

ВК **является** самой популярной социальной сетью в России. Больше всего она пользуется популярностью у сторонников текущего политического режима, вследствие своих очевидных связей с государством. Поэтому наличие большого количества провоенных постов, отражающих нарративы пропаганды, не является чем-то неожиданным.

Но насколько в действительности мнение рядовых пользователей отражает нарративы, которые мы видим в российских СМИ? Какие аргументы являются наиболее важными для людей, поддерживающих войну? И остались ли в ВК люди, выражающие антивоенную позицию?

Понимать, как люди выражают свою позицию по поводу войны и какие аргументы их убеждают больше всего, важно не только из исследовательского интереса, но и чтобы составлять эффективные стратегии борьбы с пропагандой и дезинформацией. Мы попытались найти наиболее популярные провоенные и антивоенные нарративы с помощью машинного анализа текстов, которые пишут пользователи в ВК.

Методология

Код здесь: https://github.com/AlesyaSokol/VK_and_war

Датасет предоставляю по запросу

Датасет

Я искала посты сторонников и противников войны с помощью API VK по хэштегам.

- Антивоенные: '#нетвойне', '#нет_войне', '#nowar', '#stoprussianaggression'
- Провоенные: '#своихнебросаем', '#занаших', '#за_наших', '#запобеду'

Всего удалось найти 32830 антивоенных и 489600 провоенных постов, размещенных на стене пользователей. Я выбрала только те посты, у авторов которых в качестве страны проживания указана Россия.

На хакатоне мы:

- Убрали посты, являющиеся репостами
- Убрали абсолютно идентичные посты, даже если они опубликованы разными пользователями (оставили только один из всех идентичных)
- Оставили только посты, опубликованные с 1 февраля 2022 по 1 марта 2023
- Сделали предобработку текста постов:
 - Удалили хэштеги
 - Удалили ссылки
 - Удалили пунктуацию
 - Выполнили леммантизацию с помощью `mystem`
- Убрали посты, в которых после предобработки осталось меньше 6 слов

После этого осталось 20448 антивоенных и 248313 провоенных поста.

Тематическое моделирование

Я использовала предобученную ML-модель [BERTopic](#) для того, чтобы разбить посты на группы по темам.

Я сделала 3 отдельных датасета и обучила 3 отдельных модели:

- Со всеми провоенными постами - чтобы посмотреть нарративы сторонников войны и их динамику во времени
- Со всеми антивоенными постами - чтобы посмотреть нарративы противников войны и их динамику во времени
- Со всеми антивоенными постами и 20448 случайными провоенными постами - чтобы сравнить нарративы сторонников и противников войны. Я сделала количество постов сторонников и противников равным, потому что иначе нарративы противников войны терялись из-за того, что провоенных постов сильно больше.

При повторных запусках обучения разбивки по темам часто не полностью совпадают. Вероятно, это не значит, что что-то не так - просто некоторые темы очень близки по смыслу, и провести границу между ними не так легко. Поэтому неудивительно, что при разных запусках это происходит немного по-разному.

После обучения большая часть постов попадает в "-1" тему - это те посты, которые модель не смогла классифицировать. Туда попадает около 50-60% постов. Можно заставить модель все же раскидать их по темам - это получается не очень плохо. Для этого я использовала функцию `reduce_outliers` с параметром `threshold = 0.3`. В некоторых случаях использован `threshold = 0.2`. После этого в "-1"-й теме остается около 5-10% постов.

В каждом случае определяется несколько сотен (400-500) тем. Автоматически объединять их в более широкие темы получается не очень хорошо. Поэтому я объединила похожие темы вручную, и в итоге в каждом из трех случаев осталось 100-300 тем. На графиках ниже показаны наиболее популярные из них.

Чтобы описания тем (ключевые слова) были более читаемыми, я добавила туда популярные слова и словосочетания в рамках каждой темы (с максимальным `tf idf`). Местами описания после этого немного подкорректированы вручную (изменены грамматические формы, удалены лишние слова).

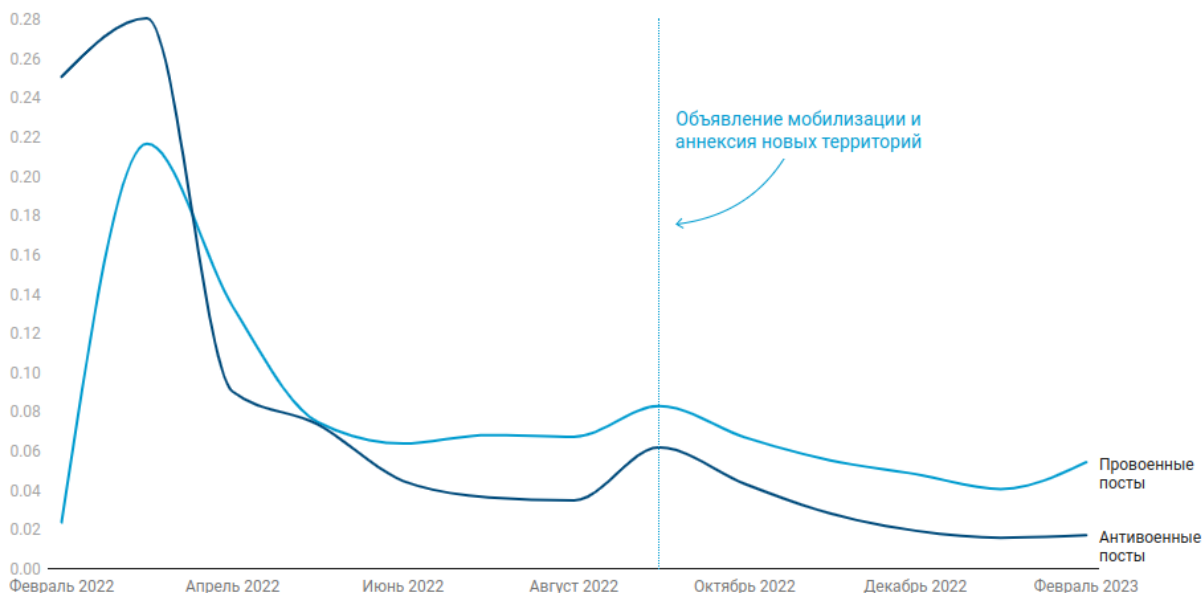
Не все из полученных тем имели смысл (выглядит, что модель иногда объединяла в темы посты, не связанные по смыслу). Темы, в которые, по-видимому, попали случайные и не очень связанные по смыслу посты, на графиках не показаны.

