Multimodal Telegram Message Analysis: An Exploratory Study

Sara Morrell

Does the subject of Telegram messages impact message views? In this study, I examined the impact of text and image content on message view counts for messages sent between 15 February 2022 and 15 March 2022 in a Telegram channel associated with the Russian Imperial Movement (RIM). My goal was to examine how multimodal computational tools can be used to inductively explore online discourse associated with an extremist group.

Dr. Sunny Yang provided mentorship, feedback, and methodological assistance for developing this project. Datasets and code will be made available upon request.

In this exploratory study, I applied a variety of computer vision and natural language processing analysis techniques to a small dataset. While the small size of the dataset limits the conclusions that can be drawn from the results of this analysis, applying these methods yields insights into which methods might be useful for future research on online discourse. I used the programming language Python to perform image and text analysis. To analyze images, I used the Google Cloud Vision API and K-Means clustering with the Python libraries Keras and Scikit-learn. I also used the Scikit-learn library, along with the spaCy library, to analyze text. In order to understand the combined impact of images and text on engagement, I ran a mixed effects regression using the Statsmodels module. While the regression provided limited evidence of the impact of text or image themes on message views, this project was still a useful learning experience for exploring different methods. Python provided me the flexibility to conduct data cleaning and multimodal analysis all in the same place. Data and code is available upon request.

Data

My dataset includes 53 images sent as messages in a Telegram channel associated with the Russian Imperial Movement (RIM), 31 of which are accompanied by text. These messages were sent between 15 February 2022 and 15 March 2022. At the time of data retrieval, the channel had approximately 3,000 subscribers. I chose to retrieve data from a Telegram channel associated with the RIM because the online communication of extremist groups is relevant to my dissertation work on the Internet and violence. An off-shoot of the Russian Imperial Movement, the Imperial Legion, has been involved in the Russia-Ukraine war as well as other conflicts and acts of violence abroad. I chose the time period 15 February 2022 to 15 March 2022 as it includes Russia's full scale invasion of Ukraine and escalation of the Russia-Ukraine war on 24 February 2022. Given this major event, I assume that a Telegram channel associated with a group involved in the war would be particularly active.

To retrieve image and text data, I use the Telegram desktop application to export messages posted between 15 February 2022 and 15 March 2022, along with their attached images. To retrieve the view counts for the same messages, I used the [4Cat Capture and Analysis Toolkit](#) to gather message metadata. I then merged the two datasets on the basis of their unique message IDs into a single table in Python. Images were kept in a separate folder, but their file names were included in the table along with the message text and view counts. To clean the text, I removed numbers, emojis, and stopwords, and converted all text to lowercase.

Image Analysis: Computer Vision

I first explored how the Google Vision API could be used to identify various image features. To do this, I adapted code provided by Dr. Sunny Yang for accessing the API and applying it to my images. I used the API capabilities for detecting faces, entity labels (e.g. military uniform, history), and logos. While the face and logo detection did not work as well as I had hoped they would, based on this exploratory trial run I think that the label detection could be useful for future analysis as it seems to provide an effective representation of the main objects and themes in a photograph. For example, a black and white photograph of a group of people in what appear to be 19th century military uniforms was annotated with the labels Crew, Military person, Suit, Military uniform, Military organization, Event, Uniform, History, Team, and Stock photography.

To identify themes across the images, I used the unsupervised machine learning algorithm k-means clustering. To calculate the k-means image clusters, I used the Keras model VGG 16 and the KMeans and PCA packages from Scikit-learn to extract the image features, reduce the number of dimensions, and calculate the clusters.

The 53 images were grouped into the clusters People and Political/Religious Symbols, Military Vehicles, Crowds and Protests, Memorial Plaque, Old Building Photo/Illustration and Maps, Documents and Illustrations, Russian Foreign Ministry Spokeswoman, Buildings and Political/Religious Symbols, Antique Photographs of People, and Military Truck. I manually labeled these clusters after reviewing their contents. While there were a few outliers in some of the clusters, for example a picture of a submarine in the Crowds and Protests cluster, the clusters do appear to accurately represent themes in the photos. In the data, clusters are labeled 0 through 9. In this paper, I label them 1 through 10 for clarity.

Text Analysis: Natural Language Processing

In the next phases of analysis, I applied natural language processing methods to analyze the message text. To automatically extract named entities such as locations or organizations, I used the spaCy package and the pretrained pipeline [ru_core_news_sm](#). The extracted named entities included Russia; first Russia; Novorossiya; Little Russia; Moscow; Ukraine; Kiev; Mariupol;

Turiis'k; the Donetsk People's Republic; Grozny; Ichkeria; Europe; Serbia; Serbian; Belgrade; Sarajevo; Yugoslavia; Auschwitz; the United States; NATO; the names of religious, political, and military figures; the names of other Telegram channels; the names of military, law enforcement, and government organizations; mass media; wolf; pandemic; and motor rally. These named entities largely align with the focus of the Russian Imperial Movement on military, political, and religious themes. In future research, I think the extraction of place names from text using named entity recognition could be useful to incorporate in a geographically focused analysis.

