

Секция «Технологии искусственного интеллекта в предоставлении государственных и муниципальных услуг»

## **Использование искусственного интеллекта для создания контента**

**Научный руководитель – Мансурова Мадина Есимхановна**

*Абдихан Жармахан Берикович*

*Студент (бакалавр)*

Казахский национальный университет имени аль-Фараби, Механико-математический факультет, Алматы, Казахстан  
*E-mail: abdihan.jarmahan@gmail.com*

Целью работы является создание системы распознавания казахской речи с участием искусственного интеллекта.

Актуальность темы состоит в том, что возрастает казахскоязычный сегмент Интернета. По данным исследовательского института «Общественное мнение», в 2017 году доля населения, владеющего государственным языком, составила 83,1%, в 2018 году - 85,9%. По динамике роста знания языка заметно, что доля казахскоязычных пользователей растет.

В цифровую эпоху распознавание речи считают одной из самых сложных технических задач, связанных с системами искусственного интеллекта. Машинное распознавание речи и аудиоответы пользователям уже широко применяются в жизни и поставлены на коммерческий поток.

Поэтому для журналистов, создающих контент, это приложение будет востребованным, оно также аккумулирует производство медиаконтента на национальном языке. В данном исследовании приведены первичные результаты группы учёных Казахстана по международному проекту "Построение системы распознавания казахской речи" с компанией PERCENT (Китай).

Методологией работы явилась подборка тематического контента для дальнейшей обработки данных. Были определены группы спикеров, которые озвучивали предоставленный контент по определенной схеме. Даны готовые тексты (250 предложений) для каждого спикера. Разъяснялся механизм работы по озвучиванию и сохранению речи в предложенном формате. Производились аудиозаписи продолжительностью примерно 30-35 минут. В инструкциях был описан процесс работы с соответствующими приложениями для Android и iOS. Для тех, кто участвовал в проекте в качестве спикера, были разработаны и предоставлены общие положения. Озвучивание материала проводилось монотонно, умеренно медленно и без художественной выразительности. В каждом файле содержалось 250 строк, в каждой строке 10-20 слов. 1 строка - 1 аудиофайл. Длительность озвучки 15 слов рассчитана на 9-10 секунд. При озвучивании важно было соблюдать монотонность речи без звуков «ыыы», «ыыым ii», «yiiъ» «ыхых».

Группа завершила работу с входными данными - файлами / потоками с аудиоданными. В данном формате имеется два типа файлов. Первый - заголовки с информацией об аудиопотоке: битрейте, частоте, количестве каналов, длине файла и т.д. Второй блок состоит из данных - цифрового сигнала, набора значений амплитуд.

Данные по небольшим временным промежуткам разделены на фреймы. Первой задачей, которую приходится решать при распознавании речи, является разбивка этой самой речи на отдельные слова. Для простоты процесса требуется работа с «разделителями» слов.

В конечном результате проекта будет разработана первая система распознавания казахской речи с участием искусственного интеллекта.

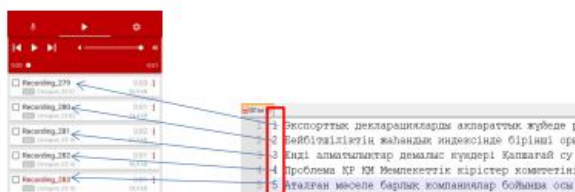
## Источники и литература

- 1) В Казахстане долю населения, владеющего госязыком, планируют увеличить до 95% к 2025 году / <https://tass.ru/obschestvo/7503381>
- 2) Система распознавания речи. <https://polygant.net/ru/ai/sistema-raspoznavaniya-rechi/>

## Иллюстрации

### II. Дыбыс жазу барысы:

Әрбір спикер үшін келесі бейнедегі 1 мәтіндік файл беріледі:



- Мұндагы алгашкы сан сөйлөмнүн реттік номері, **ОЛ ОҚЫЛМАЙДЫ**.
  - Әрбір сөйлөм жеке аудиофайлга жазылады, яғни жаңа аудио басталып жазылады.
  - **Аудио жазбасын бастамае бұрын бұған дейінгі болған бүкіл аудиофайлдарды жойып тастаңыз!!!**
- Оқу темпі әдеттегі ыргакпен болады, яғни спикердің өз сөйлөу темпі: тым жылдам емес, баяуырақ. Мәнерлеп оқу қажет емес (табиғи өз ыргағыңызды сақтаңыз, тым жылдам сөйлөр болсаңыз біршама баяулаңыз).

Оқу темпіне мысал:

**[KAZ]**

**[RUS]**

Рис. 1. Рисунок № 1 – Инструкция для приложения Android

<p>11. Выбираем Archive в «output» и отправляем через удобный вам способ</p>	<p>Мәнерлеп емес, өз табиғи дауысыңызды сақтау жеткілікті. Тыныш бөлме;</p> <ol style="list-style-type: none"> <li>2. 1 файлда 250 жол, әр жолда 10-20 сөзден бар: 1 жол=1 аудиофайл</li> <li>3. егер оқып келе жатып 1 сөзде мүдіріп қателесіп кеткен жағдайда, сол жолды басынан оқу қажет, яғни бастан аяқ қатесіз оқылуы қажет.</li> <li>4. Темп шамалы баяуырақ болады. мысалы: 15 сөз 9-10 секундтай созылады.</li> <li>5. "ыыы", "ыыым иі" "ыій", "ыхых" деп тамақ қырған дыбыстар болмауы керек;</li> <li>6. Біткенде келесі почтаның біріне жіберсеңіз болады: <a href="mailto:marzhan.kyrgyzbaeva@gmail.com">marzhan.kyrgyzbaeva@gmail.com</a> <a href="mailto:nurqaliqadyrbek@gmail.com">nurqaliqadyrbek@gmail.com</a></li> </ol>
--	--

Рис. 2. Рисунок № 2 – Инструкция для приложения iOS