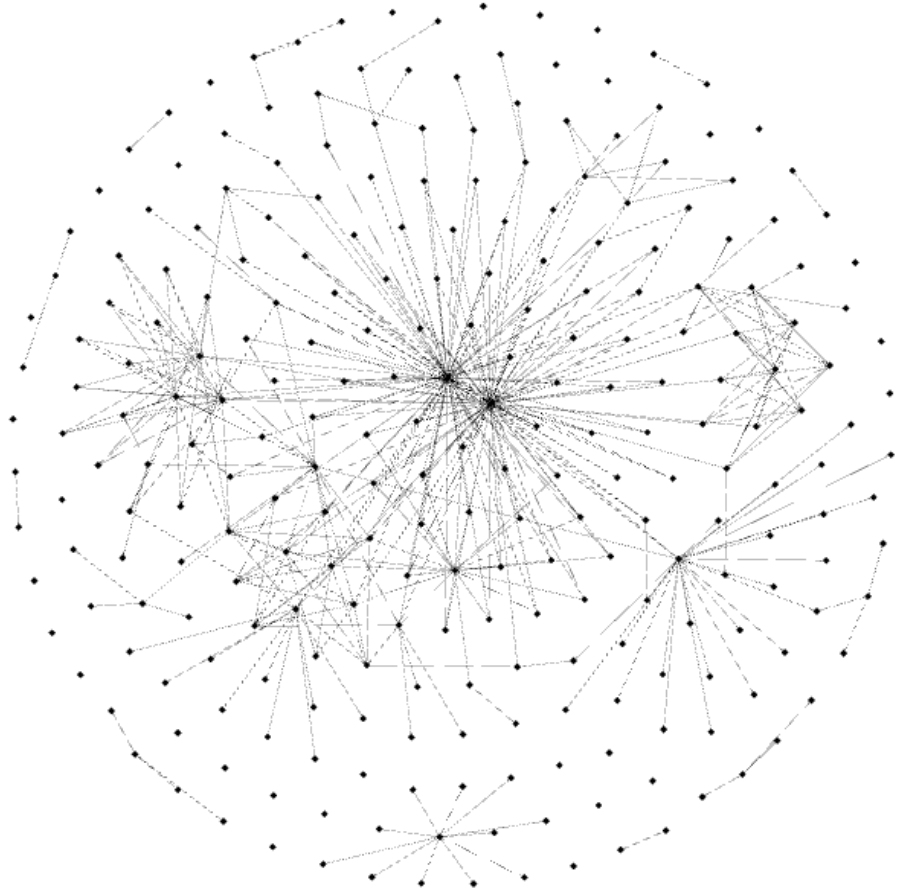
Gephi


Gephi 0.9.2 - Project 1
File Workspace View Tools Window Help
Overview
Data Laboratory
Preview
Workspace 1
Appearance
Layout
ForceAtlas 2
Run
Threads
Threads number
7
Performance
Tolerance (speed)
1.0
Approximate Repulsion
Approximation
1.2
Tuning
Scaling
2.0
Stronger Gravity
Gravity
1.0
Behavior Alternatives
Dissuade Hubs
LinLog mode
Prevent Overlap
Edge Weight Influence
1.0
Graph
Dragging
(Configure)

Gephi

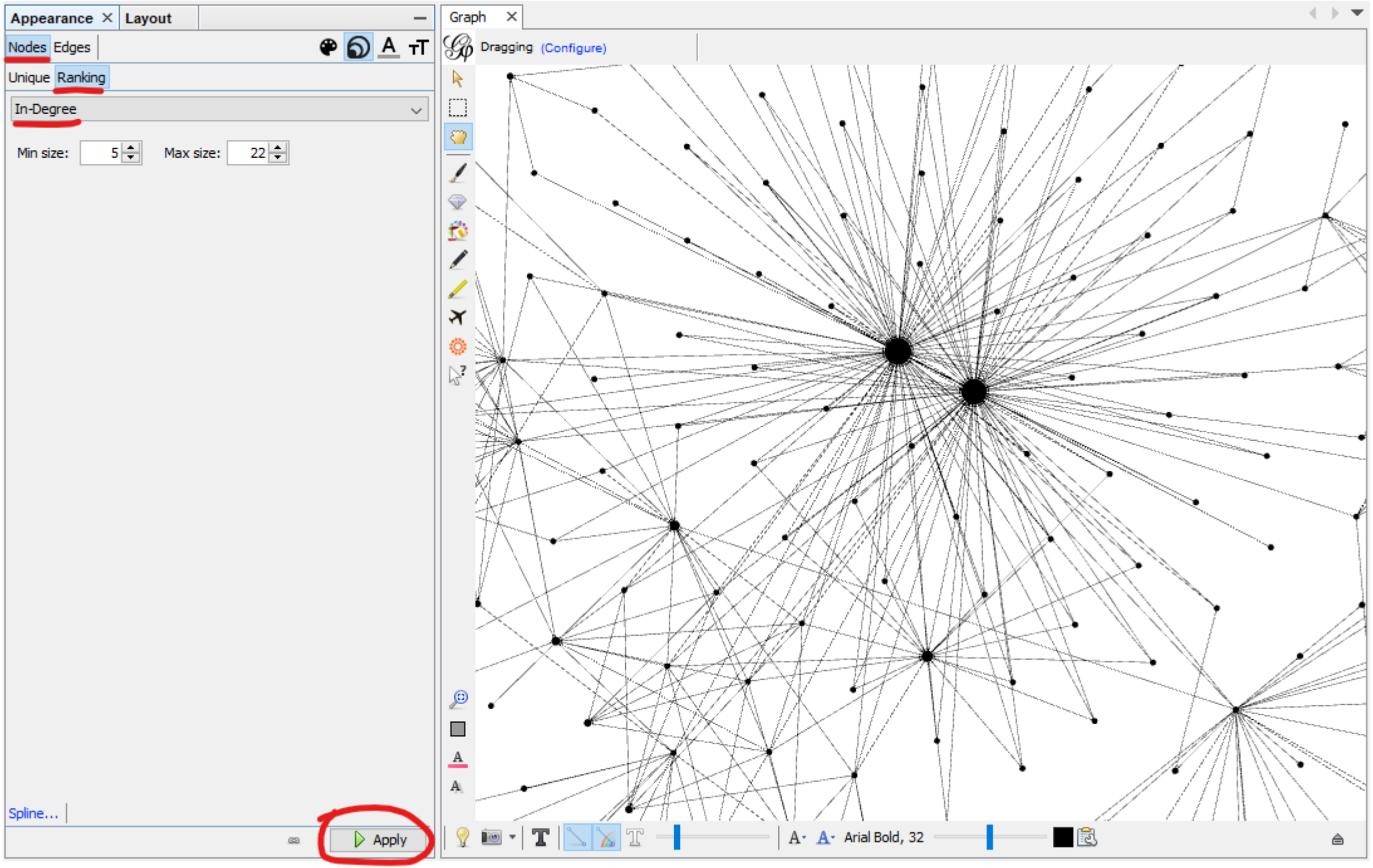


ForceAtlas 2 Layout

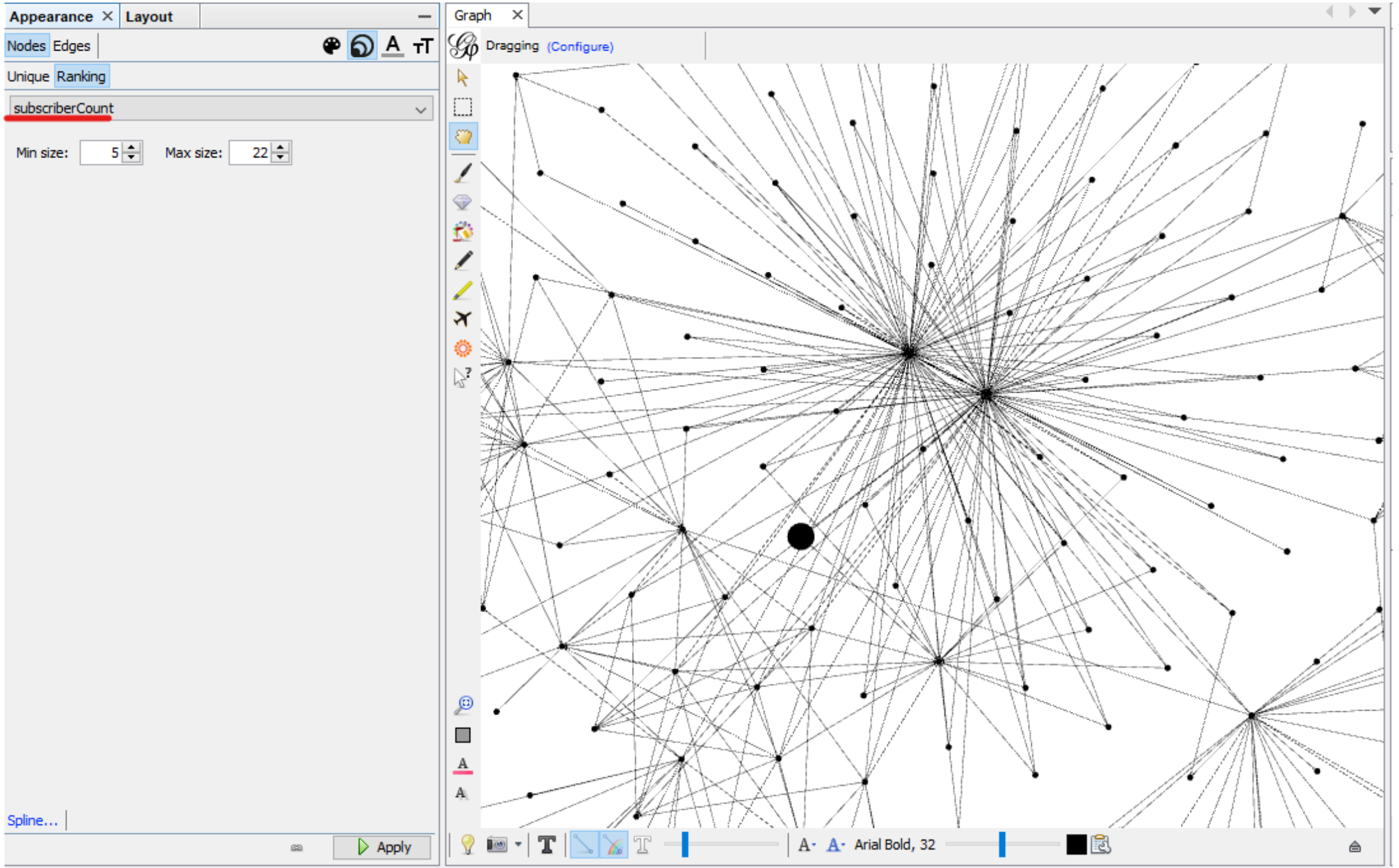


Fruchterman Reingold Layout

Gephi


Appearance
Layout
Nodes
Edges
Unique
Ranking
In-Degree
Min size:
5
Max size:
22
Spline...
Apply
Graph
Dragging
(Configure)
Arial Bold, 32

Gephi


Appearance
Layout
Nodes
Edges
Unique
Ranking
subscriberCount
Min size:
5
Max size:
22
Spline...
Apply
Graph
Dragging
(Configure)
Arial Bold, 32

Gephi

Workspace 1

Appearance | Layout

Nodes | Edges

Unique | Partition | Ranking

node_type

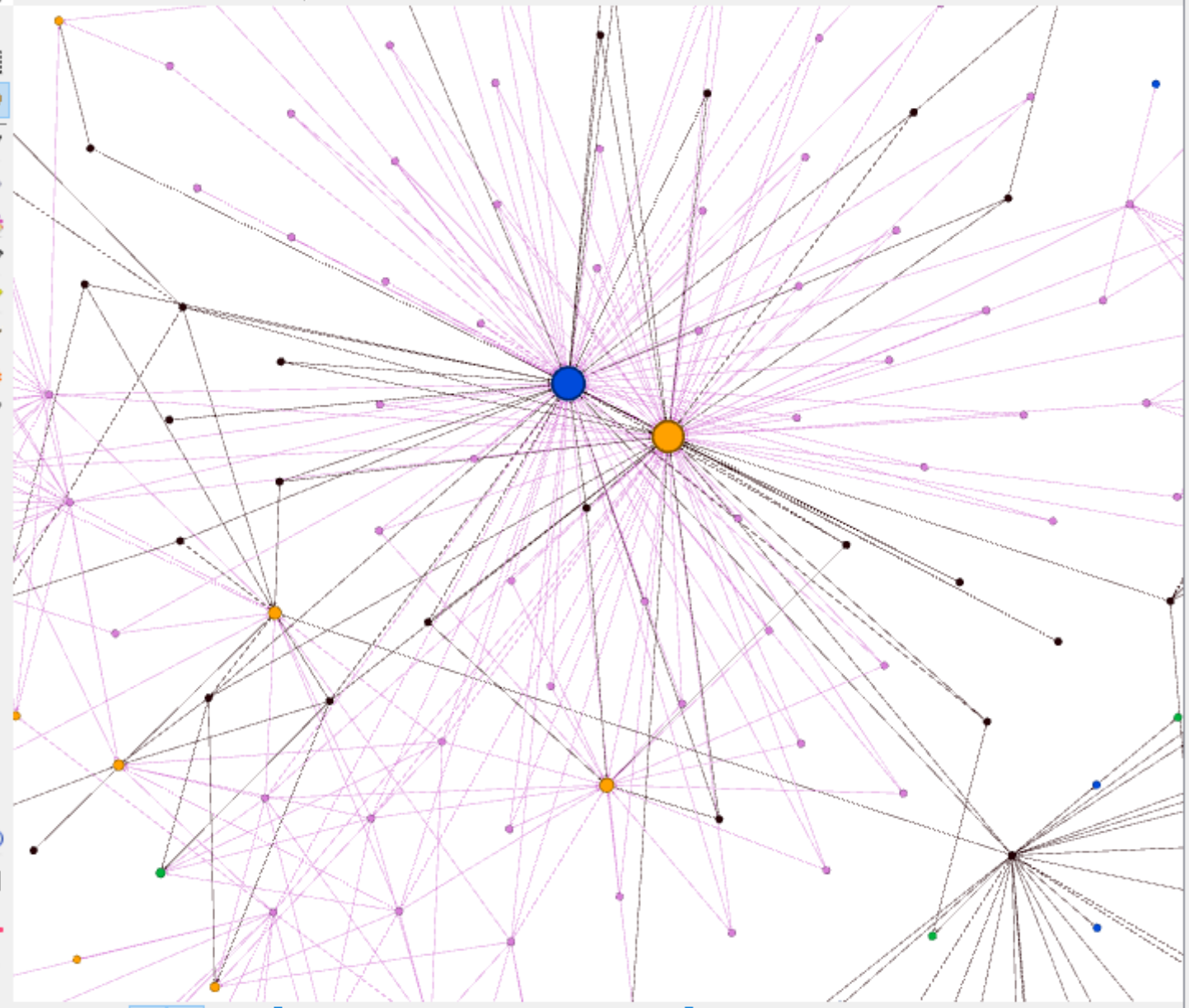
ACCOUNT: Facebook Group	(39.77%)
ACCOUNT: Facebook Page	(15.83%)
NAMED ENTITY: PERSON	(15.06%)
NAMED ENTITY: ORG	(14.67%)
NAMED ENTITY: LOC	(14.67%)

Palette...

Apply

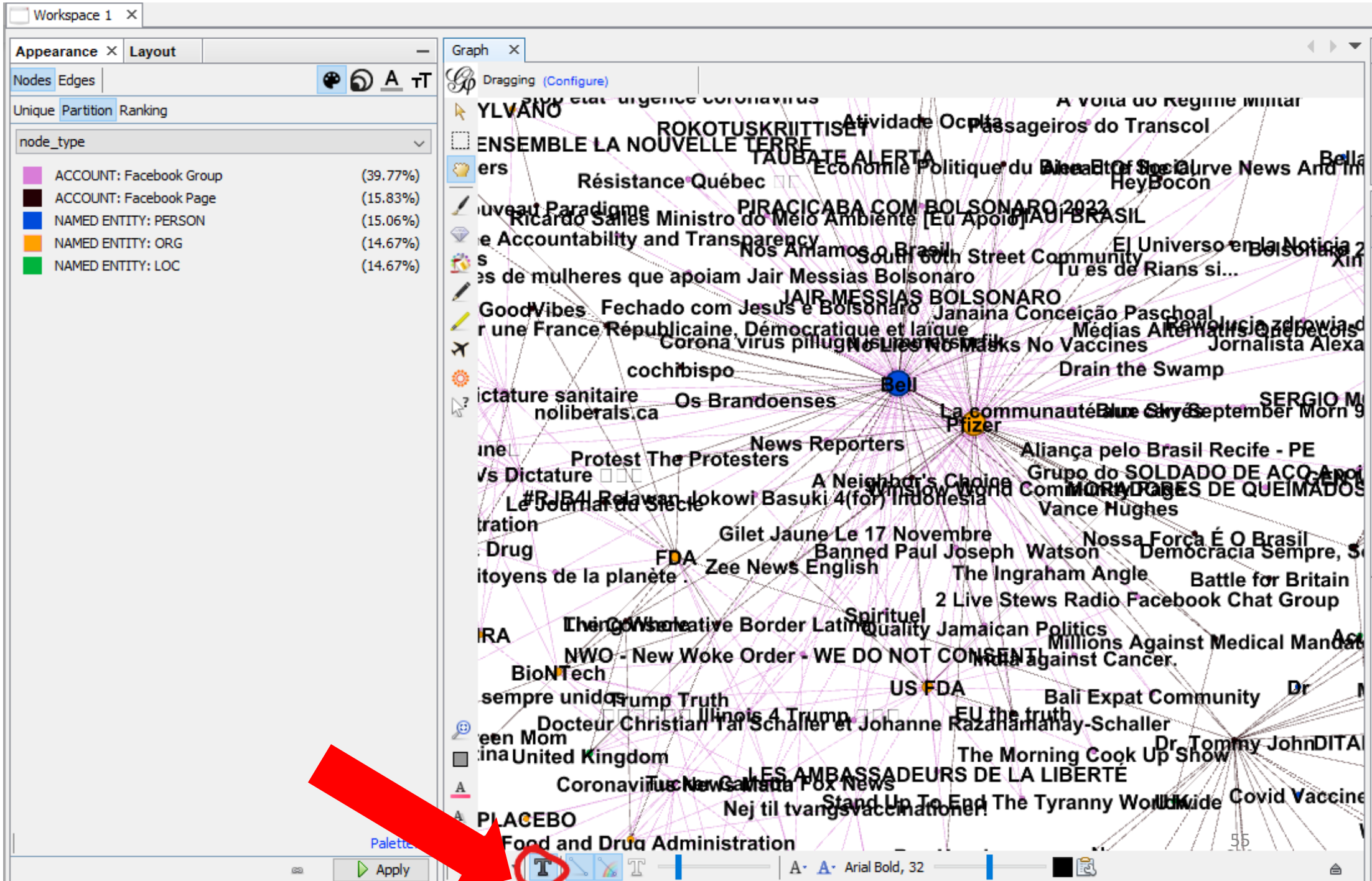
Graph

Dragging (Configure)

