



**Team:**

Yeshvendra Singh

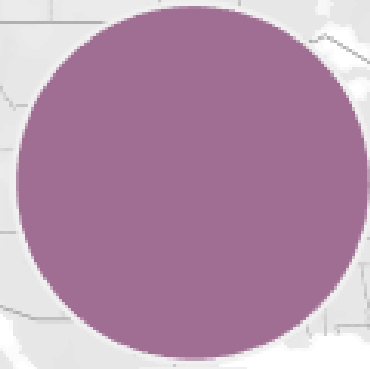
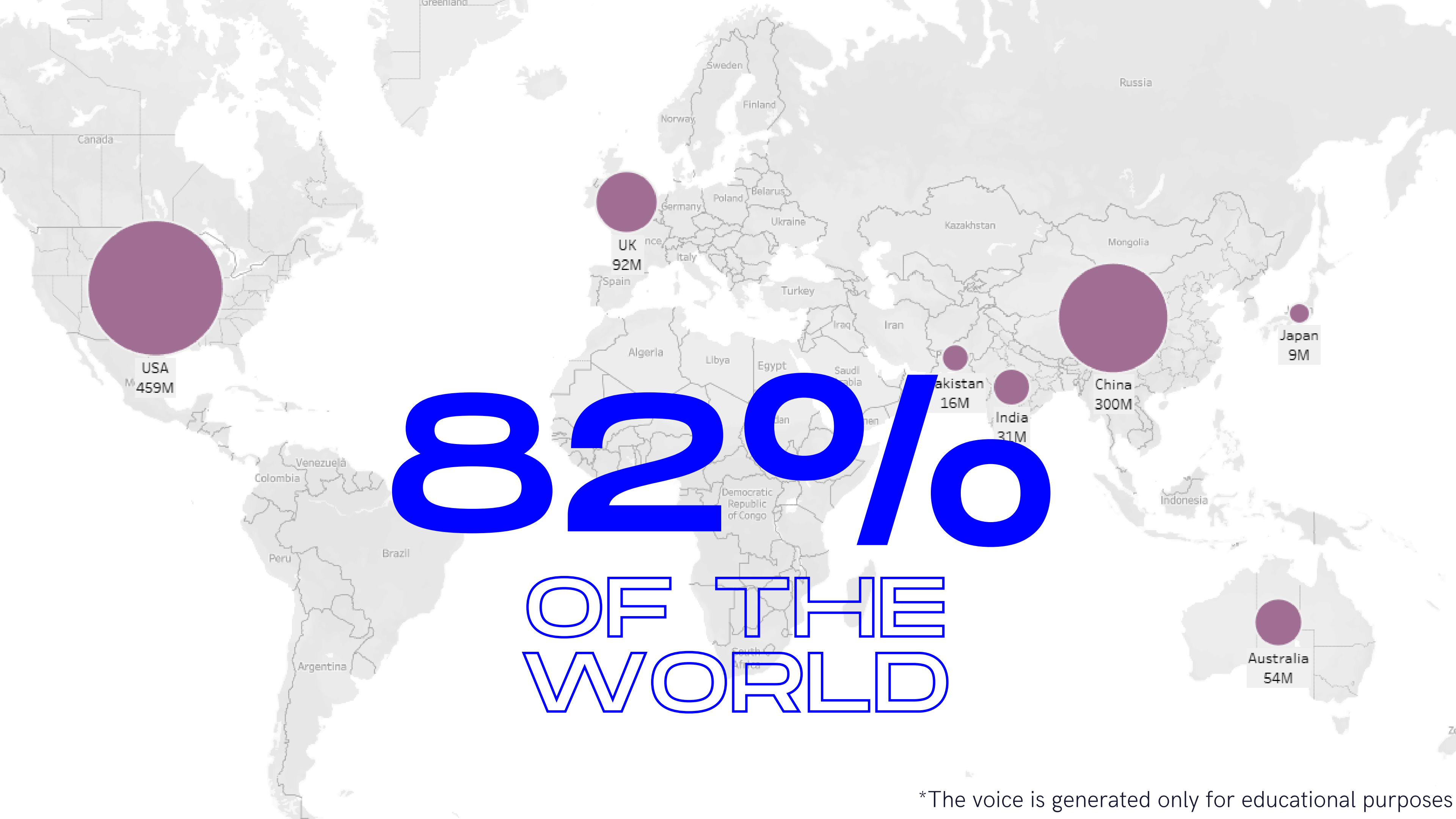
Zelina Naozer Wankadiya

