Appendix 1: Methodology

To analyse tailored propaganda, we identified 12 clusters of VK users. These clusters represent the main socio-demographic groups in Russia and their media consumption patterns. In selecting VK users, we aimed to match the demographic characteristics of the Russian population based on Rosstat data. A similar methodology is used in the survey studies by the Chronicles project and the Levada Centre.

After a preliminary check of the sample for bots, we analysed the content in the VK communities most popular with the identified clusters. As an additional parameter, we used the level of user support for the war in Ukraine. We trained a separate model to determine this level.

Sampling Methodology

In constructing a sample that reflects the main socio-demographic groups of the Russian population, we relied on data from Rosstat: population numbers, age group distribution, and the number of women per 1,000 men in each age group.

Using this data, we ensured the sample's compliance with the socio-demographic composition of Russia by adhering to the following parameters:

- The proportion of regions in the country's population;
- The ratio of urban and rural populations in each specific region;
- The distribution of cities by size in each region;
- The age distribution nationwide;
- The ratio of men to women in each age category nationwide.

Settlements were grouped into the following categories for each region:

- Cities with a population of more than 1 million (all cities with over 1 million inhabitants are included in the sample as self-representative units);
- Cities with a population from 500,000 to 1 million;
- Cities with a population from 100,000 to 500,000;
- Urban settlements with a population of up to 100,000;
- Rural settlements and villages.

When selecting users, we set demographic characteristics in the search filters without filling in the search fields. In some cases, we specified a random day and/or month of birth. This was done to reduce the number of popular users (with many friends/subscribers) in the search results and ensure greater sample representativeness. In the search parameters, we specified gender, city, and age range (5-year intervals for ages up to 70 years, then one interval up to 100 years).

The additional filtering algorithm was applied at the following thresholds:

- If the search results included more than 50,000 profiles, we additionally specified a random day and month of birth;
- For 20,000 to 50,000 profiles — only a random day of birth;
- For 5,000 to 20,000 profiles — only a random month of birth.

In addition, we set an offset as a random number from 300 to the final digit of the search results to exclude the most popular users from the sample. In cases where the search results showed fewer than 1,000 users, no offset was set — considering the proportion of inactive profiles on VK, a result of 1,000 users is too small for additional filtering.

We considered only profiles where the users had logged into VK at least once in 2023 (profile collection was conducted in September 2023). Since the subject of analysis is the content in the communities to which VK users are subscribed, we included only users with 15 or more subscriptions in the sample. In total, 30,955 users meeting the above criteria were selected.

Determining Attitudes Towards the War

To determine attitudes towards the war based on VK community subscriptions, we trained a model using a sample of users who posted pro-war and anti-war content.

This additional sample was constructed from over 10,000 profiles based on the linguistic characteristics of two groups — supporters and opponents of the war. We identified war supporters by searching for phrases like “special military operation,” anti-Ukrainian pejoratives used in Russian propaganda (“Banderites,” “Khokhols”), and popular pro-war hashtags (#СвоихНеБросаем, #ZаНаших, etc.). Opponents were identified by phrases such as “Russian invasion,” the pejorative language they use against war supporters (“rashists,” “zombies,” etc.), and popular anti-war hashtags (#нетвойне, #stoprussianaggression, etc.).

As a result, we selected 16,383 users, of whom 6,752 used anti-war language, and 9,631 used pro-war language. After filtering out profiles where Russia was not indicated as the country of residence, 10,551 users remained, with 3,267 against the war and 7,284 supporting it. Naturally, many supporters and opponents of the war do not use VKontakte, so our sample is not a quantitative representation of all Russians. It reflects only the users of this specific social network who have publicly expressed their views on the war.

The ratio of anti-war to pro-war users may be skewed, as fewer anti-war posts are published on VK. VKontakte is an unsafe platform; its administration cooperates with law enforcement and state agencies, and large communities posing “inconvenient” questions are blocked.

After automatic searching, we manually reviewed 311 posts from the identified users to assess the proportion of correctly classified posts. We found that "pro-war" and "anti-war" language does not always literally reflect a user's position. For instance, some opponents of the war used the word “Banderites” ironically, highlighting their negative attitude towards propaganda language. Some posts did not concern the war in Ukraine at all — the phrase “Russian invasion” was used to describe other historical events.

Ultimately, about 67% of the posts classified as anti-war were genuinely anti-war. Among pro-war posts, this accuracy reached 96%. This proportion of correctly identified posts was sufficient for training the model.