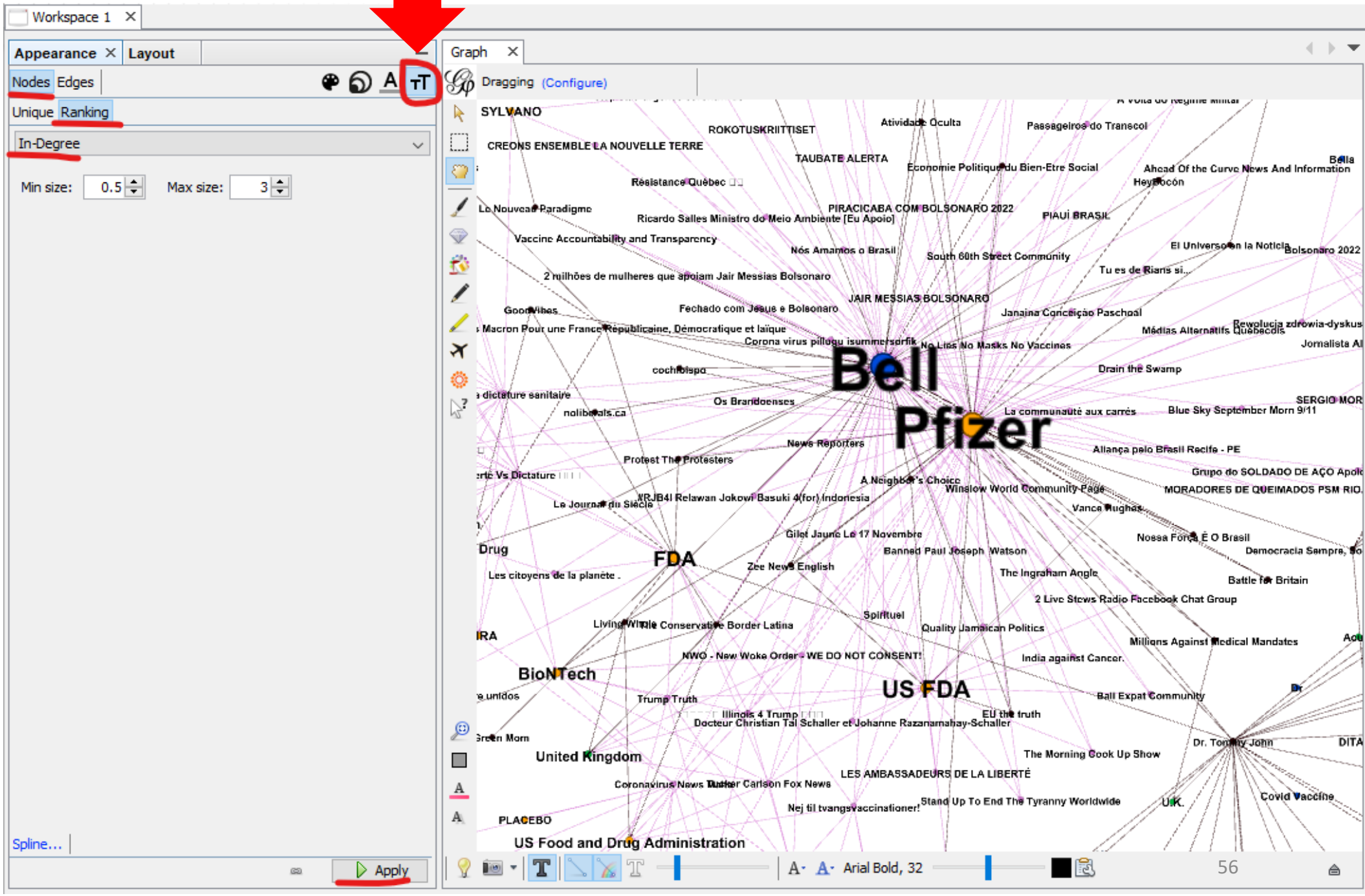


Arial Bold, 32

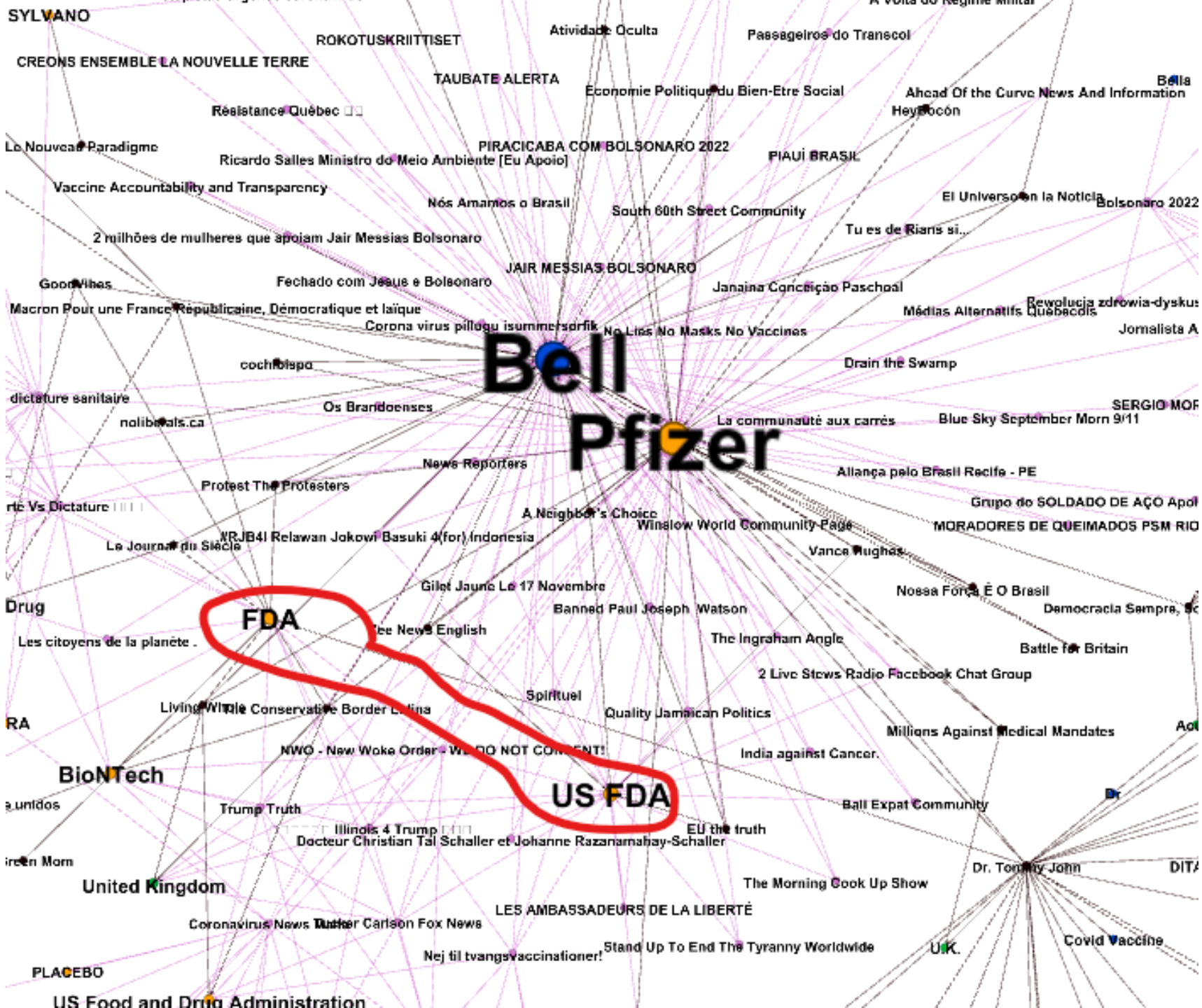
Gephi


Workspace 1
Appearance
Layout
Nodes
Edges
Unique
Partition
Ranking
node_type
ACCOUNT: Facebook Group (39.77%)
ACCOUNT: Facebook Page (15.83%)
NAMED ENTITY: PERSON (15.06%)
NAMED ENTITY: ORG (14.67%)
NAMED ENTITY: LOC (14.67%)
Graph
Dragging (Configure)
Apply
Arial Bold, 32

Gephi

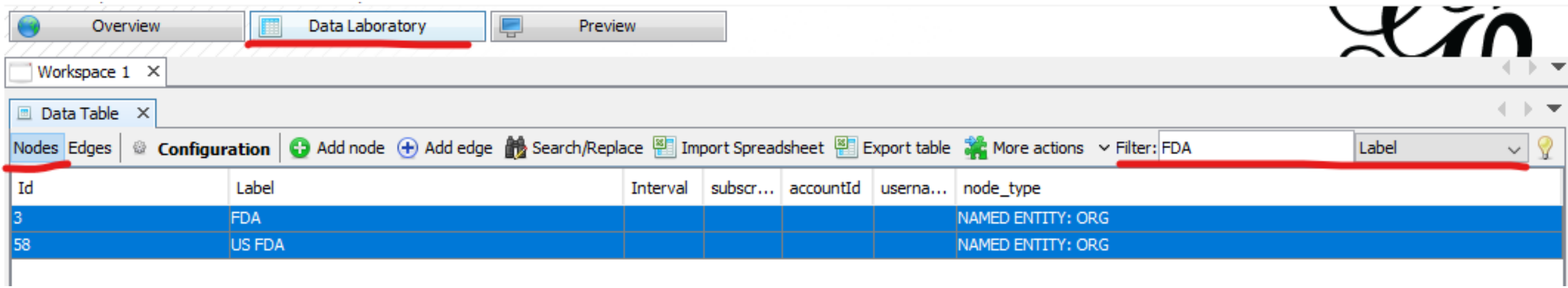

Workspace 1
Appearance
Layout
Nodes
Edges
Unique
Ranking
In-Degree
Min size: 0.5
Max size: 3
Spline...
Apply
Graph
Dragging (Configure)
Bell
Pfizer
FDA
US FDA
BioNTech
United Kingdom
US Food and Drug Administration
Arial Bold, 32
56

Gephi

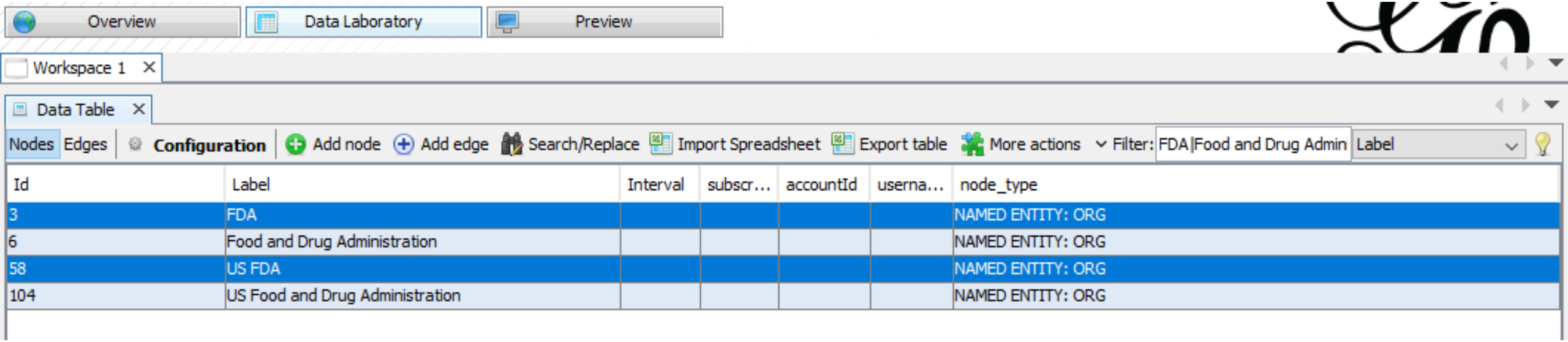


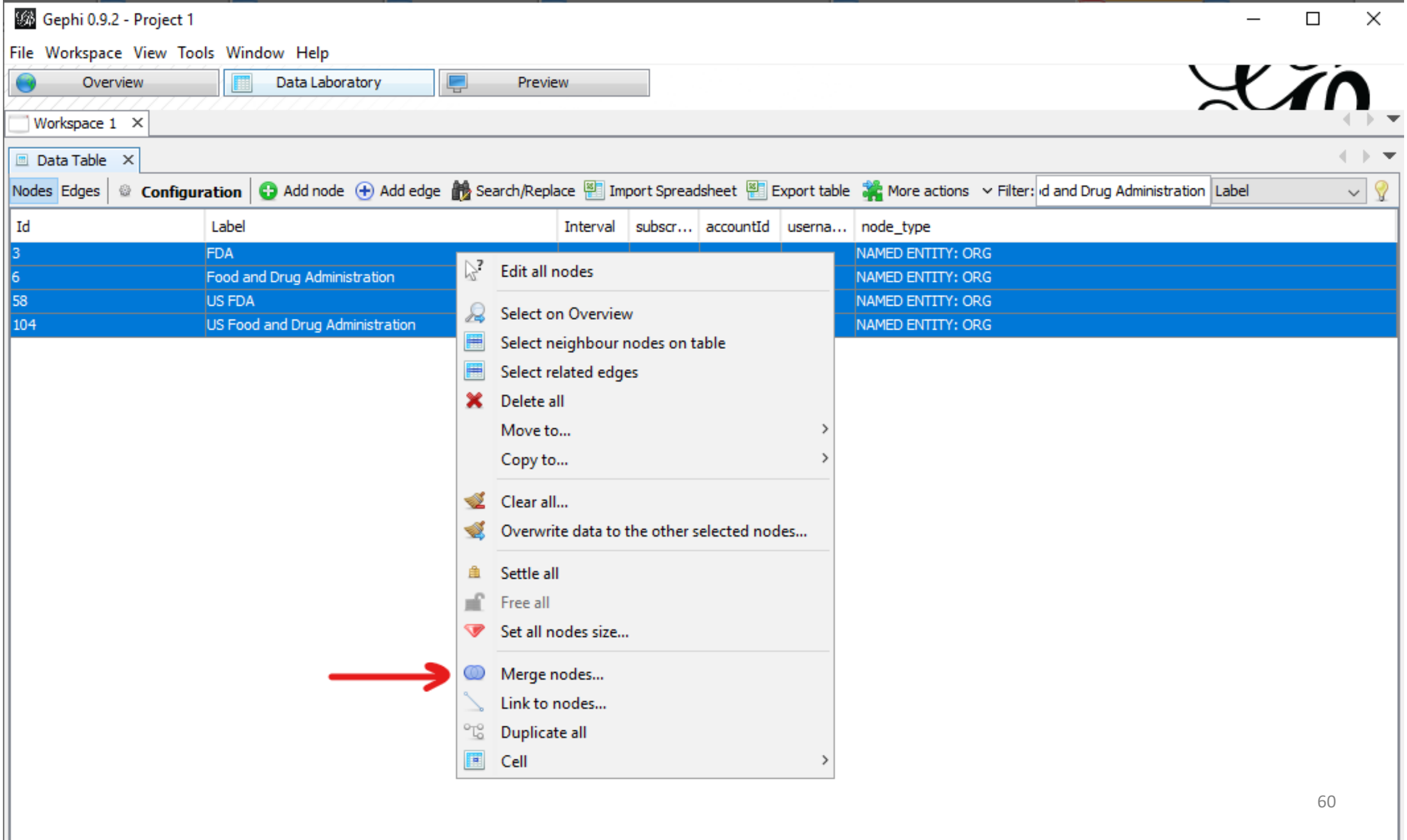
SYLVANO
ROKOTUSKRIITTISET
Atividade Oculta
Passageiros do Transcol
CREONS ENSEMBLE LA NOUVELLE TERRE
TAUBATE ALERTA
Economie Politique du Bien-Etre Social
Ahead Of the Curve News And Information
Bella
Résistance Québec
Le Nouveau Paradigme
PIRACICABA COM BOLSONARO 2022
Ricardo Salles Ministro do Meio Ambiente [Eu Apoio]
PIAUÍ BRASIL
Vaccine Accountability and Transparency
Nós Amamos o Brasil
South 60th Street Community
El Universo en la Noticia
Bolsonaro 2022
2 milhões de mulheres que apoiam Jair Messias Bolsonaro
Tu es de Rians si...
JAIR MESSIAS BOLSONARO
GoodVibes
Fechado com Jesus e Bolsonaro
Janaina Conceição Paschoal
Macron Pour une France Républicaine, Démocratique et laïque
Médias Alternatifs Québecois
Corona virus pillugu isummersorfik
No Lies No Masks No Vaccines
Jornalista A
Bell
cochrolspo
Drain the Swamp
dictature sanitaire
Os Brandoenses
SERGIO MOR
noliberals.ca
La communauté aux carrés
Blue Sky September Morn 9/11
Pfizer
News Reporters
Aliança pelo Brasil Recife - PE
Protest The Protesters
Grupo do SOLDADO DE AÇO Apol
A Neighbor's Choice
Winslow World Community Page
MORADORES DE QUEIMADOS PSM RIO
Le Journal du Siècle
Vance Hughes
Gilet Jaune Le 17 Novembre
Nossa Força É O Brasil
Drug
FDA
Banned Paul Joseph Watson
Democracia Sempre, So
Les citoyens de la planète .
The Ingraham Angle
Battle for Britain
2 Live Stews Radio Facebook Chat Group
Spirituel
Quality Jamaican Politics
Millions Against Medical Mandates
India against Cancer.
BioNTech
US FDA
Trump Truth
Ball Expat Community
Illinois 4 Trump
EU the truth
Docteur Christian Tal Schaller et Johanne Razanamahay-Schaller
Dr. Tommy John
United Kingdom
The Morning Cook Up Show
LES AMBASSADEURS DE LA LIBERTÉ
Coronavirus News
Tucker Carlson Fox News
Stand Up To End The Tyranny Worldwide
U.K.
Covid Vaccine
Nej til tvangsvaccinationer!
PLACEBO
US Food and Drug Administration

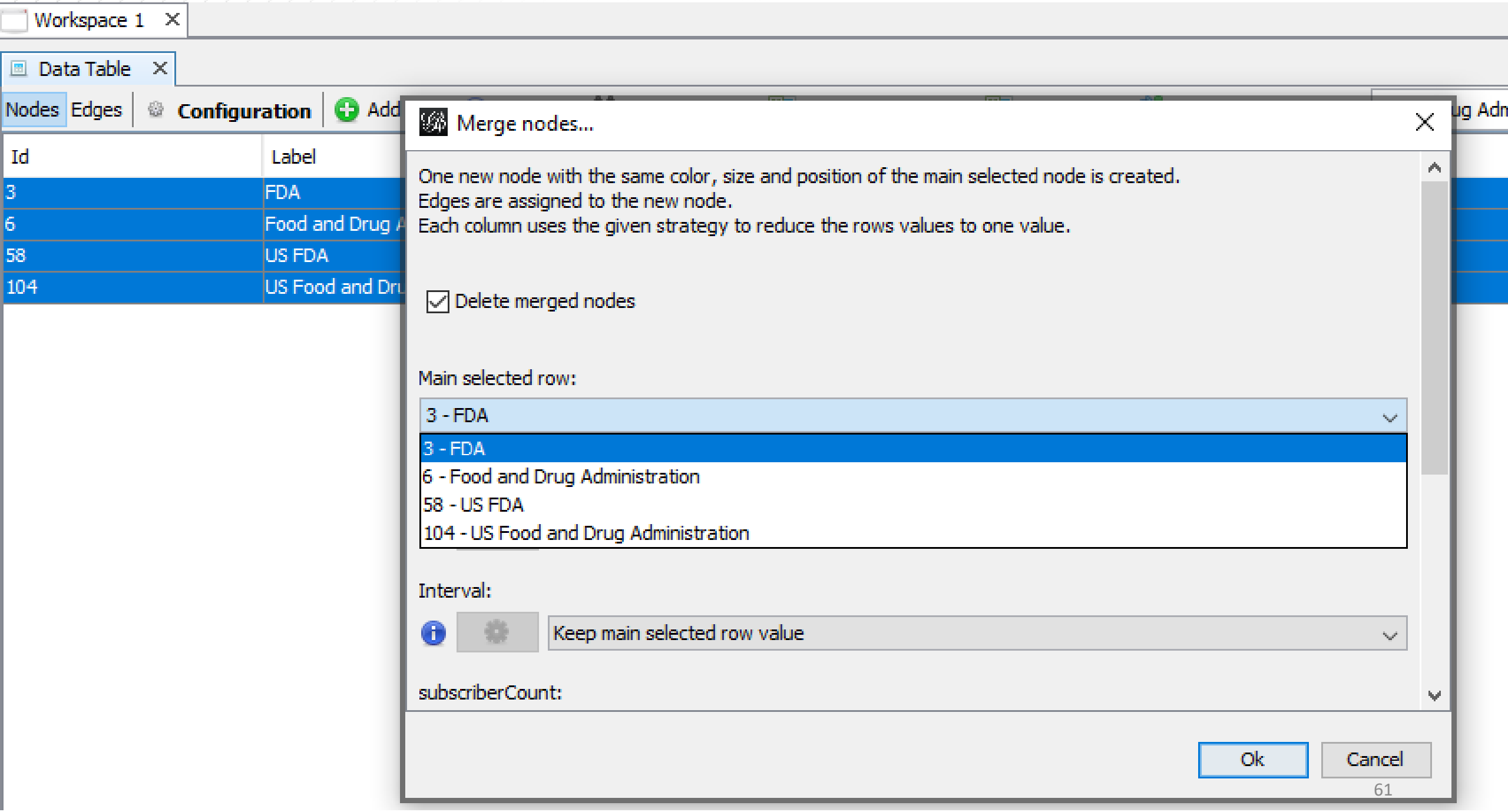
Gephi

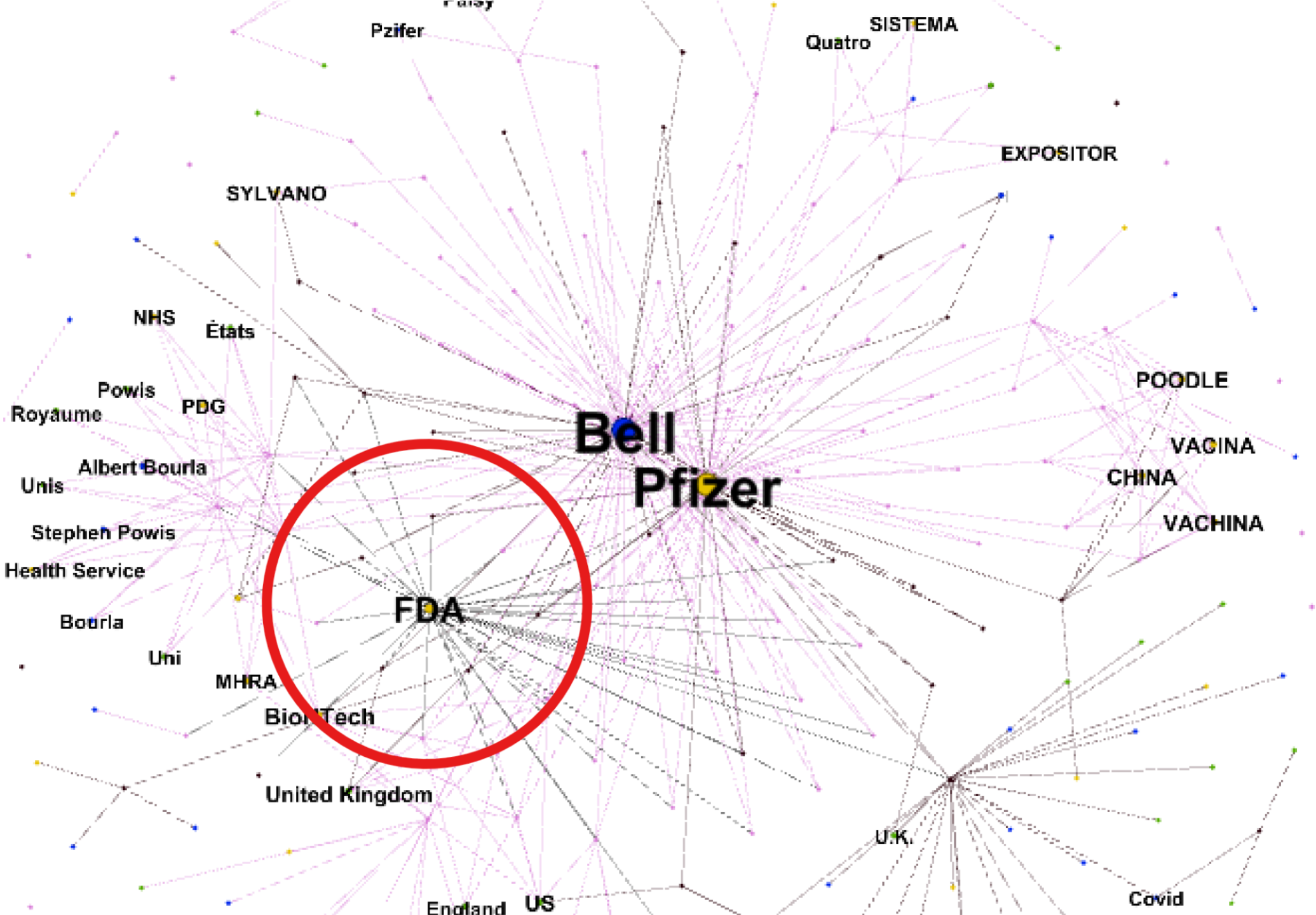

Overview
Data Laboratory
Preview
Workspace 1
Data Table
Nodes
Edges
Configuration
Add node
Add edge
Search/Replace
Import Spreadsheet
Export table
More actions
Filter:
FDA
Label
Id
Label
Interval
subscr...
accountId
userna...
node_type
3
FDA
NAMED ENTITY: ORG
58
US FDA
NAMED ENTITY: ORG