Оказалось, что в некоторых темах значительный вклад в посты вносит какой-то один пользователь. Есть люди (и среди сторонников, и среди противников войны), которые написали на своей стене порядка 100 очень похожих, но не абсолютно одинаковых постов. В связи с этим посты пользователей, которые написали более 15% постов в какой-либо теме, из этой темы были удалены уже после тематического моделирования.

Результаты

Сторонники войны

Изменение количества постов со временем

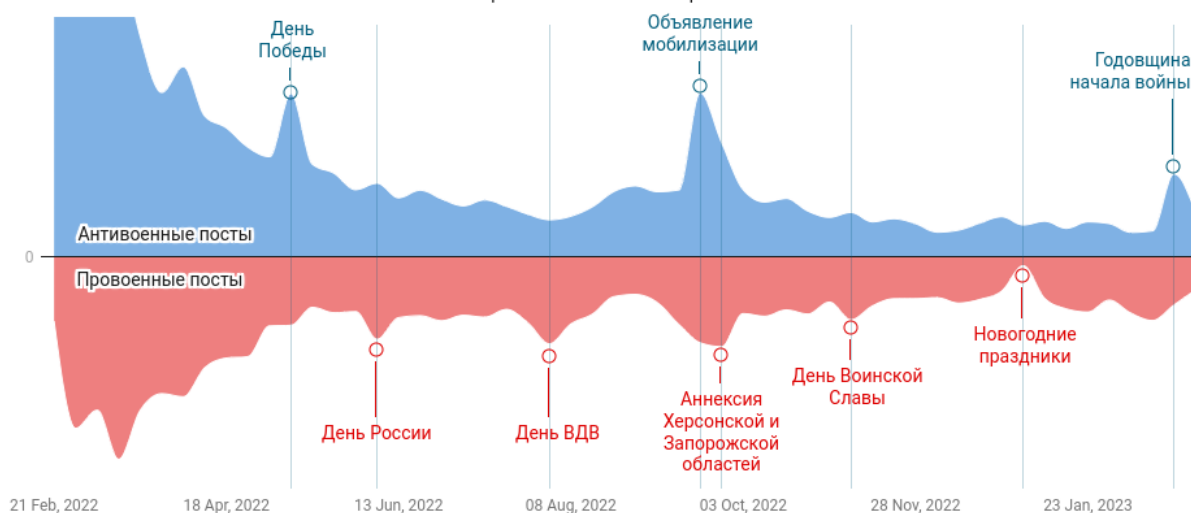


Интерактивный график: <https://datawrapper.dwcdn.net/i2jIT/1/>

По горизонтали месяца с февраля 2022 по февраль 2023. Показано относительное количество постов (процент от общего числа постов).

