



Social Media Trends and Sentiment analysis

Russia – Ukraine War (2022)

Team 5 – Camille Nigon, Vladimira Gabor, Hendrik Croce

The background of the slide is a light gray map with blue lines representing rivers or roads. There are several blue dots and 'x' marks scattered across the map. A white rectangular box with a blue border is centered on the slide, containing the text.

Business Goals & Research Questions

Business Goals

- Provide an informative **social media monitor** as a non-profit open-source solution
- We provide quantitative and qualitative overview of trends and sentiments of social media reactions:
 - Quantitative – density of keywords, trending buzzwords
 - Qualitative – sentiment analysis
- With our solution, users can understand faster sentiments, compare trends on social media and take more informed actions based on our analysis
- Our main target customer segments include reporters, NGOs and interested users
- We analyse and represent human language computationally as a service and house a data lake with directly extracted information from Twitter & Reddit

Research Questions

Question #1

Which buzzwords are used across social media (Twitter & Reddit) related to Ukraine-Russia war?

Question #2

What is the daily development of word choice over time in the Ukraine-Russia war?

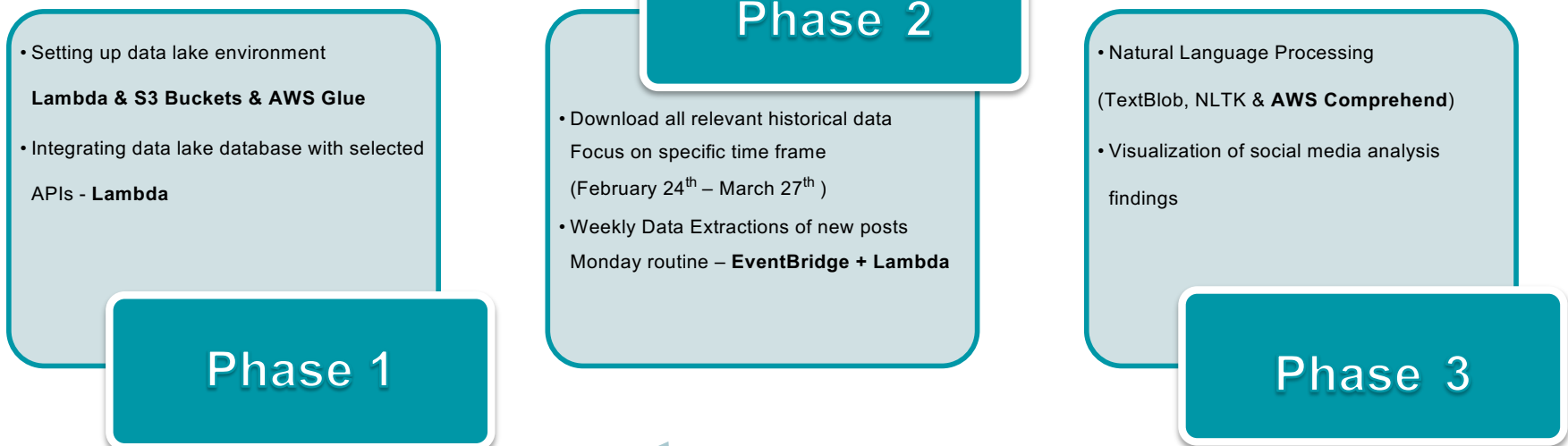
Question #3

What is the sentiment development on the topic of Russia-Ukraine war over time?

The background of the slide is a light gray map. It features several blue contour lines, some solid and some dashed. There are five blue dots placed at various locations on the map, and six small 'x' marks, also in blue, scattered across the background. A white rectangular box with a blue border is centered on the slide, containing the section header.

2. Methodology

Methodology



The background of the slide is a light gray map with blue contour lines. There are five blue dots placed at various locations on the map, and several small 'x' marks are scattered across the background.

3. API Sources

Twitter API



- Academic Research Access
- Hashtags:
 - #StandWithUkraine
 - #UkraineRussiaWar
 - #UkraineWar
- Tweepy Python Library
- 3200 most liked tweet per request