To identify themes across the messages, I used latent dirichlet allocation topic modeling from the Gensim library. Topic modeling identifies themes in text based on word co-occurrence. Since my dataset is relatively small and the messages are generally short I will likely have relatively few topics. To evaluate an appropriate number of topics for my model, I ran several different visualizations using the package pyLDAvis. I used the visualizations to subjectively approximate topic exclusivity, or how distinct each topic is from other topics. I also examined the words most strongly associated with each topic to determine which set of topics was most interpretable. Topic exclusivity is represented in the visualization using an intertopic distance map; exclusive topics do not overlap on the map. I tested models with 10, 7, 5, 4, and 3 topics. The visualizations for each model are displayed below.
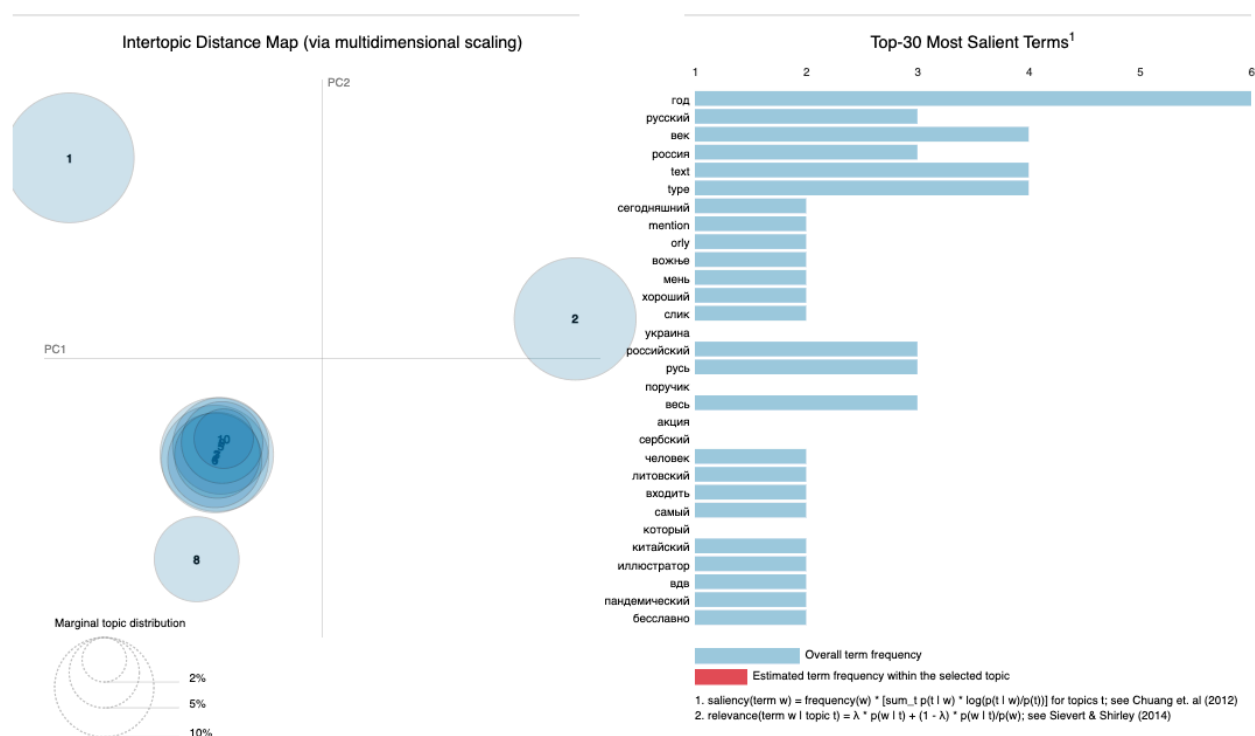
Figure 1
Topic Model with 10 Topics
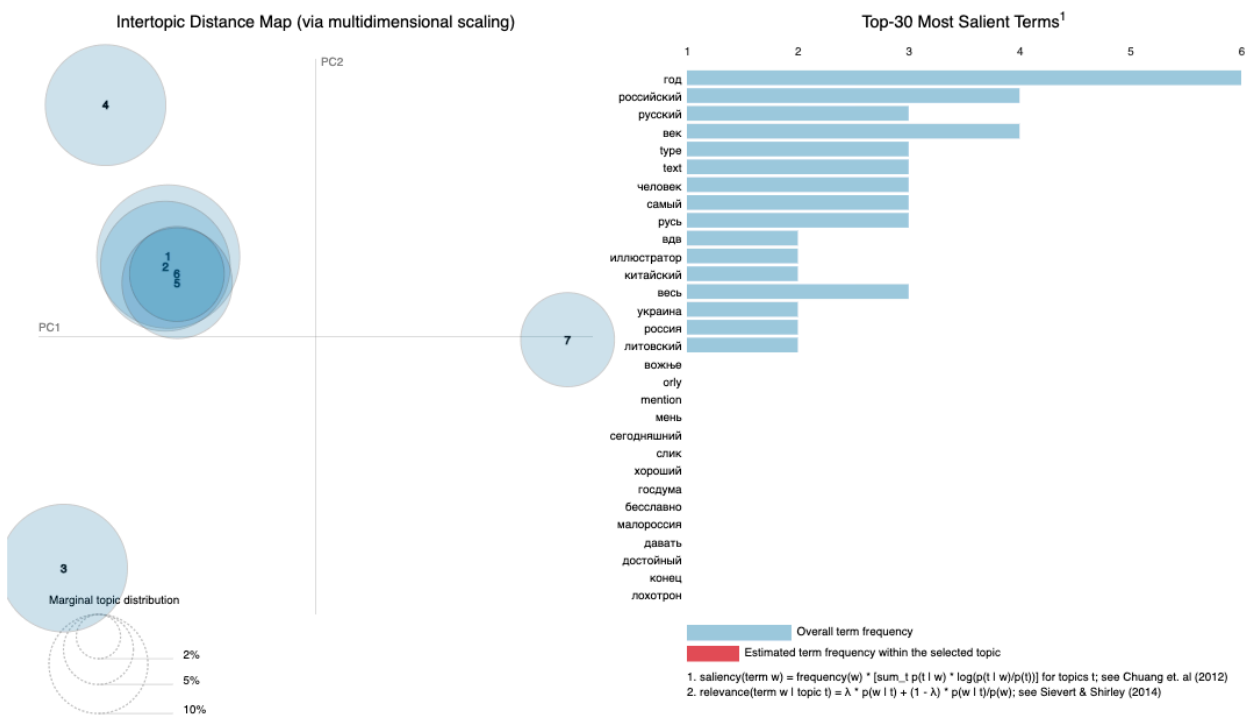
Figure 2
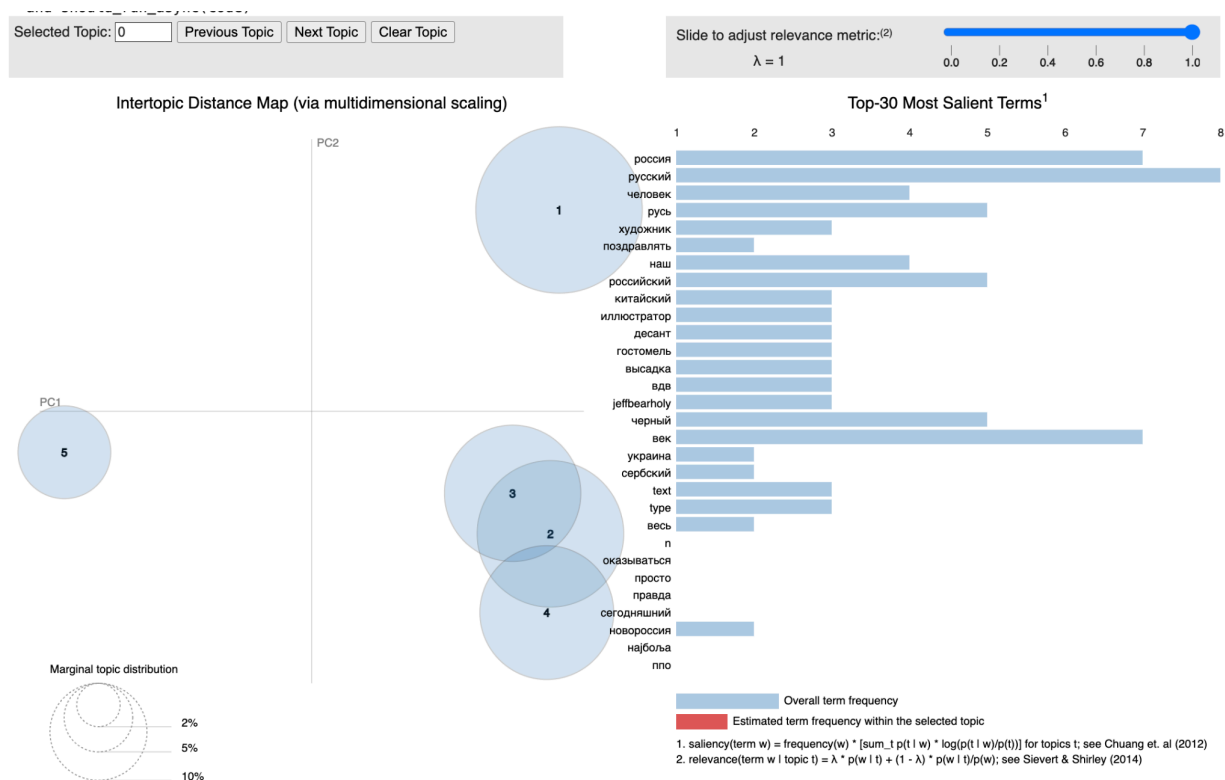*Topic Model with 7 Topics*