Gephi

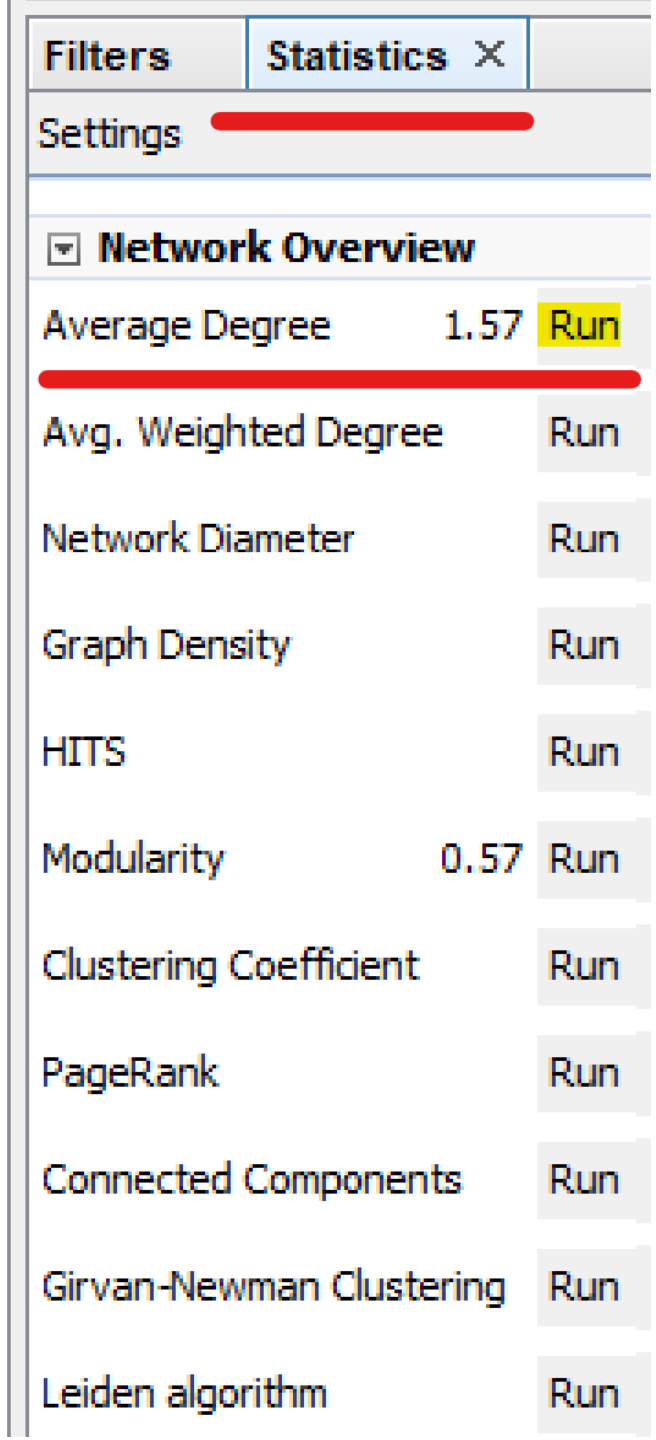

Overview
Data Laboratory
Preview
Workspace 1
Data Table
Nodes
Edges
Configuration
Add node
Add edge
Search/Replace
Import Spreadsheet
Export table
More actions
Filter:
FDA|Food and Drug Admin
Label
Id
Label
Interval
subscr...
accountId
userna...
node_type
3
FDA
NAMED ENTITY: ORG
6
Food and Drug Administration
NAMED ENTITY: ORG
58
US FDA
NAMED ENTITY: ORG
104
US Food and Drug Administration
NAMED ENTITY: ORG


Gephi 0.9.2 - Project 1
File Workspace View Tools Window Help
Overview
Data Laboratory
Preview
Workspace 1
Data Table
Nodes Edges
Configuration
Add node
Add edge
Search/Replace
Import Spreadsheet
Export table
More actions
Filter:
Label
Id
Label
Interval
subscr...
accountId
userna...
node_type
3
FDA
NAMED ENTITY: ORG
6
Food and Drug Administration
NAMED ENTITY: ORG
58
US FDA
NAMED ENTITY: ORG
104
US Food and Drug Administration
NAMED ENTITY: ORG
Edit all nodes
Select on Overview
Select neighbour nodes on table
Select related edges
Delete all
Move to...
Copy to...
Clear all...
Overwrite data to the other selected nodes...
Settle all
Free all
Set all nodes size...
Merge nodes...
Link to nodes...
Duplicate all
Cell

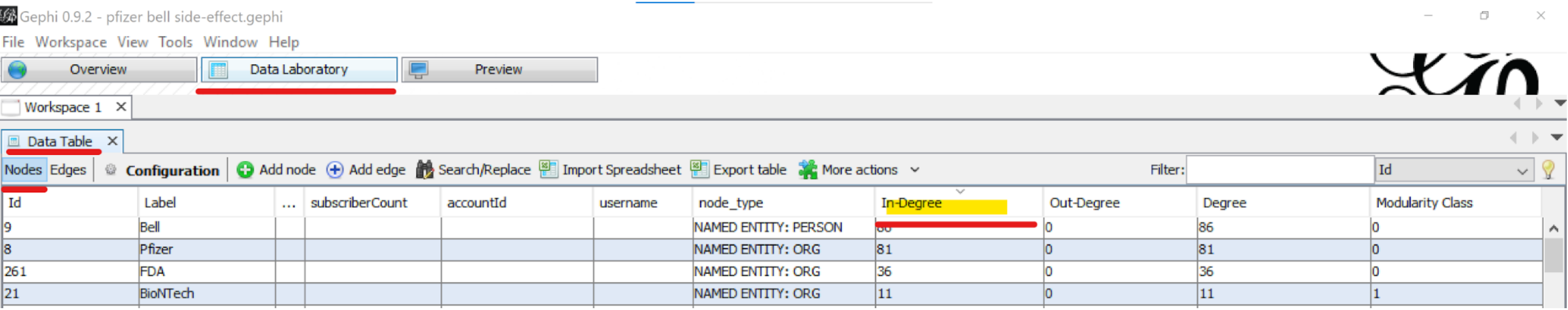

Workspace 1
Data Table
Nodes
Edges
Configuration
Id
Label
3
FDA
6
Food and Drug A
58
US FDA
104
US Food and Dru
Merge nodes...
One new node with the same color, size and position of the main selected node is created.
Edges are assigned to the new node.
Each column uses the given strategy to reduce the rows values to one value.
Delete merged nodes
Main selected row:
3 - FDA
3 - FDA
6 - Food and Drug Administration
58 - US FDA
104 - US Food and Drug Administration
Interval:
Keep main selected row value
subscriberCount:
Ok
Cancel


Pzifer
SISTEMA
Quatro
EXPOSITOR
SYLVANO
NHS
Etats
Powis
PDG
Royaume
Albert Bourla
Unis
Stephen Powis
Health Service
Bourla
Uni
MHRA
BioNTech
United Kingdom
England
US
Bell
Pfizer
FDA
POODLE
VACINA
CHINA
VACHINA
U.K.
Covid

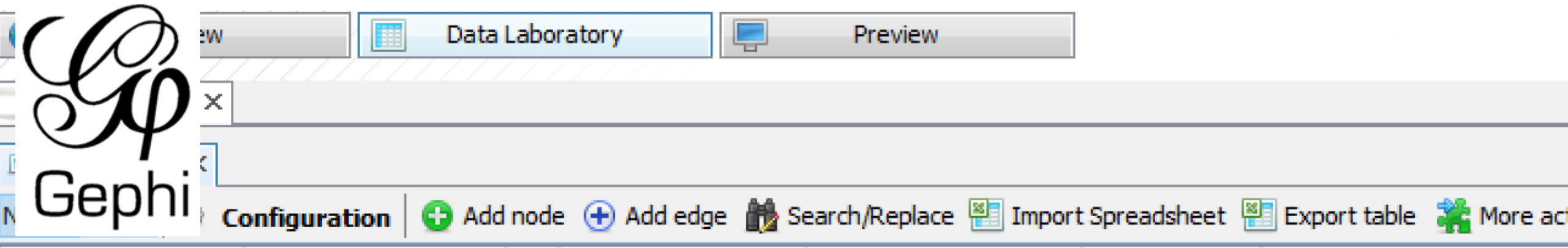
Gephi


Filters
Statistics
Settings
Network Overview
Average Degree 1.57 Run
Avg. Weighted Degree Run
Network Diameter Run
Graph Density Run
HITS Run
Modularity 0.57 Run
Clustering Coefficient Run
PageRank Run
Connected Components Run
Girvan-Newman Clustering Run
Leiden algorithm Run

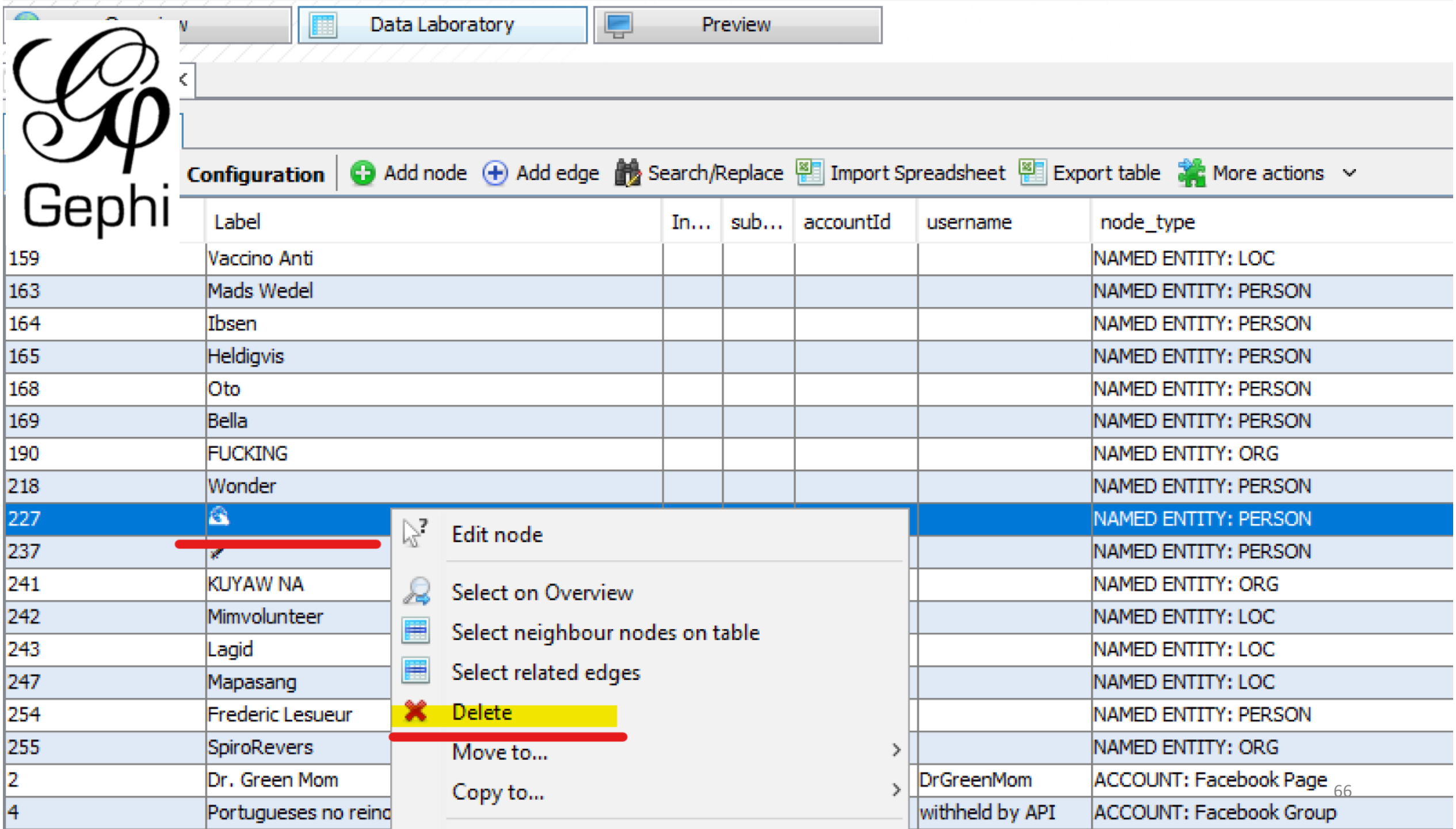
Gephi

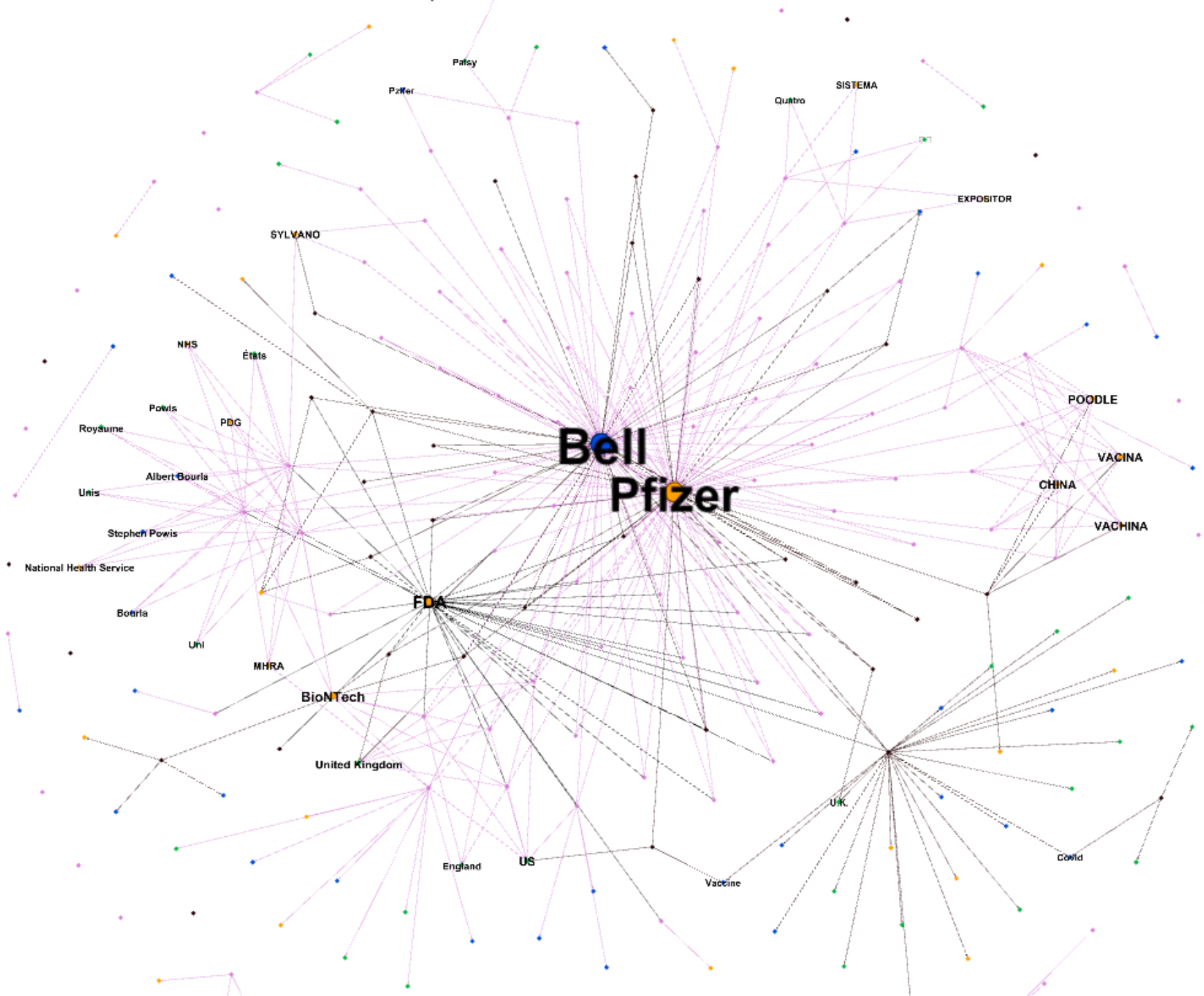


Id	Label	...	subscriberCount	accountId	username	node_type	In-Degree	Out-Degree	Degree	Modularity Class
9	Bell					NAMED ENTITY: PERSON	86	0	86	0
8	Pfizer					NAMED ENTITY: ORG	81	0	81	0
261	FDA					NAMED ENTITY: ORG	36	0	36	0
21	BioNTech					NAMED ENTITY: ORG	11	0	11	1



Id	Label	...	subscriberCount	accountId	username	node_type
9	Bell					NAMED ENTITY: PERSON
8	Pfizer					NAMED ENTITY: ORG
261	FDA					NAMED ENTITY: ORG
21	BioNTech					NAMED ENTITY: ORG
16	US					NAMED ENTITY: LOC
53	VACINA					NAMED ENTITY: ORG
54	VACHINA					NAMED ENTITY: ORG
55	POODLE					NAMED ENTITY: ORG
56	CHINA					NAMED ENTITY: ORG
103	United Kingdom					NAMED ENTITY: LOC
20	MHRA					NAMED ENTITY: ORG
30	Food & Drug					NAMED ENTITY: ORG
252	SYLVANO					NAMED ENTITY: ORG

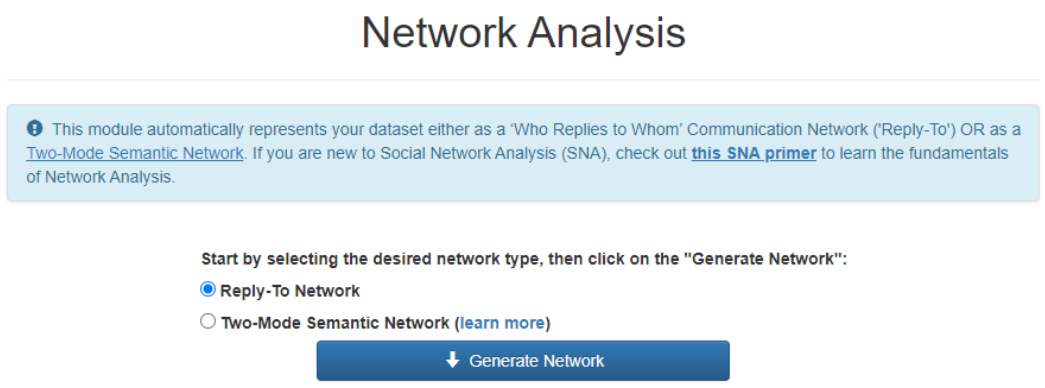

Gephi
Data Laboratory
Preview
Configuration
Add node
Add edge
Search/Replace
Import Spreadsheet
Export table
More actions
Label
In...
sub...
accountId
username
node_type
159 Vaccino Anti NAMED ENTITY: LOC
163 Mads Wedel NAMED ENTITY: PERSON
164 Ibsen NAMED ENTITY: PERSON
165 Heldigvis NAMED ENTITY: PERSON
168 Oto NAMED ENTITY: PERSON
169 Bella NAMED ENTITY: PERSON
190 FUCKING NAMED ENTITY: ORG
218 Wonder NAMED ENTITY: PERSON
227 NAMED ENTITY: PERSON
237 NAMED ENTITY: PERSON
241 KUYAW NA NAMED ENTITY: ORG
242 Mimvolunteer NAMED ENTITY: LOC
243 Lagid NAMED ENTITY: LOC
247 Mapasang NAMED ENTITY: LOC
254 Frederic Lesueur NAMED ENTITY: PERSON
255 SpiroRevers NAMED ENTITY: ORG
2 Dr. Green Mom DrGreenMom ACCOUNT: Facebook Page
4 Portugueses no reino withheld by API ACCOUNT: Facebook Group
Edit node
Select on Overview
Select neighbour nodes on table
Select related edges
Delete
Move to...
Copy to...