Fatima-Ezzahra Areqti

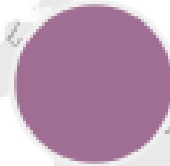
Dayana Kartieva

Chinese presence on  
Social Media Platforms

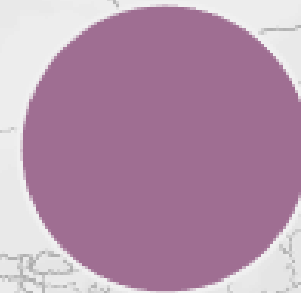
**influence and  
interference**



USA  
459M



UK  
92M



China  
300M



Japan  
9M



Pakistan  
16M



India  
31M



Australia  
54M

82%

OF THE  
WORLD

\*The voice is generated only for educational purposes

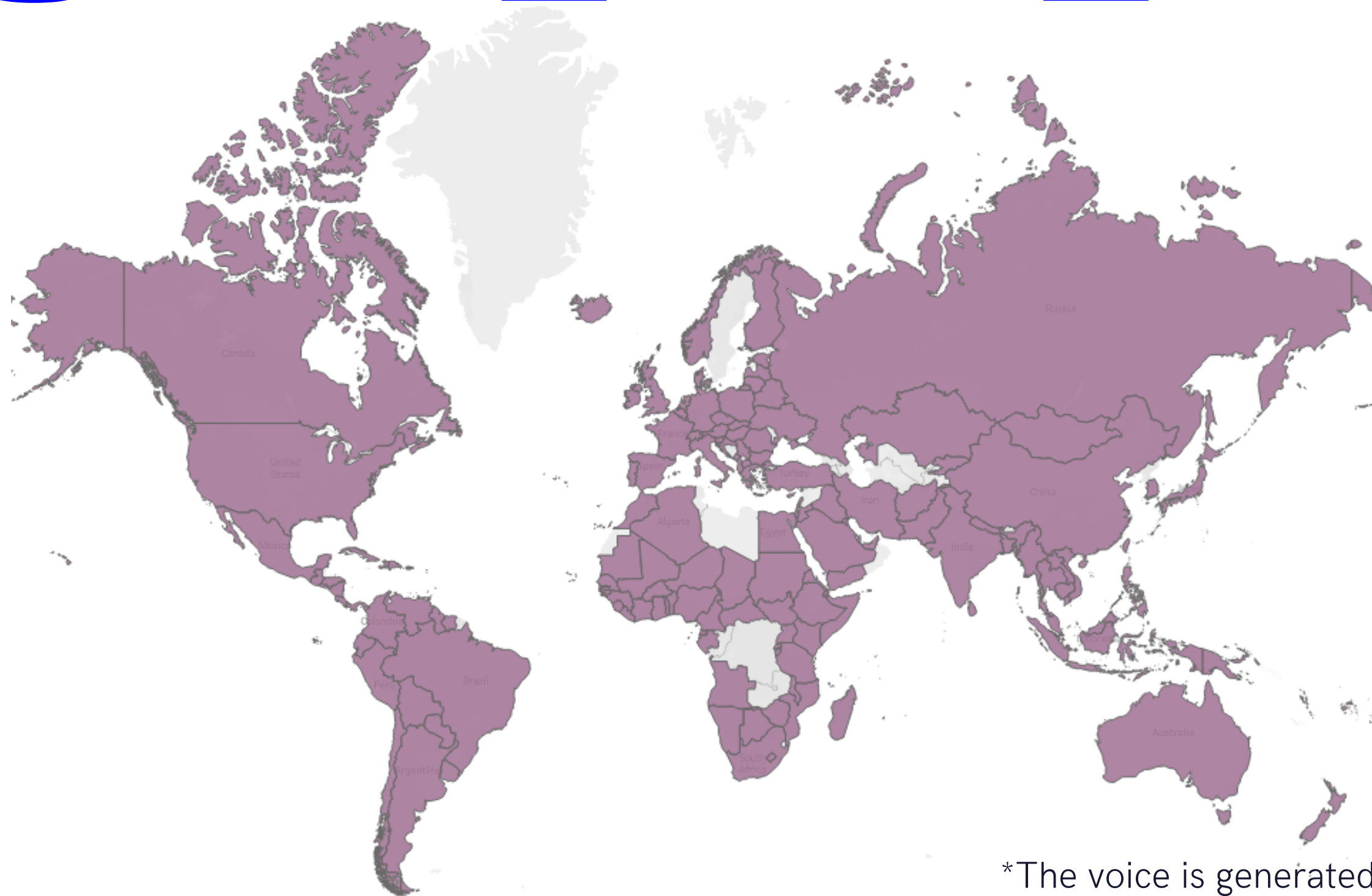
CAN YOU SEE



\*The voice is generated only for educational purposes

# HOW WE

## SCALE EFFORT EFFECT



\*The voice is generated only for educational purposes

# Scalability by Languages

English  
61.78%

Chinese  
27.26%

French  