In the next stage, we adjusted the sample so that the demographics in the groups of war supporters and opponents were the same. We selected war supporters who matched the size of the settlement, age, and gender distribution of the opponents. This reduced the number of war opponents in the sample to 2,753 and supporters to 1,626. The final sample consisted of 4,379 users. This helped to minimise the influence of other factors on user subscription choices.

We then constructed a matrix where users were placed in rows and communities in columns. Each cell in the matrix could take a value of 0 or 1, indicating whether the user was subscribed to the community. We reduced the dimensionality of the resulting matrix from the initial 11,110 (all communities in the sample with at least 20 users subscribed) to 100 using the TruncatedSVD method.

Using the resulting data, we trained a logistic regression model to determine whether a user supports the war. For the test sample not used in training, the model correctly predicted opposition to the war for anti-war users in 76% of cases. The average probability of opposition to the war for anti-war users in this sample was 58%. For pro-war users, the characteristics of predicting war support were 61% and 57%, respectively.

This model cannot accurately predict whether a user supports the war but can infer it based on the user's content consumption. It outputs a continuous value representing this characteristic, which allows for the inclusion of more neutral users. Logistic regression was chosen as the simplest and most interpretable model that can also correctly produce a continuous value.

To check the meaningfulness of our model, we extracted characteristic communities for supporters and opponents of the war. For war supporters, the top 10 communities were:

- UNITED RUSSIA (ЕДИНАЯ РОССИЯ)
- Government of Russia (Правительство России);
- State Duma (Государственная Дума);
- Valentina Matvienko (Валентина Матвиенко);
- SVO Reports | Army | Russia (Сводки СВО | Армия | Россия);
- Ramzan Kadyrov (Рамзан Кадыров);
- Kirill Zhigulin (Кирилл Жигулин);
- Supporters of «United Russia» (Сторонники «Единой России»);
- Ministry of Defence of Russia (Минобороны России);
- Poddubny |Z|O|V| (Поддубный |Z|O|V|).

For opponents, the top 10 communities were:

- Navalny's Team (Команда Навального);
- Alexei Navalny (Алексей Навальный);
- Ateo (Ateo);
- Mikhail Khodorkovsky (Михаил Ходорковский);
- TV Rain (Телеканал Дождь);
- OVD-Info (ОВД-Инфо);
- Yoshkin Krot (Ёшкин Крот);
- Ilya Yashin (Илья Яшин);
- POLITKUKHNYA OF THE CRIMINAL AUTHORITY (ПОЛИТКУХНЯ ПРЕСТУПНОЙ ВЛАСТИ);
- Anti-Corruption Foundation (FBK) (Фонд борьбы с коррупцией (ФБК)).

We applied the trained model to our representative sample of over 30,000 users and obtained for each user a degree of support for the war, expressed as the probability that the user is a war supporter based on logistic regression. The average value of this probability across the entire sample was 50.5%.

A notable limitation of our method is that the training sample is formed based on VK posts, which inherently considers only active users who express their views publicly. This approach excludes users who do not publicly express their position on social media. Additionally, a user's real-life stance may not align with what is declared on social media. Moreover, as noted earlier, the use of marker words does not always directly correlate with a position on the war.

However, we observe a logical correlation between a user's political stance and their political subscriptions. War supporters are predictably more often subscribed to United

Russia and pro-war bloggers, while opponents are more often subscribed to Alexei Navalny and other opposition politicians.

Clustering Users

The main hypothesis of this study is that a significant amount of information about a user's media consumption can be extracted from their subscription data. Therefore, the primary stage of the research is clustering users based on their subscriptions.

To do this, we constructed a user matrix in the same way we did for determining attitudes towards the war: rows represent users, columns represent communities, and each cell can take a value of 0 or 1, indicating whether the user is subscribed to that community. We only considered communities that had 10 or more subscribers from the sample.

For clustering, we reduced the dimensionality from 46,395 (the number of all communities with 10 or more subscribers) to 100 using PCA, and then applied k-means clustering with 50 clusters. We then built a hierarchy of clusters using the linkage method, based on their centroids. Clusters with fewer than 100 users and/or a narrow topic of interest were filtered out. Similar clusters were merged based on this hierarchy. This resulted in 12 clusters corresponding to distinct socio-demographic groups.