Pfizy
Pzifer
SISTEMA
Quatro
EXPOSITOR
SYLVANO
NHS
Etats
Powis
PDG
Royaume
Albert Bourla
Unis
Stephen Powis
National Health Service
Bourla
Uni
MHRA
BioNTech
United Kingdom
England
US
Vaccine
UK
Covid
POODLE
VACINA
CHINA
VACHINA
Bell
Pfizer
FDA

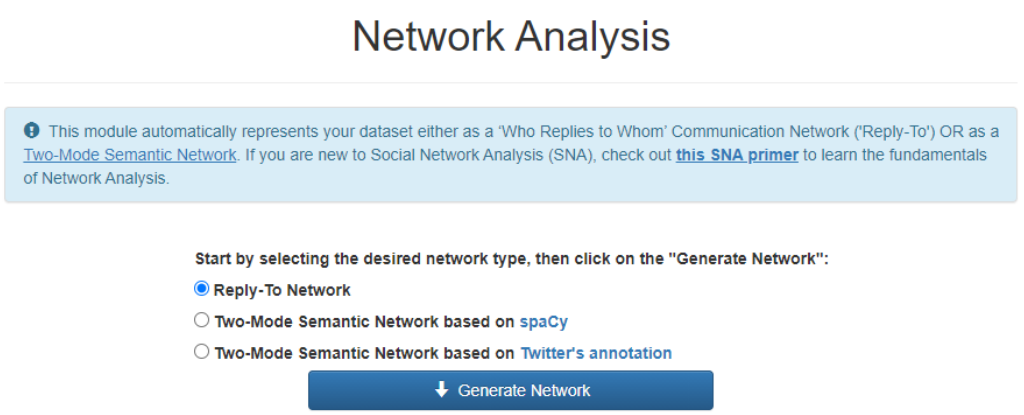
Semantic Networks for Other Dataset Types

communalytic

Reddit/CSV Datasets



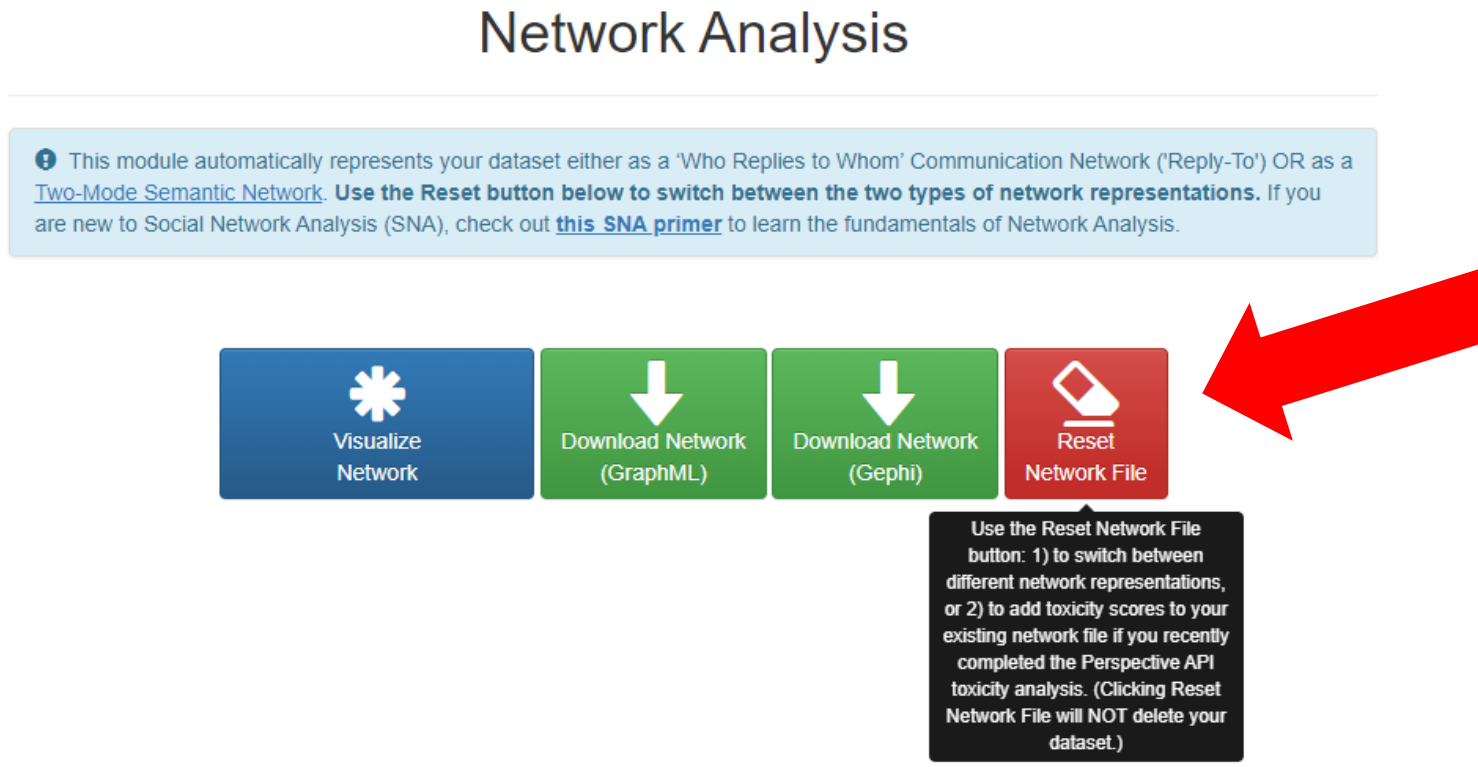
Twitter Datasets



Semantic Networks for Other Dataset Types

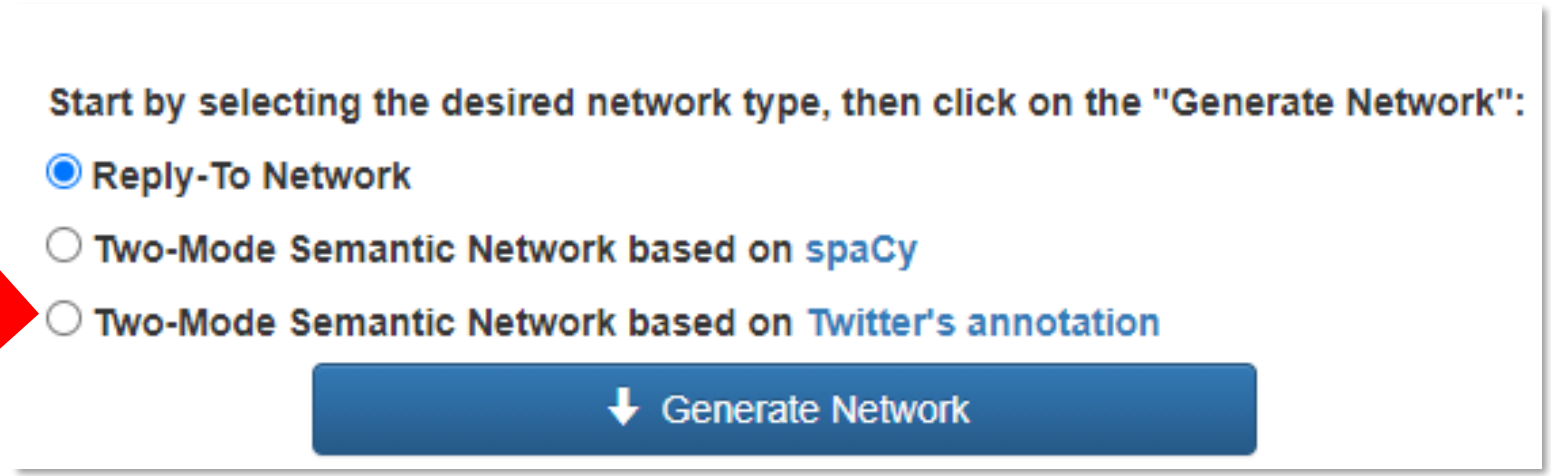
communalytic

Use the **Reset Network File button** to switch between the two types of network representations



communalytic

Additional option to discover Named Entities for Twitter datasets



Tweet annotation types

Entities

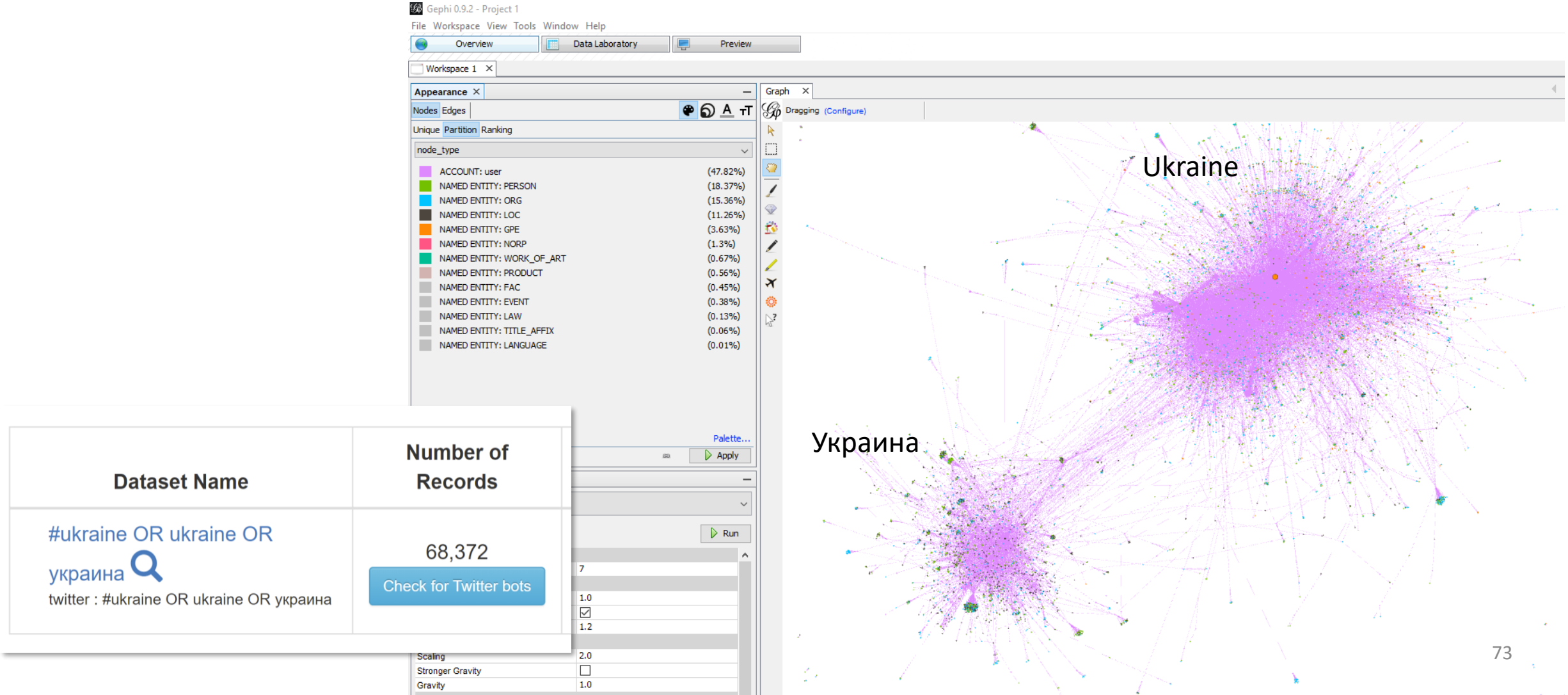
Entity annotations are programmatically defined entities that are nested within the entities field and are reflected as annotations in the payload. Each annotation has a confidence score, and an indication of where in the Tweet text the entities were identified (start and end fields).

The entity annotations can have the following types:

- **Person** - Barack Obama, Daniel, or George W. Bush
- **Place** - Detroit, Cali, or "San Francisco, California"
- **Product** - Mountain Dew, Mozilla Firefox
- **Organization** - Chicago White Sox, IBM
- **Other** - Diabetes, Super Bowl 50

https://developer.twitter.com/en/docs/twitter-api/annotations

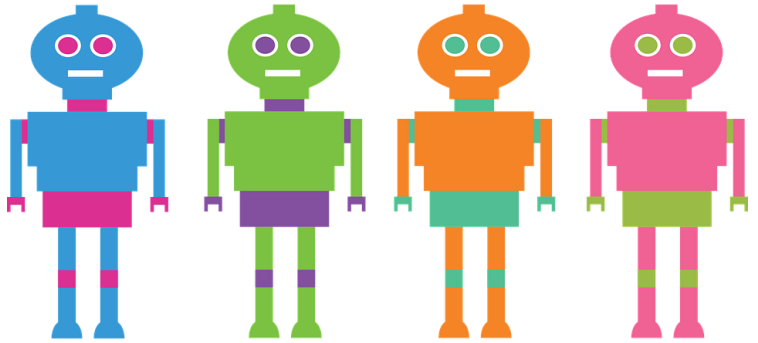
Example: Two-mode Semantic Network of a Twitter dataset mentioning **Ukraine OR Украина**



Dataset Name	Number of Records
#ukraine OR ukraine OR украина twitter : #ukraine OR ukraine OR украина	68,372 Check for Twitter bots

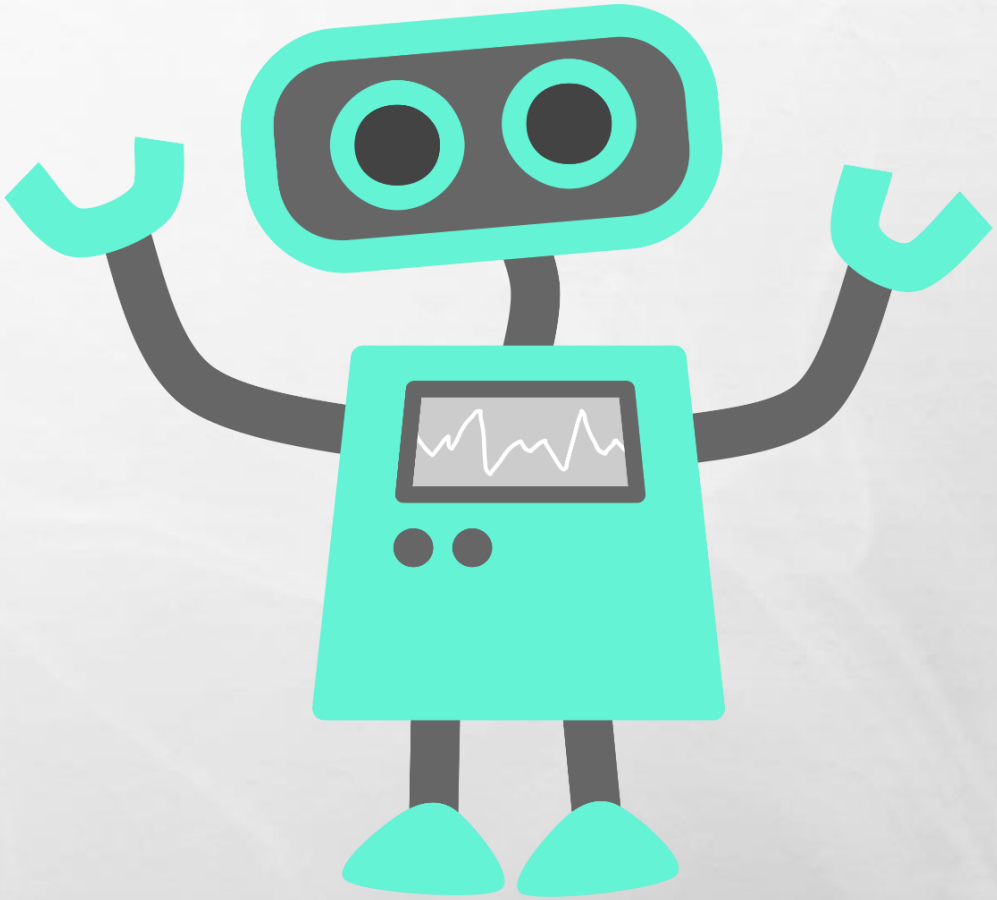
Outline

1. Data Collection using CrowdTangle (Facebook & Instagram)
2. Two-mode Semantic Networks
3. **Twitter Bot Detection**

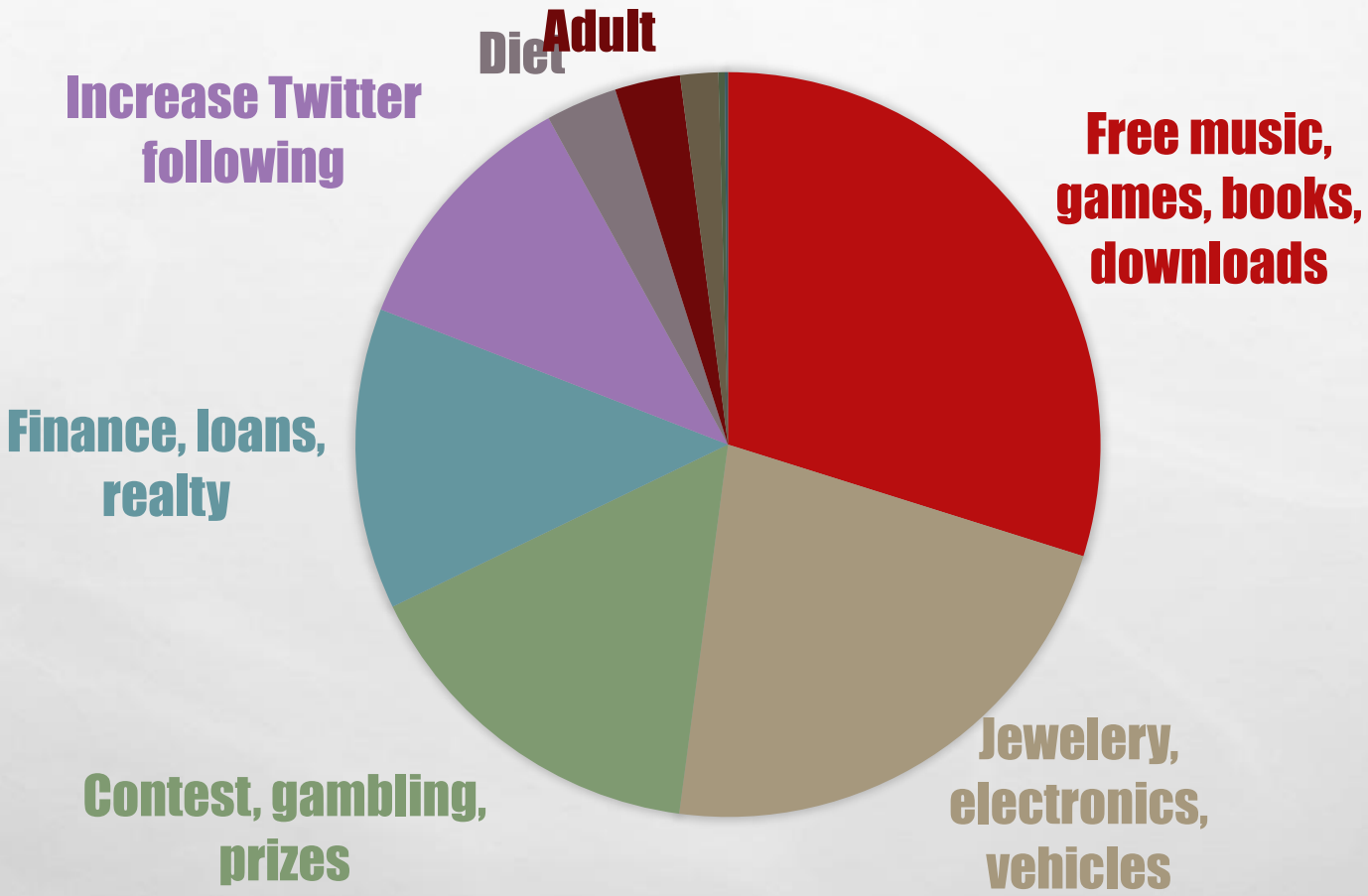


SOCIAL BOT

SOFTWARE DESIGNED TO ACT ON THE INTERNET WITH SOME LEVEL OF AUTONOMY

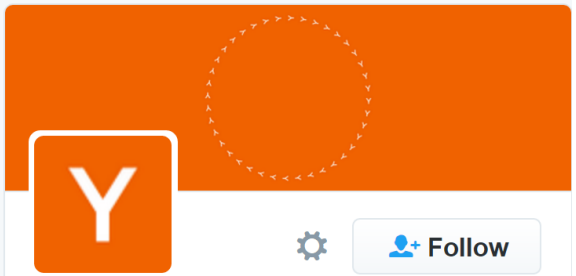


DIFFERENT TYPES OF BOTS



(Grier et al, 2010)

Sample Twitter Bots

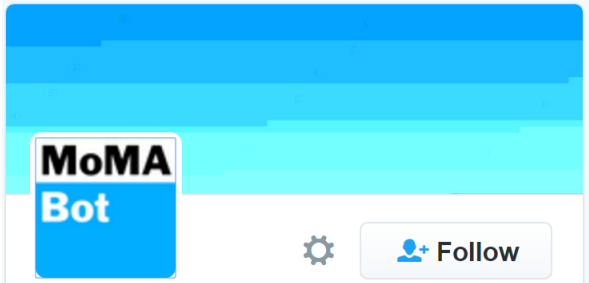


Follow

Hacker News

@newsycombinator

I'm a news.ycombinator bot, get the latest from Hacker News! Built by the lovely @riklomas from @superhi_ – sport…



Follow

MoMA Bot

@MoMARobot

I tweet a random object from the collection of the Museum of Moderm Art four times a day. Bot by …

Follow

The Lesser Bot

@ebooks_goetia

The Lesser Bot of Solomon offers you endless pages from a text in the style of Ars Goetia and the Pseudomonarchia Daemonum. Courtesy @digitalsqand…

Follow

Netflix Bot

@netflix_bot

Tweets new releases on Netflix Instant. Created by @zak123. Check out @Redbox_bot as well. Updates between 5-7AM MST Every Morning.

Follow

Lowpoly Bot

@Lowpolybot

A Twitter bot by @Quasimondo that creates random low-polygon versions of pictures it receives.…

Follow

Generate A Cat Bot

@GenerateACat

Yet another procedural cat generating bot. FOLLOW and TWEET AT ME to generate a cat! Generator by @mousefountain. Consciousness by …

Follow

Yayfrens Bot

@yayfrens

Hello! I am a loving friend bot! @mossdogmusic and @inurashii created me (with help from @cirne) and I love you all very much!