Динамика количества постов

Относительное изменение количества антивоенных и провоенных постов со временем



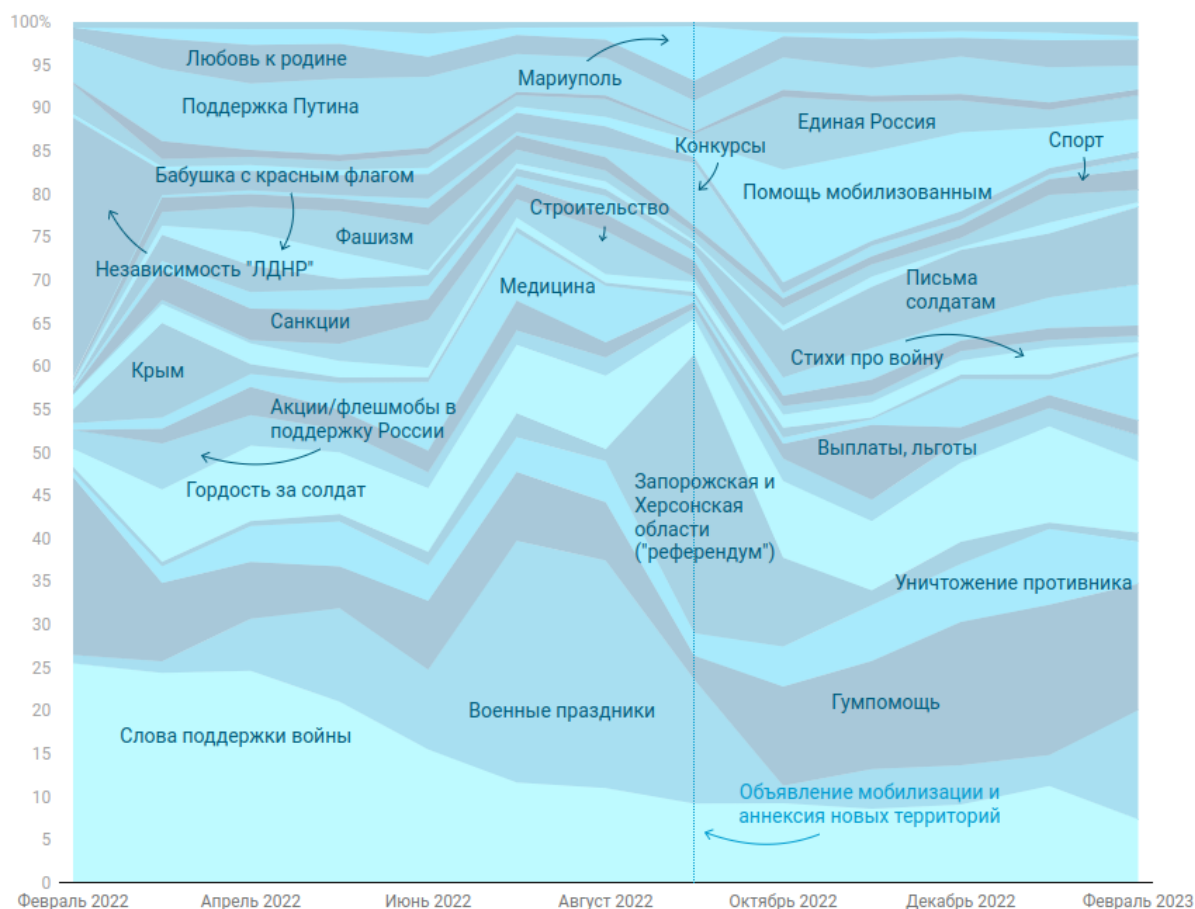
Здесь посты отнормированы на медианное значение.

Интерактивный график: <https://datawrapper.dwcdn.net/2KZe5/1/>

Интересные наблюдения

- Антивоенные посты начали в большом количестве появляться сразу после начала войны (в феврале 2022), в то время как провоенные - только в марте 2022
- Со временем антивоенных постов становится меньше быстрее, чем провоенных (что может быть связано со страхом выражать свою позицию в ВК)
- Мобилизация и аннексия новых территорий - единственный момент, когда было заметное значимое увеличение количества постов о войне (как провоенных, так и антивоенных)