Intertopic Distance Map (via multidimensional scaling)

Top-30 Most Salient Terms[1]

Figure 3
Topic Model with 5 Topics

Figure 4
Topic Model with 4 Topics

Figure 5
Topic Model with 3 Topics

Intertopic Distance Map (via multidimensional scaling)

Top-30 Most Salient Terms[1]

Marginal topic distribution

2%
5%
10%

■ Overall term frequency
■ Estimated term frequency within the selected topic
1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)
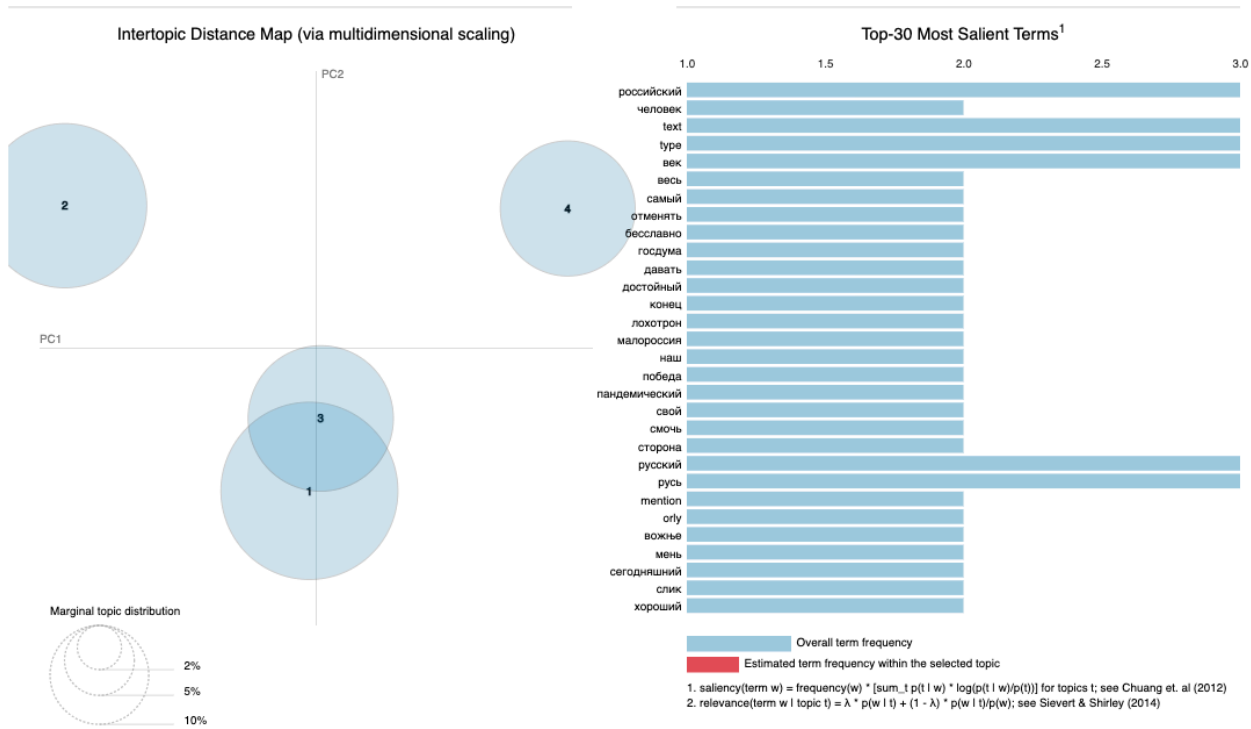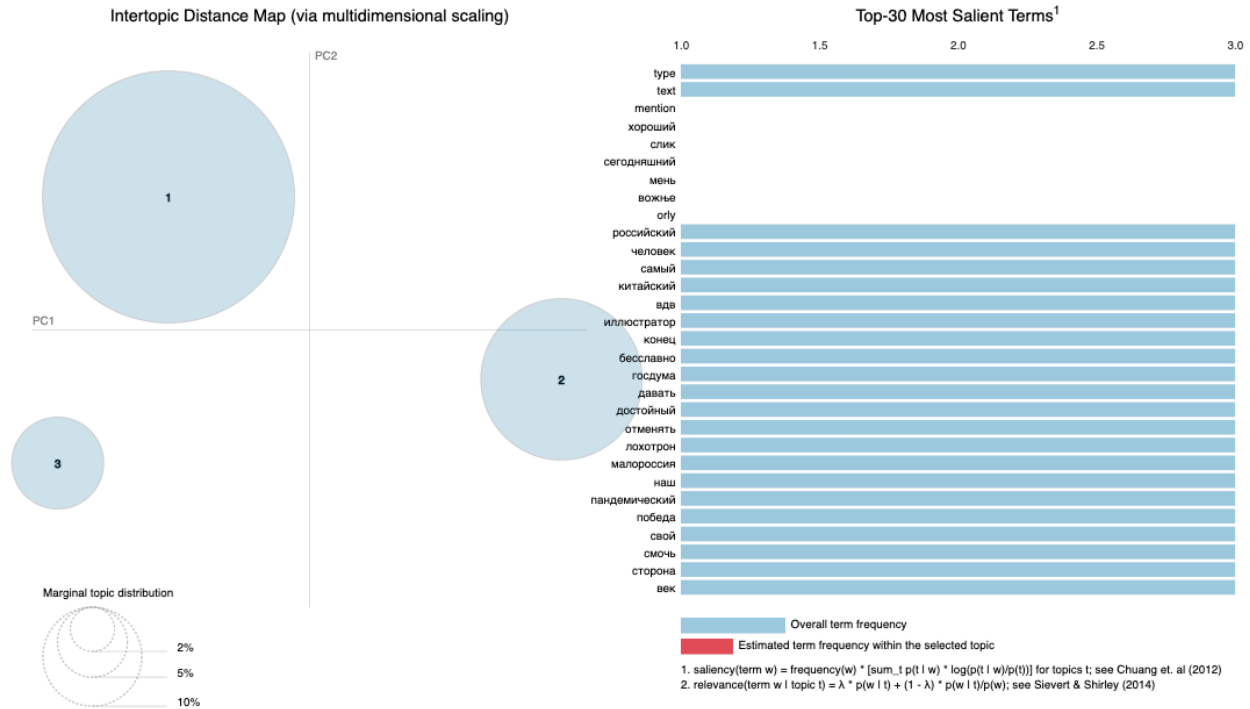
Based on these visualizations, models with 10 and 7 topics clearly do not yield sufficiently distinct topics. Based on the words associated with each topic, I do not think the topics from the models with 3 or 4 topics are interpretable. Therefore, I will use the model with 5 topics, as the topics are relatively distinct and interpretable.

Table 1 displays the words most strongly associated with each of the topics in the 5 topic model. I subjectively interpret and label each topic based on the words associated with that topic. If the most strongly associated words did not form a clear topic, as in topic 5, I also reviewed the messages associated with that topic to help assign a label. In the data, topics are labeled 0 through 4. In this paper, I label them 1 through 5 for clarity.

