



ІТМО

**ПРИМЕНЕНИЕ ИНСТРУМЕНТОВ АНАЛИЗА
TELEGRAM КАНАЛОВ ДЛЯ ГОРОДСКИХ И
СОЦИАЛЬНЫХ ПРОЕКТОВ**

Борис Низомутдинов
Анна Углова

Источники информации

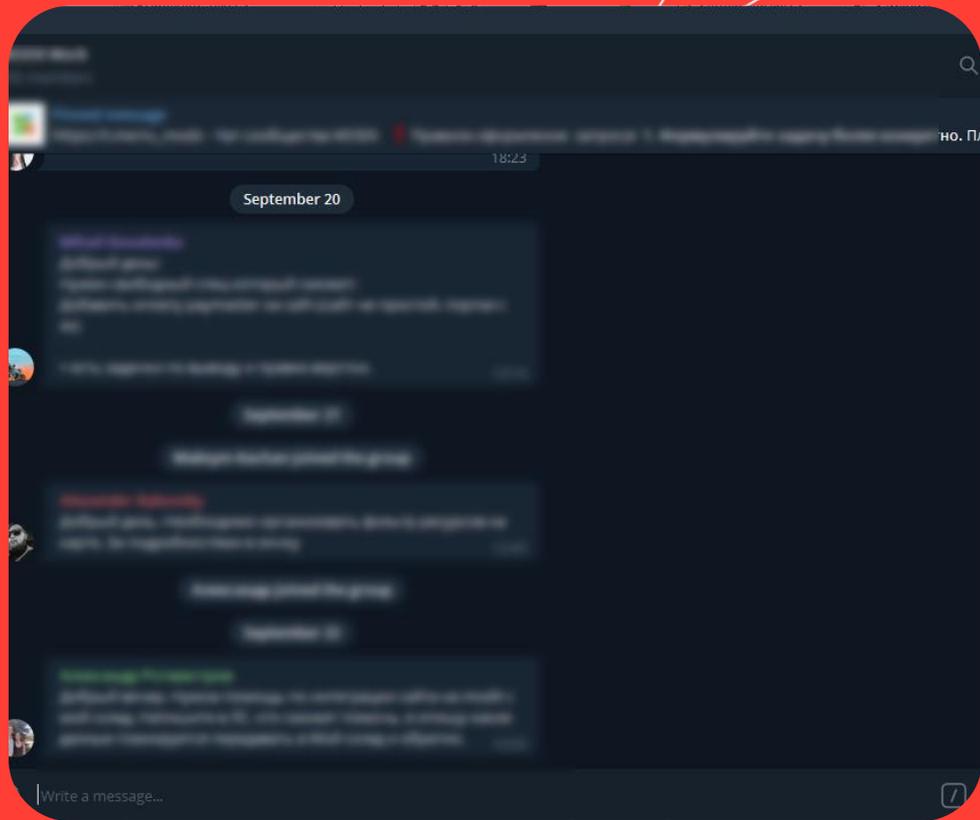


Google Maps



Введение

ІТМО



Для сбора информации были использованы встроенные функции мессенджера Telegram, в интерфейсе программы для ПК имеется возможность выгрузки сообщений из канала или публика без использования API.