2.59%

Tamil  
1.69%

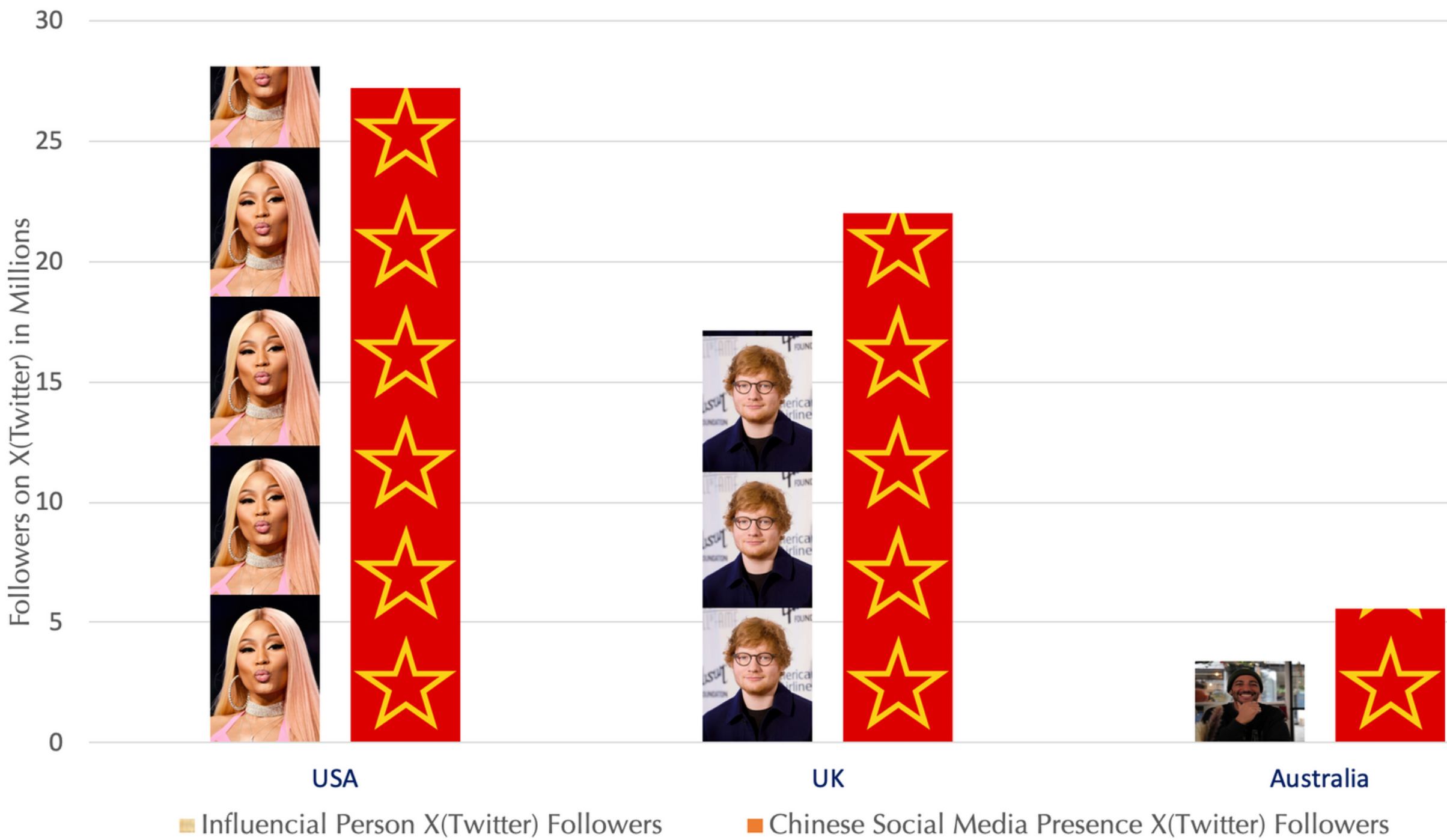
Arabic  
1.80%

Urdu  
1.33%





Influence VS Interference in Top 3 countries



Scalability by Followers



CANIS Data Visualization and Foreign Interference

# Scalability in regions



**1 Adult in 5  
on the Planet  
follows  
Chinese social  
media**