Table 1
Topic Words and Labels

Topic	Label (Interpretation)	Words
1	Historical Kievan Rus	*русь* (Rus)**,** *черный* (black), *век* (century, age, time)**,** *год* (year)**,** *литовский* (Lithuanian), xiii, *состав* (composition, compound, makeup)**,** *называть* (call, name)**,** *княжество* (principality, princedom)**,** *xiv*

2	Celebratory Motor Rally	*поздравлять* (congratulate, compliment, felicitate), *правда* (really, truth, sooth, verity)**,** *сегодняшний* (today, today's), *данашње* (today, present-day), rs, *најбоља* (the best, good)**,** *пожалуй* (perhaps, rather, very likely)**,** *вожње* (driving, drive, ride), orly, *автопробег* (rally, motor rally)**,**
3	Russian Nation	*россия* (Russia), *русский* (Russian), *человек* (human, person, man)**,** *наш* (our), *век* (century, age, time)**,** *год* (year)*,* *население* (population, people)**,** *народ* (people, the people, nation, folk)*,* *помнить* (remember, keep in mind, mind)**,** *первый* (first, top)
4	Russian Military Deployment	*год* (year)**,** *русский* (Russian), *грозный* (fearsome, formidable, redoubtable, Grozny), *солдат* (soldier, enlisted man)*,* *станция* (station, post), *владикавказский* (Vladikavkaz)**,** *дорога* (road, way, route)**,** *железный* (iron, irony, ferreous)**,** *барс* (leopard), *лодка* (boat)
5	Russian Forces in Ukraine	*российский* (Russian), *художник* (artist, painter, craftsman)**,** *вдв* (airborne forces)**,** *китайский* (Chinese)**,** jeffbearholy**,** *высадка* (landing), *гостомель* (guest house, Hostomel')*,* *десант* (landing, decent)**,** *иллюстратор* (illustrator)**,** *украина* (Ukraine)

Note. Translations performed using the [Cambridge Dictionary](). If the translation provided by Cambridge Dictionary seemed inaccurate or incomplete, then the word was Googled and [Google translate]() was used to assist in a correct translation. Errors are the author's own. If a word could have multiple meanings in English, multiple potential translations are provided in the parenthesis.

Engagement

For this research project, I focused on the impact of image cluster and text topic on the level of engagement with a message (measured through view count). I first present descriptive statistics on these three variables and then present the results of a mixed effects regression testing the significance of the impact of text topics and image clusters on message views.

Descriptive Statistics

Table 2 displays the number of messages and average view count for each topic and cluster. This data aggregation was done using pivot tables in the Google sheets.

Table 2

Topic and Cluster Message Counts and Average Views

Subject	Message Count	Average Views
Text Topic		
1: Historical Kievan Rus	8	501.75
2: Celebratory Motor Rally	3	421.67
3: Russian Nation	6	528.50
4: Russian Military Deployment	7	446.43
5: Russian Forces in Ukraine	7	564.86
Text Total	**Total Text Message Count**	**Total Text Average Views**
	31	492.64
Image Cluster		
1: People and Political/Religious Symbols	9	583.44
2: Military Vehicles	2	608
3: Crowds and Protests	7	579.57
4: Memorial Plaque	1	526
5: Old Building Photo/Illustration and Maps	6	606.17
6: Documents and Illustrations	9	533.11
7: Russian Foreign Ministry Spokeswoman	1	541
8: Buildings and Political/Religious Symbols	6	610.83
9: Antique Photographs of People	12	443.58
10: Military Truck	1	526
Image Total	**Total Image Message Count**	**Total Image Average Views**
	54	555.77
Overall Total	**Total Message Count**	**Total Average Views**
	54	547.04

Interestingly, the messages with no text, but with an image, had an average of 609.17 views, higher than the average for any text topic.

To demonstrate how the text topics and image clusters overlap, figure 6 displays a co-occurrence matrix of topics and clusters.

Figure 6

Topic and Cluster Co-Occurrence Matrix
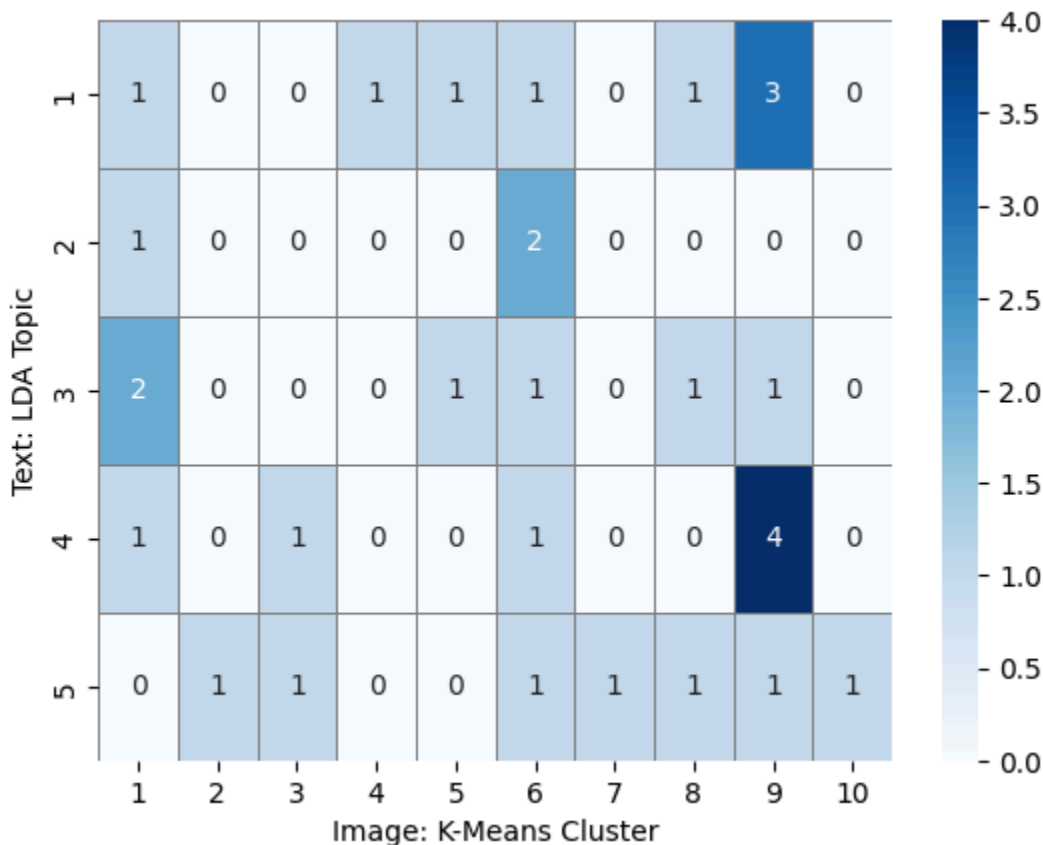
Given the small dataset and low number of messages for many of the topics and clusters, this matrix is fairly sparse. The topic and cluster pairs with the greatest number of co-occurrences are topic 4 (Russian Military Deployment) and cluster 9 (Antique Photographs of People). The pair with the second greatest number of co-occurrences are topic 1 (Historical Kievan Rus) and cluster 9 (Antique Photographs of People). Other pairs with more than 1 co-occurrence are topic 2 (Celebratory Motor Rally) and cluster 6 (Documents and Illustrations) and topic 3 (Russian Nation) and cluster 1 (People and Political/Religious Symbols).