Нарративы сторонников войны

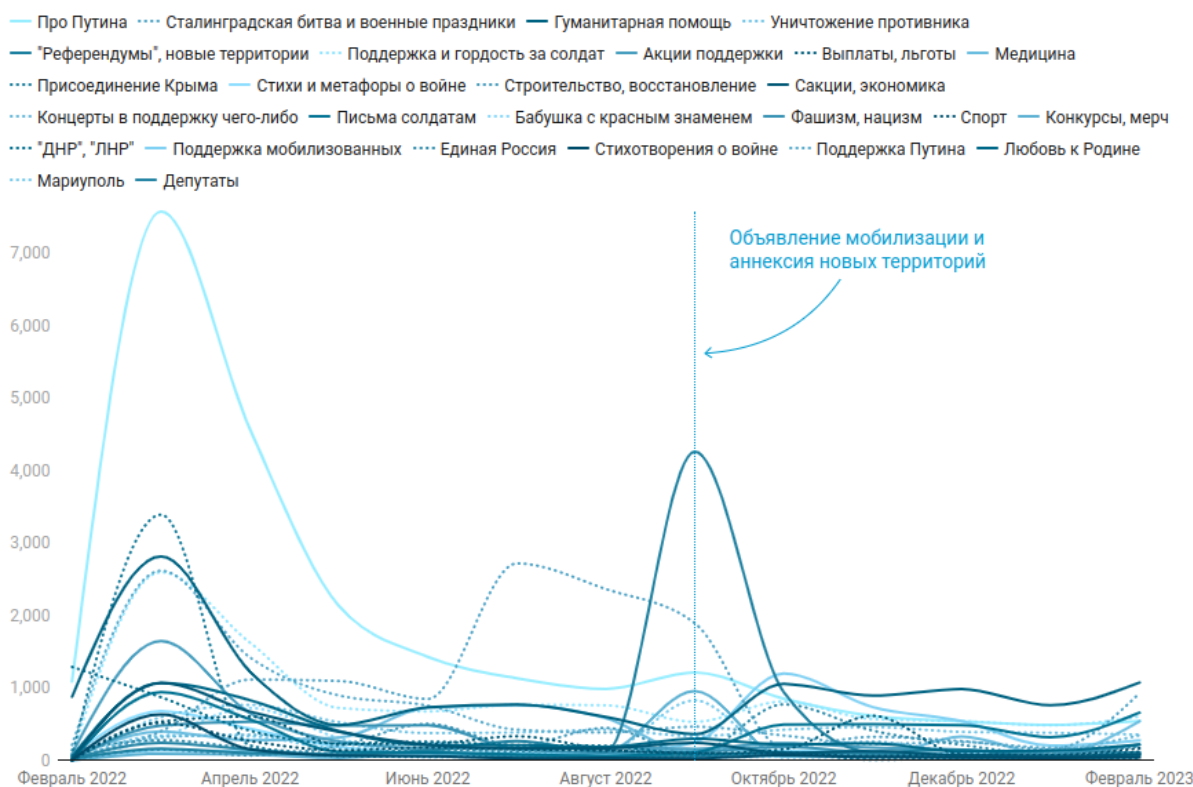


Интерактивный график: <https://datawrapper.dwcdn.net/rTuLR/6/>

Здесь за 100% взято количество постов, опубликованных в данный месяц, попадающих в самые популярные темы. Нужно учитывать, что 100% - это не вообще все посты, опубликованные в этот месяц - те, что не попадают в популярные темы, в анализ не включены.

Нарративы сторонников войны

По вертикальной оси количество постов из данной категории, по горизонтальной - месяца с февраля 2022 по февраль 2023



Интерактивный график: <https://datawrapper.dwcdn.net/vPwpw/7/>

Примеры постов

🇷🇺 Провоенные темы

В таблице приведены 3 случайных примера постов из каждой категории. Категория не всегда определяется корректно, поэтому некоторые посты неидеально соответствуют теме.

Интересные наблюдения

- Количество постов про письма на фронт и гумпомощь заметно увеличилось после начала мобилизации
- Гордость за солдат и любовь к родине со временем практически не изменились
- При этом количество постов с поддержкой Путина несколько уменьшилось
- "Референдумы" о присоединении Херсонской и Запорожской областей были значимыми инфоповодами лишь в течение двух месяцев, потом люди перестали об этом говорить
- Мобилизация была важной темой постов, и этот эффект все еще не сошел на нет - люди до сих пор о ней пишут
- Про фашизм больше всего писали с марта по июнь 2022, потом количество постов резко уменьшилось и с тех пор не поднималось

Аргументация сторонников войны

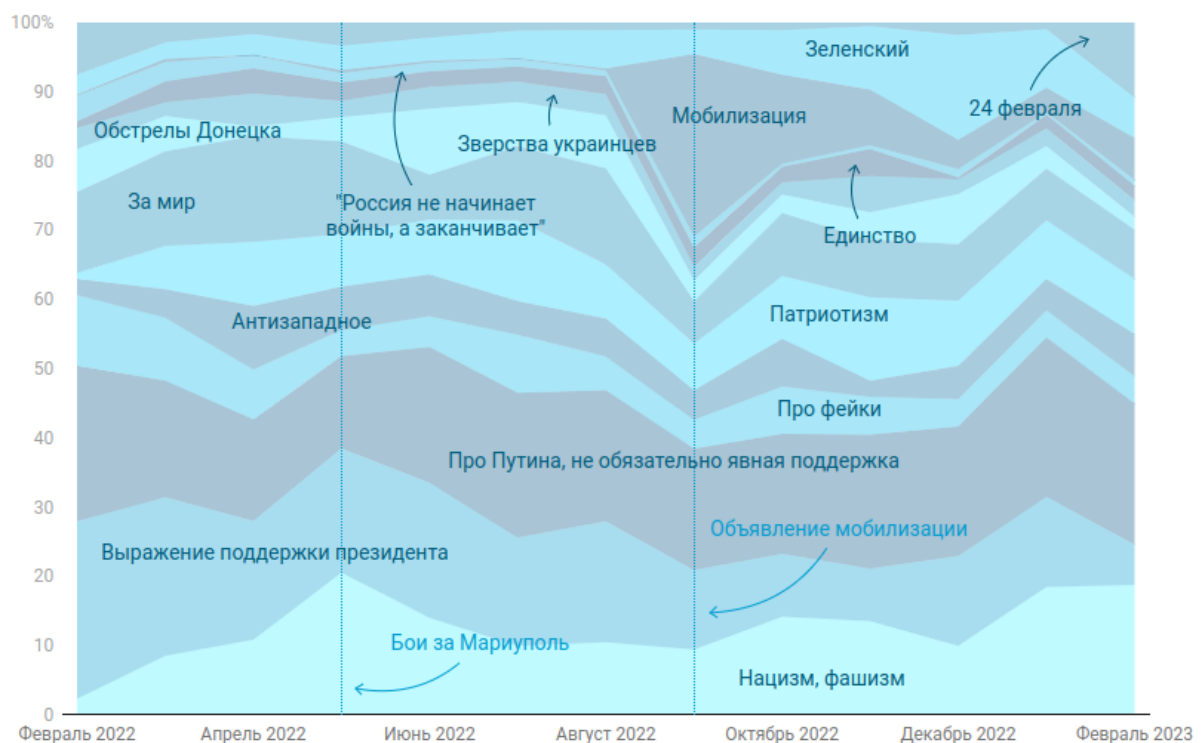
Большую часть провоенных постов составляют посты с описанием некоторых событий или фактов (сбор гумпомощи, "референдумы" в Херсонской и Запорожской областях и т.д.). Интересно выделить те посты, которые содержат непосредственно слова поддержки войны и аргументацию этой поддержки. Я собрала такие посты в отдельный датасет, в них вошли те посты, которые были классифицированы как:

- Слова поддержки войны
- Посты про фашизм
- Слова поддержки Путина
- Патриотизм, любовь к родине

Таких постов оказалось 38875.