**Are you of those Ten?**



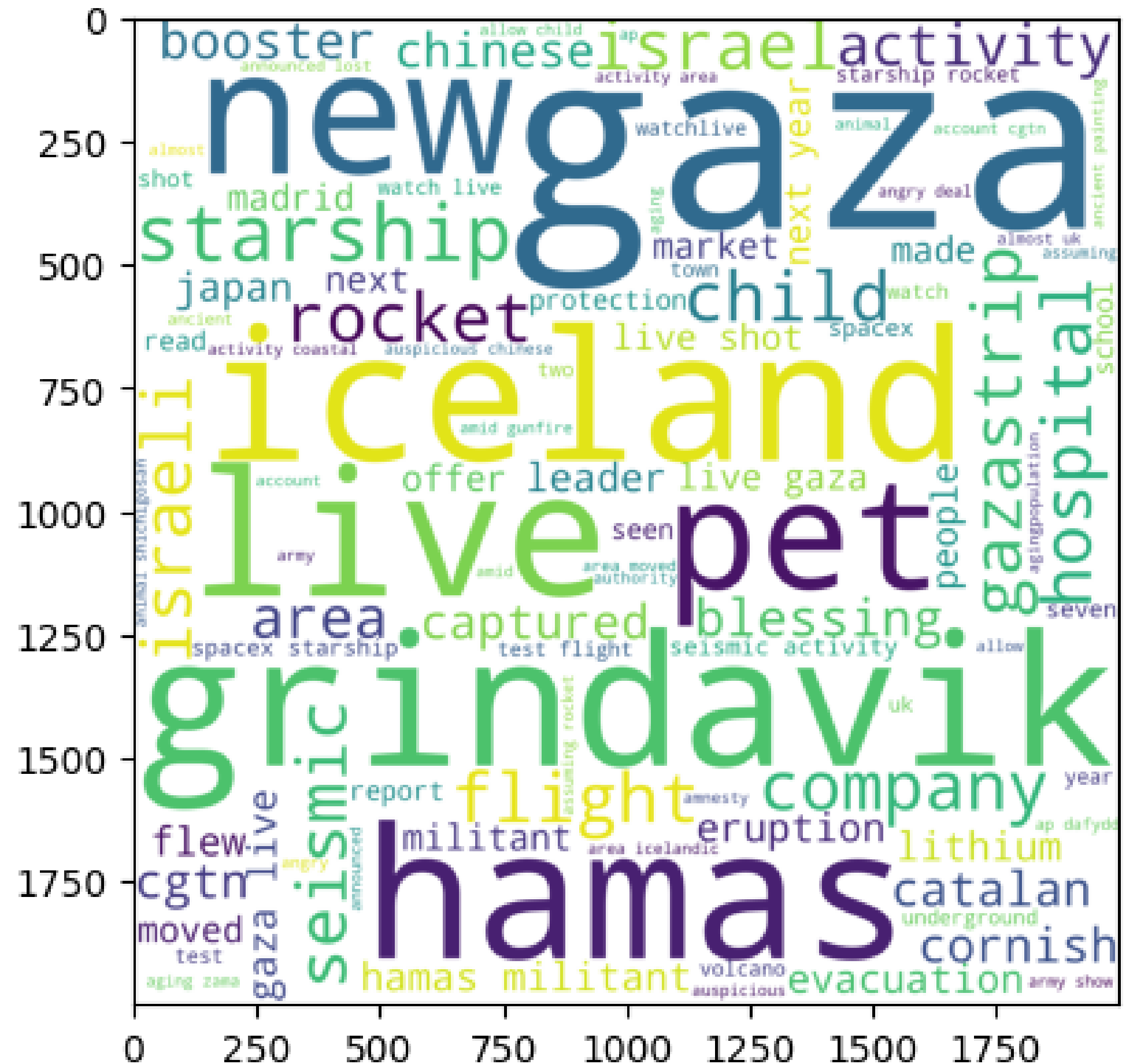
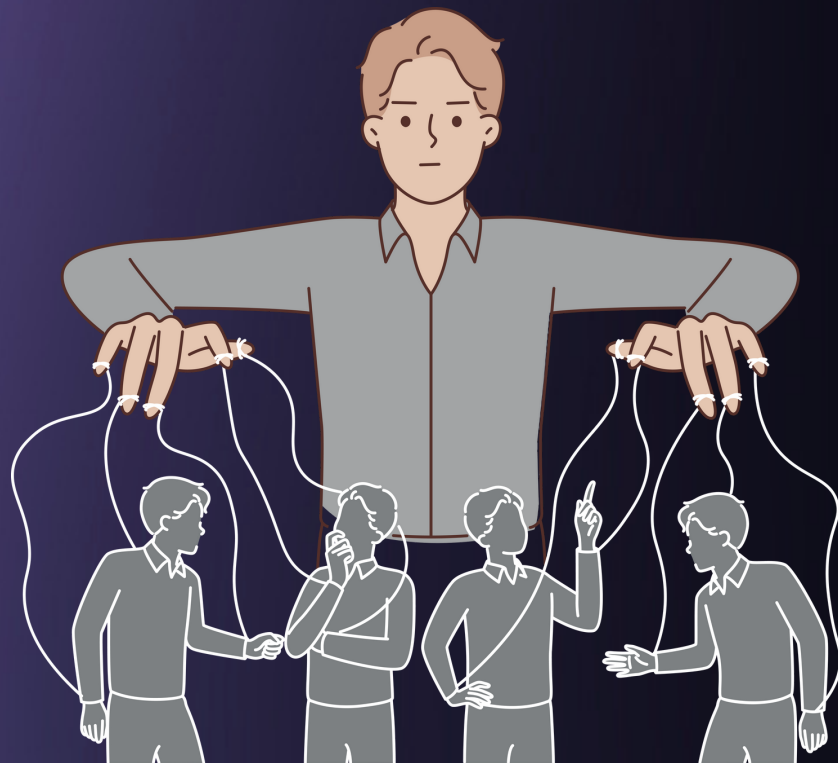
# While you slept on November 18

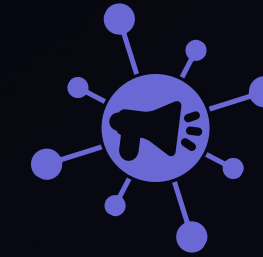
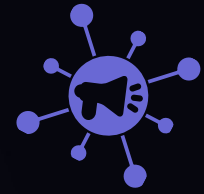


**17 h of video**  
**20 posts**

were posted by one  
account **CGTN Europe**  
on Facebook  
(2,1M followers)