Another thirteenth cluster included about half of all users who could not be categorised. These constituted 49% of the sample. Likely, users in the uncategorised cluster did not belong to any defining community for the 12 clusters (or could belong to multiple clusters).

In terms of their demographic parameters and overlapping subscriptions, the users we selected into the 12 thematic clusters almost perfectly match the rest of the sample. Therefore, the clustering likely correctly reflects the main social groups characteristic of the entire sample. The demographic parameters of the sample and the clustered users are presented in the graphs below. Among the clustered users, 59% were women, while in the entire sample, 54% were women.

After reducing the dimensionality to 100 using PCA, we applied UMAP to project the results onto a two-dimensional plane.

Distribution of users by settlement size

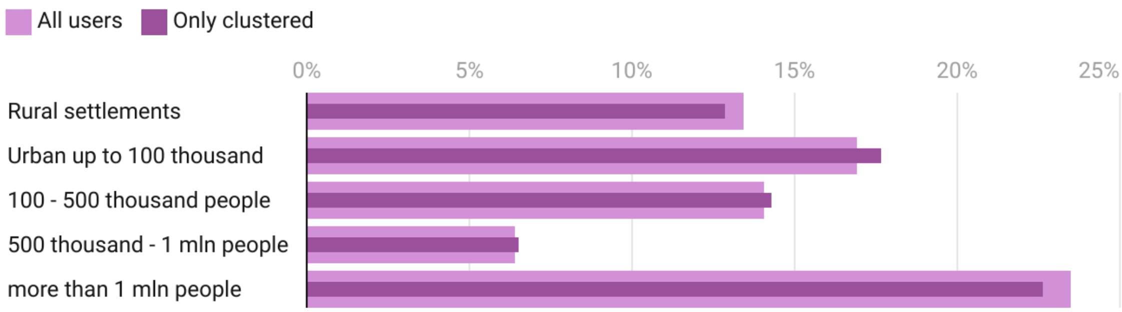


Chart: Cedar • Source: Editorial calculations based on vk.com data • Created with Datawrapper

Distribution of users by age

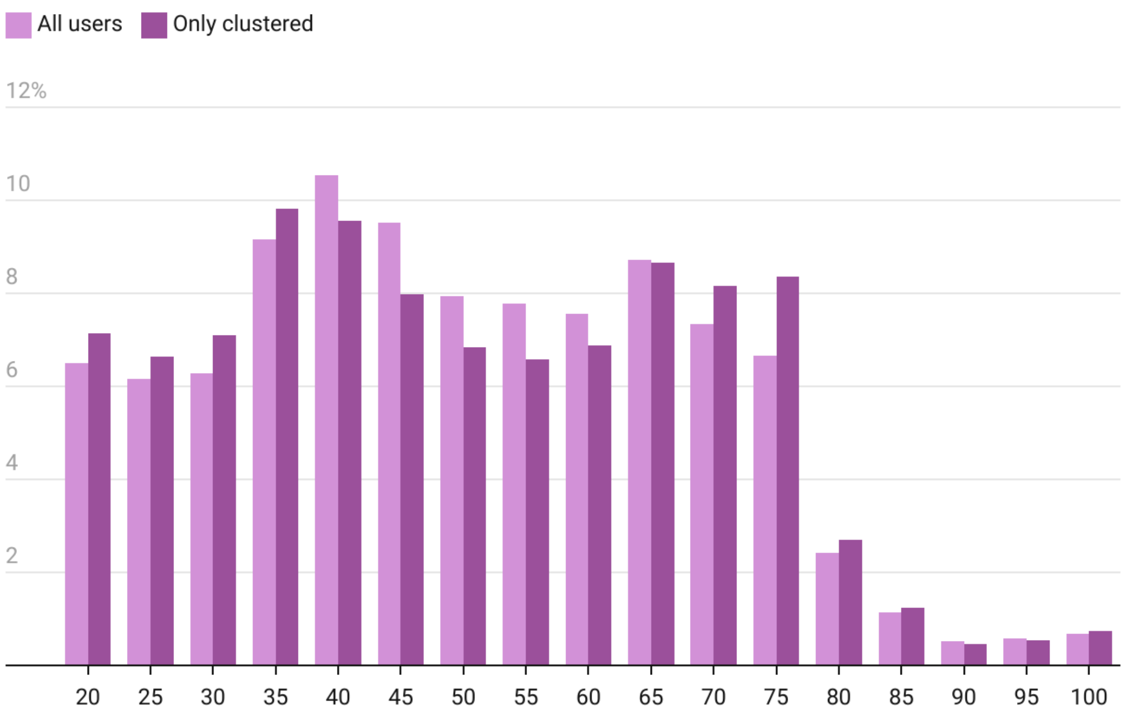


Chart: Cedar • Source: Editorial calculations based on vk.com data • Created with Datawrapper