Follow

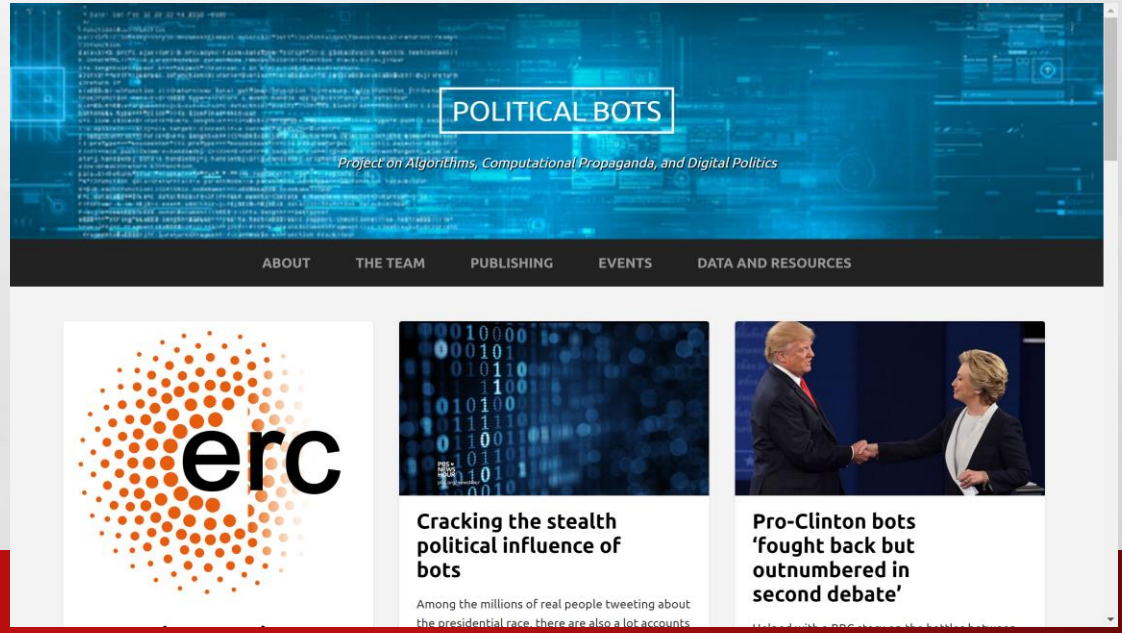
Reverse OCR

@reverseocr

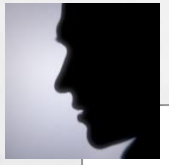
A bot that picks a word and then draws randomly until an OCR library (antimatter15.com/ocrad.js/demo.…) reads that word. 4x daily. Also on…

“BAD” BOTS

- SOCIAL BOTNET DESIGNED TO EXPOSE PRIVATE INFORMATION
- HINDER THE ADVANCEMENT OF PUBLIC POLICY
- ARTIFICIALLY ENLARGING THE AUDIENCE OF SOME PEOPLE
- RUIN THE REPUTATION OF A COMPANY



DETECTING BOTS…



Photo

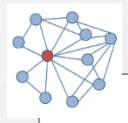
- Color & Edge histograms
- Color & Edge Directivity Descriptor (CEDD)
- Image Similarity

Message

-

Poster

-

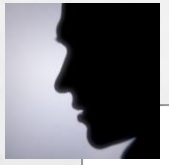


Social Network

-

(Yardi et al, '09; Grier et al, '10; Wang, '10; Jin et al, '11)

DETECTING BOTS…



Photo

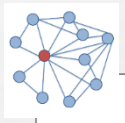
- Color & Edge histograms
- Color & Edge Directivity Descriptor (CEDD)
- Image Similarity

Message

- Sensitive words
- URL
- Duplicates
- #hashtags
- @replies

Poster

-

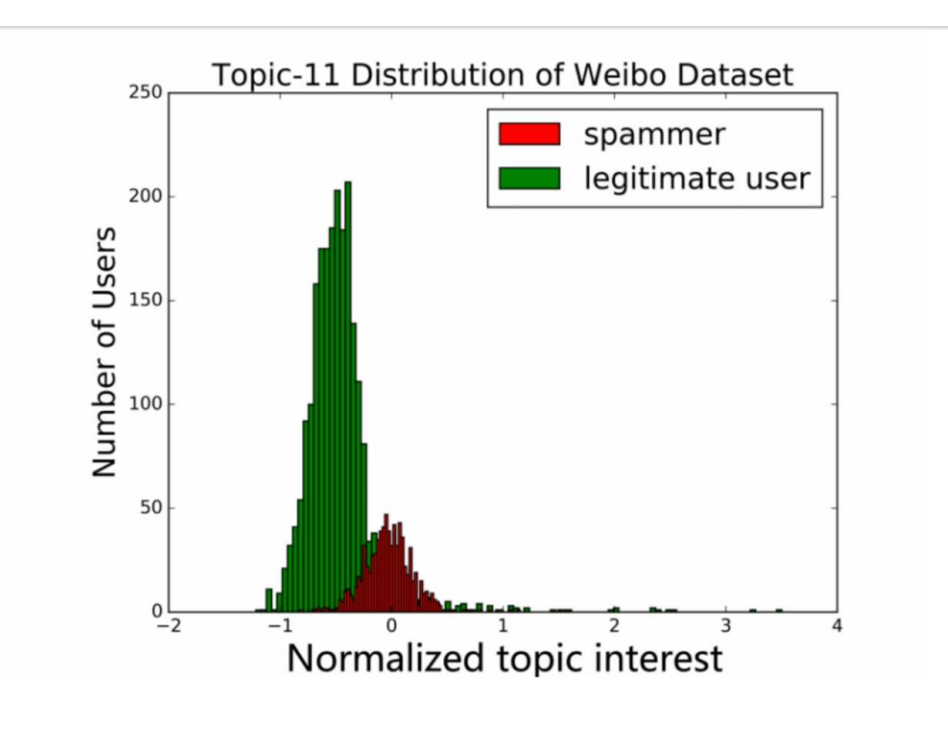


Social Network

-

(Yardi et al, '09; Grier et al, '10; Wang, '10; Jin et al, '11)

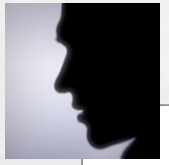
BOT DETECTION BASED ON TOPIC MODELING



- "LEGITIMATE" USERS MAINLY SHARE ON LIMITED TOPICS OF INTEREST.
- "FAKE ACCOUNTS" POST ON A WIDE RANGE OF TOPICS, LIKELY BECAUSE THEY RANDOMLY COPY & RT OTHER USERS.

(LIU ET.AL., 2016)

DETECTING BOTS…



Photo

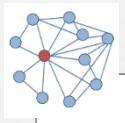
- Color & Edge histograms
- Color & Edge Directivity Descriptor (CEDD)
- Image Similarity

Message

- Sensitive words
- URL
- Duplicates
- #hashtags
- @replies

Poster

- Username
- Engagement level
- Creation date

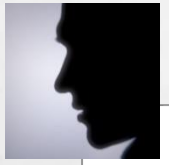


Social Network

-

(Yardi et al, '09; Grier et al, '10; Wang, '10; Jin et al, '11)

DETECTING BOTS…



Photo

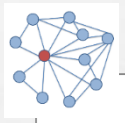
- Color & Edge histograms
- Color & Edge Directivity Descriptor (CEDD)
- Image Similarity

Message

- Sensitive words
- URL
- Duplicates
- #hashtags
- @replies

Poster

- Username
- Engagement level
- Creation date



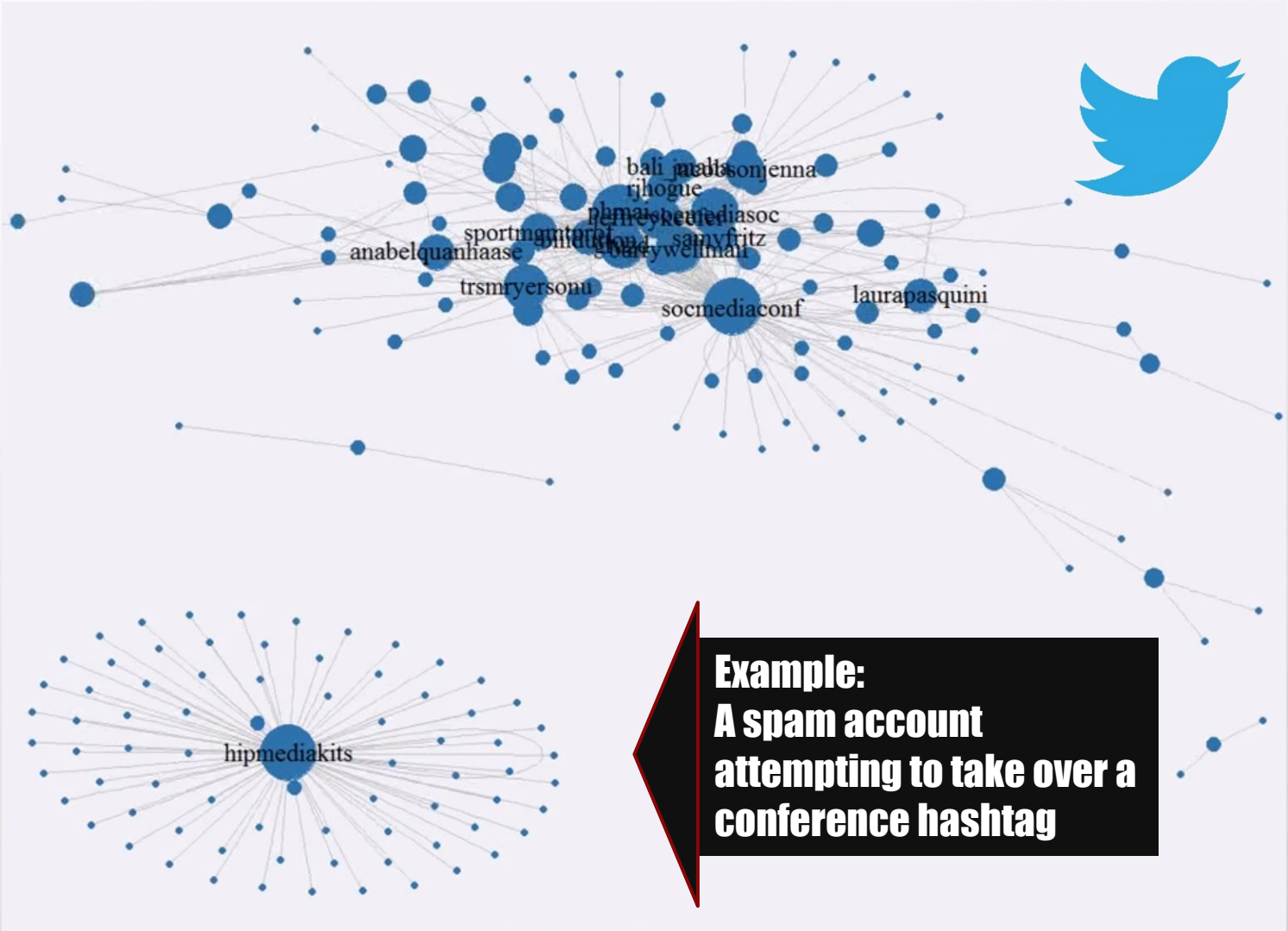
Social Network

- # Friends
- # Following
- In/out degree centrality
- Clustering

(Yardi et al, '09; Grier et al, '10; Wang, '10; Jin et al, '11)

Detecting Bots…

Using Social Network Analysis



Example:
A spam account attempting to take over a conference hashtag

TWITTER BOT DETECTION UTILIZING MULTIPLE MACHINE LEARNING ALGORITHMS

BY MRP STUDENT JIEFAN LUO

SUPERVISOR: ANATOLIY GRUZD

Features	Accuracy (Cross-validation)
followers_count	0.6736118
followers_count, friends_count	0.6864886
followers_count, friends_count, favourites_count	0.7188525
followers_count, friends_count, favourites_count, statuses_count	0.7640931
followers_count, friends_count, favourites_count, statuses_count, created_at	0.8691169
followers_count, friends_count, favourites_count, statuses_count, created_at, description_en	0.8870175

Twitter Bot Detection in Communalytic with Botometer API

What is Botometer API?

How to use Twitter Bot Analysis in Communalytic

How to get a Botometer API Key

What is Botometer API?

- ➔ Botometer API is a tool created by the Observatory of Social Media at Indiana University
- ➔ Checks the activity of a Twitter user's account and detects the potential use of automation (shows bot like activity)
- ➔ Differentiate between different types of bot behaviours

https://botometer.osome.iu.edu/

What Types of Bot behaviours can it detect?

6 main bot account behaviours Botometer detects:

- **Echo-chamber -** “accounts that engage in follow back groups and share and delete politic content in high volume”
- **Fake follower -** “bots purchased to increase follower counts”
- **Financial -** “bots that post using cashtags”
- **Self declared -** “bots from botwiki.org”
- **Spammer -** “accounts labeled as spambots from several datasets”
- **Other -** “miscellaneous other bots obtained from manual annotation, user feedback, etc.”

Definitions taken from Botometer’s FAQ page: https://botometer.osome.iu.edu/faq#which-score

How does Botometer differentiate these behaviours?

- ➔ Botometer uses different machine learning models to determine whether the pattern of behaviour is more like a human or a bot.

- ➔ It generates **probability scores** for 6 types of bot-like behaviours. (See slide 3)

- ➔ It then compare “models trained on different kinds of bots and on human accounts” to calculate an **Overall** probability score that corresponds to the model(s) with which it has the highest confidence in.

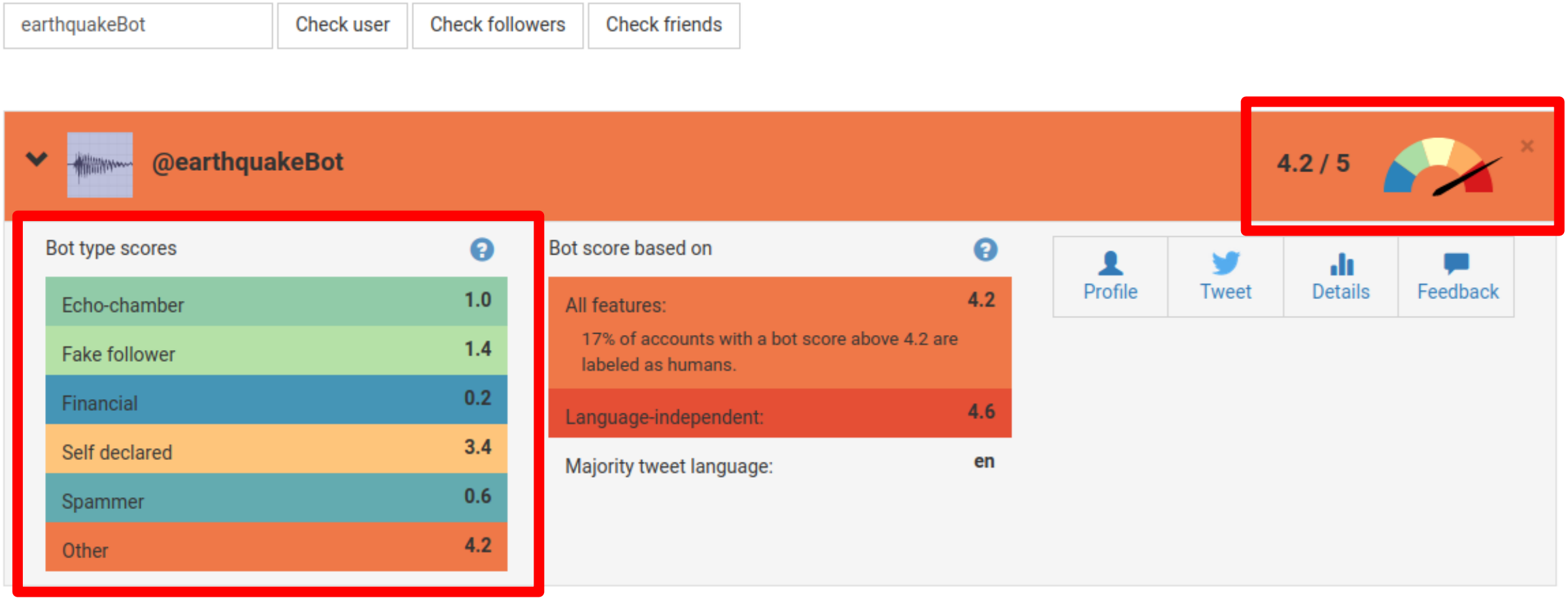
Botometer's Web Interface

https://botometer.osome.iu.edu/

Let's take a look at a self proclaimed bot to see how it is scored by Botometer...

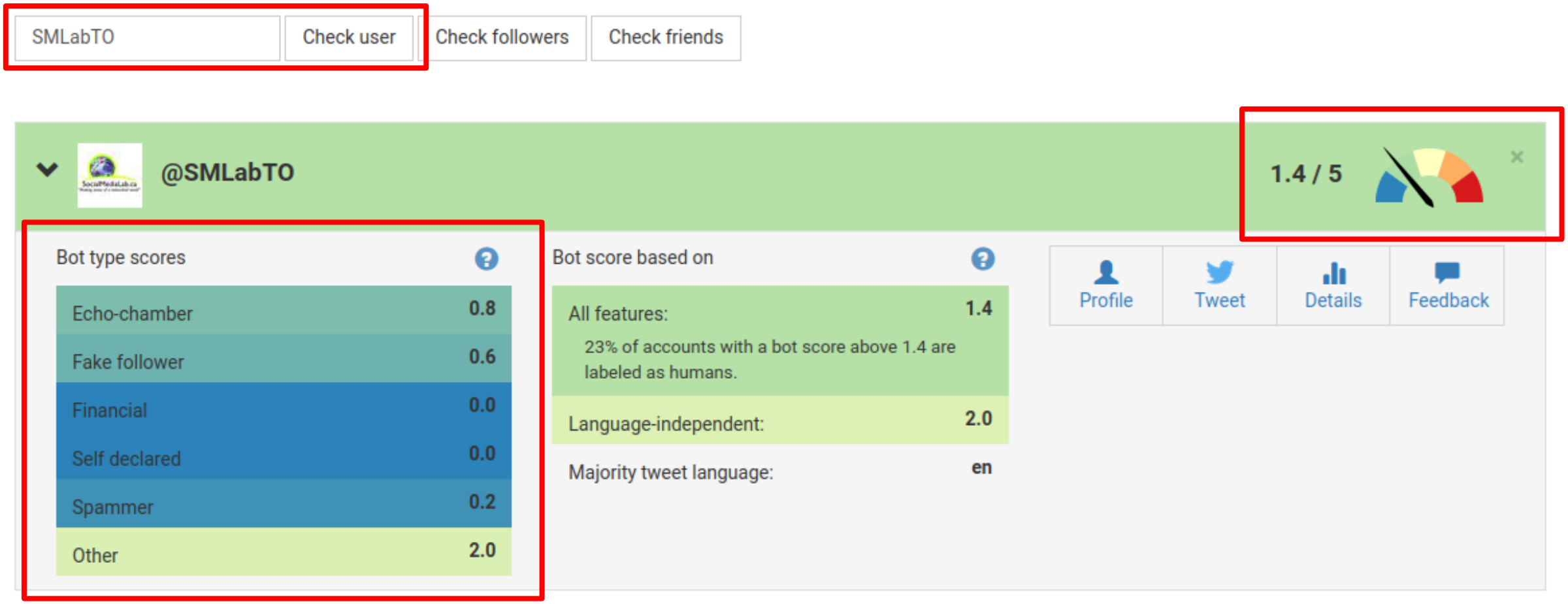
Botometer's Web Interface

https://botometer.osome.iu.edu/



Botometer's Web Interface

https://botometer.osome.iu.edu/



Outline

What is Botometer API?

How to use Twitter Bot Analysis in Communalytic

How to get a Botometer API Key

communalytic (EDU)

FAQ Tutorials Publications My Datasets

My Datasets

Server Time: July 19, 2021 17:31 UTC

Collect data from

Reddit (LIVE) Twitter Thread CrowdTangle CSV File

My Datasets Shared With Me

Total records in the account: 469 / 30000

2%

Dataset Name	Number of Records	Toxicity Analysis	Network Analysis	Download Dataset	Collaborators	Delete
Justin Trudeau - Nelson Mandela Tweet twitter : https://twitter.com/JustinTrudeau/status/1416855674189930496	Records: 469 Check for Twitter bots					

communalytic (EDU)

FAQ Tutorials Publications My Datasets

Back to My Datasets

Overview

Dataset Name: **Justin Trudeau - Nelson Mandela Tweet**

Platform: **twitter**

Search Criteria: **https://twitter.com/JustinTrudeau/status/1416855674189930496**

Collection started: **2021-07-19 17:30**

Collection ended: **2021-07-19 17:30**

Records: **469**

Posts Per Day

Word Cloud

Emoji Cloud

Top Ten Posters

Toxicity Analysis

Twitter Bot Analysis

Twitter Bot Analysis

The Twitter Bot Analysis Module is designed to detect potential use of automation. It analyzes accounts in your dataset and generates the probability scores for a variety of different types of automated and/or fraudulent activities (from 0=not likely to 1=highly likely). Once the analysis is complete, the probability scores for each account will be added to your dataset.

Analyze Top 100 Accounts — This option will only analyze 100 most active posters in your dataset.

OR

Analyze All (357) Accounts — This option will analyze all accounts in your dataset, which will take **~1 days** to complete using the Botometer Basic plan (500 requests per day). Also since there are >100 unique accounts in your dataset, to use this option, you'll need to request and add your own Botometer API key to your Communalytic account under My Profile.

- The Twitter Bot Analysis Module is based on a machine learning API called Botometer, a project by the Observatory on Social Media (OSoMe) at the University of Indiana.
- If you already have an API key for Botometer, it can be added to your Communlaytic account under My Profile; otherwise, you can requests one here.
- Botometer generates probability scores for a variety of automated and/or fraudulent activities including Echo Chamber, Fake Follower, Financial, Self Declare, Spammer and Others.
- Twitter accounts that have been deleted, banned or privated on Twitter will not be analyzed by Botometer.
- To ensure that you are not flagged as a security/excessive user of the Botometer API, Communalytic will only use up to 80% of your daily API limit. When the 80% threshold is reached, the analysis will pause for a day.