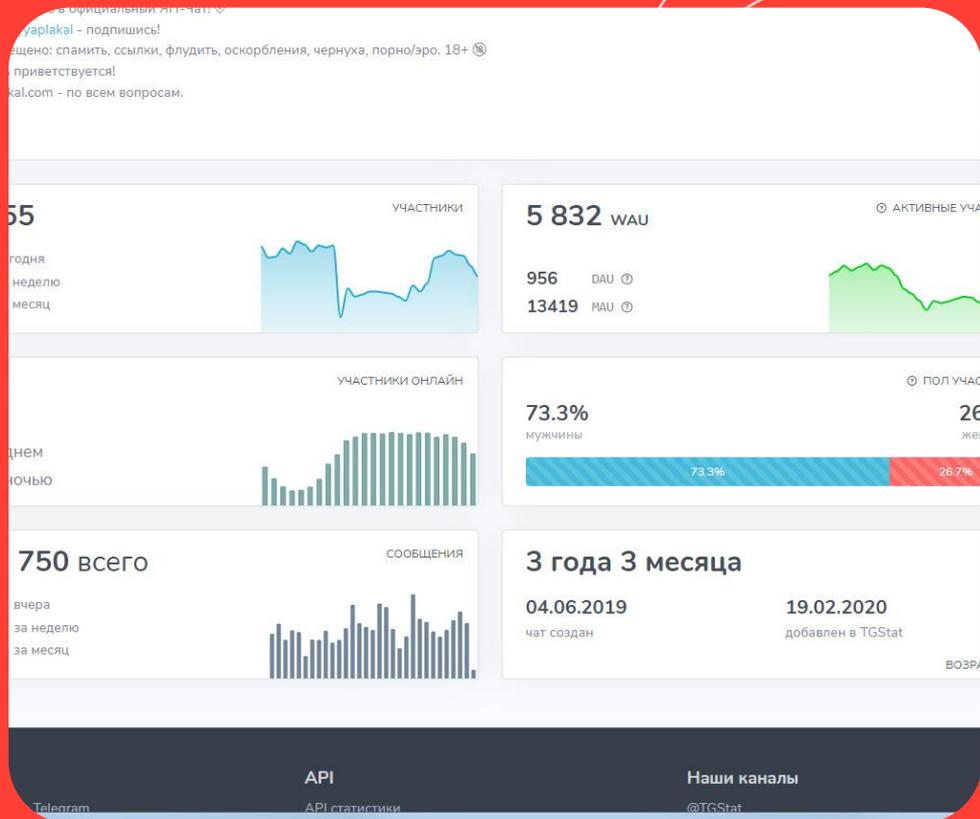
Пример 1

В работе рассмотрена возможность использования результатов автоматизированного анализа деструктивного контента телеграм каналов для прогнозирования рисков развития саморазрушающего поведения у пользователей.

Для сбора информации использовались встроенные функции мессенджера телеграм, для обработки данных была разработана система оценки тональности и поиска сообщений по заданной теме.

Исследование проводилось на примере анализа 269 каналов по 29 тематикам.

TG Stat



На первом шаге, были отобраны каналы в Telegram, по ключевым словам для каждой категории, всего выбрано 269 каналов, далее, с помощью встроенной функции были сохранены все сообщения из этих каналов, итоговая база представляет собой набор json файлов, разделенных на категории. Для обработки полученных файлов была разработана система для отбора сообщений по словарю и подсчет по разным категориям.



Data collection



В ходе обработки отобранных json-файлов использовалась python-библиотека `pandas`, которая предоставляет специальные структуры данных и операции для манипулирования числовыми таблицами и временными рядами (в нашей задаче использовалась для чтения json-файлов и работы с `dataframe`), а также модуль `re`, который предоставляет определенные методы для регулярных выражений (в данной задаче использовался такой метод как `re.split()` для работы со словарями).



Обработка данных



В написанном скрипте отбираются сообщения пользователей, которые содержат слова из заданных терминологических словарей и по каждой категории подсчитывается общее количество и процент отобранных сообщений, а затем данная информация сохраняется в итоговый CSV файл.

На этапе качественного анализа полученных данных, мы посчитали объем деструктивного контента для разных категорий каналов

На основе проанализированной информации были составлены словари по 9 категориям:

- Категория 1. Суицид;
- Категория 2. Кризисные состояния
- Категория 3. Симптом суицидального поведения;
- Категория 4. Факторы, приводящие к суициду;
- Категория 5. Депрессия.
- Категория 6. Апатия/грусть.
- Категория 7. Факторы приводящие к депрессии.
- Категория 8. Потеря Смысла жизни.
- Категория 9. Аутоагрессия.