Bot Detection

Comments Bots

To verify our sample for the presence of pro-government bots, which typically leave comments on posts, we utilised the database of the Botnadzor project, which monitors bots on VKontakte. According to Botnadzor, our sample contained 34 bots—0.1% of all pages.

Botnadzor checks about 2,500 different communities on VK daily, collecting information on comments to all posts made on these pages. This list includes news and political pages (e.g., “RIA Novosti”, “TASS”, “RBC”, “Lentach”, “PostNews”), regional groups and media, as well as other groups popular among bots (e.g., the “Overheard” network of pages). Additionally, Botnadzor monitors posts on news and political topics not related to a fixed list of pages: the war in Ukraine, the Arab-Israeli conflict, news about drone strikes, and more.

The project does not disclose the detailed methodology of bot detection, but it is known that bots are identified based on similar activity patterns, after which most bot accounts are validated manually. The activity patterns considered by Botnadzor include the number of comments written, the groups where they are posted, the parameters of the posts to which comments are made, the registration dates of the pages (bot accounts are often registered in batches), and other criteria.

We independently verified the validity of some of these criteria using all the comments downloaded by Botnadzor from 102,955 pages over a week—1,651,993 user comments made from 804,001 accounts and 102,533 bot comments (from 7,623 accounts).

For a conceptual check of the correctness of Botnadzor’s classification, we clustered all accounts that wrote comments based on the groups where they posted them. We used a clustering method analogous to our user audience clustering in this study (described above). We created a matrix where rows represent users, columns represent the pages where they posted comments, and each cell contains the number of comments a specific user made on a specific page. To simplify the analysis, we excluded pages with fewer than four comments per week and accounts that wrote fewer than five comments. This resulted in 39,814 user accounts, 3,551 bots, and 11,601 pages.

We reduced the dimensionality of the resulting matrix from 11,601 to 2 using the UMAP method (n_neighbors = 5). On the resulting projection (where each point represents one account), it is clear that bots (marked in red) are almost exclusively located in one cluster, while almost all real users (marked in green) are outside this cluster.

This visualisation shows that bot activity is concentrated in a narrower range of pages than that of regular users. Additionally, bots leave roughly the same number of comments, unlike real people. On average, real users in the sample wrote two comments per week, while bots wrote 13.5. Thus, the verification results confirm the validity of identifying bots based on their activity patterns used by Botnadzor.