Given that the message and image subjects, as well as view counts, might be influenced by external global events, I also graphed view count over time, along with topic and cluster frequency over time. Figure 7 displays the view counts for messages sent each day.

Figure 7

Message View Count Over Time
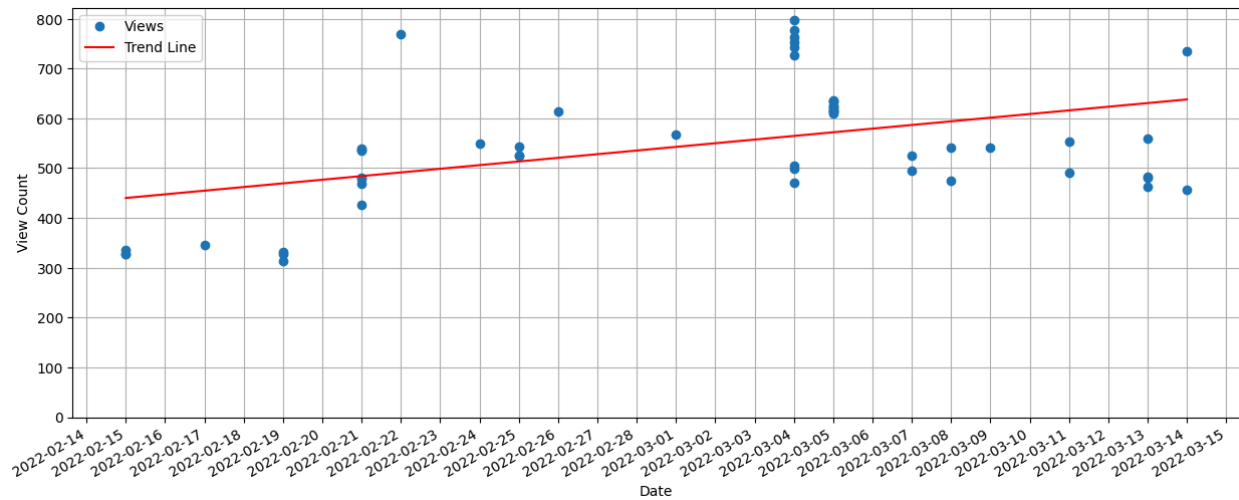
The message view counts do increase significantly (Trend slope: 7.33, P-value: 0.0006) over time. This is likely because messages that have been in the chat for longer would have more time to be viewed, leading to a higher view count. This significant increase could lead to correlation among the view counts. Figure 8 shows the changes in topic count over time.
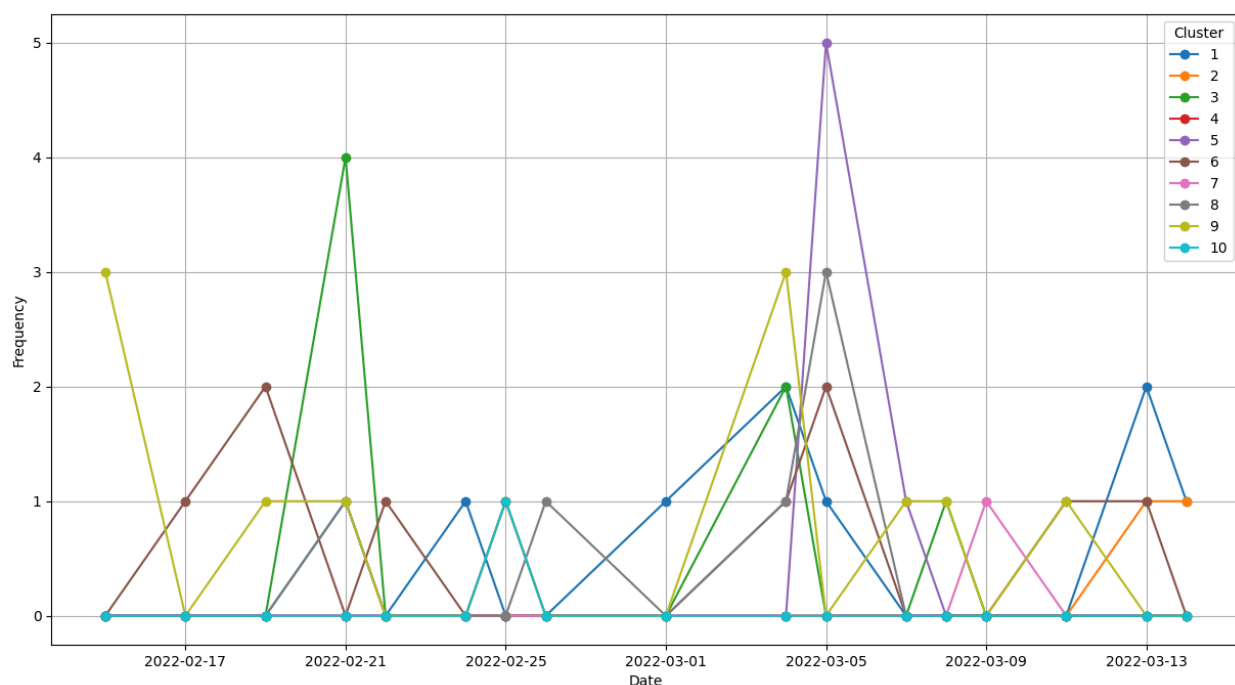
Figure 8