Я выполнила обучение отдельно на этом датасете. Как и раньше, после автоматического определения тем, я сгруппировала некоторые темы вручную и убрала бессмысленные. Результаты:

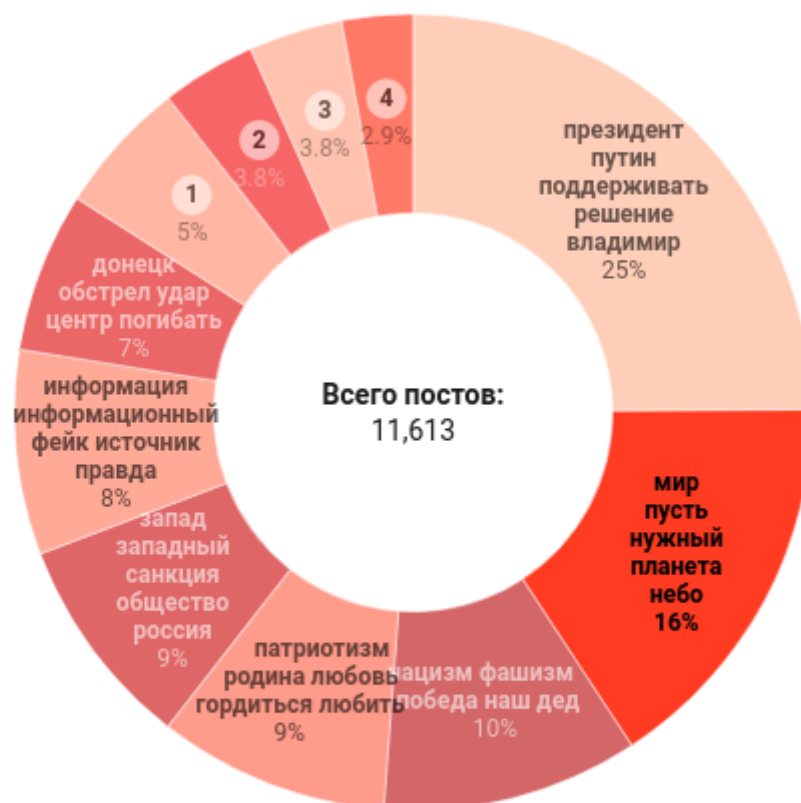
Нарративы постов о поддержке войны



Интерактивный график: <https://datawrapper.dwcdn.net/pAlh3/11/>

Здесь все еще много описаний фактических событий, которые не являются аргументацией поддержки войны (а именно, темы про мобилизацию, 24 февраля и Зеленского). Если их убрать, то получается, что со временем нарративы почти не меняются. Можно сравнить в абсолютных числах, каких нарративов больше. На диаграмме указаны не все темы, а только самые популярные.

Аргументация поддержки войны



- 1 акция поддержка поддерживать проходить участие
- 2 единый вместе народ единство друг
- 3 видео украинский фейк буча тело
- 4 война донбасс не заканчивать житель