Reddit API

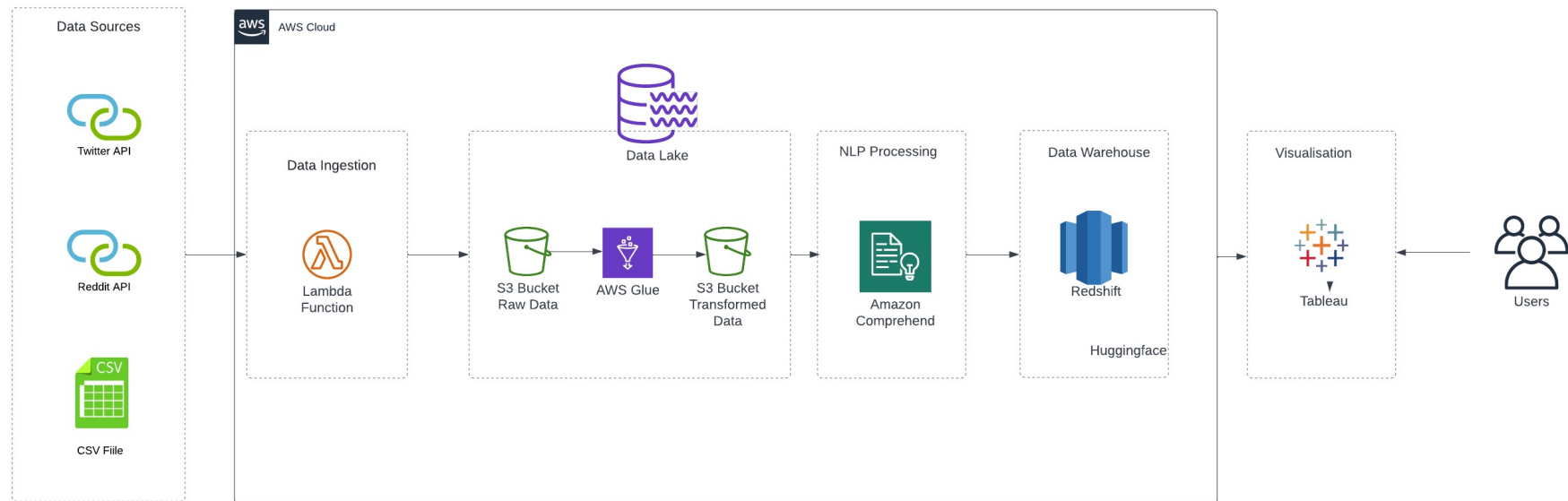


- Subreddit – “UkrainianConflict”
330K members
- Focusing on Headlines
- 20 top posts **this weeks**
(EventBridge)
- PRAW & PushShift Python Library

The background of the slide is a light gray map. It features several blue lines, some solid and some dashed, which represent geographical features like rivers or roads. There are also several small blue dots scattered across the map. A white rectangular box with a blue border is centered on the map, containing the text '4. Architecture'.

4. Architecture

Architecture



The background of the slide is a light gray map. It features several blue lines, some solid and some dashed, which represent geographical features like rivers or roads. There are also several small blue dots scattered across the map, possibly indicating specific locations or data points. The overall style is clean and modern.

5. Business Insights

Insufficient Accuracy of Sentiment Analysis - Multipolarity

Having only the “total result” of the analysis can be misleading

Challenge is that we assign polarity to sentence aspects, ignoring the broad contextual meanings

Common analytical challenge when assessing **political sentiment analysis**.

Inssufficient Accuracy of Sentiment Analysis - NLTK

"Ukrainian forces **destroy** an ammunition depot in the Belgorod region of Russia"

Compound -0.5423 (negative)

"The **good news** from the Russia-Ukraine talks in Istanbul **is not that Putin** is suddenly acting in **good faith**, but that **heroic** Ukrainian resistance is making him look for off-ramps and diplomatic disguise"

Compound +0.8020 (highly positive)

"Russian oligarch Deripaska **loses** U.S. court battle to lift sanctions"

Compound -0.5994 (negative)

*Compound Assesed Value determines whether the analyzed headline is positive or negative.

Keyword Density – Redefinition of Desired Outcomes



The background of the slide is a light gray map. It features several blue contour lines, some solid and some dashed. There are five blue dots placed at various locations on the map, and four small 'x' marks, also in blue, scattered across the background. A white rectangular box with a blue border is centered on the slide, containing the text '6. Next Steps'.

6. Next Steps

Next Steps

- Reframe Sentiment Analysis approach - research possible extensions such as Anger, Fear, and Happiness.
- Alternatively, try to label and categorize sentiments as “Pro Russia” vs “Pro Ukraine”
- Excluding hashtags and mentions such as “Ukraine, Russia, War” from the keyword density & buzzwords analysis
- Peer review of our sentiment analysis with third party
<https://twitter-sentiment-csv.herokuapp.com/>

The background of the slide is a light gray map. It features several blue contour lines, some solid and some dashed. There are five blue dots placed at various locations on the map: one on a dashed line on the left, one on a solid line in the upper middle, one on a solid line in the lower middle, one on a solid line on the right, and one on a solid line in the bottom right. There are also several small 'x' marks scattered across the map, some on dashed lines and some in open areas. A white rectangular box with a blue border is centered on the slide, containing the text '7. Lessons Learned'.

7. Lessons Learned

Lessons Learned

- Limited capacity of Cloud9 Lambda layer - unable to sustain large grouping of Python packages
- Adjusting Twitter data extraction to 3200 Tweets per request not to overreach the Lambda boundaries (15mins)
- Multi-layer challenges of Natural Language Processing



**Thank you for your
attention!**