Twitter Bot Analysis

The Twitter Bot Analysis Module is designed to detect potential use of automation. It analyzes accounts in your dataset and generates the probability scores for a variety of different types of automated and/or fraudulent activities (from 0=not likely to 1=highly likely). Once the analysis is complete, the probability scores for each account will be added to your dataset.

Analysis In Progress ...

(You may close this window and visit later to download the bot analysis results)

27%

Progress: 27 / 100 Users Analyzed

Check progress in 30 sec

Cancel Analysis

communalytic (EDU)

FAQ Tutorials Publications My Datasets

Back to My Datasets

Overview

Dataset Name: **Justin Trudeau - Nelson Mandela Tweet**

Platform: **twitter**

Search Criteria: **https://twitter.com/JustinTrudeau/status/1416855674189930496**

Collection started: **2021-07-19 17:30**

Collection ended: **2021-07-19 17:30**

Records: **469**

Posts Per Day

Word Cloud

Emoji Cloud

Top Ten Posters

Toxicity Analysis

Twitter Bot Analysis

Twitter Bot Analysis

The Twitter Bot Analysis Module is designed to detect potential use of automation. It analyzes accounts in your dataset and generates the probability scores for a variety of different types of automated and/or fraudulent activities .

100 out of **100** accounts from your dataset were analyzed by Botometer API.

Reset Analysis

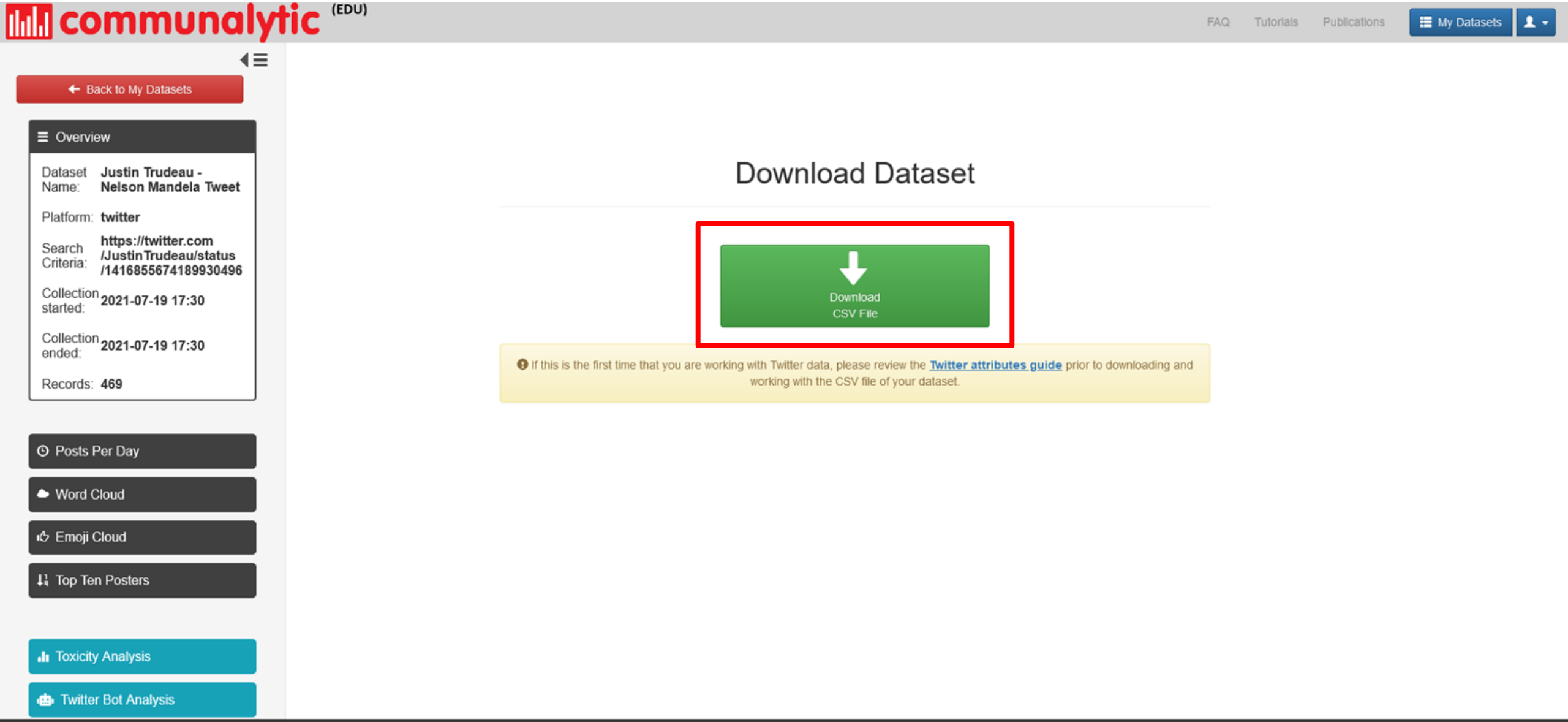
Download a CSV of your dataset with the Botometer probability scores for the various types of automated and/or fraudulent activities (from 0=not likely to 1=highly likely).

Username	Overall	Echo Chamber	Fake Follower	Financial	Self Declared	Spammer	Other
1979HAB	0.04	0.14	0.03	0.00	0.00	0.03	
7otnauh	0.34	0.50	0.16	0.01	0.00	0.01	
AdbudMan	0.08	0.05	0.14	0.07	0.01	0.09	
Ahmad86815198	0.78	0.14	0.28	0.06	0.41	0.07	
Al_Mahbook	0.30	0.12	0.30	0.06	0.01	0.14	
Ange04920249	0.40	0.18	0.32	0.01	0.23	0.33	
Angie_CRK	0.71	0.09	0.52	0.05	0.71	0.25	
BBKing94554247	0.40	0.16	0.15	0.00	0.03	0.16	
BeanPlayer	0.06	0.08	0.02	0.00	0.00	0.03	

Bot type scores

Echo-chamber	1.0
Fake follower	1.4
Financial	0.2
Self declared	3.4
Spammer	0.6
Other	4.2

© 2021 Communalytic | FAQ | Terms of service | Privacy Policy


communalytic (EDU)
FAQ
Tutorials
Publications
My Datasets
Back to My Datasets
Overview
Dataset Name: Justin Trudeau - Nelson Mandela Tweet
Platform: twitter
Search Criteria: https://twitter.com/JustinTrudeau/status/1416855674189930496
Collection started: 2021-07-19 17:30
Collection ended: 2021-07-19 17:30
Records: 469
Posts Per Day
Word Cloud
Emoji Cloud
Top Ten Posters
Toxicity Analysis
Twitter Bot Analysis
Download Dataset
Download CSV File
If this is the first time that you are working with Twitter data, please review the Twitter attributes guide prior to downloading and working with the CSV file of your dataset.
© 2021 Communalytic | FAQ | Terms of service | Privacy Policy

Viewing Botometer Scores in Excel

botometer_ checked_on	user_status	botometer_ overall	botometer_ echo_chamber	botometer_ fake_follower	botometer_ self_declared	botometer_ spammer	botometer_ other

text	name	usernar	botometer_checked_on	user_status	botometer_overall	botometer_echo_chamber	botometer_fake_follo	botometer_financial	botometer_self_decla	botometer_spamr	bot
2377	Merry Szymc	MerrySzyr	7/19/2021 20:52	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@Cherish	Patriotgirl45	Patriotgirl	7/19/2021 20:52	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@MerrySa	Jennifer Hunt	jhunter23	7/19/2021 20:52	ACCOUNT_NOT_FOUND	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ngzamo	Patriotgirl45	Patriotgirl	7/19/2021 20:52	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@Reyn84	lisa hannah #I	lisaalma	7/19/2021 21:07	ACCOUNT_AVAILABLE	0.11	0.27	0.25	0.02	0.05	0.01	
@monica	AdotKitz	AdotKitz	7/19/2021 20:37	ACCOUNT_AVAILABLE	0.28	0.09	0.2	0.06	0.23	0.1	
@vlogs_k	Trisha Layma	Cherish19	7/19/2021 20:52	ACCOUNT_NOT_FOUND	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@jhunter	Merry Szymc	MerrySzyr	7/19/2021 20:52	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@Madam	Robin Burris	RobinBurr	7/19/2021 20:37	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@Gnome	Sheri Cotner	CotnerShe	7/19/2021 21:22	ACCOUNT_AVAILABLE	0.09	0.13	0.21	0.02	0.05	0.02	
@acablart	Robin Burris	RobinBurr	7/19/2021 20:37	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@brotha_	Sigma Owen :	TryCrying	7/19/2021 20:52	ACCOUNT_AVAILABLE	0.01	0.12	0.04	0	0.03	0	
@MerrySa	Jennifer Hunt	jhunter23	7/19/2021 20:52	ACCOUNT_NOT_FOUND	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@POTUS	Eddard Snark	WokeMay	7/19/2021 20:37	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@Segidrea	Dracaris	Akylle7	7/19/2021 20:51	ACCOUNT_AVAILABLE	0.27	0.15	0.21	0	0.05	0.22	
@Mozelle	Chris.janda	Chrisjanda	7/19/2021 20:36	ACCOUNT_AVAILABLE	0.81	0.04	0.35	0.08	0.76	0.18	
@CotnerS	Gnome Stonk	GnomeStc	7/19/2021 20:37	ACCOUNT_AVAILABLE	0.07	0.04	0.18	0.03	0.11	0.02	
@JamesM	Nancy Hart	nhart543	7/19/2021 20:52	ACCOUNT_AVAILABLE	0.26	0.51	0.24	0	0.23	0.02	
@crosby_	MJPalau	PalauMj	7/19/2021 20:36	ACCOUNT_AVAILABLE	0.93	0.93	0.22	0	0.01	0.03	
2377	Merry Szymc	MerrySzyr	7/19/2021 20:52	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
It's great	Phillip- #Bider	Phillip202	7/19/2021 20:52	ACCOUNT_AVAILABLE	0.95	0.95	0.48	0	0.03	0	
@WokeM	Mozelle Hedg	MozelleHe	7/19/2021 21:07	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@vixereii	Amanda81	AmandaN	7/19/2021 21:22	ACCOUNT_AVAILABLE	0.17	0.01	0.12	0.01	0.07	0.13	
@MerrySa	Jennifer Hunt	jhunter23	7/19/2021 20:52	ACCOUNT_NOT_FOUND	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2377	Merry Szymc	MerrySzyr	7/19/2021 20:52	ACCOUNT_SUSPENDED	N/A	N/A	N/A	N/A	N/A	N/A	N/A
@mlissa1	Vivian Kay	wildviv	7/19/2021 21:07	ACCOUNT_AVAILABLE	0.66	0.02	0.32	0.11	0.45	0.05	
@vixereii	Amanda81	AmandaN	7/19/2021 21:22	ACCOUNT_AVAILABLE	0.17	0.01	0.12	0.01	0.07	0.13	
@Sarafina	Fillers Diller	krokuskris	7/19/2021 20:52	ACCOUNT_AVAILABLE	0.04	0.37	0.1	0.02	0.01	0.07	

Outline

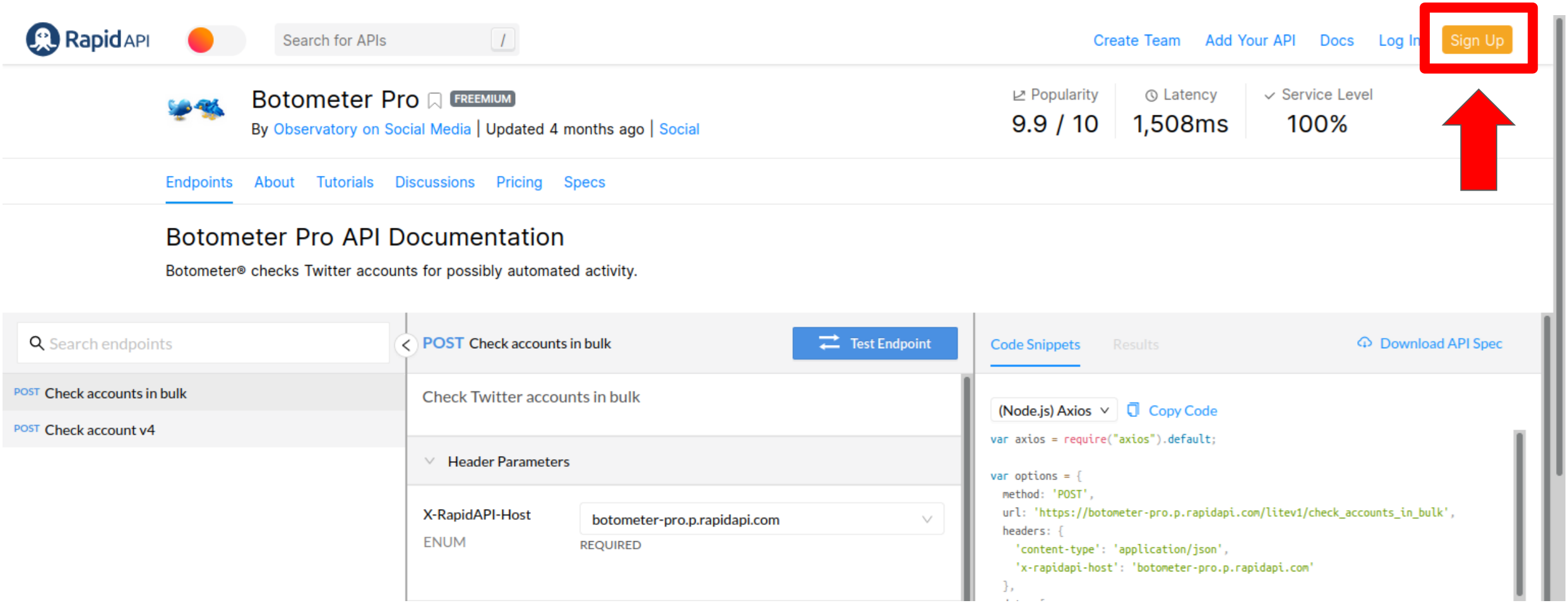
What is Botometer API?

How to use Twitter Bot Analysis in Communalytic

How to get a Botometer API Key

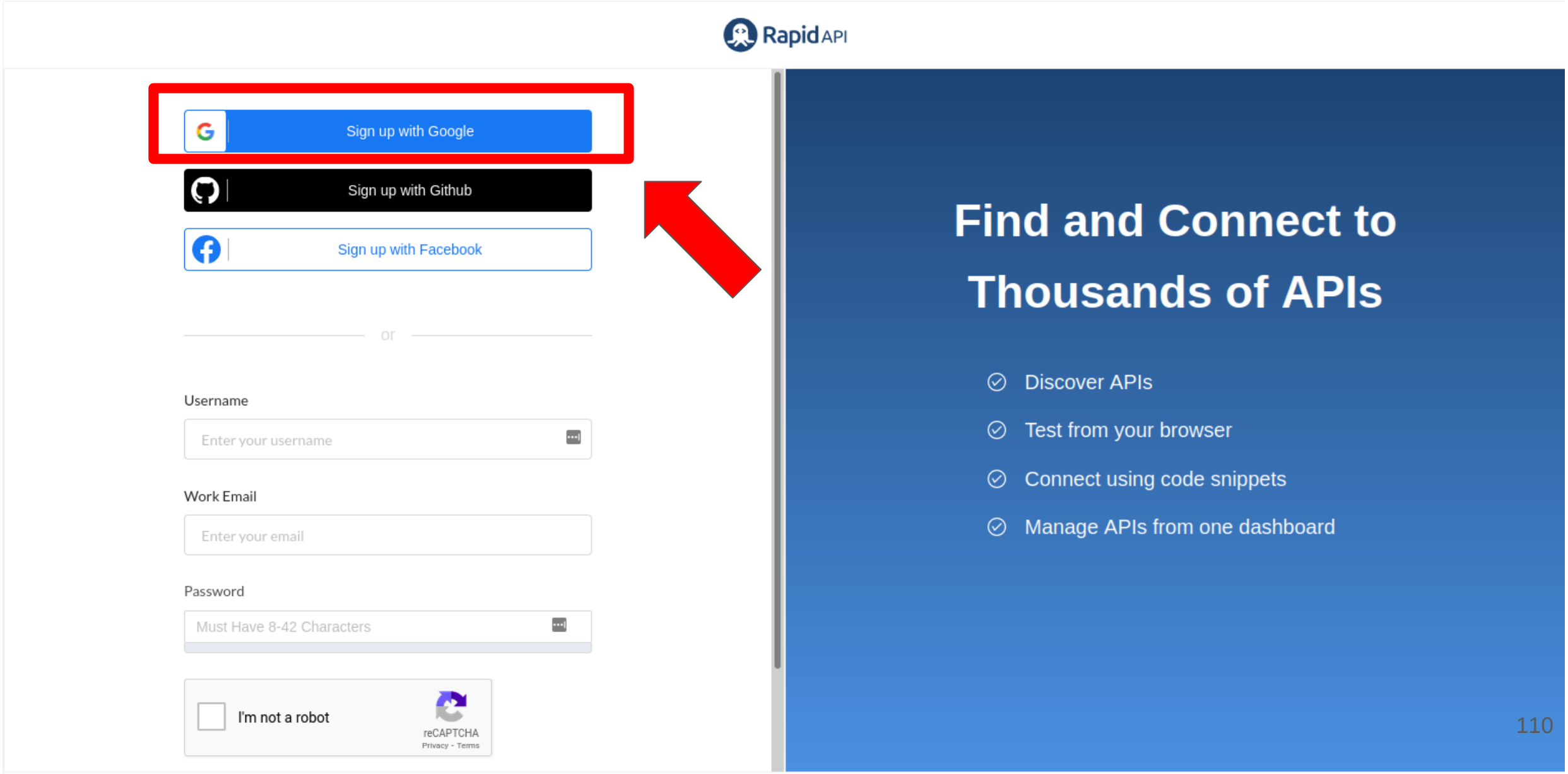
How to request your own Botometer API key