Интерактивный график: <https://datawrapper.dwcdn.net/u3gnL/1/>

Здесь 2 темы про Путина я объединила в одну

Примеры постов

📌 Посты - аргументация сторонников войны

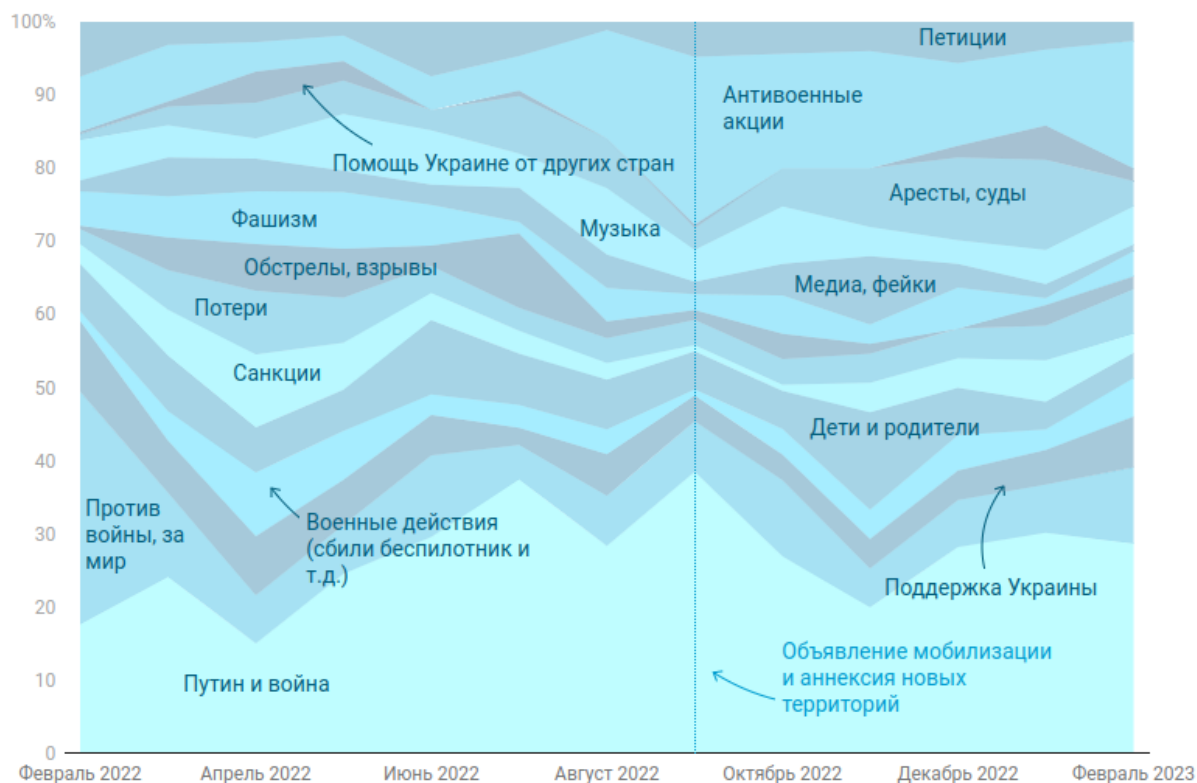
Интересные наблюдения

- Самым популярным нарративом в постах о поддержке войны является поддержка Путина
- Среди аргументов поддержки войны самый популярный аргумент - борьба за мир
- Следующий по популярности аргумент - борьба с фашизмом и нацизмом
- Защита Донбасса входит в список популярных аргументов о поддержке войны, но не является преобладающим аргументом
- Нарративы поддержки войны не сильно меняются со временем

Противники войны

Нарративы противников войны

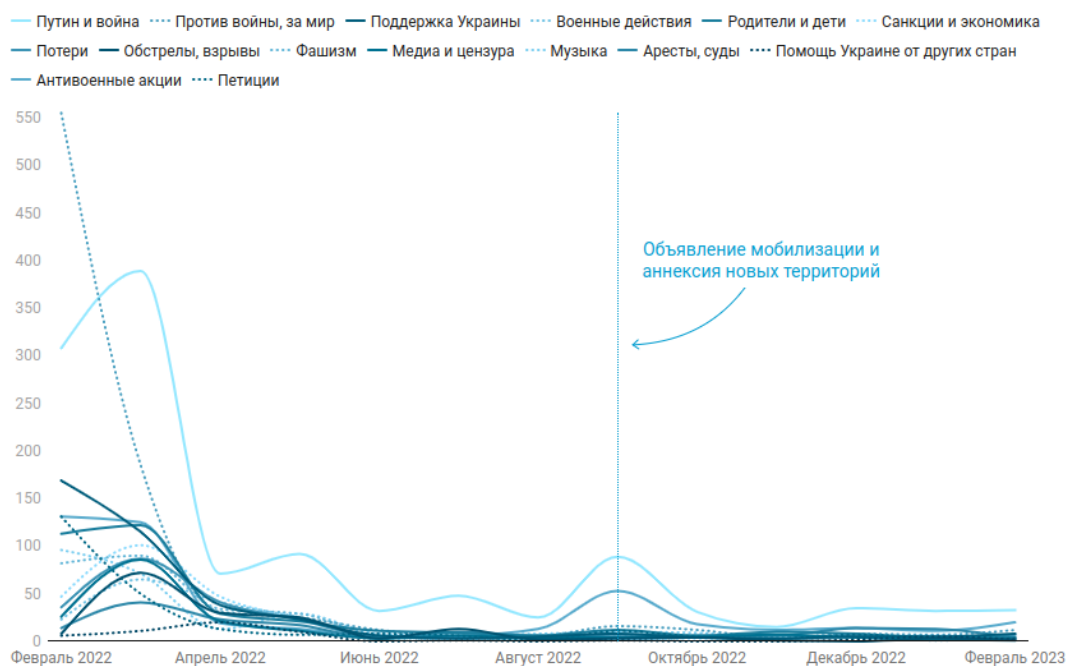
Динамика тем антивоенных постов во времени. По горизонтальной оси месяца от февраля 2022 по февраль 2023