# The content of those activities





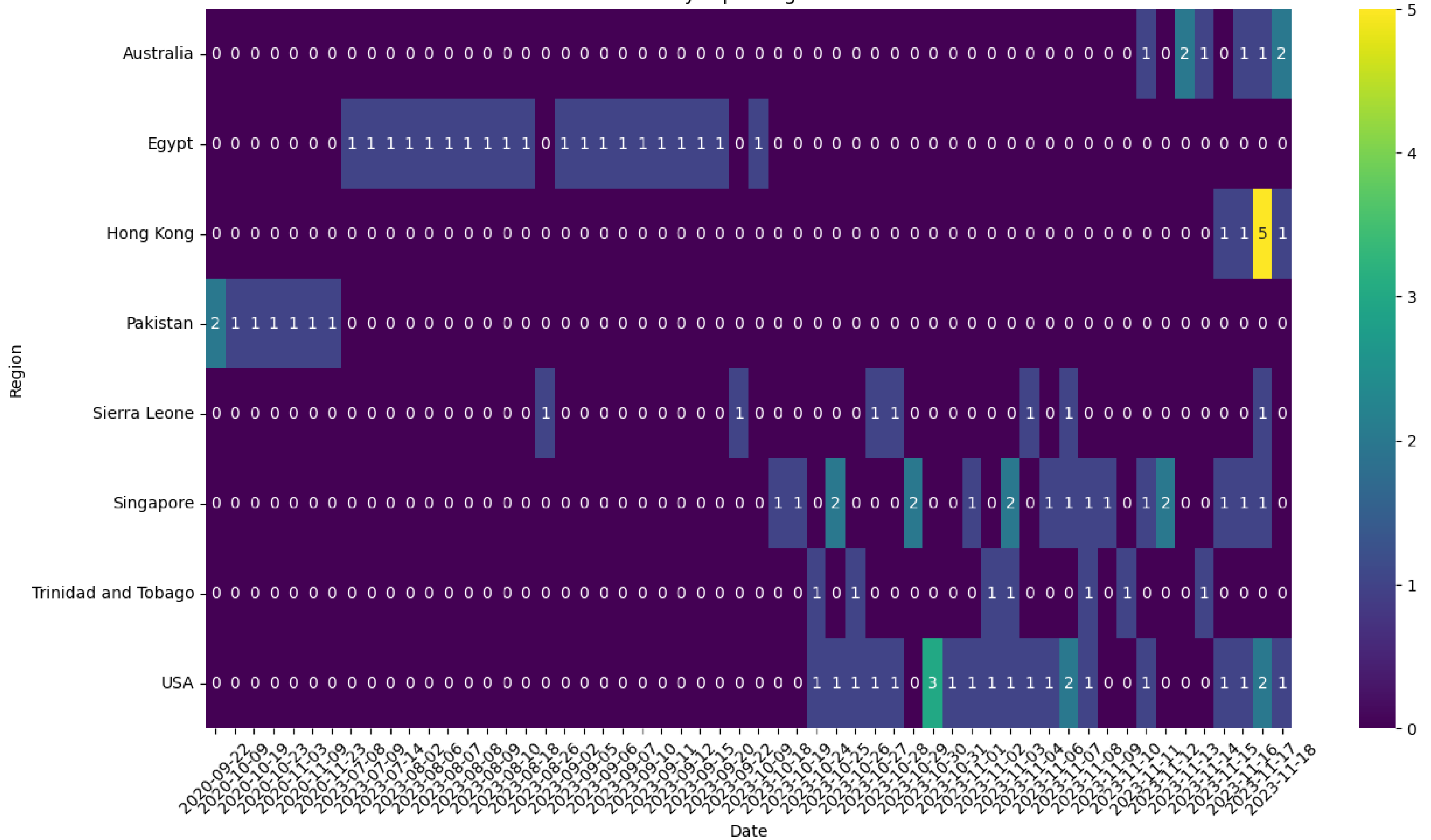
# Let's focus on posts that went VIRAL



## Criteria

- likes
- shares
- comments

### Viral Posts by Top 8 Regions and Date





Methodology

# Tools that we used

TOOL 01

**Tableau**

TOOL 02

**Excel**

TOOLS 03

**Python**



# Data cleaning pipeline

## Technique 1

Handled all the NAN values

## Technique 2

Filtered out ASCII code related values, removed stop words, applied lemmatization

## Technique 3

Filtered dataset based on regions, languages and SMPs specific pages

# Challenges

1. Handling the data in multiple languages
2. Handling different region-related data
3. Limitation of data points provided leading to web scraping problems