Clustering of accounts of users who left comments

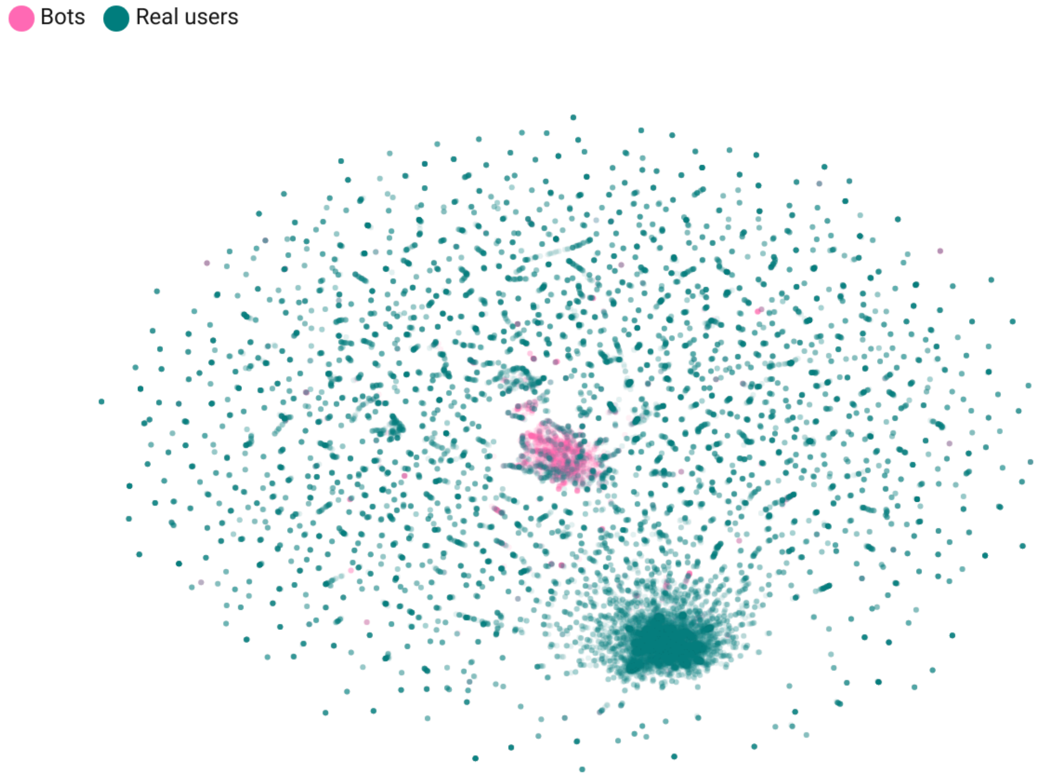


Chart: Cedar • Source: Editorial calculations based on vk.com and "Botnadzor" data • Created with Datawrapper

Subscriber Bots

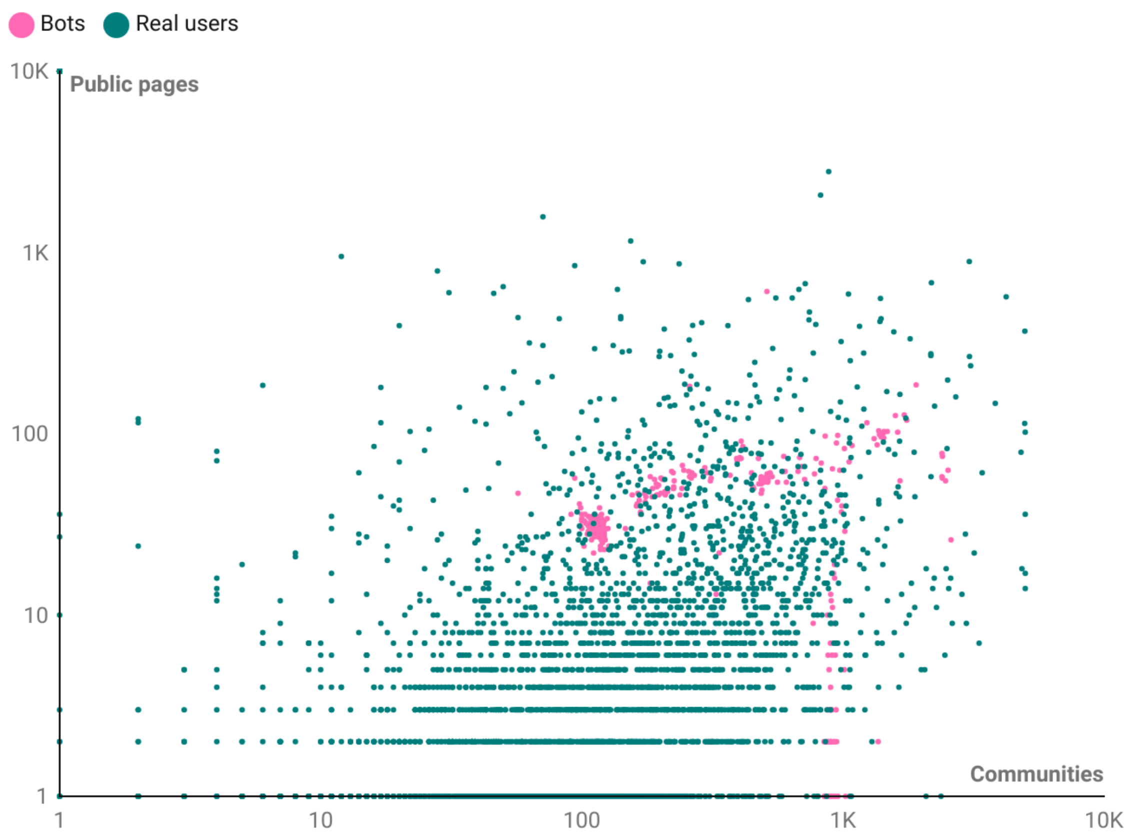
A separate type of bots is used to inflate the number of subscribers in communities on VK. These accounts do not leave comments, so most of them are not in the Botnadzor database. There should be significantly more of such bots on VK, as subscribers in communities are inflated by tens of thousands for a small fee.