Интерактивный график: <https://datawrapper.dwcdn.net/DpTV9/8/>

Нарративы противников войны

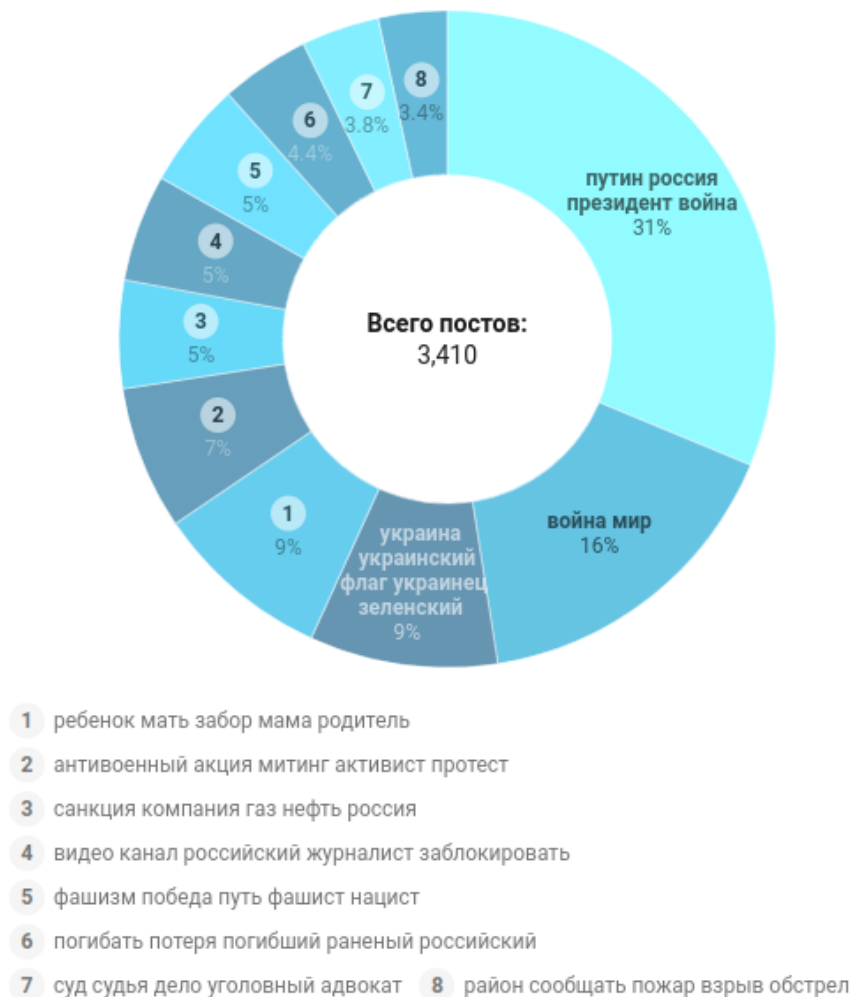
По вертикальной оси количество постов из данной категории, по горизонтальной - месяца с февраля 2022 по февраль 2023



Интерактивный график: <https://datawrapper.dwcdn.net/VJXP3/12/>

Если убрать темы с описанием фактических событий, а не аргументацией (военные действия, петиции, поставки оружия Украине), и оставить только то, что похоже на аргументы, получится так:

Аргументация неподдержки войны



Интерактивный график: <https://datawrapper.dwcdn.net/zd7um/2/>

Примеры постов

🚩 Антивоенные темы

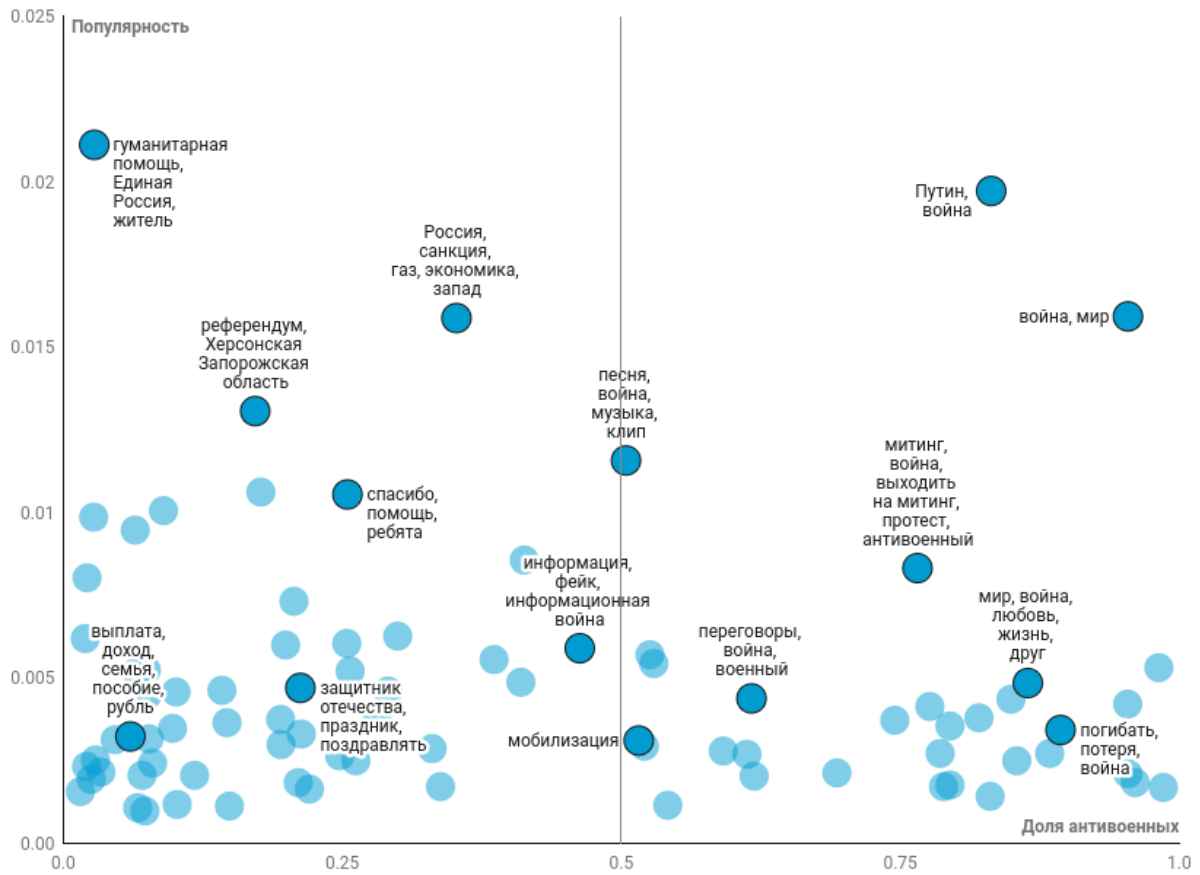
Интересные наблюдения

- На фоне объявления мобилизации произошло заметное увеличение количества постов об антивоенных акциях
- Преобладающим нарративом в осуждении войны, как и в случае с провоенными постами, является обсуждение Путина
- Основным аргументом против войны является ценность мира