https://rapidapi.com/OSoMe/api/botometer-pro/



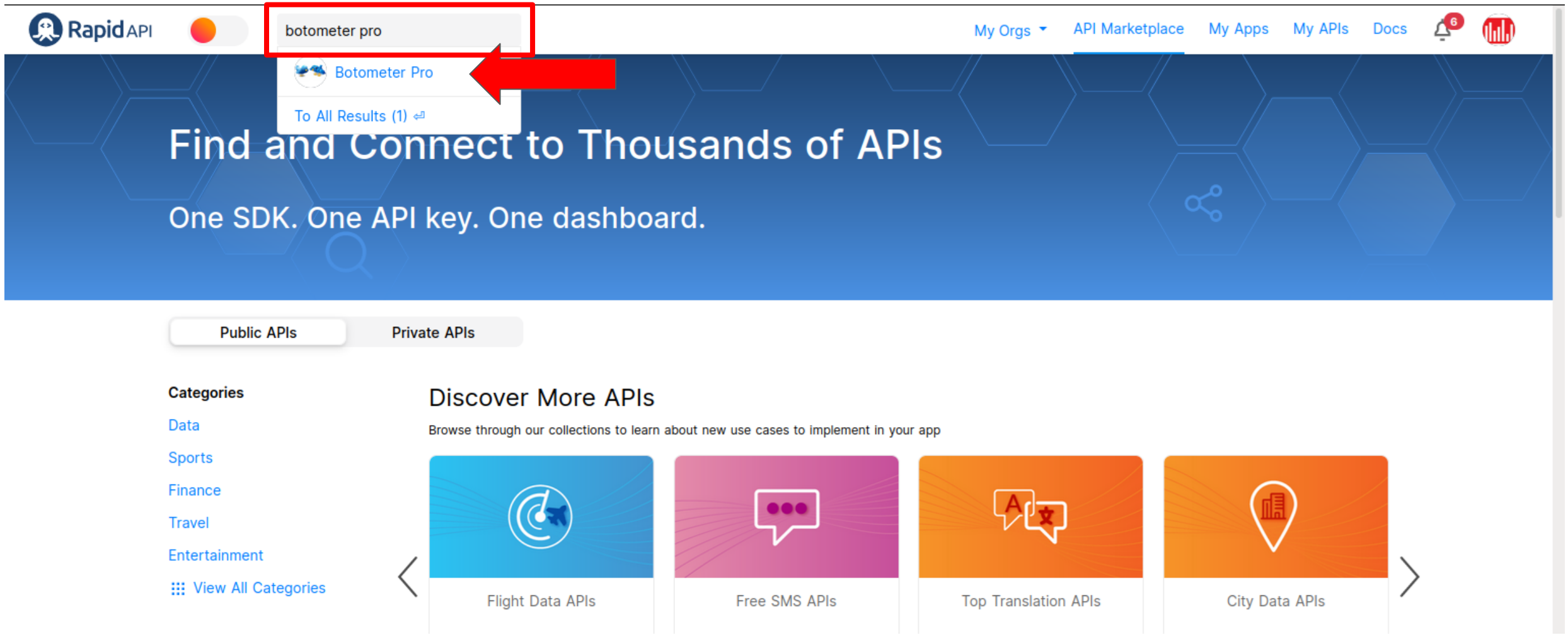
Create a new RapidAPI account

https://rapidapi.com/

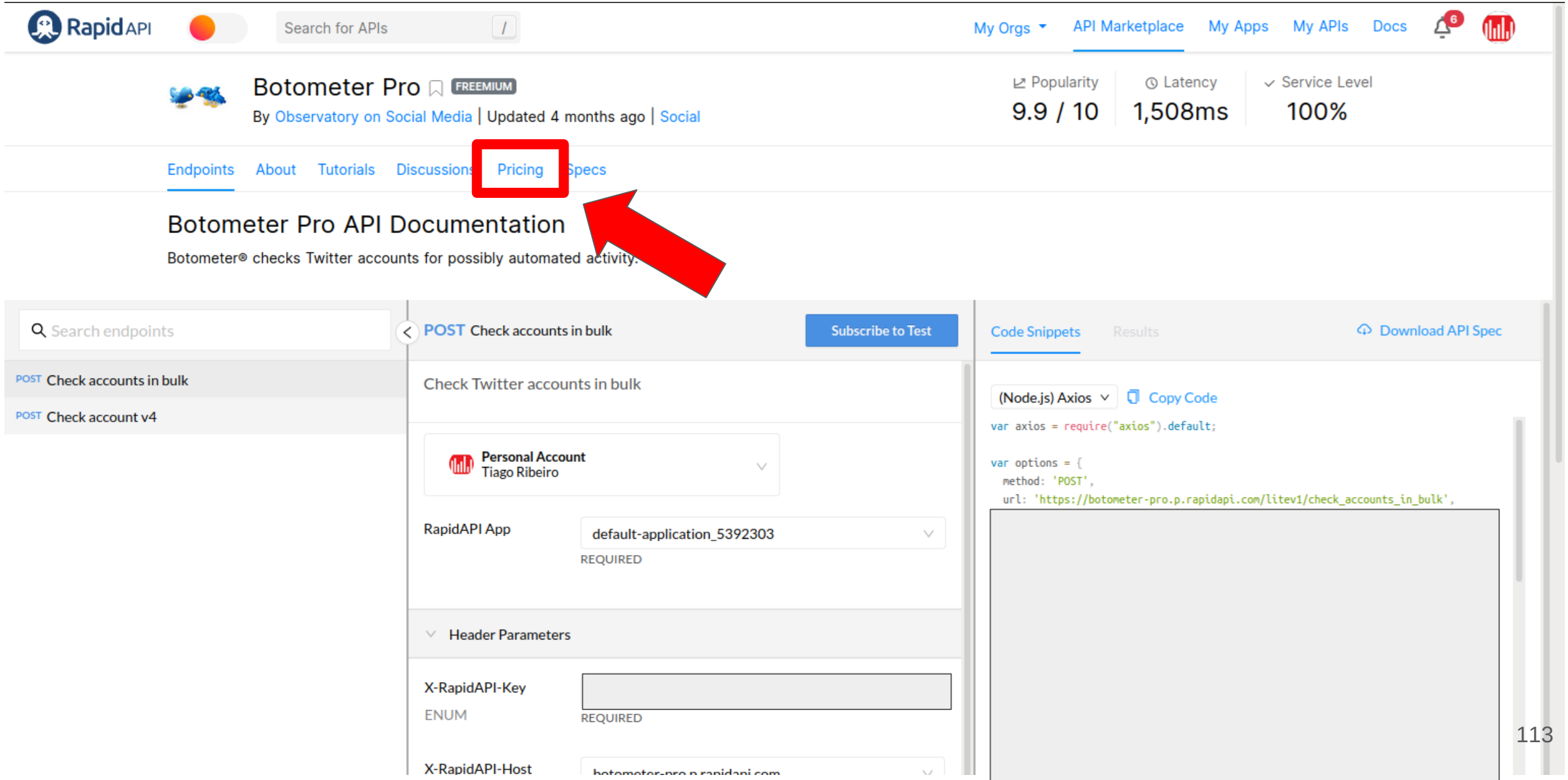


Search for ‘Botometer’ API on the RapidAPI site

https://rapidapi.com/



View & Select ‘Botometer’ API Plan



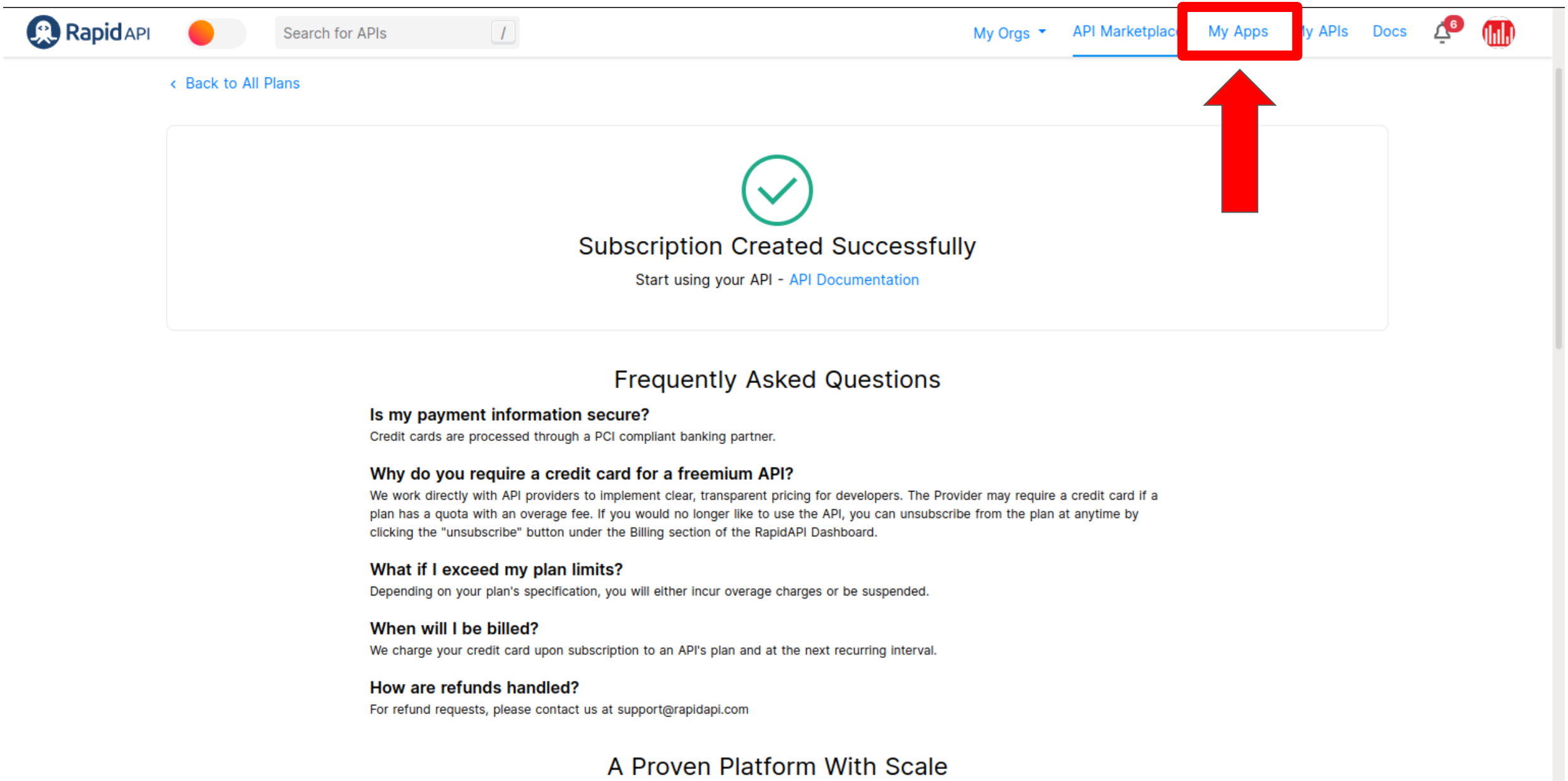
View & Select ‘Botometer’ API Plan

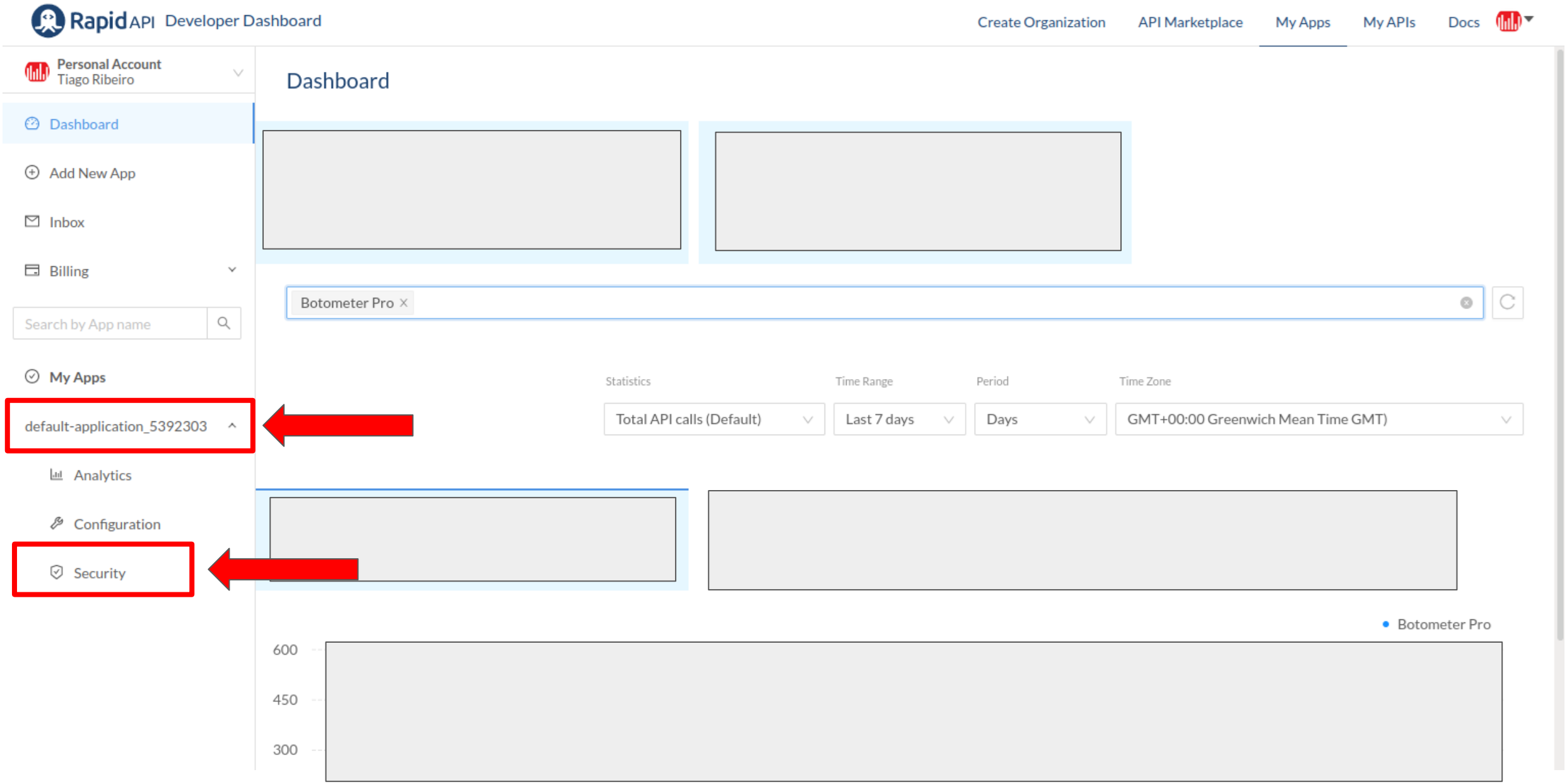
Choose the Right Plan For You

RapidAPI partners directly with API providers to give you no-fuss, transparent pricing. Find a plan that best matches the scale you need for your application.

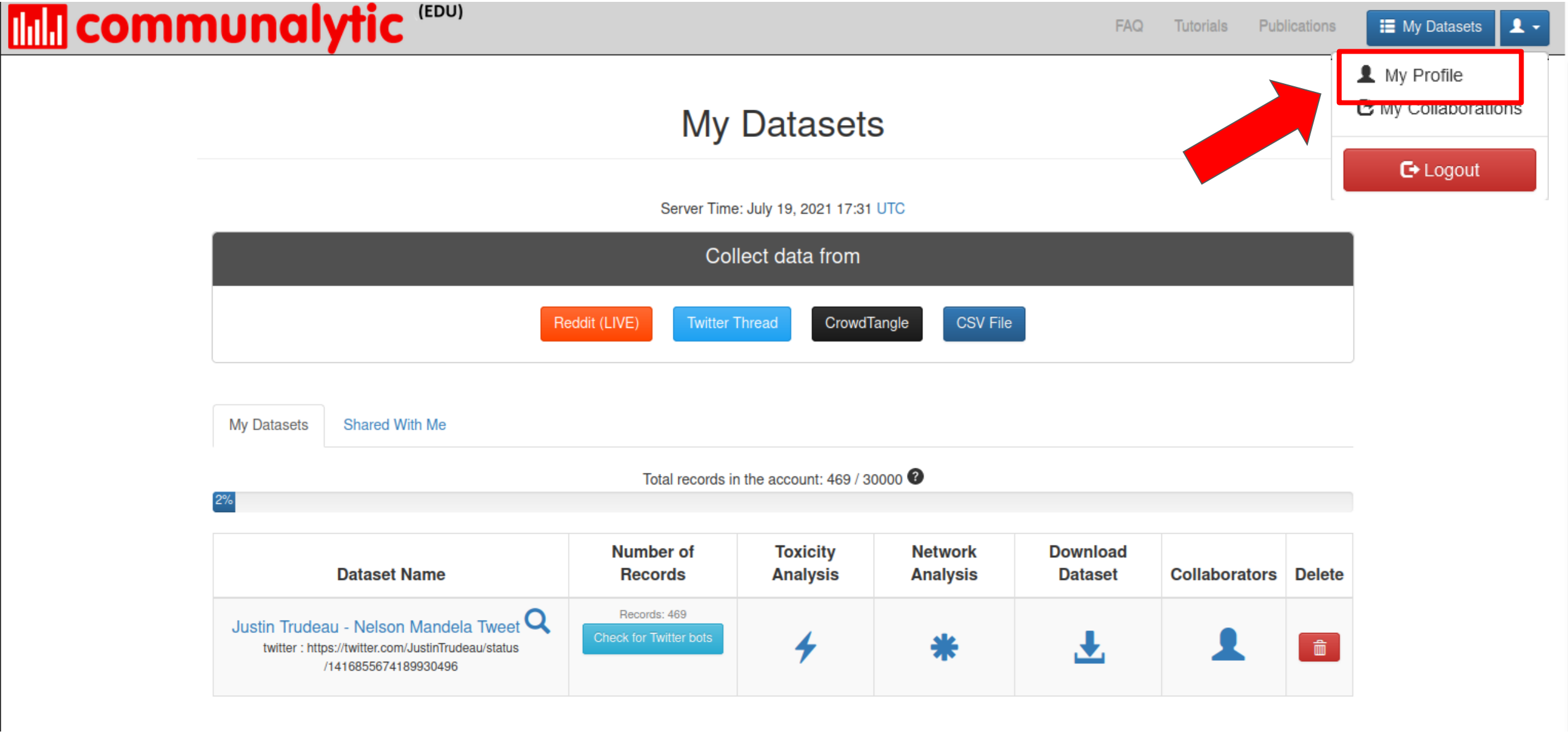
Tiago Ribeiro

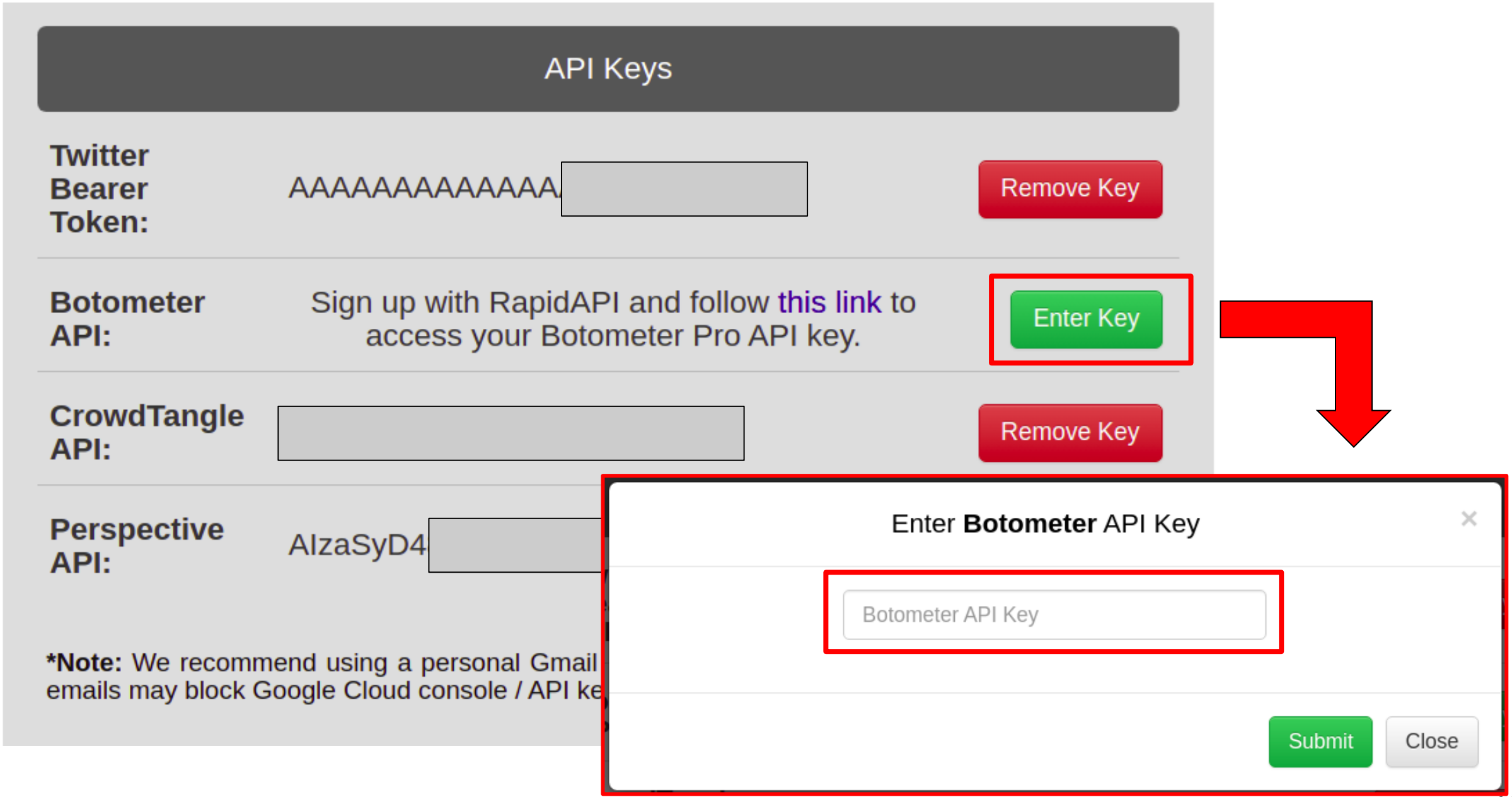
Objects	Basic **$0.00 / mo** Subscribe	Recommended Pro **$0.00 / mo** Subscribe	Ultra **$50.00 / mo** Subscribe
BotometerLite			**200 / day** + $0.01 each other
Check account	**500 / day** Hard Limit	**2,000 / day** + $0.001 each other	**17,280 / day** + $0.001 each other
Features			
BotometerLite ⓘ	✗	✗	✓
Check account ⓘ	✓	✓	✓
Rate Limit	one request per second		

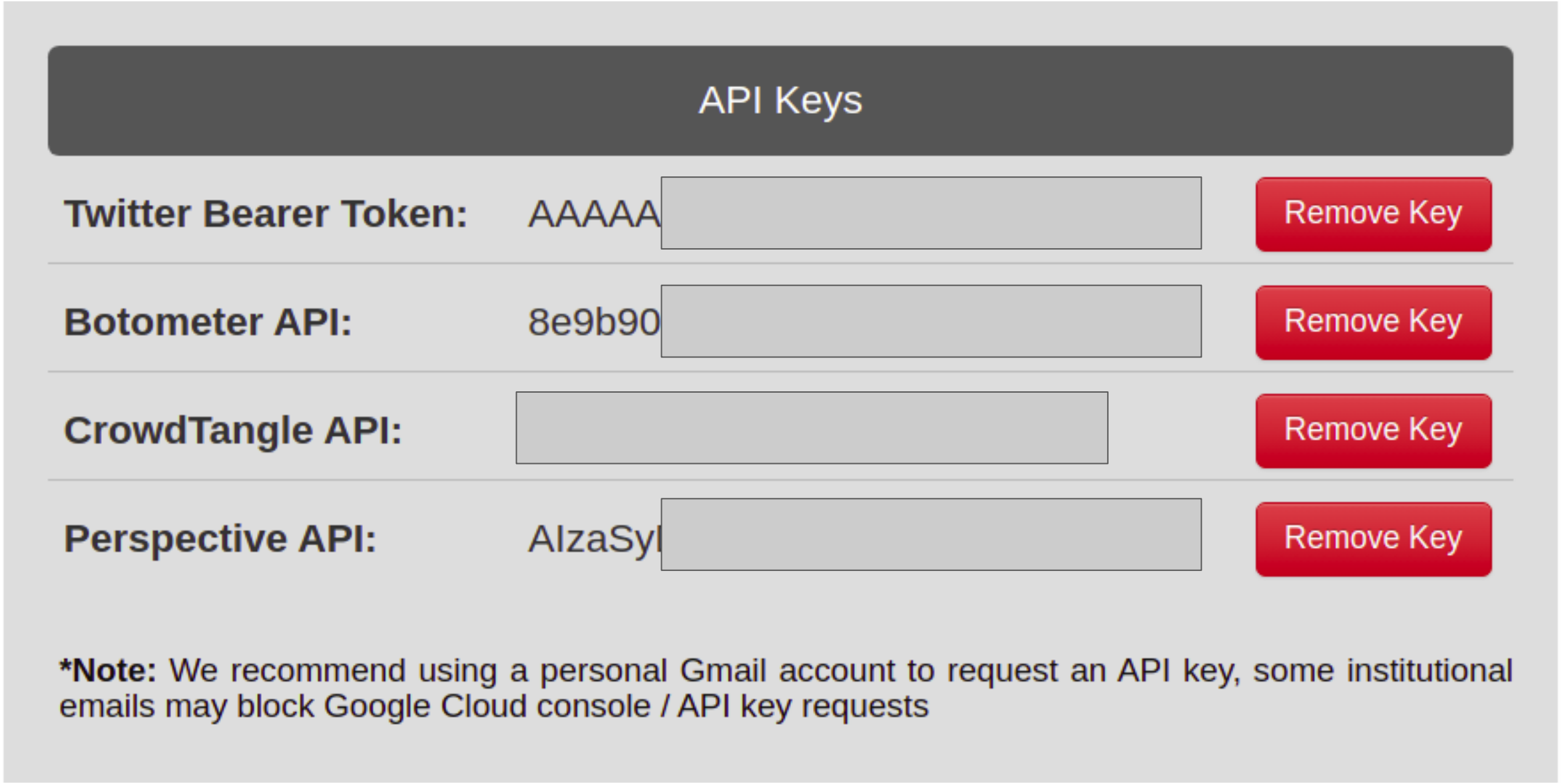

Rapid API
Search for APIs
/
My Orgs
API Marketplac
My Apps
y APIs
Docs
6
Back to All Plans
Subscription Created Successfully
Start using your API - API Documentation
Frequently Asked Questions
Is my payment information secure?
Credit cards are processed through a PCI compliant banking partner.
Why do you require a credit card for a freemium API?
We work directly with API providers to implement clear, transparent pricing for developers. The Provider may require a credit card if a plan has a quota with an overage fee. If you would no longer like to use the API, you can unsubscribe from the plan at anytime by clicking the "unsubscribe" button under the Billing section of the RapidAPI Dashboard.
What if I exceed my plan limits?
Depending on your plan's specification, you will either incur overage charges or be suspended.
When will I be billed?
We charge your credit card upon subscription to an API's plan and at the next recurring interval.
How are refunds handled?
For refund requests, please contact us at support@rapidapi.com
A Proven Platform With Scale