Обработка данных

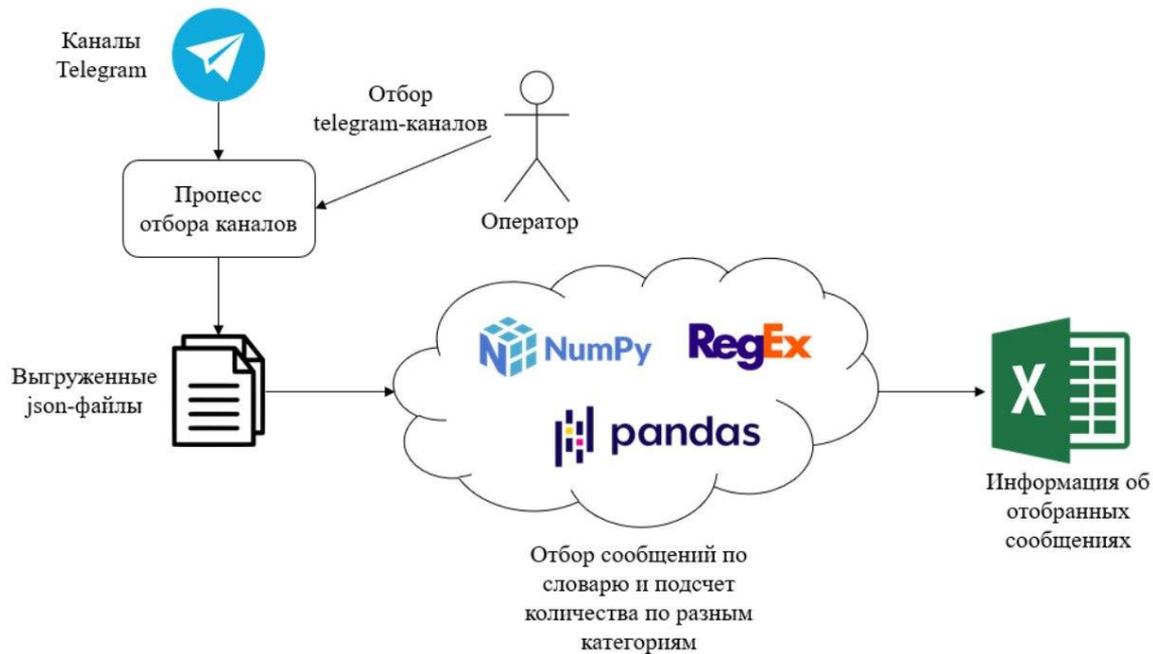


Figure 1

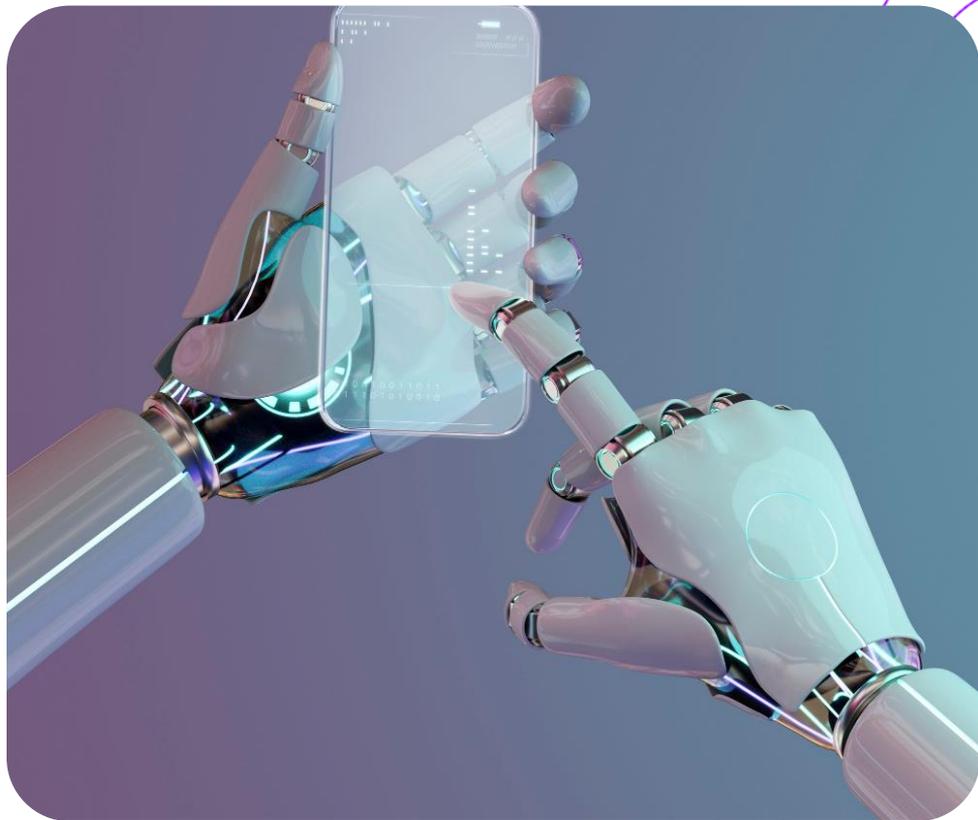
Обработка данных

Таблица 1. Объем деструктивного контента в телеграм-каналах разных категорий

Категории каналов	Общий % деструктивного контента	Описание факторов приводящих к саморазруш. поведению	Описание симптомов суицидального поведения	Описание симптомов депрессии	Описание симптомов аутоагрессии
Психология	54,69	39,60	9,15	12,56	1,66
Медицина	54,53	38,82	8,86	10,22	1,88
Право	49,86	37,61	7,57	10,98	1,86
Эзотерика	45,92	33,08	7,55	10,75	2,00
Бизнес и стартапы	36,93	27,15	6,00	9,00	1,95
Курсы и гайды	33,91	26,77	4,79	5,39	0,11
Здоровье и фитнес	30,79	23,04	4,44	5,86	0,82
Карьера	29,09	22,99	4,94	7,85	0,39
Семья и дети	27,19	21,61	3,53	4,69	0,11
Даркнет	26,11	16,80	8,56	3,80	0,24
Блоги	21,04	15,84	3,58	4,46	0,65
Новости/СМИ	17,73	12,34	3,51	2,53	0,82
Образование	17,12	14,03	2,56	2,71	0,08
Искусство	15,83	11,47	2,98	3,25	0,20
Спорт	14,37	11,06	1,75	2,67	0,53
Путешествия	14,33	10,21	3,03	2,24	0,33
Познавательное	13,53	10,44	2,02	1,91	0,41
Шок-контент	12,87	8,26	2,92	1,97	1,01
Природа	12,29	9,94	1,54	1,16	0,23

Из приведенной таблицы видно, что общий объём деструктивного контента присутствует на каналах всех категорий. В большей степени (30%-54%) он представлен на специализированных каналах связанных с психологией, медициной, правом, а также на каналах с разной тематикой: эзотерика, бизнес, стартапы, курсы, гайды, здоровье и фитнес. Большая часть контента связана с описанием различных факторов, которые провоцируют дистресс и симптомов депрессии, которая может стать причиной возникновения саморазрушающего поведения.

Заключение



Современные социально-психологические исследования демонстрируют важность изучения контента, выкладываемого в социальных сетях, в связи с возросшим количеством аутодеструктивного поведения среди современной молодежи и влиянием информационной среды на эмоционально-психологическое состояние пользователей.

Работа выполнена при поддержке совета по грантам Президента Российской Федерации для государственной поддержки молодых российских ученых № МК-1883.2022.2



Благодарю за внимание

ITMO *re than a*
UNIVERSITY

Boris Nizomutdinov

boris@itmo.ru