To identify these bots, we conducted an experiment: we created a fake community on VK and purchased 1000 bot subscribers to determine their characteristics. Most of these bots were banned within the first few hours, and ultimately, 443 bot subscribers remained available for analysis. As a training dataset for the bot subscriber detection model, we added a sample of 10,000 random real users who commented on the same posts as the bots. We then collected quantitative parameters from the profiles of 10,000 users and 443 bots: number of photos, audio recordings, subscribers, subscriptions, groups, and so on, a total of 13 parameters. Each of these parameters was used as one of the coordinates.

To interpret this method, we can look at the results for two of the 13 coordinates: the number of subscriptions to communities and groups. Communities and groups are different types of pages on VK, but functionally they are not significantly different. In this section, we distinguish between them to illustrate bot activity patterns; in the rest of the text, the terms “public page,” “group,” and “community” will be used as synonyms.

It is evident that in these coordinates, bots are quite tightly grouped (due to the artificial nature of filling profiles), but the coordinate values are not extreme, so it is not always possible to determine at a glance whether a profile is a bot.

Number of user communities and public pages



Logarithmic scale

Chart: Cedar • Source: Editorial calculations based on vk.com and "Botnadzor" data • Created with Datawrapper

The dimensionality of the profile parameter space was reduced from 13 to 5 using UMAP (n_neighbors = 15), and clustering was performed in this space using HDBSCAN. Clusters, where more than 50% of the accounts were bots, were considered bot clusters, and this result was used to predict whether an account was a bot. For visualisation purposes, the dimensionality was further reduced to two using TSNE.

Result of the model training

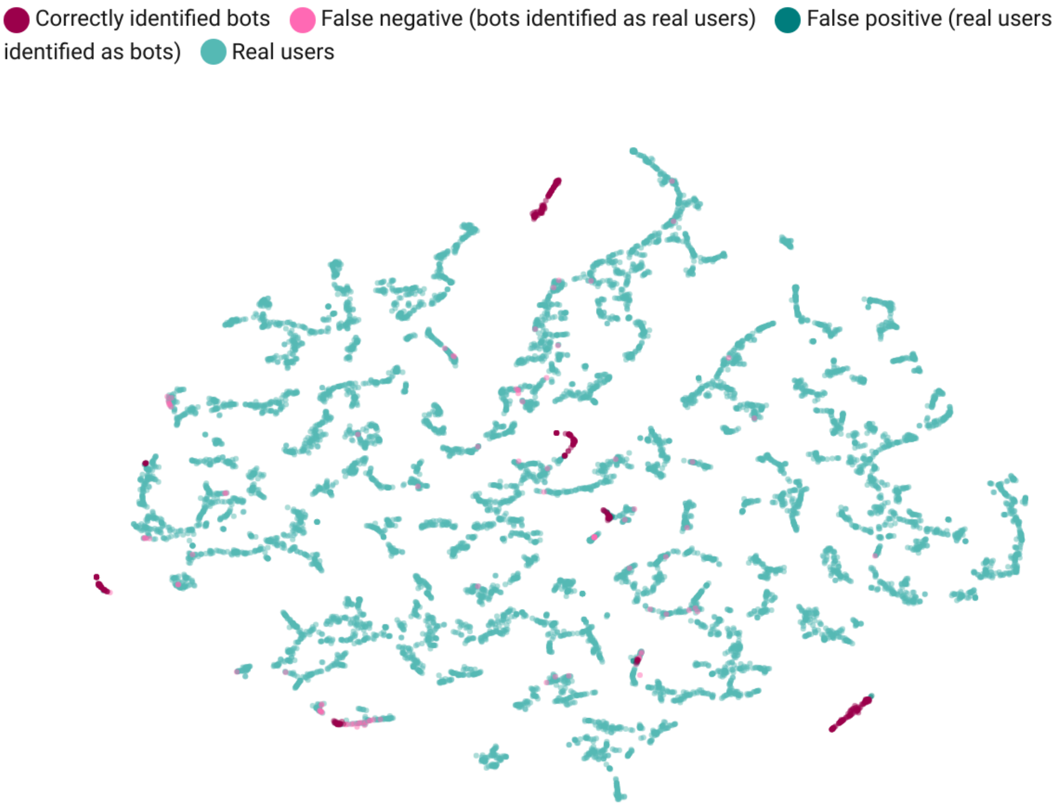


Chart: Cedar • Source: Editorial calculations based on vk.com and "Botnadzor" data • Created with Datawrapper

The precision of the method on the training dataset was 93%, sensitivity (recall) was 79%, and specificity was 99.7%. However, a limitation of our method is that we cannot guarantee the absence of bot groups with different profile completion patterns.