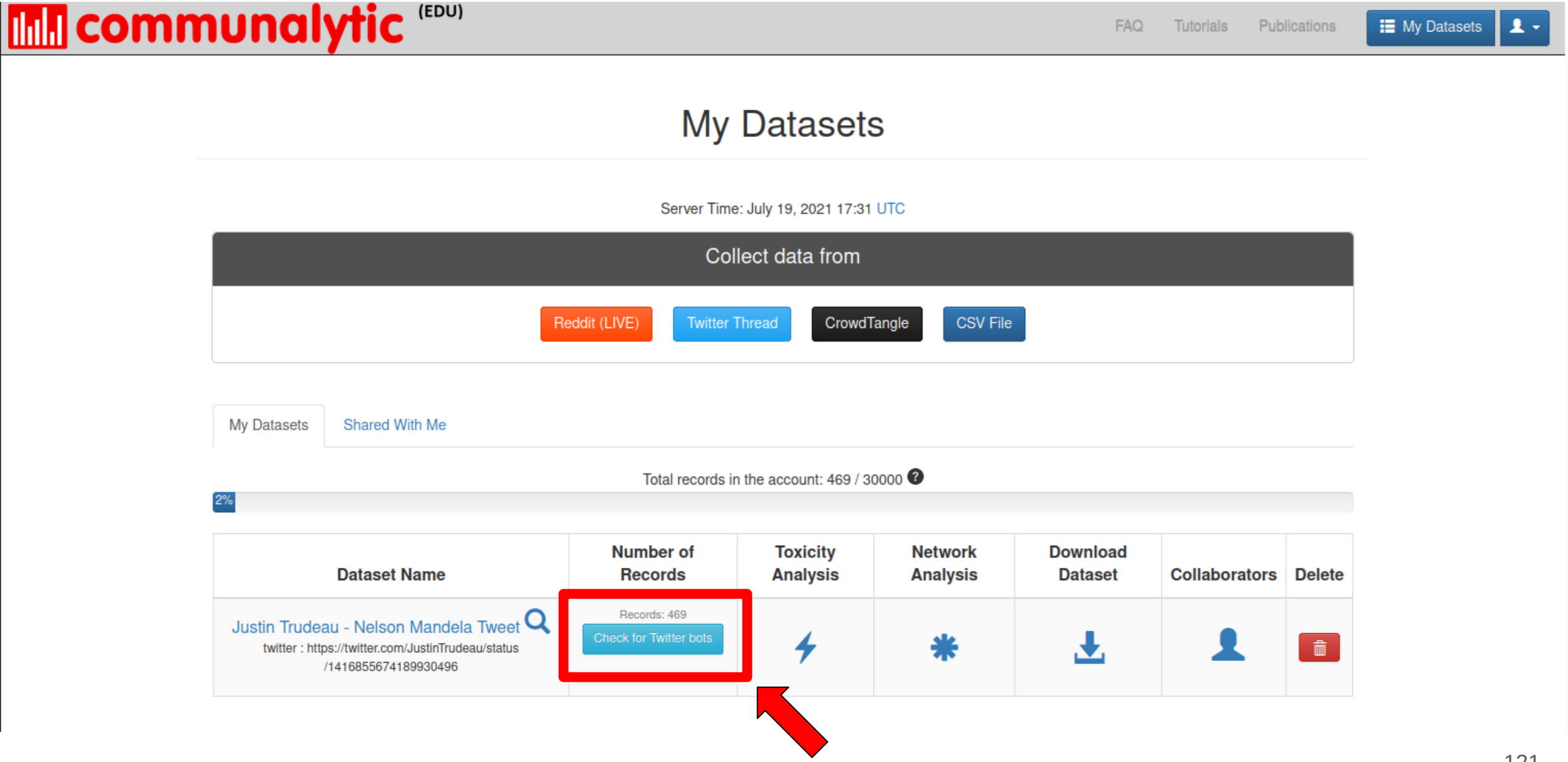

RapidAPI Developer Dashboard
Create Organization
API Marketplace
My Apps
My APIs
Docs
Personal Account
Tiago Ribeiro
Dashboard
Add New App
Inbox
Billing
Search by App name
My Apps
default-application_5392303
Analytics
Configuration
Security
Dashboard
Botometer Pro
Statistics
Time Range
Period
Time Zone
Total API calls (Default)
Last 7 days
Days
GMT+00:00 Greenwich Mean Time GMT)
Botometer Pro
600
450
300

default-application_5392303 - Security
Add New Key
Name
Key
Created On
Application Key
Edit


communalytic (EDU)
FAQ
Tutorials
Publications
My Datasets
My Profile
My Collaborations
Logout
My Datasets
Server Time: July 19, 2021 17:31 UTC
Collect data from
Reddit (LIVE)
Twitter Thread
CrowdTangle
CSV File
My Datasets
Shared With Me
Total records in the account: 469 / 30000
2%
Dataset Name
Number of Records
Toxicity Analysis
Network Analysis
Download Dataset
Collaborators
Delete
Justin Trudeau - Nelson Mandela Tweet
twitter : https://twitter.com/JustinTrudeau/status/1416855674189930496
Records: 469
Check for Twitter bots


API Keys
Twitter Bearer Token:
AAAAAAAAAAAAA
Remove Key
Botometer API:
Sign up with RapidAPI and follow this link to access your Botometer Pro API key.
Enter Key
CrowdTangle API:
Remove Key
Perspective API:
AIzaSyD4
*Note: We recommend using a personal Gmail
emails may block Google Cloud console / API ke
Enter Botometer API Key
Botometer API Key
Submit
Close


API Keys
Twitter Bearer Token:
AAAAA
Remove Key
Botometer API:
8e9b90
Remove Key
CrowdTangle API:
Remove Key
Perspective API:
AIzaSyI
Remove Key
*Note: We recommend using a personal Gmail account to request an API key, some institutional emails may block Google Cloud console / API key requests


communalytic (EDU)
FAQ
Tutorials
Publications
My Datasets
My Datasets
Server Time: July 19, 2021 17:31 UTC
Collect data from
Reddit (LIVE)
Twitter Thread
CrowdTangle
CSV File
My Datasets
Shared With Me
Total records in the account: 469 / 30000
2%
Dataset Name
Number of Records
Toxicity Analysis
Network Analysis
Download Dataset
Collaborators
Delete
Justin Trudeau - Nelson Mandela Tweet
twitter : https://twitter.com/JustinTrudeau/status/1416855674189930496
Records: 469
Check for Twitter bots

Back to My Datasets

Overview

Dataset Name: **Justin Trudeau - Nelson Mandela Tweet**

Platform: **twitter**

Search Criteria: **https://twitter.com/JustinTrudeau/status/1416855674189930496**

Collection started: **2021-07-19 17:30**

Collection ended: **2021-07-19 17:30**

Records: **469**

Posts Per Day

Word Cloud

Emoji Cloud

Top Ten Posters

Toxicity Analysis

Twitter Bot Analysis

Twitter Bot Analysis

The Twitter Bot Analysis Module is designed to detect potential use of automation. It analyzes accounts in your dataset and generates the probability scores for a variety of different types of automated and/or fraudulent activities .

100 out of **100** accounts from your dataset were analyzed by Botometer API.

Reset Analysis

Download a CSV of your dataset with the Botometer probability scores for the various types of automated and/or fraudulent activities (from 0=not likely to 1=highly likely).

Username	Overall	Echo Chamber	Fake Follower	Financial	Self Declared	Spammer	Other
1979HAB	0.04	0.14	0.03	0.00	0.00	0.03	0.18
7otnauh	0.34	0.50	0.16	0.01	0.00	0.01	0.40
AdbudMan	0.08	0.05	0.14	0.07	0.01	0.09	0.17
Ahmad86815198	0.78	0.14	0.28	0.06	0.41	0.07	0.78
Al_Mahbook	0.30	0.12	0.30	0.06	0.01	0.14	0.31
Ange04920249	0.40	0.18	0.32	0.01	0.23	0.33	0.39
Angie_CRK	0.71	0.09	0.52	0.05	0.71	0.25	0.61
BBKing94554247	0.40	0.16	0.15	0.00	0.03	0.16	0.28
BeanPlayer	0.06	0.08	0.02	0.00	0.00	0.03	0.27

© 2021 Communalytic | FAQ | Terms of service | Privacy Policy

communalytic (EDU)

FAQ Tutorials Publications My Datasets

Back to My Datasets

Overview

Dataset Name: **Justin Trudeau - Nelson Mandela Tweet**

Platform: **twitter**

Search Criteria: **https://twitter.com/JustinTrudeau/status/1416855674189930496**

Collection started: **2021-07-19 17:30**

Collection ended: **2021-07-19 17:30**

Records: **469**

Posts Per Day

Word Cloud

Emoji Cloud

Top Ten Posters

Toxicity Analysis

Twitter Bot Analysis

Twitter Bot Analysis

The Twitter Bot Analysis Module is designed to detect potential use of automation. It analyzes accounts in your dataset and generates the probability scores for a variety of different types of automated and/or fraudulent activities (from 0=not likely to 1=highly likely). Once the analysis is complete, the probability scores for each account will be added to your dataset.

Analyze Top 100 Accounts — This option will only analyze 100 most active posters in your dataset.

OR

Analyze All (357) Accounts — This option will analyze all accounts in your dataset, which will take **~1 days** to complete using the Botometer Basic plan (500 requests per day). Also since there are >100 unique accounts in your dataset, to use this option, you'll need to request and add your own Botometer API key to your Communalytic account under My Profile.

- The Twitter Bot Analysis Module is based on a machine learning API called Botometer, a project by the Observatory on Social Media (OSoMe) at the University of Indiana.
- If you already have an API key for Botometer, it can be added to your Communlaytic account under My Profile; otherwise, you can requests one here.
- Botometer generates probability scores for a variety of automated and/or fraudulent activities including Echo Chamber, Fake Follower, Financial, Self Declare, Spammer and Others.
- Twitter accounts that have been deleted, banned or privated on Twitter will not be analyzed by Botometer.
- To ensure that you are not flagged as a security/excessive user of the Botometer API, Communalytic will only use up to 80% of your daily API limit. When the 80% threshold is reached, the analysis will pause for a day.

Twitter Bot Analysis

The Twitter Bot Analysis Module is designed to detect potential use of automation. It analyzes accounts in your dataset and generates the probability scores for a variety of different types of automated and/or fraudulent activities (from 0=not likely to 1=highly likely). Once the analysis is complete, the probability scores for each account will be added to your dataset.

Analysis In Progress ...

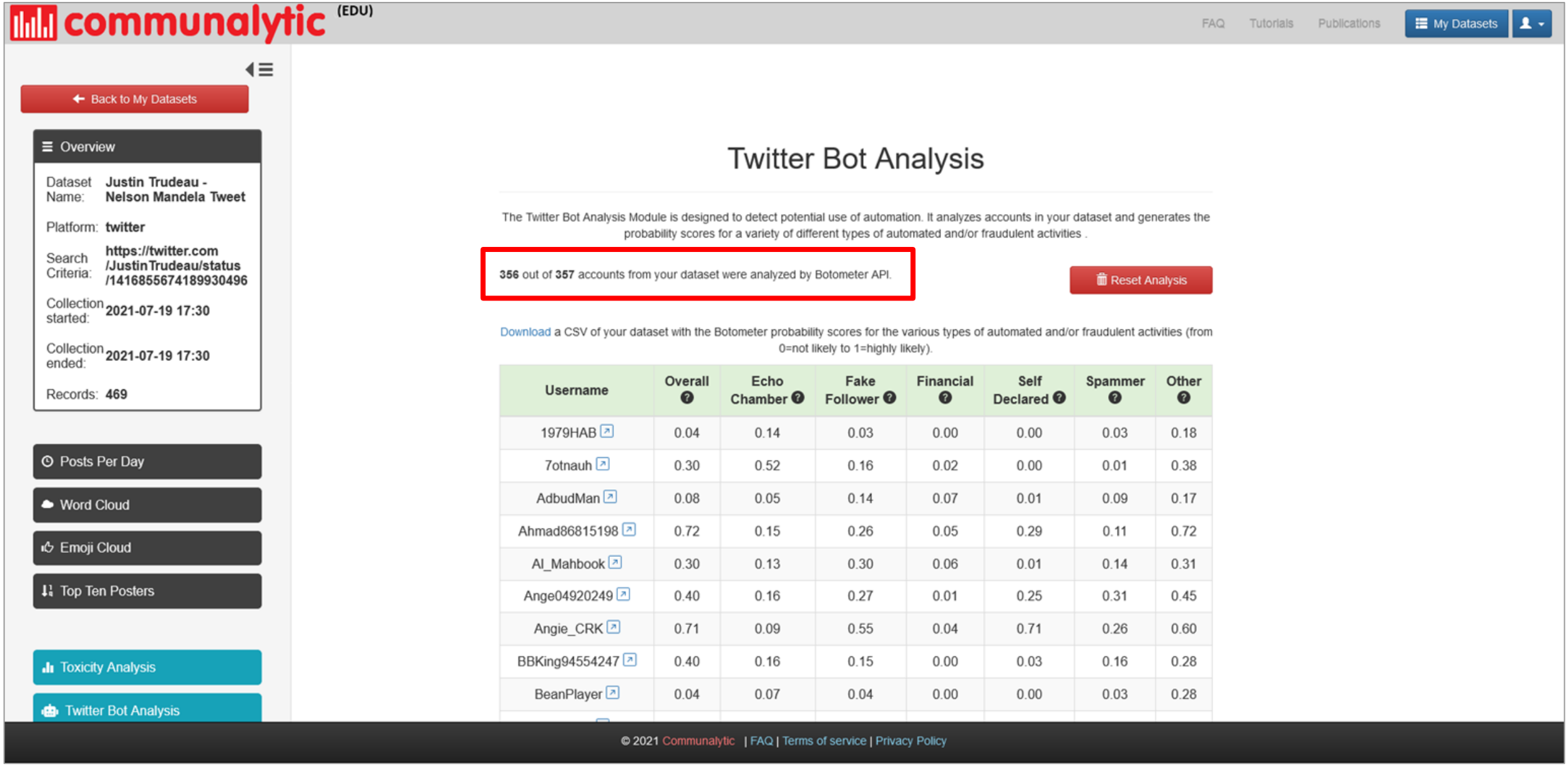
(You may close this window and visit later to download the bot analysis results)

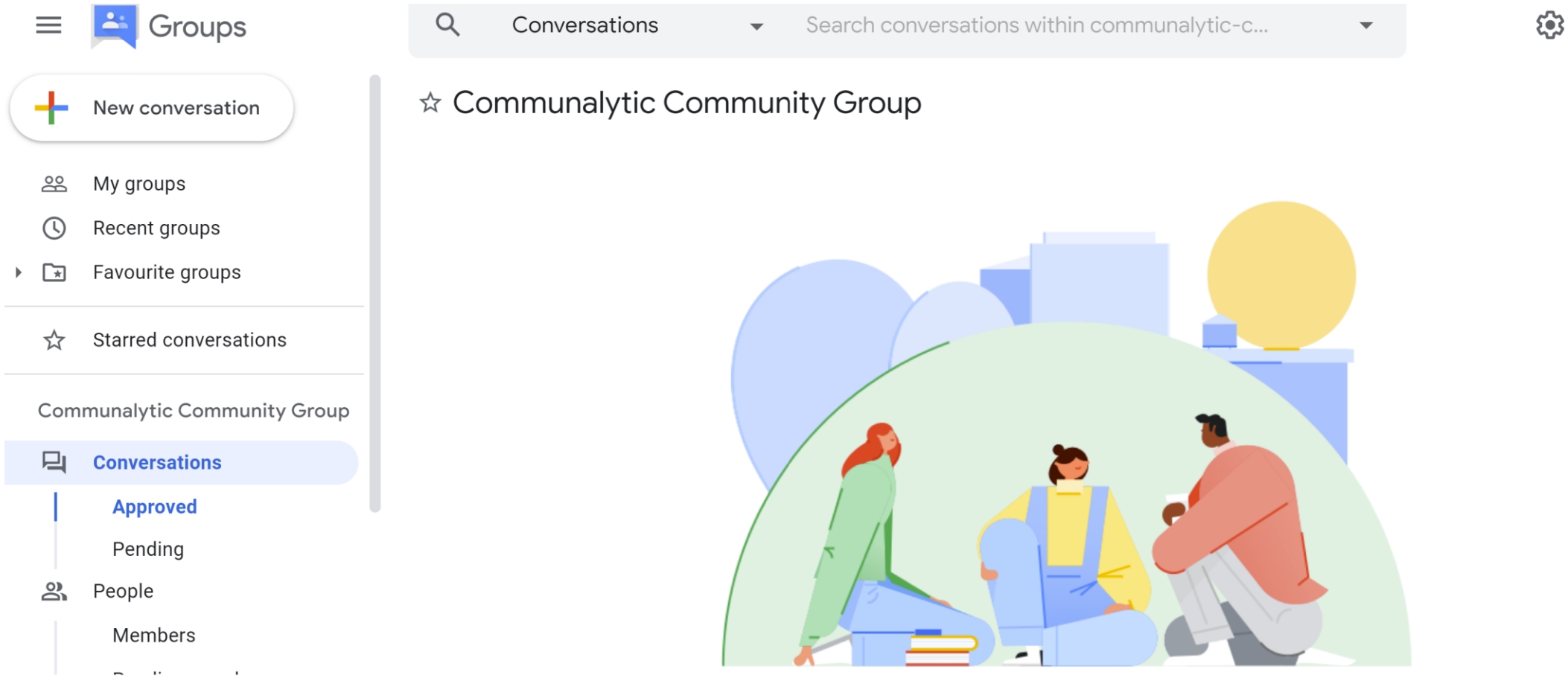
34%

Progress: 122 / 357 Users Analyzed

Check progress in 19 sec

Cancel Analysis


communalytic (EDU)
FAQ
Tutorials
Publications
My Datasets
Back to My Datasets
Overview
Dataset Name: Justin Trudeau - Nelson Mandela Tweet
Platform: twitter
Search Criteria: https://twitter.com/JustinTrudeau/status/1416855674189930496
Collection started: 2021-07-19 17:30
Collection ended: 2021-07-19 17:30
Records: 469
Posts Per Day
Word Cloud
Emoji Cloud
Top Ten Posters
Toxicity Analysis
Twitter Bot Analysis
Twitter Bot Analysis
The Twitter Bot Analysis Module is designed to detect potential use of automation. It analyzes accounts in your dataset and generates the probability scores for a variety of different types of automated and/or fraudulent activities .
356 out of 357 accounts from your dataset were analyzed by Botometer API.
Reset Analysis
Download a CSV of your dataset with the Botometer probability scores for the various types of automated and/or fraudulent activities (from 0=not likely to 1=highly likely).
Username
Overall
Echo Chamber
Fake Follower
Financial
Self Declared
Spammer
Other
1979HAB 0.04 0.14 0.03 0.00 0.00 0.03 0.18
7otnauh 0.30 0.52 0.16 0.02 0.00 0.01 0.38
AdbudMan 0.08 0.05 0.14 0.07 0.01 0.09 0.17
Ahmad86815198 0.72 0.15 0.26 0.05 0.29 0.11 0.72
Al_Mahbook 0.30 0.13 0.30 0.06 0.01 0.14 0.31
Ange04920249 0.40 0.16 0.27 0.01 0.25 0.31 0.45
Angie_CRK 0.71 0.09 0.55 0.04 0.71 0.26 0.60
BBKing94554247 0.40 0.16 0.15 0.00 0.03 0.16 0.28
BeanPlayer 0.04 0.07 0.04 0.00 0.00 0.03 0.28
© 2021 Communalytic | FAQ | Terms of service | Privacy Policy

Join the Communalytic Discussion Group

https://groups.google.com/u/1/g/communalytic-community-group