Text Topic Frequency Over Time
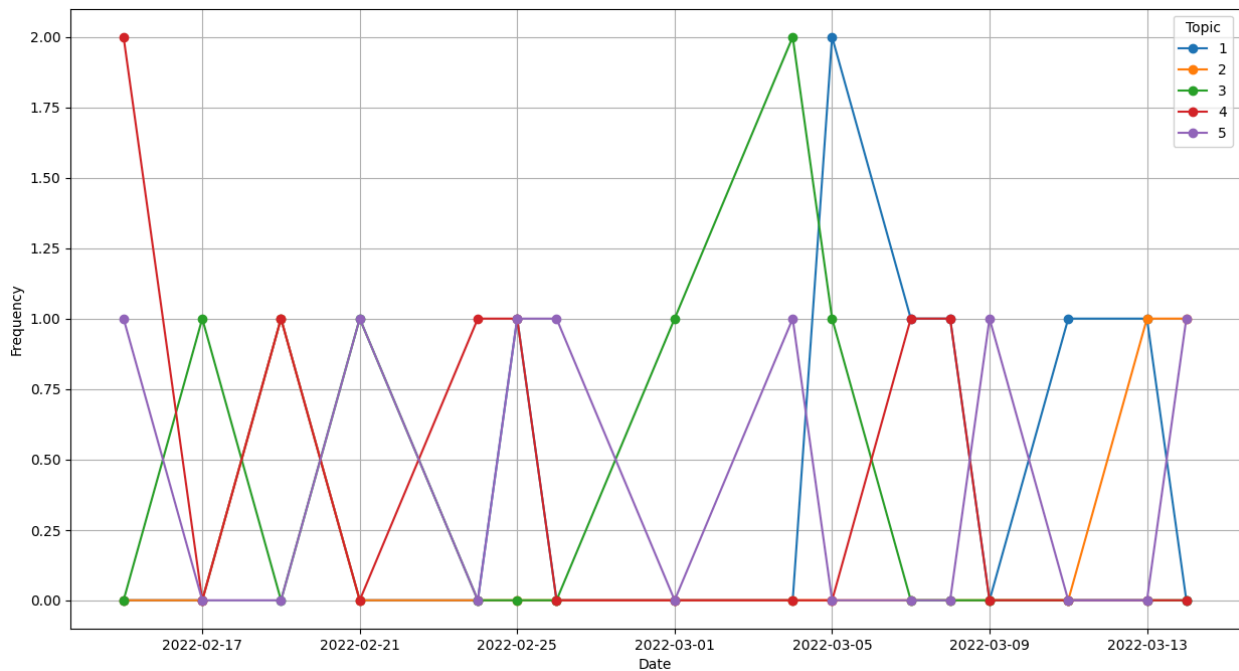
None of the changes in topic count over time constitute a statistically significant trend. Figure 9 shows the cluster count over time.

**Figure 9**

Image Cluster Frequency Over Time

The only statistically significant trend in cluster count frequency over time is for cluster 2 (Trend slope: 0.02, P-value: 0.02).

Regression Analysis

In order to test the relationship between the independent variables image cluster and text topic and the dependent variable view count I ran a mixed effects regression model. Integrating machine learning prior to running a regression has been discussed in prior work. I selected a mixed effects model to account for the significant autocorrelation among view counts at lags 1 and 2, which indicate that the data points are likely not independent. Figure 10 displays the results of the autocorrelation function. These results indicate that the messages may be clusters of related statistical units, which mixed-effects models are particularly suitable for. In order to run the analyses, I also had to limit my data to only include messages with an image and text (n=31).