The trained model was applied to the main user dataset used in the study. To visualise the model's performance, we plotted this dataset in the same coordinates as the training dataset.

The model predicted that in the collected representative sample, there are 0.9% bot-subscribers, totalling 258 accounts. This proportion varies across different clusters. The maximum number of bots was found in the "z-patriots" cluster, amounting to 2.4%. Such a number of bots should not affect the analysis results.

Predictions of the model on the selection

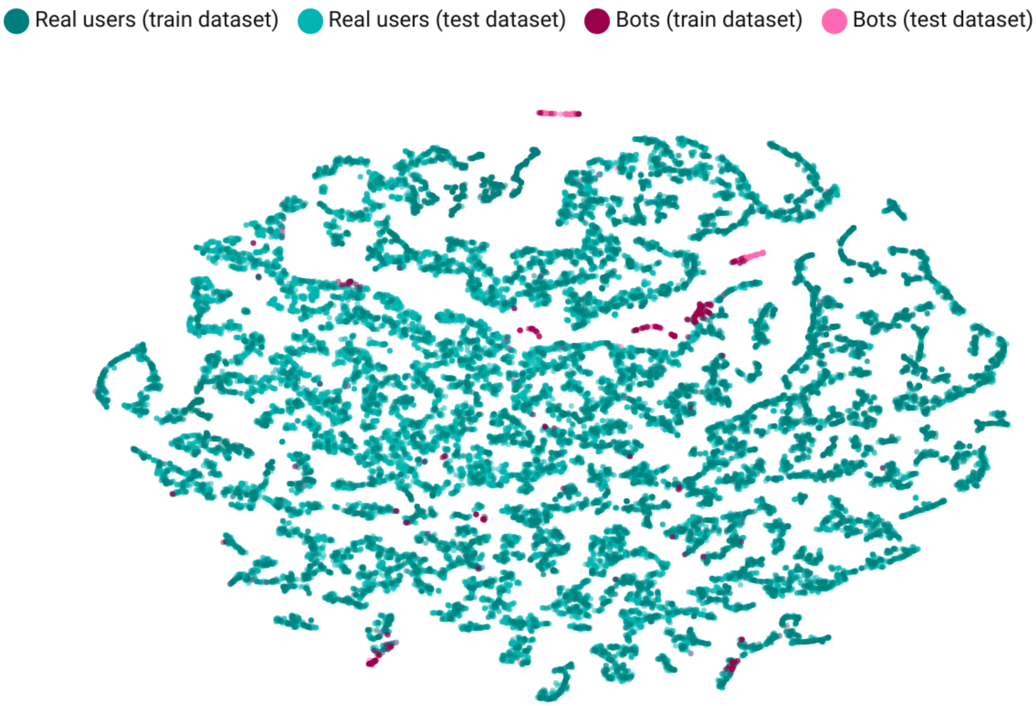


Chart: Cedar • Source: Editorial calculations based on vk.com and "Botnadzor" data • Created with Datawrapper