Сравнение антивоенных и провоенных постов

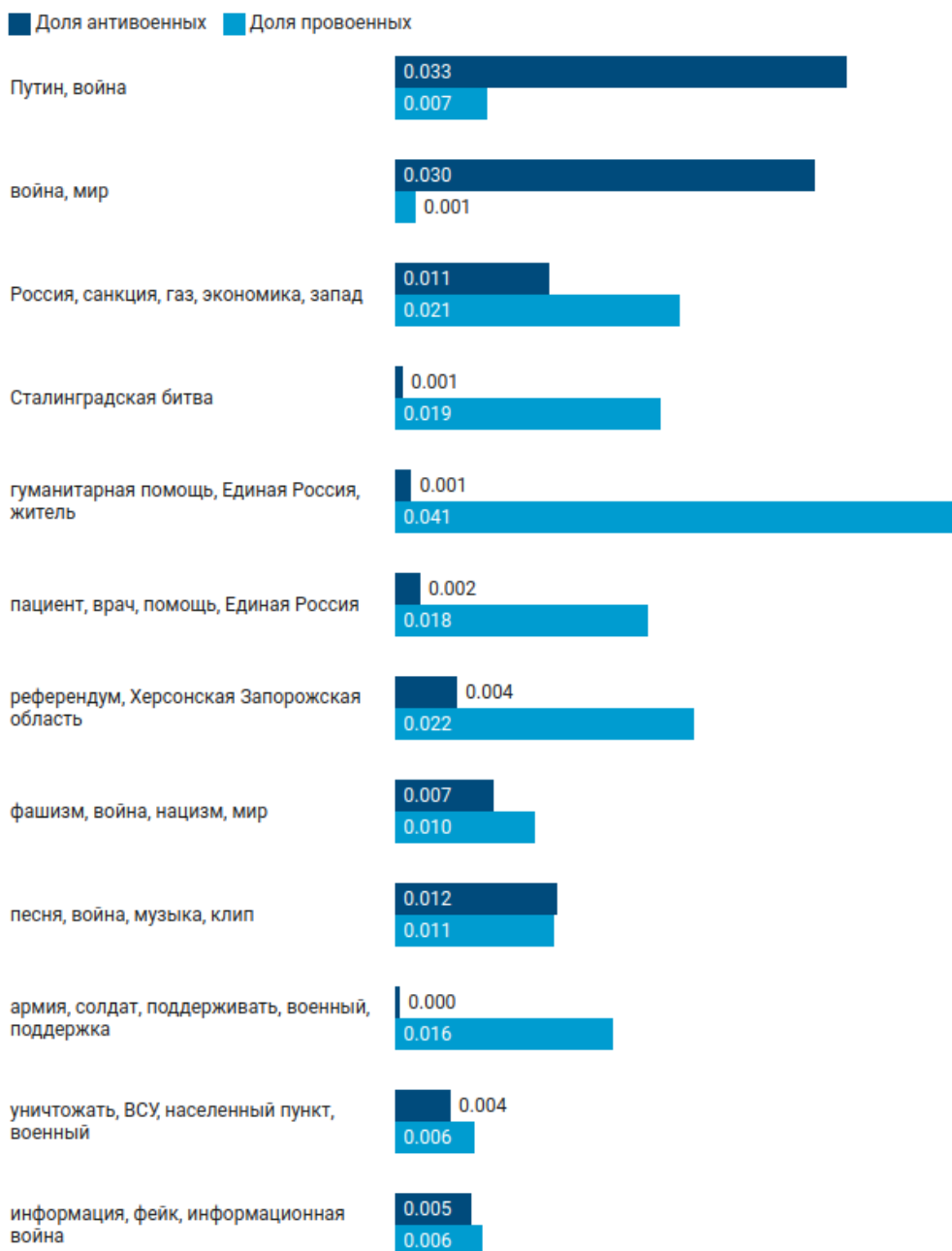
Нарративы сторонников и противников войны

Справа посты противников войны, слева - сторонников



Интерактивный график: <https://datawrapper.dwcdn.net/0a1O9/10/>

Нарративы сторонников и противников войны



Интерактивный график: <https://datawrapper.dwcdn.net/7Ca0r/6/> - там много других тем, которые не поместились на эту картинку.

Здесь показана доля постов, относящихся к данной теме, от всех антивоенных/провоенных постов.

Примеры постов

+ Темы (антивоенные и провоенные)

Интересные наблюдения

- Для сторонников войны важной темой обсуждения являются сборы гуманитарной помощи Российской армии, пособия и льготы
- У сторонников выделена тема про соцопросы и доверие президенту по опросам ВЦИОМ, в то время как противники войны о соцопросах не говорят
- У противников войны есть отдельная тема про стресс и страх, в то время как у сторонников ее нет
- Противники сильно больше говорят про потери, военные действия, обстрелы
- Сторонники войны больше говорят о санкциях, чем противники
- Сторонники и противники войны примерно в одинаковой степени в связи с войной обсуждают музыку (провоенную или антивоенную соответственно), фейки и дезинформацию, Зеленского, Россию и русский мир, мобилизацию и Крымский мост