Figure 10

Autocorrelation Function (ACF)
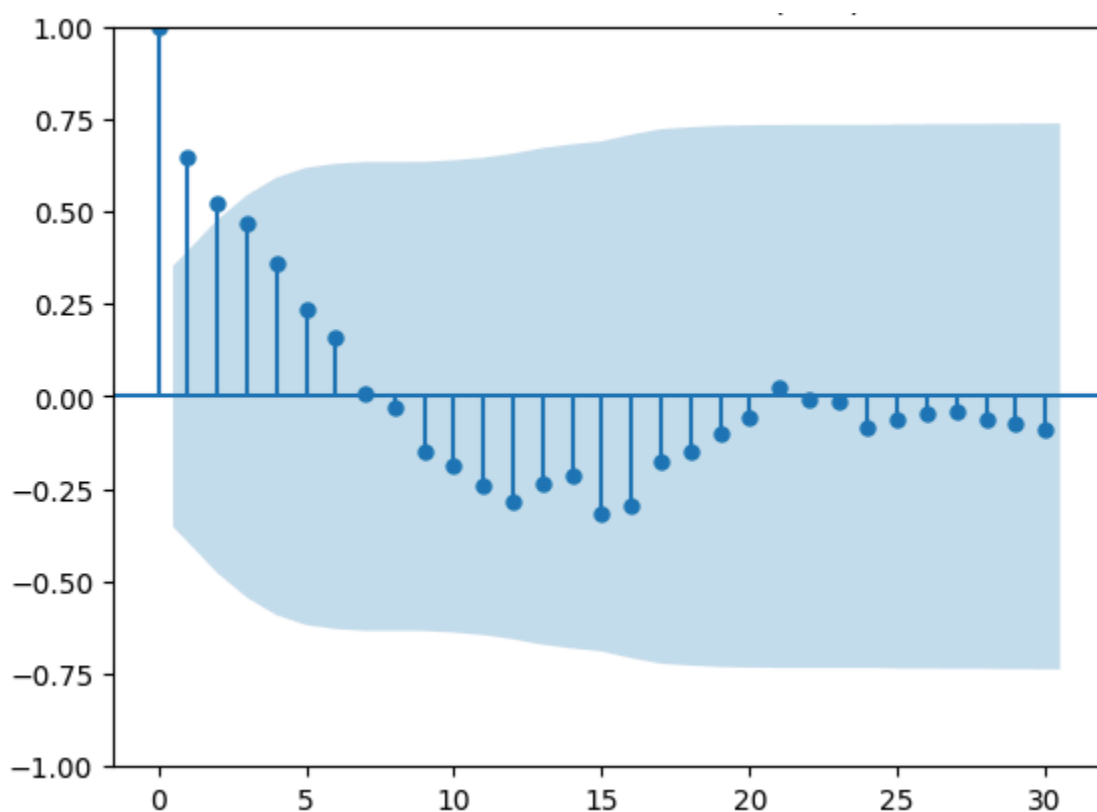
While my data does not meet the other assumptions of a linear mixed-effects model, likely due to the small size of the dataset, this type of regression still seems the most effective way of evaluating the relationship between image cluster, text topic, and view count.

The results of the linear mixed-effects regression model are displayed in table 3. Topic 1 and cluster 1 are excluded from these results as reference categories. The results should be interpreted as compared to topic 1, or cluster 1, depending on the variable.

Table 3

Linear Mixed-Effects Regression Model

Variable	Coef.	*Std.Err.*	*z*	Confidence Interval		*p*
				0.025	*0.975*	
Intercept	497.732	54.381	9.153	391.148	604.316	0.000
Topic 2: Celebratory Motor Rally	-32.903	44.160	-0.745	-119.455	53.649	0.456

Variable	Coef.	*Std.Err.*	*z*	Confidence Interval		*p*
				0.025	*0.975*	
Topic 3: Russian Nation	-9.280	35.918	-0.258	-79.678	61.118	0.796
Topic 4: Russian Military Deployment	24.658	36.478	0.676	-46.838	96.154	0.499
Topic 5: Russian Forces in Ukraine	69.216	40.695	1.701	-10.545	148.977	0.089
Cluster 2: Military Vehicles	164.216	74.130	2.215	18.924	309.509	0.027
Cluster 3: Crowds and Protests	-25.737	61.984	-0.415	-147.224	95.750	0.678
Cluster 4: Memorial Plaque	-19.750	72.821	-0.271	-162.478	122.977	0.786
Cluster 5: Old Building Photo/Illustrati on and Maps	11.679	49.547	0.236	-85.431	108.789	0.814
Cluster 6: Documents and Illustrations	4.531	37.832	0.120	-69.618	78.680	0.905
Cluster 7: Russian Foreign Ministry Spokeswoman	-54.179	79.082	-0.685	-209.176	100.818	0.493
Cluster 8: Buildings and Political/Religi ous Symbols	-6.428	45.565	-0.141	-95.733	82.877	0.888

Variable	Coef.	*Std.Err.*	*z*	Confidence Interval		*p*
				0.025	*0.975*	
Cluster 9: Antique Photographs of People	-56.091	35.999	-1.558	-126.649	14.466	0.119
Cluster 10: Military Truck	-88.967	70.551	-1.261	-227.245	49.311	0.207
Group Variable	9680.347	144.618				

Note. Number of Observations = 31, Number of Groups = 6, Min. Group Size = 2, Max. Group Size = 7, Mean Group Size = 5.2.

These results show that the only topic or cluster that significantly impacts view count is image cluster 2, Military Vehicles (p = 0.027). Messages with images in this cluster get approximately 164.216 more views than messages with images in cluster 1, People and Political/Religious Symbols, holding the other variables in the regression constant. However, there are also only two messages in the dataset with images in this cluster, therefore, these results may not be meaningful.

Conclusion

This exploratory project allowed me to apply various methods to understand the impact of the subjects of Telegram messages on how many views they received, providing insight into their impact on online discourse. While the size of the dataset, and its failure to meet regression assumptions, limits the conclusions that can be drawn from this paper, several methods do appear particularly useful for future research. These include using topic modeling and image clustering to quantify themes in messages. These quantitative values can then be used in further statistical analysis, such as the mixed effects regression model presented in this paper. Automatically extracting named entities also appears to be surprisingly useful for identifying place names and geographic locations in messages.

Future research could apply the methods used in this project to a larger dataset to see if the mixed effects regression model can identify themes that influence how many views a message received. Named entity extraction may also be useful for future geographically oriented projects that may wish to associate instances of online discourse with particular places.