Share of bots in user clusters

predicted by the model

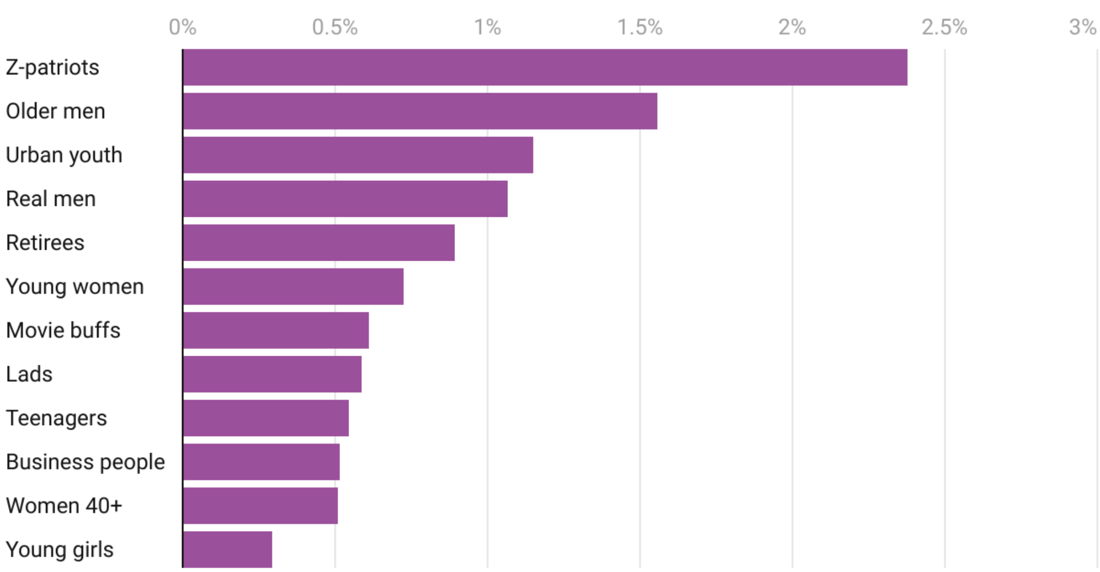


Chart: Cedar • Source: Editorial calculations based on vk.com data • Created with Datawrapper

A relatively small proportion of bots in our sample may be due to the fact that VK bans bot-subscribers quickly, resulting in few remaining searchable bot accounts. Out of the 1000 subscribers we purchased, only 56 remained after 3 days. The rest were banned by the social network's administration.

Media Consumption Analysis

To analyse user media consumption, we extracted the last 500 posts from each of the top 20 communities for each cluster, totalling 10,000 posts per cluster.

A significant portion of VK content consists of images with embedded text. Among the extracted posts, 20% contained no text at all, 90% included photos or videos, and 54% contained images with embedded text. This necessitated the extraction of text from images using the pytesseract module in Python, which was then analysed alongside textual content from the posts.

For topic modelling, we employed BERTopic. Approximately half of all posts couldn't be categorised. After filtering out political posts (which were analysed separately), we calculated the proportion of each non-political topic in the dataset.

Identification of Political Content

To analyse content, we selected posts containing more than 10 words (including text extracted from images), which constituted 57% of the entire dataset.

Identification of political content proceeded through several stages:

1. Zero-shot classification using NLI to filter out some promotional posts (identified by the word "sales").
2. Checking for the presence of keywords like "Putin" and "Ukraine" in any form. Posts containing these keywords were automatically classified as political.
3. For posts without these markers, another round of zero-shot classification using NLI was performed to check for affiliations with categories such as "politics," "Ukraine," and "combat operations."

Initially, 13% of posts were classified as political, but manual verification revealed that not all of them were political. Therefore, all political posts were additionally checked using the OpenAI API:

1. ChatGPT was used to summarise each post's content and identify words related to military operations, wars, politics, government structures, states, legislation, and ideology.
2. Summarised messages were further analysed by ChatGPT to determine if they contained any words related to military operations, wars, politics, government structures,

states, legislation, and ideology. Posts with affirmative responses were classified as political.

3. For negative responses, an additional question was posed: "Does this post include any (even indirect) mentions or expressions of political support, patriotism, or protests?" This was necessary to further filter cases where ChatGPT did not classify messages expressing support for Russia as political.

Ultimately, 9% of posts containing more than 10 words were classified as political, totalling 6691 political posts. It's worth noting that this method of identifying political posts provides a conservative estimate, as some political posts may not be identified due to insufficient contextual understanding by the models.

For a detailed analysis of tailored propaganda, we selected communities with at least 10 political posts (out of 500 original posts) identified by our method. From these, we chose only entertainment groups where news wasn't the main theme based on the title and description of the community. Opposition community “Lentach” and narrow-topic communities (about weapons and military actions) were excluded because almost all posts in them were classified by the model as political, despite not being propaganda in reality.

Thus, there were 30 entertainment communities with tailored propaganda. From these selected entertainment communities, we downloaded 3000 posts each and identified political posts using the aforementioned algorithm.

This specific dataset was used to analyse propaganda content. For this purpose, we applied BERTopic again. In this case, 92% of posts were successfully classified, while the topic remained undefined for the remaining 8%.