

БӨЛІМ: FARABI UNIVERSITY / UNIVER / СТУДЕНТ

Қазақ тіліндегі дауыстық AI-көмекші: цифрлық трансформация жолындағы жаңа қадам

ЖАРИЯЛАНДЫ
13.04.2026**ТИРЕК СӨЗДЕР**

Google Gemini, NLP, Speech-to-Text, TTS, веб-қосымша, дауыстық AI-көмекші, Жасанды интеллект, қазақ тілі, сөйлеуді тану

СІЛТЕМЕ<https://bilimger.kz/188092/>**Шаяхмет Шахризада Абдрашитқызы**

Эл-фарابي атындағы Қазақ Ұлттық Университетінің Ақпараттық технологиялар факультетінің Ақпараттық жүйелер мамандығының 4-ші курс студенті

Ғылыми жетекші: **А.Н. Шормакова**, PhD, доцент м.а

АННОТАЦИЯ

Мақалада қазақ тілін толық қолдайтын дауыстық AI-көмекші веб-қосымшасын жасаудың нәтижелері сипатталады. Google Cloud Speech-to-Text, Google Gemini және Narakeet TTS технологияларын біріктіре отырып, тану дәлдігі 95%-дан асқан жүйе іске асырылды. Жұмыс отандық IT саласындағы тіл технологияларының дамуына маңызды үлес қосады.

Зерттеу өзектілігі. Бүгінгі таңда Siri, Google Assistant, Amazon Alexa сияқты дауыстық технологиялар бүкіл әлемде кеңінен қолданылуда. Алайда осы жетекші жүйелердің барлығы үшін қазақ тілі — жартылай немесе мүлдем қолданылмайтын тіл. ЮНЕСКО-ның мәліметтеріне сәйкес, тілдік алуантүрлілікті сақтау цифрлық дәуірдің маңызды міндеттерінің бірі болып табылады.

Дипломдық жұмыстың мақсаты — қазақ тілін толық қолдайтын, дауыстық командалар арқылы жасанды интеллектпен өзара әрекет ету мүмкіндігін беретін веб-қосымша жасау. Бұл зерттеу мемлекеттік қызметтерді ана тілінде цифрландыруға, жаңа буын үшін қазақ тілін насихаттауға және мүмкіндігі шектеулі адамдарға технологияға қол жетімділікті арттыруға тікелей үлес қосады.

Қазақ тіліне NLP жүйелерін бейімдеудің күрделілігі

Қазақ тілі — агглютинативті тіл. Бұл оның НЛП жүйелерімен жұмыс жасауды ерекше

қиын ететін бірнеше фундаменталды сипаттамасын алдын ала анықтайды:

Морфологиялық күрделілік: Бір сөзге бірнеше аффикс жалғана отырып, бүтін сөйлемнің мағынасын бере алады — бұл токенизация мен лемматизацияны айтарлықтай қиындатады.

Үндесім заңы: Дыбыс үндестігі аффикстер таңдауын анықтайды, сондықтан дұрыс синтез үшін тек грамматикалық емес, фонетикалық заңдарды да ескеру қажет.

Оқыту деректерінің шектеулігі: Ағылшын немесе орыс тілдерімен салыстырғанда, қазақ тілі корпустары айтарлықтай шағын — бұл модельдердің дәлдігіне тікелей әсер етеді.

Екі алфавит: Кириллица мен латын алфавитінің қатар қолданылуы мәтін нормализациясы кезінде қосымша қиындықтар туғызады.

Жүйенің техникалық шешімі. Жұмыс барысында клиент-сервер архитектурасына негізделген толыққанды веб-қосымша іске асырылды. Жүйенің үш негізгі AI-сервисті біртұтас тізбекке байланыстыратын архитектурасы бар:

Компонент Технология Нәтиже

Сөйлеуді тану STT Google Cloud Speech-to-Text (kk-KZ) Дәлдігі 95%+

AI жауап генерациясы Google Gemini 2.0 Flash 32 000 токен контекст

Сөйлеу синтезі TTS Narakeet (Marzhan дауысы) 22kHz, табиғи үн

Frontend React.js 18.2 + Vite 5.0 Адаптивті интерфейс

Backend Node.js 20 + Express.js REST API, HTTPS

Жүйе қалай жұмыс жасайды? Пайдаланушы браузердегі микрофон батырмасын басады — бұл дауыстық өзара әрекеттестік тізбегін іске қосады. Тізбек төрт кезеңнен тұрады:

1. Дыбыс жазу — Браузердің MediaRecorder API арқылы WEBM/OPUS форматында аудио жазылады (максимум 30 секунд).

2. Сөйлеуді мәтінге айналдыру — Google Cloud Speech-to-Text серверіне жіберіліп, kk-KZ тіл кодымен қазақша мәтінге аударылады.

3. AI жауап генерациясы — Google Gemini моделі тек қазақ тілінде жауап беруге бағдарланған жүйелік нұсқаулар негізінде сұраққа жауап дайындайды.

4. Дауысты синтездеу — Narakeet сервиси Marzhan дауысы арқылы жауапты табиғи қазақша сөйлеуге айналдырып, пайдаланушыға аудио ретінде жеткізеді.

Сынақ нәтижелері және қол жеткізілген көрсеткіштер. Жүйені функционалдық және

функционалдық емес сынақтардан өткізу жоғары нәтижелер берді. Дауысты тануды сынаудың орташа дәлдігі 98,3%-ды құрады, ал толық өңдеу циклінің жалпы уақыты — дауыстық сұраудан аудио жауапқа дейін — 5,5-тен 8,5 секундқа дейін болды, яғни белгіленген 10 секундтық талапты орындады.

Жүйе Chrome, Firefox, Safari браузерлерінде, сондай-ақ iPhone мен Android смартфондарында сәтті жұмыс жасады. Интерфейс WCAG 2.1 қолжетімділік стандарттарына сәйкес жасалды.

Практикалық маңызы және болашақ бағыттары. Жасалған жүйенің қолданылу аясы кең. Ең алдымен, ол білім беру мекемелерінде қазақ тілін үйрету үшін пайдаланылуы мүмкін. Мемлекеттік мекемелерде ақпараттық қызметтерді ана тілінде ұсыну үшін де негіз бола алады. Мүгедектігі бар адамдарға технологияға қол жетімділікті арттыру — тағы бір маңызды бағыт.

Болашақтағы дамыту жоспарлары арасында WebSocket технологиясы арқылы нақты уақыттағы өзара әрекеттестік, офлайн-режимде жергілікті модельдерді қолдану, диалектілерді тану мүмкіндігі, сондай-ақ React Native негізінде мобилдік қосымша жасау бар.

Қорытынды

Бұл дипломдық жұмыс — тек техникалық жоба ғана емес, қазақ тілінің цифрлық кеңістіктегі болашағына деген нақты үлес. Дауыстық AI технологияларын ана тіліне бейімдеу арқылы жасанды интеллектті қазақ тілінде сөйлетуге болатынын дәлелдейді. Ашық бастапқы кодпен жарияланған жоба одан әрі зерттеулер мен коммерциялық әзірлемелер үшін берік негіз болып табылады.

Түйін сөздер: дауыстық AI-көмекші, қазақ тілі, NLP, сөйлеуді тану, Google Gemini, Speech-to-Text, TTS, веб-қосымша, жасанды интеллект

ҚМ АА Куәлік нөмірі: **KZ45VPY00102718** — ҚР Мәдениет және Ақпарат министрлігі

© 2026 **Bilimger.kz** Ақпараттық-танымдық білім порталы. Барлық мазмұн авторлық құқықпен қорғалған.