

ИНТЕНСИВ
**Архипелаг
2121**

**Интеллектуальная
система фильтрации
обсценной лексики
на основе глубоких
нейронных сетей**

АГЕНТСТВО
СТРАТЕГИЧЕСКИХ
ИНИЦИАТИВ

20.35
УНИВЕРСИТЕТ

ПЛАТФОРМА НТИ



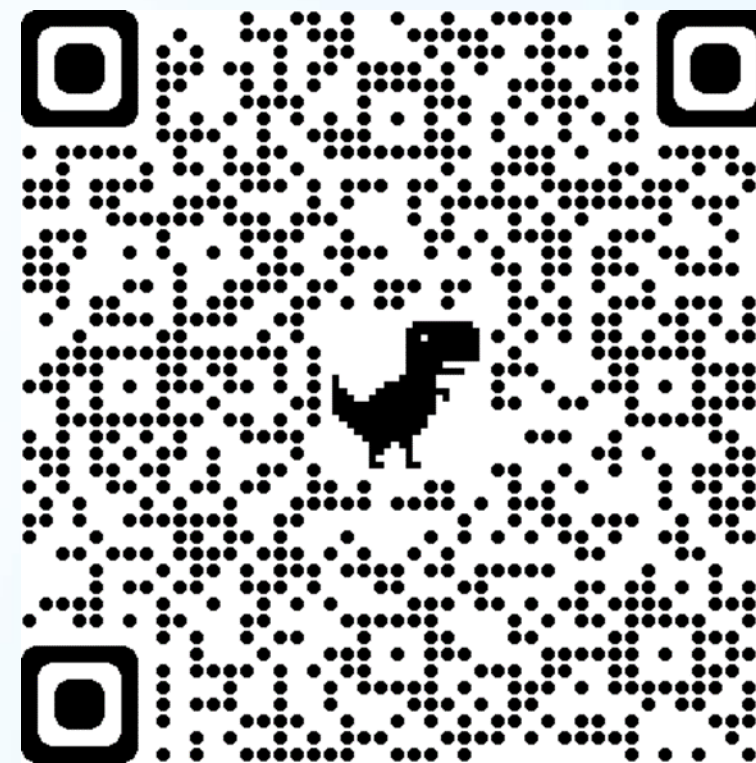
МИНИСТЕРСТВО НАУКИ
И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ



Актуальность проекта

Интеллектуальная система фильтрации обсценной лексики на основе глубоких нейронных сетей приведёт к уменьшению количества административных правонарушений, связанных с обсценной лексикой и её запрете в социальных сетях и позволит уменьшить количество такой лексики и её производных в прямом эфире. Это согласовывается с Законопроектом № 223849-7 «О внесении изменений в Федеральный закон "Об информации, информационных технологиях и о защите информации"» (в части установления особенностей распространения информации в социальной сети)).

Русская обсценная лексика, пример:



Проблема

Почему существующих вариантов решения не достаточно?

Механизмы распознавания речи и обработки аудиозаписей на русском языке мало предобучены, так как формальная семантика русского языка является одной из самых тяжёлых для обучения нейронных сетей. В Российской Федерации для системы фильтрации обсценной лексики на основе глубоких нейронных сетей мало конкурентов: одна из по популярных моделей ASR (Automatic Speech Recognition) от ПАО «Сбербанк России» не может предоставить свою разработку массовому пользователю, так как происходит обучение системы на личных данных клиента.

Проблема клиента, которую вы решаете.

Попадание обсценной лексики в прямой эфир социальных сетей не было регламентировано законами до 2021 года. По данным электронного издания «Тинькофф» и других СМИ, количество вариаций просторечных и сильных окрашенных эмоционально слов резко возросло с момента принятия законопроекта о запрете нецензурной брани в социальных сетях.

Решение

Что вы предлагаете, уникальные преимущества и выгоды для клиента.

Разработать интеллектуальную систему фильтрации обсценной лексики на основе глубоких нейронных сетей, решающую задачи переобучения методом fine-tuning модели NeMo ASR для русского языка, создания графической оболочки, встраивания данной нейронной сети в аудиопоток после выхода микрофона и цензурной обработки аудиодорожки.

РЫНОК

Рынок, на котором вы работаете, его объем, рост и уровень конкуренции.

Массовый рынок, потребители информации в виде видео-контента. Российский рынок предполагает малый уровень конкуренции в сфере фильтрации обсценной лексики.

Бизнес-модель

Бизнес-модель – как вы зарабатываете или планируете.

Ключевые партнёры: государство, госкорпорации

Ключевые виды деятельности: производство (разработка и внедрение нового продукта)

Ключевые ресурсы: интеллектуальные (программное обеспечение)

Ценностные предложения: свежий проект по фильтрации лексики русского языка

Взаимоотношения с клиентами: оперативность, постоянная работа сервисов

Каналы сбыта: сайт проекта, плагины

Потребительские сегменты: массовый рынок

Структура издержек: расходы на создание, воплощение и тестирование системы фильтрации обценной лексики на основе глубоких нейронных сетей, поддержание взаимоотношений с клиентами, получение прибыли.

Потоки поступления дохода: доходы от предоставления системы заказчикам и пользователям

Текущие результаты

Текущие результаты: успешные кейсы, клиенты или предварительные договоренности, привлеченные инвестиции и др.

Проект находится на стадии предобработки и сбора данных, освоения и выявления новых методологий механизмов распознавания речи.



ФГБОУ ВО «Курский государственный университет» оказывает проекту информационную и материальную поддержку.

Команда

Ключевые члены вашей команды (СЕО, СТО и СМО), опыт и компетенции;



Мария Бородина

Анализ данных и дата-инженерия
ИТ-разработка
Систематизация и методология



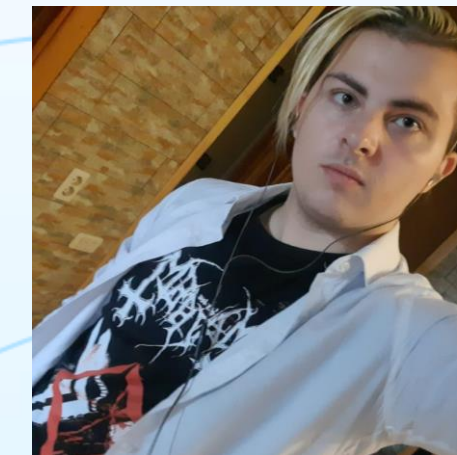
Анастасия Сафронова

PR и маркетинг
Продакт менеджмент
ИТ-разработка



Максим Исаев

Руководитель проекта
ИТ-разработка



Максим Павлов

Систематизация и методология

Планы развития

Планы развития, потребности и предложение для того, кому вы адресуете презентацию.

Разработка
прототипа



Привлечение
инвестиций



Популяризация
продукции



Получение заказов от
госкорпораций



ИНТЕНСИВ
**Архипелаг
2121**

АГЕНТСТВО
СТРАТЕГИЧЕСКИХ
ИНИЦИАТИВ

20.35
УНИВЕРСИТЕТ

ПЛАТФОРМА НТИ



МИНИСТЕРСТВО НАУКИ
И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Контакты

Сайт <https://vk.com/progksu>

Телефон +7 (4712) 70-05-56

Email poiais@kursksu.ru

